



ESTUDOS BIOQUÍMICO-COMPUTACIONAIS EM ASPARTIL PROTEASES
CATEPSINA D-SÍMILE DE *SCHISTOSOMA MANSONI* E EM SERINIL PROTEASES
ATIVAS SOBRE A COAGULAÇÃO SANGUÍNEA

Floriano Paes Silva Junior

Tese de Doutorado apresentada ao Programa de Pós-graduação em Química Orgânica, Instituto de Química, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Química Orgânica.

Orientadores: Ricardo Bicca de Alencastro
& Salvatore Giovanni De Simone

Rio de Janeiro
Dezembro de 2005

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

ESTUDOS BIOQUÍMICO-COMPUTACIONAIS EM ASPARTIL PROTEASES
CATEPSINA D-SÍMILE DE *SCHISTOSOMA MANSONI* E EM SERINIL PROTEASES
ATIVAS SOBRE A COAGULAÇÃO SANGUÍNEA

Floriano Paes Silva Junior

Orientadores: Ricardo Bicca De Alencastro & Salvatore Giovanni De Simone

Tese de Doutorado submetida ao Programa de Pós-graduação em Química Orgânica,
Instituto de Química, da Universidade Federal do Rio de Janeiro - UFRJ, como parte
dos requisitos necessários à obtenção do título de Doutor em Química Orgânica.

Aprovada por:

Presidente, Prof. Dr. Ricardo Bicca De Alencastro

Co-orientador, Prof. Dr. Salvatore Giovanni De Simone

Prof. Dr. Pierre Mothé Esteves

Prof. Dr. Glaucius Oliva

Prof. Dr. Walter Filgueira de Azevedo Junior

Prof. Dr. Jose Daniel Figueroa Villar

Rio de Janeiro
Dezembro de 2005

Silva-Jr, Floriano Paes.

Estudos Bioquímico-Computacionais em Aspartil Proteases Catepsina D-Símile de *Schistosoma mansoni* e em Serinil Proteases Ativas sobre a Coagulação Sanguínea / Floriano Paes Silva Junior. - Rio de Janeiro: UFRJ/ IQ, 2005.

xxix, 503f.: il.; 31 cm.

Orientadores: Ricardo Bicca De Alencastro e Salvatore Giovanni De Simone

Tese (doutorado) – UFRJ/ Instituto de Química/
Programa de Pós-graduação em Química Orgânica, 2005.

Referências Bibliográficas: f. 322-366.

1. SmCD1. 2. Hemoglobina. 3. Proteases 4. *Schistosoma mansoni*. 5. Coagulação 6. Veneno de serpentes. I. Alencastro, Ricardo Bicca De & De Simone, Salvatore Giovanni. II. Universidade Federal do Rio de Janeiro, Instituto de Química, Programa de Pós-graduação em Química Orgânica. III. Título.

RESUMO

ESTUDOS BIOQUÍMICO-COMPUTACIONAIS EM ASPARTIL PROTEASES CATEPSINA D-SÍMILE DE *SCHISTOSOMA MANSONI* E EM SERINIL PROTEASES ATIVAS SOBRE A COAGULAÇÃO SANGUÍNEA

Floriano Paes Silva Junior

Ricardo Bicca De Alencastro & Salvatore Giovanni De Simone

Resumo da Tese de Doutorado submetida ao Programa de Pós-graduação em Química Orgânica, Instituto de Química, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Doutor em Química Orgânica.

Proteases são enzimas ubíquas nos organismos vivos, de vírus até mamíferos. Neste trabalho, duas classes de enzimas proteolíticas envolvidas em sistemas biológicos distintos foram abordadas por técnicas bioquímicas, físicas e computacionais. O primeiro destes sistemas diz respeito às aspartil proteases (APs) do trematódeo *Schistosoma mansoni*, agente etiológico da doença parasitária de segunda maior importância em número de mortes anuais que impõe à humanidade. Primeiramente, são relatados resultados de quimioterapia experimental com inibidores de HIV-1 protease e estudos de modelagem molecular de uma AP catepsina D-símile envolvida na degradação da hemoglobina por este parasito (SmCD1). A seguir neste capítulo, são descritos o fracionamento cromatográfico e análises de bancos de dados visando a caracterização de novas APs de *S. mansoni* (SmAPs). Por fim, a clonagem, expressão heteróloga (*Pichia pastoris* e *Escherichia coli*) de duas variantes e um fragmento C-terminal de SmCD1 e os anti-soros policlonais construídos contra estes recombinantes são relatados. No segundo sistema biológico estudado, serinil proteases (SPs) envolvidas na coagulação sanguínea, são descritas a purificação e a caracterização bioquímica e estrutural de duas enzimas trombina-símile de venenos de serpentes (SVTLEs), BJ-48 e LMR-47. Ainda neste capítulo, estudos de modelagem molecular são descritos sobre: (i) o reconhecimento molecular no subsítio S1 de SVTLEs e outras enzimas tripsina-símile e (ii) o sítio de ligação ao sódio e alosteria em SPs da cascata da coagulação humana.

Palavras-chave: SmCD1, Hemoglobina, Proteases, Esquistossomose, Coagulação, Toxinas de venenos de serpentes.

Rio de Janeiro
Dezembro de 2005

ABSTRACT**BIOCHEMICAL-COMPUTATIONAL STUDIES ON CATHEPSIN D-LIKE
ASPARTYL PROTEASES FROM *SCHISTOSOMA MANSONI* AND ON SERINYL
PROTEASES INVOLVED IN BLOOD COAGULATION**

Floriano Paes Silva Junior

Ricardo Bicca De Alencastro & Salvatore Giovanni De Simone

Abstract da Tese de Doutorado submetida ao Programa de Pós-graduação em Química Orgânica, Instituto de Química, da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Doutor em Química Orgânica.

Proteases are ubiquitous enzymes in living organisms, from viruses to mammals. In this work, two classes of proteolytic enzymes involved in distinct biological systems were investigated by biochemical, physical and computational techniques. The first of these systems is related to aspartyl proteases (APs) from the trematode *Schistosoma mansoni*, the etiological agent of the second most important parasitic disease in number of deaths imposed to human kind. Firstly, the experimental chemotherapy with HIV-1 protease inhibitors and molecular modeling studies on a cathepsin D-like AP involved in hemoglobin degradation by this parasite (SmCD1) are reported. Next in the chapter, the chromatographic fractionation and database analysis aiming the characterization of novel APs from *S. mansoni* (SmAPs) are described. Finally, cloning, heterologous expression (*Pichia pastoris* e *Escherichia coli*) of two variants and a C-terminal fragment from SmCD1 along with anti-sera raised against these recombinants are reported. In the second biological system studied, i.e. serinyl proteases (SPs) involved in blood coagulation, the purification and the biochemical and structural characterization of two snake venom thrombin-like enzymes (SVTLEs), BJ-48 and LMR-47, are described. Still in the chapter, molecular modeling studies are described on: (i) molecular recognition in S1 subsite of SVTLEs and other trypsin-like enzymes and (ii) the sodium binding site and allostery in SPs from human coagulation cascade.

Key-words: SmCD1, Hemoglobin, Proteases, Schistosomiasis, Coagulation, Snake venom toxins.

Rio de Janeiro
December 2005

Como não podia deixar de ser, este trabalho é dedicado às pessoas que mais amo em minha vida:

Aos meus pais, Maiber e Floriano, não somente por terem doado o material genético necessário à escritura de minha vida, mas por terem cuidado deste livro com todo o carinho e dedicação que se é possível, permitindo que esta estória caminhasse cada vez mais para um final triunfante e feliz.

À minha companheira Sabrina, parte fundamental desta estória, pela generosidade e compreensão com todos meus momentos de angústia, mal-humor e ausência que este trabalho nos impôs.

“A vida esguicha como uma fonte para aqueles que perfuram a rocha da inércia”

Carrel, Alexis.

AGRADECIMENTOS

Preciso iniciar agradecendo ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, à Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, ao Laboratório Nacional de Luz Síncrotron e à Fundação Oswaldo Cruz pelo apoio financeiro e material sem o qual a realização deste trabalho não seria possível. Em seguida, gostaria de expressar minha profunda gratidão às pessoas que de alguma forma participaram de mais esta conquista:

Ao meu orientador e amigo, Dr. Salvatore Giovanni De Simone pela confiança depositada ao longo destes anos e por toda a liberdade concedida para que eu pudesse me desenvolver como pesquisador;

Ao Dr. Ricardo Bicca de Alencastro pela orientação, consideração e pela importante participação em minha formação científica;

Ao Dr. Octávio Antunes cuja orientação não foi oficializada neste documento, mas a quem dedico profunda admiração e sou grato por dividir seu vasto conhecimento;

Ao Dr. Nilson Zanchin (CEBIME/LNLS) pela orientação em biologia molecular, por ceder espaço e material em seu laboratório para os diversos experimentos de clonagem;

Ao Dr. Fabio Ribeiro (CPqRR) e ao Dr. Arnaldo Maldonado-Jr (IOC/FIOCRUZ) pela cessão de amostras de vermes de *S. mansoni* e pela colaboração nos testes in vivo com fármacos;

À Dra. Eleonora Kurtenbach (CCS/UFRJ) e seu então aluno de mestrado Iuri Bastos pela colaboração com os experimentos de expressão em *Pichia pastoris*;

Ao Dr. Luiz Juliano e à Dra. Maria Juliano (EPM/UNIFESP) pela doação dos substratos fluorogênicos peptídicos;

Ao Dr. Enrico Di Cera (WUSM, St. Louis) pela colaboração nos estudos cinéticos com serinil proteases;

À Dra. Íris Torriane (LNLS) e seu então aluno de doutorado, Me. Cristiano Oliveira, pela colaboração nos experimentos de SAXS;

Às Me. Tereza Lima e Me. Adriana Alves (CEBIME/LNLS) pelo apoio, não apenas técnico, mas também fraternal, que me foi concedido em minhas diversas estadas no LNLS;

Ao Me. Filipe Quirino e à Me. Cláudia Conceição (INCQS/FIOCRUZ) pela colaboração com as análises de carboidratos e disponibilização de glicosidases e colunas cromatográficas;

À Dra. Débora Foguel (CCS/UFRJ) por permitir a utilização do espectropolarímetro de dicroísmo circular;

Ao Dr. Sérgio Verjovski-Almeida e seu então aluno de doutorado Me. Ricardo DeMarco por cederem vermes de *S. mansoni* para a extração do RNA;

À Dra. Yara Traub-Cseko (IOC/FIOCRUZ) e seu aluno de doutorado Marcel Villa por ceder diversos reagentes, animais de laboratório e permitir o uso do termociclador;

Ao Me. Jorge Mendonça (FAR/FIOCRUZ) por disponibilizar os inibidores comerciais de HIV-1 protease na forma de fármaco puro e ao Dr. José Liporage (IPEC/FIOCRUZ) por fornecer as formas farmacêuticas;

Ao Dr. Marcelo Pelajo (Dpto. Patologia, IOC/FIOCRUZ) pela colaboração na microscopia de fluorescência confocal;

Ao Laboratório de Toxinologia (IOC/FIOCRUZ) pelo apoio com técnicas proteômicas em geral;

Aos Laboratórios de Reativos e de Tecnologia de Expressão Recombinante (BIO/FIOCRUZ), por permitirem a utilização do sonicador;

Aos professores e colegas do Laboratório de Físico-química Orgânica e Modelagem Molecular: Dra. Magaly Girão pela orientação e Dr. César Oliveira, Dra.

Rita Cristina, Me. Elaine Martins, Rodrigo Amorim e Felipe Fleming pelos agradáveis momentos de convivência;

Aos colegas do Laboratório de Bioquímica de Proteínas e Peptídeos: Dra. Cristiane Martins pelo apoio irrestrito; Dra. Giani Alexandre pela aquisição dos espectros de fluorescência; Me. Herbert Guedes pela colaboração em muitas etapas deste trabalho e pelas intrigantes discussões científicas; Me. André Souza pela síntese de peptídeos e pela experiência compartilhada; Me. Lúcia Brahim pelo auxílio nas tarefas burocráticas; Mônica Alcon, Luiz Buscaccio, Eliane Nascimento e Carlos Correa pelo apoio técnico; e aos demais alunos e técnicos que permanecem ou passaram pelo laboratório durante a execução deste trabalho, pelo auxílio e pelos prazerosos momentos de descontração dos quais partilhei;

Aos membros da banca pela paciência e compreensão sobre as infundáveis tentativas de marcação de uma data para a defesa e pela generosidade de doarem uma parte de seu precioso tempo para ler um documento tão volumoso como este.

Por fim, aos meus parentes e amigos, Silvio Paulo, Ruth e Mark, Rodrigo Rangel, Rodney e Camille, Fábio (Mad), Fernando e Renata, Marcinho, Leonardo, Marcelo e aos muitos outros cujos nomes sinto em ter omitido pela memória que me falha, por se esforçarem em preservar nossa amizade apesar de meus frequentes momentos de ausência;

“Se eu vi mais longe, foi por estar de pé sobre ombros de gigantes”

Newton, Isaac.

SIGLAS E ABREVIATURAS

2-ME: β -mercaptoetanol

ACN: ACN

BSA: Albumina sérica bovina

CHAPS: 3-[(3-Colamidopropil)dimetilamonio]-1-propano sulfonato

CPK: Esquema de cores utilizado em modelos moleculares (Corey, Pauling, Koltun)

CPqRR: Centro de Pesquisas René Rachou

D.O.^{xxx}: Densidade ótica a xxx nm

DNA: Ácido desoxirribonucléico

DTT: ditioneitol

GSH: glutationa reduzida

GSSG: glutationa oxidada

IDA: ácido iminodiacético

IMAC: Cromatografia de afinidade com metal imobilizado

IOC: Instituto Oswaldo Cruz

IPTG: isopropil-tio- β -galactosídeo

Kb: 1000 pb

LB: Luria-Bertani

LNLS: Laboratório Nacional de Luz Síncrotron

MALDI-TOF: Ionização por desorção a laser auxiliado pela matriz-tempo de vôo

Mb: 10⁶ pb

MES: Ácido 2-morfolinoetanossulfônico

NaOAc: acetato de sódio

NC: Nitrocelulose

NTA: ácido nitriloacético

OG: octil- β -glicosídeo

Pb: pares de base

PBS: Salina tamponada com fosfato

PCR: Reação em cadeia da polimerase

QSAR: Relação estrutura-atividade quantitativa

RNA: Ácido ribonucléico

SDS: dodecilsulfato de sódio

SDS-PAGE: Eletroforese em gel de poliacrilamida-SDS

SEC: cromatografia de exclusão molecular

TCA: Ácido tricloroacético

TFA: Ácido trifluoroacético

LISTA DE TABELAS

Tabela 1: Sítios de clivagem na hemoglobina humana, reconhecidos pelas catepsinas D de esquistossomos (SjCD1.1 e SmCD1.1) e humana (hCD).	59
Tabela 2: Doses e posologia dos fármacos.....	71
Tabela 3: Matriz de planejamento do desenho fatorial fracionado empregado na triagem de condições favoráveis à renaturação das formas longa e curta de SmCD1.	88
Tabela 4: Níveis dos fatores pesquisados no desenho fatorial fracionado.	89
Tabela 5: Avaliação do tratamento experimental de camundongos infectados com <i>S. mansoni</i>	107
Tabela 6: Lista de transcritos semelhantes à SmCD1.1 encontrados no “ <i>Schistosome Gene Index</i> ”.....	109
Tabela 7: Lista de transcritos semelhantes à SmCD1.1 encontrados no “ <i>Schistosoma mansoni</i> EST genome project”	109
Tabela 8: Genes parálogos de aspartil proteinases em <i>Schistosoma mansoni</i> (SmAPs) identificados por análise <i>in silico</i> de transcritos do parasito.....	111
Tabela 9: Rendimento em mg de proteína obtido das preparações de corpos de inclusão (CI) das formas longa e curta de SmCD1.	125
Tabela 10: Constantes de especificidade (K_s) das enzimas Lmr-47 e BJ-48 contra tripeptídeos variando em P1.	284
Tabela 11: Constantes cinéticas das enzimas Lmr-47 e BJ-48 contra moléculas ativas sobre o sistema hemostático.....	286
Tabela 12: Avaliação do efeito do carboidrato na redução da atividade amidásica de BJ-48 mediante tratamento térmico.....	303
Tabela 13: Avaliação do efeito do carboidrato na resistência de BJ-48 à inibição da atividade amidásica por inibidores protéicos.....	304

LISTA DE ILUSTRAÇÕES

ESQUEMAS

Esquema 1	18
Esquema 2	19
Esquema 3	22
Esquema 4	27
Esquema 5	38

FIGURAS

Fig. 1: Representação esquemática do reconhecimento molecular de proteases.....	4
Fig. 2: Ciclo de desenvolvimento de novos fármacos baseados na estrutura de seus alvos biológicos.	9
Fig. 3: Alinhamento de aspartil proteases ortólogas na região do segundo aspartato catalítico (Asp ²¹⁵).....	14
Fig. 4: Enovelamento de aspartil proteinases eucarióticas e retrovirais..	15
Fig. 5: Elementos estruturais e simetria no enovelamento de aspartil proteases do clã da pepsina.....	16
Fig. 6: Rede de ligações de hidrogênio (pontilhado verde) formadas por átomos da cadeia principal e das cadeias laterais dos resíduos do sítio catalítico da aspartil proteinase do vírus HIV-1 (código PDB 5HVP)..	17
Fig. 7: Os dois grupos carboxílicos de Asp ³⁵ (D32 na pepsina) e Asp ²¹⁸ (D215 na pepsina) com sua molécula de água associada..	19
Fig. 8: Isómeros de estado de transição (IET). IETs mimetizam o intermediário tetraédrico formado na hidrólise da ligação peptídica por aspartil proteases (topo).	20
Fig. 9: Mecanismo catalítico proposto para rhizopuspepsina (SUGUNA <i>et al.</i> , 1987).....	21
Fig. 10: Mecanismo catalítico de aspartil proteases eucarióticas proposto por Andreeva & Rumsh (2001).....	23
Fig. 11: Aparato catalítico simétrico de aspartil proteases, conforme proposto por Northrop (2001).	24
Fig. 12: Modelagem da interação do composto 12 no sítio ativo da HIV-1 protease nativa..	28
Fig. 13: Interação de AG-1343 com a protease do HIV (código PDB 1OHR).....	29
Fig. 14: Distribuição e situação global do controle da esquistossomose.....	31
Fig. 15: Esquistossomose mansônica em dois meninos de área endêmica no Brasil.....	32

Fig. 16: Série histórica dos percentuais de positividade para esquistossomose no Brasil de 1977 a 2004.	33
Fig. 17: Anatomia de vermes adultos de <i>S. mansoni</i>	34
Fig. 18: Ciclo evolutivo de <i>Schistosoma sp.</i>	36
Fig. 19: Cromossomos de <i>S. mansoni</i>	45
Fig. 20: Digestão da hemoglobina por esquistossomos.	53
Fig. 21: Relação estrutural entre as duas formas possíveis da proteinase catépsina D-símile de <i>S. mansoni</i>	56
Fig. 22: Especificações do desenho fatorial fracionado empregado na renaturação de SmCD1.1 e SmCD1.2.	87
Fig. 23: Cromatografia de afinidade da fração EAV em coluna pepstatina-agarose.	104
Fig. 24: Análise por SDS-PAGE 12% em condições redutoras das frações da cromatografia de EAV em pepstatina-agarose.	105
Fig. 25: 2DE 15% de FLPA. Foram aplicados 150 µg de proteína à fita de 7 cm (IPG strip 3-10) para a primeira dimensão.	106
Fig. 26: Eficácia no tratamento experimental da esquistossomose mansônica com inibidores de HIV-1 protease.	108
Fig. 27: Análise dos transcritos de <i>S. mansoni</i> semelhantes à SmCD1.1 presentes nos bancos “Schistosoma Gene Index” (SmGI) e “ <i>Schistosoma mansoni</i> EST genome project” (SmEST).	110
Fig. 28: Alinhamento múltiplo entre SmCD1 e fragmentos (seqüências parciais) dos demais genes pertencentes à família de SmAPs.	113
Fig. 29: Histograma dos genes da família de aspartil proteinases de <i>S. mansoni</i> (SmAPs) de acordo com o número de ESTs encontrados no SmGI Release 5.	114
Fig. 30: Análise em gel de agarose (0,8%) de amostras do teste de amplificação com o kit Access RT-PCR System (Promega).	116
Fig. 31: Análise em gel de agarose (0,8%) de amostras do teste de amplificação com o kit Thermoscript RT-PCR System (Life Technologies).	117
Fig. 32: Análise em gel de agarose (0,8%) de amostras de PCR de colônia para confirmação dos insertos pro-SmCD1 em clones de <i>E. coli</i> DH5α/pGEM-t.	119
Fig. 33: Análise em gel de agarose (0,8%) de amostras da digestão com Nde I para determinação da orientação dos insertos de pro-SmCD1 em clones <i>E. coli</i> DH5α/pGEM-t.	119
Fig. 34: Análise em gel de agarose (0,8%) de amostras da subclonagem das isoformas de proSmCD1 em pET28a.	121
Fig. 35: Análise em gel de agarose (0,8%) de amostras referentes à correção da deleção na região 3' no cDNA da forma longa de SmCD1.	121

Fig. 36: Análise em gel de agarose (0,8%) de amostras de PCR de colônia para a confirmação dos insertos proSmCD1-corr em clones de <i>E. coli</i> DH5 α /pET28a.....	122
Fig. 37: Análise em gel de agarose (0,8%) de amostras dos clones corrigidos (5C e 3C) de pET28a/proSmCD1.1 clivados com enzimas de restrição.....	123
Fig. 38: Análise por SDS-PAGE (10%, redutora) de amostras do teste de indução da expressão das formas longa e curta de SmCD1 em <i>E. coli</i> BL21(DE3).....	124
Fig. 39: Análise por SDS-PAGE (12%, redutora) de amostras de preparações utilizadas nos ensaios de solubilização de corpos de inclusão (CI) de proSmCD1.....	125
Fig. 40: Efeito do pH na solubilização dos corpos de inclusão de proSmCD1 por 8M uréia.	126
Fig. 41: Efeito de uréia, sarcosil e guanidina-HCl (em pH 7,5) na solubilização dos corpos de inclusão das formas longa (A) e curta (B) de proSmCD1.....	127
Fig. 42: Efeito de aditivos na solubilização dos corpos de inclusão das formas longa (A) e curta (B) de proSmCD1 por 8M uréia, pH 7,5.....	128
Fig. 43: Purificação das isoformas de (6xHis)proSmCD1 por cromatografia de afinidade em Ni-IDA-agarose (5 mL) em condições desnaturantes.....	129
Fig. 44: Análise do estado de agregação das isoformas de SmCD1 por cromatografia de exclusão molecular (SEC).....	130
Fig. 45: Análise por SDS-PAGE (12%, redutora) de amostras do ensaio de renaturação das isoformas de proSmCD1.	132
Fig. 46: Gráficos da D.O. ²⁸⁰ observada versus predita pelos modelos de regressão linear da triagem de condições de renaturação da forma curta (A) e longa (B) de proSmCD1	133
Fig. 47: Perfis preditivos dos 5 fatores mais importantes para a recuperação das isoformas de proSmCD1 solúvel dos ensaios de renaturação.	134
Fig. 48: Gráfico dos efeitos principais normalizados para cada um dos dez fatores avaliados no desenho fatorial fracionado 2 ¹⁰⁻⁶ com resolução III.....	135
Fig. 49: Análise da produção de soros policlonais em coelhos imunizados com as formas longa e curta de proSmCD1.....	136
Fig. 50: Representação esquemática da construção do vetor de expressão pB _{Uf5} /PepC-term derivado de pET21a.	137
Fig. 51: Análise por SDS-PAGE (12%, redutora) das frações do extrato solúvel (S) e de corpos de inclusão (P) do teste de indução da expressão de PtnD _{-pepC-term} por 1 mM de IPTG a 37 °C.	138
Fig. 52: Análise por SDS-PAGE (15%, redutora) das frações de 3h de indução da expressão de PtnD _{-pepC-term} em <i>E. coli</i> BL21(DE3).	139

Fig. 53: Análise por SDS-PAGE (15%, redutora) das frações da cromatografia do extrato solúvel de <i>E. coli</i> BL21(DE3) expressando PtnD/pepC-term em resina Ni-NTA superflow.....	141
Fig. 54: Análise por SDS-PAGE (12%, redutora) das amostras do teste de clivagem com protease TEV.....	142
Fig. 55: Análise por SDS-PAGE (15%, redutora) das frações da cromatografia de PtnD/pepC-term na coluna Econopac High-Q.....	144
Fig. 56: Análise por SDS-PAGE (12%, redutora) das amostras de PtnD/pepC-term cromatografadas em coluna Resource S acoplada a sistema de HPLC no 1º teste (A) e 2º teste (B).	145
Fig. 57: Análise por SDS-PAGE (15%, redutora) das frações da 2ª IMAC para isolamento de pepC-term a partir da proteína de fusão digerida com a protease TEV..	145
Fig. 58: Análise por espectrometria de massas (MALDI-TOF) da fração não-ligada na 2ª IMAC	146
Fig. 59: Análise por espectrometria de massas (MALDI-TOF) da fração não-ligada na 2ª IMAC submetida à hidrólise com tripsina. Os íons relevantes estão assinalados com setas.	147
Fig. 60: Cromatografia em fase reversa (coluna C18 VP-ODS) da amostra de pepC-term (não-ligado da 2ª IMAC).....	148
Fig. 61: Cromatografia de exclusão molecular (coluna Superdex Peptide HR 10/30) da preparação de pepC-term.....	149
Fig. 62: Análise por espectrometria de massas (MALDI-TOF) do pico 4 da SEC.....	150
Fig. 63: Análise por espectrometria de massas (MALDI-TOF) do pico 4 digerido com tripsina.	151
Fig. 64: Modelo ab initio de pepC-term predito pelo programa SCRATCH.....	153
Fig. 65: Predição de estrutura secundária de pepC-term.....	154
Fig. 66: Modelo do segmento 19-46 de pepC-term predito pelo programa 3D-PSSM.....	155
Fig. 67: Espectros de CD de 18N_pepC-term realizados em água pura (A), tampão NaOAc pH 3,5 (B) e tampão fosfato pH 7,5 (C).	157
Fig. 68: Determinação do título e especificidade do soro anti-18N_pepC-term por <i>Immunoblots</i>	158
Fig. 69: Análise do potencial imunogênico da extensão C-terminal da forma longa de SmCD1.....	159
Fig. 70: Análise em gel de agarose (0,8%) dos fragmentos lineares gerados para a clonagem de proSmCD1.1 e proSmCD1.2 em pPIC9k.....	160

Fig. 71: Análise em gel de agarose (0,8%) das amostras de PCR para detecção da presença dos insertos correspondendo às formas longa e curta de proSmCD1 em pPIC9k.	161
Fig. 72: Análise em gel de agarose (0,8%) dos fragmentos gerados pelas enzimas Eco RI e Not I para confirmação da clonagem de pro-SmCD1.1 e pro-SmCD1.2 em pPIC9k.	161
Fig. 73: Análise em gel de agarose (0,6%) das amostras dos vetores pPIC9k (1), pPIC9k/proSmCD1.1 (2) e pPIC9k/proSmCD1.2 (3), submetidas a linearização com Sal I.	162
Fig. 74: Análise em gel de agarose (0,8%) das amostras de PCR de colônia para confirmação dos integrantes His+.	163
Fig. 75: Análise em gel de agarose (0,8%) das amostras de PCR de colônia (<i>primers</i> AOX1) para confirmação dos integrantes His ⁺ selecionados em geneticina 1 mg/mL (1G), 2 mg/mL (2G) e 4 mg/mL (4G).	164
Fig. 76: Análise por SDS-PAGE (12%, redutora) das amostras de sobrenadantes de culturas de <i>P. pastoris</i> induzidas com 0,5% de metanol por 0h (A), 12h (B) e 48h (C) para expressão da forma curta de pro-SmCD1.	165
Fig. 77: Análise por SDS-PAGE (12%, redutora) de amostras de sobrenadantes de culturas de <i>P. pastoris</i> induzidas com 0,5% de metanol por 72h para expressão da formas curta e longa de proSmCD1.	166
Fig. 78: Curvas de crescimento de culturas de <i>P. pastoris</i> GS115 induzidas com 0,5% de metanol em meio mínimo (BM) para expressão das formas longa e curta de pro-SmCD1.	167
Fig. 79: Análise por SDS-PAGE (12%, redutora) de amostras das frações celulares de clones de <i>P. pastoris</i> GS115 expressando as formas longa e curta de SmCD1.	168
Fig. 80: <i>Immunoblots</i> para análise da reatividade dos soros anti-proSmCD1.1, anti-proSmCD1.2 e anti-18N_pepC-term.	169
Fig. 81: Alinhamento estrutural de serinil proteases pertencentes à família S1A.	204
Fig. 82: Enovelamento de enzimas da família da quimiotripsina.	206
Fig. 83: Mecanismo catalítico em serinil proteases da família da quimiotripsina.	208
Fig. 84: Mecanismo cinético da hidrólise da ligação peptídica catalisada por serinil proteases da família S1A.	209
Fig. 85: Mecanismos de inativação de serinil proteases da família da quimiotripsina por modificação covalente com compostos sintéticos.	211
Fig. 86: Mecanismo cinético e aspectos estruturais da inibição por SERPINS.	212
Fig. 87: Modo de ligação canônico de inibidores protéicos reversíveis de serinil proteases conforme ilustrado pelo complexo entre tripsina suína e SBTI (PDB 1AVW).	214

Fig. 88: Sítio de ligação e modo de interação de derivados de benzamidina com enzimas tripsina-símile.	215
Fig. 89: Localização das oito alças na superfície responsáveis pela determinação da especificidade de substrato em enzimas da família da quimiotripsina.	224
Fig. 90: Cascata da coagulação.	225
Fig. 91: Localização celular de processos envolvidos na ativação da coagulação.	227
Fig. 92: Alinhamento estrutural das principais serinil proteases da cascata da coagulação humana.	228
Fig. 93: Sítios alostéricos no domínio catalítico de FVIIa (A) e estrutura de domínios em FVIIa em complexo com fator tecidual (B).	230
Fig. 94: Sítios de interação entre FIXa e FVIIIa mapeados nas estruturas 3D.	233
Fig. 95: Determinantes de especificidade e sítios alostéricos em FXa.	236
Fig. 96: Exosítio de ligação a ânions em aPC.	239
Fig. 97: Determinantes estruturais da especificidade de substrato na α -trombina.	243
Fig. 98: Componentes do veneno de serpentes afetando a hemóstase.	246
Fig. 99: Alinhamento entre SVSPs e serinil proteases de mamíferos selecionadas.	248
Fig. 100: Estrutura de uma serinil protease de veneno de serpentes.	254
Fig. 101: Modo de ação de inibidores indiretos e diretos de trombina. Vide texto para maiores detalhes.	265
Fig. 102: Perfil de variação da atividade amidásica de Lmr-47 com a temperatura.	284
Fig. 103: Análise por SDS-PAGE 12% redutora de amostras do teste de ligação de BJ-48 em coluna Con A.	287
Fig. 104: Análise por SDS-PAGE 12% redutora de amostras da deglicosilação enzimática de Lmr-47 em condições desnaturantes.	288
Fig. 105: Espectros de CD de Lmr-47 em soluções aquosas tamponadas em pH 6,0 e 7,5.	289
Fig. 106: Espectros de fluorescência intrínseca de Lmr-47 em soluções aquosas tamponadas em pH variando de 5,0 a 8,0.	290
Fig. 107: Localização espacial de resíduos potencialmente envolvidos nas propriedades de fluorescência de Lmr-47.	291
Fig. 108: Efeito do pH na intensidade de fluorescência de Lmr-47.	292
Fig. 109: Análise de dados de SAXS para Lmr-47 em pH 6,0 e 10 °C.	293
Fig. 110: Modelo <i>ab initio</i> de Lmr-47 calculado a partir dos dados experimentais de SAXS.	295

Fig. 111: Comparação entre o envelope molecular de Lmr-47 obtido por SAXS e o modelo teórico da cadeia polipeptídica da enzima.	295
Fig. 112: Análise por SDS-PAGE 12% redutora de amostras do ensaio piloto de deglicosilação de BJ-48 e Lmr-47 com PNGase F.	297
Fig. 113: Análise por SDS-PAGE 12% redutora de amostras da deglicosilação de BJ-48 e Lmr-47 nativas.	299
Fig. 114: Análise cromatográfica de monossacarídeos neutros liberados pela deglicosilação não-desnaturante de BJ-48 (A) e Lmr-47 (B) no sistema Dionex Biol C.	301
Fig. 115: Análise das amostras da deglicosilação nativa de BJ-48 (A) e Lmr-47 (B) fracionadas por HIC na coluna Phenyl 5PW.	302

SUMÁRIO

RESUMO	iv
ABSTRACT	v
DEDICATÓRIA.....	vi
AGRADECIMENTOS	vii
SIGLAS E ABREVIATURAS	x
LISTA DE TABELAS	xi
LISTA DE ILUSTRAÇÕES	xii

PRELÚDIO

1 Proteases: função catalítica importância biológica e regulação	2
2 Estratégias catalíticas e esquema de classificação corrente	5
3 Proteases como alvo de desenvolvimento de novos fármacos	7

CAPÍTULO 1

1 INTRODUÇÃO.....	12
1.1 Estrutura e função de aspartil proteases do clã da pepsina.....	12
3.1.1 Aspectos gerais, evolução e enovelamento	12
1.1.1 Mecanismo catalítico e reconhecimento de substrato revelados por estudos com inibidores	17
3.1.2 Ativação e papel do pro-peptídeo na estabilização do enovelamento.....	24
3.1.3 Ativação e papel do pro-peptídeo na estabilização do enovelamento.....	25
1.2 Aspartil proteases catépsina D-símile de <i>S. mansoni</i>	30
1.2.1 Esquistossomose.....	30
1.2.1.a Agentes etiológicos e distribuição.....	30
1.2.1.b Impacto global da doença.....	31
1.2.1.c Situação da doença no Brasil.....	32
1.2.1.d Biologia básica e ciclo evolutivo dos esquistossomos	34
1.2.1.e Patogênese	36
1.2.1.f Diagnóstico	37

1.2.1.g	Controle farmacoterapêutico da esquistossomose mansônica	37
1.2.1.h	Resistência aos fármacos em uso e novas alternativas terapêuticas	41
1.2.2	Projeto genoma de <i>S. mansoni</i> : fonte potencial de novos alvos terapêuticos.....	44
1.2.2.a	Genoma nuclear.....	44
1.2.2.b	Programa de descoberta de novos genes: análise do transcriptoma	45
1.2.2.c	Sequenciamento total do genoma (STG).....	47
1.2.2.d	Genômica funcional.....	48
1.2.3	Digestão da hemoglobina por <i>S. mansoni</i>	49
1.2.4	A hemoglobinase catepsina D-símile	53
1.2.4.a	Duas formas da mesma enzima?	53
1.2.4.b	Eficácia da vacinação com SjCD.....	57
1.2.4.c	Especificidade de substrato	58
1.2.5	Inibição de enzimas do catabolismo da hemoglobina como estratégia de desenvolvimento de novos fármacos esquistossomicidas	60
2	OBJETIVO	62
2.1	Geral	62
2.2	Específicos.....	62
3	MATERIAIS E MÉTODOS.....	63
3.1	Reagentes.....	63
3.2	Procedimentos gerais.....	63
3.2.1	Métodos empregados em proteínas	63
3.2.1.a	SDS-PAGE.....	63
3.2.1.b	Immunoblots.....	63
3.2.1.c	Estimativa do conteúdo protéico	64
3.2.1.d	Precipitação de proteínas	64
3.2.1.e	Hidrólise trípica em solução.....	65
3.2.1.f	Hidrólise trípica <i>in situ</i> de proteínas em gel.....	65
3.2.1.g	Sequenciamento N-terminal por degradação de Edman automatizada.....	65
3.2.2	Métodos empregados em ácidos nucléicos.....	66
3.2.2.a	Eletroforese em gel de agarose.....	66

3.2.2.b	Estimativa do conteúdo	66
3.2.2.c	Sequenciamento de DNA	66
3.2.3	Espectropolarimetria de dicroísmo circular (CD)	67
3.2.4	MALDI-TOF	67
3.2.5	Produção de anti-soro policlonal em coelho	67
3.2.6	Bioinformática	68
3.3	Fracionamento de extratos de vermes adultos por cromatografia de afinidade em pepstatina-agarose.....	69
3.3.1	Preparo do extrato solúvel	69
3.3.2	Cromatografia de afinidade em pepstatina-agarose.....	69
3.3.3	Atividade enzimática	69
3.3.4	Eletroforese bidimensional (2DE).....	70
3.4	Quimioterapia experimental com inibidores comerciais de HIV-1 protease.....	70
3.4.1	Geral	70
3.4.2	Infecção experimental.....	71
3.4.3	Grupos e tratamento experimental.....	71
3.4.4	Forma de administração.....	72
3.4.5	Análise dos resultados	72
3.5	Pesquisa in silico de transcritos pertencentes à família gênica de aspartil proteases em <i>S. mansoni</i> (SmAPs)	72
3.5.1	Bancos de dados	72
3.5.2	Metodologia.....	72
3.6	Clonagem dos genes das formas longa e curta de SmCD1	73
3.6.1	Desenho dos oligonucleotídeos (<i>primers</i>).....	73
3.6.2	Vermes.....	74
3.6.3	Extração do RNA	74
3.6.4	Otimização das condições de RT-PCR.....	75
3.6.5	Clonagem em vetores de passagem (pGEM-T).....	76
3.6.5.a	Purificação dos fragmentos e ligação	76
3.6.5.b	Transformação, seleção dos clones positivos e determinação da orientação	76
3.7	Produção das formas longa e curta de proSmCD1 em <i>E. coli</i>	77
3.7.1	Clonagem em vetores de expressão pET28a	77
3.7.1.a	Preparação dos plasmídeos.....	77

3.7.1.b	Digestão com enzimas de restrição	77
3.7.1.c	Purificação dos fragmentos e ligação	78
3.7.1.d	Transformação em <i>E. coli</i> DH5 α , seleção dos clones positivos e preparação dos plasmídeos	78
3.7.1.e	Transformação em <i>E. coli</i> BL21(DE3)	78
3.7.2	Correção do clone pET28a/proSmCD1.1	78
3.7.2.a	Subclonagem da forma longa corrigida no vetor pET28a.....	79
3.7.3	Teste de indução da expressão em <i>E. coli</i> BL21(DE3)	79
3.7.3.a	Indução e coleta das amostras	79
3.7.3.b	Extratos	80
3.7.4	Teste de solubilização de corpos de inclusão (CI)	80
3.7.4.a	Preparação dos CI.....	80
3.7.5	Teste de purificação por IMAC	82
3.7.6	Análise do estado de agregação em 4M de guanidina-HCl por cromatografia de exclusão molecular (SEC).....	82
3.7.7	Teste de renaturação por diluição	83
3.7.8	Triagem de condições de renaturação	83
3.7.8.a	Ensaio preliminares.....	83
3.7.8.b	Desenho fatorial fracionado.....	84
3.7.8.c	Metodologia de renaturação	89
3.7.8.d	Respostas – hidrólise de substrato fluorogênico e concentração de proteína	89
3.7.8.e	Planejamento e análise estatística.....	90
3.7.9	Teste de auto-ativação	90
3.7.10	Tentativa de renaturação por diálise.....	90
3.8	Produção do domínio COOH-terminal de SmCD1 (pepCterm) em <i>E. coli</i>	91
3.8.1	Clonagem em vetor pBUF5	91
3.8.1.a	<i>Primers</i> e vetores	91
3.8.1.b	Clonagem.....	91
3.8.2	Estudos de expressão da fusão PtnD/PepC-term.....	92
3.8.3	Padronização da purificação de PtnD/PepC-term por IMAC.....	92
3.8.3.a	Expressão e preparação do extrato solúvel.....	92
3.8.3.b	Cromatografia.....	92
3.8.4	Testes de purificação de PtnD/PepC-term.....	93

3.8.4.a	Resina catiônica.....	93
3.8.4.b	Resina aniônica.....	93
3.8.5	Purificação inicial de PepC-term liberado da fusão	94
3.8.5.a	Teste de clivagem com TEV protease	94
3.8.5.b	Purificação de PepC-term por IMAC	94
3.8.6	Testes de purificação de PepC-term	95
3.8.6.a	Cromatografia de fase reversa	95
3.8.6.b	Gel filtração	95
3.8.7	Caracterização Estrutural.....	95
3.8.7.a	Métodos computacionais	95
3.8.7.b	CD de 18N_PepC-term.....	97
3.9	Produção das formas longa e curta de proSmCD1 em <i>Pichia pastoris</i>	97
3.9.1	Subclonagem em vetor de expressão pPIC9k.....	97
3.9.1.a	Vetor e estratégia	97
3.9.2	Integração no genoma de <i>P. pastoris</i>	99
3.9.2.a	Preparação do DNA linear para transformação.....	99
3.9.2.b	Preparação de células competentes de <i>P. pastoris</i> GS115	100
3.9.2.c	Eletroporação.....	100
3.9.2.d	PCR de colônia	101
3.9.3	Seleção de clones múltipla-cópia com geneticina	101
3.9.4	Teste de expressão em meio mínimo.....	102
3.9.4.a	Pequena escala com tempo de indução máximo de 72h.....	102
3.9.4.b	Média escala com tempo de indução máximo de 120h.....	102
4	RESULTADOS	104
4.1	Análise da fração ligada a pepstatina-agarose de EAV.....	104
4.2	Avaliação do efeito do tratamento com inibidores comerciais de HIV-1 protease sobre a infecção experimental de camundongos com <i>S. mansoni</i>	106
4.3	Caracterização in silico da família gênica de SmAPs	108
4.4	Clonagem das formas longa e curta do zimogênio de SmCD1 (proSmCD1)	115
4.4.1	Otimização das condições de amplificação dos transcritos por RT-PCR.....	115
4.4.2	Clonagem em pGEM-t e sequenciamento	118
4.5	Formas longa e curta de pro-SmCD1 produzidas em <i>E. coli</i>	120

4.5.1	Clonagem em vetor de expressão pET28a	120
4.5.2	Estudo da expressão das isoformas de SmCD1 em <i>E. coli</i> BL21(DE3).....	123
4.5.3	Triagem de condições de solubilização de pro-SmCD1 presente em corpos de inclusão	124
4.5.4	Purificação por IMAC e determinação do estado de agregação em 4M de guanidina-HCl	128
4.5.5	Identificação de condições favoráveis à renaturação de pro-SmCD1 empregando desenho fatorial fracionado.....	130
4.5.6	Obtenção de anti-soros policlonais contra frações semi-purificadas de SmCD1 desnaturada.....	135
4.6	Domínio COOH-terminal da forma longa de SmCD1 produzido em <i>E. coli</i>	136
4.6.1	Construção pBuff5/pep_C-term e teste de indução em BL21(DE3).....	136
4.6.2	Otimização da superprodução solúvel da fusão contendo pep_C-term em <i>E. coli</i> BL21(DE3)	138
4.6.3	Desenvolvimento de uma estratégia de purificação de pepC-term.....	140
4.6.4	Caracterização de um subproduto de degradação proteolítica de pep_C-term por espectrometria de massas	149
4.6.5	Estudos estruturais em pepC-term.....	152
4.6.5.a	Métodos teóricos	152
4.6.5.b	Avaliação do grau de estruturação de 18N_pepC-term em solução por CD	155
4.6.6	Título e especificidade do anti-soro policlonal contra 18N_pep_C-term	158
4.7	Clonagem das formas longa e curta de SmCD1 em vetor de expressão em <i>P. pastoris</i>	159
4.7.1	Seleção por geneticina de clones múltipla cópia de <i>P. pastoris</i> /pro-SmCD1.....	163
4.7.2	Testes de expressão de pro-SmCD1 secretada no meio de cultura de <i>P. pastoris</i> GS115.....	164
4.8	Deteção de isoformas de SmCD1 em extratos de vermes adultos por immunoblot.....	168
5	DISCUSSÃO.....	170
5.1	Especificidade e Perspectiva da inibição de SmCD1 como estratégia esquistossomicida	170

5.1.1	Especificidade de substrato de SmCD1.2 analisada por modelagem molecular	170
5.1.2	Desempenho de inibidores comerciais de HIV-1 protease na quimioterapia experimental da esquistossomose mansônica	173
5.1.3	Identificação de novos inibidores	175
5.2	Complexidade na família de SmAPs	177
5.2.1	Cromatografia em pepstatina-agarose como estratégia de isolamento de aspartil proteases em <i>S. mansoni</i> e outros parasitos.....	177
5.2.2	Espécies detectadas por <i>immunoblots</i>	178
5.2.3	Análise dos transcritos de SmAPs	179
5.3	Estudos de renaturação de SmCD1 produzida a partir de corpos de inclusão em <i>E. coli</i>	180
5.3.1	Estratégias de renaturação de aspartil proteases.....	180
5.3.2	Solubilização de CIs de SmCD1 por uréia e guanidina-HCl.....	183
5.3.3	Planejamento estatístico de experimentos, aplicado ao renovelamento de proteínas	185
5.3.4	Viabilidade de produção de isoformas de SmCD1 funcionais a partir da renaturação de corpos de inclusão produzidos em <i>E. coli</i>	188
5.4	Estudos de expressão de SmCD1 em <i>P. pastoris</i>	189
5.4.1	Expressão de proteínas heterólogas secretadas	189
5.4.2	Viabilidade de produção das isoformas de SmCD1	191
5.5	A extensão C-terminal de SmCD1.1	194
5.5.1	Função e papel biológico	194
5.5.2	Expressão em <i>E. coli</i> e análises estruturais	196
5.5.3	Potencial de exploração no diagnóstico da esquistossomose	198
6	CONCLUSÕES	199

CAPÍTULO 2

1	INTRODUÇÃO.....	203
1.1	Estrutura e função de serinil proteases da família da quimiotripsina	203
1.1.1	Aspectos gerais e enovelamento.....	203
1.1.2	Mecanismo catalítico e cinética enzimática	207
1.1.3	Mecanismos de inativação.....	210
1.1.3.a	Inibidores covalentes sintéticos	210

1.1.3.b	Inibidores covalentes naturais: SERPINS	211
1.1.3.c	Inibidores não-covalentes protéicos	213
1.1.3.d	Inibidores não-covalentes sintéticos: afinidade de ligação ao subsítio S1	214
1.1.4	Origens da especificidade de substrato.....	217
1.1.4.a	Bases cinéticas e energéticas	217
1.1.4.b	Bases estruturais	220
1.2	Serininil proteases ativas sobre a coagulação sanguínea.....	224
1.2.1	A coagulação sanguínea	224
1.2.2	Estrutura e função de serininil proteases da cascata da coagulação	227
1.2.2.a	Fator VII	229
1.2.2.b	Fator IX	231
1.2.2.c	Fator X.....	234
1.2.2.d	Proteína C	238
1.2.2.e	Trombina e a modulação alostérica de SPCs pelo íon Na ⁺	240
1.2.3	Serininil proteases de venenos de serpentes (SVSP) ativas sobre a coagulação	245
1.2.3.a	Componentes do veneno interagindo com o sistema hemostático	245
1.2.3.b	SVSPs e enzimas trombina-símile do veneno de serpentes (SVTLEs)....	247
1.2.3.c	As SVTLEs de surucucu (<i>Lachesis muta</i>): Lmm-TLE e Lmr-47.....	250
1.2.3.d	Relações estrutura-função em SVSPs.....	252
1.2.4	Serininil proteases como fonte de novas abordagens para o tratamento e diagnóstico de distúrbios da coagulação	255
1.2.4.a	Importância das patologias afetando os fatores da coagulação... ..	255
1.2.4.b	Aplicações práticas de SVSPs na hemóstase.....	258
1.2.4.c	Potencial de serininil proteases no desenvolvimento de terapias para doenças do sistema hemostático.....	260
2	OBJETIVOS.....	268
3	MATERIAIS E MÉTODOS.....	269
3.1	Reagentes.....	269
3.2	Procedimentos gerais.....	269
3.2.1	SDS-PAGE	269

3.2.2	Estimativa do conteúdo protéico	270
3.2.3	Sequenciamento N-terminal por degradação de Edman automatizada.....	270
3.2.4	Bioinformática	270
3.3	Obtenção de Lmr-47 e BJ-48.....	270
3.3.1	Cromatografia de afinidade	270
3.3.2	Gel filtração	271
3.3.3	Ensaio de atividade	271
3.4	Efeito da temperatura na atividade de Lmr-47	272
3.4.1	Temperatura ótima de atividade amidásica	272
3.4.2	Estabilidade térmica	272
3.5	Caracterização da estrutura em solução de LMR-47.....	272
3.5.1	Espectropolarimetria de dicroísmo circular (CD)	272
3.5.2	Espectroscopia de Emissão da Fluorescência Intrínseca.....	273
3.5.3	Espalhamento de raios-X a baixos ângulos (SAXS).....	273
3.5.3.a	Preparo da amostra	273
3.5.3.b	Configuração experimental, coleta e tratamento preliminar de dados.....	273
3.5.3.c	Análise dos dados	274
3.5.3.d	Modelagem <i>ab initio</i>	275
3.6	Caracterização bioquímica adicional de LMR-47 e BJ-48.....	276
3.6.1	Especificidade do subsítio S1 e concentração do sítio-ativo.....	276
3.6.2	Interação com moléculas ativas sobre o sistema hemostático.....	276
3.6.2.a	Hirudina	277
3.6.2.b	Proteína C	277
3.6.2.c	PAR1	277
3.6.2.d	FpA	278
3.6.2.e	ATIII.....	278
3.7	Caracterização dos carboidratos de LMR-47 e BJ-48.....	278
3.7.1	Teste de ligação em Con A.....	278
3.7.2	Determinação enzimática do conteúdo e tipo de carboidratos em Lmr-47.....	279
3.7.3	Deglicosilação enzimática em condições nativas.....	279
3.7.3.a	Estudos preliminares	279
3.7.3.b	Deglicosilação “semi-preparativa”	279

3.7.3.c	Deglicosilação preparativa	280
3.7.4	Análise química dos carboidratos liberados	280
3.7.4.a	Carboidratos sialilados	281
3.7.4.b	Carboidratos neutros.....	281
3.7.5	Fracionamento do produto de deglicosilação nativa	281
3.7.5.a	Teste com cromatografia de afinidade.....	281
3.7.5.b	Teste com cromatografia de exclusão molecular (SEC)	281
3.7.5.c	Fracionamento por cromatografia de interação hidrofóbica (HIC).....	282
3.7.6	Caracterização das frações deglicosiladas de BJ-48.....	282
3.7.6.a	Estabilidade térmica	282
3.7.6.b	Inibição por SBTI e Infestina	282
4	Resultados.....	283
4.1	Determinação da temperatura ótima de atividade amidásica de Lmr-47.....	283
4.2	Especificidade do subsítio S1 de BJ-48 e Lmr-47.....	284
4.3	Interação com moléculas ativas sobre o sistema hemostático.....	285
4.4	Caracterização da fração carboidratos em Lmr-47 e BJ-48	286
4.4.1	Interação com Con A.....	286
4.4.2	Análise enzimática da porção carboidrato em Lmr-47.....	287
4.5	Análise da estrutura em solução de Lmr-47	289
4.5.1	CD.....	289
4.5.2	Fluorescência	289
4.5.3	SAXS	292
4.6	Análise do papel do carboidrato na estrutura e função de Lmr-47 e BJ-48.....	296
4.6.1	Deglicosilação enzimática de BJ-48 e Lmr-47 em condições não-desnaturantes	296
4.6.2	Análise química dos carboidratos liberados	300
4.6.3	Resolução das formas parcialmente deglicosiladas de Lmr-47 e BJ-48 por HIC	301
4.6.4	Estabilidade térmica	303
4.6.5	Resistência a inibidores protéicos	304
5	DISCUSSÃO	305
5.1	Especificidade das SVSPs BJ-48 e Lmr-47.....	305

5.1.1	Subsítio S1	305
5.1.1.a	BJ-48 e Lmr-47 são enzimas específicas para Arg em P1	305
5.1.1.b	Evolução da especificidade de substrato em SVSPs: similaridades com calicreínas teciduais.....	306
5.1.1.c	Bases estruturais para a especificidade por Arg na família da quimiotripsina	308
5.1.1.d	Uma possível aplicação biotecnológica para BJ-48 e Lmr-47?.....	310
5.1.1.e	Cromatografia de afinidade baseada em benzamidina: uma técnica para a purificação seletiva de amidases Arg-específicas de venenos de serpentes?.....	310
5.1.2	Interação com ligantes estendidos: Ações fisiológicas de Lmr-47 e BJ-48.....	312
5.2	Estrutura em solução de Lmr-47	315
5.3	Papel do carboidrato na estrutura e função de Lmr-47 e BJ-48	316
6	CONCLUSÕES	320
	REFERÊNCIAS	322
	ANEXO 1	367
	ANEXO 2	375
	ANEXO 3	393
	ANEXO 4	402
	ANEXO 5	415
	ANEXO 6	461
	Curriculum Vitae	472

Prelúdio

1 Proteases: função catalítica, importância biológica e regulação

Peptidase é o termo recomendado pela União Internacional de Bioquímica e Biologia Molecular para designar todas as enzimas proteolíticas. Entretanto, a denominação “protease” refere-se ao mesmo grupo de enzimas cuja função catalítica é hidrolisar ligações peptídicas. Estas últimas nada mais são do que ligações amida em meio a um arcabouço protéico. Amidas são estáveis à hidrólise sob condições ordinárias. A velocidade de hidrólise espontânea em água pura (k_{non}) varia de $1,25 \times 10^{-9} \cdot s^{-1}$ em pH 9 e 37 °C até 3×10^{-11} em pH neutro e 25 °C, enquanto que a reação catalisada pela enzima conversora de angiotensina (uma metaloprotease), apresenta proficiência catalítica, definida como $(k_{cat}/K_m)/k_{non}$ (expressa em unidades M^{-1}) de $5,3 \times 10^{13}$ (BABINE & BENDER, 1997). Este valor corresponde a uma estabilização do estado de transição da hidrólise da ligação amida de 18,8 kcal/mol. As peptidases recebem o prefixo endo- quando clivam ligações no meio de cadeias polipeptídicas, podendo neste caso serem chamadas também de proteinases, e o prefixo exo- quando clivam ligações em uma das extremidades da cadeia polipeptídica, compreendendo as chamadas amino- e carboxipeptidases.

Historicamente, o estudo das proteases iniciou-se com o interesse de fisiologistas e bioquímicos na digestão protéica em seres humanos. Por este motivo, as proteases digestivas contidas nas secreções do pâncreas e estômago estão entre as enzimas melhor caracterizadas e muito do conhecimento atual sobre estrutura e função de enzimas se deve ao estudo destas proteases. O interesse em proteases evoluiu a partir do conhecimento de que além das ações digestivas, estas enzimas estão envolvidas na regulação de muitos processos fisiológicos tais como a remoção de proteínas danificadas, maturação protéica e o processamento preciso de proteínas regulatórias (NEURATH, 1999). Exemplos recentes de descobertas do envolvimento de proteases em importantes processos patofisiológicos incluem: a via de processamento da proteína precursora amilóide na doença de Alzheimer (SCHIMMOLLER *et al.*, 2002; EHRMANN & CLAUSEN, 2004), vias neurodegenerativas no organismo modelo *Caenorhabditis elegans* (SYNTHICHAKI *et al.*, 2002), a via ubiquitina-proteasoma (DeMARTINO & SLAUGHTER, 1999), e caspases na apoptose (SOLARY *et al.*,

1998). Além disto, muitas enzimas proteolíticas de origem animal, vegetal e, principalmente, microbiana têm recebido importantes aplicações biotecnológicas, tais como: na formulação de detergentes de lavanderia, no processamento do couro e em diversos ramos da indústria alimentícia (laticínios, panificação, processamento da soja e obtenção de hidrolisados protéicos) (RAO *et al.*, 1998). Outras proteases têm sido usadas na terapêutica, tais como o fator IX na hemofilia (HORTELANO & OFUSO, 2000) e ativadores de plasminogênio como reteplase e alteplase na trombólise (MARTIN *et al.*, 1999; HANNA & SMALLING, 1998).

As implicações biológicas das reações proteolíticas são tão grandes que talvez, nenhuma outra classe de proteínas esteja envolvida numa diversidade tão grande de processos fisiológicos. Para isto, muitas enzimas proteolíticas mostram um alto grau de especificidade na identificação (ligação) e processamento de seus substratos. Entretanto, apesar de serem necessárias fisiologicamente, a maior parte das proteases são potencialmente perigosas ao seu microambiente natural, de forma que a atividade proteolítica precisa ser cuidadosamente controlada pela célula. Este controle é normalmente obtido através de diversos mecanismos, entre eles: (i) regulação da expressão/secreção e/ou ativação dos zimogênios; (ii) degradação específica das enzimas maduras; (iii) pelo bloqueio da atividade proteolítica por inibidores; (iv) compartimentalização, ou seja, co-localização da protease e do substrato no mesmo compartimento celular e (v) associação com domínios adaptadores, isto é, domínios não-proteolíticos que conferem especificidade na interação enzima-substrato.

Proteases não atacam seus substratos protéicos aleatoriamente. Ao contrário, o reconhecimento de substrato depende de uma variedade de fatores, dentre os quais o mais marcante é a energia livre de ligação entre este e a protease. Entretanto, independentemente do mecanismo catalítico empregado por uma protease, alguns princípios fundamentais parecem ditar o reconhecimento molecular efetuado por estas enzimas. Talvez o mais importante destes princípios seja a idéia de complementaridade entre reentrâncias na superfície da protease (bolsas ou subsítios de especificidade) e as cadeias laterais dos substratos peptídicos, conforme formulado por I. Schechter e A. Berger, há quase quarenta anos atrás (SCHECHTER & BERGER, 1967). Utilizando o sítio ativo da papaína como modelo na demonstração deste conceito, estes autores

cunharam os termos P4-P4' e S4-S4' para designar, respectivamente as cadeias laterais dos resíduos da proteína-alvo e os subsítios de ligação na protease. Conforme esquematizado na **Figura 1**, as cadeias laterais do substrato são designadas a partir da ligação lábil, P1-P1', recebendo números maiores quanto mais distantes estiverem desta. O acúmulo de estruturas 3D de inúmeros complexos protease-inibidor permitiu aprofundar os princípios de reconhecimento molecular em proteases através da observação de que estas enzimas universalmente reconhecem uma conformação estendida (em fita- β) em todos seus inibidores peptídicos (análogos de substratos) e não-peptídicos, salvo raras exceções (TYNDALL & FAIRLIE, 1999). Conseqüentemente, os subsítios de especificidade das proteases estão aproximadamente dispostos alternadamente na fenda de ligação ao substrato (**Fig. 1B**). Não por acaso, a formação de uma folha- β com a protease fornece máxima exposição das cadeias laterais e de grupos receptores/doadores de ligações de hidrogênio presentes na cadeia principal do substrato protéico.

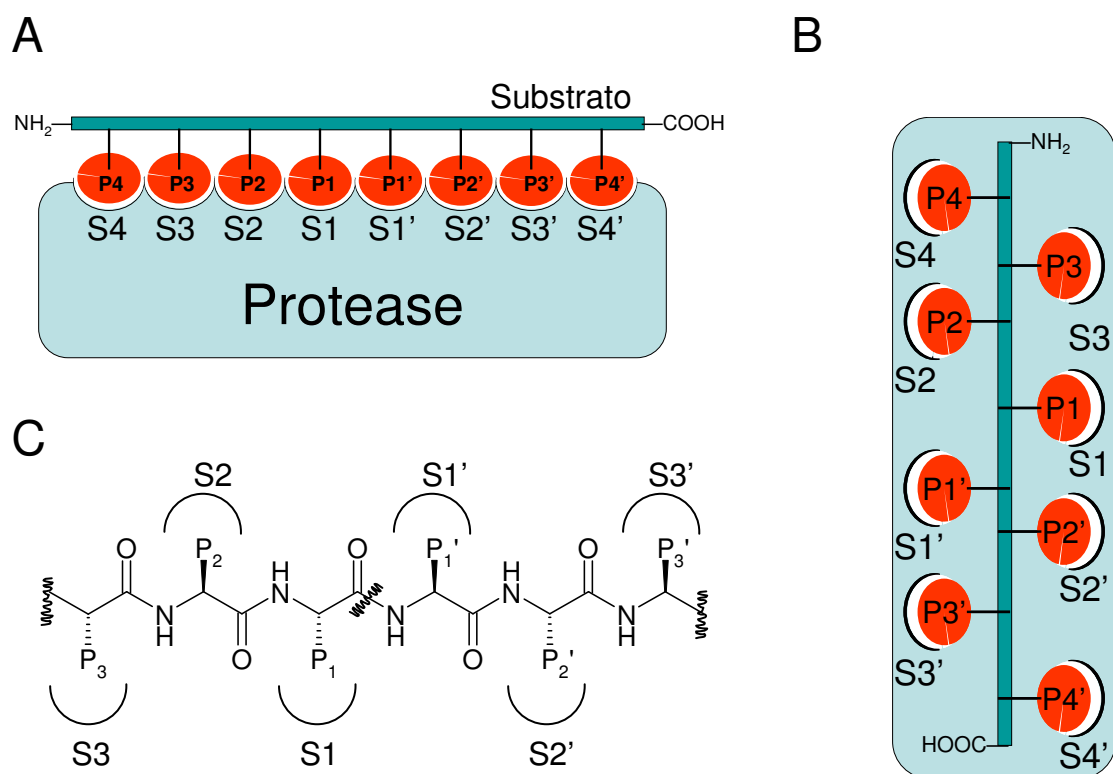


Fig. 1: Representação esquemática do reconhecimento molecular de proteases. A- modelo original e nomenclatura adotada por Shechter e Berger (1967). B- incorporação no modelo em A do modo de ligação estendido de substratos na fenda de ligação de proteases. C- Estrutura de um substrato peptídico P1-P3.

Outro pré-requisito óbvio para um substrato ser atacado por uma protease é a acessibilidade de superfície dos sítios potenciais de clivagem. Sítios de clivagem que estão normalmente escondidos na conformação nativa de uma proteína podem se tornar acessíveis em função de desnaturação ou perda de interação com um ligante qualquer. Este princípio fundamental da regulação da atividade proteolítica constitui um método elegante de prevenir a degradação de proteínas enveladas. Outro fenômeno relevante é a seqüencialidade dos eventos de degradação. Proteases envolvidas em processos regulatórios, por exemplo, frequentemente trabalham de forma combinada quando mais de uma protease está envolvida no processamento de um dado substrato (EHRMANN & CLAUSEN, 2004).

Conforme mencionado acima, um meio importante de controle da atividade proteolítica ocorre ao nível da ativação dos zimogênios. Estes são formas precursoras inativas das enzimas que são convertidos à forma madura ativa através da remoção proteolítica do pro-peptídeo. A ativação do zimogênio foi por muito tempo considerada uma resposta rápida e irreversível a um sinal relevante que leva à protease constitutivamente ativa. Entretanto, relatos recentes indicam ativação reversível do zimogênio onde outras proteínas ou fatores externos tais como sais ou temperatura podem ligar e desligar a atividade proteolítica (EHRMANN & CLAUSEN, 2004).

Mais recentemente, um outro papel tem sido reivindicado para os pro-peptídeos. Mais do que simples inibidores intramoleculares, tem sido demonstrado que muitos destes pequenos domínios são essenciais para o correto envelamento do domínio catalítico (BRYAN, 2002). A maior parte dos pro-peptídeos ocorrem como extensões no N-terminal longas o suficiente para ter um envelamento independente do restante da proteína.

2 Estratégias catalíticas e esquema de classificação corrente

Enzimas proteolíticas podem ser divididas em duas grandes classes de acordo com a estratégia catalítica adotada. Uma classe emprega uma molécula de água ligada à enzima como nucleófilo no ataque à carbonila amídica da ligação lábil. Esta molécula de água pode ser tanto ativada pela coordenação a um cátion de metal de transição, como Zn^{2+} , no sítio ativo de metaloproteases, quanto pela polarização imposta por dois

aspartatos no sítio ativo de aspartil proteases. A segunda grande classe de enzimas proteolíticas atua ativando um átomo nucleofílico de um resíduo de aminoácido para dar início à quebra da ligação amida do substrato. Até o momento, somente os grupos hidroxílicos de um resíduo de serina ou treonina e sulfidril de um resíduo de cisteína foram identificados como nucleófilos ativados, dando origem às classes das serinil, treonil e cisteinil proteases, respectivamente. Estas classes de enzimas catalisam a hidrólise da ligação peptídica em duas etapas. Primeiro, o átomo nucleofílico ativado da enzima ataca a carbonila amídica da ligação lábil para quebrar a ligação C-N, formando um intermediário acilado (um éster ou tioéster). Este intermediário acilado da enzima é hidrolisado pela água numa segunda etapa pra completar o processo de hidrólise. Apesar de se apresentarem fisiologicamente como peptidases, a maior parte destas enzimas também pode catalisar a hidrólise de ésteres, tioésteres, anidridos e cloretos de acila. O contrário, contudo, não é verdadeiro, pois enzimas que se apresentam fisiologicamente como amidases são incapazes de clivar ligações peptídicas, apesar de estas enzimas apresentarem um ancestral comum às aspartil proteases (KOBAYASHI *et al.*, 1997).

Talvez, o maior progresso no estudo de proteases tenha sido o estabelecimento de uma classificação racional baseada no conhecimento das classes catalíticas e na homologia reconhecida pela similaridade na seqüência, proposta por Rawlings & Barret (1993). Nesta classificação, clãs reúnem as famílias evolutivas de peptidases assim derivadas, que mostram algum sinal de relação (homologia) distante, a julgar por semelhanças na estrutura 3D ou na disposição dos resíduos catalíticos ao longo da seqüência. Originalmente, apenas 4 tipos catalíticos foram considerados (M-metalo, A-aspartil, S-serinil e C-cisteinil proteases) mais um quinto grupo para mecanismo desconhecido (U). Treonil proteases envolvidas no proteasoma, foram identificadas posteriormente (revisado por DeMARTINO & SLAUGHTER, 1999; GROLL *et al.*, 2005). As serinil peptidases reuniram o maior número de famílias, com a grande maioria pertencendo ao clã SA. Este clã, também conhecido como o clã da quimiotripsina é caracterizado por uma tríade catalítica do tipo Asp, His e Ser. As aspartil proteases, compreendendo o terceiro maior grupo de enzimas (após cisteinil peptidases) puderam ter suas duas grandes famílias (família A1 de enzimas pepsina-

símile e família A2 de enzimas retrovirais) reunidas no clã AA (dois aspartatos catalíticos).

A classificação de Rawlings & Barret deu origem ao banco de dados de peptidases denominado MEROPS ([RAWLINGS & BARRET, 1999](#)). Para a construção deste banco, os autores consideraram que normalmente há uma relação um pra um entre peptidases e seus genes, e trataram produtos parálogos de diferentes genes num organismo como peptidases diferentes. Critérios subjetivos, como semelhanças nas propriedades catalíticas (pH e temperaturas ótimas, preferência de substrato, etc.) foram utilizados para classificar genes ortólogos como variações de uma mesma peptidase em diferentes organismos. Outro aspecto importante da construção do banco de dados é que somente a parte da proteína responsável pela atividade catalítica (a unidade peptidase) foi utilizada nas comparações de seqüências que deram origem às famílias evolutivas. Isto é relevante, pois muitas peptidases são estruturas quiméricas, contendo domínios adicionais não-catalíticos (incluindo pré- e pro-peptídeos) que são compartilhados com outros grupos de proteínas. Com o sequenciamento de genomas inteiros de muitos organismos complexos, o número de seqüências de peptidases disponível aumentou enormemente. Estima-se que cerca de 2% dos genes estruturais codifiquem proteases ([RAWLINGS & BARRET, 2000](#)). Desta forma, a construção e manutenção do banco de dados MEROPS representou uma grande contribuição para o entendimento da estrutura e função de proteases.

3 Proteases como alvo de desenvolvimento de novos fármacos

Tendo em vista a diversidade de processos patofisiológicos em que estão envolvidas, as vantagens de se concentrar em proteases como alvos para quimioterapia são evidentes. Além disso, há uma substancial quantidade de informações sobre o desenvolvimento de inibidores da protease do vírus da imunodeficiência humana tipo 1 (HIV-1) que apresentam uma extensiva sobreposição metodológica com o desenho de inibidores de proteases humanas e de outros agentes infecciosos (revisado por [BRIK & WONG, 2003](#)). O emprego de ciclos de desenho de fármacos baseado na estrutura de seus alvos, como o da **Figura 2** tem o potencial de resultar no desenho de inibidores de proteases cada vez mais potentes e seletivos.

A inibição seletiva de proteases de mamíferos que se tornam aberrantemente expressas ou desreguladas durante doenças é uma importante meta no desenvolvimento e desenho de novos fármacos. As principais razões para este interesse incluem os dados acumulados na enzimologia, bioquímica, fisiologia, estrutura 3D e mecanismo de inibição destas enzimas. O número de proteases humanas sob investigação como alvos para drogas aproximadamente dobrou entre 1998 e 2000. O projeto genoma humano certamente teve participação crucial neste aumento recente de alvos disponíveis. Uma anotação preliminar do genoma humano mostrou a existência de 498 seqüências codificando proteases (1,6 % do total de proteínas preditas), superando outras classes tradicionais de proteínas exploradas pela indústria farmacêutica, como as proteína-quinases (437), receptores acoplados à proteína G (369) e receptores de esteróides (35) (SOUTHAN, 2001). Atualmente, 14 % das proteases humanas são investigadas como alvos potenciais para quimioterapia, incluindo: proteasoma no câncer (treonil proteases, ZAVRSKI *et al.*, 2005), catepsinas S e K na artrite e osteoporose (cisteinil proteases; YASUDA *et al.*, 2005), Caspase-3 (cisteínil proteinase) na isquemia cerebral (LOETSCHER *et al.*, 2001), TACE (enzima conversora do fator de necrose tumoral, uma metaloprotease) na artrite, câncer e diabetes (MOSS *et al.*, 2001), metaloproteinases da matriz no câncer (MMPs; FODA & ZUCKER, 2001), catepsina D na doença de Alzheimer (PAPASSOTIROPOULOS *et al.*, 2001) e ativador uroquinase de plasminogênio (u-PA) no câncer (JEDINAK & MALIAR, 2005).

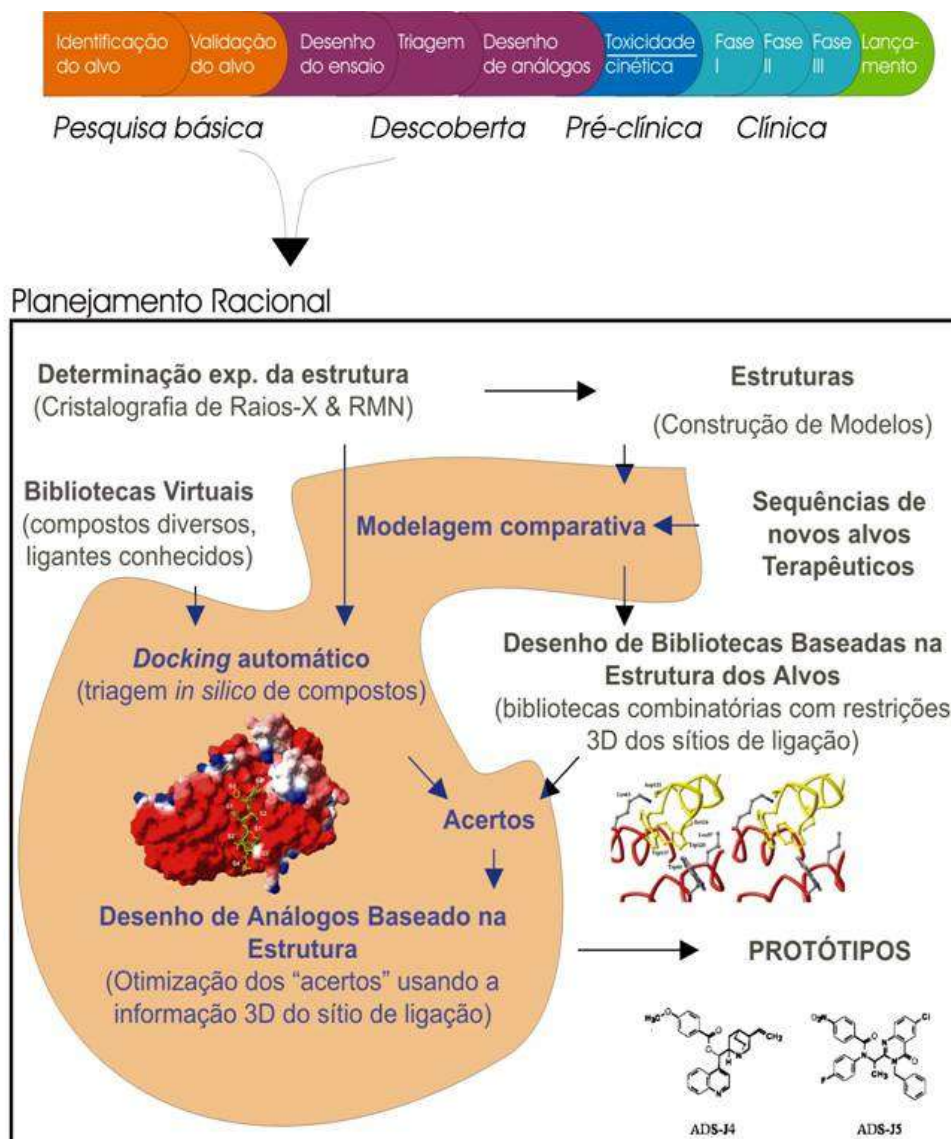


Fig. 2: Ciclo de desenvolvimento de novos fármacos baseados na estrutura de seus alvos biológicos.

A inibição seletiva de proteases de agentes infecciosos atuando nas células ou no espaço extracelular do hospedeiro é uma estratégia interessante no combate de infecções por diversos tipos de microorganismos. Ao retardar as velocidades de replicação destes organismos invasores, acaba por assistir os mecanismos imunológicos de defesa envolvidos em sua erradicação. Entre as doenças infecciosas que apresentam proteases microbianas sendo investigadas como alvos para quimioterapia estão: hepatite C (serinil protease, [BRETNER, 2005](#)), candidíase (aspartil proteinases; [HOEGL *et al.*, 1996](#)), malária (aspartil proteinases; [CARROLL *et al.*, 1998](#); [RATHORE *et al.*, 2005](#)) e outras parasitoses, e.g. doença de chagas (cisteinil proteinases; [McKERROW, 1999](#)). Também foi proposto que a inibição de enzimas da classe das serinil, cisteinil e metalo proteases

poderia ser utilizada no combate de infecções causadas por bactérias de diversas espécies como, por exemplo, *Clostridium sp.*, *Escherichia coli* e *Pseudomonas aeruginosa* (TRAVIS & POTEPA, 2000). Alguns pesquisadores têm investigado o papel da região pré-pró de zimogênios de proteases como inibidores naturais específicos destas enzimas (KHAN *et al.*, 1999). Este tipo de inibição parece envolver mecanismos diferentes dos apresentados pelos inibidores existentes, baseados na ligação sob conformação estendida no sítio ativo de proteases (TYNDALL & FAIRLIE, 1999).

Em suma, o estudo de proteases ainda tem muito a revelar sobre a função biológica destas enzimas e isto certamente contribuirá para a descoberta de novas abordagens terapêuticas para muitas das doenças que afligem o homem.

Capítulo 1

Aspartil proteases catepsina D- símile de *S. mansoni*

1 INTRODUÇÃO

1.1 *Estrutura e função de aspartil proteases do clã da pepsina*

3.1.1 Aspectos gerais, evolução e enovelamento

As aspartil proteases (EC 3.4.23) perfazem uma classe de enzimas amplamente distribuída, sendo encontrada em vertebrados, fungos, plantas e retrovírus (MEROPS; RAWLINGS & BARRET, 2000). Todas as aspartil proteases conhecidas são endopeptidases (proteínases) caracterizadas por um pH ótimo na faixa ácida (pH 3-5), inibição pela pepstatina, o uso de duas cadeias laterais de ácido aspártico no mecanismo catalítico e especificidade por substratos contendo resíduos hidrofóbicos em P1-P1' (TANG & WONG, 1987; DAVIES, 1990; KAY & DUNN, 1992; DUNN & HUNG, 2000; DUNN, 2002). Além disso, todas as aspartil proteases de que se tem conhecimento são sintetizadas como zimogênios inativos, contendo um segmento N-terminal adicional, de aproximadamente 45-55 resíduos de aminoácidos, que é clivado e separado durante a ativação (TANG & WONG, 1987; DAVIES, 1990; DUNN, 2002).

Como em qualquer outra família de proteinases, muitas funções distintas são executadas por seus membros. As enzimas gástricas pepsina, gastrisina e quimosina (o principal componente do coalho na fabricação de queijo), encontradas nas secreções gástricas de mamíferos e envolvidas na digestão de muitos vertebrados, são as mais conhecidas (SOGAWA *et al.*, 1983; FOLTMAN *et al.*, 1977; revisado por FOLTMAN *et al.*, 1981). Outras aspartil proteases encontradas em humanos e outros vertebrados incluem catepsinas D e E, beta-secretase (BACE) e renina. Catepsina D, uma importante enzima lisossomal, também degrada proteínas, mas o faz intracelularmente (FAUST *et al.*, 1985). Catepsina E é não-lisossomal e intracelular, estando associada ao processamento de antígenos e prohormônios (AZUMA *et al.*, 1989). BACE está envolvida no processamento da proteína precursora amilóide no cérebro humano (SINHA *et al.*, 1999). Renina (HOBART *et al.*, 1984), realiza uma clivagem específica no angiotensinogênio para produzir o decapeptídeo que é o precursor da angiotensina II (um octapeptídeo que é um importante fator no controle da pressão sanguínea). Aspartil proteases também são encontradas em invertebrados, plantas, platelmintos, protozoários, fungos e vírus, mas não em bactérias. Nos parasitas causadores da malária,

Plasmodium sp., aspartil proteases, conhecidas como plasmepsinas estão envolvidas no metabolismo da hemoglobina (FRANCIS *et al.*, 1997). Nos fungos, algumas aspartil proteinases desempenham um papel na esporulação e atuam como fatores de virulência em algumas espécies infecciosas (REICHARD *et al.*, 1994; HOEGL *et al.*, 1996). Nos retrovírus, muitas das proteínas são sintetizadas como poliproteínas, que durante a ativação do vírus são clivadas pela protease retroviral (TOH *et al.*, 1985).

A maior parte das aspartil proteinases eucarióticas são proteínas monoméricas, com um peso molecular de 35-40 kDa. Estas enzimas possuem seqüências características na região dos dois resíduos catalíticos de ácido aspártico: (hidrofóbico, geralmente Phe)-Asp²⁵-Thr-Gly-Ser, no domínio N-terminal e um domínio C-terminal correspondente (hidrofóbico)-Asp²¹⁵-Thr-Gly-Ser/Thr (numeração conforme a pepsina suína) (Fig. 3). As proteinases retrovirais são homodiméricas, com cada monômero consideravelmente menor do que o domínio catalítico de enzimas eucarióticas, tendo menos que 130 aminoácidos de comprimento. Estas também contêm a seqüência Asp-Ser/Thr-Gly, o que levou Toh *et al.* (1985) a proporem que estas proteinases pertencessem ao clã da pepsina. Pearl & Taylor (1987) alinharam as seqüências de 16 proteinases retrovirais com cada um dos dois domínios de proteinases pepsina-símiles, baseados em padrões similares de conservação de resíduos e predições de estrutura secundária. Estes autores concluíram que proteinases retrovirais poderiam ser modeladas como “versões parsimoniosas” das proteinases eucarióticas e associadas em dímeros para fornecer uma enzima ativa.

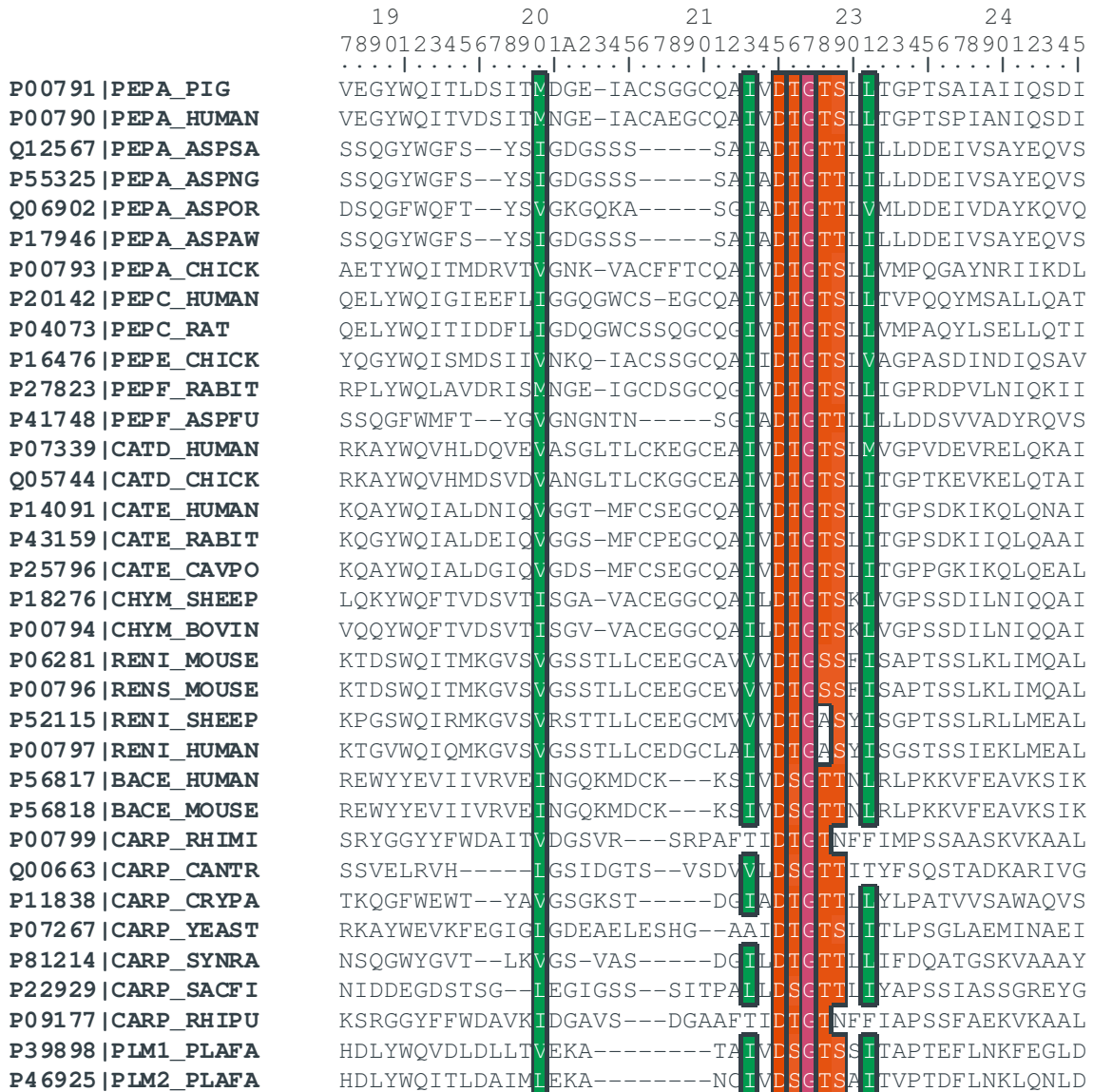


Fig. 3: Alinhamento de aspartil proteases ortólogas na região do segundo aspartato catalítico (Asp²¹⁵). Proteases são identificadas pelo código de acesso no SwissProt e numeradas de acordo com a pepsina suína. Acrônimos das proteases: PEPPI, C, E e F = pepsina A, C, E e F; CATD, E = catepsina D e E; CHYM = quimosina; RENI, S = renina comum e isoforma submandibular; BACE = beta-secretase; CARP = proteases fúngicas; PLM1, 2 = plasmepsinas 1 e 2. Acrônimos das espécies: PIG = *Sus scrofa* (porco); HUMAN = *Homo sapiens*; ASPFU = *Aspergillus fumigatus*; ASPSA = *A. saitoi*; ASPNG = *A. niger*; ASPOR = *A. oryzae*; ASPAW = *A. awamori*; CHICK = *Gallus gallus* (galinha); CAVPO = *Cavia porcellus* (porco da Índia); SHEEP = *Ovis aries* (ovelha); BOVIN = *Bos taurus* (buzina); MOUSE = *Mus musculus* (camundongo); RHINI = *Rhizomucor miehei*; CANTR = *Candida tropicalis*; YEAST = *Saccharomyces cerevisiae* (levedo de pão); SYNRA = *Syncephalastrum racemosum*; SACFI = ; RHIPU = *Rhizomucor pusillus*. Resíduos coloridos são conservados em mais de 90% das sequências.

Atualmente, perto de 400 estruturas de aspartil proteinases foram depositadas no *Protein Data Bank* (PDB; BERMAN *et al.*, 2000), sendo mais de 50% deste número contribuído por complexos HIV-1 protease-inibidor e apenas pouco mais de 10% de mamíferos (DUNN, 2002). A estrutura secundária geral dos membros da família das aspartil proteinases consiste quase inteiramente de folhas β -pregueadas com poucas α -hélices. A molécula é bilobulada com dois domínios de estrutura similar relacionados por um eixo de pseudosimetria (ou simetria, no caso da HIV-1 protease) C_2 (Fig. 4A). Uma ampla fenda, com aproximadamente 40 Å de extensão, atravessa inteiramente a molécula e separa os dois domínios. Sob estes, uma extensa folha β -pregueada de 6 fitas forma a base da molécula. A HIV protease possui esta estrutura β -pregueada típica, mas apresenta 4 ao invés de 6 fitas interdigitantes (Fig. 4B).

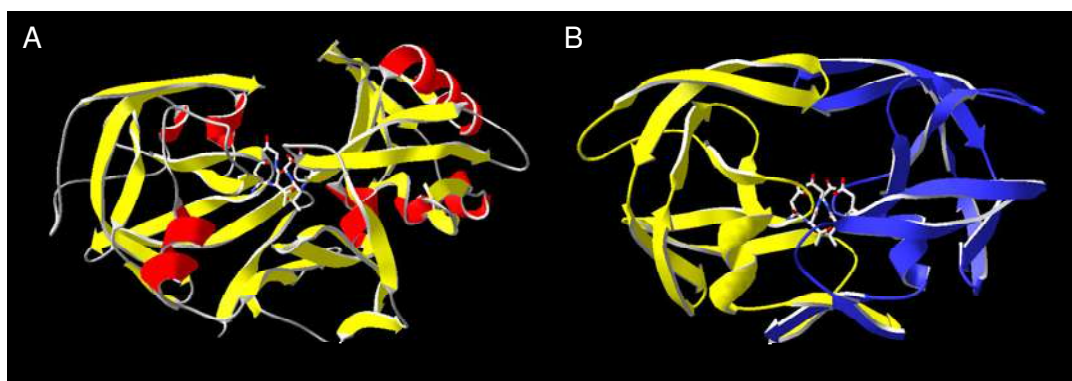


Fig. 4: Enovelamento de aspartil proteinases eucarióticas e retrovirais. A- Pepsina suína (código PDB 4PEP); fitas- β em amarelo e α -hélices em vermelho; os resíduos Asp-Thr-Gly do sítio ativo estão representados em “palitos” coloridos em CPK; o assoalho do sítio ativo, ao centro, está em baixo e os *flaps* acima. B- Aspartil protease do HIV (código PDB 1OHR); os monômeros estão coloridos em azul e amarelo; notar estrutura simétrica do homodímero, com o eixo de pseudo-simetria C_2 , vertical e no plano da página.

Como o alinhamento da **Figura 3** denota, a seqüência de aminoácidos ao redor dos resíduos do sítio ativo é muito semelhante em todas as aspartil proteinases. Além disto, conforme mencionado acima, há um eixo de pseudo-simetria dividindo a molécula em dois lóbulos enovelados semelhantemente. Esta simetria no enovelamento pode ser estendida à molécula inteira para todos os membros da família de estrutura conhecida. Mais ainda, dentro de cada lóbulo, outro eixo de simetria C_2 pode ser encontrado, apesar de menos evidente. Desta forma, mesmo antes da elucidação da estrutura homodimérica da HIV-1 protease (NAVIA *et al.*, 1989; WLODAVER *et*

al.,1989) já se acreditava que os genes de aspartil proteases tinham surgido de um gene ancestral com aproximadamente $\frac{1}{4}$ do tamanho dos genes atuais, através de pelo menos dois eventos de duplicação e fusão de genes separados (TANG *et al.*, 1978; TANG & WONG, 1987). Desta forma, a estrutura de cada domínio pode ser descrita como uma série de elementos de estrutura secundária que são repetidos dentro do domínio para construir a estrutura, conforme ilustrado com a pepsina suína na **Figura 5**.

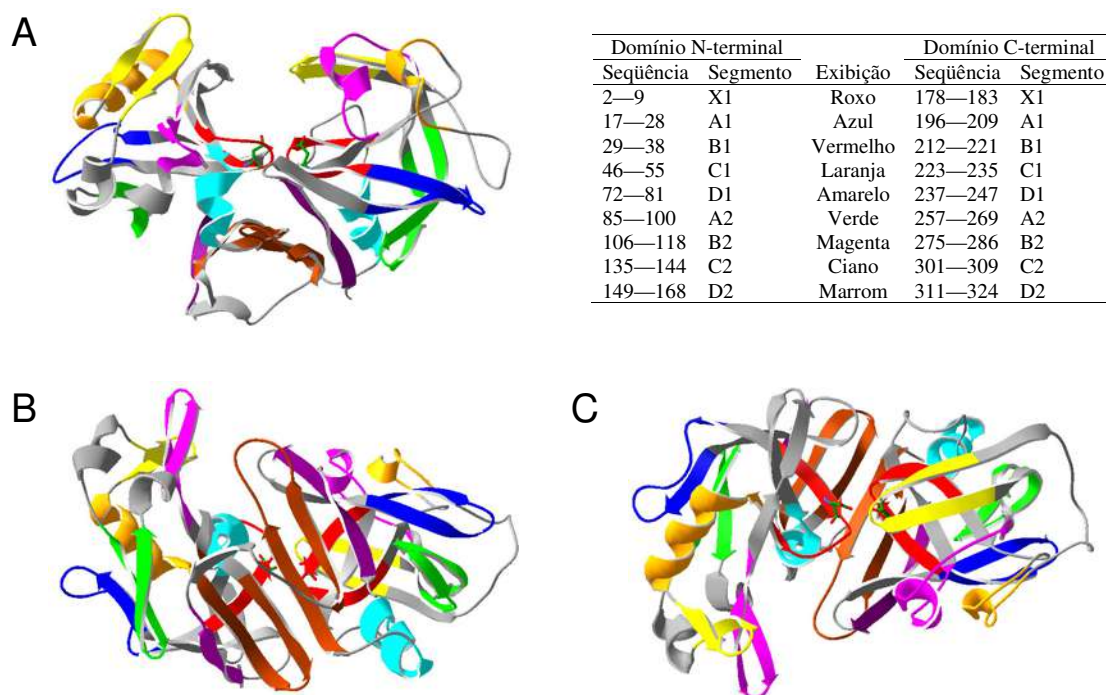


Fig. 5: Elementos estruturais e simetria no enovelamento de aspartil proteases do clã da pepsina. A estrutura nativa da pepsina suína (código PDB 4PEP) foi colorida conforme indicado no quadro no canto superior direito, com o domínio N-terminal à direita de cada painel. Elementos equivalentes (por simetria) dentro do próprio domínio recebem a mesma letra, e.g. A1 e A2, enquanto que elementos equivalentes entre os domínios N- e C-terminal recebem a mesma cor. A- Visão ao longo da fenda de ligação ao substrato, destacando o *flap* (à direita); B- Visão de baixo para destacar o assoalho do sítio ativo; C- Visão de cima para destacar a fenda de ligação ao substrato.

Os resíduos de ácido aspártico catalíticos estão localizados nos términos das alças em formato da letra grega psi (ψ), que se estendem de cada domínio e estão relacionadas umas às outras por um eixo pseudo-C₂. Os dois grupos carboxílicos no sítio ativo estão conectados por uma complexa rede de ligações de hidrogênio que faz uso da seqüência Asp-Thr-Gly (**Fig. 6**). As duas treoninas desempenham um papel crucial, onde o grupo hidroxila de uma faz ligações de hidrogênio com o nitrogênio do grupo amida da outra. Estas também estão em distância apropriada para formar ligações

de hidrogênio com os oxigênios carbonílicos dos dois resíduos hidrofóbicos precedendo os aspartatos catalíticos. Poder-se-ia esperar que esta rede de ligações de hidrogênio tornasse a região do sítio ativo da enzima bem rígida e os baixos fatores B cristalográficos observados nesta região confirmam este fato. Uma exceção é o *flap*, uma alça unindo duas fitas- β (em inglês, *hairpin loop* ou grampo de cabelo), que na enzima nativa possui um alto fator B, indicando alta mobilidade. O *flap* projeta-se sobre a fenda do sítio ativo fechando-se sobre inibidores ligados à protease. Vale ressaltar que diferentemente das aspartil proteases de mamíferos e fungos, que apresentam somente um *flap*, a protease de HIV apresenta dois destes.

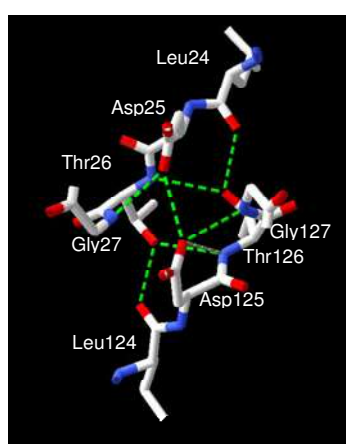
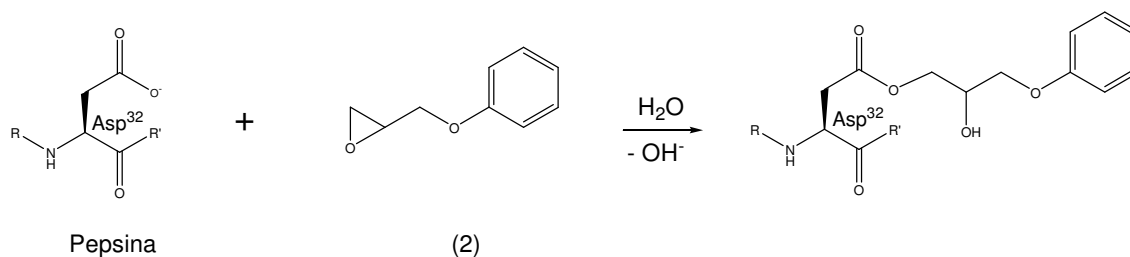
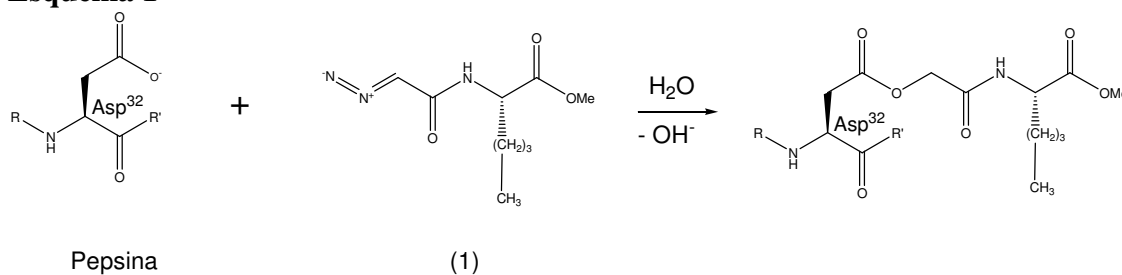


Fig. 6: Rede de ligações de hidrogênio (pontilhado verde) formadas por átomos da cadeia principal e das cadeias laterais dos resíduos do sítio catalítico da aspartil protease do vírus HIV-1 (código PDB 5HVP). Resíduos em palito, coloridos em CPK (vermelho – oxigênio, azul – nitrogênio e branco carbono).

1.1.1 Mecanismo catalítico e reconhecimento de substrato revelados por estudos com inibidores

Conforme ilustrado no **Esquema 1**, os resíduos de aspartato pertencentes ao aspartato catalítico de aspartil proteases foram originalmente identificados pela reação com éster metílico de diazoacetilnorleucina ou DAN (**1**; [RAGAGOPOLAN *et al.*, 1966](#)) e com 1,2-epoxi-3-(p-nitrofenoxy)propano ou EPNP (**2**; [CHEN & TANG, 1972](#)).

Esquema 1



Pepstatina (3), um hexapeptídeo isolado de *Streptomyces sp.* (UMEZAWA *et al.*, 1970), contendo o aminoácido incomum, ácido (3S,4S)-4-amino-3-hidróxi-6-metileptanóico ou simplesmente estatina (4), inibe universalmente todas as aspartil proteases do clã da pepsina. A análise de inibidores ligados às primeiras estruturas determinadas de aspartil proteinases fúngicas forneceu a base para a modelagem do substrato, o que por sua vez pôde levar a uma avaliação de possíveis mecanismos de ação em termos tridimensionais. A **Figura 1.4** ilustra a disposição dos dois grupos carboxílicos dos aspárticos catalíticos, em relação à molécula de água co-cristalizada em três enzimas fúngicas (SUGUNA *et al.*, 1987; JAMES & SIELECKI, 1985; PEARL & BLUNDELL, 1984). Neste arranjo, a molécula de água central faz ligações de hidrogênio com o átomo de oxigênio externo (O δ 2) de um grupo carboxilato e com o átomo de oxigênio interno (O δ 1) do outro grupo carboxilato.

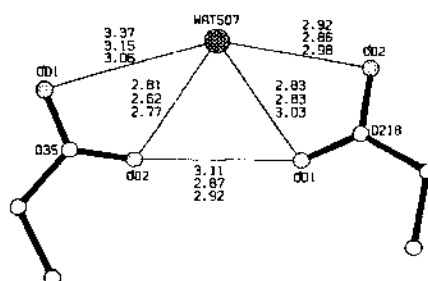
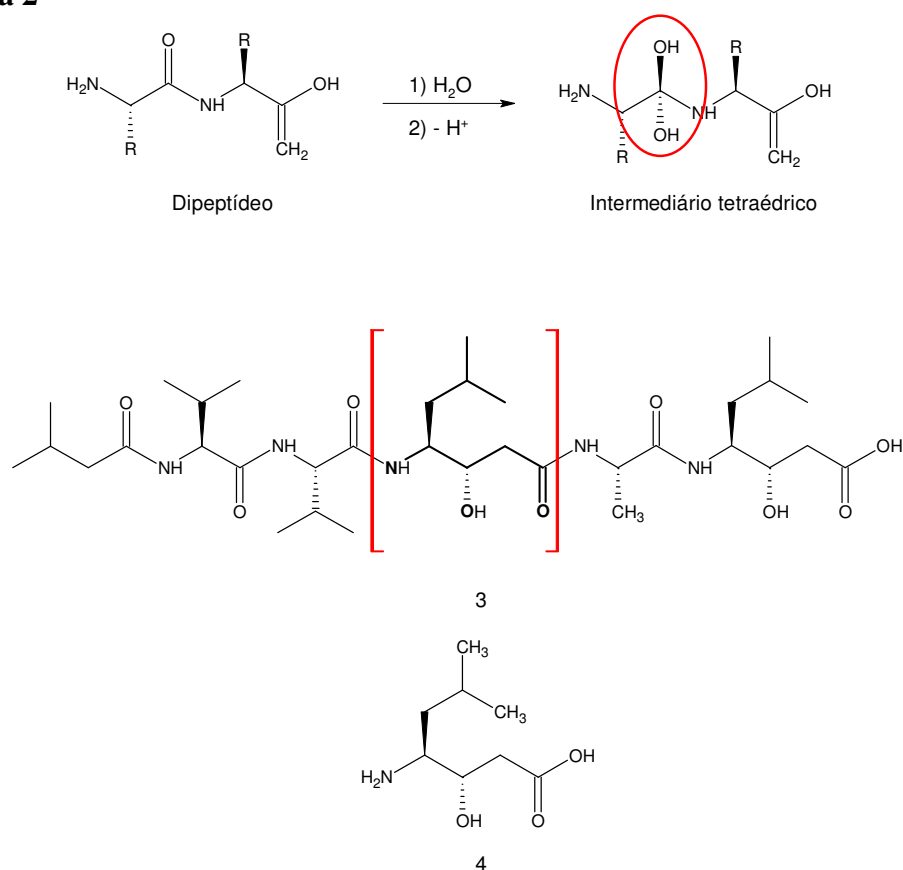


Fig. 7: Os dois grupos carboxílicos de Asp³⁵ (D32 na pepsina) e Asp²¹⁸ (D215 na pepsina) com sua molécula de água associada. Os três conjuntos de distâncias (em Å) provém de rhizopuspepsina (SUGUNA *et al.*, 1987), penicillopepsina (JAMES & SIELECKI, 1985) e endothiapepsina (PEARL & BLUNDELL, 1984), respectivamente. Adaptado de Vacca *et al.* (1994).

Os resultados de cristalografia dos complexos enzima-inibidor de proteases fúngicas mostraram que a hidroxila do primeiro resíduo de estatina liga-se entre os Oδ2 de ambos resíduos aspartato (32 e 215), efetivamente substituindo a molécula de água observada nesta posição das enzimas nativas. Em função destes resultados, logo se propôs que o mecanismo de ação da pepstatina estivesse associado à semelhança do resíduo de estatina com um suposto intermediário tetraédrico na hidrólise da ligação peptídica por aspartil proteinases (**Esquema 2**).

Esquema 2



Alguns exemplos tradicionais de isómeros do estado de transição (IET) que foram desenvolvidos como inibidores da renina estão ilustrados na **Figura 8**. Todas as estruturas mimetizam o intermediário tetraédrico que ocorre durante a hidrólise de um dipeptídeo. No caso da estatina, há um encurtamento de um átomo de carbono na cadeia

principal em relação ao estado de transição. Isto faz com que as demais cadeias laterais do inibidor sofram um deslocamento de aproximadamente 1,5-2,0 Å em relação às posições esperadas para a porção C-terminal do inibidor (e.g., P1'-P4').

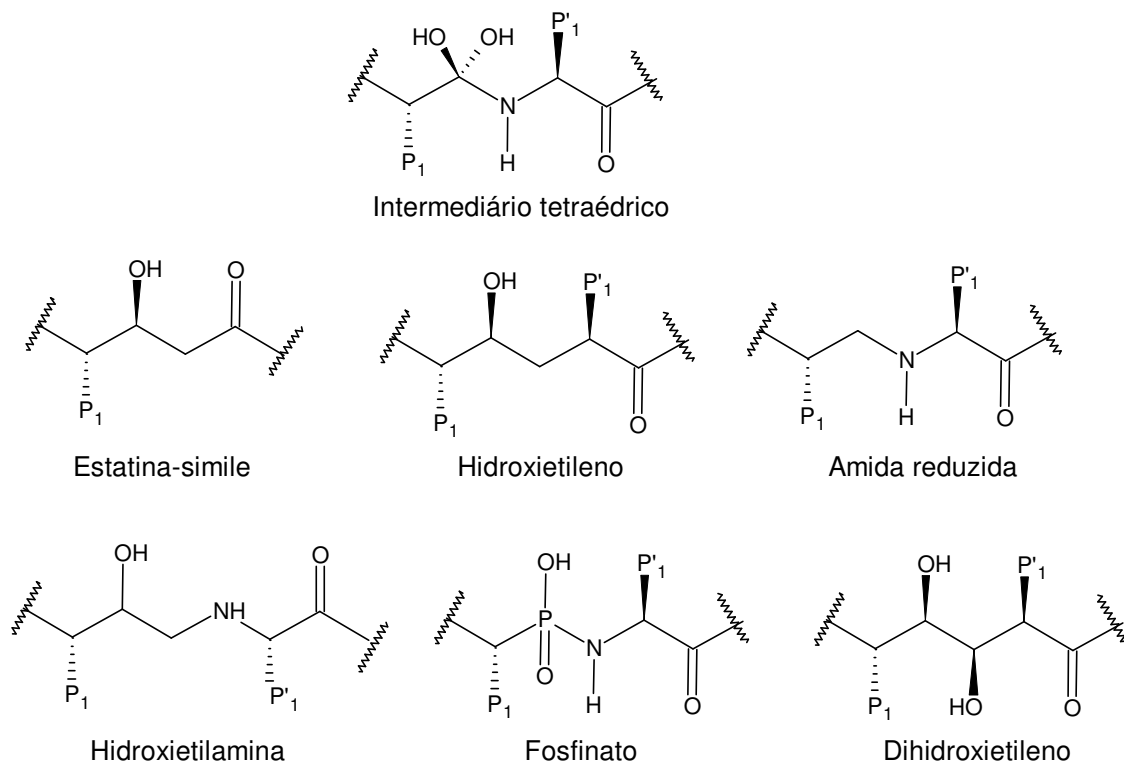


Fig. 8: Isósteros de estado de transição (IET). IETs mimetizam o intermediário tetraédrico formado na hidrólise da ligação peptídica por aspartil proteases (topo).

A partir dos dados estruturais disponíveis, [Suguna *et al.* \(1987\)](#) propuseram um mecanismo de hidrólise da ligação peptídica baseado na premissa de que se o substrato liga-se à enzima semelhantemente ao inibidor, a ligação lábil teria que ser torcida para fora da planaridade para que o ataque à carbonila ocorresse. Esta torção reduziria o caráter de ligação dupla da ligação peptídica, causando a adoção de um arranjo mais piramidal pelo átomo de nitrogênio (como numa amina secundária), tornando-o mais susceptível a aceitar um próton. O mecanismo procederia então com a molécula de água central, a qual é tornada mais nucleofílica pelas ligações de hidrogênio dela para os oxigênios carboxílicos, atacando o grupo carbonílico do peptídeo (**Fig. 9**). O átomo de nitrogênio seria subsequente ou simultaneamente protonado, possivelmente por um próton transferido do Asp²¹⁵, seguido de rearranjo de carga, clivagem da ligação peptídica e saída dos produtos (**Fig. 9**).

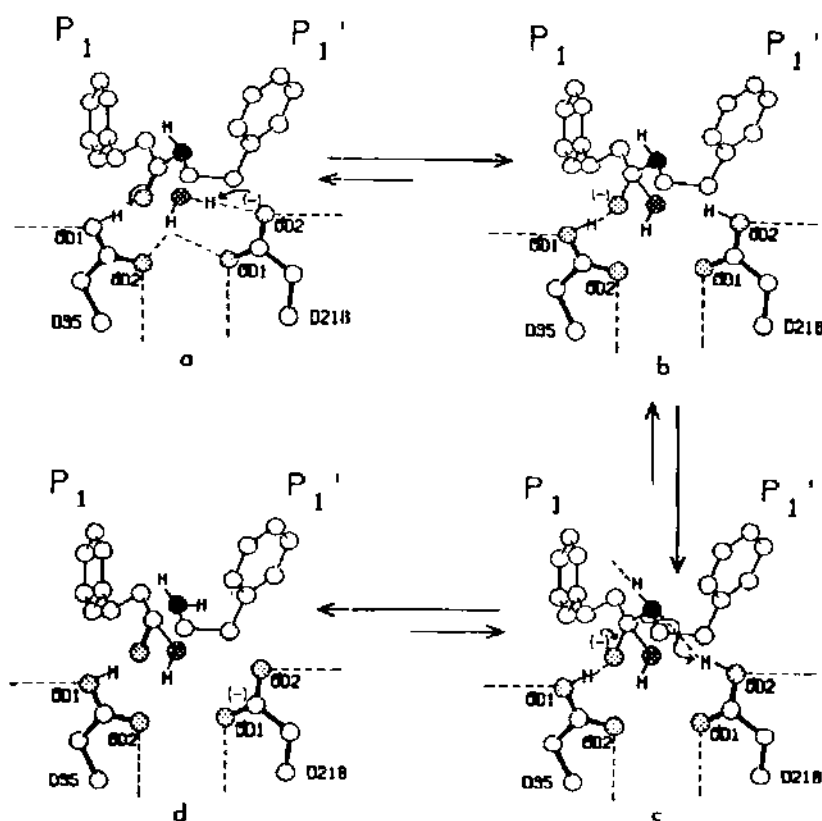
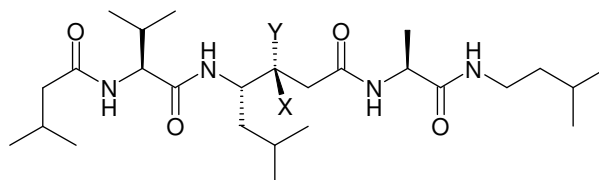


Fig. 9: Mecanismo catalítico proposto para rhizopuspepsina (SUGUNA *et al.*, 1987). . (a) O substrato, representado por um dipeptídeo Phe-Phe, com um grupo peptídico torcido para fora da planaridade, é atacado pela molécula de água central, com transferência de um próton para Asp²¹⁵ (b); (c) Este próton é transferido para o nitrogênio da amida, resultando no rearranjo e quebra da ligação peptídica seguida da saída dos produtos (d). Reproduzido de Vacca *et al.*, 1994.

Importantes contribuições para a elucidação do mecanismo catalítico de aspartil proteases surgiram da série de difluoroestatonas sintetizadas por Rich *et al.* (1985). Previamente, observou-se que a estatona **6** é menos potente que o derivado de estatina **5**. Estes autores demonstraram que a causa de inibidores baseados no cerne difluoroestatona [StoF₂ = ácido (S)-4-amino-2,2-difluoro-3-oxo-6-metil-heptanóico] como **9** serem cerca de 50 vezes mais potentes do que estatonas como **6** e 10 vezes mais potentes que a respectiva estatina **8** residia na maior facilidade de formação do hidrato **10** (Esquema 3). Estes hidratos (**7** e **10**) funcionam como perfeitos análogos do estado de transição, proporcionando alta afinidade de ligação à aspartil proteases. Muitos inibidores potentes foram desenvolvidos baseados neste conceito e os estudos realizados por Veerapandian *et al.* (1992) com o inibidor CP-81,282 (morfolino-4-carbonil-Phe-Nle-ciclo-hexildifluoroestatona-N-metilamida) e de James *et al.* (1992) com Iva-Val-

Val-StoF₂NHCH₃ resultaram em um mecanismo catalítico concordante com o proposto anteriormente por [Suguna et al. \(1987\)](#). Em todas as propostas, Asp²¹⁵ atua como uma base geral para remover um próton de uma molécula de água, enquanto Asp³² doa um próton ao oxigênio carbonílico da ligação peptídica lábil. Estas propostas foram posteriormente corroboradas por um estudo de difração de nêutrons com um IET hidroxietileno em endothiapepsina ([COATES et al., 2001](#)). Os dados demonstraram claramente que Asp²¹⁵ não estava protonado, mas formava ligações de hidrogênio com o grupo hidroxílico do IET ligado ao sítio ativo e que Asp³² deveria estar protonado no átomo de oxigênio externo (Oδ2).

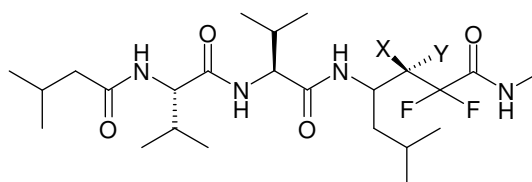
Esquema 3



5: X = OH, Y = H $K_i = 1$ nM

6: X, Y = O $K_i = 56$ nM

7: X = OH, Y = OH



8: X = OH, Y = H $K_i = 10$ nM

9: X, Y = O $K_i = 1$ nM

10: X = OH, Y = OH

Um detalhamento deste mecanismo catalítico baseado nas informações estruturais de complexos enzima-inibidor foi fornecido por [Andreeva & Rumsh \(2001\)](#). Através da análise de 82 complexos, estes autores descobriram mais uma molécula de água conservada, designada W2, entre Tyr⁷⁵, Asn³⁷ e Ser³⁵. Com a ligação do substrato e conseqüente fechamento do “flap”, a hidroxila de Ser³⁵ sofre uma rotação, posicionando-se para formar uma ligação de hidrogênio com Oδ2 de Asp³², enquanto W2 forma uma nova ligação com OH de Tyr⁷⁵ (**Fig. 10**). Ao mesmo tempo, Thr²¹⁸, do outro lado do flap, previne a protonação de Asp²¹⁵, enquanto o substrato não se liga, através de uma ligação de hidrogênio com Oδ2 deste resíduo (**Fig. 10**).

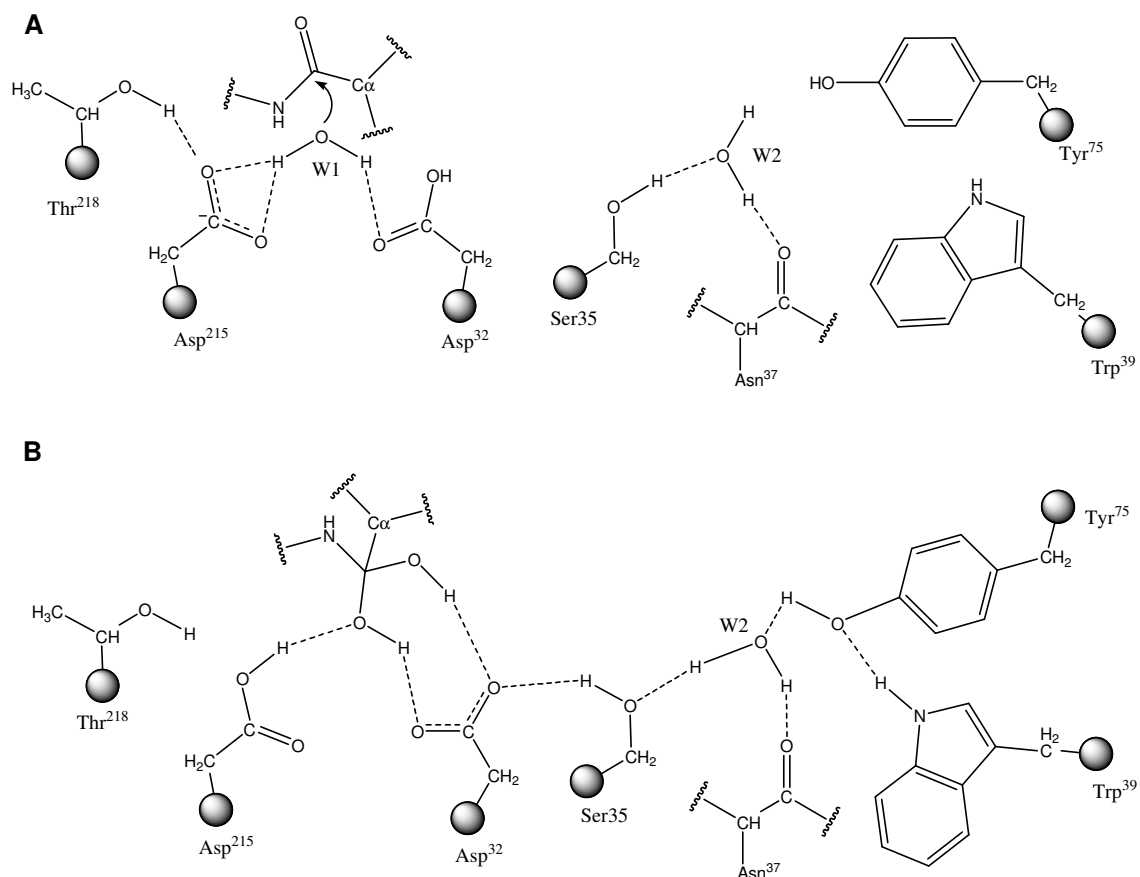


Fig. 10: Mecanismo catalítico de aspartil proteases eucarióticas proposto por [Andreeva & Rumsh \(2001\)](#). A- aproximação do substrato à enzima livre, com ligações de hidrogênio entre os aspartatos catalíticos e molécula de água ligada (W1). B- com a formação do complexo enzima-substrato, a ligação de hidrogênio entre Thr²¹⁸ e Asp²¹⁵ é desfeita, procedendo com o ataque à ligação peptídica para formar o intermediário tetraédrico mostrado. Mudanças conformacionais no flap (Tyr⁷⁵) e em Ser³⁵ alteram o padrão de ligações de hidrogênio com uma segunda molécula de água ligada ao sítio ativo (W2). Isto resulta numa estabilização do intermediário tetraédrico.

Além dos estudos estruturais, cálculos *ab initio* de alto nível também contribuíram para a elucidação do mecanismo catalítico operando em aspartil proteases ([PARK *et al.*, 2000](#); [PIANA *et al.*, 2001](#); [PIANA & CARLONI, 2001](#)). [Park *et al.* \(2000\)](#) investigaram a possibilidade de um ataque nucleofílico direto de um dos aspartatos catalíticos diretamente na carbonila da ligação lábil utilizando cálculos MP2 6-31G**//RHF 6-31G**. Os autores concluíram que considerando-se apenas superfície de energia potencial da reação, a catálise nucleofílica é mais favorecida que a usualmente invocada catálise geral ácida / geral básica (GA/GB). Entretanto, quando a contribuição entrópica fornecida pela movimentação atômica é incluída nos cálculos, o

mecanismo GA/GB mostra-se mais favorável. Os resultados destes cálculos, na ausência do solvente, realçam a importância da pré-organização do sítio ativo para alcançar eficiência catalítica por aspartil proteinases.

Posteriormente, a discussão dos detalhes no mecanismo catalítico de aspartil proteinases resumiu-se à posição de um próton. Contrariando os dados estruturais disponíveis, apontando para o próton no oxigênio externo de Asp³², [Northrop \(2001\)](#) propôs um mecanismo baseado nos estudos teóricos de [Piana & Carloni \(2001\)](#) que encontraram a forma monoprotonada simétrica como a mais estável em HIV-1 protease, com a distância de 2,5 Å entre os átomos Oδ1. A principal característica do mecanismo de Northrop é a formação de uma ligação de hidrogênio de baixa barreira entre os dois resíduos de aspartato, numa estrutura cíclica de 10 membros completada pelas conhecidas ligações de hidrogênio entre a molécula de água ligada ao sítio ativo e os átomos Oδ2 dos aspartatos catalíticos (**Fig. 11**).

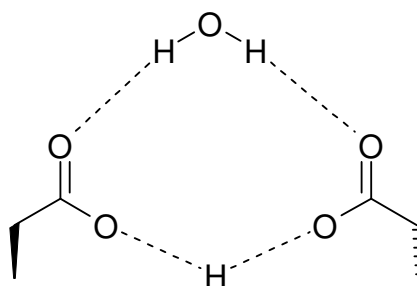


Fig. 11: Aparato catalítico simétrico de aspartil proteases, conforme proposto por [Northrop \(2001\)](#). Foi postulado que a ligação de hidrogênio entre os dois aspartatos é de baixa-barreira.

3.1.2 Ativação e papel do pro-peptídeo na estabilização do enovelamento

Aparentemente, todas as aspartil proteases são sintetizadas na forma de seus zimogênios inativos. A conversão na forma ativa envolve a remoção do pro-peptídeo (ou peptídeo de ativação) de cerca de 45 resíduos do N-terminal. A maior parte dos zimogênios é ativada por acidificação até pH 1-4,5 ([revisado por TANG & WONG, 1987](#)). Neste mecanismo intramolecular, o primeiro passo é uma desnaturação local do pro-peptídeo que acarreta na exposição do sítio ativo e clivagem entre os resíduos 16-17 do peptídeo de ativação. A estrutura do pepsinogênio revela que o pro-peptídeo enovela-se independentemente do restante da proteína, cujo domínio catalítico apresenta conformação muito semelhante à enzima madura ([JAMES & SIELECKI, 1986](#)). Os

autores sugeriram que a inibição ocorre pelo bloqueio do sítio ativo pelo segmento 11-44 do pro-peptídeo, mais especificamente pela interação de Lys^{36P} com os aspartatos catalíticos. Posteriormente, experimentos de mutagênese sítio-dirigida confirmaram esta proposta (RICHTER *et al.*, 1999). Pro-renina não obedece à este mecanismo, muito provavelmente por não apresentar este resíduo de lisina na posição 36 do pro-peptídeo (TANG & WONG, 1987).

O pro-peptídeo de pepsinogênio e de muitas outras aspartil proteinases passíveis de autoativação em pH ácido é rico em resíduos básicos. Isto levou alguns autores a postular que a interação eletrostática entre o peptídeo de ativação e o sítio catalítico pudesse ter um papel importante (DAVIES, 1990). Entretanto, o trabalho de mutagênese de Tanaka & Yada (2001) desmentiu esta hipótese demonstrando que a estabilização conseguida inserindo mais resíduos básicos no pro-peptídeo da pepsina de galinha foi desprezível.

O papel da região pro- na estabilização do processo de enovelamento durante a síntese de aspartil proteases começou a ser demonstrado nos estudos de van den Hazel *et al.* (1993, 1995) com a peptidase A de levedura. Os autores mostraram que a expressão em *trans* do pro-peptídeo era capaz de reverter a degradação que era observada quando a enzima madura fusionada à seqüência pré (peptídeo sinal) era expressa isoladamente. Fukuda *et al.* (1996) demonstraram que a presença da seqüência pro da aspartil protease de *R. niveus* é necessária para a secreção na enzima ativa em sistema de expressão em levedura. Por outro lado, Fortenberry & Chirgwin (1995) haviam demonstrado que o pro-peptídeo de Catepsina D não era estritamente necessário para a expressão da enzima em células de mamífero. Os autores também demonstraram que o direcionamento para o lisossoma foi perdido na ausência do pro-peptídeo, resultando em um nível de expressão 66% menor da proteína recombinante.

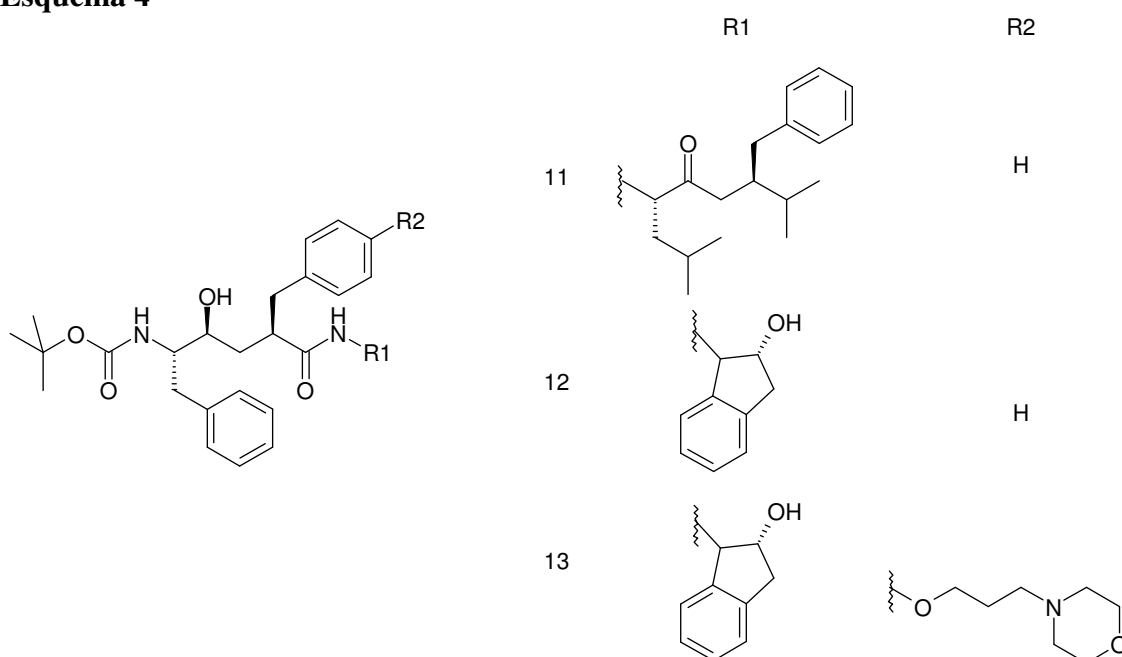
3.1.3 Aspartil proteinases como alvo de desenvolvimento de fármacos

Os esforços de desenvolvimento de inibidores da família de aspartil proteinases iniciaram-se no final dos anos 70 e estiveram grandemente ativos durante os anos 80 quando a indústria farmacêutica perseguia a enzima renina como alvo no tratamento da hipertensão (revisado por LIN & FRISHMAN, 1996). A maior parte dos primeiros

inibidores da renina foram desenvolvidos tendo por base a seqüência mínima do substrato reconhecido (resíduos 6-13 do angiotensinogênio) e a utilização dos IETs mostrados na **Figura 1.5**. Um dos primeiros exercícios de modelagem por homologia foi justamente a predição de modelos da estrutura 3D da renina que pudessem auxiliar no desenho de novos inibidores (BLUNDELL *et al.*, 1983). Posteriormente, a disponibilidade de uma série de inibidores de renina complexados com aspartil proteases fúngicas permitiu que novos compostos fossem desenhados. Isto perdurou até o início dos anos 90 quando as primeiras estruturas da renina foram resolvidas (SIELECKI *et al.*, 1988; DHANARAJ *et al.*, 1992). Uma característica importante da renina é sua extrema especificidade, sendo capaz de clivar única e exclusivamente a ligação Leu¹⁰-Val¹¹ de seu substrato fisiológico. Entretanto, o desenvolvimento de diversos inibidores potentes de renina assemelhando-se a uma grande variedade de seqüências peptídicas e de diferentes tamanhos, sugere que a força de ligação do substrato ao sítio ativo da protease não seja fator determinante para a eficiência catalítica. Isto possivelmente se deve ao fato de que somente poucas seqüências de substratos seriam capazes de posicionar corretamente a ligação peptídica no sítio catalítico para que a hidrólise ocorra. Alguma evidência para esta hipótese foi fornecida pelo estudo de Cooper & Bailey (1994) com 21 complexos de inibidores peptidomiméticos com endothiapepsina.

Muitos dos potentes inibidores de HIV protease existentes hoje são peptidomiméticos baseados no conceito de IET. Muitos grupos identificaram inibidores da renina ativos contra a HIV protease e muitos destes compostos foram subsequentemente desenvolvidos em inibidores potentes e seletivos. Um outro fator importante que influenciou o desenho de inibidores de HIV protease foi a ampla disponibilidade de estruturas cristalográficas de ambas as proteases nativas e inibidas (revisado por VACCA *et al.*, 1994). Por exemplo, o pentapeptídeo **11** (**Esquema 4**) L-682,679) mostrou ser um bom arcabouço para aumentar a solubilidade aquosa, mantendo a potência intrínseca (SOLMS *et al.*, 1991).

Esquema 4



O composto **12** (L-684,434) foi obtido por substituição dos resíduos em P2 e P3, sendo em seguida modelado no sítio ativo da HIV-1 protease (**Esquema 4**). Através da inspeção da estrutura ficou evidente que as posições 4 de ambos os grupos fenílicos em P1 e P1' estavam orientados para a borda do sítio ativo (**Fig. 12**). Acreditava-se que grupos capazes de aumentar a solubilidade em água poderiam ser inseridos em qualquer das posições P1 e/ou P1' sem resultar em perda de potência devido a menor interação desses grupos com a protease. Subseqüentemente, muitos compostos foram sintetizados com este desenho característico e o mais potente da série, **13** (L-689,502), contém um grupo morfolina etóxido na posição 4 do grupo fenil em P1 (**Esquema 4**). O composto **13** foi co-cristalizado com a HIV protease e forneceu o primeiro exemplo de um complexo enzima-inibidor não-peptídico ([THOMPSON *et al.*, 1992](#)).

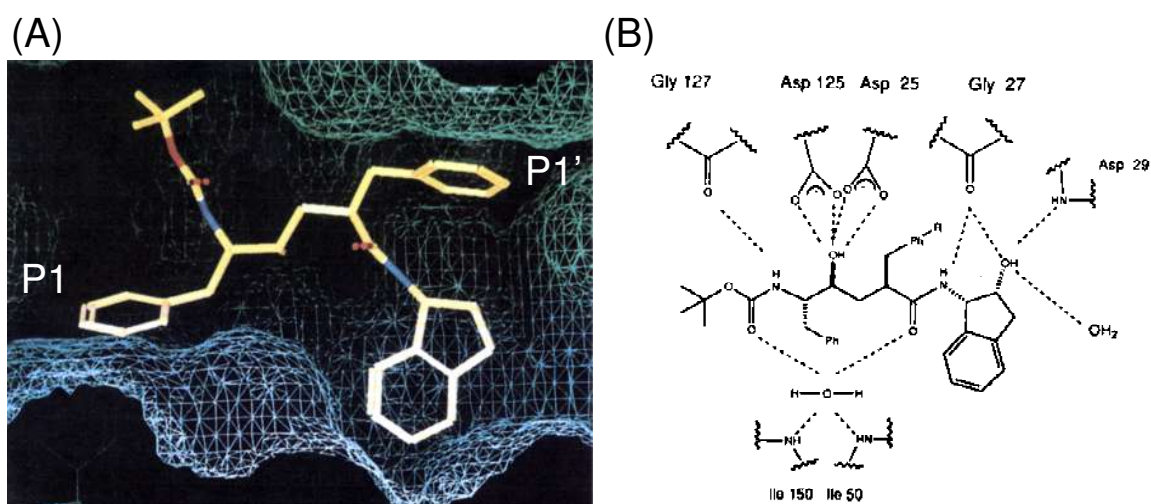
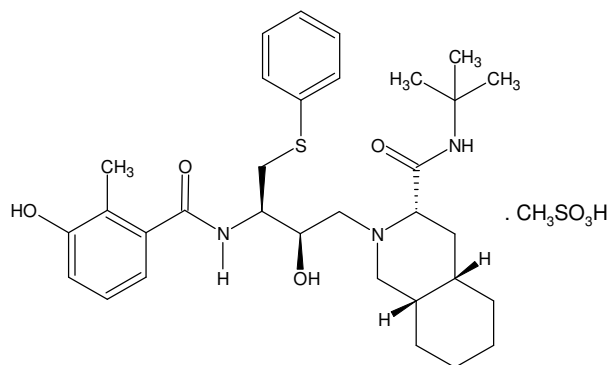


Fig. 12: Modelagem da interação do composto 12 no sítio ativo da HIV-1 protease nativa. A- Superfície molecular dos subsítios S1 e S1'. B- Diagrama esquemático das possíveis interações por ligações de hidrogênio entre 12 e HIV-1 protease como encontrado no complexo enzima-inibidor. Adaptado de [Vacca *et al.* \(1994\)](#).

Nelfinavir (utilizado na clínica como o sal mesilato; **14**) foi o primeiro inibidor comercial de HIV-1 protease a ter a estratégia de desenvolvimento inteiramente baseada na estrutura do alvo molecular. Nelfinavir é um potente inibidor de protease ($IC_{50} = 2$ nM) e agente antiviral ($ED_{50} = 14$ nM). A primeira menção na literatura ao Nelfinavir, ainda batizado com o código AG-1343, foi em 1995 no trabalho de [Gehlhaar *et al.* \(1995\)](#), que descreveram uma simulação computacional do ancoramento (em inglês, *docking*) na HIV-1 protease. Posteriormente, [Kaldor *et al.* \(1997\)](#) relataram a determinação por cristalografia de raios-X do complexo enzima-inibidor, o que permitiu analisar o modo de ligação de AG-1343 ao sítio ativo da protease (**Fig. 13**). A estrutura do complexo revela como os substituintes tiofenil éter e fenolamida deste inibidor não peptídico, interagem com os subsítios S1 e S2 da enzima, respectivamente.



14

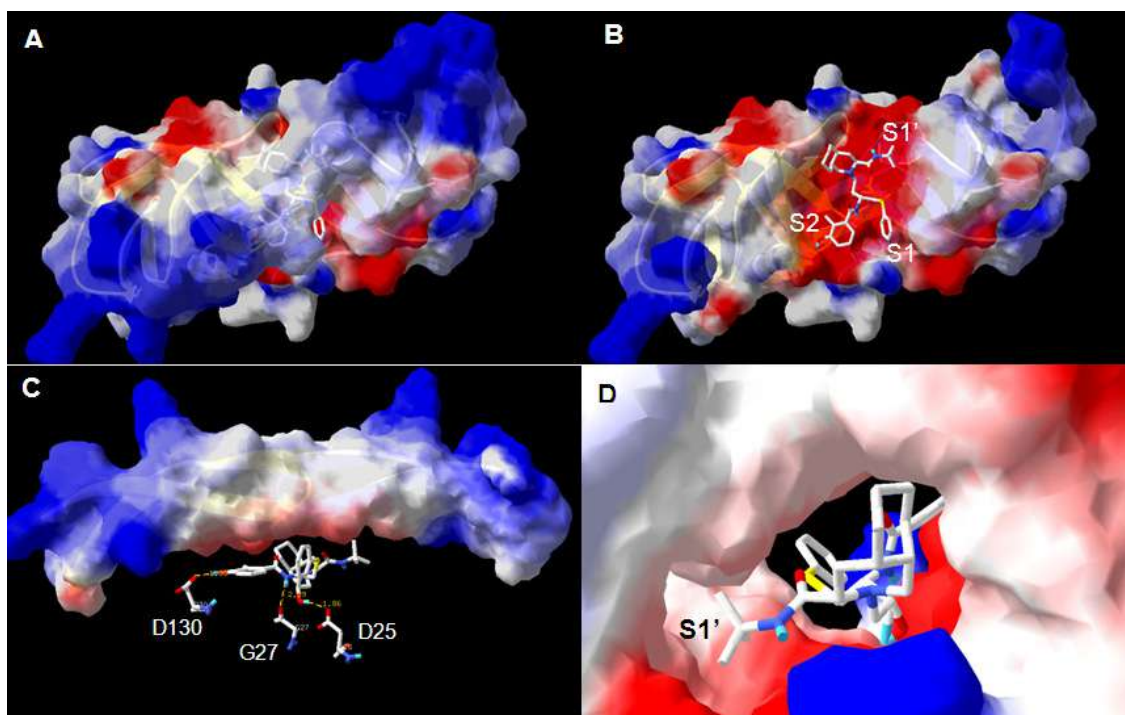


Fig. 13: Interação de AG-1343 com a protease do HIV (código PDB 1OHR). A- Visão com o eixo de simetria C_2 perpendicular ao plano da página, mostrando a região de *flap* encobrindo o inibidor. B- A região de *flap* foi extraída para permitir a observação da interação dos grupos tiofenil (P1), fenolamido (P2) e *t*-butilperidroisoquinolina (P1min) com os subsítios da enzima. C- Visão com o eixo de simetria no plano da página e pontes de hidrogênio de AG-1343 com os resíduos Asp²⁵, Gly²⁷ e Asp¹³⁰. D- Superfície do subsítio S1' em detalhe.

Alguns conceitos gerais aplicáveis ao desenvolvimento de inibidores de aspartil proteases surgiram da experiência previa de desenvolvimento de inibidores de HIV-1 protease e renina (COOPER, 2002): (i) os grupos CO e NH, desde P3 até P3' estão quase sempre envolvidos em ligações de hidrogênio com grupos delineando a fenda de ligação ao substrato e *flaps*. Assim, diferenças nas afinidades de ligação de inibidores

do mesmo tamanho devem ser devidas à interações seqüência-específicas com os subsítios que compõe o sítio ativo; (ii) a ligação de algumas cadeias laterais em S1 pode influenciar na conformação adotada pelo componente do inibidor interagindo com sítios distais, como por exemplo quando o grupo ciclo-hexilmetil em P1 força Phe em P3 a adotar um conformação menos favorável em um inibidor de renina; (iii); A redução do número total de conformações acessíveis a um peptídeo em solução pode ser alcançada por ciclização. Isto também é importante para favorecer a afinidade de ligação pela protease, devido ao subsítios de reconhecimento do substrato serem fracamente definidos em aspartil proteases. De fato, inibidores potentes têm sido relatados pela aplicação deste conceito (TYNDALL & FAIRLIE, 2001); (iv) sabidamente, peptídeos dispõem de baixa biodisponibilidade devido a problemas de absorção intestinal e hidrólise por enzimas digestivas e plasmáticas. Algum sucesso no aumento da estabilidade de inibidores de renina foi conseguido através de bloqueio dos grupos N- e C-terminal, substituição de resíduos envolvidos em sítios de clivagem conhecidos para proteases digestivas (Phe para quimiotripsina, Arg ou Lys para tripsina) e, logicamente, alteração de ligações peptídicas para ligações não hidrolizáveis.

1.2 *Aspartil proteases catepsina D-símile de S. mansoni*

1.2.1 Esquistossomose

1.2.1.a Agentes etiológicos e distribuição

As esquistossomoses também são conhecidas por bilharzioses em homenagem a Theodor Bilharz, que identificou pela primeira vez os parasitas no Egito em 1851. Historicamente, estas parasitoses são associadas principalmente às regiões da Ásia, África, América do Sul e algumas ilhas do Caribe. Atualmente, sabe-se que as doenças são causadas por trematódeos do gênero *Schistosoma* e têm como principais agentes etiológicos no homem as espécies (REY, 1991): *S. mansoni* (África e é a única espécie na América Latina), *S. haematobium* (espécie mais prevalente na África e Oriente Médio), *S. intercalatum* (restrito a 10 países na África), *S. japonicum* (encontrado em toda a região do Pacífico, incluindo China e Filipinas) e *S. mekongi* (restrito a áreas limitadas do Laos e Camboja). Os números globais da esquistossomose (WHO, 1996; CHITSULO et al., 2000) comumente citados incluem: 600 milhões de pessoas sob

risco, mais de 200 milhões de pessoas infectadas (10% destes severamente infectados) e 120 milhões de pacientes sintomáticos. Uma reavaliação recente do impacto da doença indica que a esquistossomose é a segunda doença parasitária em causa de mortes prematuras, perdendo apenas para a malária (CHITSULO *et al.*, 2004). Estratégias de controle da morbidade através da quimioterapia resultaram no controle efetivo da doença no Brasil, Sul da África, Oriente Médio, China e Filipinas; o desenvolvimento econômico no Caribe e Ilhas Maurício erradicou a doença nestas regiões (Fig. 14).

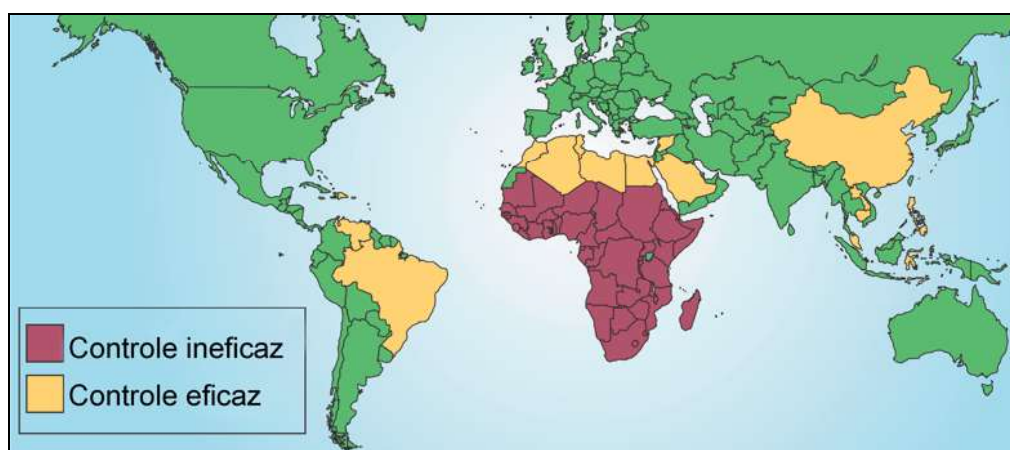


Fig. 14: Distribuição e situação global do controle da esquistossomose. Adaptado de Chitsulo *et al.*, 2004.

1.2.1.b Impacto global da doença

A gravidade que assume a doença em muitos casos e o déficit orgânico que produz, fazem das esquistossomoses um dos mais sérios problemas de saúde pública, em escala mundial, e pesado fardo para as populações das áreas endêmicas. Recentemente, o impacto da esquistossomose foi reavaliado, sugerindo que 70 milhões de indivíduos sofrem de hematúria associada à infecção por *S. haematobium*, 18 milhões são afetados por complicações da bexiga e 10 milhões sofrem de hidronefrose (FENWICK *et al.*, 2003). A morbidade devida à *S. mansoni* inclui hepatoesplenomegalia, fibrose hepática e ascite, e cerca de 130.000 pessoas morrem por ano de hematêmese devido à hipertensão portal associada (WHO, 2002). A Figura 15 mostra a gravidade da doença em crianças, as quais são especialmente vulneráveis à esquistossomose e as em idade escolar apresentam-se frequentemente física e intelectualmente comprometidos pela anemia associada (GUYATT *et al.*, 2001).



Fig. 15: Esquistossomose mansônica em dois meninos de área endêmica no Brasil. . À esquerda, o menino de 11 anos exibe aumento maciço no fígado e baço com marcante expansão da circulação abdominal colateral. O outro menino, com 9 anos, encontra-se num estágio bem avançado da infecção crônica inicial, com sinais de fase aguda (dores de cabeça, náusea, anorexia, febre e diarreia). Fonte: WHO/TDR (http://www9.who.int/tropical_diseases/databases/imagelib.pl?imageid=9101069).

Enquanto a maior parte das parasitoses vai diminuindo de importância, em função do desenvolvimento econômico e dos métodos de controle disponíveis, a esquistossomose encontra-se ainda em expansão em muitas regiões do mundo. Isto está vinculado ao desenvolvimento de recursos hídricos para a irrigação ou para a produção de energia hidroelétrica, especialmente na África, mas que pode tornar-se relevante no cenário brasileiro em razão da recente crise energética atravessada pelo país.

1.2.1.c Situação da doença no Brasil

No Brasil, onde a doença é conhecida popularmente por xistossomose, xistosa ou doença dos caramujos, assim como por “barriga d’água” (devido à ascite que acompanha as formas mais graves, com fibrose hepática), admite-se existir mais de seis milhões de indivíduos infectados (REY, 1991). A esquistossomose mansônica tem ampla distribuição geográfica no Brasil, com maior intensidade de transmissão na região nordeste do país e norte de Minas Gerais.

Segundo dados da Divisão de Esquistossomose da Secretaria de Vigilância Sanitária do Ministério da Saúde (MS/SVS/ESQUISTOSSOMOSE), desde o início da década de 1950 até o presente ano tem sido observada redução nas prevalências de infecção, mediante inquéritos coproscópicos populacionais. Entretanto, tem ocorrido

uma maior distribuição espacial, com o processo de urbanização e migração. Os indicadores de morbidade hospitalar e mortalidade também têm evidenciado redução em análises realizadas nos sistemas de informações existentes. Nos últimos 20 anos o percentual de positividade detectado em inquéritos coproscópicos realizados inicialmente pelo Ministério da Saúde e mais recentemente pelas Secretarias Estaduais e Municipais de Saúde, tem se situado entre 5,5% a 11,6% da população examinada (**Fig. 16**). Observam-se variações anuais na positividade, em função da heterogeneidade das áreas que são submetidas aos inquéritos em cada estado. Dentre os estados destacam-se Alagoas, Pernambuco, Sergipe, Paraíba, Bahia, Minas Gerais e Espírito Santo, com os maiores percentuais de positividade.

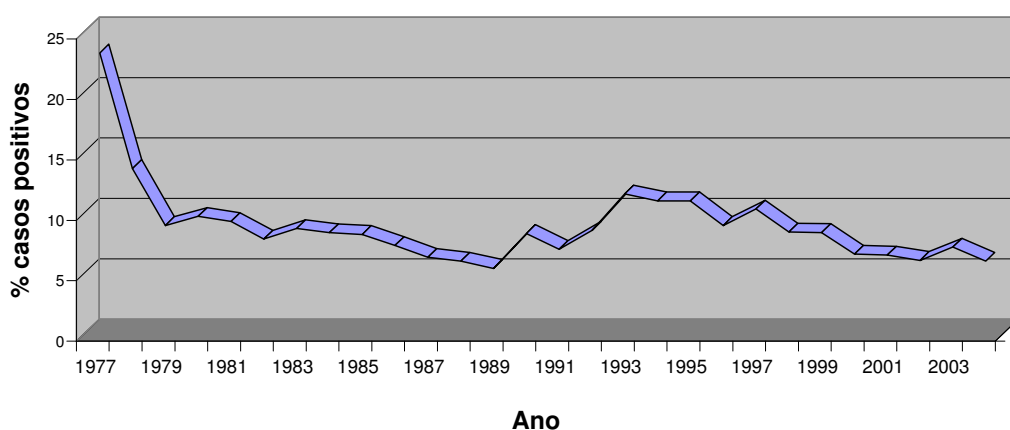


Fig. 16: Série histórica dos percentuais de positividade para esquistossomose no Brasil de 1977 a 2004. Fonte: MS/SVS/ESQUISTOSSOMOSE.

Os dados do MS/SVS também indicam que nos estados de maior percentual de positividade, também têm sido registrados os maiores coeficientes de mortalidade e proporção de internações por esquistossomose. A tendência histórica aponta para uma redução na morbidade e mortalidade por formas graves pela doença no país. Entretanto, nas áreas endêmicas da região nordeste do Brasil, a ocorrência da forma hepatoesplênica ainda representa um importante problema de saúde. Desde meados da década de 1970, vem sendo desenvolvido um programa de controle da esquistossomose no País, baseado na realização de inquéritos coproscópicos e tratamento de portadores e, em menor escala, ações de saneamento, tratamento de criadouros e educação em saúde. Nos municípios onde o programa tem alcançado uma boa cobertura e regularidade, os indicadores têm sido impactados positivamente. Desde 1993, o Ministério da Saúde tem trabalhado para a implementação da descentralização das ações de controle da

esquistossomose. Essas ações foram intensificadas a partir de 1999 e 2000, quando as Secretarias Municipais passaram a assumir gradativamente o controle desta endemia em seu território.

1.2.1.d Biologia básica e ciclo evolutivo dos esquistossomos

Os helmintos trematódeos da classe Digenea são em geral achatados dorsoventralmente, alongados e de aspecto foliáceo; outras vezes são piriformes ou longos e delgados, chegando as fêmeas de *S. mansoni* a serem cilíndricas e filiformes. O tamanho varia de um milímetro até vários centímetros de comprimento. Os esquistossomos, assim como todos os trematódeos da classe Digenea, são parasitas heteroxenos, isto é, requerem mais de um hospedeiro para completar sua evolução. O nome trematódeo (do grego *trematodes*, significando “furado”) faz referência a estruturas conspícuas da superfície do corpo, as ventosas, das quais uma situada na superfície ventral é particularmente desenvolvida e denominada acetábulo. Os trematódeos do gênero *Schistosoma* distinguem-se dos outros Digenea por apresentarem os sexos separados com acentuado dimorfismo sexual (**Fig. 17**). Na forma adulta, a fêmeas são normalmente encontradas acasaladas ao macho através de uma estrutura na face ventral deste último, conhecida como canal ginecóforo. No *S. mansoni* e em outras espécies que parasitam o homem, uma segunda ventosa, denominada ventosa oral, dispõe-se em torno da boca, na extremidade anterior do corpo (**Fig. 17**).

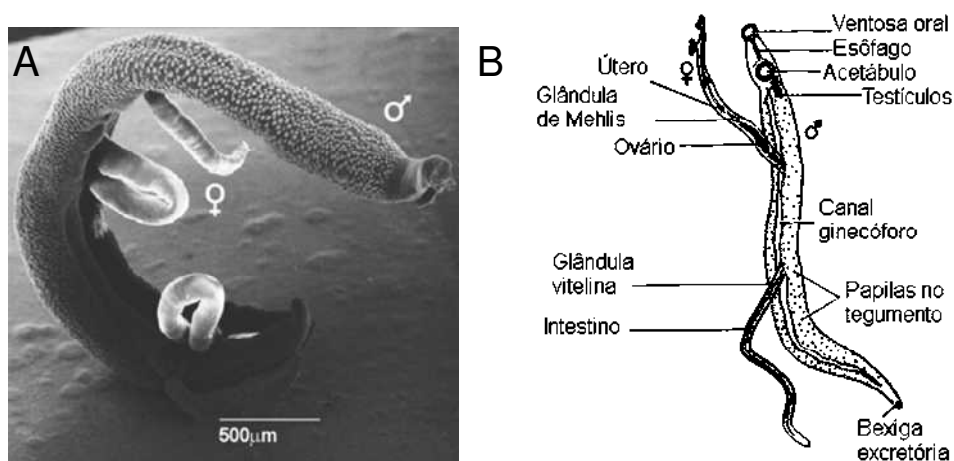


Fig. 17: Anatomia de vermes adultos de *S. mansoni*. A – Fotomicrografia eletrônica de varredura mostrando a fêmea dentro do canal ginecóforo do macho. B- Esquema destacando as principais características anatômicas.

O ciclo de vida de esquistossomos é bem descrito (REY, 1991) e algumas características da infecção causada pelas espécies *S. mansoni*, *S. haematobium* e *S. japonicum* estão ilustradas na **Figura 18**. Em sua fase adulta, o *S. mansoni* caracteriza-se como parasito da luz dos vasos sangüíneos do homem e de outros mamíferos, habitando preferencialmente as vênulas do plexo hemorroidário superior e as ramificações mais finas das veias mesentéricas, particularmente da mesentérica inferior, onde põe seus ovos. Depois de atravessarem a mucosa intestinal, os ovos são eliminados com as fezes e, quando chegam a ser arrastados em tempo útil para dentro de águas superficiais, libertam seus miracídios, que nadam durante algumas horas até encontrar certos moluscos aquáticos dos gêneros *Biomphalaria* (*S. mansoni*), *Bulinus* (*S. haematobium*) ou *Oncomelania* (*S. japonicum*). Penetrando nos tecidos do molusco, os miracídios transformam-se em esporocistos que, por poliembrionia, geram esporocistos-filhos e depois cercárias. Voltando ao meio líquido, as cercárias que abandonam o hospedeiro invertebrado ficam nadando livremente na água, quase sempre em direção à superfície, enquanto não têm oportunidade de entrar em contato com a pele do hospedeiro vertebrado (homem ou outro animal suscetível).

Assim que penetram ativamente a pele do homem, as cercárias logo perdem sua cauda, diferenciando-se em esquistossômulos. Portanto, as cercárias (larvas) constituem a forma infectante para o homem e outros vertebrados. Depois de permanecer na pele por um tempo variável (de minutos a um dia inteiro), o esquistossômulo empreende a migração através do corpo de seu novo hospedeiro, penetrando nos vasos cutâneos e através da circulação, chega ao coração direito, depois aos pulmões e, em seguida, ao fígado. Chegando ao sistema porta intra-hepático, os esquistossômulos desenvolvem-se e alcançam a fase adulta. Os vermes adultos acasalam-se e migram para as vênulas da parede intestinal, caminhando contra a corrente sangüínea da veia porta e das veias mesentéricas, onde a fêmea deposita seus ovos. Muitos dos ovos são eliminados nas fezes, mas uma parte migra até o fígado pelo sistema porta, onde provocam a reação imunológica que caracteriza a doença.

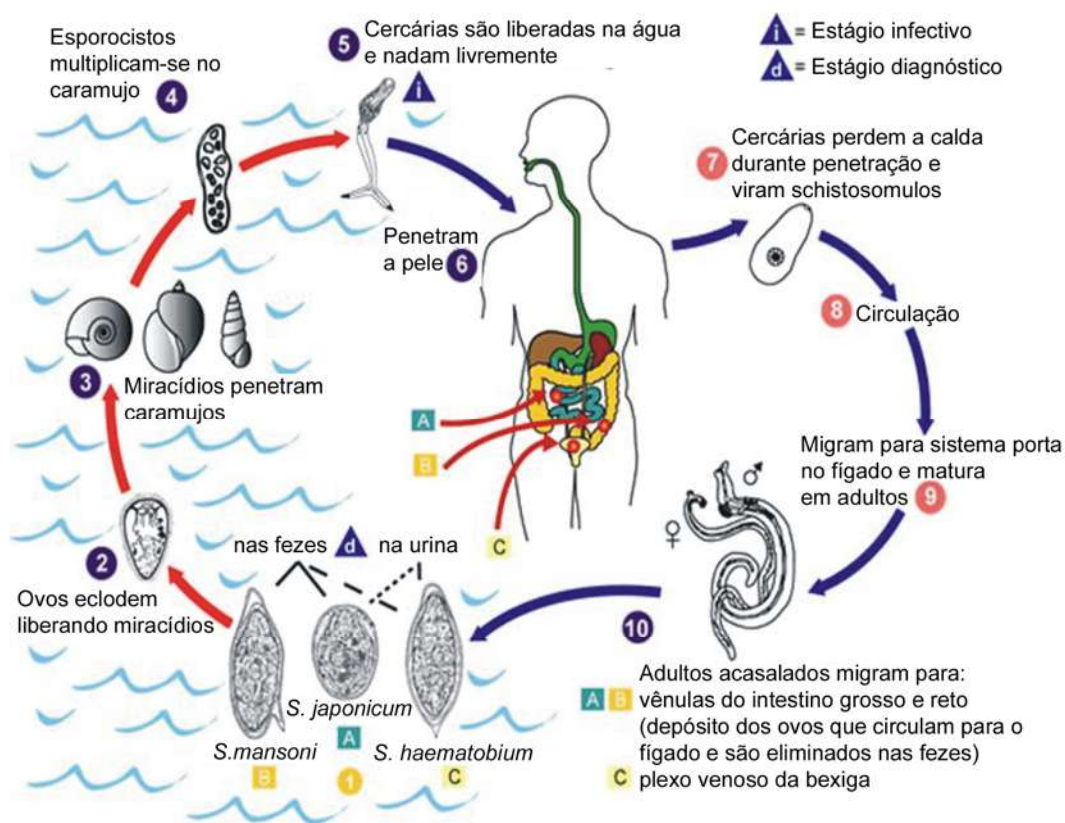


Fig. 18: Ciclo evolutivo de *Schistosoma sp.*

1.2.1.e Patogênese

Didaticamente, a esquistossomose pode ser dividida em uma fase aguda e uma fase crônica (REY, 1991). A fase aguda (ou inicial) é caracterizada por processos discretos de lesão na pele (exantemas, prurido e outras reações alérgicas locais), pulmão ou fígado, de origem inflamatória. Por vezes, a morte de alguns vermes produz obstrução embólica do vaso, seguida de necrose do tecido em torno, mais tarde substituído por tecido de cicatrização. Dependendo do número de parasitas e da sensibilidade do paciente, pode-se desenvolver um quadro descrito como forma toxêmica da esquistossomose, caracterizada por febre, eosinofilia, linfadenopatia, esplenomegalia e urticária. Admite-se constituir esta, uma forma de reação imunológica, semelhante à doença do soro e provocada por imunocomplexos. A fase crônica (ou tardia) da esquistossomose é acompanhada pela formação de granulomas, fibrose periportal, hepatoesplenomegalia, lesões cardiopulmonares e lesões diversas como tumores, lesões renais e do sistema nervoso central devido a localizações ectópicas dos ovos de *S. mansoni*.

1.2.1.f Diagnóstico

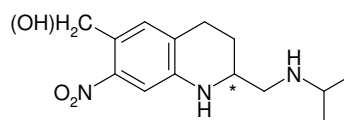
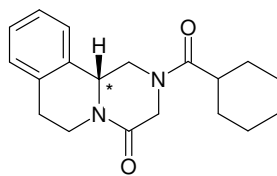
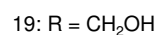
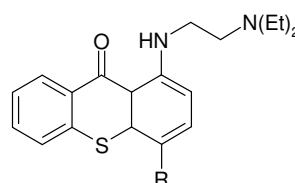
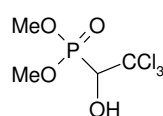
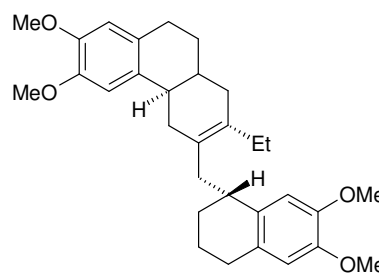
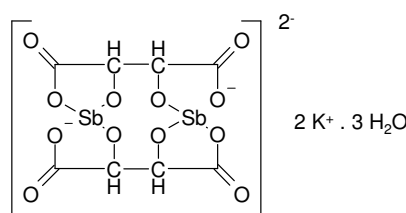
Os vermes adultos permanecem escondidos na circulação sanguínea ao longo da infecção por esquistossomos, portanto o diagnóstico parasitológico é baseado na evidência indireta da presença destes no paciente, através da detecção dos ovos. Assim, o diagnóstico da esquistossomose é feito por técnicas envolvendo filtração da urina na suspeita de infecção por *S. haematobium* (FELDMEIER, 1993) e esfregaço fecal nos casos de infecção por *S. mansoni* e *S. japonicum* (KATZ *et al.*, 1972). Estas técnicas são relativamente simples e de baixo custo, mas requerem treinamento técnico especializado e acesso a um microscópio. Além disso, estes métodos sofrem de rapidez, praticidade e podem ser considerados esteticamente desagradáveis. Métodos sorológicos podem vir a ser úteis para detecção dos antígenos em áreas endêmicas e testes da presença de anticorpos podem auxiliar o diagnóstico em áreas não-endêmicas. Os níveis de antígeno circulante no sangue do paciente são frequentemente bem menores que o de anticorpos, mas teoricamente podem ser diretamente relacionados com o número de vermes presentes (STURROK, 2001). Atualmente, o acesso a laboratórios especializados para o imunodiagnóstico já não mais é uma necessidade, pois testes simples foram desenvolvidos para o uso em campo (VAN ETTEN *et al.*, 1994), apesar de ainda não estarem disponíveis comercialmente.

1.2.1.g Controle farmacoterapêutico da esquistossomose mansônica

Historicamente, antimoniais (15) e desidroemetina (16) (Esquema 5) foram introduzidas durante a primeira grande guerra (~1917) como os primeiros anti-helmínticos efetivos contra a esquistossomose (revisado por CIOLI *et al.*, 1995). Estes foram suplantados durante a segunda guerra mundial (1942-1945) pela xantenona conhecida como Miracil D ou lucantona (18). Cerca de 20 anos se passaram até que o derivado mais potente, hicantona (19) fosse obtido por fermentação microbiana. Ainda, em 1960 foi introduzido o organofosforado denominado metrifonato (17), o qual ainda está em uso com *S. haematobium*. Entretanto, com exceção deste último, nenhum destes compostos provocou um impacto significativo na esquistossomose e hicantona mostrou-se mutagênico, prevalecendo as estratégias de controle através da erradicação do vetor. No início da década de 70, dois novos fármacos foram introduzidos no tratamento da esquistossomose mansônica: praziquantel (PZQ; 20) e oxamniquina

(OXA; 21). Desde então, a estratégia de controle da esquistossomose mudou das tentativas de redução da transmissão usando moluscidas para a farmacoterapia (STURROK, 2001). Isto ocorreu devido ao alto custo e difícil manutenção do controle dos caramujos aliado à possibilidade de efeitos negativos sobre o meio ambiente que esta estratégia poderia trazer. Além disso, os fármacos que passaram a estar disponíveis eram seguros, eficazes e acessíveis a um baixo custo. Entretanto, custos crescentes causados pelas elevadas taxas de re-infecção e as repetidas intervenções comumente requeridas mesmo após tratamento em massa, têm limitado o impacto do controle baseado somente na farmacoterapia. Uma vez que o desenvolvimento de uma vacina ainda é um objetivo elusivo, o sucesso no controle da esquistossomose está vinculado ao comprometimento político e à implementação de um pacote de controle integrado, onde a farmacoterapia atua como principal fator operacional (SIBLEY & HUNT, 2003).

Esquema 5



OXA é produto de uma segunda geração de análogos estruturais da hancantona (CIOLI *et al.*, 1995). Há um centro quiral na molécula (indicado com asterístico), mas

as preparações comerciais constituem uma mistura racêmica. Este fármaco é eficiente apenas contra *S. mansoni*, sendo incapaz de reverter a infecção nas espécies prevalentes na África e Ásia. Além disso, OXA é mais ativo em vermes machos e menos ativo entre a 3ª e 7ª semana de infecção. O mecanismo de ação da OXA envolve a inibição da síntese do DNA, possivelmente através da intercalação deste composto no lugar das bases nitrogenadas corretas. Evidências de cruzamentos genéticos entre cepas resistentes e suscetíveis à OXA, indicam que o fator responsável pela resistência é um gene recessivo (CIOLI *et al.*, 1995). Isto sugere que a razão para a ineficácia de OXA em outras espécies de esquistossomos é a falta de uma atividade que é essencial para o desenvolvimento do efeito esquistossomicida. Posteriormente, foi sugerido que esta atividade seria pertencente à uma sulfotransferase presente unicamente em *S. mansoni* e ausente em outras espécies não-suscetíveis.

A OXA, é usada na clínica na forma de cápsulas de 250 mg ou como um xarope de 250 mg em 5 mL. Uma dose única de 15-20 mg/kg é eficiente na América do Sul, África ocidental e na ilhas do Caribe (CIOLI *et al.*, 1995). O fármaco é bem absorvido na parede intestinal e é metabolizada no fígado à metabólitos inativos, que são excretados na urina. Possui meia vida curta de 1-2 horas, sendo eliminada no plasma em 10-12 horas. Os efeitos indesejáveis consistem em tonteira e cefaléia transitórias, cuja ocorrência em vários estudos é relatada em 30-95% dos pacientes e em distúrbios gastrintestinais em 10-20% dos pacientes. Podem ocorrer sintomas causados pela estimulação do sistema nervoso central, incluindo alucinações e episódios convulsivos. Manifestações alérgicas e outros sintomas, que surgem vários dias após a interrupção do tratamento, podem estar relacionados com a liberação de produtos do trematódeo morto (RANG *et al.*, 2001). Desde que a patente do PZQ expirou sua validade, OXA deixou de ser o medicamento esquistossomicida de menor custo, perdendo para formulações genéricas do primeiro. Desta forma, PZQ tornou-se o fármaco de escolha para o tratamento da esquistossomose, inclusive a mansônica.

O PZQ é um anti-helmíntico de amplo-espectro, sendo eficaz também na cisticercose, que até a introdução deste medicamento não tinha tratamento eficaz (RANG *et al.*, 2001). Apesar de ser efetivo contra os estágios infectivos do parasito (cercárias e jovens esquistossômulos), uma proporção significativa das formas imaturas

dos vermes são capazes de sobreviver à administração de PZQ da 1ª à 5ª semana de infecção (CIOLI *et al.*, 1995). O grupo pirazino-isoquinolina da molécula foi inicialmente explorado pela Merck (Darmstadt, Alemanha) no desenvolvimento de tranqüilizantes. Desde que propriedades anti-helmínticas (anti-cestódeos) foram detectadas nestas substâncias pela Bayer, mais de 400 derivados foram sintetizados, dentre os quais estava o PZQ (CIOLI *et al.*, 1995). A excelente atividade biológica codificada na estrutura extremamente simples deste fármaco apresenta uma “faca de dois gumes” para a química medicinal, pois ao mesmo tempo em que sua síntese é extremamente simples, pouco espaço existe para modificação do cerne ativo da molécula. Assim como a OXA, PZQ também é produzido comercialmente como uma mistura racêmica, apesar de saber-se atualmente que apenas o isômero R(-) ou *levo* é ativo (vide estereoquímica indicada na estrutura 20).

Pouco se sabe sobre o mecanismo de ação do PZQ. Nas menores doses efetivas, causa aumento na atividade muscular do parasito, seguido de contração, paralisia espástica e destacamento dos vermes das paredes dos vasos sanguíneos, resultando no rápido deslocamento das veias mesentéricas para o fígado (TRACY & WEBSTER, 2001). Foi proposto que o alvo deste fármaco fosse uma glutathione S-transferase de 26 kDa, um alvo de vacina em *S. japonicum*, baseado na elucidação do complexo cristalográfico com esta enzima (McTIGUE *et al.*, 1995). Entretanto, pouco tempo depois foi demonstrado que PZQ era incapaz de inibir a atividade desta enzima (revisado por CIOLI, 1998). O evento central na atividade do PZQ em esquistossomos é possivelmente o fenômeno de influxo de Ca^{2+} no tegumento que é observado em poucos minutos após exposição ao fármaco. No entanto, nada é sabido sobre o mecanismo pelo qual esta alteração na homeostasia do Ca^{2+} é operada. Por outro lado, tem sido demonstrado que a eficiência do PZQ é prejudicada em animais imunodeficientes, porém podendo ser reestabelecida pela administração de anticorpos específicos contra alguns antígenos de superfície (CIOLI *et al.*, 1995; REDMAN *et al.*, 1996; CIOLI, 1998). Estas duas linhas de evidência, combinadas, deram origem à proposta de que o mecanismo de ação de praziquantel pudesse estar associado à exposição de antígenos presentes sob a bicamada lipídica que recobre o tegumento mediante perturbações da membrana induzidas pela contração muscular disparada pelo aumento do influxo de

Ca²⁺ (REDMAN *et al.*, 1996).

O PZQ originalmente produzido pela Bayer é encontrado na forma de tabletes de 600 mg, que podem ser divididos em 4 segmentos de 150 mg, a um custo de US\$2-5 (FENWICK *et al.*, 2003). Atualmente, mais de 20 companhias farmacêuticas formulam PZQ de boa qualidade (DOENHOFF *et al.*, 2000) a partir de matérias-primas produzidas na Coreia do Sul e China e o fornecem a um custo médio de US\$0,1 em pacotes de 1000 tabletes de 600 mg (KUSEL & HAGAN, 1999). Uma única dose oral de 40 mg/kg ou três doses de 20 mg/kg cada, administradas em intervalos de 4-6h, geralmente produzem taxas de cura de 70-95% e consistente redução (85%) na contagem de ovos nas fezes (TRACY & WEBSTER, 2001). O PZQ não exerce nenhum efeito farmacológico no homem, na posologia indicada (RANG *et al.*, 2001). Sofre rápida absorção após administração oral; grande parte da droga é rapidamente convertida a metabólitos inativos durante a primeira passagem pelo fígado (sofre “efeito de primeira passagem”), sendo estes excretados pela urina (TRACY & WEBSTER, 2001). A meia vida plasmática do composto original é de 60-90 minutos. Ocorrem efeitos indesejáveis leves, que costumam ser transitórios e raramente de importância clínica (RANG *et al.*, 2001). Estes últimos consistem em distúrbios gastrintestinais, tonteira, dor muscular e articular, erupções cutâneas e febre baixa (provavelmente devido aos produtos tóxicos liberados pela morte dos parasitos).

1.2.1.h Resistência aos fármacos em uso e novas alternativas terapêuticas

Diferenças na complexidade biológica e ecológica nos parasitos e seus vetores fazem com que helmintos como *S. mansoni* sejam menos sujeitos aos problemas de resistência à farmacoterapia do que os parasitos unicelulares, como por exemplo, as espécies de *Plasmodium* causadoras da malária. Por causarem doenças de desenvolvimento mais lento e que normalmente são diagnosticadas meses após a infecção inicial, helmintos são expostos aos fármacos menos frequentemente e somente na fase crônica da infecção. Do ponto de vista dessa menor exposição, há uma menor tendência ao desenvolvimento de resistência, pois o tratamento exerce menor pressão sobre a população de parasitas (SIBLEY & HUNT, 2003). Adicionalmente, organismos diplóides como *S. mansoni* são menos sujeitos à expressão de um fenótipo mutante, pois somente mutações dominantes à resistência são manifestadas por mutantes em um único

alelo. Por outro lado, a presença de um ciclo sexual no hospedeiro humano favorece o desenvolvimento repentino de resistência por recombinação genética, ao contrário de parasitas clonais cuja resistência evolui lentamente.

Atualmente, um número considerável de evidências experimentais foi acumulado demonstrando que resistência à PZQ e OXA pode ser induzida em cepas mantidas em laboratório pressionadas por estes fármacos (revisado por [BRINDLEY, 1994](#); [CIOLI, 2000](#)). Maior debate tem sido dedicado ao aparecimento de resistência em isolados de campo, principalmente contra o fármaco de escolha no tratamento da esquistossomose — PZQ ([CIOLI, 2000](#); [FENWICK et al., 2003](#)). Apesar de não terem relevância imediata à situação na clínica, a seleção de cepas resistentes no laboratório serve para dar uma indicação que o desenvolvimento de resistência no campo pode realmente ocorrer e com que velocidade.

Excluindo-se relatos ocasionais de redução na eficiência de PZQ, duas grandes investigações em larga-escala, foram realizadas a fim de verificar a presença ou não de cepas de *S. mansoni* resistentes a este fármaco (revisado por [CIOLI, 2000](#)). Os estudos foram independentemente conduzidos a partir de dados da infecção de um grupo de indivíduos de um intenso foco recentemente identificado no Senegal e de outro grupo proveniente da região do delta do Nilo no Egito. Em ambos os casos, quedas de até 50% nas taxas de cura com a dose padrão (40 mg/kg) foram bem documentadas. Doses de PZQ de até 60 mg/kg foram administradas a indivíduos refratários ao tratamento, mas a cura só foi obtida pelo tratamento com doses usuais de OXA. Em relação ao foco no Senegal, fatores ligados à alta taxa de transmissão e à intensidade de infecção extremamente elevada, além da ingenuidade imunológica da população relacionada ao recente estabelecimento do foco, foram invocados para explicar os achados ([CIOLI, 2000](#)).

Compatível com simulações computacionais recentemente publicadas ([DANSO-APPIAH & DE-VLAS, 2002](#)), acredita-se que a elevada intensidade de infecção poderia de fato, sozinha, explicar taxas de cura inferiores a 40%, mesmo para tratamentos com 90-95% de eficiência. Com relação à alta taxa de transmissão na região do Senegal, é possível que uma grande proporção dos indivíduos tivessem sido re-infectados nas 4-5 semanas que antecederam ao tratamento. Assim, PZQ seria intrinsecamente ineficaz

contra uma fração considerável da população de vermes presentes nestes indivíduos. Esta proposta foi apoiada pelo sucesso obtido quando indivíduos da mesma região foram submetidos a dois tratamentos seguidos com três semanas de intervalo, abordagem esta que alcançou os níveis normais de cura. Se cepas resistentes estivessem presentes, seria esperado que o segundo tratamento fosse tão ineficaz quanto o primeiro. Além disso, nenhum sucesso foi alcançado em sucessivas tentativas de demonstrar em laboratório a resistência à PZQ por cepas isoladas a partir do estudo no Senegal. Assim, as evidências contrariando o aparecimento de resistência nesta região são mais convincentes do que às apontado a favor.

Diferentemente, na infecção experimental de camundongos com cercárias derivadas dos isolados de campo da região do Delta do Nilo, no Egito, mostrou-se claramente um aumento de 2-6 vezes na dose necessária para reduzir em 50% o número de vermes (ED50) em relação à cepa suscetível. Colocado em perspectiva, este resultado não chega a ser alarmante, uma vez que cepas resistentes à hancantona, por exemplo, mostram uma ED50 até 1000 vezes maior, que as cepas suscetíveis a este fármaco. No entanto, os dados são importantes, pois são corroborados por dados da caracterização molecular da cepa resistente e podem ser correlacionados com a suscetibilidade em pacientes (CIOLI, 2000).

Em função da expansão dos programas de controle da esquistossomose, principalmente na África, a pressão a qual estarão sujeitas as cepas selvagens presentes nestas áreas endêmicas continuará a aumentar. Estima-se que somente em 4 países da África, mais de 100 milhões de pessoas sejam tratadas nos próximos 10 anos, podendo o consumo atual de praziquantel chegar a 40 milhões de tabletes ao ano, a partir de 2005 (FENWICK *et al.*, 2003). Uma forma de contrabalancear esta pressão é instituir programas de vigilância farmacológica para detectar os primeiros sinais de resistência e restringir a farmacoterapia aos grupos de maior risco, como mulheres, pescadores e crianças em idade escolar (SIBLEY & HUNT, 2003).

O reduzido arsenal atualmente disponível para o controle farmacoterapêutico da esquistossomose é razão de preocupação para muitos especialistas da área, assim como deveria ser para as autoridades públicas. Na possibilidade da resistência ao PZQ chegar ao ponto de inviabilizar o uso deste fármaco no tratamento da esquistossomose, a única

alternativa disponível seria OXA, mesmo assim somente na infecção por *S. mansoni*; as esquistossomoses por *S. haematobium* e *S. japonicum* estariam descobertas. Apesar de nenhum indicativo existir de quanto tempo este quadro aterrorizador — cerca de 200 milhões de pessoas acometidas por uma doença sem tratamento — levará para acontecer, é evidente que não se pode continuar com apenas um (ou nenhum) fármaco disponível para o tratamento da esquistossomose. Estratégias em andamento para a descoberta de novas alternativas terapêuticas incluem: estudos clínicos em fase III com o antimalárico artemisinina; produção estereoespecífica do isômero *levo*(R) de PZQ; síntese de análogos de OXA; e triagem aleatória de novas moléculas esquistossomicidas (FENWICK *et al.*, 2003). Excetuando-se os estudos com artemisinina, nenhuma destas estratégias tem sido significativamente perseguidas na prática e todas, exceto o desenho de análogos de OXA, carecem de um componente racional. Programas de descoberta de novos genes e projetos para o seqüenciamento do genoma de *S. mansoni* (FRANCO *et al.*, 2000) têm o potencial de contribuir significativamente com a descoberta de novos alvos terapêuticos, abastecendo assim os programas (ainda escassos) de desenvolvimento racional de novos fármacos esquistossomicidas.

1.2.2 Projeto genoma de *S. mansoni*: fonte potencial de novos alvos terapêuticos

1.2.2.a Genoma nuclear

O genoma de *S. mansoni* é rico em AT (34% G+C) e aproximadamente 40% repetitivo, com o restante dos genes sendo cópia única ou pertencendo a pequenas famílias gênicas (revisado por WILLIAMS & JOHNSTON, 1999; LoVERDE *et al.*, 2004; OLIVEIRA *et al.*, 2004; EL-SAYED *et al.*, 2004). O cariótipo de *S. mansoni* consiste em oito pares de cromossomos, sete autossômicos e um sexual, com a fêmea formando um par heterogamético WZ e os machos um par homogamético ZZ (Fig. 19). Cada cromossomo contém entre 15 e 64 Mb dos 270 Mb de DNA nuclear total (SIMPSON *et al.*, 1982) e podem ser distinguidos por diferenças no tamanho, tamanho relativos dos braços e conteúdo de heterocromatina (padrão de bandamento-C).

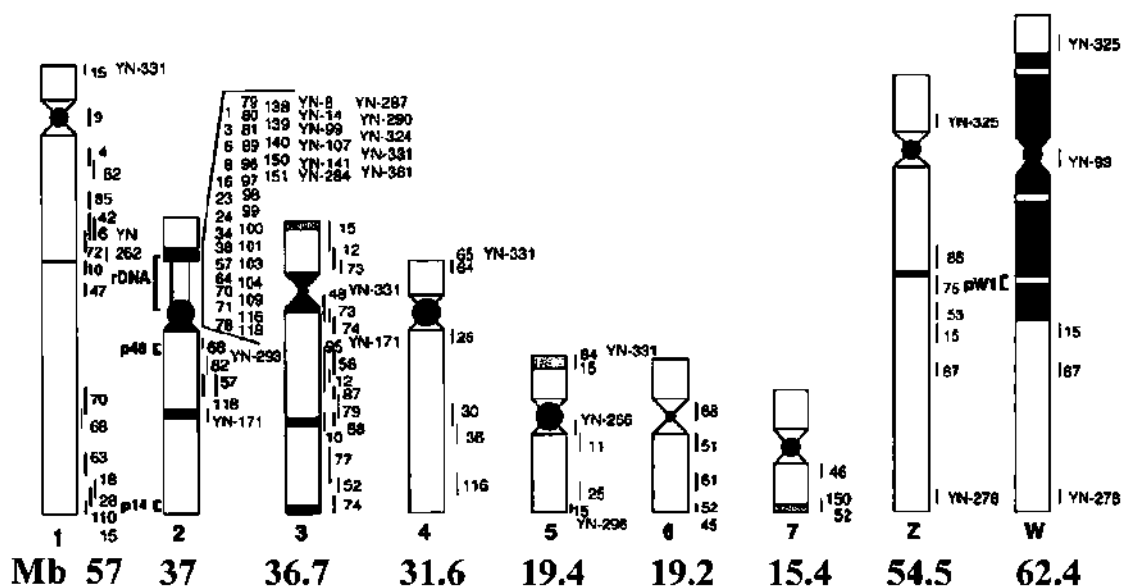


Fig. 19: Cromossomos de *S. mansoni*. Ideograma mostrando a localização de genes (p14 e p48), elementos repetitivos (rDNA e pW1) e fragmentos de DNA clonados em cromossomos artificiais de levedura (YACs; número a direita). Adaptado de [LoVerde et al. \(2004\)](#).

1.2.2.b Programa de descoberta de novos genes: análise do transcriptoma

O projeto genoma de *S. mansoni* (PGS) iniciou-se em 1992 como uma iniciativa brasileira em colaboração com “The Institute for Genomic Research” (TIGR, EUA) para apoio logístico e treinamento de pessoal. A partir de 1994, o PGS passou a receber auxílio do “WHO/UNDP/World Bank Special Program for Research and Training in Tropical Diseases (TDR)”, tornando-se a rede genoma de esquistossomos (RGE). Os objetivos principais, inicialmente perseguidos pela RGS constituíam (i) descoberta e caracterização de novos genes de *S. mansoni* e *S. japonicum*; (ii) desenvolvimento de um mapa físico de baixa resolução para *S. mansoni* e (iii) sequenciamento e análise do genoma mitocondrial de ambas as espécies. Em virtude do grande tamanho do genoma de *S. mansoni*, com 40% de elementos repetitivos, e dos recursos limitados disponíveis, um projeto de sequenciamento genômico em larga escala foi inicialmente preterido a uma estratégia de identificação de novos genes mais direta, que tirasse proveito das diversas bibliotecas de cDNA previamente disponíveis para *S. mansoni* ([WILLIAMS & JOHNSTON, 1999](#)). A decisão de não adotar nenhuma linhagem de referência para os estudos genômicos permitiu que o PGS pudesse ser posto em prática imediatamente. Tal estratégia consistia na geração de etiquetas de seqüências expressas (ESTs), que são

seqüências curtas de 100-400 bases obtidas pelo sequenciamento de uma única passagem de um clone de biblioteca de cDNA. Este processo é realizado repetidas vezes com clones selecionados aleatoriamente, normalmente a partir da extremidade 5' do DNA. As seqüências são submetidas a diversos filtros de bioinformática para eliminar segmentos oriundos do vetor e trechos de baixa qualidade. As seqüências processadas são então submetidas a um processo de agrupamento (*clustering*) de forma a reduzir a redundância e facilitar a análise dos dados (PROSDOCIMI *et al.*, 2002).

Antes de 1994, existiam apenas 220 seqüências em bancos de dados para todas as espécies de esquistossomos. Franco *et al.* (1995) deram início ao PGS gerando um total de 607 ESTs a partir de uma biblioteca de cDNA de vermes adultos, representando 169 genes, 154 dos quais previamente desconhecidos. Esta única iniciativa foi capaz de aumentar em mais de 50% o número de genes conhecidos de *S. mansoni*, demonstrado o potencial desta estratégia para a descoberta de novos genes. Com o passar dos anos e apoio da Organização Mundial de Saúde, ESTs foram gerados a partir de bibliotecas de cDNA de todas as formas evolutivas do parasito, chegando a 16.296 seqüências no final de 2002 (MERRICK *et al.*, 2003). Estes ESTs estão depositados no GenBank (BENSON *et al.*, 2005), como usual, mas estão melhor organizados no banco abrigado no TIGR, conhecido como “*Schistosoma mansoni* Gene Index” (SmGI). O objetivo do SmGI é fornecer, dentro de uma plataforma pública, uma visão não-redundante de todos os genes de *S. mansoni* e dados sobre os padrões de expressão, papéis celulares, funções e inter-relações evolutivas (MERRICK *et al.*, 2003). Na construção de SmGI, além dos ESTs gerados pelo projeto genoma, transcritos expressos de *S. mansoni* (ETs), i.e., seqüências codificantes completas ou mRNAs parciais, são incorporados a partir do processamento de registros no GenBank. Seqüências comparadas par-a-par com o programa WU-BLAST (ADAMS *et al.*, 1992) são ordenadas em *clusters* de seqüências pareadas com $\geq 95\%$ de identidade em uma região com pelo menos 40 pb de comprimento, com pontas dessemelhantes < 20 pb. A partir daí, seqüências compreendendo cada *cluster* são reunidas em seqüências consenso tentativas (TCs) usando ferramentas de bioinformática usuais. Qualquer EST não atribuído a um TC é denominado um *singleton*. Após este processo, verificou-se que o total de ESTs no SmGI representava 7491 novas seqüências, ao final de 2002.

Em outubro de 2003, uma contribuição muito relevante foi dada por [Verjovski-Almeida et al. \(2003\)](#) que geraram 125.000 seqüências a partir do mRNA de diversos estágios do desenvolvimento de *S. mansoni*, através da técnica ORESTES ([DIAS-NETO et al., 2000](#)). O método ORESTES utiliza pequenas quantidades de mRNA (15-40 ng) para a construção de mini-bibliotecas, que associadas à baixa “estringência” e utilização de *primers* aleatórios na reação de transcrição reversa, resulta numa preferência estatisticamente demonstrada pela porção central dos genes. Além de focalizar na região de maior probabilidade de encontrar a região codificante de um gene, esta abordagem é complementar às estratégias usuais de geração de ESTs, que são direcionadas para as extremidades dos genes. As seqüências ORESTES foram compiladas em 30.988 contigs (ou SmAEs, como denominado pelos autores; similares aos TCs do SmGI). Através de dois métodos distintos, o número total de genes de *S. mansoni* foi estimado em 14.000, um número ligeiramente inferior à estimativa anterior de 15.000-20.000 genes ([SIMPSON et al., 1982](#); [FRANCO et al., 2000](#)). Um resultado usual relatado em muitos projetos transcriptoma é a distribuição de genes por categorias funcionais. Por exemplo, a análise de [Verjovski-Almeida et al. \(2003\)](#) mostrou que uma fração significativa dos genes de *S. mansoni* está envolvida no metabolismo de proteínas (25%), sendo 50% destes genes específicos de metazoários. Os autores também destacaram SmAEs de potenciais alvos de fármacos, como por exemplo genes parálogos de canais de cálcio que podem relacionados com o mecanismo de ação do praziquantel, receptores de superfície e prováveis fatores de resistência por analogia com outros organismos.

1.2.2.c Sequenciamento total do genoma (STG)

Conforme afirmado por Phil LoVerde, no primeiro encontro da RGE em 1994, ninguém ousava suspeitar que o sonho do STG de *S. mansoni* pudesse se tornar real ([LoVERDE et al., 2004](#)). Entretanto, os melhoramentos contínuos nas tecnologias empregadas para o sequenciamento de DNA, avanços nas ferramentas de bioinformática e a contínua redução nos custos, fizeram o sonho tornar-se realidade. Espera-se que o STG resultará na descoberta de novos genes, identificação de marcadores microsatélite, além de fornecer informação essencial sobre a organização gênica e genômica. Como passo inicial, um grande número de seqüências terminais de BACs (cromossomos

artificiais de bactérias) foram geradas com o intuito de descobrir novos genes e servir de marcadores para a construção de mapas físicos de alta resolução para agilizar o sequenciamento do genoma. Um total de 28 Mb de seqüências descontínuas foi obtido, fornecendo em média um marcador de 600-700 pb a cada ~8500 pb (LoVERDE *et al.*, 2004; EL-SAYED *et al.*, 2004). A estratégia original de sequenciamento a ser empregada seria a “*Map-as-you-go*”, onde regiões genômicas são selecionadas iterativamente de forma a minimizar a quantidade de sobreposição das seqüências obtidas. Entretanto, muitas das seqüências terminais de BACs geradas anteriormente caíam em extensas regiões repetitivas do genoma, tornando-se inúteis como ponto de referência para o sequenciamento.

Atualmente, a estratégia ótima adotada para o STG de *S. mansoni* é a técnica “*shotgun*”, onde um número muito grande de sequenciamentos de passagem única (GSS) são realizados a partir de ambas as pontas de clones selecionados aleatoriamente a partir de bibliotecas de diferentes tamanhos de inserto. Iniciando no TIGR em 2002, o total de seqüências obtidas até o momento (2,8 milhões de GSSs) equivale a uma cobertura de 7,5 vezes o tamanho do genoma (EL-SAYED *et al.*, 2004). Baseando-se em modelos estatísticos, estima-se que menos de 0,5% do genoma ainda está por ser seqüenciado (EL-SAYED *et al.*, 2004). Atualmente, os centros envolvidos estão empenhados no processamento das seqüências para disponibilização na forma de um banco de dados.

1.2.2.d Genômica funcional

Com a era genômica de *S. mansoni* chegando ao seu final, alguns grupos já estão envolvidos com abordagens pós-genômicas, como a técnica de micro-arranjos (*microarrays*) e a proteômica. Hoffmann *et al.* (2002a) aplicaram micro-arranjos para estudar a expressão diferencial de genes em vermes machos e fêmeas. Através desta abordagem, conseguiu-se dobrar o número genes sexo-específicos conhecidos. Mais recentemente, espectrometria de massas MALDI-TOF acoplada à eletroforese bidimensional foi empregada no estudo do proteoma solúvel de quatro formas evolutivas de *S. mansoni*, nominalmente cercárias, esquistossômulos de pulmão, vermes adultos e ovos (CURWEN *et al.*, 2004). Os autores observaram que muitas das proteínas mais abundantes nos mapas eletroforéticos bidimensionais que puderam ser

identificadas foram previamente relatadas, muitas das quais como antígenos majoritários. Devido à presença ubíqua dessas proteínas nos quatro estágios do crescimento do parasito analisados, os autores argumentaram a necessidade de selecionar proteínas realmente estágio-específicas como marcadores.

1.2.3 Digestão da hemoglobina por *S. mansoni*

Os trematódeos possuem aparelho digestivo simples, seus tubos digestivos apresentam uma única comunicação com o meio exterior (a boca na extremidade da ventosa anterior), pois não possuem ânus. Esta cavidade afunilada continua-se com uma região muscular, o esôfago. Este último desemboca no ceco bifurcado que depois volta a se unir na junção cecal. Depois de digeridos os alimentos e de absorvidos os nutrientes, os resíduos são regurgitados pela abertura oral, graças a ondas antiperistálticas que percorrem o órgão (REY, 1991).

As formas adultas dos trematódeos do gênero *Schistosoma* alimentam-se do sangue de seu hospedeiro (CHEEVER & WELLER, 1958). O consumo de um macho de *S. mansoni* é calculado em cerca de 30.000 hemácias por hora, enquanto que a fêmea consome dez vezes mais glóbulos vermelhos (LAWRENCE, 1973). A hemoglobina (Hb) proveniente das hemácias é essencial para o desenvolvimento, crescimento e reprodução dos parasitos do gênero *Schistosoma* (CHEEVER & WELLER, 1958). Os eritrócitos são lisados pela ação da hemolisina contida no esôfago destes parasitos (KASSCHAU *et al.*, 1986; BOGITSH, 1989). A Hb liberada dos eritrócitos flui para o trato digestivo do esquistossomo, onde é catabolizada a dipeptídeos ou aminoácidos livres. Estes peptídeos difundem-se para as células gastrodermais da parede do intestino (ou são captados por estas) a fim de poderem ser, em última análise, incorporados nas proteínas do esquistossomo (ZUSSMAN *et al.*, 1970; BOGITSH, 1989). Enquanto que a porção globina é utilizada, hematina, o produto final oxidado do componente heme é regurgitado pelo parasito e acumula-se no fígado dos hospedeiros infectados (LAWRENCE, 1973).

O primeiro relato sobre a digestão da Hb por estes parasitas foi feito há mais de 40 anos por Timms & Bueding (1959) demonstrando que as necessidades nutricionais *in vitro* do *S. mansoni* adulto poderiam ser satisfeitas pela adição de aminoácidos

derivados da Hb ao meio de cultura. Este autores sugeriram que, *in vivo*, “estes aminoácidos estariam disponíveis ao parasita pela digestão das células sanguíneas vermelhas e pela ação de uma ou mais enzimas proteolíticas presentes no canal alimentar dos vermes”. No mesmo trabalho, foi identificada uma atividade proteolítica com um pH ótimo de 3,9 nos extratos de *S. mansoni* adultos. Um importante avanço na caracterização das peptidases envolvidas no catabolismo da hemoglobina foi conseguido por Dresden & Deeler (1979), onde a partir de extratos aquosos de *S. mansoni* adultos, foi parcialmente purificada uma atividade proteolítica com pH ótimo de atividade ácido e sensível a compostos contendo tiol. Posteriormente, foi demonstrado que a enzima em questão tratava-se de uma cisteinil proteinase catepsina B-símile, pertencente ao clã da papaína (DRESDEN *et al.*, 1981; CHAPPELL *et al.*, 1986).

No início dos anos 80, estudos imunoquímicos (RUPPEL *et al.*, 1987) detectaram a presença de atividades proteolíticas em frações protéicas de 31 kDa (Sm31) e 32 kDa (Sm32) no trato digestivo de esquistossomos. Estas frações foram subsequentemente atribuídas a uma hemoglobinasase catepsina B-símile (SmCB; GOTZ & KLINKERT, 1993) e a uma asparaginil proteinase incapaz de digerir eficientemente a Hb (SmAE; DALTON *et al.* 1995), respectivamente. Um transcrito de catepsina L (SmCL1) foi seqüenciado por SMITH *et al.* (1994). A partir de uma biblioteca de cDNA enriquecida com transcritos de vermes fêmeas de *S. mansoni*, um segundo gene de catepsina L (SmCL2) foi clonado por Michel *et al.*, (1995). Baseados nos experimentos que localizaram SmCL2 no sistema reprodutivo de fêmeas, estes autores descartaram um envolvimento desta enzima no catabolismo de hemoglobina. As duas enzimas catepsina L-símile de *S. mansoni* apresentam considerável diferença nas seqüências de aminoácidos, exibindo provavelmente funções distintas. Posteriormente, utilizando-se substratos supostamente capazes de discriminar catepsinas L, B e H, foi sugerido que a principal atividade cisteinil proteinase exercida nos extratos de *S. mansoni* era referente a uma catepsina L ao invés de B como se acreditava (DALTON *et al.*, 1996).

Butler *et al.* (1995) relataram a seqüência de uma enzima similar à catepsina C em *S. mansoni* (SmCC). Catepsinas C são cisteinil dipeptidil aminopeptidases,

consequentemente a enzima identificada em esquistossomos poderia estar envolvida com as etapas finais no catabolismo da hemoglobina. Assim, até meados da década de 90, apenas cisteinil proteinases eram conhecidas no trato digestivo de esquistossomos, com as catepsinas-símile L e B possivelmente atuando diretamente sobre a Hb, enquanto que a enzima cathepsina C-símile poderia estar atuando em etapas finais de degradação (DALTON *et al.*, 1995b). Baseado na homologia com legumaínas de plantas, acreditava-se que SmAE estivesse envolvida na ativação em *trans* de outros zimogênios de proteases digestivas de esquistossomos.

O grupo de P.J. Brindley da Universidade de Queensland (Austrália) relatou a clonagem do primeiro gene de uma aspartil hemoglobinase de esquistossomos a partir de uma biblioteca de cDNA de *S. japonicum* e a seqüência nucleotídica obtida mostrou tratar-se de uma enzima similar à catepsina D (SjCD1; BECKER *et al.*, 1995). Os autores foram provavelmente motivados por resultados prévios de experimentos de localização ultra-estrutural com pepstatina marcada com mercúrio e com soro heterólogo anti-catepsina D bovina em *S. japonicum*, que mostraram a localização de uma enzima homóloga na luz intestinal e em vesículas na gastroderme, coerente com uma papel na digestão, e também no tegumento (BOGITSH & KIRSCHNER, 1986; BOGITSH & KIRSCHNER, 1987). Ainda, Bogitsh *et al.* (1992) já haviam demonstrado a importância de uma enzima homóloga à catepsina D para a digestão da hemoglobina em *S. japonicum*, através de experimentos de inibição *in vitro* e *in vivo* com o soro heterólogo contra a enzima ortóloga bovina. Portanto, os esquistossomos empregam, sem dúvida, uma variedade de enzimas proteolíticas para levar a cabo a digestão da Hb. De fato, há a necessidade de classificar as proteinases envolvidas neste processo por suas propriedades estruturais e catalíticas, além do termo “hemoglobinase”.

Nos últimos 10 anos, duas excelentes revisões foram publicadas sintetizando o conhecimento sobre as proteases associadas ao trato digestivo de esquistossomos, principalmente em relação à especificidade de substrato e papel de cada enzima na digestão da hemoglobina (BRINDLEY *et al.*, 1997; CAFFREY *et al.*, 2004). Neste período, um novo gene codificando uma enzima catepsina B-símile (SmCB2) foi identificado, mas este não está presente no trato digestivo e sim no tegumento. Na revisão de Caffrey *et al.* (2004) foi exposto resultado de análise dos transcritos de

cisteinil proteases digestivas presentes em SmGI revelando um segundo gene estreitamente relacionado a SmCB1 (como foi rebatizada a primeira catepsina B identificada), com 95% de identidade ao nível da seqüência de aminoácidos. Entretanto, a questão se este segundo gene seria um variante polimórfico codificado por um alelo não foi abordada. Polimorfismo também foi identificado no gene de SmAE, onde uma mutação no resíduo de cisteína catalítico (C197) gera um variante inativo da enzima. Quanto à função de SmAE, foi recentemente demonstrado que esta enzima é realmente capaz de ativar o zimogênio de uma protease digestiva em *S. mansoni*, SmCB1.

O predomínio do tipo de atividade cisteinil protease em *S. mansoni* foi reavaliado e atualmente acredita-se que a atividade catepsina B seja dominante, em acordo com experimentos de marcação do sítio ativo, inibição seletiva e coerente com o nível de expressão deste gene, conforme inferido pelo número de transcritos presentes no SmGI (CAFFREY *et al.*, 2004). SmCB1 pode atuar ainda como uma dipeptidil exopeptidase, devido à presença de um motivo estrutural de 20 resíduos em sua superfície. Este tipo de atividade também é encontrado em SmCC, a qual demonstrou-se necessitar de uma catepsina L ou S para a sua ativação a partir do zimogênio. Curiosamente, a localização de SmCL2 no parasito foi reavaliada e atualmente acredita-se que juntamente com seu parólogo, SmCL1, encontra-se na gastroderme de *S. mansoni* e podem ser as responsáveis pela ativação de SmCC. Desta forma, o conhecimento atual sobre a digestão de hemoglobina em esquistossomos avançou consideravelmente na última década, podendo ser resumido conforme a **Figura 20**. Vale ressaltar que a maioria absoluta das evidências experimentais apoiando este esquema é proveniente de ensaios bioquímicos *in vitro*. Claramente, a demonstração *in vivo* da existência ou não de uma via catabólica ordenada para a digestão de hemoglobina ainda está faltando.

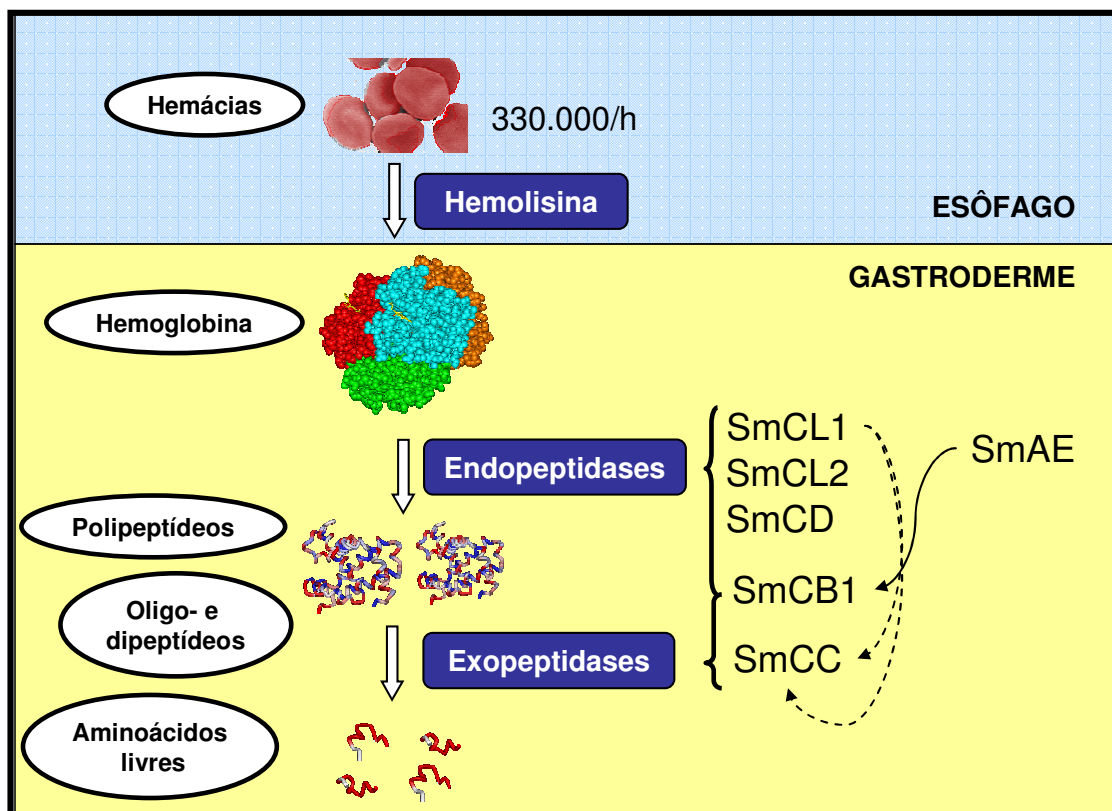


Fig. 20: Digestão da hemoglobina por esquistossomos.

1.2.4 A hemoglobinase catepsina D-símile

1.2.4.a Duas formas da mesma enzima?

Originalmente, [Becker et al. \(1995\)](#) relataram a seqüência do cDNA do zimogênio da aspartil proteinase clonada (batizada de Sjasp; SjCD1 no sistema de [CAFFREY et al., 2004](#)) com 380 resíduos de aminoácidos. A seqüência de aminoácidos, traduzida a partir do cDNA, apresentava 50-55% de identidade com outras catepsinas D previamente descritas, incluindo a humana. Através de alinhamentos com outras enzimas melhor caracterizadas, os autores propuseram que a protease madura teria 36 kDa resultantes de seus 329 resíduos de aminoácidos. Além das seqüências $D^{33}TG$ e $D^{231}TG$, características do sítio ativo, todos os 4 pares de cistina encontrados na enzima ortóloga humana também foram conservados em Sjasp. A análise também revelou que o sítio de processamento responsável pela clivagem da catepsina D humana (resíduos 94-107) em uma cadeia leve e outra pesada, não foi mantido em Sjasp. A análise da seqüência indicou ainda a presença de 4 sítios potenciais de N-glicosilação (resíduos 70, 172, 199 e 212), dois dos quais compartilhados pela catepsina D humana

(N70 e N199). Os autores também anteciparam que alguns dos resíduos distintos entre a enzima do parasito e a humana poderiam resultar numa especificidade de substrato diferenciada.

Becker *et al.* (1995) também relataram uma atividade peptidásica sobre o substrato de catepsina D Phe-(Ala)₂-Phe(NO₂)-Phe-Val-Leu-metoxi-4-piridil e uma atividade hemoglobinolítica, ambas na faixa de pH 3,0 a 5,5, com um máximo em pH 3,5. Os autores relataram que estas atividades puderam ser significativamente inibidas por pepstatina, embora não tenham especificado em que pH. Isto é importante, pois a atividade hemoglobinolítica de cisteinil proteases como SmCB1, SmCL1 e SmCL2 são importantes em pH > 4, onde a atividade de Sjasp é irrelevante, conforme mostrado pelos autores. Na época os autores atribuíram a maior parte da atividade hemoglobinolítica à Sjasp e o restante a uma atividade catepsina-L. Vale ressaltar que a proposta atualmente melhor fundamentada é que SmCB1 responde pela maior parte desta atividade em pH > 4 (CAFFREY *et al.*, 2004). Além disso, baseados em resultados de *Southern blot* utilizando o cDNA completo de Sjasp (codificando os 380 resíduos do zimogênio) como sonda, os autores concluíram que o gene de Sjasp estaria presente numa única cópia no genoma de *S. japonicum* e que um gene homólogo deveria estar presente no genoma de *S. mansoni*.

Uma comunicação do grupo australiano por Wong *et al.* (1997) relatou a seqüência de uma aspartil proteinase de *S. mansoni* (originalmente denominada Smasp por esses autores, SmCD no sistema de CAFFREY *et al.*, 2004) identificada a partir da triagem de uma biblioteca de cDNA de vermes adultos, utilizando o cDNA completo de Sjasp como sonda. Um clone positivo com um inserto de 428 códons foi encontrado, codificando uma pauta de leitura livre (ORF) completa, exceto pelas duas primeiras bases do códon de iniciação ATG. Conforme esperado, a seqüência de aminoácidos de Smasp apresentou elevada similaridade com a enzima ortóloga Sjasp (84% id. em 379 resíduos) e semelhança mediana (50-55%) com enzimas ortólogas de vertebrados, como humanos e aves, e invertebrados, como *Aedes aegypti*. Entretanto, uma diferença crucial foi encontrada na seqüência de Smasp em relação à enzima de *S. japonicum* e demais catepsinas D conhecidas, uma extensão de 43 resíduos no C-terminal. Assim, foi predita a massa de 41,2 kDa para os 377 resíduos da protease madura.

Em função deste interessante resultado, Wong *et al.* resolveram pesquisar o número de cópias do gene de S_{masp} no genoma de *S. mansoni* e verificar a presença de um gene homólogo no genoma de *S. japonicum* como uma segunda forma de S_{jasp}. Três sondas foram utilizadas, correspondendo ao cDNA de S_{masp} completo, ao domínio catalítico conservado e à extensão C-terminal. Baseados nos padrões de hibridização obtidos com os genomas digeridos com 3 diferentes enzimas de restrição, os autores concluíram que o gene de S_{masp} estaria presente em uma ou poucas cópias em *S. mansoni*. Por outro lado, os autores concluíram que um gene homólogo também estaria presente no genoma de *S. japonicum*, aparentemente contrariando os resultados relatados anteriormente pelo mesmo grupo (BECKER *et al.*, 1995), uma vez que a “forma longa” agora encontrada do gene S_{jasp} deveria ter hibridizado também com a sonda da “forma curta”, utilizada anteriormente.

Mesmo com esta inconsistência, os autores concluíram que “esquistossomos expressam ao menos duas formas de aspartil proteases catepsina D-símile; uma forma com uma extensão C-terminal e outra sem esta extensão”. Tornando a situação mais confusa ainda, uma errata ao artigo original de Becker *et al.* (1995) foi publicada afirmando que a seqüência anteriormente publicada para S_{jasp} tinha na verdade 428 resíduos e não 380, como dito anteriormente. Mais ainda, a protease madura seria tão longa quanto a seqüência de S_{masp}, incluindo os 43 códons homólogos referentes à extensão C-terminal. Os autores comentaram que o engano ocorreu em função de um erro no eletroesferograma do sequenciamento do DNA de S_{jasp}. Isto põe em questão a real existência de duas formas dos genes das aspartil proteases catepsina D-símile identificadas em esquistossomos. Como veremos adiante, em trabalhos subseqüentes do grupo australiano e de nenhum outro grupo, a questão das formas longa e curta destas enzimas não mais foi abordada. Assim, para que as supostas formas das enzimas catepsina D-símile de *S. mansoni* e *S. japonicum* possam ser corretamente identificadas, deste ponto em diante será adotado o sistema de nomenclatura de CAFFREY *et al.*, (2004), onde as formas longas serão denominadas SmCD1.1 e SjCD1.1 e as formas curtas SmCD1.2 e SjCD1.2, respectivamente (Fig. 21). O prefixo pro será usado para designar o zimogênio destas enzimas, e.g., proSmCD1.1.

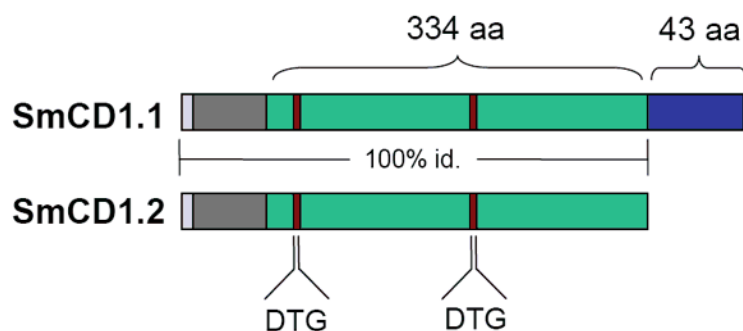


Fig. 21: Relação estrutural entre as duas formas possíveis da proteinase catepsina D-símile de *S. mansoni*. A forma longa (SmCD1.1) distingue-se da forma curta (SmCD1.2) exclusivamente pela extensão C-terminal de 43 resíduos (azul escuro). O restante da molécula, incluindo as regiões pré (ou peptídeo sinal, azul claro), pro (cinza) e a protease madura (verde), incluindo o sítio catalítico (vermelho) são idênticas entre as duas formas.

Posteriormente, uma comunicação por [Cesari et al. \(1998\)](#) relatou os usuais ensaios de atividade hemoglobinolítica com extratos ácidos (AE; 1M NaCl e pH 3,8) e salinos (SE; 0,85% NaCl) de vermes machos, fêmeas e de ambos os sexos da cepa venezuelana JL de *S. mansoni*. Experimentos em solução na ausência de tióis, indicaram uma atividade específica 2,5-4,0 vezes maior para AE em relação à SE, passível de inibição por pepstatina (7 μ M) e DAN (12 μ M). Estes autores observaram uma clivagem específica do monômero das Hbs bovina e humana (16 kDa), produzindo fragmentos de 10 e 4-5 kDa, com uma maior atividade sobre a proteína bovina. Este resultado destoa do relatado anteriormente para as cepas filipinas (*S. japonicum*) ou porto-riquenhas (*S. mansoni*) por [Becker et al. \(1995\)](#), que observaram uma degradação extensa e inespecífica da Hb humana. Cesari et al. relataram uma incapacidade da fração IgG de soro anti-catepsina D bovina de inibir a atividade hemoglobinolítica de AE, apesar destes anticorpos reconhecerem uma espécie em *Immunoblots*. Ainda, os autores comentaram a inabilidade de soros de pacientes infectados por *S. mansoni* de identificar esta mesma espécie. A falta de inibição de degradação *in vitro* da Hb por soro heterólogo anti-catepsina D contraria o descrito anteriormente com extratos de *S. japonicum* por [Bogitsh et al. \(1992\)](#). Por fim, os autores também relataram o isolamento de uma espécie de 45 kDa a partir de AE usando cromatografia de afinidade em pepstatina-agarose. A atividade específica e demais detalhes da purificação não foram

descritos, mas os autores comentaram sobre a facilidade de degradação desta preparação em polipeptídeos de menor massa molecular.

Outra comunicação, desta vez pelo grupo australiano, relatou atividade catepsina D-símile em 3 formas evolutivas de *S. japonicum*: ovos, miracídios e vermes adultos machos, fêmeas e de sexos misturados, utilizando como substratos o-aminobenzoil-Ile-Glu-Phe(NO₂)-Phe-Arg-Leu-NH₂ e Hb humana (VERITY *et al.*, 1999). *Imunoblots* com soro de coelho elaborado contra proSmCD1.1 recombinante (46 kDa), produzida como corpos de inclusão (CI) em *E. coli*, foram capazes de detectar bandas de 47 e 41 kDa em todos os estágios de crescimento de *S. japonicum* testados. *Southern blots* usando sonda de 530 pb de região do gene de SjCD1.1 abrangendo parte do domínio catalítico e parte da extensão C-terminal foram realizados contra cDNA obtido do mRNA das 3 formas evolutivas pesquisadas, mostrando a presença marcante do transcrito nas formas adultas, em menor extensão em miracídios e em baixa quantidade em ovos. Estes resultados foram compatíveis com os níveis de atividades hemoglobinolítica e peptidásica detectadas, onde também se puderam diferenciar os níveis de atividade das formas adultas: fêmea > sexos misturados > machos.

1.2.4.b Eficácia da vacinação com SjCD

Verity *et al.* (2001) verificaram que a vacinação de camundongos com SjCD1.1 expressa em *E. coli* e em células de inseto resulta numa redução significativa, porém singela no número de vermes total (21-38%) ou em fêmeas separadamente (22-40%). A vacinação foi, no entanto, incapaz de reduzir a fecundidade. Anticorpos dos isotipos IgG1, IgG2a e IgG2b foram encontrados em níveis elevados nos animais vacinados, mas nenhuma correlação entre estes níveis e a eficácia da proteção foi observada. Mais ainda, o soro imune dos camundongos vacinados não foi capaz de inibir a degradação *in vitro* da Hb humana pela protease recombinante e a transferência passiva do soro ou anticorpos dos animais vacinados antes e depois do desafio com os parasitas, não reduziu significativamente o número de vermes ou ovos nos animais que receberam a transfusão. Em função destes resultados, os autores concluíram que anticorpos podem não ter um papel contundente na proteção (moderada) observada e que este efeito pode ser devido à uma combinação de respostas humorais e celulares.

1.2.4.c Especificidade de substrato

Brindley *et al.* (2001) fizeram uma importante contribuição para o estudo da especificidade de substrato das enzimas catepsina D-símile de esquistossomos ao determinar os sítios de clivagem na Hb humana. Para isto, as proteases de *S. mansoni* e *S. japonicum* foram produzidas em sistema de expressão em células de inseto, obtendo-se espécies migrando com ~ 40 kDa em SDS-PAGE, a partir da fração do sobrenadante de cultura ligada à pepstatina-agarose. Estas espécies mostraram-se ativas contra substrato fluorogênico de catepsina D e Hb humana e reagiram com soro de coelho elaborado contra uma forma truncada de 366 resíduos (Glu³⁷ até Gly³²⁹) de SjCD1.1. Desta forma, apesar dos autores não terem especificado na publicação original, pode-se inferir que as espécies de ~40 kDa purificadas por cromatografia de afinidade em pepstatina-agarose correspondam às formas maduras de SjCD1.1 e SmCD1.1. Os autores interpretaram que as enzimas puderam se auto-ativar após secreção no sobrenadante de cultura (26 °C) ou durante a permanência no tampão de ligação na coluna de afinidade (pH 3,6 a 4 °C). O sequenciamento N-terminal das enzimas maduras (SjCD1.1 = NVEPRPE e SmCD1.1 = GVDPQPE) revelou que a pro-região era constituída por 35 aminoácidos, dois a menos que o predito antes pela análise da seqüência (BECKER *et al.*, 1995). Através de microscopia confocal de cortes de vermes adultos marcados com soro de coelho anti-SjCD1.1, a enzima foi localizada na gastroderme de ambos os sexos, porém com maior intensidade na fêmea. Ao nível subcelular, as eletromicrografias revelaram a presença das enzimas nas cisternas do retículo endoplasmático rugoso (RER), em vacúolos digestivos superficiais e estruturas membranares associadas.

Treze sítios de hidrólise nas cadeias α e β de Hb humana foram determinados para a enzima de *S. japonicum*, contra 15 para a enzima de *S. mansoni*. A maior parte dos 17 sítios preferidos de hidrólise na Hb humana foi compartilhada entre as enzimas de esquistossomos (11 ao todo), mas não pela enzima ortóloga humana (**Tabela 1**). Os sítios de clivagem iniciais reconhecidos pelas enzimas de esquistossomos são α Phe33/Leu34, α Leu109/Ala110 e β Phe41/Phe42. Como a maior parte das aspartil proteases, resíduos hidrofóbicos são preferidos nas posições P1 e P1' dos sítios de hidrólise na Hb humana reconhecidos por SjCD1.1 e SmCD1.1. Entretanto, ambas as

enzimas exibem uma preferência por prolina em P1' (sítio de clivagem α Phe36/Pro37) incomum para aspartil proteases eucarióticas e caracteristicamente exibida por proteases retrovirais.

A habilidade de SjCD1.1 recombinante de degradar especificamente IgG humana em 3 sítios principais e indiscriminadamente as proteínas séricas albumina e componente C3 do complemento sugere uma função desta protease em mecanismos de defesa dos esquistossomos contra o sistema imune do hospedeiro e na digestão de proteínas séricas como fonte de aminoácidos para seu desenvolvimento (VERITY *et al.*, 2001). A enzima utilizada nestes estudos *in vitro* foi a forma longa recombinante expressa em células de inseto. Os sítios de clivagem identificados compreendiam resíduos hidrofóbicos em P1 e P1', de acordo com o observado anteriormente com a Hb humana (BRINDLEY *et al.*, 2001).

Tabela 1: Sítios de clivagem na hemoglobina humana, reconhecidos pelas catepsinas D de esquistossomos (SjCD1.1 e SmCD1.1) e humana (hCD).

Sítio	SjCD1.1	SmCD1.1	hCD*
Cadeia- α			
V ¹ -LSPA			✗
AGEY ²⁴ -GAEA			✗
AEAL ²⁹ -ERMF	✗	✗	
LERM ³² -FLSF			✗
ERMF ³³ -LSFP	✗	✗	
FLSF ³⁶ -PTTK	✗	✗	
SHCL ¹⁰⁵ -LVTL	✗	✗	
LLVT ¹⁰⁸ -LAAH	✗	✗	✗
LVTL ¹⁰⁹ -AAHL	✗	✗	
LDKF ¹²⁸ -LASV	✗		
ASVS ¹³³ -TVLT		✗	✗
STVL ¹³⁷ -TSKY			✗
Cadeia- β			
VHL ³ -TPEE		✗	
TPEE ⁷ -KSAS			✗
VTAL ¹⁴ -WGKV	✗	✗	✗
TALW ¹⁵ -GKVN		✗	
LGRL ³¹ -LVVY	✗	✗	✗
GRL ³² -VVYP	✗	✗	✗
TQRF ⁴¹ -FESF	✗	✗	✗
FESF ⁴⁵ -GDLS	✗		✗
PDAV ⁵⁴ -MGNP			✗
KGTF ⁸⁵ -ATLS		✗	
VQAA ¹²⁹ -YQKV	✗	✗	

* Fruitier *et al.* (1998).

1.2.5 Inibição de enzimas do catabolismo da hemoglobina como estratégia de desenvolvimento de novos fármacos esquistossomicidas

Desde a identificação da “atividade hemoglobinasase” no intestino de esquistossomos por [Timms & Bueding \(1959\)](#), compostos capazes de inibir a degradação de hemoglobina vêm sendo buscados ([ZUSSMAN & BAUMAN, 1971](#)). A elucidação das enzimas envolvidas neste processo foi crucial, para a consideração do catabolismo da Hb como fonte de novos alvos para o desenho de compostos esquistossomicidas. O bloqueio da proteólise da Hb do hospedeiro através de inibidores de aspartil e cisteinil proteases resulta em marcantes efeitos anti-esquistossomais *in vitro* e *in vivo* ([BOGITSH *et al.*, 1992](#); [WASILEWSKI *et al.*, 1996](#)), demonstrando assim o papel essencial desta via no metabolismo do parasito e indicando que inibidores seletivos direcionados para as proteases chaves nesta via poderiam constituir a base para uma nova estratégia de quimioterapia da esquistossomose. Para isto são essenciais estudos em que a especificidades destas proteases são determinadas e comparadas com as enzimas ortólogas humanas ([BRINDLEY *et al.*, 2001](#); [MATHIEU *et al.*, 2002](#)). Estes dados podem ser subsequentemente analisados à luz das peculiaridades na estruturas 3D das proteases do parasito e do hospedeiro humano, obtidas por cristalografia de raios-X, RMN ou modelagem por homologia (e.g. [BRADY *et al.*, 2000](#)).

Avanços recentes na biologia molecular de eucariotos agora permitem o bloqueio da expressão de genes específicos em esquistossomos, conforme demonstrado com SmCB1.1 por [Skelly *et al.* \(2003\)](#) utilizando a técnica de interferência de RNA. Este tipo de abordagem é importante, pois fornece um meio de questionar a função de genes específicos na biologia do parasito. Entretanto, a ferramenta de interferência de RNA pode trazer revelações desagradáveis sobre a complexidade de vias metabólicas em eucariotos superiores, como esquistossomos. Por exemplo, [Skelly *et al.* \(2003\)](#) observaram por diversos métodos que apesar da expressão de SmCB1.1, uma das proteases mais importantes, senão a mais importante, na degradação da Hb por *S. mansoni*, ter sido efetivamente bloqueada, nenhum efeito adverso foi detectado em esquistossômulos cultivados *in vitro*. Isto indica que intervenções quimioterápicas almejando a inibição de proteases envolvidas na digestão de Hb por esquistossomos

precisam considerar a redundância nesta via metabólica, provavelmente proporcionada por genes parálogos destas enzimas.

Além de esquistossomos, outros parasitas que dependem da digestão da Hb para a sobrevivência têm empregado aspartil proteases neste processo, e.g. *Plasmodium falciparum* (plasmepsinas; BERRY *et al.*, 1994) e *Haemonchus contortus* (LONGBOTTOM *et al.*, 1997). Outros agentes infecciosos também já tiveram aspartil proteases identificadas, apresentando, porém, funções diversas tais como maturação de outras proteases e hormônios, além da digestão proteolítica da Hb do hospedeiro vertebrado: *Ancylostoma caninum* (Acasp; HARROP *et al.*, 1996), *Onchocerca volvulus* (JOLODAR & MILLER, 1998) e *Strongyloides stercoralis* (GALLEGO *et al.*, 1998). Invariavelmente, estas proteases apresentaram similaridade considerável com as aspartil proteases lisossomais (catepsinas D e E), renina e pepsina. Enquanto que as plasmepsinas parecem formar um grupo mais divergente destas enzimas, as proteases de *A. caninum* e, em particular, as de *Schistosoma sp.* são mais próximas evolucionariamente da catepsina D, do que de qualquer outra aspartil protease (BECKER *et al.*, 1995). De fato, as catepsinas D de mamíferos participam entre outras coisas, do catabolismo da Hb de eritrócitos senescentes (FRUITIER *et al.*, 1998; BEYER & DUNN, 1996).

Plasmepsinas têm sido exploradas como alvo no desenvolvimento de antimaláricos há algum tempo (revisado por KESAVULU *et al.*, 2005). Estes estudos têm sido em muito beneficiados pela disponibilidade de estruturas cristalográficas de diferentes plasmepsinas de algumas espécies de *Plasmodium* (revisado por MEHLIN, 2005). Até o momento, estruturas 3D de alta resolução não estão disponíveis para as enzimas catepsina D-símile de esquistossomos. Entretanto, a similaridade significativa com as seqüências de aspartil proteases de estrutura 3D conhecida, como as enzimas humanas catepsina D, renina e pepsina, viabiliza a obtenção de modelos teóricos confiáveis das aspartil proteases destes parasitos. Este tipo de abordagem teria importância fundamental na elucidação da relação estrutura-função na especificidade de substrato exibida por estas enzimas, como primeiro passo para o desenho de uma nova estratégia de desenvolvimento de compostos esquistossomicidas.

2 OBJETIVO

2.1 *Geral*

Aprofundar o conhecimento atual sobre enzimas catepsina D-símile de *S. mansoni* usando, principalmente, abordagens bioquímicas e computacionais.

2.2 *Específicos*

- Obter um modelo por homologia da estrutura 3D forma curta de SmCD1 e aplicá-lo no estudo da especificidade de substrato desta enzima considerando os sítios de hidrólise identificados na hemoglobina humana;
- Caracterizar bioquimicamente aspartil proteases purificadas do extrato aquoso de vermes adultos de *S. mansoni*;
- Avaliar as eficácias de inibidores comerciais de HIV-1 protease na quimioterapia experimental da esquistossomose mansônica em camundongos;
- Identificar *in silico* genes de aspartil proteases presentes nos bancos de dados dos projetos genoma de *S. mansoni*;
- Clonar e expressar os genes das formas longa e curta de proSmCD1 em *Escherichia coli* e *Pichia Pastoris*;
- Clonar e expressar em *E. coli* a extensão carboxi-terminal de 43 resíduos de SmCD1;
- Produzir antisoros policlonais específicos contra as formas longa, curta e o C-terminal de SmCD1, visando estudos imunocitoquímicos;
- Caracterizar estruturalmente o domínio C-terminal de SmCD1;
- Caracterizar bioquimicamente as isoformas recombinantes de SmCD1, visando o estudo da especificidade de substrato;
- Racionalizar, através de modelagem molecular, a especificidade de substrato determinada experimentalmente.

3 MATERIAIS E MÉTODOS

Esta seção é consubstanciada pelo ANEXO I, onde são encontradas as metodologias empregadas na modelagem por homologia da forma curta de SmCD1 e no estudo de modelagem molecular da especificidade por substrato desta enzima.

3.1 Reagentes

Membranas de nitrocelulose foram adquiridas de Bio-Rad Laboratories (Hercules, CA-EUA). Tripsina grau proteoma, agarose de baixa EEO e os demais reagentes grau P.A. foram obtidos de Sigma, Inc. (St Louis, MO-EUA). Reagentes gerais para biologia molecular, quando não especificados, foram obtidos de Promega (Madison, WI-EUA). Enzimas de restrição foram adquiridas junto a Invitrogen (Carlsbad, CA-EUA) ou Promega. Kits para purificação de DNA foram adquiridos de Qiagen (Venlo, KJ-Holanda) ou Promega.

3.2 Procedimentos gerais

3.2.1 Métodos empregados em proteínas

3.2.1.a SDS-PAGE

Para as análises por SDS-PAGE utilizou-se uma adaptação do sistema de tampões de Laemmli (1970) e o sistema de eletroforese em gel vertical Mini-protean II (Bio-Rad Laboratories, California). Resumidamente, alíquotas de amostras de proteínas eram diluídas em tampão de aplicação de amostra 4x concentrado (80 mM Tris-HCl pH 6,8 contendo 2 % SDS, 12 % glicerol e 0,015 % azul de bromofenol) em condições redutora (5% β -mercaptoetanol adicionado ao tampão de amostra) ou não-redutora. As corridas eram rotineiramente feitas aplicando-se uma diferença de potencial constante de 120 V. As proteínas eram reveladas com o corante Coomassie Brilliant Blue G-250 ou com um kit comercial de revelação por impregnação com prata (Bio-Rad Laboratories, California).

3.2.1.b Immunoblots

Para a detecção imunológica de amostras separadas eletroforéticamente (*immunoblots*), as proteínas presentes nos geles de poliacrilamida eram inicialmente eletrotransferidas para membranas de nitrocelulose. As transferências eram realizadas

com auxílio do sistema Transblot semi-dry transfer cell (Bio-Rad Laboratories, California) usando voltagem constante de 15V, por 40-60min, dependendo da faixa de tamanho da proteína de interesse e da espessura (0.75 - 1.5 mm) e porosidade (concentração de acrilamida de 10-15%) do gel. Em seguida, a eficiência da transferência era verificada corando-se as proteínas na membrana com o corante Ponceau S e comparando-se as proteínas remanescentes no gel reveladas como usual. Em seguida, a etapa de bloqueio era realizada incubando-se as membranas em solução salina tamponada com fosfato (PBS: 140 mM NaCl + 2,7 mM KCl + 10 mM Na₂HPO₄ + 1,8 mM KH₂PO₄, pH 7,4) ou Tris-HCl (TBS: 50 mM Tris-HCl pH 7,5 + 150 mM NaCl) contendo 5% de leite em pó desnatado e 0,05% tween-20 (polisorbato 80) por 2h, 25 °C, sob agitação ou por 16-18h (*overnight*) em geladeira. A etapa de ligação do anticorpo primário era realizada retirando-se as membranas da lavagem, transferindo-se diretamente para solução diluída do soro policlonal (1:25 até 1:2.000) em TTBS (TBS + 0,05% tween-20) e incubando-se por 2h, 25 °C, sob agitação. Após 3 ciclos de 5min de lavagem com TTBS, as membranas eram incubadas a 25 °C, sob agitação, por 1-2h em solução diluída em TTBS do anticorpo secundário conjugado à fosfatase alcalina (1:2.000 para revelação com substrato cromogênico ou 1:20.000 para revelação com substrato quimioluminescente). Após 5 ciclos de 5min de lavagem com TBS (sem detergente) a membrana estava pronta para a revelação. A detecção era feita com kit de detecção colorimétrico (Bio-Rad Laboratories, California) ou quimioluminescente (Pierce), seguindo as recomendações dos respectivos fabricantes.

3.2.1.c Estimativa do conteúdo protéico

O conteúdo protéico das amostras foi rotineiramente estimado pelos métodos de [Lowry \(1951\)](#) ou de [Bradford \(1976\)](#), conforme implementado nos kits “DC Protein Assay” e “Protein Assay” (Bio-Rad Laboratories, California), respectivamente, ou acompanhado pela medida da D.O.²⁸⁰, dependendo da agilidade requerida ou da compatibilidade com o tampão em que a amostra encontrava-se dissolvida.

3.2.1.d Precipitação de proteínas

Amostras de proteínas a serem analisadas por eletroforese ou espectrometria de massas, contendo altas concentrações de sais, principalmente cloretos como os de guanidina e sódio, eram rotineiramente precipitadas com TCA, numa adaptação do

método de [Ouellette et al. \(2003\)](#). Brevemente, adicionava-se 10% v/v de TCA 6 M gelado à amostra de proteína em tubo de microcentrifuga. Agitava-se vigorosamente e depois se mantinha em gelo por 15min – 30min. Adicionava-se, então, 3 volumes de TCA 0,6 M gelado e agitava-se vigorosamente. Em seguida, centrifugava-se por 10min a 10.000g, 25 °C e descartava-se o sobrenadante. Para a lavagem do precipitado, adicionava-se 1 mL de acetona a -20 °C, seguido de agitação vigorosa para ressuspender todo o precipitado. Então, centrifugava-se como anteriormente, reduzindo-se o tempo para 5min e ao final descartava-se o sobrenadante. A etapa de lavagem era repetida pelo menos mais uma vez. O precipitado era seco ao ar livre por cerca de 18h ou em estufa a 37 °C por 2h.

3.2.1.e Hidrólise trípica em solução

A hidrólise enzimática de amostras de proteínas visando identificação por busca em bancos de dados usando o mapa de peptídeos trípticos detectado por espectrometria de massas é conhecida como *peptide mass fingerprinting* (PMF). As hidrólises trípicas visando PMF foram rotineiramente realizadas em tubos de microcentrifuga siliconizados incubando-se a amostra em tampão 10 mM (NH₄)₂CO₃ (pH 8) contendo 1–10 µg de proteína com 0,1 µg de tripsina (grau análise proteômica) por 1-4 h a 37 °C. Posteriormente, as amostras eram secas em sistema Speed-vac (Thermo Lifescience, Basingstoke, UK) e finalmente ressuspendidas em 10 µL de água deionizada.

3.2.1.f Hidrólise trípica *in situ* de proteínas em gel

A hidrólise trípica diretamente em amostras de proteínas em gel de poliacrilamida foi feita com ligeira modificação do método de [Shevchenko et al. \(1996\)](#).

3.2.1.g Sequenciamento N-terminal por degradação de Edman automatizada

O seqüenciamento da região N-terminal de proteínas foi realizado num seqüenciador de fase gasosa PSQ-1 (Shimadzu), através do método de microsseqüenciamento das proteínas imobilizadas em um suporte de PVDF ([DE-SIMONE et al., 1994](#)).

3.2.2 Métodos empregados em ácidos nucleicos

Técnicas usuais de trabalho com DNA e RNA foram realizadas de acordo com Sambrook *et al.* (1989).

3.2.2.a Eletroforese em gel de agarose

Geles de agarose para análise de amostras de DNA eram rotineiramente preparados em concentrações de acrilamida variando de 0,6 a 1,0 % agarose (baixa eletroendosse) em tampão TAE (40 mM Tris base + 0,115% v/v ácido acético glacial + 1 mM EDTA, pH 8,0) ou TBE (45 mM Tris-borato + 1 mM EDTA). As corridas eram feitas com voltagem constante (80V). A revelação era realizada incubando-se, após a corrida, o gel em solução 1 µg/mL de brometo de etídio por 10min. O gel podia então ser visualizado em transiluminador com radiação U.V. e fotografado.

3.2.2.b Estimativa do conteúdo

Amostras de RNA foram estimadas espectrofotometricamente usando a relação $1 \text{ D.O.}^{260} = 40 \mu\text{g/mL RNA}$. A leitura das amostras a 260 nm, diluídas 100 – 1000 vezes em água DEPC (água deionizada + 0,05% v/v dietilpirocarbonato, autoclavada), já descontada da leitura a 320 nm era efetivamente usada na relação previamente citada. A qualidade do RNA era avaliada pela razão $\text{O.D.}^{260}/\text{O.D.}^{280}$, onde amostras de RNA pura fornecem uma razão próxima de 2,0.

As amostras de DNA foram quantificadas espectrofotometricamente usando a relação $1 \text{ D.O.}^{260} = 50 \mu\text{g/mL}$. Alternativamente, a quantidade de DNA numa amostra era estimada pela intensidade da banda em geles de agarose revelados com brometo de etídio.

3.2.2.c Sequenciamento de DNA

As reações de seqüenciamento foram realizadas em equipamentos ABI prism 377 (PE Biosystems) usando o kit de reação BigDye Terminator Cycle Sequencing, com DNA polimerase AmpliTaq (enzima FS) (PE Biosystems) seguindo-se o protocolo fornecido pelo fabricante.

3.2.3 Espectropolarimetria de dicroísmo circular (CD)

Para este experimento, utilizou-se um espectropolarímetro J-720 (JASCO Inc., Easton, EUA). As medidas foram realizadas a T.A em cubetas de 0,2 cm de caminho ótico e os dados de dicroísmo coletados na faixa espectral de 260 – 195 nm com resolução de 0,5 – 1,0 nm. Os espectros obtidos foram resultantes de 3 acumulações e subtração do espectro do respectivo tampão. O sinal de CD em mdeg foi convertido em elipsidade ($\theta_{mrw,\lambda}$) ou absortividade molar residual média ($\Delta\epsilon$) para usar programas de predição de estrutura secundária. A conversão em $\theta_{mrw,\lambda}$ ($\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$) foi feita a partir da fórmula abaixo:

$$\theta_{mrw,\lambda} = \frac{mrw * CD}{10 * L * c}, \text{ onde } mrw \text{ é o "peso médio do resíduo"} = MM/(NR-1), \text{ com}$$

MM = massa molecular (g/mol) e NR = número de resíduos; CD é o sinal de CD (mdeg), L é o caminho ótico (cm) e c a concentração (mg/mL). Para obter os dados em $\Delta\epsilon$ ($\text{M}^{-1}\cdot\text{cm}^{-1}$) usou-se a relação $\theta_{mrw,\lambda} = 3298 * \Delta\epsilon$.

3.2.4 Espectrometria de massas por ionização/desorção a laser assistida pela matriz (MALDI) acoplada à detecção por tempo de voo (MALDI-TOF)

As análises foram feitas em espectrômetros ABI-4700 Proteomics Analyser (MALDI-TOF/TOF), instalados no Departamento de Farmacodinâmica da FIOCRUZ e no CEBIME/LNLS. As amostras de proteínas eram inicialmente dissolvidas em solução 50% ACN + 0,3% TFA e posteriormente diluídas com a solução da matriz (5 mg/mL ácido α -ciano-4-hidroxicinâmico em 50% ACN + 0,3% TFA) para concentrações de 10 - 0,1 pmol. Usualmente, 1 μL de cada amostra misturada à matriz era aplicada à placa de MALDI. Os dados foram adquiridos com laser Nd:YAG com frequência de repetição de 200 Hz e até 4000 disparos foram acumulados para cada espectro. Utilizou-se apenas calibração externa e os demais parâmetros de operação do instrumento foram os fornecidos pelo fabricante.

3.2.5 Produção de anti-soro policlonal em coelho

Para a produção dos anti-soros, usou-se as dependências do biotério de experimentação do Instituto Oswaldo Cruz. Os coelhos albinos com 2 semanas de vida usados nos experimentos eram obtidos do Centro de Criação de Animais de Laboratório

(CECAL) da FIOCRUZ. Tipicamente, um protocolo de imunização consistia em 4 inóculos subcutâneos seguidos com intervalo de 7 dias, onde o primeiro e segundo inóculos eram feitos com antígeno emulsificado em adjuvante completo de Freund (500 µL do antígeno em PBS + 500 µL do adjuvante) e os dois últimos com antígeno emulsificado em adjuvante incompleto. A partir da 3ª semana de imunização, o título do anti-soro era acompanhado por *immunoblot*, empregando amostras da proteína usada como antígeno. Para antígenos com mais de 20 kDa, usava-se de 10 - 100 µg de proteína pura ou parcialmente purificada (> 90% de pureza) por inóculo. Alternativamente, para proteínas insolúveis em tampão aquoso, macerava-se as bandas das proteínas recortadas de um gel de poliacrilamida, que em seguida eram ressuspensas direto na emulsão. Para peptídeos puros com menos de 10 kDa, usava-se de 100 - 500 µg por inóculo. Ao término da imunização, os animais eram anestesiados e sacrificados por punção cardíaca, obtendo-se rendimentos de 50 – 100 mL de soro hipermune.

3.2.6 Bioinformática

As buscas em bancos de dados foram realizadas usando programas da família BLAST (ALTSCHUL *et al.*, 1990), implementados no NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>) ou ExPASy Proteomics Server (<http://ca.expasy.org/>). Os alinhamentos múltiplos foram executados usando-se o programa ClustalX para Windows (TOMPSON *et al.*, 1997). Alinhamentos de seqüências foram manipulados no programa BIOEDIT (TOM HALL, Ibis Pharmaceuticals). As ferramentas de análise proteômica usadas foram as disponíveis no ExPASy. Predição de sítios de glicosilação foram realizadas nos servidores WWW NetNGlyc 1.0 e NetOGlyc 3.5 (GUPTA *et al.*, 1999). A análise da frequência de utilização de códons por um dado organismo foi realizada no programa “*Graphical Codon Usage Analyser*” de T. Schödl (FUHRMANN *et al.*, 2004). A manipulação e análise de estruturas de proteínas foi conduzida nos programas WebLab Viewer Lite (Accelrys, San Diego, CA) e SwissPDB Viewer (SCHWEDE *et al.*, 2003).

3.3 Fracionamento de extratos de vermes adultos por cromatografia de afinidade em pepstatina-agarose

3.3.1 Preparo do extrato solúvel

Vermes adultos liofilizados (160 mg cedidos pelo Dr. Fábio Ribeiro, do Laboratório de Esquistossomose, CPqRR, FIOCRUZ, Minas Gerais) foram macerados em gral e pistilo sob adição periódica de nitrogênio líquido. Adicionou-se Tampão de Solubilização (TS: 100 mM tampão NaOAc pH 3,5 + 1 M NaCl) suficiente para avolumar a amostra para 50 mL. Em seguida, o material suspenso em TS foi submetido a ultracentrifugação a 100.000 g por 30min a 4 °C. O sobrenadante, denominado extrato aquoso de vermes (EAV) foi filtrado a 0,45 µm e reservado para a cromatografia de afinidade.

3.3.2 Cromatografia de afinidade em pepstatina-agarose

Os 50 mL de EAV foram aplicados a um fluxo de 0,5 mL/min em coluna contendo 5 mL de resina pepstatina-agarose, previamente equilibrada com 100 mL de tampão de solubilização/ligação. Recolheu-se o material não-ligado e lavou-se com tampão de ligação até obter uma D.O.²⁸⁰ do eluato inferior a 0,01 U.A. O efluente da lavagem foi recolhido em frações de 10 mL para posterior análise por SDS-PAGE. Em seguida, a eluição do material ligado especificamente à coluna foi realizada com tampão 100 mM Tris-HCl pH 8,2 + 1 M NaCl (tampão de eluição), recolhendo-se frações de 1,0 - 1,5 mL. As frações correspondendo ao pico de D.O.²⁸⁰ foram reunidas (totalizando aproximadamente 6 mL) e dializadas por 18h em câmara fria (4 - 8 °C) contra 1 L de 10 mM (NH₄)₂CO₃ (pH 8).

3.3.3 Atividade enzimática

A atividade enzimática da fração ligada à pepstatina-agarose foi testada incubando-se 1 a 5 µg de proteína com 12,5 µM do substrato de catepsina D, Phe-Ala-Ala-Phe(4-NO₂)-Phe-Val-Leu(4-piridilmetil) éster (Sigma), em tampão 100 mM NaOAc pH 3,5, num volume final de 200 µL. Transcorridos 30min de incubação a 37 °C, a reação foi parada com aquecimento (10min / 90 °C) e a D.O.³¹⁰ lida. Um ensaio em branco foi realizado parando-se a reação imediatamente após ter sido disparada pela

adição do substrato. A atividade específica foi expressa em μM de substrato hidrolisado por minuto por mg de proteína usando o $\Sigma^{310} = 800 \text{ M}^{-1} \cdot \text{cm}^{-1}$.

3.3.4 Eletroforese bidimensional (2DE)

A fração de EAV ligada à pepstatina-agarose (FLPA) foi liofilizada e ressuspendida em tampão de solubilização para 2DE contendo 8 M uréia + 4% CHAPS + 60 mM DTT + 0,002% azul de bromofenol (O'FARRELL *et al.*, 1975). A amostra solúvel (125 μL) foi carregada em fita de gradiente imobilizado de pH 3-10 não-linear de 7 cm (IPG strip, Amersham Biosciences), *overnight*, durante a etapa de reidratação em tampão 8 M uréia + 0,5% (p/v) CHAPS + 0,2% (p/v) DTT + 0,5% (v/v) tampão IPG + 0,002% azul de bromofenol. A primeira dimensão da separação das proteínas foi realizada em um sistema de focalização isoeétrica IPGphor (Amersham Biosciences), usando o seguinte programa: 0-500V em 30min e manutenção por 0,25 kV/h; 500-100V em 30min e manutenção por 0,5 kV/h; 1000-5000V em 100min e manutenção por 7,5 kV/h. Para a segunda dimensão, a fita foi equilibrada em 10 mL de tampão de equilíbrio (50 mM Tris-HCl pH 8,8 + 2% SDS + 6 M uréia + 30% (v/v) glicerol + 0.002% azul de bromofenol) + 1% DTT por 15min. Em seguida, a fita foi re-equilibrada em 10 mL de tampão de equilíbrio + 2,5% iodoacetamida. Finalmente, a fita foi aplicada no topo de um mini-gel de poliacrilamida a 15% e a SDS-PAGE conduzida como descrito no item 3.2.1.a. O gel foi revelado pelo método da impregnação pela prata.

3.4 ***Quimioterapia experimental com inibidores comerciais de HIV-1 protease***

3.4.1 Geral

Estes experimentos foram realizados em colaboração com o Dr. Fabio Ribeiro (Lab. Esquistossomose, CPqRR) e com o Dr. Arnaldo Maldonado-Jr (Medicina Tropical, IOC/FIOCRUZ). Foi utilizada a cepa LE mantida nestes laboratórios. Os fármacos testados foram doações do Instituto de Tecnologia de Fármacos (Farmanguinhos, FIOCRUZ-RJ) ou do Instituto de Pesquisas Evandro Chagas (IPEC, FIOCRUZ-RJ). Os medicamentos foram manipulados no Laboratório de Bioquímica de Proteínas e Peptídeos para o preparo das doses. Códigos foram atribuídos para cada

fármaco, incluindo um controle positivo (PZQ) e um negativo (NaCl), e estes foram enviados para o Laboratório de Esquistossomose do CPqRR.

3.4.2 Infecção experimental

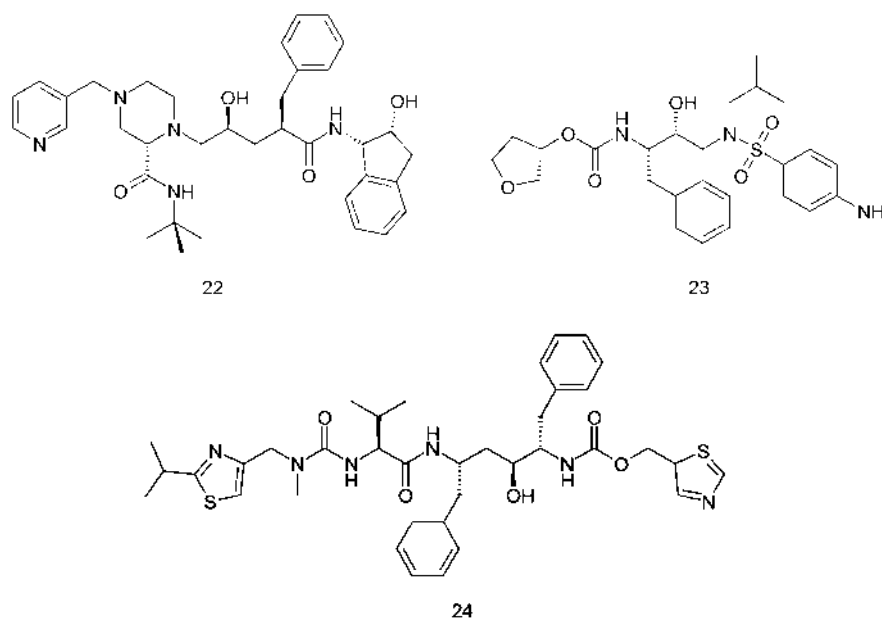
Camundongos de 4 semanas de vida foram infectados experimentalmente por exposição da calda a 100 ± 10 cercárias.

3.4.3 Grupos e tratamento experimental

Cada fármaco foi testado em três doses diferentes, com diferentes posologias (**Tabela 2**). O efeito de cada dose foi avaliado em grupos de 10 camundongos. O tratamento era iniciado 45 dias após a infecção e aplicado por 5 dias. Foi feito 1 grupo controle, com 10 camundongos cada, i.e., para cada tipo de veículo empregado (solução de ácido cítrico, cremofor EL e DMSO).

Tabela 2: Doses e posologia dos fármacos.

Fármaco	Forma Farmacêutica	Doses (mg/kg)	Posologia (vezes ao dia)
Indinavir (22)	Sólido (pó)	20; 100	3
Amprenavir (23)	Líquido (gel)	450	2
Ritonavir (24)	Líquido (gel)	100	2
Nelfinavir	Sólido (pó)	50; 250	2
PZQ	Sólido (pó)	25; 125; 250	1



3.4.4 Forma de administração

Para os fármacos em pó, dissolvia-se a dose em volume de solução de ácido cítrico 0,05 M referente a 5 mL/kg de peso do camundongo e administrava-se por sonda oral. Os fármacos veiculados em gel eram suspensos em volume de solução a 1% de Cremophor EL referente a 5 mL/kg de peso do camundongo e administrados por sonda oral.

3.4.5 Análise dos resultados

Quinze dias após o tratamento, realizava-se perfusão do fígado e mesentério para coleta e contagem dos vermes. Também era contabilizado o número de vermes que permaneciam no fígado por exame microscópico deste material. A eficácia de cada fármaco/dose foi avaliada através da redução no número de vermes em relação ao grupo controle (camundongos tratados com o veículo para administração dos fármacos testados).

3.5 Pesquisa in silico de transcritos pertencentes à família gênica de aspartil proteases em S. mansoni (SmAPs)

3.5.1 Bancos de dados

Para este levantamento de transcritos de SmAPs foram pesquisados os bancos de ESTs do programa de descoberta de genes de *S. mansoni* — SmGI 5.0, liberado em 7 de Novembro de 2003 e do projeto transcriptoma de *S. mansoni* liderado pelo Dr. Sérgio Verjovski (USP, São Paulo) — “*Schistosoma mansoni EST genome project*” (SmEST), contendo as seqüências publicadas no final de 2003 (VERJOVSKI *et al.*, 2003). O primeiro destes bancos foi desenvolvido pela WHO/UNDP/World Bank *Schistosoma Genome Network*, enquanto que o segundo foi abastecido por uma rede de laboratórios do Estado de São Paulo, Brazil (ONSA network).

3.5.2 Metodologia

Utilizando os serviços disponíveis na página eletrônica de cada banco, buscas com os programas blastn e balstp ou tblastn foram realizadas usando, respectivamente, as seqüências do cDNA (código de acesso no Genebank U60995) e de aminoácidos de

pré-proSmCD1. As seqüências homólogas identificadas usando como critério $E < 0,1$ foram retiradas e catalogadas quanto à sua composição (número de ESTs compondo as seqüências-consenso) e origem (estágio do ciclo do parasita usado na construção da biblioteca). Buscas também foram realizadas através das palavras-chaves: “peptidase”, “protease”, “proteínase”, “aspartic”, “aspartyl”, “pepsin”, “cathepsin” e “renin”, visando obter eventuais seqüências que não apresentam similaridade estatisticamente significativa com SmCD1, mas que aparecem anotadas como aspartil proteases nos bancos de dados. Para os agrupamentos de ESTs interrelacionados (*clusters*) idênticos à seqüência do cDNA de SmCD1 depositado no Genbank, foi observado se a região alinhada entre o EST e o cDNA correspondia ao domínio N-terminal ou ao domínio C-terminal de SmCD1, com o intuito de identificar diferenças nos níveis de expressão das formas longa e curta da enzima.

Diferenças no nível de expressão entre diferentes estágios do desenvolvimento do parasito também foram pesquisadas para os transcritos de SmAPs mais abundantes presentes no SmGI. Isto foi realizado usando ferramenta disponível no próprio SmGI que busca por transcritos observados com freqüências significativamente (usando como critério estatístico o parâmetro $R > 9$) distintas entre dois grupos definidos pelo usuário. Cada grupo representa uma ou mais bibliotecas de cDNA estágio-específicas (de onde os ESTs foram derivados) selecionadas. Para o propósito desta análise foram feitas duas comparações: (i) macho versus fêmea e (ii) vermes adultos (bibliotecas “8 semanas”, “adulto”, “macho”, “fêmea” e “maduro sexualmente”) versus os demais estágios do crescimento (bibliotecas “cercaria”, “ovos”, “3 horas”, “sétimo dia”, “larva”, “pulmão”, “miracídio” e “esporocisto”).

3.6 Clonagem dos genes das formas longa e curta de SmCD1

3.6.1 Desenho dos oligonucleotídeos (*primers*)

Usando-se a seqüência de DNA depositada no Genbank sob código U60995.1 codificando a pre-pro-enzima SmCD1 (WONG *et al.*, 1997), *primers* foram desenhados visando a amplificação das formas longa e curta da pro-enzima (proSmCD1) a partir de amostras de RNA. O *primer* 5', comum às duas formas, foi desenhado com 33 bases, das quais 27 anelam com a região pro do gene codificando a enzima: SmCDEcoRI/NdeI

5' GAA TTC **CAT ATG** GTG GTT AGG ATC CCT CTT CAT CCG TTA 3'. O oligonucleotídeo apresenta um sítio de Eco RI (sublinhado) para subsequente clonagem em vetor pPIC9K para expressão em *P. pastoris*, sítio de Nde I (emoldurado) para clonagem em vetores pET23 ou pET28 para expressão em *E. coli* e ainda um códon de metionina (negrito) para início da expressão da pro-enzima. Para clonagem da forma longa, desenhou-se oligonucleotídeo com 32 bases anelando com o mRNA da enzima contendo a seqüência C-terminal completa: SmCD-longa 5' TCA AAC TTC ATC TGA AAG AAG TTT AGA AAA GG 3'. Este *primer* não contém sítios de restrição, pois será utilizado o sítio de Not I do vetor pGEM-T (Promega) para subclonar o cDNA nos vetores de expressão (vide abaixo). Como a região é rica em AT, foi necessário desenhar um *primer* mais longo p/ ter uma temperatura de anelamento mais alta e aumentar as chances de anelamento específico. Para clonagem da forma curta, utilizou-se *primer* anelando com a região do mRNA codificando o COOH-terminal da enzima, menos os 43 últimos resíduos de aminoácido: SmCD-curta 5' **TTA** TGG ATG AAG AGC TTT CGC AAA ACC GAC AC 3'. Em negrito está o códon de terminação inserido para truncar o gene da forma longa.

3.6.2 Vermes

Os vermes usados na preparação de RNA foram cedidos pelo Dr. Sérgio Verjovski (Depto. Bioquímica/USP) a partir dos ciclos biológicos mantidos no Instituto Adolfo Lutz (São Paulo). Espécimes adultos foram obtidos a partir de perfusão hepática de hamsters e camundongos com 45 dias de infecção. Os vermes, em PBS, foram acondicionados em gelo por 10h antes da extração do RNA.

3.6.3 Extração do RNA

Para extração do RNA total utilizou-se uma adaptação do método de [Chomczynski & Sacchi \(1987\)](#). Cerca de 2500 vermes, adultos e de ambos os sexos, (700 mg peso úmido) foram transferidos para homogeneizador tipo Potter-Elvehjem de vidro boro-silicato e pistilo de teflon, com capacidade de 5 mL. Os vermes foram congelados em nitrogênio líquido e submetidos a dois ciclos de congelamento-descongelamento. Em seguida, adicionou-se 7 mL do reagente Trizol (Gibco) e procedeu-se a homogeneização do material. Transferiu-se o homogenato para tubo de

centrífuga e a partir daí seguiu-se as instruções fornecidas pelo fornecedor do reagente para extração de RNA total. O precipitado de RNA obtido ao final do processo foi ressuspensão em 100 µL de água DEPC, quantificado e armazenado em alíquotas. Alíquotas de trabalho foram mantidas a -20 °C até o uso, enquanto que as demais foram armazenadas a -80 °C.

3.6.4 Otimização das condições de RT-PCR

A reação da transcriptase reversa acoplada à reação em cadeia da polimerase (RT-PCR) foi usada para obtenção e amplificação do cDNA das formas longa e curta de proSmCD1 a partir do RNA total de vermes adultos e de ambos os sexos de *S. mansoni*. Dois kits comerciais foram utilizados na tentativa de amplificar os transcritos desejados: “Access RT-PCR System” (Promega) e “Thermoscript RT-PCR System” (Life Technologies). O primeiro kit faz a síntese e amplificação do cDNA em um único passo (single-step), enquanto que o último trabalha as duas etapas separadamente. Em cada kit seguiu-se as recomendações gerais de uso do fabricante. Para o kit “single-step” utilizou-se as seguintes condições de PCR: 1 ciclo 48 °C, 45min (síntese do cDNA) / 1 ciclo 94 °C, 2min (inativação da RT e desnaturação do DNA) / 40 ciclos 94 °C, 30s (desnaturação do DNA); 60 °C, 1min (anelamento); 68 °C, 2min (extensão) / 1 ciclo 68 °C, 7min (extensão final) / 1 ciclo 4 °C, ∞ (fim da reação). Com o kit thermoscript, primeiramente desnaturou-se o RNA a 65 °C por 5min (desnaturação do RNA), seguido de resfriamento em banho de gelo. Logo após, o RNA foi submetido à transcrição reversa de acordo com as instruções do kit, incubando-se a 55 °C, 60min e parando-se a reação por desnaturação térmica a 85 °C, 5min. Após tratamento com Rnase H (37 °C, 20min) para remoção do excesso de RNA, a amostra de cDNA foi amplificada por PCR, seguindo as instruções do kit e usando a seguinte programação do termociclador: 1 ciclo 94 °C, 2min (desnaturação inicial do DNA) / 35 ciclos 94 °C, 30s (desnaturação do DNA); 55 °C, 2min (anelamento); 72 °C, 90s (extensão) / 1 ciclo 72 °C, 7min (extensão final) / 1 ciclo 4 °C, ∞ (fim da reação). Além dos controles dos reagentes fornecidos em cada kit (oligos e RNA), utilizou-se como controles da preparação de RNA, 3 pares de nucleotídeos desenhados para amplificar fragmentos de aproximadamente 0,5 kb de genes de *S. mansoni* com alto (actina), médio (catepsina S) e baixo (receptor de retinol) nível de transcrição. Além dos próprios kits, os parâmetros variados para se otimizar a

amplificação dos genes de interesse foram: quantidade de RNA e cDNA, temperatura de anelamento, quantidade dos *primers*, tempo de extensão e número de ciclos de amplificação na reação de PCR.

3.6.5 Clonagem em vetores de passagem (pGEM-T)

3.6.5.a Purificação dos fragmentos e ligação

Os produtos de RT-PCR foram purificados por precipitação em NaOAc/etanol ou utilizando-se o kit QIAquick (Qiagen) previamente à ligação em plasmídeo pGEM-T. Este vetor carrega uma deoxitimidina não-pareada em suas extremidades 3'. Isto permite que se tire proveito da adição de uma única deoxiadenosina ao terminal 3' dos fragmentos amplificados por certas DNA polimerase termoestáveis, como a que foi usada na RT-PCR deste trabalho. Seguindo-se as instruções do kit pGEM-T\$Vector System I (Promega), para a reação de ligação utilizou-se 10 µL do produto de RT-PCR purificado (30-150 ng), 2 µL plasmídeo (100 ng), 4 µL T4 DNA ligase (12u Weiss) e 4 µL do tampão 5x concentrado fornecido no kit. Incubou-se overnight a 16 °C.

3.6.5.b Transformação, seleção dos clones positivos e determinação da orientação

O produto de ligação (5 µL) foi usado na transformação de células de *E. coli* DH5α por eletroporação, seguindo-se protocolo padrão na literatura (SAMBROOK *et al.*, 1989). As células transformadas foram plaqueadas em meios LBAMP (LB + ampicilina sódica 50 µg/mL) + X-Gal (80 µg/mL) + 0,5 mM IPTG, e incubadas overnight a 37 °C. Os clones positivos foram inicialmente identificados pela coloração branca das colônias e a presença dos insertos das formas longa e curta de SmCD1 posteriormente confirmadas por PCR de colônia (SAMBROOK *et al.*, 1989). Culturas (5 mL) incubadas por 12-16h (37 °C, 200 rpm) dos clones confirmados por PCR (4 - 6 clones) foram usadas na preparação de amostras de DNA plasmidial (miniprep), conforme protocolo de lise alcalina descrito na literatura (SAMBROOK *et al.*, 1989). Os plasmídeos assim obtidos (pGEM-T/proSmCD1.1 e pGEM-T/proSmCD1.2) foram submetidos a digestão com Nde I (Invitrogen) a fim de se determinar a orientação dos insertos (100 ng do plasmídeo + 2,5 u Nde I em tampão "React3 10x"). Os clones com orientação correta (sítio de Not I em 5' e sítio de Nde I em 3') tiveram amostras de DNA plasmidial submetidas ao sequenciamento.

3.7 Produção das formas longa e curta de proSmCD1 em *E. coli*

3.7.1 Clonagem em vetores de expressão pET28a

3.7.1.a Preparação dos plasmídeos

Para subclonagem em vetor pET28a tirou-se proveito do sítio de Nde I inserido pelo *primer* 5' e o sítio de Not I presente na região de clonagem de pGEM-T, 3' ao inserto. O kit “QIAprep Spin Miniprep” (Qiagen) foi usado na preparação de pGEM-T/proSmCD1.1 e pGEM-T/proSmCD1.2 a partir de 5 mL de culturas (12-16h / 37 °C em LB-AMP) de clones *E. coli* DH5 α carregando estes plasmídeos. Usando o kit “Plasmid Midi” (QIAGEN), obteve-se a preparação do plasmídeo pET28a a partir de 250 mL de cultura (12-16h / 37 °C em LB-AMP) de *E. coli* DH5 α em LB-KAN (LB + 50 mg/mL sulfato de kanamicina). Os protocolos indicados pelo fabricante de ambos os kits baseiam-se no método da lise alcalina (SAMBROOK *et al.*, 1989) e usam resinas catiônicas para purificação do DNA em colunas, por centrifugação ou pressão manual. O rendimento de cada preparação foi avaliado pela intensidade da banda revelada em geles de agarose 0,8%.

3.7.1.b Digestão com enzimas de restrição

Primeiramente, digeriu-se o vetor pET28a (1 μ g em 30 μ L) e os vetores pGEM-T/proSmCD1.1 e pGEM-T/proSmCD1.2 (250-500 ng em 30 μ L) com as duas enzimas (15u de Nde I e 30u de Not I), ambas as reações em um volume final de 100 μ L em tampão react3 (Invitrogen), por 4h a 37 °C. Controles da digestão de pET28a com cada uma das enzimas em separado, também foram realizados. Após análise do progresso da reação por eletroforese em gel de agarose 0,8%, complementou-se as digestões de pET28a e dos vetores baseados em pGEM-T com mais 15u de Nde I e 30u de Not I, respectivamente. Além disso, adicionou-se 25u de Nde I à digestão dos vetores pGEM-T/proSmCD1.1 e pGEM-T/proSmCD1.2. Procedeu-se com a incubação a 37 °C por mais 16-18h.

3.7.1.c Purificação dos fragmentos e ligação

Os produtos das reações de digestão (~ 120 µL) foram submetidos a uma eletroforese em gel de agarose 0,8% preparativo e as bandas referentes aos insertos liberados (1,4 kb para a forma longa e 1,2 kb para a forma curta) e ao vetor pET28 duplamente digerido (5-6 kb) foram recortadas e processadas com o kit QIAquick para extração do DNA. Para a reação de ligação dos fragmentos referentes aos genes das formas longa e curta em pET28a, usou-se uma proporção 1:1 (vetor:inserto) e incubou-se com a enzima T4 DNA ligase por 1h a 25 °C ou 16-18h/16 °C, num volume final de 10 µL em tampão da enzima (Gibco).

3.7.1.d Transformação em *E. coli* DH5α, seleção dos clones positivos e preparação dos plasmídeos

Os produtos das reações de ligação foram usados na transformação de células de *E. coli* DH5α termocompetentes, seguindo protocolo padrão na literatura (SAMBROOK *et al.*, 1989). As células expostas ao DNA foram plaqueadas em LB-KAN e cultivadas a 37 °C durante a noite. As colônias obtidas foram submetidas à confirmação da presença do inserto por PCR de colônia. Os clones positivos foram inoculados em 5 mL de meio LB-KAN e cultivados durante a noite (37 °C, 200 rpm) para subsequente purificação do DNA plasmidial por miniprep, usando o kit “QIAprep Spin Miniprep”.

3.7.1.e Transformação em *E. coli* BL21(DE3)

Os plasmídeos (pET28a/proSmCD1.1 e pET28a/proSmCD1.2) obtidos como descrito no item anterior foram usados para transformar células de *E. coli* BL21(DE3) por eletroporação, as quais foram então plaqueadas em LB-KAN e cultivadas a 37 °C durante a noite. As colônias obtidas foram inoculadas em 2 mL de meio LB-KAN e cultivadas por 12h /37 °C para preparação de estoques em 15% de glicerol e utilização no teste de expressão.

3.7.2 Correção do clone pET28a/proSmCD1.1

Para a correção da mutação da base 995 e da deleção do códon 1228-1230 no clone pET28a/proSmCD1.1 os seguintes *primers* foram desenhados: SmCD-AccI 5' CTC GGA GCC ACT CGT CTA CCC GG 3' e SmCD-NotI/corr 5' GCGGCCGC TCA AAC TTC ATC TGA **AAG AAG** TTT AGA AAA GGC GAA T 3' (reverso). No

primeiro oligonucleotídeo, o sítio de restrição com Acc I aparece sublinhado, enquanto que no segundo, os códons relativos à correção da deleção na região 3' do gene aparecem em negrito. A estratégia para correção da forma longa consistiu em primeiramente amplificar a partir do sítio de Acc I (posição 840, contando a partir da 1ª base do gene da pro-enzima) até o C-terminal com os *primers* acima descritos usando o vetor pGEM-t/proSmCD1.1 (clone 3A) como substrato da PCR e em seguida clonar em vetor de passagem pGEM-t (obtendo o clone pGEM-t/CatD_Acc). As condições da PCR e os procedimentos para clonagem em pGEM-t foram semelhantes aos itens 3.6.4 e 3.6.5. Com isso, as duas mutações seriam corrigidas, reduzindo-se as chances de inserir nova mutação pela ação da DNA polimerase, devido ao menor tamanho do fragmento amplificado. Ao mesmo tempo, ter-se-ia um clone com a extensão C-terminal correta para futura subclonagem em vetor de expressão pBuff5, conforme descrito no item 3.8.1 adiante.

3.7.2.a Subclonagem da forma longa corrigida no vetor pET28a

A subclonagem do gene corrigido da forma longa de SmCD1 em vetor de expressão pET28a foi conseguida adotando-se procedimentos análogos aos descritos no item 3.7.1. Resumidamente, digeriu-se o vetor pET28a/SmCD1.2 com Nde I e Acc I (Invitrogen), e isolou-se o fragmento (nucleotídeos 1-840) a partir da banda em gel de agarose. Um segundo fragmento (nucleotídeos 841-1245) foi semelhantemente isolado após digestão de pGEM-t/CatD_Acc com as enzimas Acc I e Not I. O terceiro e último fragmento foi obtido pela digestão de vetor pET28a com Nde I e Not I. Os três fragmentos de DNA linear foram ligados e usados para transformar células competentes de *E. coli* DH5 α . Os clones corretos foram identificados por PCR de colônia seguido de análise de restrição e finalmente seqüenciados para confirmar as correções efetuadas.

3.7.3 Teste de indução da expressão em *E. coli* BL21(DE3)

3.7.3.a Indução e coleta das amostras

Para o teste de indução, culturas de 16-18h a 37 °C de *E. coli* BL21(DE3) pET28a/SmCD1.1 e pET28a/SmCD1.2 em LB-KAN foram usadas como pré-inóculo de culturas de 50 mL em meio LB-KAN. Para o teste de indução a 30 °C usou-se 0,7 mL de inóculo, enquanto que para as culturas induzidas a 37 °C foram feitos inóculos de 0,5

mL. As culturas em Erlenmeyers de 250 mL foram cultivadas em agitadores orbitais (200 rpm) termostatizados nas referidas temperaturas até que a D.O.⁶⁰⁰ atingisse 0,7 - 0,9 U.A.. Neste momento, uma alíquota de 15 mL de cada cultura foi centrifugada a 4 °C / 5000g / 10min, o sobrenadante descartado, o pellet de células rotulado como “Tº de indução” e armazenado a -20 °C. Imediatamente após a retirada desta alíquota, adicionou-se IPTG 1M suficiente para obter uma concentração final de 1 mM. A partir daí, amostras de diferentes volumes foram coletadas em diferentes tempos (1h – 10 mL, 2h – 7 mL, 3h – 5 mL e 4h – 5 mL) e processadas como descrito acima.

3.7.3.b Extratos

Os extratos das amostras dos diferentes tempos de indução foram obtidos como descrito a seguir. Retirou-se as células do freezer e ressuspendeu-se em 1 ml de tampão A (10 mM Tris-HCl pH 7,4 + 100 mM NaCl + 1 mM de EDTA + 0,5 mM de PMSF). Adicionou-se 5 µL de lisozima (10 mg/ml) e incubou-se em banho de gelo por 20-30min. Após verificar a lise das células através do aumento da viscosidade do meio, procedeu-se a homogeneização e fragmentação do DNA liberado com 3 ciclos de 15s/60s de sonicação/descanso em gelo, com o intuito de evitar a desnaturação protéica ao superaquecer a amostra. As amostras foram então centrifugadas a 20.000g / 4 °C / 10min. Nos casos em que os pellets apresentaram-se muito grandes, verificando-se que a lise não foi boa, ressuspendeu-se novamente em tampão A e repetiu-se o procedimento. Os sobrenadantes das amostras dos diferentes tempos e temperaturas de indução testados foram transferidos para novo tubo e rotulados como “extrato solúvel”. Os pellets obtidos foram ressuspendidos em 0,5 ml de tampão B (tampão A + 2% Triton X-100), com auxílio de pipeta. Sonicou-se, como descrito acima, por mais 2 vezes e centrifugou-se como anteriormente. O sobrenadante foi removido e procedeu-se com a lavagem dos pellets em tampão B por mais uma vez. O pellet lavado foi ressuspendido em 50 µl de tampão A e misturado a 50 µl de tampão de amostra de SDS-PAGE.

3.7.4 Teste de solubilização de corpos de inclusão (CI)

3.7.4.a Preparação dos CI

As amostras de corpos de inclusão das formas longa e curta de proSmCD1 usadas no teste de solubilização foram preparadas conforme descrito a seguir: 1 L de

LB-KAN foi inoculado com 25 mL de cultura 16-18h a 37 °C e cultivado na mesma temperatura sob agitação constante (200 rpm) até O.D.⁶⁰⁰ entre 0,7 e 0,9 U.A.. A expressão das proteínas recombinantes foi então induzida pela adição de IPTG 1 mM e a cultura incubada nas mesmas condições por mais 4h. As células foram coletadas por centrifugação a 5.000g / 10min / 4 °C e lavadas 2 vezes com 100 mL de meio LB novo. O pellet lavado de células foi então pesado e congelado a -20 °C até o uso no dia seguinte. O pellet de células (cerca de 6 g) ressuspenso em 30 mL de tampão A (50 mM Tris-HCl pH 7,5 + 100 mM NaCl + 1 mM EDTA pH 8,2 + 1 mM PMSF) foi adicionado de 150 µL de lisozima (10 mg/mL) e incubado por 1h a 4 °C. O material foi então homogeneizado em 5 ciclos de 1min sonicação / 1min banho de gelo. Em seguida, os CI foram coletados por centrifugação a 10.000g / 20min / 4 °C, ressuspendidos em 20 mL de tampão B (tampão A + 2M uréia + 1% Triton X-100) e sonicados como anteriormente. Os CI foram novamente coletados por centrifugação e procedeu-se a lavagem mais uma vez. Os CI foram finalmente ressuspendidos em tampão C (tampão A + 5% glicerol) e armazenados a -20 °C até o uso.

3.7.4.b Ensaio

Para a seleção da melhor condição de solubilização dos CI de proSmCD1, foram avaliados o efeito do pH, agente solubilizante e aditivos, totalizando 32 tampões diferentes: 8M uréia em pH 3-10 (50 mM glicina-HCl pH 3 e 4; 50 mM NaOAc pH 5 e 6; 50 mM Tris-HCl pH 7, 8, 9 e 10); 4, 5 e 6M de guanidina-HCl, 0,5%, 1,0% e 2,5% de sarcosil e 4, 6 e 8M uréia em 50 mM Tris-HCl pH 7,5; 0,1M, 0,5M e 0,75M NaCl ou 0,5%, 1,0% e 1,5% de detergente não-iônico (sarcosil, octilglicosídeo, Brij-35 ou CHAPS) + 8M uréia em 50 mM Tris-HCl pH 7,5. O teste consistiu em misturar 50 µL da suspensão de CI (15 mg/mL) a 1 mL do tampão de solubilização, agitar vigorosamente e incubar por 30min a 25 °C sob leve agitação. O material solubilizado era obtido após centrifugação a 20.000g / 10min / 16 °C. A concentração de proteína solúvel obtida em cada ensaio era estimada pela D.O.²⁸⁰. As amostras também foram analisadas por SDS-PAGE. A eficiência de cada condição de solubilização foi expressa em função de uma condição padrão (50 mM Tris-HCl pH 7,5 + 1% SDS).

3.7.5 Teste de purificação por IMAC

Para o ensaio de purificação por cromatografia de afinidade com metal imobilizado (IMAC), um pellet de CI (30 mg de proteína), preparados como descrito no item anterior, foi ressuspensão em 4 mL de tampão A (50 mM Tris-HCl pH 8,5 + 100 mM NaCl + 6M guanidina-HCl) + 10 mM de imidazol (tampão de equilíbrio/solubilização) e permaneceu sob agitação suave por 6h a 25 °C. O sobrenadante da centrifugação a 20.000g / 10min / 16 °C foi filtrado a 0,22 um (filtro PSU Whatman) e posto sob agitação suave por 20min a 25 °C com 5 mL de resina Ni-IDA agarose (Probond, Invitrogen), previamente equilibrada com tampão A + 10 mM de imidazol. Ao final, transferiu-se a mistura para coluna plástica de 10 mL (Bio-rad) e deixou-se escoar todos os 8 mL do material não-ligado. A coluna foi então lavada com aproximadamente 100 mL de tampão de lavagem (tampão A + 20 mM imidazol) equilíbrio até que a D.O.²⁸⁰ do lavado fosse inferior a 0,05 U.A.. As proteínas ligadas especificamente à resina foram eluídas com 100 mM imidazol em tampão 50 mM Tris-HCl pH 8 + 100 mM NaCl + 4M guanidina-HCl + 5% (v/v) glicerol. O eluato (20 mL) foi coletado em frações de 1 mL e a D.O.²⁸⁰ monitorada. As frações tiveram alíquotas de 60 uL precipitadas por TCA/acetona que foram posteriormente analisadas por SDS-PAGE a 12%. As frações contendo as pro-enzimas das formas longa e curta de SmCD1 foram reunidas e o conteúdo protéico estimado pelo método de Bradford (vide item 3.2.1.c).

3.7.6 Análise do estado de agregação em 4M de guanidina-HCl por cromatografia de exclusão molecular (SEC)

As preparações das formas longa e curta de proSmCD1 obtidas por cromatografia em Ni-IDA, conforme descrito no item anterior, foram submetidas a SEC em condições desnaturantes para avaliação do estado de agregação em 4M de guanidina-HCl. A coluna Superdex-200 HR 10/30 (Pharmacia) foi previamente equilibrada a um fluxo de 0,5 mL/min com tampão de corrida (50 mM Tris-HCl pH 8 + 100 mM NaCl + 4M guanidina-HCl + 5% glicerol). As amostras de proSmCD1 (250 µL), solubilizadas neste mesmo tampão, foram aplicadas e a corrida cromatográfica desenvolvida por 50min a 0,5 mL/min em um sistema Äkta purifier (Pharmacia),

monitorando-se a D.O.²⁸⁰. As proteínas anidrase carbônica (29 kDa) e BSA (66 kDa) foram usadas como padrões para estimativa da massa molecular.

3.7.7 Teste de renaturação por diluição

Para este ensaio e os descritos nos itens 3.7.8.a e 3.7.8.b, a preparação de C.I. e a cromatografia em Ni-IDA para cada uma das preparações de proSmCD1 (longa e curta) foram tipicamente realizados como descrito nos itens 3.7.4.a e 3.7.5, respectivamente. Quatro mililitros de amostras de proSmCD1.1 e proSmCD1.2 eluídas da coluna Ni-IDA (1 mg/mL) foram diluídas, gota a gota, em 1 L de tampão 50 mM fosfato de sódio pH 7,0. O material foi mantido sob agitação em câmara-fria por 48h, quando foi concentrado até 10 mL em sistema Amicon de 500 mL de capacidade, pressurizado com nitrogênio e em banho de gelo, usando-se membrana de 30 kDa.

3.7.8 Triagem de condições de renaturação

3.7.8.a Ensaio preliminares

Para ambas as formas de proSmCD1, as variáveis experimentais mais críticas foram testadas: (i) forma de renaturação: diluição ou diálise; (ii) tempo de diálise necessário para retirar quase a totalidade dos aditivos (500 mM guanidina-HCl, 200 mM L-arginina, 2 mM CaCl₂, 2 mM MgCl₂) e agentes desnaturantes (4 M guanidina) remanescentes; (iii) concentração inicial de proteína (0,6 e 0,06 mg/mL). Os aditivos empregados foram escolhidos com base no potencial de atuarem como interferentes nas etapas subsequentes da análise.

Para as amostras renaturadas por diálise, 250 µL da solução de proteína, já na concentração a ser testada, foram transferidos para o aparato adaptado de microdiálise. Este último consistia em tubos de microcentrífuga vazados na tampa e cobertos por membrana de diálise (10 – 12 kDa). Após aplicação da amostra os tubos eram vedados com auxílio da própria tampa e filme plástico tipo *parafilm*. Os tubos de microdiálise, assim preparados em duplicata, eram imersos de ponta-cabeça em 12,5 mL do tampão (50 mM Tris-HCl pH 8,2) contendo os aditivos acima descritos em tubos cônicos de 50 mL. A diálise era realizada durante a noite, por cerca de 12h, em câmara-fria. Para as amostras renaturadas por diluição, 50 µL foram pipetados para 950 µL do tampão de renaturação contendo os aditivos em tubo de microdiálise. A seguir, juntamente com as

amostras dializadas durante a noite, os tubos de microdiálise eram transferidos para 500 mL de tampão 0,1 M NaOAc pH 3,5 + 0,1 M NaCl (TAC) e a microdiálise retomada. Com 4h de diálise, uma alíquota de 50 µL foi retirada de cada microtubo e congelada até o seu processamento. Ao fim de 8h de diálise, os 200 µL restantes das amostras e as alíquotas de 50 µL retiradas com 4h foram submetidos a centrifugação por 30min a 20.000g / 4 °C. O sobrenadante foi reservado para análise por SDS-PAGE a 12% e ensaios enzimáticos sobre substrato fluorogênico, como descrito no item 3.7.8.d adiante, e sobre Hb bovina.

Para os ensaios de atividade proteolítica sobre Hb bovina, 25 µL de amostra eram adicionados a 100 µL de TAC e incubado a 37 °C por 30min. A seguir, a reação era disparada pela adição de 125 µL de solução aquosa de Hb 2% (p/v). Ao fim de 30min de incubação a 37 °C, paralisava-se a reação pela adição de 250 µL de TCA 20% gelado e as amostras mantidas em banho de gelo por 20min. A seguir, as amostras eram centrifugadas a 10.600g por 10min a 18 °C e o sobrenadante lido a 280 nm para quantificação dos peptídeos derivados da Hb. Ensaios em branco eram feitos parando-se a reação de cada amostra no tempo zero de incubação.

3.7.8.b Desenho fatorial fracionado

Na tentativa de se identificar condições mais favoráveis para a renaturação das isoformas de proSmCD1 utilizou-se um desenho fatorial fracionado baseado em [Armstrong *et al.* \(1999\)](#). Os dez fatores avaliados, em dois níveis cada, em 16 condições experimentais estão descritos nas **Tabelas 3 e 4**. Nestas tabelas é seguida a notação padrão onde “+” e “-” denotam um fator em seus níveis “alto” e “baixo”, respectivamente. O conjunto de símbolos “-” e “+” na **Tabela 3** é denominado “matriz de análise” ou “matriz do modelo”. A parte da tabela formada pelas colunas X1-X10 é denominada “tabela de planejamento” ou “matriz de planejamento”.

Num desenho experimental chamado “fatorial completo”, todas as combinações dos níveis de um fator são testadas e o número total de experimentos necessários é $N = k \cdot L \cdot n$, onde k é o número de fatores, L é o número de níveis e n corresponde ao número de réplicas. Para desenhos em dois níveis, há $n2^k$ ensaios. Um desenho fatorial fracionado também apresenta um número de experimentos que é potência de dois: $N =$

$n \cdot 2^{k-p}$, onde $p < k$. Estes desenhos experimentais são ditos ortogonais, o que significa que as estimativas dos efeitos não são correlacionadas. Em outras palavras, ao remover um efeito da análise, os efeitos restantes permanecem os mesmos. O efeito principal de um fator XI é calculado fazendo-se a média das respostas em todos os ensaios em que este tinha um valor “+” e subtraindo-se deste a média das respostas de todos os ensaios em que este tinha um valor “-”. O desenho experimental codificado na **Tabela 3** é ortogonal e isto pode ser constatado verificando-se que o produto de quaisquer duas colunas é zero. Além disso, o desenho é balanceado, pois há um número idêntico de experimentos para cada nível de um dado fator. Isto é importante, pois maximiza a sensibilidade de testes estatísticos subsequentes.

Caso todos os dez fatores descritos na **Tabela 3** fossem avaliados em seus dois níveis e sem réplicas, seriam necessários $2^{10} = 1024$ experimentos. No desenho fatorial fracionado descrito neste trabalho, almejou-se avaliar os mesmos 10 fatores com apenas 16 ensaios, ou seja, 1/64 dos experimentos necessários ao fatorial completo. Desta forma, o desenho poderia ser descrito como 2^{10-6} , o qual é numericamente igual a $2^4 = 16$. Esta informação forneceu o primeiro passo para a construção da **Tabela 3** que foi começar com uma matriz de planejamento de um fatorial 2^4 completo. A matriz de análise correspondente foi construída conforme a chamada “ordem padrão”. De acordo com esta ordem, a primeira coluna dos fatores ($X1$) começa com “-” e alterna em sinal através de todos os 2^{k-p} ensaios. A segunda coluna ($X2$) inicia com “-” repetido duas vezes, então alterna com dois seguidos no sinal contrário até que todas as 2^{k-p} posições sejam preenchidas. A terceira coluna ($X3$) começa com “-” repetido 4 vezes, depois “+” repetido 4 vezes e assim por diante. Em termos gerais, a i -ésima coluna (X_i) inicia com 2^{i-1} repetições de “-” seguido de 2^{i-1} repetições de “+”.

Com as colunas dos 4 fatores principais montadas ($X1 - X4$), as colunas dos 6 fatores restantes foram construídas introduzindo-se interações de 1ª e 2ª ordem das colunas $X5 - X10$. Por exemplo, a coluna de interação $X7$ foi obtida multiplicando-se algebricamente a entrada na linha relevante em $X1$ pelo valor em $X2$ para obter o valor na linha correspondente em $X7 = X1 * X2$. Em outra notação, o “gerador” para esta coluna foi $7 = 12$. Num desenho fatorial com 2^{k-p} experimentos normalmente há p geradores para definir p colunas adicionais de fatores. Definindo-se I como uma coluna feita

somente de “+”, decorre que o produto de quaisquer duas colunas num desenho fatorial é igual a I . Assim, pode-se escrever o gerador $7 = 12$ na forma “ $I = 127$ ”, multiplicando-se ambos os lados da equação por 7. Nesta notação, os demais geradores do planejamento experimental descrito neste trabalho foram: $I = 1235$, $I = 1246$, $I = 348$, $I = 149$ e $I = 2410$. A coleção total de geradores do desenho 2^{10-6} , conhecida como “a relação de definição”, compreendeu 11 (ou $2^p - 1$) “palavras” que puderam ser formadas multiplicando-se quaisquer duas, três, quatro, cinco ou seis destas 6 palavras originais. Isto é importante, pois o tamanho da menor palavra numa relação de definição é conhecido como a “resolução” do desenho experimental.

A resolução descreve o grau com que a estimativa dos “efeitos principais” é confundida com as estimativas das interações de 2ª, 3ª e de mais alta ordem. Quanto maior a resolução de um desenho experimental, mais livre de erros é a estimativa dos efeitos principais e dos efeitos de menor importância (e de maior ordem). A resolução do desenho fatorial fracionado empregado neste trabalho foi III, o que significa que os fatores principais não foram confundidos com outros fatores principais, mas interações de dois fatores (como em X7) foram confundidas com estes. Desta forma, somente os fatores principais e as interações de 2ª ordem não confundidas com estes poderiam ser incluídos no modelo. Interações de mais alta ordem foram consideradas nulas ou desprezíveis. Assim sendo, na matriz de planejamento contida na **Tabela 3** há 4 fatores principais, 4 interações de dois fatores e 2 interações de três fatores. O modelo de regressão linear completo relaciona estes fatores da seguinte forma: $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_{321}X_5 + \beta_{421}X_6 + \beta_{12}X_7 + \beta_{34}X_8 + \beta_{14}X_9 + \beta_{42}X_{10}$, onde $\beta_{321} = \beta_5$, $\beta_{421} = \beta_6$, $\beta_{12} = \beta_7$, $\beta_{34} = \beta_8$, $\beta_{14} = \beta_9$ e $\beta_{42} = \beta_{10}$. Optou-se por gerar um modelo inicial incluindo apenas os fatores principais e agregar interações de 2ª ordem somente no caso de haver uma indicação pela análise do modelo inicial.

Matrizes de análise construídas como descrito anteriormente são sempre ortogonais. A “ordem de execução” que aparece na **Tabela 3** reflete a ordem randômica com que os experimentos foram realizados. Na verdade, o planejamento experimental descrito na **Tabela 3** é uma simplificação, pois os experimentos foram realizados em duplicata, totalizando 32 condições em ordem aleatória. Uma representação gráfica das

especificações completas do planejamento experimental adotado é encontrada na **Figura 22**.

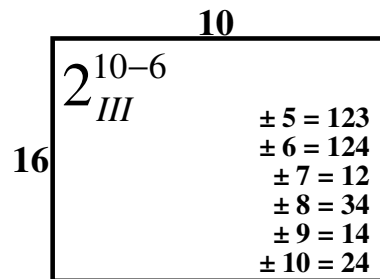


Fig. 22: Especificações do desenho fatorial fracionado empregado na renaturação de SmCD1.1 e SmCD1.2.

Tabela 4: Níveis dos fatores pesquisados no desenho fatorial fracionado.

Fatores	Níveis	
	(-)	(+)
Tampão/pH	50 mM MOPS pH 5,8	50 mM Tris-HCl pH 8,2
Força iônica	10 mM NaCl + 0,40 mM	KCl 250 mM NaCl + 10 mM KCl
Poliol	—	4% glicerol
Cátions divalentes	1 mM EDTA	2 mM MgCl ₂ + 2 mM CaCl ₂
Aditivo polar	—	500 mM L-arginina
Aditivo não-polar	—	400 mM sacarose
Ag. Caotrópico	—	1 M guanidina-HCl
Detergente	—	0,5 mM octilglicosídeo
Ag. redox ativo	1 mM DTT	1 mM GSH : 0,1 mM SSG
Conc. de proteína	0,1 mg/mL	1 mg/mL

3.7.8.c Metodologia de renaturação

Para este ensaio, 50 μ L das amostras de alta (1 mg/mL) ou de baixa concentração de proteína (0,1 mg/mL) em 4M de guanidina-HCl, eram instantaneamente diluídos em 950 μ L de um dos 16 diferentes tampões contendo os aditivos descritos nas **Tabelas 3 e 4**. Em seguida, os aditivos eram retirados por microdiálise por 10 - 12h contra tampão 0,1 M NaOAc pH 3,5 + 0,1 M NaCl, similarmente ao descrito acima. O sobrenadante das amostras centrifugadas a 20.000g / 30min / 4 °C eram transferidos para novos tubos e reservados para avaliação das respostas.

3.7.8.d Respostas – hidrólise de substrato fluorogênico e concentração de proteína

As amostras renaturadas como descrito acima tinham a concentração de proteína estimada pela D.O.²⁸⁰ e eram então testadas quanto à atividade hidrolítica sobre o substrato peptídico fluorogênico Abz-AIKFFSRQ-EDDnp ([PIMENTA et al., 2001](#)) pela emissão de fluorescência a 420 nm (320 nm excitação). Sucintamente, 20 μ L da amostra de proteína eram misturados a 176 μ L de tampão de atividade (0,1 M NaOAc pH 3,5 + 0,1 M NaCl) e após 30min de incubação à 37 °C, a reação era disparada por adição de 4 μ L de solução estoque do substrato, obtendo-se uma concentração final durante o ensaio de 1,5 μ M. As amostras eram incubadas a 37 °C e ao término de 30min as reações eram paralisadas pela adição de 1,8 mL de NaOH 1,6 M. As medidas eram realizadas em um espectrofluorímetro digital Shimadzu RF, utilizando cubeta de 1 cm de caminho ótico (2 mL volume mínimo). Ensaio em branco eram preparados

substituindo-se os 20 μL de amostra no ensaio por tampão ou, alternativamente, parando-se a reação de cada amostra no tempo zero. Também foram realizados controles positivos com catepsina D bovina e controles negativos com amostras incubadas com o inibidor pepstatina A (1 μM).

3.7.8.e Planejamento e análise estatística

O planejamento do desenho factorial fracionado e a análise estatística dos resultados foram realizados no programa JMP In release 5.1 (SAS Institute, Cary, NC). O cálculo do efeito principal (EP) de cada fator foi realizado no programa MS-Office Excel 2003 (Microsoft) de acordo com a fórmula: $EP = \frac{\sum RF^+ - \sum RF^-}{16}$, onde RF^+ é a resposta de cada condição onde o fator em questão está presente no nível “+” e RF^- onde está presente no nível “-”.

3.7.9 Teste de auto-ativação

Amostras provenientes de ensaios de renaturação eram pré-incubadas em tampão 0,1 M NaOAc pH 3,5 + 0,1 M NaCl por 30 e 60min antes do ensaio com substrato fluorogênico como descrito no item anterior. Alíquotas destas amostras também eram analisadas por SDS-PAGE 12%.

3.7.10 Tentativa de renaturação por diálise

As frações eluídas da coluna de afinidade foram reunidas totalizando 10 mL (1,5 mg/mL). As amostras de proSmCD1.1 e proSmCD1.2 foram inicialmente dializadas (membranas de 10 - 12 kDa de limite de exclusão do poro) durante a noite por 16-18h em câmara-fria (4 – 8 °C) contra 1,5 L de tampões 50 mM NaOAc pH 5,8 + 2 mM MgCl_2 + 2 mM CaCl_2 + 1 mM GSH / 0,1 mM GSSG e 50 mM Tris-HCl pH 8,2 + 10 mM NaCl + 10 mM KCl + 1 mM GSH / 0,1 mM GSSG + 200 mM L-arginina, respectivamente. Em seguida, os tampões foram trocados e retomou-se a diálise por mais 4h. Então, os tampões de renaturação com aditivos foram trocados por 1 L de tampão 50 mM fosfato de sódio pH 7,0 e as amostras dializadas por 4h. A seguir, o tampão foi renovado e dializou-se durante a noite, por 16-18h. As amostras foram retiradas dos sacos de diálise, centrifugadas a 20.000g / 30min / 4 °C. Os pellets e os sobrenadantes foram analisados por SDS-PAGE a 12%.

3.8 Produção do domínio COOH-terminal de SmCD1 (PepC-term) em E. coli

3.8.1 Clonagem em vetor pBUf5

3.8.1.a Primers e vetores

Primers foram desenhados para clonar a região COOH-terminal (últimos 43 resíduos) codificada pelo gene da forma longa de SmCD1 a partir do vetor pGEM-t carregando o referido gene: SmCD/C-term-BamHI 5' GGA TCC ATG CAT CCA GAT TCT GTT CAT CGT ACC 3' e SmCD/C-term-HindIII 5' AAGCT TCA AAC TTC ATC TGA AAG AAG TTT AGA AAA GGC 3' (*primer* reverso). Os códons de iniciação, ATG, e terminação, TCA, da transcrição estão destacados em negrito. Sítios de restrição das enzimas Bam HI e Hind III (sublinhados nas seqüências acima) foram inseridos nas regiões 5' e 3' do fragmento amplificado através de PCR, respectivamente, permitindo a clonagem direcional no vetor pBUf5. Este vetor é um derivado do plasmídeo comercial pET21a (Invitrogen) obtido pela inserção, entre os sítios de Nco I e Bgl II, do gene codificando um fragmento de 14 kDa da proteína D do capsídio do fago lambda. O gene foi inserido na mesma pauta de leitura que a seqüência de início da tradução, logo abaixo da região codificando a cauda de hexahistidina. Além disso, entre o sítio de Bgl II e Bam HI, este último na região de clonagem múltipla, foi inserido uma seqüência codificando um pequeno peptídeo reconhecido pela protease de vírus do tabaco, conhecida como protease TEV. O vetor pBUf5 foi gentilmente cedido pelo Dr. Alberto Spisni (CeBiME/LNLS, Campinas-SP).

3.8.1.b Clonagem

As etapas de clonagem foram semelhantes às já detalhadas para os genes das formas longa e curta de proSmCD1, conforme o item 3.6.5. Resumidamente, o DNA codificando o peptídeo C-terminal da forma longa (PepC-term) foi amplificado a partir do plasmídeo pGEM-t/CatD_Acc, o produto de PCR (~ 120 pb) isolado a partir de gel de agarose e clonado em pGEM-t. Os clones positivos foram selecionados quanto à orientação correta do inserto por análise de restrição. A partir do clone selecionado procedeu-se a subclonagem em vetor pBUf5, como descrito brevemente a seguir: os vetores pBUf5 e pGEM-t/PepC-term foram digeridos com Bam HI e Hind III. O inserto liberado de pGEM-t/PepC-term (~ 120 pb) foi purificado a partir da banda no gel de agarose. Os dois fragmentos lineares foram então ligados e o produto usado na

transformação de *E. coli* DH5 α . Clones positivos foram identificados pela amplificação do inserto PtnD/PepC-term de ~ 650 pb, usando *primer* T7 5' (anela com segmento do vetor) e *primer* reverso específico. Após verificação por sequenciamento da correção da construção, o plasmídeo foi purificado e usado para transformar células competentes da cepa BL21(DE3) de *E. coli*.

3.8.2 Estudos de expressão da fusão PtnD/PepC-term

Os testes de expressão da fusão PtnD/PepC-term por *E. coli* BL21(DE3) transformada com o vetor pBUf5/PepC-term foram realizados semelhantemente aos testes descritos no item 3.7.3. Em particular, foram avaliadas 3 temperaturas de indução da expressão: 25 °C, 30 °C e 37 °C, e em cada temperatura foram avaliadas 4 concentrações do agente indutor IPTG: 0,10; 0,25; 0,50 e 1,0 mM.

3.8.3 Padronização da purificação de PtnD/PepC-term por IMAC

3.8.3.a Expressão e preparação do extrato solúvel

Inoculou-se 500 mL de meio LB-AMP (100 μ g/mL de ampicilina) com 1% v/v de cultura de cerca de 16h a 37 °C. Realizou-se o cultivo sob agitação constante (200 rpm) a 30 °C até que a D.O.⁶⁰⁰ atingisse 0,7 - 0,9 U.A.. Neste momento, induziu-se a expressão de PtnD/PepC-term com adição de IPTG 0,5 mM e retomou-se o cultivo por mais 3h nas mesmas condições. Ao término deste período, as células foram coletadas por centrifugação a 5.000g / 10min / 4 °C e lavadas com 200 mL de LB, antes de serem armazenadas a -80 °C. As células foram suspensas em 20 mL de tampão 50 mM Tris-HCl pH 7,5, contendo 5% glicerol, 300 mM NaCl e 5 mM imidazol (tampão 1 + 5 mM imidazol) e transferidas para 2 tubos de teflon. Adicionou-se PMSF para obter uma concentração final de 1 mM e em seguida, adicionou-se 80 μ L de lisozima (10 mg/mL) em cada tubo. A lise procedeu por incubação em banho de gelo por 30min. O lisado foi homogeneizado por sonicação em gelo (5 ciclos de 30s) e centrifugado por 1h a 30.000g (4 °C). O sobrenadante foi filtrado (0,22 μ m) e reservado para a cromatografia de afinidade.

3.8.3.b Cromatografia de Afinidade

Cerca de 0,5 mL da resina Ni-NTA superflow (Qiagen) foi transferido para coluna plástica de 10 mL (Bio-Rad) e procedeu-se a lavagem com 6 mL de água deionizada (12 volumes de coluna). Em seguida, a resina foi equilibrada com 2,5 mL de

tampão 1. Aplicou-se 9,5 mL do extrato solúvel obtido no item anterior, vedou-se a coluna e incubou-se sob agitação suave a 4 °C por 12-14h. O material não-ligado foi recolhido e lavou-se a coluna com 3 x 2,5 mL de tampão 1 + 5 mM de imidazol (frações 1-3). A partir daí, passagens sucessivas de 2 x 2,5 mL de tampão 1 contendo 20, 50, 100 e 200 mM de imidazol foram realizadas (frações 4-12). As frações foram recolhidas, o conteúdo protéico estimado pela D.O.²⁸⁰.

3.8.4 Testes de purificação de PtnD/PepC-term

3.8.4.a Resina catiônica

As frações 4-10 eluídas da coluna Ni-NTA superflow (~ 15 mL) foram reunidas e transferidas para membrana de diálise. Trocou-se o tampão para 50 mL Tris-HCl pH 7,5 (tampão A) em dois passos de diálise de pelo menos 6h contra 500 mL de tampão a 4 °C. A amostra foi congelada a -20 °C até a cromatografia. O material particulado resultante do congelamento foi retirado por filtração a 0,22 µm e a amostra aplicada a um fluxo de 0,5 mL/min em coluna Econopac High-Q (Bio-Rad) de 1 mL, previamente equilibrada em tampão A, acoplada a um sistema de cromatografia líquida de alta eficiência (HPLC). Após coleta do material não-ligado, realizou-se a lavagem com tampão A até que a D.O.²⁸⁰ retornasse a zero, recolhendo-se 3 frações de 10 mL. Em seguida, fez-se um gradiente de 0-50% de tampão B (tampão A + 1 M de NaCl) em 25min e seis frações de 4 mL foram recolhidas. Continuou-se a corrida com 50% de tampão B por mais 60min, recolhendo-se 15 frações de 2 mL. Por fim, fez-se um gradiente de 50-100% B em 25min.

3.8.4.b Resina aniônica

O material não-ligado e a 1ª fração recolhida da lavagem da coluna High-Q foram reunidos e dializados contra 1,5 L de 50 mM tampão citrato de sódio pH 4,9 (tampão A) a 4 °C por 18h. O material foi filtrado (0,22 µm) e aplicado em uma coluna Resource S (Pharmacia) de 1 mL, previamente equilibrada em tampão A, a um fluxo de 0,5 mL/min. A cromatografia foi realizada em um sistema HPLC e o efluente da coluna monitorado a 280 nm. Após coleta do não-ligado, o fluxo foi elevado para 1 mL/min e lavou-se com 30 mL de tampão A. Aplicou-se então um gradiente de 0-50% B (tampão A + 1M NaCl) por 20min, seguido de 5min de 50% B. A partir daí, iniciou-se novo gradiente de 50-100% B em 15min, seguido de 5min a 100% B.

Em um segundo teste, um extrato solúvel de células expressando PtnD/PepC-term foi preparado como descrito no item 3.8.3.a e aplicado em coluna Ni-NTA superflow. O material ligado especificamente à resina foi eluído em tampão de equilíbrio da coluna Resource S, acrescido de 50 mM imidazol e 300 mM NaCl. Frações de 1 mL foram coletadas e a D.O.²⁸⁰ lida. As frações 1-8 foram reunidas e aplicadas sob fluxo de 0,2 mL/min na coluna Resource S, previamente equilibrada no mesmo tampão. Neste ponto, o fluxo foi elevado para 1 mL/min e lavou-se a coluna com tampão de equilíbrio até a D.O.²⁸⁰ voltar à linha de base. Quatro frações de 4 mL foram coletadas. Iniciou-se um gradiente de 0-25% B (tampão de equilíbrio + 1 M de NaCl) em 5min, seguido de 5min a 25% B. A seguir, aplicou-se gradiente de 25-100% B em 25min, seguido de 5min a 100% B. Frações de 1 mL foram coletadas e posteriormente analisadas por SDS-PAGE 13,5%.

3.8.5 Purificação inicial de PepC-term liberado da fusão

3.8.5.a Teste de clivagem com protease TEV

Para este teste e todas as digestões de PtnD/PepC-term para liberação do peptídeo PepC-term foi utilizada protease TEV recombinante produzida pelo grupo do Dr. Nilson Zanchin (CeBiME, LNLS, Campinas-SP). As frações 1-4 do lavado do 2º teste de purificação em coluna Resource S foram reunidas (~ 15 mL) e dializadas contra 1,5 L de tampão 50 mM Tris-HCl pH 8,0 + 0,5 mM EDTA (tampão da TEV), por 16h a 4 °C. Após a diálise, a concentração de proteína do material foi determinada. Em cada ensaio, 15 µL da amostra de PtnD/PepC-term (5 µg de proteína) foi misturada a diferentes quantidades da preparação de protease TEV (0 - 3 µL) e o volume final ajustado para 20 mL em tampão da TEV suplementado com 1 mM DTT. Incubou-se por 2h a 30 °C. As reações foram paralisadas por fervura em tampão de amostra e analisadas por SDS-PAGE 15%.

3.8.5.b Purificação de PepC-term por IMAC

Seguindo-se a reação de digestão com protease TEV, a amostra contendo PepC-term livre era aplicada em coluna Ni-NTA superflow previamente equilibrada com tampão da TEV sem DTT (50 mM Tris-HCl pH 8,0 + 0,5 mM EDTA). O material não-ligado contendo PepC-term era recolhido e reservado para purificação posterior. O material ligado à resina, incluindo os contaminantes da preparação de PtnD/PepC-term e

a protease TEV era eluído pela adição de tampão da TEV suplementado com 200 mM de imidazol.

3.8.6 Testes de purificação de PepC-term

3.8.6.a Cromatografia de fase reversa

À amostra do não-ligado da coluna Ni-NTA (ver item 2.8.6) adicionou-se TFA suficiente para uma concentração final de 0,1%. A amostra foi filtrada a 0,22 μm e aplicada sob fluxo de 0,2 mL/min em coluna Shimpack VP-ODS (Shimadzu), previamente equilibrada em 0,1% de TFA. A corrida foi realizada em um sistema HPLC, monitorando-se o efluente da coluna a 220 nm. Após recolher-se o não-ligado, elevou-se o fluxo para 1 mL/min e iniciou-se um gradiente de 0-100% de B (80% v/v ACN + 0,1% TFA) em 30min, seguido de 10min de 100% B. Os picos foram coletados e reservados para análise por espectrometria de massas.

3.8.6.b Gel filtração

Uma amostra de não-ligado à coluna Ni-NTA, contendo PepC-term, foi concentrada de 5 mL para 500 μL usando filtro Centricon-3 (7.000g / 4 °C / ~ 4h). Foram aplicados 250 μL em coluna Superdex peptide 10/30 HR (Pharmacia) equilibrada em 40 mM de tampão fosfato de sódio pH 7,0 + 0,25M NaCl. A cromatografia foi desenvolvida em sistema HPLC, neste mesmo tampão, a um fluxo de 0,5 mL/min e a absorvância monitorada a 220 e a 280 nm, em corridas independentes. Posteriormente, o protocolo foi otimizado ao se substituir o tampão de TEV usual pelo tampão volátil 150 mM de bicarbonato de amônio pH 8,0. Assim, as amostras do não-ligado da Ni-NTA eram liofilizadas e ressolubilizadas no tampão volátil para aplicação na coluna de gel filtração previamente equilibrada neste mesmo tampão.

3.8.7 Caracterização Estrutural

3.8.7.a Métodos computacionais

A predição de estrutura secundária de pepC-term foi gerada pelo método de predição “consenso” (SILVA-JR *et al.*, 2002). De acordo com este método, uma bateria de 6 programas disponíveis na forma de servidores WWW (HNN, SOPMA, PHD, PSI-PRED, J-PRED e SSPRO) foi empregada na análise da seqüência e a estrutura secundária gerada manualmente pela inspeção da concordância entre as predições

individuais. HNN (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html) é baseado numa rede neural hierárquica e está disponível no pacote NPS@ (COMBET *et al.*, 2000). SOPMA (<http://www.ibcp.fr/predict.html>) foi desenvolvido por Geourjon & Deleage (1995) e consiste na otimização recursiva de parâmetros do algoritmo baseado na comparação de seqüências de proteínas com estrutura conhecida. O autores estimam que o programa é capaz de alcançar até 82% de acerto quando combinado com outro método de predição baseado em redes neurais, PHD (ROST & SANDER, 1993). PHD é encontrado no servidor PredictProtein (<http://www.embl-heidelberg.de/predictprotein/predictprotein.html>). O servidor PSI-PRED encontrado em <http://bioinf.cs.ucl.ac.uk/psipred/psiform.html> (MCGUFFIN *et al.*, 2000) é baseado no método de matrizes de escore posição-específica desenvolvido por Jones (1999). J-PRED é uma rede neural desenvolvida pelo grupo de Barton em Oxford (CUFF *et al.*, 1998) com índices de acerto acima de 70% e está disponível no endereço <http://www.compbio.dundee.ac.uk/~www-jpred/submit.html>. SS-PRO é mais um método baseado em redes neurais (POLLASTRI *et al.*, 2002) e se destaca por utilizar um conjunto de 11 redes bidirecionais recorrentes.

Duas estratégias foram empregadas na predição da conformação espacial de pepC-term: modelagem *ab initio* e reconhecimento do enovelamento. Dentro da primeira estratégia, empregou-se o servidor SCRATCH (<http://www.igb.uci.edu/tools/scratch/>) para obter um modelo de baixa resolução, construído a partir de previsões de mapas de contato e predições de estrutura secundária (POLLASTRI & BALDI, 2002). Para a metodologia de “reconhecimento de enovelamento”, dois servidores WWW foram utilizados: 3D-PSSM (<http://www.sbg.bio.ic.ac.uk/~3dpssm/index2.html>) e GenTHREADER, o qual é oferecido junto ao servidor de PSI-PRED. Em geral, nesta metodologia conhecida como “threading”, a seqüência de uma proteína é “enovelada” sobre um banco de dados de estruturas e avaliada quanto à compatibilidade desta sobreposição. No programa 3D-PSSM (FISHER *et al.*, 1999), a homologia com proteínas distantes é reconhecida usando a informação tridimensional para gerar uma matriz de escore posição-específica. GenTHREADER foi originalmente desenvolvido por Jones (1999) para a predição em larga-escala da estrutura de seqüências genômicas. Mais recentemente, a rede neural simples empregada anteriormente na combinação de escores de alinhamentos com técnicas usuais de *threading* foi sofisticada para incorporar matrizes de escore posição-

específica (PSI-BLAST) e predições de PSI-PRED para alcançar predições mais confiáveis (McGUFFIN & JONES, 2003).

3.8.7.b CD de 18N_PepC-term

O pico da gel filtração caracterizado como um fragmento de 2,1 kDa correspondendo aos primeiros 18 resíduos de PepC-term foi liofilizado e ressuspendido em 500 μ L de água deionizada (~ 6 mg/ mL). A amostra foi dividida em duas partes iguais: uma parte teve o volume dobrado com tampão 20 mM NaOAc, pH 3,5 e a outra com tampão 20 mM fosfato de sódio pH 7,0. O efeito de 200 mM NaCl e 0,5% SDS foi avaliado em cada um dos tampões através da adição de volume desprezível (< 5% do volume total) de soluções concentradas de NaCl (5 M) e SDS (20%).

3.9 Produção das formas longa e curta de proSmCD1 em *Pichia pastoris*

3.9.1 Subclonagem em vetor de expressão pPIC9k

3.9.1.a Vetor e estratégia

Através do *primer 5'* usado na transcrição reversa dos mRNA das formas longa e curta de SmCD1 foram introduzidos dois sítios de restrição, Eco RI e Nde I, para facilitar as clonagens em vetores de expressão em *E. coli* (pET28a) e *P. pastoris* (pPIC9k), respectivamente. O plasmídeo pPIC9k (Invitrogen) é um vetor de 9276 pb para expressão de proteínas em *P. pastoris* na forma secretada através da fusão com o gene de 269 pb codificando o sinal de secreção do fator- α (Manual do Multi-copy Pichia Expression Kit, Invitrogen). Este vetor contém o gene bacteriano de resistência à kanamicina (*kan*) para seleção *in vivo* de insertos múltipla cópia. Diferente de bactérias, em *Pichia sp.* o gene *kan* confere resistência à geneticina, cujo nível é aproximadamente proporcional ao número de cópias deste gene integrado ao genoma. Esta última característica, aliada à transcrição do gene clonado sob controle do forte promotor AOX1 favorece a obtenção de altos níveis de expressão do gene de interesse. Além disso, a presença do gene HIS4 codificando a enzima histidinol desidrogenase permite a complementação de cepas deficientes em HIS4, como GS115 e KM71, permitindo a seleção de clones positivos em meios deficientes em histidina.

Assim, a estratégia de subclonagem em pPIC9k consistia em isolar o fragmento liberado da digestão dos clones pGEM-t/proSmCD1 com Eco RI e Not I, e ligar em pPIC9k, previamente digerido com as mesmas enzimas, para clonagem na mesma pauta de leitura do códon de iniciação da seqüência sinal. Contudo, uma dificuldade adicional foi criada para a subclonagem do gene da forma longa entre os sítios de Eco RI e Not I de pPIC9k, pois o gene corrigido da forma longa só estava disponível em pET28a, conforme descrito no item 3.7.2. Isto porque o sítio de Eco RI, anterior ao sítio de Nde I no gene clonado em pGEM-t, foi perdido na subclonagem em pET28a entre os sítios de Nde I e Not I (este último sítio provém da região múltipla de clonagem do pGEM-t). Desta forma, a subclonagem do gene da forma longa teve que passar por uma ligação de três fragmentos lineares: fragmento de 0,8 kb Eco RI/Acc I de pGEM-t/proSmCD1.2, fragmento de 0,4 kb Acc I/Not I de pET28a/proSmCD1.1-corrigida e pPIC9k Eco RI/Not I.

Os fragmentos lineares necessários para clonagem dos genes das formas longa e curta de SmCD1 em pPIC9k foram preparados por reações de digestão a 37 °C, por 12-16h, com as enzimas apropriadas. Os vetores foram preparados a partir de culturas em pequena escala de clones de *E. coli* DH5 α , usando-se kit comercial para purificação em coluna (SV Wizard Plus, Promega) ou o método da lise alcalina (SAMBROOK *et al.*, 1989). Para liberação do fragmento Eco RI/Acc I de pGEM-t/proSmCD1.2 (F1), 10 μ L do vetor (~500 ng) foram digeridos com 12 unidades de Eco RI (Invitrogen) e 10 unidades de Acc I (Invitrogen) em 30 μ L de tampão React 1 (Invitrogen). Semelhantemente, para liberação do fragmento Eco RI/Not I de pGEM-t/proSmCD1.2 (F3), usou-se 12 unidades de Eco RI e 30 unidades de Not I (Invitrogen) em 30 μ L do tampão React 3 (Invitrogen). O fragmento Acc I/Not I (F2) foi liberado do vetor pET28a/pro-CatD-longa-corrigido digerindo-se 25 μ L do vetor (~500 ng) com 30 unidades de Not I e 15 unidades de Acc I em 50 μ L de tampão React 3. Por fim, o vetor pPIC9k (V1) foi preparado incubando-se 25 μ L de sua preparação (~1000 ng) com 18 unidades de Eco RI e 22,5 unidades de Not I em 50 μ L de tampão React 3.

As amostras provenientes da digestão do vetor pGEM-t/proSmCD1.2 foram incubadas com 3 μ g de RNase A por 30min a 37 °C. Subsequentemente, as amostras de digestão foram aplicadas em geles de agarose (0,8% para os vetores derivados de pGEM-t e pET28a e 0,5% para pPIC9k) e os fragmentos de DNA linear apropriados recortados do gel, pesados e purificados com o kit Kiaquick (QIAGEN), seguindo-se as

instruções do fabricante. Ao final do procedimento, as amostras foram eluídas com 50 μ L de água deionizada, exceto F2 que foi eluído com 35 μ L. As reações de ligação foram realizadas a 16 °C por cerca de 18h. Para a subclonagem dos genes das formas longa e curta de SmCD1 em pPIC9k, 5 μ L de V1 e 1 unidade de T4 DNA ligase (Promega) foram adicionados ou a 5 μ L de F1 + 5 μ L de F2 ou a 5 μ L de F3 apenas, respectivamente, em um volume final de 20 μ L de tampão fornecido com a enzima (tampão ligase, Promega). Um controle, suprimindo-se F1+F2 ou F3, foi feito em paralelo.

Para a transformação por choque térmico, 60 μ L das células competentes de *E. coli* DH5 α (SAMBROOK *et al.*, 1989) foram adicionados aos tubos das reações de ligação, dentro de câmara de fluxo laminar, em seguida incubando-se por 30min em gelo. Ao término deste período, os tubos foram transferidos para banho-maria a 42 °C e incubados por 90s. Os tubos foram então retornados ao gelo e mantidos por mais 5min até adição de 1 mL de meio LB (dentro da câmara de fluxo laminar). Subsequentemente, os tubos foram incubados sob agitação por 1h a 37 °C. As células foram então coletadas por centrifugação a 10.000g / 10min / 25 °C e ressuspensas, com auxílio de pipeta em 100 μ L do sobrenadante. As células foram espalhadas com auxílio de alça de Drigalski em placas LB-AMP (50 μ g/mL de ampicilina sódica) e incubadas por 12-16h a 37 °C. As colônias obtidas foram testadas quanto a presença dos insertos de interesse por PCR de colônia (SAMBROOK *et al.*, 1989) usando *primers* AOX1 5' e 3' (Manual do Multi-copy Pichia Expression Kit, Invitrogen). Confirmação posterior foi conseguida por análise de restrição e sequenciamento.

3.9.2 Integração no genoma de *P. pastoris*

3.9.2.a Preparação do DNA linear para transformação

Células provenientes de 4 x 10 mL de culturas 12-16h a 37 °C de clones *E. coli* DH5 α , confirmados quanto à presença de pPIC9k/proSmCD1 (longa e curta) e pPIC9k, foram obtidas por centrifugação (5.000g / 10min / 25 °C) e processadas para obtenção do DNA plasmidial com o kit Wizard Plus (Promega), seguindo-se as instruções do fabricante para volumes de 5-10 mL de cultura. Para cada clone, os eluatos de 4 colunas (em água deionizada) foram reunidos, totalizando 200 μ L. Os vetores (~ 40 μ g) foram linearizados através de digestão por cerca de 16h a 37 °C com 50 unidades de Sal I em tampão D (Promega) e na presença de 20 μ g de acetil-BSA. Após confirmação que

virtualmente todo DNA havia sido digerido por Sal I, procedeu-se a purificação do DNA plasmidial com o kit Wizard SV gel/PCR (Promega), tratando-se os 200 μ L da reação de digestão como um “produto de PCR”. O DNA linear purificado foi eluído em 20 μ L de água deionizada estéril.

3.9.2.b Preparação de células competentes de *P. pastoris* GS115

Inoculou-se 5 mL de meio YPD (extrato de levedura, peptona, dextrose; [Manual do Multi-copy Pichia Expression Kit, Invitrogen](#)) em um tubo cônico de centrífuga de 50 mL com uma colônia de placa recente de *P. pastoris* GS115 e incubou-se a 30 °C sob agitação constante por aproximadamente 18h, quando a D.O.₆₀₀ chegou a 2,1 U.A.. Neste momento, alíquotas de 1,13 e 0,565 mL foram retiradas e usadas para inocular 250 mL de meio YPD em frascos Erlenmeyer de 1L. As culturas foram incubadas sob agitação constante (180 rpm) por cerca de 18h a 30 °C. A cultura proveniente do volume de pré-inóculo = 0,565 mL (D.O.⁶⁰⁰ = 9,85) foi diluída com YPD para uma D.O.⁶⁰⁰ = 0,7. Cerca de 1,5h depois de nova incubação nas condições já descritas, conseguiu-se a cultura na D.O.⁶⁰⁰ requerida de 1,4-1,5 U.A.. As células foram coletadas por centrifugação a 5000g por 5min a 4° C, descartou-se o sobrenadante e lavou-se as células com 250 mL de água deionizada estéril e gelada. As células foram novamente coletadas e repetiu-se a lavagem com um volume de 125 mL de água. Uma lavagem final foi realizada em 10 mL de sorbitol 1M estéril e gelado. Finalmente, as células competentes foram coletadas e ressuspendidas em 350 μ L de sorbitol 1M estéril e gelado.

3.9.2.c Eletroporação

Adicionou-se 80 μ L de suspensão de células para 4 cubetas de eletroporação de 0,2 cm de espessura, estéreis e geladas. Em seguida, 10 μ L de água deionizada ou 10 μ L das preparações de pPIC9k, pPIC9k/proSmCD1.1 e pPIC9k/proSmCD1.2 (~ 10 μ g de DNA) foram adicionados às cubetas, em condições estéreis. Adicionou-se sorbitol 1M estéril e gelado q.s.p. 400 μ L e incubou-se as cubetas em gelo por 5min. As suspensões foram homogeneizadas com pipeta estéril e um pulso de 2,43 kV (400 Ω / 25 μ F) aplicado em intervalos que variaram de 7,4 a 8,8 ms. A seguir, avolumou-se as suspensões a 1 mL com sorbitol 1M gelado e alíquotas de 12,5 – 200 μ L foram retiradas para inocular (por espalhamento) placas de agar MD (meio mínimo de dextrose; [Manual](#)

do Multi-copy Pichia Expression Kit, Invitrogen). As placas foram mantidas a 30 °C e após 48-72h de incubação colônias His4⁺ foram isoladas.

3.9.2.d PCR de colônia

A presença do cassete de expressão P-AOX1/sinal fator- α /proSmCD1 (1,7 kb na forma curta e 1,9 kb na forma longa) nas colônias isoladas das placas inoculadas com 12,5 μ L da transformação foi confirmada por PCR, usando *primers* específicos para o gene de proSmCD1 ou *primers* AOX1 5' e 3', conforme descrito a seguir: com auxílio de uma palito de madeira estéril, as colônias foram pescadas, riscadas em uma placa MD e o restante dissolvido em 100 μ L de solução de lise (50 mM Tris-HCl pH 7.5 + 1% SDS) em um tubo de microcentrífuga. Os tubos foram então agitados vigorosamente e aquecidos a 95° C por 5min. Subsequentemente, os tubos foram centrifugados (20.000g / 5min / 25 °C) e o sobrenadante usado como substrato nas ampliações por PCR. Para 20 reações de 25 μ L foi feita uma pré-mistura com 50 μ L de tampão 10x livre de íons Mg²⁺ (Promega), 50 μ L de MgCl₂ 25 mM, 25 μ L de mistura de dNTPs 10 mM, 25 μ L de cada *primer* AOX1 (5' e 3') a 5 μ M, 10 μ L de Taq DNA polimerase (1 u/ μ L; Promega) e água deionizada q.s.p. 500 μ L. Para a reação final, 23 μ L da pré-mistura eram misturados a 2 μ L do substrato. As condições do termociclador foram: 1 ciclo 94° C, 5min / 30 ciclos 94 °C, 1min; 54 °C, 1min; 72 °C, 2min / 1 ciclo 72 °C, 7min / 1 ciclo 4 °C, ∞ .

3.9.3 Seleção de clones múltipla-cópia com geneticina

Colônias das placas inoculadas com 50, 100 e 200 μ L da suspensão de células transformadas como descrito no item anterior, foram coletadas e ressuspensas em 1 mL de água deionizada estéril com auxílio de alça de Drigalski. A concentração de células foi determinada a partir da D.O.⁶⁰⁰ da suspensão usando a relação 1 U.A.⁶⁰⁰ = 5 x 10⁷ células/mL. A partir de uma alíquota diluída 500 vezes, 100 μ L (cerca de 10⁶ células) foram retirados e distribuídos em placas YPD contendo as seguintes concentrações de geneticina: 0,25; 0,50; 1,0; 2,0 e 4,0 mg/mL. As placas foram incubadas a 30 °C por 3 a 5 dias, quando as primeiras colônias puderam ser visualizadas. Colônias selecionadas das placas de YPD-geneticina (de 1,0 e 4,0 mg/mL para pPIC9k/proSmCD1.2 e de 2,0 e 4,0 mg/mL para pPIC9k/proSmCD1.1) foram repicadas para placas MD com auxílio de palito estéril e o restante homogeneizado em 50 μ L de solução de lise. Clones contendo

os cassetes de expressão corretos foram confirmados por PCR de colônia, como descrito no item 3.9.2.d.

3.9.4 Teste de expressão em meio mínimo

3.9.4.a Pequena escala com tempo de indução máximo de 72h.

Para o teste, 8 clones de *P. pastoris* GS115 carregando cada um dos vetores pPIC9k/proSmCD1.2 e pPIC9k/proSmCD1.1, e 1 clone de pPIC9k foram selecionados. Dos 8 clones, 1 era proveniente de placa sem geneticina, outro de placa com 1 mg/mL (forma curta) ou 2 mg/mL (forma longa) e os demais de placas com 4 mg/mL deste antibiótico. As colônias foram inoculadas em 5 mL de meio BMG (meio tamponado, mínimo com glicerol; [Manual do Multi-copy Pichia Expression Kit, Invitrogen](#)) contido em tubos cônicos de 50 mL. As culturas foram incubadas a 30 °C, sob agitação vigorosa e constante (200 rpm) por cerca de 20-24h até que a D.O.⁶⁰⁰ atingisse 4-5 U.A. em seguida, as células foram coletadas por centrifugação (5.000g / 10min / 4 °C) e ressuspensas em 1 mL de meio BM (BMG sem glicerol). A D.O.⁶⁰⁰ foi lida e alíquotas retiradas para obter 20 mL de culturas a 1,0 U.A. em meio BM. A expressão de proSmCD1 foi então induzida pela adição de 0,5% de metanol (grau sequenciamento de proteínas, Sigma). As culturas foram incubadas em frascos Erlenmeyer de 125 mL, vedados com 2 camadas de gaze estéril, por até 72h a 30°C, sob agitação vigorosa e constante (220 rpm). Alíquotas de 1 mL eram retiradas a cada 24h, centrifugadas a 10.000g / 5min / 4° C e o sobrenadante reservado para análise posterior. A concentração de metanol foi mantida constante durante a indução, adicionando-se volume suficiente a cada 24h.

3.9.4.b Média escala com tempo de indução máximo de 120h.

Para o teste, 2 clones de *P. pastoris* GS115 carregando cada um dos vetores pPIC9k/proSmCD1.2 e pPIC9k/proSmCD1.1, 1 clone de pPIC9k (controle negativo) e 1 clone de pPIC9k/rPsd1 (controle positivo) foram selecionados. Para os clones de proSmCD1, um era proveniente de placa sem geneticina e o outro de placa com 4 mg/mL deste antibiótico. As colônias foram inoculadas em 50 mL de meio BMG (meio tamponado, mínimo com glicerol; [Manual do Multi-copy Pichia Expression Kit, Invitrogen](#)) e as culturas em frascos erlenmeyer de 500 mL foram incubadas a 30 °C, sob agitação vigorosa e constante (200 rpm) por cerca de 20-24h até que a D.O.⁶⁰⁰ ~ 10 U.A.. Em seguida, alíquotas das células foram coletadas por centrifugação (5.000g /

10min / 4 °C) e ressuspendidas em 200 mL de meio BM (BMG sem glicerol), de forma a obter uma D.O.⁶⁰⁰ inicial de 2,0 U.A.. A expressão de proSmCD1 foi então induzida pela adição de 0,5% de metanol (grau sequenciamento de proteínas, Sigma). As culturas foram incubadas em frascos Erlenmeyer de 2 L, vedados com 2 camadas de gaze estéril, por até 120h a 30°C, sob agitação vigorosa e constante (220 rpm). Alíquotas de 10 mL eram retiradas a cada 24h, as células coletadas por centrifugação a 10.000g / 5min / 4° C e o sobrenadante liofilizado. A concentração de metanol foi mantida constante durante a indução, adicionando-se volume suficiente a cada 24h.

As análises efetuadas consistiam inicialmente em medir a D.O.⁶⁰⁰ e o pH das alíquotas das culturas. A partir das células eram preparados extratos conforme [Hoffman et al. \(2002b\)](#) e os sobrenadantes liofilizados eram ressolubilizados em 1 mL de tampão 50 mM Tris-HCl pH 7.5. As amostras dos extratos celulares e sobrenadantes de cultura foram analisados quanto ao conteúdo protéico, atividade peptidásica com substrato cromogênico para catepsina D, SDS-PAGE e *immunoblot*, conforme descrito acima.

4 RESULTADOS

Esta seção é consubstanciada pelos resultados referentes à modelagem por homologia da forma curta de SmCD1 (SmCD1.2) e o estudo de modelagem molecular da especificidade por substrato desta enzima descritos no ANEXO 1.

4.1 Análise da fração ligada à pepstatina-agarose de EAV

A partir de 160 mg de vermes adultos liofilizados obteve-se 1,7 mg de proteína no EAV em tampão 100 mM NaOAc pH 3,5 + 1 M NaCl. Este material foi aplicado em coluna de pepstatina-agarose e o material especificamente ligado (FLPA) foi eluído por elevação do pH a 8,5, de acordo com o perfil mostrado na **Figura 23**. Obteve-se 150 µg de proteína em FPLA com uma atividade específica de 499 µM/min/mg sobre o substrato cromogênico Phe-Ala-Ala-Phe(4-NO₂)-Phe-Val-Leu(4-piridilmetil) éster.

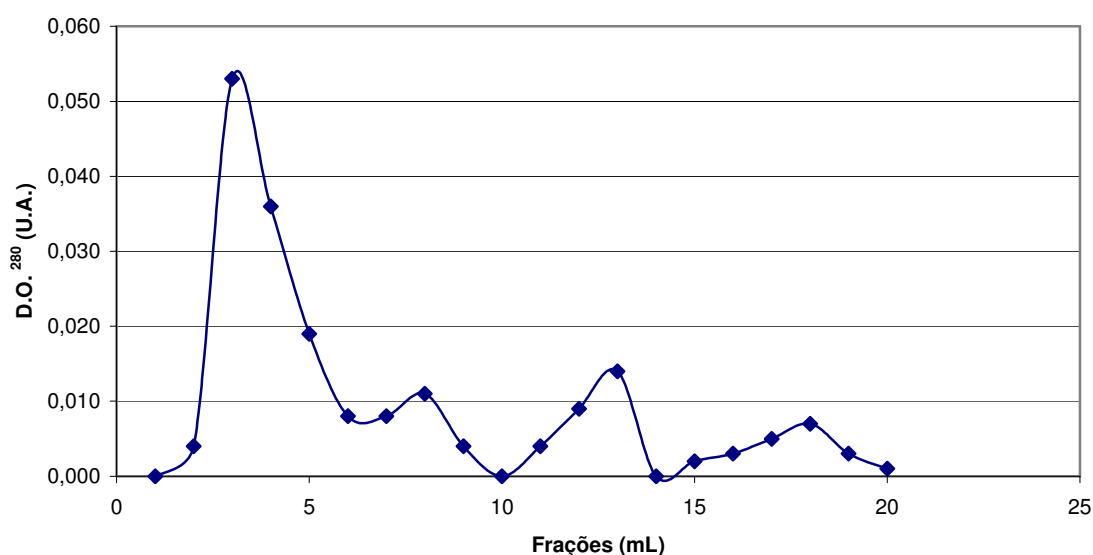


Fig. 23: Cromatografia de afinidade da fração EAV em coluna pepstatina-agarose. Foram aplicados 1,7 mg de proteína à coluna e a eluição foi realizada em tampão 100 mM Tris-HCl pH 8,2 + 1 M NaCl.

A análise por SDS-PAGE 12% da FPLA revelou um duplete de bandas majoritárias em ~36 kDa e bandas menos intensas na faixa entre 97-45 kDa e principalmente abaixo de 29 kDa (**Fig. 24**, 2). O perfil da fração não-ligada à pepstatina revelado por SDS-PAGE (**Fig. 24**, 3), comparado ao perfil de um extrato bruto dos vermes (macerados diretamente em tampão de amostra da SDS-PAGE), sugeriu que a extração em pH ácido forneceu um EAV já bem mais pobre em contaminantes. A comparação das canaletas 2 e 3 da **Figura 24** também revelou proteínas na fração não-

ligada à pepstatina-agarose possivelmente idênticas às bandas de ~36 kDa e ~29 kDa em FLPA. De fato, verificou-se que a fração não-ligada à pepstatina-agarose apresentava atividade específica ligeiramente superior à FLPA com 587 $\mu\text{M}/\text{min}/\text{mg}$. Este resultado indica possivelmente que as condições de ligação à coluna não foram ideais, possibilitando que a atividade de SmAPs pudesse ser encontrada na fração não-ligada. Alternativamente, peptidases ácidas não pertencentes à classe das aspartil proteases presentes na fração não-ligada poderiam ser responsáveis pela degradação do substrato de catepsina D.

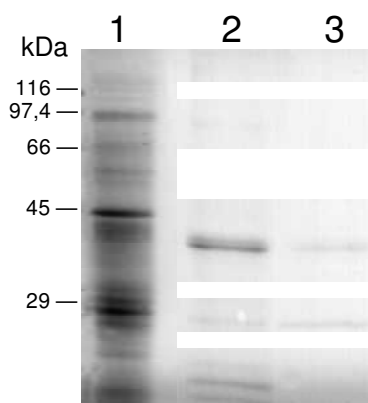


Fig. 24: Análise por SDS-PAGE 12% em condições redutoras das frações da cromatografia de EAV em pepstatina-agarose. 1- Homogenato de *S. mansoni* em tampão de amostra da SDS-PAGE (50 μg); 2- FLPA (5 μg); 3- fração não-ligada (2 μg). Revelação com coomassie blue R-250.

A complexidade da FLPA foi completamente revelada por 2DE (**Fig. 25**). Pode-se perceber pelo gel bidimensional que há uma maior concentração de proteínas na faixa ácida de pH, com espécies migrando desde 97 até 14,4 kDa. Famílias de *spots* de mesma massa molecular foram reveladas entre 45 e 20,1 kDa (destacadas por elipses de diferentes cores em **Fig. 25**), num perfil compatível com as bandas majoritárias reveladas anteriormente por SDS-PAGE 12% (**Fig. 24, 2**). Os *spots* de cada uma destas famílias foram recortados do gel, digeridos com tripsina e submetidos à análise por espectrometria de massas MALDI-TOF. Infelizmente, os espectros obtidos não apresentaram qualidade suficiente, principalmente devido à baixa intensidade dos picos, para permitir a identificação das proteínas por PMF.

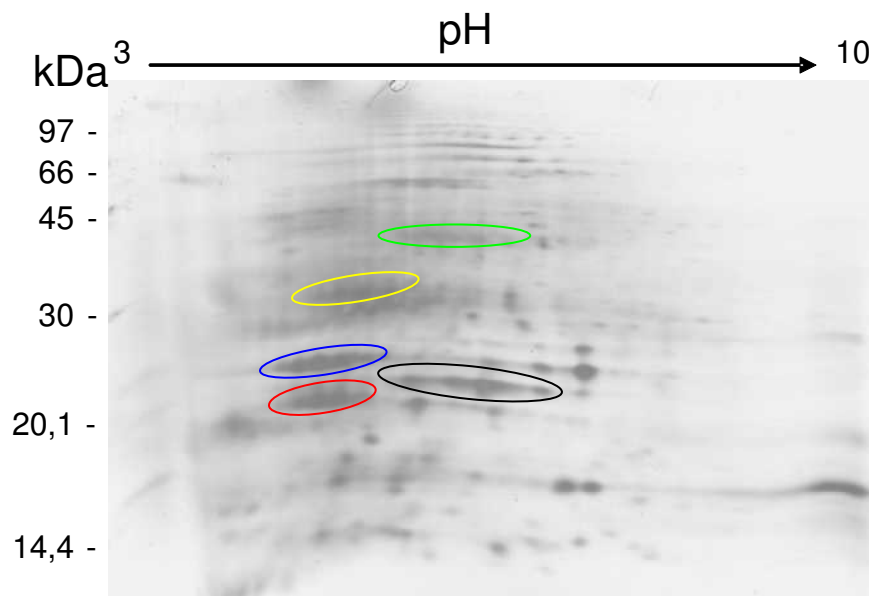


Fig. 25: 2DE 15% de FLPA. Foram aplicados 150 μ g de proteína à fita de 7 cm (IPG strip 3-10) para a primeira dimensão. A 2ª dimensão foi conduzida em um sistema Mini-Protean II (Bio-Rad). As famílias de *spots* selecionadas para análise posterior estão destacadas. O gel foi revelado pela prata.

4.2 Avaliação do efeito do tratamento com inibidores comerciais de HIV-1 protease sobre a infecção experimental de camundongos com *S. mansoni*

Grupos de 10 camundongos com 45 dias de infecção por *S. mansoni* foram submetidos a cinco dias de tratamento por via oral com 4 inibidores comerciais de HIV-1 protease: indinavir, ritonavir, nelfinavir e amprenavir. Os resultados da quimioterapia experimental estão resumidos na **Tabela 5**. Nenhum dos inibidores de HIV-1 protease testados foi capaz de reduzir significativamente o número de vermes recuperados por perfusão do mesentério ou necropsia do fígado em comparação ao grupo controle. Os dados da **Tabela 5** também mostram que não houve alteração na distribuição anatômica dos vermes mediante tratamento com estes fármacos.

Tabela 5: Avaliação do tratamento experimental de camundongos infectados com *S. mansoni*.

Fármaco	Dose (mg/kg)	Posologia (n ^o vezes ao dia)	Animais examinados (n)	Vermes recuperados	Distribuição de vermes (%)		Vermes mortos no fígado (%)
					Mesentério	Fígado	
Controle	-	-	30	21,6 ± 5,2	90,5	9,5	0,0
Indinavir	20	3	8	19,5 ± 4,0	91,7	8,3	0,0
	100	3	10	23,6 ± 9,0	94,9	5,1	0,0
Ritonavir	100	2	8	40,3 ± 15	94,1	5,9	0,0
Nelfinavir	50	2	9	22,0 ± 7,4	91,9	8,1	0,0
	250	2	9	18,6 ± 5,4	94,0	6,0	0,0
Amprenavir	450	2	4	38,8 ± 22	95,5	4,5	0,0
Praziquantel	25	1	10	11,5 ± 5,5	69,5	30,5	14,0
	125	1	8	9,1 ± 5,5	52,0	48,0	42,5
	250	1	9	6,2 ± 4,2	27,3	72,7	69,0

Por outro lado, praziquantel foi capaz de inverter o tropismo do parasito para o fígado na dose mais alta testada (250 mg/kg). Nesta dose de praziquantel, quase 70% dos vermes recuperados do fígado apresentaram-se mortos (**Tabela 5**), enquanto que o oograma mostrou-se 100% alterado (dados não-mostrados). Conforme esperado, praziquantel mostrou eficácia dose-dependente no tratamento experimental (**Fig. 26**). Estes resultados validam os experimentos realizados e efetivamente demonstram que os inibidores de HIV-1 protease testados foram ineficazes no tratamento experimental da esquistossomose mansônica.

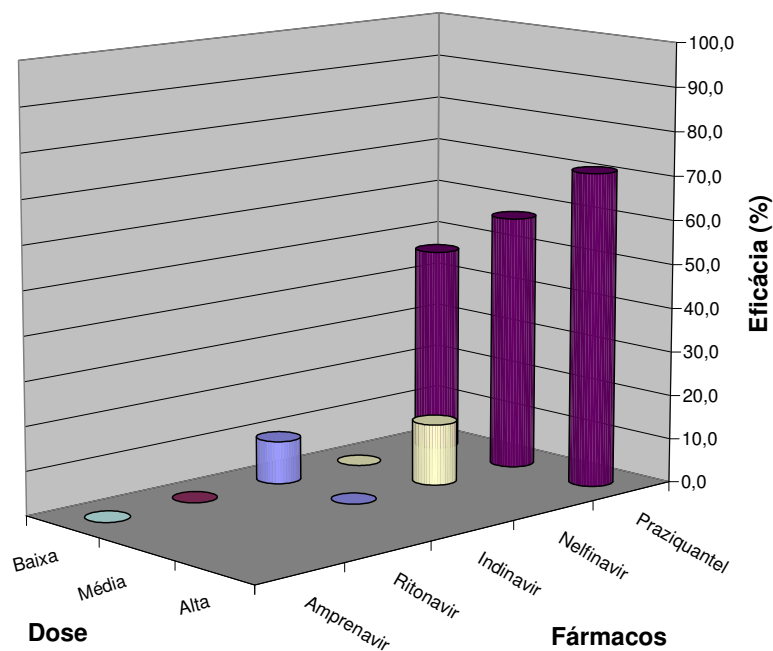


Fig. 26: Eficácia no tratamento experimental da esquistossomose mansônica com inibidores de HIV-1 protease. Eficácia é expressa como $\{[(C-E)/C]*100\}$, onde C e E correspondem ao número de vermes recuperados dos grupos controle e experimentais, respectivamente. Os fármacos foram testados em até 3 doses distintas, conforme especificado na **Tabela 5**.

4.3 Caracterização *in silico* da família gênica de SmAPs

Buscas foram realizadas em dois bancos de dados de ESTs de *S. mansoni* com o objetivo de avaliar a consistência da metodologia empregada para o levantamento de transcritos codificando aspartil proteases. Foram utilizados programas da família BLAST e a sequência de cDNA ou de aminoácidos traduzida de SmCD1.1. Com ambos os bancos de dados, maior sensibilidade na busca foi conseguida utilizando a sequência da proteína. No caso de SmEST, 6 “hits” foram retornados pela busca com blastp contra 4 retornados por blastn. Analogamente, 10 “hits” foram encontrados em SmGI através da busca com tblastn, contra apenas 7 encontrados por blastn e a sequência nucleotídica. Os resultados das buscas com tblastn em SmGI e blastp em SmEST estão resumidos nas **Tabelas 6 e 7**.

Tabela 6: Lista de transcritos semelhantes à SmCD1.1 encontrados no “*Schistosoma Gene Index*”

Transcrito	Pauta de leitura	P(N)	Extensão do Alinhamento	Identidade (%)	Similaridade (%)
TC7388	+1	3 ⁻²²⁴	414	100	100
TC18677	+2	4 ⁻⁴⁷	180	52	70
TC7393	+2	5 ⁻⁴⁶	92	98	98
TC15409	+3	1 ⁻³⁰	160	38	60
BF440111	+3	2 ⁻²³	90	48	68
TC13201	-3	2 ⁻¹⁵	99	37	59
AA169900	-1	2 ⁻¹³	65	49	69
TC14786	-3	3 ⁻⁰⁹	44	59	68
H98379	+3	0,1	16	100	100
CD078122	+2	1,0	47	29	55

Tabela 7: Lista de transcritos semelhantes à SmCD1.1 encontrados no “*Schistosoma mansoni* EST genome project”

Transcrito	Pauta de leitura	P(N)	Extensão do alinhamento	Identidade (%)	Similaridade (%)
C602095	-3	1 ⁻¹⁵⁸	281	98	98
C717461	+2	1 ⁻⁴⁵	89	100	100
C701808	+3	1 ⁻⁴⁵	89	100	100
C612412	+3	4 ⁻³⁵	72	98	100
C608391	-3	1 ⁻¹⁷	141	30	53
C602644	-2	6 ⁻⁰⁷	36	52	75

Os transcritos de *S. mansoni* listados nas **Tabelas 6 e 7** foram caracterizados quanto à sua composição de ESTs, região homóloga alinhada com a seqüência de SmCD1.1 e como estes se relacionam entre si. O resultado desta análise está condensado graficamente na **Figura 27**. Conforme esperado, todos os transcritos no banco SmEST puderam ser mapeados em TCs (*tentative consensus*) do SmGI, uma vez que este último contém todas as seqüências em SmEST. Curiosamente, a forma como os agrupamentos (clusters) foram feitos em SmEST é bem diferente do observado em SmGI. Os transcritos C602095 (com 13 ESTs), C612412 (com 11 EST) e C701808 (com 1 EST) estão contidos em TC7388 (com 40 ESTs). C717461 (com 1 EST) está contido em TC7393 (com 2 ESTs). C608391 e C602644 em SmEST correspondem a TC15409 (composto por 3 ESTs) e TC13201 (composto por 2 ESTs) em SmGI, respectivamente.

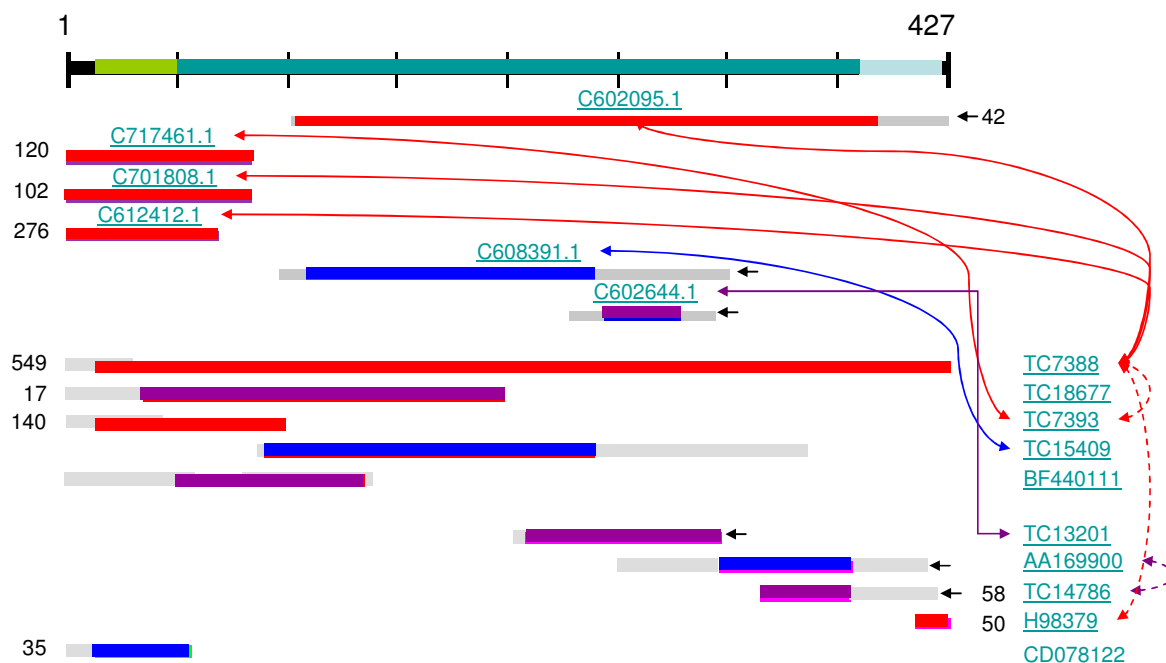


Fig. 27: Análise dos transcritos de *S. mansoni* semelhantes à SmCD1.1 presentes nos bancos “Schistosome Gene Index” (SmGI) e “*Schistosoma mansoni* EST genome project” (SmEST). No topo, a sequência da preproSmCD1.1 está representada com o peptídeo sinal em preto, a região pro em verde claro, o domínio catalítico na protease madura em verde escuro, seguida da extensão C-terminal de 43 resíduos em azul claro. Abaixo, os 6 transcritos em SmEST (com os códigos identificadores logo acima) e os 10 transcritos em SmGI (com os códigos identificadores ao lado direito) estão representados, refletindo a região homóloga à SmCD1.1. Em vermelho estão seqüências com >98%, em lilás com >40% e azul com <40% de identidade com SmCD1.1. Os segmentos em cinza representam partes dos transcritos não homólogos à SmCD1.1, enquanto os números à esquerda ou direita denotam a posição do transcrito onde o alinhamento foi truncado. Setas negras à esquerda identificam transcritos cujas fitas anti-senso foram alinhadas. As setas duplas coloridas indicam relações de parentesco entre os transcritos (vide texto).

Outra observação curiosa foi o fato de que dentro de um mesmo banco de dados observou-se que transcritos com seqüências de aminoácidos traduzidas idênticas, pelo menos dentro do segmento alinhado com SmCD1.1, apresentavam-se divididos em *clusters* ou *singletons* (*cluster* constituído por um único EST) distintos. Este foi o caso de TC7393 e H98379, cujos segmentos alinhados revelaram-se idênticos à TC7388 e consequentemente à SmCD1.1. Semelhantemente, TC14786 (com 5 ESTs) apresentou seqüência consenso idêntica ao EST denominado AA169900. Com o intuito de realizar uma estimativa mais parsimoniosa do número de membros pertencentes à família de SmAPs, considerou-se que TC7388, TC7393 e H98379 representam todos um mesmo transcrito correspondendo ao gene de SmCD1. Da mesma forma, TC14786 foi incorporado em AA169900, sendo considerados um único gene. Assim, chegou-se ao

número de pelo menos 6 genes parálogos codificando SmAPs semelhantes, porém claramente distintos de SmCD1.1. A listagem dos genes de SmAPs identificados com a nomenclatura adotada neste trabalho, juntamente com os ortólogos mais próximos é encontrada na **Tabela 8**.

Tabela 8: Genes parálogos de aspartil proteinases em *Schistosoma mansoni* (SmAPs) identificados por análise *in silico* de transcritos do parasito.

Transcritos (código/banco)*	Nome do Gene	Ortólogos mais próximos (banco/código/descrição/organismo/id%)
U60995/GB	SmCD1.1	GB/L41346/aspartic proteinase (<i>S. japonicum</i>) [345/416 (82%)] GB/AF420068/aspartic protease (<i>Clonorchis sinensis</i>) [231/411 (56%)] GB/AAZ39883/cathepsin D (<i>O. viverrini</i>) [218/411 (53%)]
CD078122/SmGI	SmCD2	GB/U60995/aspartic proteinase (<i>S. mansoni</i>) [16/47 (29%)]
TC18677/SmGI	SmCD3	GB/L41346/aspartic proteinase (<i>S. japonicum</i>) [98/208 (47%)] GB/U60995/aspartic proteinase (<i>S. mansoni</i>) [98/210 (46%)] DBJ/BAD15111/cathepsin D (<i>Todarodes pacificus</i>) [97/207 (46%)]
BF440111/SmGI	SmCD4	DBJ/BAC57431/cathepsin D (<i>Xenopus laevis</i>) [53/95 (55%)] GB/AAV90625/cathepsin D protein (<i>Sus scrofa</i>) [45/89 (50%)] GB/AF420068/aspartic protease (<i>C. sinensis</i>) [45/93 (48%)]
TC15409/SmGI	SmCD5	GB/U60995/aspartic proteinase (<i>S. mansoni</i>) [62/160 (38%)] DBJ/BAC57431/cathepsin D (<i>X. laevis</i>) [65/154 (42%)] DBJ/BAD15111/cathepsin D (<i>T. pacificus</i>) [64/156 (41%)]
TC13201/SmGI	SmCD6	EMB/ AJ876601/nothepsin (<i>Podarcis sicula</i>) [40/98 (40%)] GB/ AY792822/cathepsin D protein (<i>S. scrofa</i>) [42/99 (42%)] DBJ/BAD15111/cathepsin D (<i>T. pacificus</i>) [41/99 (41%)]
AA16900/SmGI	SmCD7	GB/AY812817/aspartic proteinase (<i>S. japonicum</i>) [60/74 (81%)] GB/AAZ39883/cathepsin D (<i>O. viverrini</i>) [33/67 (49%)] EMB/ AJ245459/necepsin II (<i>Necator americanus</i>) [35/67 (52%)]

*GB: Genebank; SmGI: *Schistosoma mansoni* Gene Index; DBJ: Japanese Nucleotide Database; EMB: European Molecular Biology Database.

Uma análise mais detalhada da seqüência dos genes de SmAPs identificados no SmGI foi realizada através de um alinhamento global com a seqüência de aminoácidos de SmCD1.1 (**Fig. 28**). Globalmente, é possível observar que nenhum dos demais genes da família de SmAPs compartilha o peptídeo sinal e a extensão C-terminal de SmCD1.1. SmCD2 alinha-se com toda a região *pro* de SmCD1.1 (14-48) e os 11 primeiros resíduos da protease madura, com maior similaridade nos segmentos 14-23 e 54-59. A seqüência consenso de SmCD3 mostrou-se homóloga a uma grande extensão

de SmCD1.1, desde a região *pro*, passando pela primeira assinatura do sítio ativo (DTG 83-85) até a posição 211. Nenhuma inserção ou deleção (*indel*) foi necessária para maximizar o alinhamento, mas diversas mutações conservativas e na sua maior parte não conservativas puderam ser observadas entre SmCD3 e SmCD1.1. Por outro lado, os resíduos de cisteína nas posições 96 e 103 foram conservados por todos os genes alinhados nesta região. O transcrito identificado como SmCD4 mostrou-se muito semelhante através de toda a região alinhada com SmCD1.1. Começando dos primeiros resíduos do N-terminal da protease madura até cerca de metade do lóbulo N-terminal da enzima, as escassas mutações observadas foram em sua maior parte não-conservativas. Destaque para a mutação G85E, inserindo uma carga negativa no sítio ativo de SmCD4. Como esta seqüência é de leitura única, é possível que esta mutação seja um artefato do sequenciamento. SmCD5 revelou-se mais semelhante a SmCD4 no segmento 96-143 do que a SmCD1.1, apresentando apenas 3 mutações em todo o intervalo. Entretanto, a seqüência de SmCD5 estende-se até a posição 255 de SmCD1.1. Blocos de maior similaridade são observados ao longo da região do alinhamento comum à SmCD1.1, SmCD3 e SmCD5. Estes blocos, em geral, coincidem com os elementos de estrutura secundária definida (α -hélice e fita- β) preditos em SmCD1 (vide Anexo 1). SmCD6 alinha-se com SmCD1.1 na região entre os resíduos 223-316. O alinhamento apresenta nível de similaridade mediano, com exceção do segmento 269-282 que compreende a 2^a assinatura do sítio ativo, no lóbulo C-terminal da enzima. Por fim, a seqüência do gene SmCD7 alinha-se no segmento 316-381 de SmCD1.

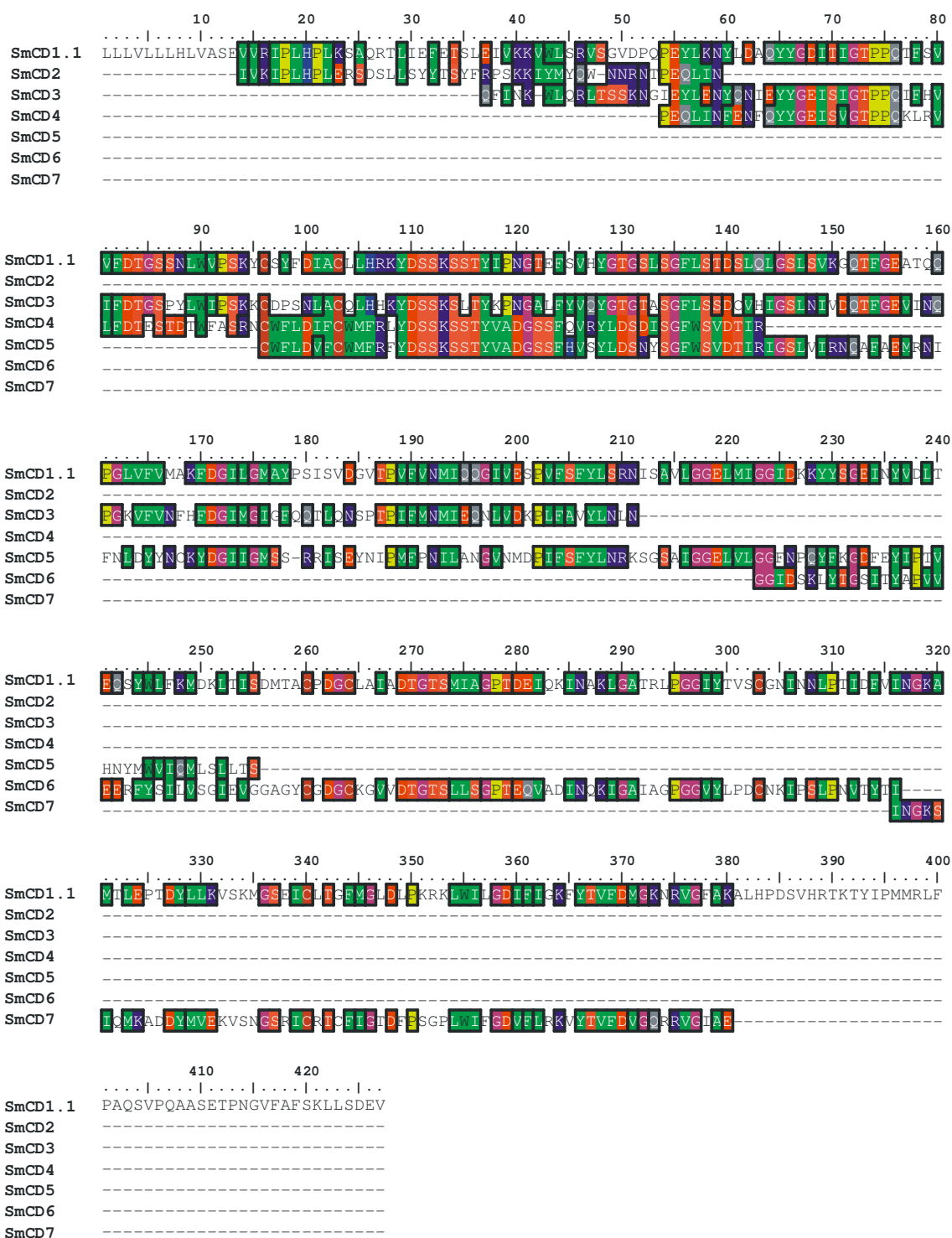


Fig. 28: Alinhamento múltiplo entre SmCD1 e fragmentos (seqüências parciais) dos demais genes pertencentes à família de SmAPs. O alinhamento foi inicialmente gerado com o programa ClustalX e em seguida editado manualmente. As seqüências foram coloridas de acordo com o limite de 20% de similaridade, usando a matriz BLOSUM62.

O nível de expressão dos genes compoendo a família de SmAPs foi estimado contabilizando-se o número de ESTs compoendo os clusters vinculados a cada gene, de acordo com a **Tabela 8**. O resultado desta análise está expresso no gráfico da **Figura 29** e indica que SmCD1 é, de longe, o gene mais expresso da família de SmAPs. Bem

abaixo, com pouco menos de 1/7 do número de ESTs encontrados até o momento, está SmCD7. Seguem em abundância SmCD3 e SmCD5, com 3 ESTs cada, SmCD6 com 2 ESTs e por fim SmCD2 e SmCD4 com apenas 1 EST cada. Foi avaliado também o nível de expressão das isoformas de SmCD1. Conforme indicado na **Figura 29**, a abundância de ESTs restritos ao domínio catalítico e de ESTs estendendo-se até a extensão C-terminal é muito semelhante. Isto indica uma possível ligação genética entre as duas formas.

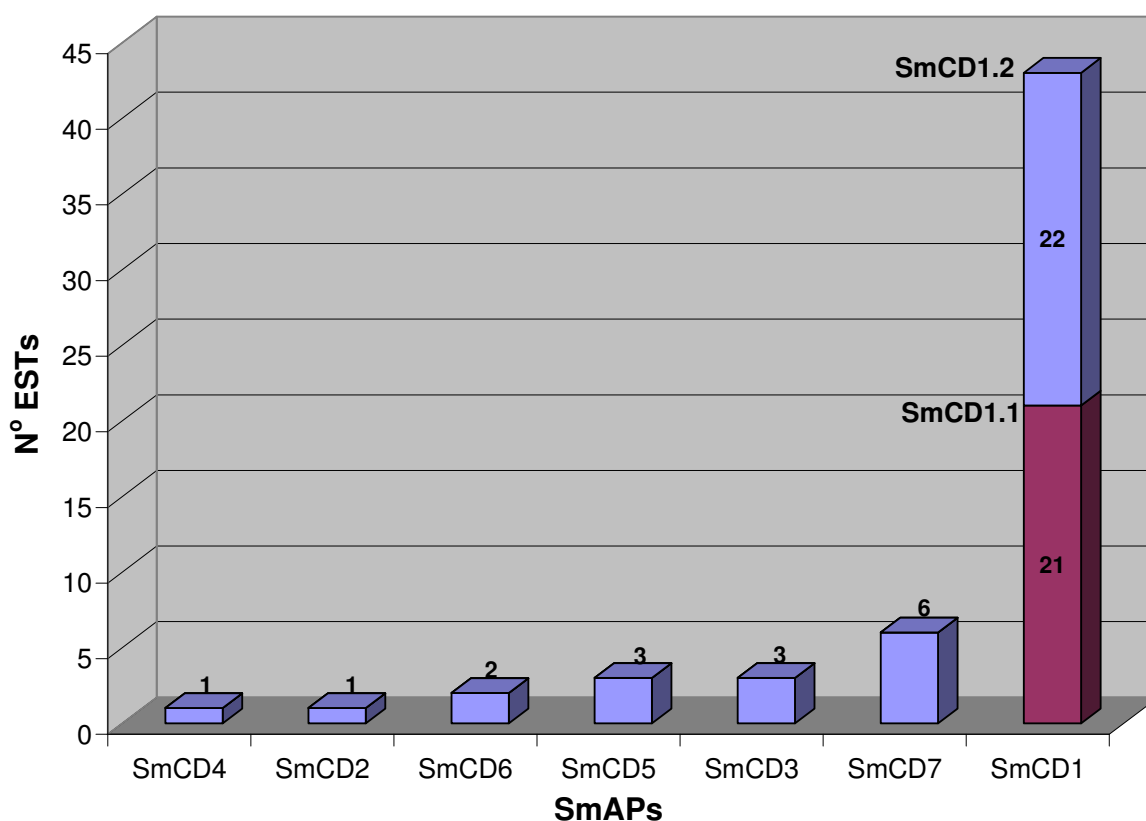


Fig. 29: Histograma dos genes da família de aspartil proteinases de *S. mansoni* (SmAPs) de acordo com o número de ESTs encontrados no SmGI Release 5.

Como um todo, os genes de SmAPs puderam ser identificados em diversos estágios do crescimento do parasito, desde ovos até formas larvais e adultas de ambos os sexos, a julgar pelas bibliotecas de origem dos ESTs. A fim de verificar a existência de uma expressão diferencial do gene mais abundante de SmAPs, SmCD1, em qualquer um dos estágios do crescimento de *S. mansoni*, foi feita uma indagação ao banco de dados com ferramenta específica disponível no SmGI. Verificou-se que não há diferença estatisticamente válida entre os níveis de expressão de SmCD1 em vermes machos ou fêmeas. De fato, somente dois transcritos foram retornados pelo programa como preferencialmente expressos em fêmeas, TC13457 e TC13518. O primeiro parece

codificar uma proteína do ovo (Chorion protein) e o segundo a subunidade 1 da NADH desidrogenase. Também foi possível verificar através da ferramenta computacional utilizada que SmCD1 não é diferencialmente expressa em vermes adultos em relação às demais formas evolutivas de *S. mansoni*.

4.4 Clonagem das formas longa e curta do zimogênio de SmCD1 (proSmCD1)

4.4.1 Otimização das condições de amplificação dos transcritos por RT-PCR

O RNA total foi extraído de 2500 vermes adultos (~ 700 mg peso úmido) recém coletados, obtendo-se 300 µg deste material. Estima-se que o conteúdo equivalente de mRNA seja de 30-100 µg. A qualidade da preparação de RNA foi avaliada pela capacidade de amplificação por RT-PCR de mensagens de genes com diferentes níveis de expressão: actina (~500 pb, abundantemente expresso), catepsina B (~650 pb, medianamente expresso) e receptor de retinol (~600 pb, raramente expresso). Fez-se um primeiro teste com 3 µg de RNA total e usando o kit em que a síntese do cDNA e amplificação são realizadas no mesmo tubo de reação (Access RT-PCR System, Promega). Neste kit, a síntese do cDNA a partir do RNA é catalizada pela transcriptase reversa do vírus da mieloblastose aviária (AMV RT) e a amplificação do cDNA é efetuada pela DNA polimerase de *Thermus flavus* (Tfl polimerase). Utilizou-se 1,25 pmol dos *primers* para amplificação de actina e catepsina B de *S. mansoni* em um volume final de reação de 50 µL. A etapa de transcrição reversa foi conduzida por 45min a 48°C e utilizou-se 60 °C como temperatura de anelamento na etapa de PCR subsequente. Em paralelo, uma amostra de RNA e *primers* para amplificação a partir desta de um fragmento de 320 pb foram utilizados como controle positivo.

Primeiramente, é interessante notar que mesmo num gel de agarose para separação de DNA, foi possível observar que o RNA de *S. mansoni* não apresenta as bandas esperadas de rRNA 28S e 18S (com a primeira aproximadamente duas vezes mais intensa que a segunda), indicativas de uma preparação de boa qualidade de RNA total da maior parte dos eucariotos (**Fig. 30**, 1). Assim, a qualidade da preparação de RNA total de *S. mansoni* só pôde ser verificada usando *primers* para testar a amplificação de mensagens de alta, média e baixa abundância. Como se pode observar na **Figura 30**, houve extensa amplificação com os *primers* de actina e do controle

positivo do kit (**Fig. 30**, 2 e 3, respectivamente). Entretanto, obteve-se amplificação bem mais fraca usando os *primers* para catepsina B (**Fig. 30**, 4), a qual é sabidamente uma mensagem menos abundante que actina. Com este resultado promissor, tentou-se a amplificação, nas mesmas condições, da mensagem de pro-SmCD1.1 usando *primers* específicos. A análise em gel de agarose de uma alíquota referente a 5% do volume total da reação de RT-PCR não revelou a banda de 1,25 kb esperada para o cDNA de pro-SmCD1.1.

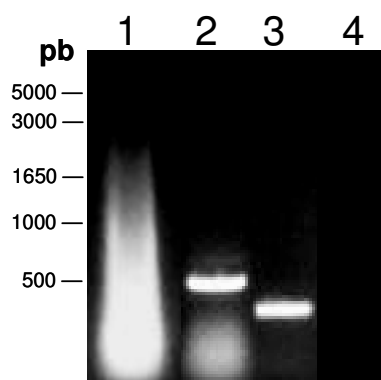


Fig. 30: Análise em gel de agarose (0,8%) de amostras do teste de amplificação com o kit Access RT-PCR System (Promega). 1- Preparação de RNA total de vermes adultos de *S. mansoni* (3 μg); 2- produto de RT-PCR usando *primers* para actina (2,5 μL); 3- produto de RT-PCR do controle positivo do kit (2,5 μL); 4- produto de RT-PCR usando *primers* para catepsina B (2,5 μL).

Um novo teste de amplificação foi realizado mantendo-se os *primers* para amplificação de catepsina B e introduzindo os *primers* de receptor de retinol (controle de amplificação de mensagens raras) e de ambas as formas de pro-SmCD1. Aumentou-se em 5 vezes a quantidade de RNA e diminuiu-se em 5 $^{\circ}\text{C}$ a temperatura de anelamento, em relação ao primeiro teste. Nestas condições, não houve amplificação com nenhum dos *primers* utilizados, inclusive de catepsina B, com os quais já se tinha obtido amplificação antes. Mais provavelmente o aumento da quantidade de RNA total usada na reação introduziu interferentes da transcrição reversa ou da ação da DNA polimerase.

Tendo em vista a ineficácia do primeiro método utilizado para a amplificação das formas longa e curta de pro-SmCD1, decidiu-se usar kit de RT-PCR que permitia a síntese e amplificação do cDNA em etapas distintas. Para a reação de síntese do cDNA, utilizou-se 1,5 μg de RNA total, *primers* (controle para amplificação do receptor de retinol e isoformas de pro-SmCD1) na concentração de 0,25 μM em um volume final de 10 μL . A transcrição reversa foi conduzida por 60min a 55 $^{\circ}\text{C}$. Na etapa posterior de

amplificação por PCR (50 μ L) utilizou-se 2 μ L da reação de síntese do cDNA, *primers* na concentração de 0,2 μ M, 55° C como temperatura de anelamento e 1,5min na etapa de extensão a 72 °C em 35 ciclos de reação. Desta vez, foi possível detectar os cDNAs amplificados de SmCD1.2 (1,1 kb) e receptor de retinol (0,6 kb) em amostras representando 20% do volume total do PCR (**Figura 31**, 1 e 2). Comparando-se as intensidades dos produtos amplificados, tem-se o indicativo que a mensagem codificando a forma curta de pro-SmCD1 é ainda menos abundante que do receptor de retinol, um transcrito considerado raro.

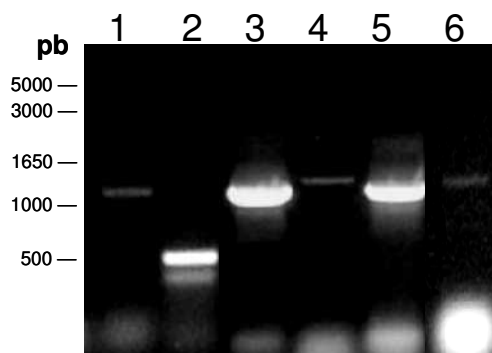


Fig. 31: Análise em gel de agarose (0,8%) de amostras do teste de amplificação com o kit ThermoScript RT-PCR System (Life Technologies). 1 e 2- produtos de PCR (10 μ L) das reação de RT (2 μ L) usando *primers* para a forma curta de pro-SmCD1 e receptor de retinol, respectivamente; 3 e 5- pro-SmCD1.2 amplificada na PCR modificada (10 μ L) usando 2 e 3 μ L da etapa de RT; 4- pro-SmCD1.1 amplificada na PCR modificada (10 μ L) usando 3 μ L da etapa de RT; 6- pro-SmCD1.1 amplificada na PCR modificada (10 μ L) usando 7 μ L da etapa de RT.

O sucesso do segundo método empregado de RT-PCR em amplificar as mensagens de mais baixa abundância pode ser fruto do sistema de enzimas (ThermoScript RT e Platinum Taq DNA polimerase High Fidelity) ou de qualquer uma das demais condições empregadas no segundo kit, uma vez que isso não foi investigado. Entretanto, nas condições acima descritas ainda não foi possível detectar o cDNA da forma longa de proSmCD1. Assim, as seguintes modificações foram introduzidas no protocolo de RT-PCR: a alíquota da reação de síntese de cDNA foi elevada para 3 μ L, o tempo de extensão foi aumentado para 2min e elevou-se o número de ciclos da PCR para 40. Com o intuito de avaliar se os efeitos observados seriam resultantes da maior quantidade de substrato ou das modificações na programação do termociclador, fez-se um controle mantendo-se os 2 μ L da reação de síntese de cDNA de proSmCD1.2 a serem utilizados nas condições modificadas de PCR. É possível observar na **Figura 31**

(3 e 5, respectivamente) que o nível de amplificação com 2 ou 3 μL de substrato da PCR foi praticamente o mesmo. Nota-se nitidamente também que as intensidades observadas no gel dos produtos de RT-PCR deste último teste foram bem superiores à intensidade anteriormente obtida (**Fig. 31**, 1). Estes resultados indicam que os 30minmin adicionais na etapa de extensão e os 5 ciclos a mais da PCR foram os responsáveis pelo maior nível de amplificação obtido. Mais importante, nestas condições foi possível detectar pela primeira vez o cDNA amplificado da forma longa de pro-SmCD1 (**Fig. 31**, 4). Pode-se observar que, nas mesmas condições, o produto final de amplificação de proSmCD1.2 (5) é bem mais abundante que pro-SmCD1.1 (4). Desconsiderando-se a diferença no tamanho dos cDNAs amplificados, este resultado indica que a mensagem de SmCD1.1 é ainda mais rara que de SmCD1.2. Aumentando-se a quantidade de cDNA usada como substrato na PCR de proSmCD1.1 para 7 μL , não foi possível obter nível maior de amplificação (**Fig. 31**, 6), confirmando os resultados anteriores com pro-SmCD1.2.

4.4.2 Clonagem em pGEM-t e sequenciamento

Os fragmentos amplificados dos genes da forma longa e curta de pro-SmCD1 foram purificados de suas bandas em gel de agarose, ligados em vetor pGEM-t e o produto de ligação usado na transformação de células competentes de *E. coli* DH5 α . Quatro colônias brancas da ligação com proSmCD1.1 e 10 colônias da ligação com proSmCD1.2 foram obtidas na placa de LB-AMP + X-Gal. As colônias brancas são indicativas da presença do inserto, pois estes clones são deficientes em β -galactosidase devido à inserção do gene clonado fora de fase com a pauta de leitura do gene desta enzima codificado em pGEM-t. Isto inativa a β -galactosidase que torna-se incapaz de converter X-gal no produto que do contrário daria coloração azul à colônia. Foram selecionadas 4 colônias brancas e uma azul (controle negativo) de cada placa para confirmação do inserto por PCR de colônia. Três das colônias brancas testadas de DH5 α /pGEM-t mostraram positivas para pro-SmCD1.1 (clones 1A, 3A e 4B) e quatro foram confirmadas quanto à presença de pro-SmCD1.2 (clones 1B-4B), conforme retratado na **Figura 32**. Curiosamente, um dos produtos amplificados de clones de DH5 α /pGEM-t/pro-SmCD1.2 apresentou pouco menos de 1kb, diferente dos demais com 1,1 kb, conforme esperado (**Fig. 32**, 8).

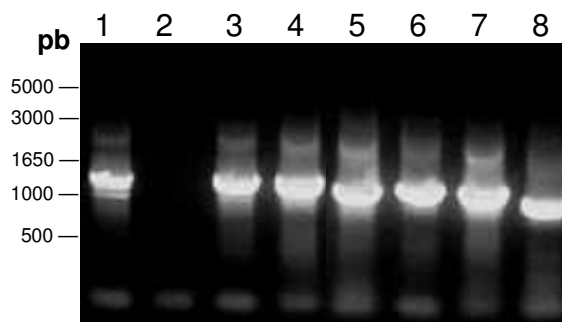


Fig. 32: Análise em gel de agarose (0,8%) de amostras de PCR de colônia para confirmação dos insertos pro-SmCD1 em clones de *E. coli* DH5 α /pGEM-t. 1 a 4- clones 1A-4A da forma longa; 5 a 8- clones 1B-4B da forma curta.

Os clones 1A-4A e 1B-4B foram cultivados e o DNA plasmidial purificado para digestão com Nde I visando a determinação da orientação dos insertos. Como pode ser verificado na **Figura 33**, todos os clones da forma longa de proSmCD1 e apenas 3 da forma curta (2B-4B) foram capazes liberar os seus insertos mediante digestão com Nde I somente. Nos clones em que o inserto foi liberado a orientação dos mesmos era apropriada para subsequente clonagem em vetor de expressão (pET28a), ou seja, sítio de Nde I em 5' e sítio de Not I em 3', logo acima do sítio Nde I do próprio vetor. É interessante notar que o clone 2A apresentou padrão de restrição coerente com os demais clones da forma longa de proSmCD1, desmentindo o resultado falso negativo do PCR de colônia. Além disso, confirmou-se que o clone 4B apresenta inserto menor que o tamanho esperado de 1,1 kb para a forma curta de SmCD1, apesar de ter o sítio de Nde I intacto.

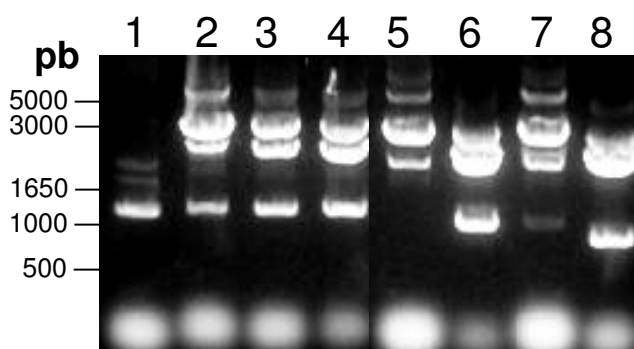


Fig 33: Análise em gel de agarose (0,8%) de amostras da digestão com Nde I para determinação da orientação dos insertos de pro-SmCD1 em clones *E. coli* DH5 α /pGEM-t. 1 a 4- clones 1A-4A da forma longa; 5 a 8- clones 1B-4B da forma curta.

Os clones 3A, 2B e 4B foram selecionados para o sequenciamento com *primers* universais M13. Dois problemas com o clone da forma longa foram revelados. A

primeira mutação foi inserida devido a um erro no *primer* reverso que teve um códon de leucina (CTT ou AAG na fita anti-senso), inadvertidamente suprimido, das posições 1228-1230 (numeração a partir da 1ª base codificando a pro-enzima), próximo à extremidade 3' do gene. A segunda mutação foi inserida naturalmente pela DNA polimerase utilizada na RT-PCR, na posição 995, onde G foi substituído por T, acarretando na mutação V→A na região comum às formas longa e curta da proteína. Mais especificamente, esta substituição ocorreu em uma alça exposta à superfície da enzima entre as fitas-b s11c e s12c no lóbulo C-terminal, distante da região de interação com o substrato conforme o modelo 3D da estrutura da forma curta da enzima (**ANEXO 1**). O clone 2B apresentou a seqüência esperada para a forma curta de SmCD1. O sequenciamento do clone 4B esclareceu que este é apenas um mutante truncado na posição 900, onde a mutação de uma base converteu um códon de Phe (TTT) em um sinal de término da transcrição (TTA). Pela análise da seqüência não foi possível determinar se este mutante surgiu do anelamento indevido com o *primer* reverso ou por uma mutação pontual, inserida naturalmente pela DNA polimerase usada na PCR.

4.5 Formas longa e curta de pro-SmCD1 produzidas em E. coli

4.5.1 Clonagem em vetor de expressão pET28a

Num primeiro momento, optou-se por trabalhar com o clone problemático da forma longa de SmCD1, uma vez que este poderia servir bem aos testes de expressão preliminares sem influenciar significativamente os resultados. Desta forma, os plasmídeos pET28a, pGEM-t/proSmCD1.1 (clone 3A) e pGEM-t/proSmCD1.2 (clone 2B) foram purificados de culturas de *E. coli* DH5α e digeridos com as enzimas Nde I e Not I. Os fragmentos lineares de 5,4kb, 1,3kb e 1,1kb foram purificados a partir de gel de agarose (**Fig. 34**). Finalmente, os cDNAs de pro-SmCD1 longa e curta foram ligados nos sítios Nde I (5') e Not I (3') de pET28a com auxílio da enzima T4 DNA ligase.

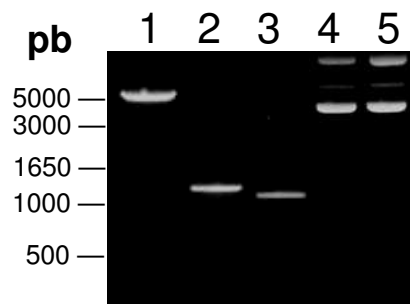


Fig. 34: Análise em gel de agarose (0,8%) de amostras da subclonagem das isoformas de proSmCD1 em pET28a. 1 a 3- fragmentos lineares purificados após digestão com Nde I e Not I. (1- pET28a; 2- proSmCD1.1 e 3- pGEM-t/proSmCD1.2, respectivamente); 4 e 5- plasmídeos purificados de clones positivos de pET28a/proSmCD1 longa (3AII) e curta (2BIV).

O produto da ligação foi usado na transformação de células competentes de *E. coli* DH5 α e os transformantes plaqueados em meio LB-Kan. Cinco colônias foram obtidas na placa de controle negativo da transformação (sem DNA) contra 5 e 7 colônias resultantes das transformações com os produtos de ligação pET28a + 3A e pET28a + 2B, respectivamente. A presença do inserto foi pesquisada por PCR de colônia em todos os clones obtidos. Apenas 1 clone positivo de pET28a/pro-SmCD1.1 (clone 3AII) e 3 de pET28a/pro-SmCD1.2 (clones 2B IV, VI e VII) foram encontrados. Os clones 3AII e 2BIV foram cultivados, o DNA plasmidial purificado (**Fig. 34**) e utilizado na transformação da cepa de expressão *E. coli* BL21(DE3).

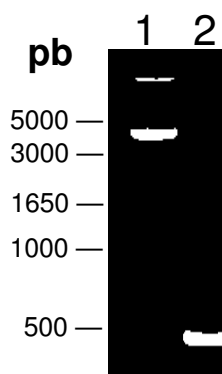


Fig. 35: Análise em gel de agarose (0,8%) de amostras referentes à correção da deleção na região 3' no cDNA da forma longa de SmCD1. 1- Preparação de pET28a/proSmCD1.1 (clone 3AII) usada na PCR; 2- Fragmento de 400 pb amplificado com *primers* SmCD-AccI (5') e SmCD-NotI/corr (3').

Para a correção do clone da forma longa de SmCD1 em pET28a (3AII), primeiramente amplificou-se o fragmento de 400 pb compreendendo o segmento que vai desde o sítio de Acc I (posição 840, contando a partir da 1ª base do gene da proenzima) até Not I na extremidade 3', usando os *primers* desenhados para a correção

(**Fig. 35**). Este fragmento foi clonado em pGEM-t obtendo-se o clone pGEM-t/CatD_Acc. O fragmento Acc I / Not I deste clone (400 pb) foi obtido por digestão com as enzimas de restrição e ligado a outros dois fragmentos lineares de 850 e 5400 pb (inserto liberado de pET28a/SmCD1.2 digerido com Nde I e Acc I e pET28a digerido com Nde I e Not I, respectivamente) para obter a construção pET28a/pro-SmCD1.1-corr. Seis colônias (1C-6C) obtidas da transformação de *E. coli* DH5 α com este produto de ligação foram pesquisadas por PCR de colônia (**Fig. 36**).

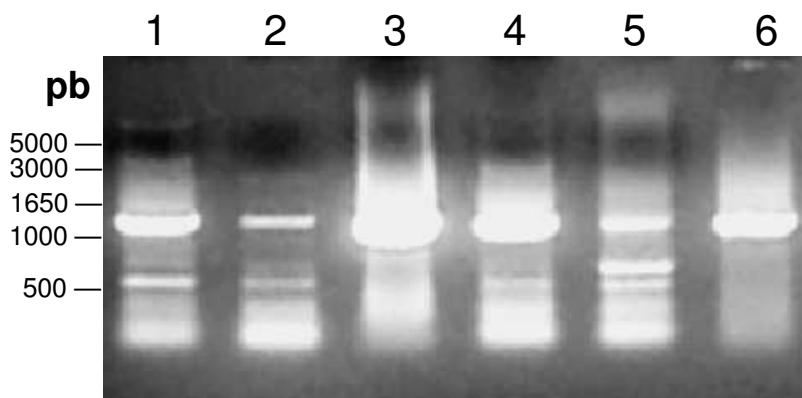


Fig. 36: Análise em gel de agarose (0,8%) de amostras de PCR de colônia para a confirmação dos insertos proSmCD1-corr em clones de *E. coli* DH5 α /pET28a. 1 a 6- clones 1C-6C.

Dois clones (3C e 5C) foram selecionados para análise de restrição e confirmação da correção por sequenciamento. Ambos os clones liberaram o fragmento esperado de 1 kb após digestão com as enzimas Bgl II (posição 401 em pET28a) e Acc I (**Fig. 37**). Entretanto, conforme pode ser observado na **Figura 37**, após digestão com as enzimas Acc I e Not I, o clone 5C apresentou um fragmento liberado menor que 400 pb. Análises posteriores também demonstraram que este clone anela com os *primers* SmCD-EcoRI/NdeI e SmCD-AccI, mas é incapaz de anelar com o *primer* SmCD-longa. Isto indica que o problema na seqüência nucleotídica deste clone ocorre entre o sítio de Acc I e a extremidade 3', antes do sítio de Not I. O clone 3C por sua vez, liberou o fragmento esperado de 400 pb a partir da digestão com Acc I e Not I e apresentou seqüência idêntica ao cDNA conhecido da forma longa de SmCD1. Este clone foi utilizado nos estudos de solubilização e renaturação desta enzima cujos resultados estão descritos nos itens 4.8 e 4.9.

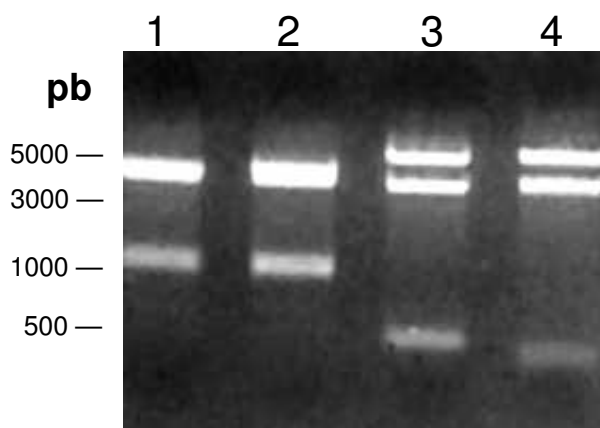


Fig. 37: Análise em gel de agarose (0,8%) de amostras dos clones corrigidos (5C e 3C) de pET28a/proSmCD1.1 clivados com enzimas de restrição. 1 e 2- clones 5C e 3C digeridos com Bgl II e Acc I; 3 e 4- clones 3C e 5C digeridos com Acc I e Not I.

4.5.2 Estudo da expressão das isoformas de SmCD1 em *E. coli* BL21(DE3)

As construções pET28a/pro-SmCD1.1 e pET28a/pro-SmCD1.2 permitem a expressão das formas longa e curta de SmCD1, respectivamente, como proteínas de fusão com uma calda N-terminal de histidina (6xHis). Mais especificamente, as proteínas são expressas com a seguinte sequência anexada ao N-terminal: “MGSSHHHHHSSGLVPRGSH”. A construção contém ainda um sítio de reconhecimento pela trombina (sublinhado, com a clivagem ocorrendo entre Arg e Gly) que permite a extração da calda de histidina da proteína expressa. Devido a esta modificação no N-terminal, as formas longa e curta do zimogênio de SmCD1 expressas em *E. coli* BL21(DE3) devem apresentar massas de 47.708,8 e 42.959,4 Daltons, respectivamente.

Os clones 3AII e 2BIV de *E. coli* BL21(DE3), correspondendo às formas longa e curta de pro-SmCD1, respectivamente, foram cultivados em pequena escala nas temperaturas de 37 °C e 30 °C. Quando a D.O.⁶⁰⁰ das culturas atingiu 0,8 U.A., a expressão das proteínas recombinantes foi induzida pela adição de 1 mM de IPTG e as culturas incubadas por mais 4h. Alíquotas das células foram coletadas ao longo da indução e processadas para obtenção de extratos aquosos (fração solúvel) e corpos de inclusão. A análise destas amostras por SDS-PAGE 10% redutora está ilustrada na **Figura 37**.

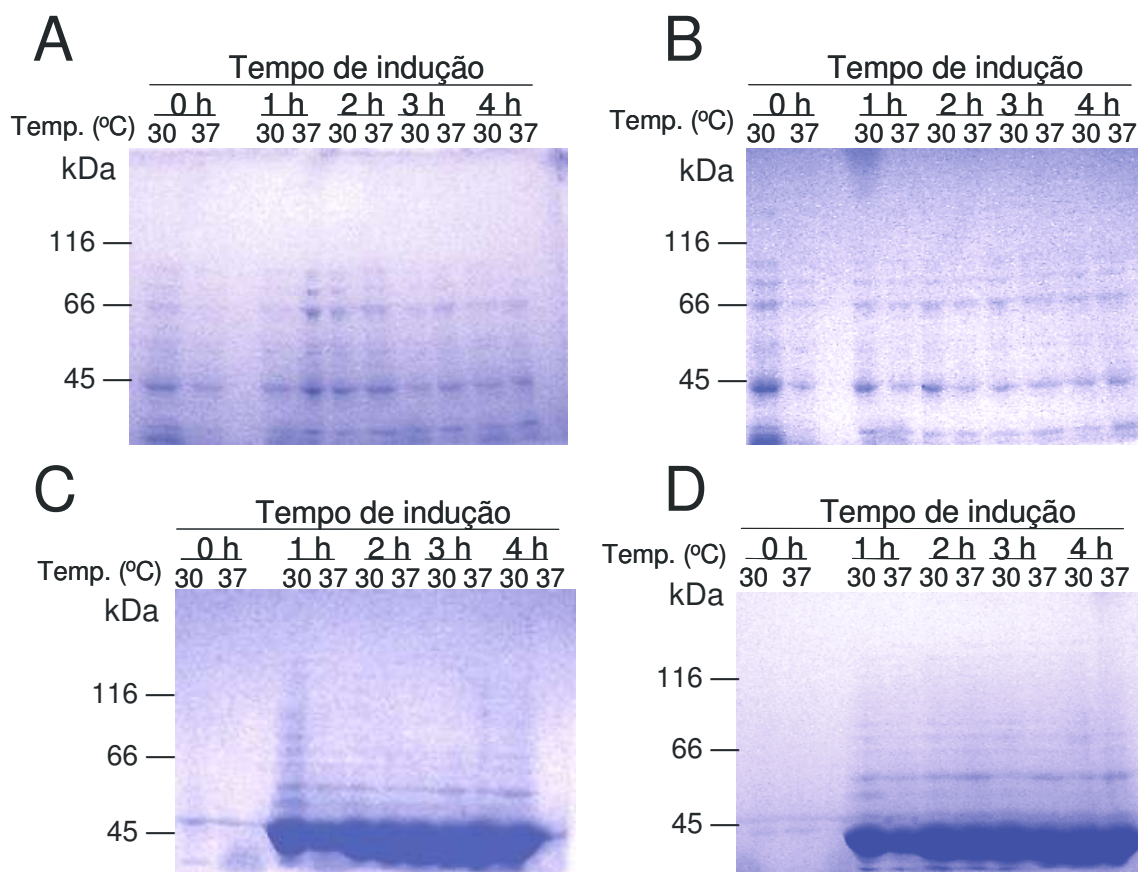


Fig. 38: Análise por SDS-PAGE (10%, redutora) de amostras do teste de indução da expressão das formas longa e curta de SmCD1 em *E. coli* BL21(DE3). A e B- amostras da fração solúvel de (6xHis)pro-SmCD1.1 e (6xHis)proSmCD1.2, respectivamente. C e D- amostras da preparação de corpos de inclusão de (6xHis)proSmCD1.1 e (6xHis)proSmCD1.2, respectivamente. Revelação por Coomassie brilliant blue G-25.

Não é possível perceber nos painéis A e B da **Figura 38** as bandas esperadas em 48 e 43 kDa nas frações solúveis dos extratos de *E. coli* BL21(DE3), expressando as formas longa e curta de pro-SmCD1, respectivamente. Diferentemente, os painéis C e D evidenciam claramente estas bandas nas preparações de corpos de inclusão. Isto indica que tanto a 37° C, como numa temperatura mais baixa (30 °C) e independentemente do tempo de indução, estas proteínas são super-expressas somente como agregados insolúveis. Pode-se concluir também, pela ausência destas bandas nas amostras do tempo 0h de indução, que as proteínas são sintetizadas somente após a adição de 1 mM de IPTG ao meio de cultura.

4.5.3 Triagem de condições de solubilização de pro-SmCD1 presente em corpos de inclusão

Culturas na escala de 1L foram produzidas para a preparação de corpos de inclusão (CI) dos clones 3AII e 2BIV de *E. coli* BL21(DE3) expressando as formas

longa e curta de pro-SmCD1. As células foram lisadas com lizozima e homogeneizadas por sonicação. Após separação da fração solúvel por centrifugação e lavagem do material precipitado, foram obtidos 144-163 mg de proteína nas preparações de CI, representando cerca de 40% da proteína celular total (**Tabela 9**). O conteúdo das proteínas recombinantes nas preparações de CI foi estimado em 70% através das intensidades relativas das bandas em um gel de poliacrilamida (**Fig. 39**, 7 e 8). Assim, estima-se que o rendimento da expressão seja 100-114 mg de pro-SmCD1 por litro de cultura.

Tabela 9: Rendimento em mg de proteína obtido das preparações de corpos de inclusão (CI) das formas longa e curta de SmCD1.

Fração	Proteína (mg)	
	(6xHis)proSmCD1 Longa	(6xHis)proSmCD1 Curta
Lisado	402,5	468,0
Solúvel	237,2	281,2
1º lavado	74,2	82,5
2º lavado	12,0	16,5
CI	163,0	144,0

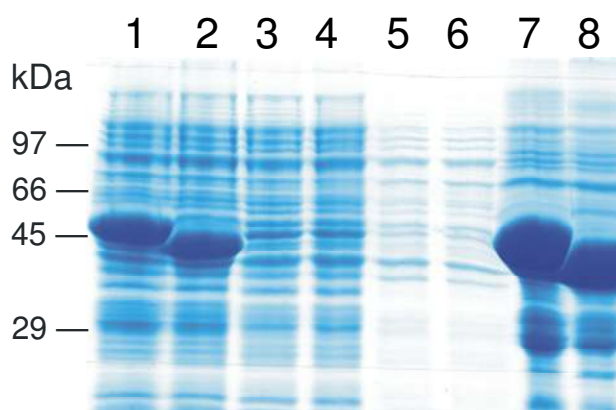


Fig. 39: Análise por SDS-PAGE (12%, redutora) de amostras de preparações utilizadas nos ensaios de solubilização de corpos de inclusão (CI) de proSmCD1. Números ímpares correspondem às frações da preparação da forma longa e pares da forma curta: 1/2- Lisado celular; 3/4- Fração solúvel; 5/6- 1º lavado; 7/8- CI. Revelação por Coomassie brilliant blue G-25.

Ensaio preliminares mostraram que 8M de uréia, por si só, não era um bom agente solubilizante dos CI de pro-SmCD1. Por outro lado, observou-se que o tampão de amostra da SDS-PAGE, contendo 0,5% de SDS, solubilizava eficientemente estas preparações. Devido aos potenciais problemas que este detergente poderia trazer nas etapas posteriores de análise da proteína, optou-se por avaliar o efeito do pH, agente

caotrópico e aditivos na solubilização dos CI de pro-SmCD1. A eficiência de cada condição de solubilização testada foi expressa em termos de quantidade de proteína solubilizada compara a uma condição “padrão” com 1% de SDS em pH 7,5. A **Figura 40** mostra que quanto mais alcalino o tampão de solubilização maior a sua eficiência, chegando a 70% dos CI solubilizados pela condição padrão em pH 10. As duas isoformas da enzima apresentaram comportamento bastante semelhante, um resultado compatível com os 404 resíduos de aminoácidos compartilhados pelas duas proteínas recombinantes com calda 6xHis.

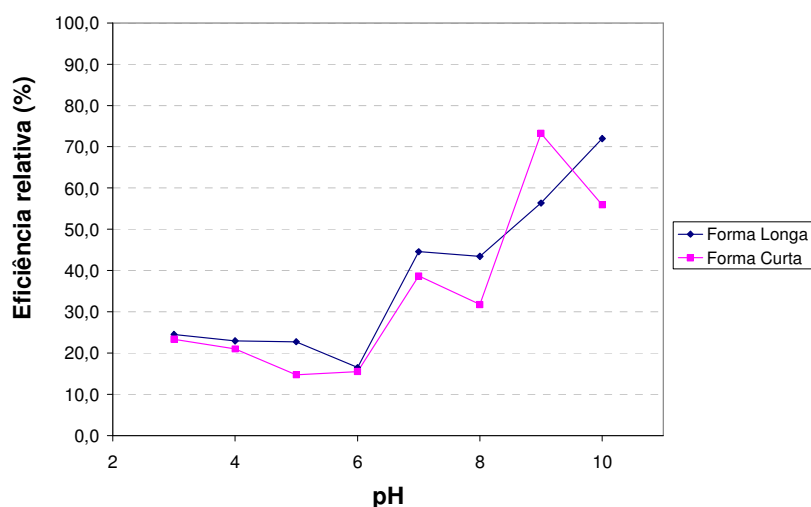


Fig. 40: Efeito do pH na solubilização dos corpos de inclusão de proSmCD1 por 8M uréia.

Três agentes solubilizantes, entre eles 2 caotrópicos, foram pesquisados em pH 7,5 quanto a capacidade de solubilização dos CI de pro-SmCD1 (**Fig. 41**). Guanidina-HCl mostrou ser o mais eficiente de todos, solubilizando de 15-20% a mais de proteínas que a condição padrão com SDS, nas concentração de 4-6M. O detergente zwitteriônico sarcosil teve um máximo de eficiência na concentração de 1% (p/v), a qual foi comparável aos cerca de 40% de eficiência relativa na solubilização dos CI conseguida com 8M de uréia. Em geral, a resposta às diferentes concentrações dos agentes solubilizantes foi a mesma para as duas formas de pro-SmCD1. Em termos absolutos, uréia foi mais eficiente para a forma curta, principalmente nas concentrações mais baixas de 4 e 6M. No caso de guanidina-HCl, somente na concentração mais baixa (4M) houve uma diferença na solubilização das duas formas, sendo esta 20% melhor na forma longa.

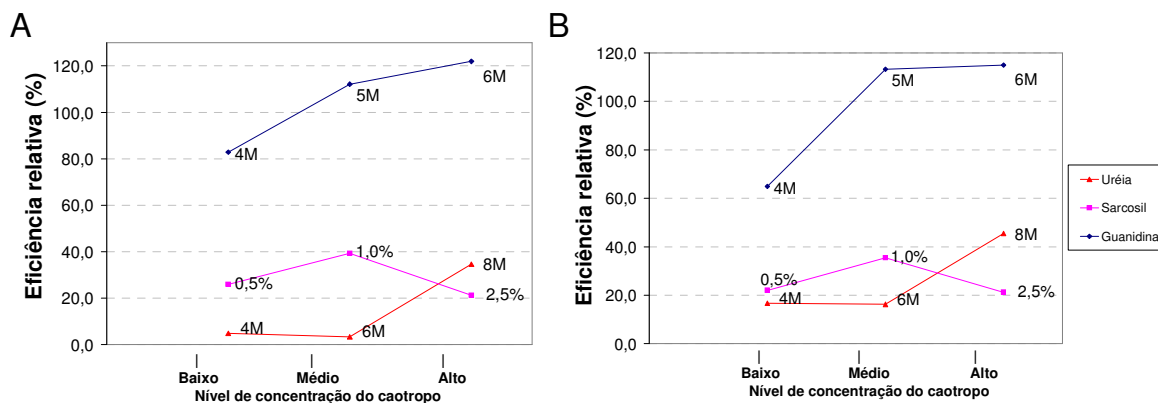


Fig. 41: Efeito de uréia, sarcosil e guanidina-HCl (em pH 7,5) na solubilização dos corpos de inclusão das formas longa (A) e curta (B) de proSmCD1.

Por fim, foram avaliados os efeitos de 5 reagentes como aditivos na solubilização de proSmCD1 conseguida por 8M de uréia em pH 7,5 (**Fig. 42**). Os reagentes testados foram NaCl e 4 detergentes, dois não-iônicos (Brij-35 e OG) e dois zwitteriônicos (sarcosil e CHAPS). Sarcosil desta vez foi usado em concentrações mais baixas em relação às concentrações utilizadas quando este foi avaliado como agente solubilizante principal. Tendo em vista a eficiência de 40% na solubilização alcançada por 8M de uréia somente, para ambas as formas de proSmCD1 (**Fig. 40**), pode-se observar na **Figura 42** que a maior parte das condições testadas com estes aditivos foi prejudicial para a solubilização da forma longa do zimogênio. O detergente CHAPS, a baixas concentrações, foi o aditivo mais eficiente em promover a solubilização da forma curta. Em relação à forma curta de pro-SmCD1, destacou-se a elevada eficiência de solubilização conseguida com sarcosil a 1%, indicando um efeito sinérgico entre este detergente e uréia. Outras condições que favoreceram significativamente a solubilização conseguida com uréia foram sarcosil a 0,5 e 1,5% para a forma longa e CHAPS a 0,1%, e Brij-35 a 0,5% para a forma curta.

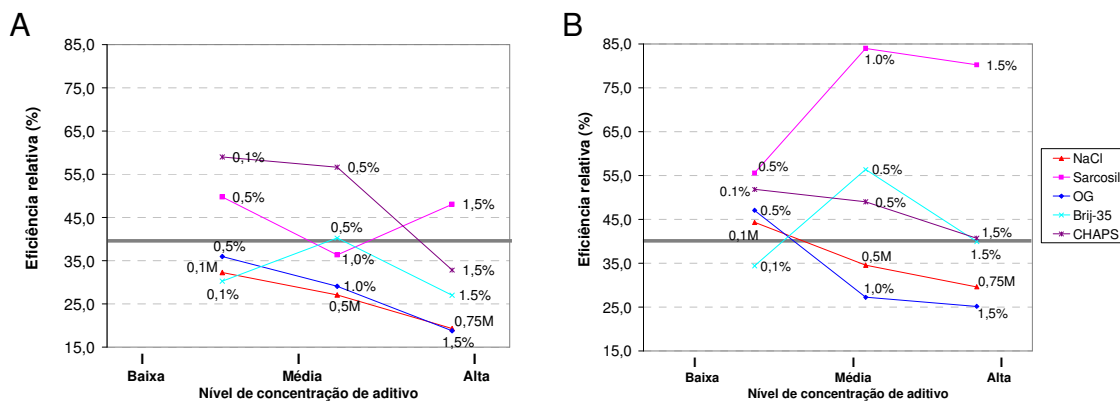


Fig. 42: Efeito de aditivos na solubilização dos corpos de inclusão das formas longa (A) e curta (B) de proSmCD1 por 8M uréia, pH 7,5. A linha em cinza demarca a eficiência de solubilização alcançada por 8M uréia em pH 7,5 sem aditivos.

4.5.4 Purificação por IMAC e determinação do estado de agregação em 4M de guanidina-HCl

As preparações de corpos de inclusão foram solubilizadas em tampão pH 8,5 contendo 6M de guanidina-HCl, 100 mM de NaCl e 10 mM de imidazol. Este material foi aplicado em coluna empacotada com a resina Ni-IDA para purificação das proteínas fusionadas com a calda N-terminal de histidina em condições desnaturantes. Após extensiva lavagem da coluna, as proteínas fortemente ligadas à resina foram eluídas em tampão pH 7,5 contendo 4M de guanidina pela presença de 100 mM de imidazol. O perfil de eluição das amostras de pro-SmCD1 longa e curta é mostrado na **Figura 43A**. A análise por SDS-PAGE das frações eluídas (**Fig. 43B**) mostrou a presença de contaminantes de alto e baixo peso molecular. Através de análise densitométrica do gel de poliácridamida estimou-se que pro-SmCD1 foi concentrada no pico da eluição até um nível de pureza de 85%.

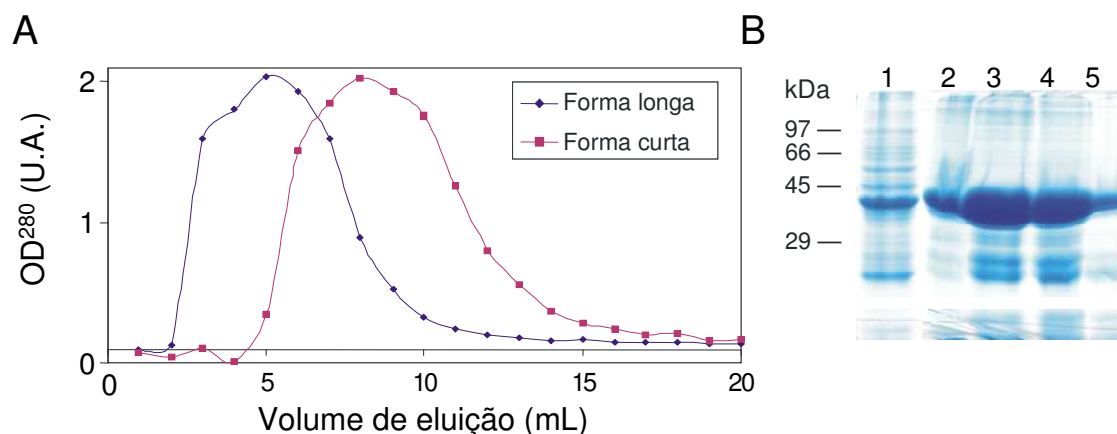


Fig. 43: Purificação das isoformas de (6xHis)proSmCD1 por cromatografia de afinidade em Ni-IDA-agarose (5 mL) em condições desnaturantes. A- Eluição com 100 mM imidazol em 4M de guanidina-HCl. B- Análise por SDS-PAGE (12%, redutora) da forma curta do zimogênio; 1- aplicado, 2 a 5- frações 6-9 da eluição. Revelação por Coomassie brilliant blue G-25.

Através da cromatografia em Ni-IDA, a concentração de guanidina-HCl foi reduzida de 6M para 4M com o intuito de facilitar as etapas posteriores de renaturação. Entretanto, com a redução da concentração do agente caotrópico é possível que alguma agregação da proteína recombinante tenha ocorrido. Desta forma, o estado de agregação das isoformas de SmCD1 purificadas em Ni-IDA foi avaliado por cromatografia de exclusão molecular (SEC) em 4M de guanidina. Como pode ser observado na **Figura 44** as isoformas de pro-SmCD1 foram eluídas em volume superior à BSA (66,6 kDa) e inferior à anidrase carbônica (29,0 kDa). Além disso, a diferença entre os volumes de eluição da forma longa e curta é compatível com a diferença de massa de 5 kDa entre as duas isoformas. A análise por SDS-PAGE das frações da SEC de ambas as formas de SmCD1 indica que estas são as proteínas majoritárias em todos os picos coletados. A integração das áreas destes picos nos cromatogramas indica que o pico principal corresponde à 90% da área total. Estes resultados mostram que ambas as formas de SmCD1 apresentam-se como partículas 90% no estado monodisperso em 4M de guanidina.

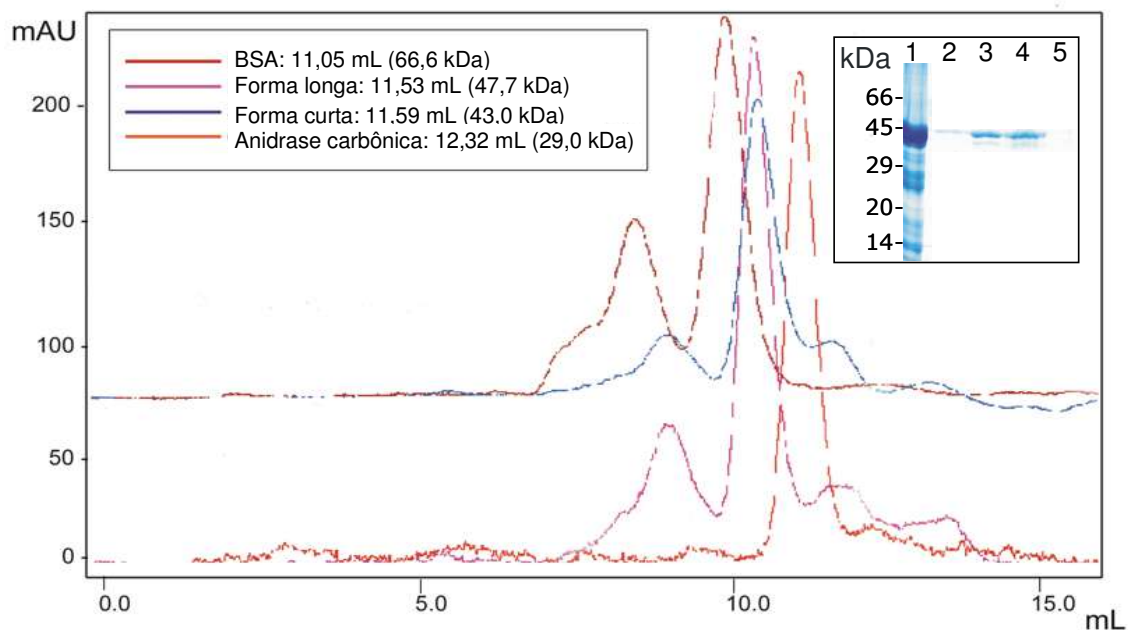


Fig. 44: Análise do estado de agregação das isoformas de SmCD1 por cromatografia de exclusão molecular (SEC). As corridas cromatográficas foram realizadas em 4M de guanidina-HCl em um sistema FPLC e registradas a 280 nm. Frações de 0,5 mL foram coletadas ao longo da corrida. No detalhe é mostrada análise por SDS-PAGE (12%, revelação por Coomassie brilliant blue G-25) das frações da SEC com a forma curta do zimogênio; 1- aplicado; 2 a 5- frações 23 a 26, correspondendo ao pico eluído.

4.5.5 Identificação de condições favoráveis à renaturação de pro-SmCD1 empregando desenho fatorial fracionado

A elevada proporção de espécies monoméricas favorece a renaturação das isoformas de pro-SmCD1, pois a reação colateral de agregação é minimizada. Desta forma, uma tentativa inicial de renaturação foi feita diluindo-se lentamente amostras de pro-SmCD1 longa e curta em volume 250 vezes maior de tampão sem agente desnaturante. A medida que as amostras eram gotejadas no tampão percebia-se o aparecimento de material floculado, indicativo da formação de agregados insolúveis das proteínas recombinantes. Após 48h, as amostras foram concentradas cerca de 100 vezes, centrifugadas a 20.000g e o sobrenadante pesquisado quanto a presença das isoformas de pro-SmCD1 renaturadas. Não foi possível detectar atividade proteolítica atribuída à aspartil proteinases, seja usando substratos cromogênicos, fluorogênicos ou naturais (hemoglobina desnaturada em pH ácido). Nem mesmo a análise por SDS-PAGE utilizando revelação pela prata pôde detectar a presença das proteínas recombinantes no material solúvel.

Com os resultados descritos acima, optou-se por investigar mais cuidadosamente condições experimentais que pudesse ser favoráveis à renaturação das isoformas de SmCD1 solúveis em 4M de guanidina-HCl. Para isto, adotou-se um desenho experimental fatorial fracionado, onde 10 fatores envolvidos no enovelamento de proteínas seriam avaliados em dois níveis, com o menor número de experimentos possível. Como preparação para estes experimentos, ensaios preliminares foram realizados com o intuito de fixar algumas variáveis experimentais consideradas críticas no processo de renaturação de uma proteína: método de renaturação (diálise ou diluição), concentração inicial de proteína (1 mg/mL ou 0,1 mg/mL) e tempo de diálise (4 ou 8h) necessário para completa retirada dos aditivos (guanidina-HCl, L-arginina e íons divalentes) com potencial para atuarem como interferentes nas etapas subsequentes de análise (SDS-PAGE, CD ou HPLC). Estes experimentos preliminares novamente evidenciaram a ausência de atividade enzimática nas amostras da renaturação. Os ensaios também mostraram que 8h de diálise foi insuficiente para retirar todos os aditivos, a julgar pela interferência com o tampão de amostra da SDS-PAGE. Entretanto, analisando amostras precipitadas com TCA foi possível detectar bandas no gel de poliacrilamida compatíveis com as massas esperadas das isoformas de pro-SmCD1.

Baseado nos resultados preliminares descritos acima, planejou-se um desenho fatorial fracionado nos moldes do desenho elaborado por Armstrong *et al.* (1999). Uma vez que nas condições testadas anteriormente não foi possível detectar atividade enzimática, mas apenas pro-SmCD1 solúvel nas amostras submetidas à renaturação, decidiu-se adotar dois tipos de medida da resposta aos fatores testados: a atividade frente ao substrato fluorogênico e a D.O.²⁸⁰ das amostras. Ao todo, 32 ensaios foram realizados para cada isoforma, correspondendo a 16 condições de renaturação em replicata e aleatoriamente testadas. As medidas de atividade enzimática foram feitas em duplicata, não sendo possível detectar atividade peptidásica em nenhuma das 32 amostras testadas. Estes ensaios foram validados através do emprego de catepsina D bovina na ausência e presença de 1 μ M de pepstatina como controles positivo e negativo, respectivamente. Entretanto, foi possível correlacionar a D.O.²⁸⁰ das amostras da renaturação com a intensidade das bandas com massa compatível com as isoformas de pro-SmCD1 (**Fig. 45**). De fato, o gel de poliacrilamida na **Figura 45** mostra que estas espécies com ~48 e ~43 kDa são majoritárias. Desta forma, mesmo na ausência de

atividade enzimática decidiu-se analisar o resultado do desenho fatorial fracionado utilizando a D.O.²⁸⁰ das amostras como resposta.

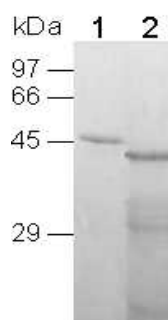


Fig. 45: Análise por SDS-PAGE (12%, redutora) de amostras do ensaio de renaturação das isoformas de proSmCD1. Alíquotas de 60 μ L das amostras que permaneceram solúveis na condição de renaturação 9 (vide **Tabela 3**) foram precipitadas com TCA, ressuspendidas em 20 μ L de tampão de amostra e aplicadas no gel. 1- forma longa e 2- forma curta. Revelação por Coomassie brilliant blue G-25.

A **Figura 46** mostra o comportamento dos modelos de regressão lineares (MRL) gerados na análise dos dados de renaturação das isoformas de SmCD1 através do gráfico das D.O.²⁸⁰ observadas experimentalmente contra as D.O.²⁸⁰ previstas pelos modelos. Pode-se perceber que o modelo da **Figura 46A**, relativo à forma curta da proenzima, apresenta uma distribuição mais regular dos dados por toda a faixa de D.O.²⁸⁰ observadas. Ao contrário, os dados coletados para a forma longa apresentam dois pontos afastados do grupo principal (**Fig. 46B**). Estes pontos correspondem ao experimento de número 7 na ordem padrão e sua replicata (**Tabela 3**). A condição de renaturação representada por este experimento apresenta elevada força iônica (250 mM NaCl + 10 mM KCl) em tampão 50 mM Tris-HCl pH 8,2 contendo 4% de glicerol. Além disso, os valores previstos pelo MRL para a forma curta apresentam um menor resíduo (a julgar pelas distâncias dos pontos da reta vermelha) e o modelo como um todo parece ser mais significativo considerando a maior inclinação da reta vermelha relativa à linha pontilhada azul. Confirmando esta análise, os parâmetros estatísticos mostram uma clara superioridade do MRL para a forma curta do zimogênio de SmCD1 ($R^2 = 0,86$; $p < 0,001$; RMSE = 0,038; $n = 32$) em comparação ao MRL gerado a partir dos dados para a forma longa ($R^2 = 0,60$; $p < 0,013$; RMSE = 0,150; $n = 32$). Entretanto, foi considerado que mesmo o modelo da forma longa ainda era útil, pois conseguia explicar a maior parte da variância nos dados e com um erro menor do que a média das respostas, indicando que este era dotado de alguma capacidade preditiva.

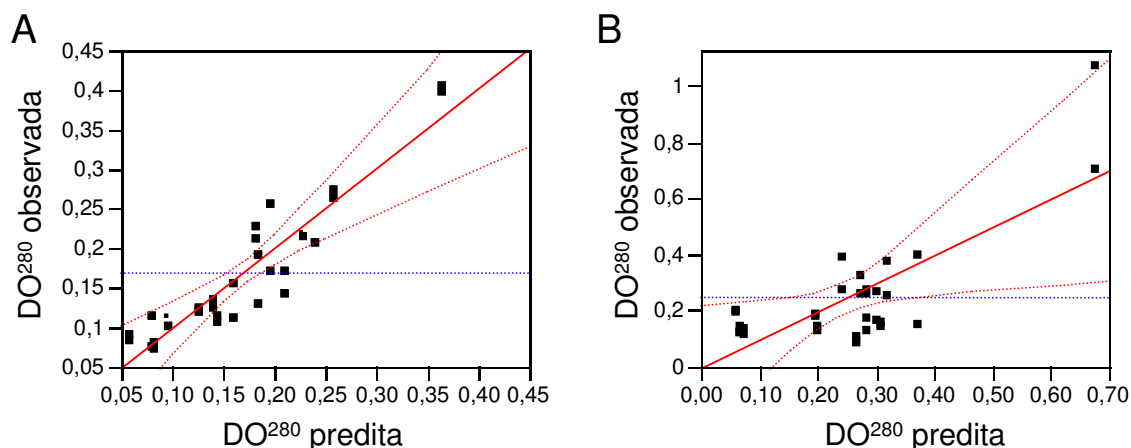


Fig. 46: Gráficos da D.O.²⁸⁰ observada versus predita pelos modelos de regressão linear da triagem de condições de renaturação da forma curta (A) e longa (B) de proSmCD1. A linha pontilhada em azul marca a média nas respostas. A linha sólida vermelha representa o modelo ideal, onde observado = predito. As linhas pontilhadas vermelhas demarcam o intervalo de 95% de confiança.

Com a qualidade dos modelos aprovada estatisticamente, passou-se à análise dos fatores influenciando o reenovelamento *in vitro* das isoformas de pro-SmCD1. No caso da forma longa, a análise estatística realizada no programa JMP In apontou 6 fatores significantes. Em ordem decrescente de importância estes fatores são: presença do agente caotrópico, aditivo polar, tipo de agente redox, presença de polioliol, força iônica e presença de cátions divalentes. No caso do modelo para a forma curta, 7 fatores mostram-se estatisticamente significantes. Em ordem decrescente de importância estes fatores são: presença do agente caotrópico, aditivo polar, tipo de agente redox, tampão/pH, concentração de proteína, força iônica e presença de polioliol. Como os 5 fatores mais importantes para a recuperação das isoformas de pro-SmCD1 solúvel dos ensaios de renaturação impactam nas respostas observadas é exibido nos gráficos da **Figura 47**. Uma característica feral destes gráficos é que quanto mais importante o fator mais inclinada é a curva. Uma inclinação positiva indica que o nível mais alto do fator favorece o aumento da resposta, no caso a D.O.²⁸⁰ das amostras solúveis após a renaturação. Assim, percebe-se que a presença de 1M de guanidina-HCl, 500 mM de L-arginina e 1 mM de DTT como agente ativo sobre oxi-redução nos tampões da renaturação favorece a recuperação para ambas as isoformas de pro-SmCD1 na forma solúvel. Este resultado é compatível com a elevada semelhança estrutural entre as duas proteínas recombinantes testadas. Entretanto, a presença da extensão de 43 resíduos de aminoácidos no C-terminal da forma longa parece confirmar alguma peculiaridade às propriedades físico-químicas desta proteína em relação à forma curta. A forma curta

tem seu processo de renaturação mais influenciado (positivamente) pelo pH alcalino (8,5) do meio e alta concentração inicial de proteína (1 mg/mL). Já a forma longa tem a recuperação na forma solúvel favorecida pela omissão de 4% glicerol e uso de menor força iônica (10 mM NaCl + 0,4 mM KCl) nos tampões de renaturação.

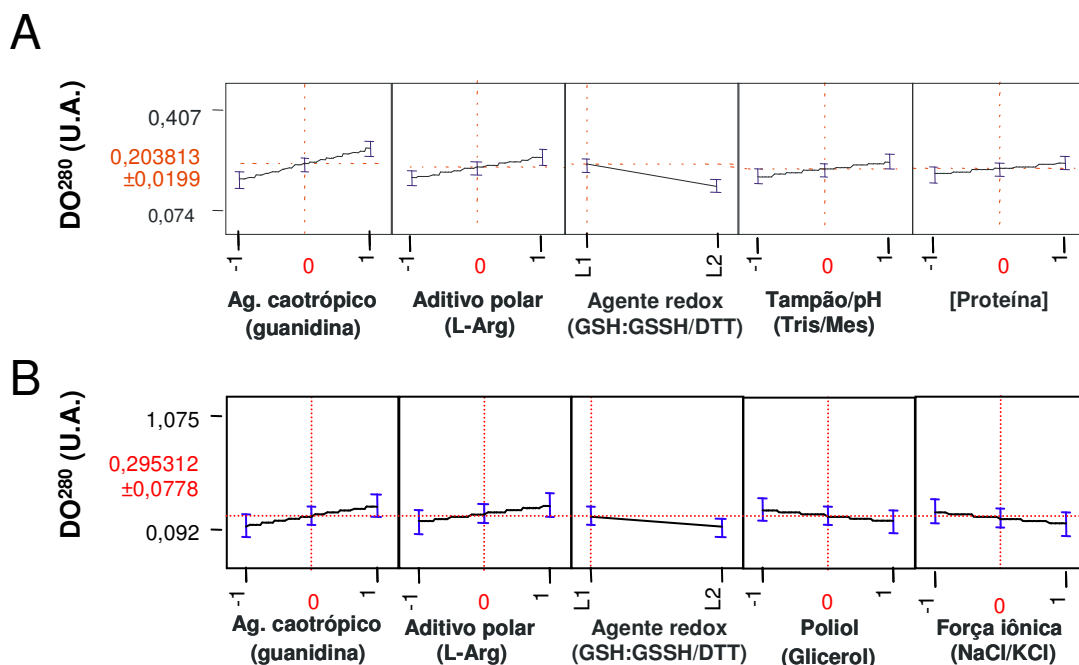


Fig. 47: Perfis preditivos dos 5 fatores mais importantes para a recuperação das isoformas de proSmCD1 solúvel dos ensaios de renaturação. Os símbolos “-1” ou L1 e “1” ou L2 representam os níveis baixo e alto dos fatores, respectivamente. As linhas pontilhadas horizontais marcam a resposta esperada (vermelho à esquerda da escala) quando todos os fatores estiverem em seus valores intermediários. A- Forma curta. B- Forma longa.

Como mostra a **Figura 48**, a análise estatística baseada nos MRL gerados para os dados de renaturação das isoformas de SmCD1 é corroborada pela análise independente dos efeitos principais normalizados dos fatores pesquisados no desenho fatorial fracionado. Novamente, agente caotrópico, aditivo polar e agente ativo sobre a oxi-redução foram os fatores mais influentes na renaturação de ambas as isoformas de pro-SmCD1. Da mesma forma, poliol e força iônica foram mais importantes para a forma longa, enquanto que concentração de proteína e pH do tampão influenciaram mais a recuperação da forma curta solúvel após o procedimento de renaturação.

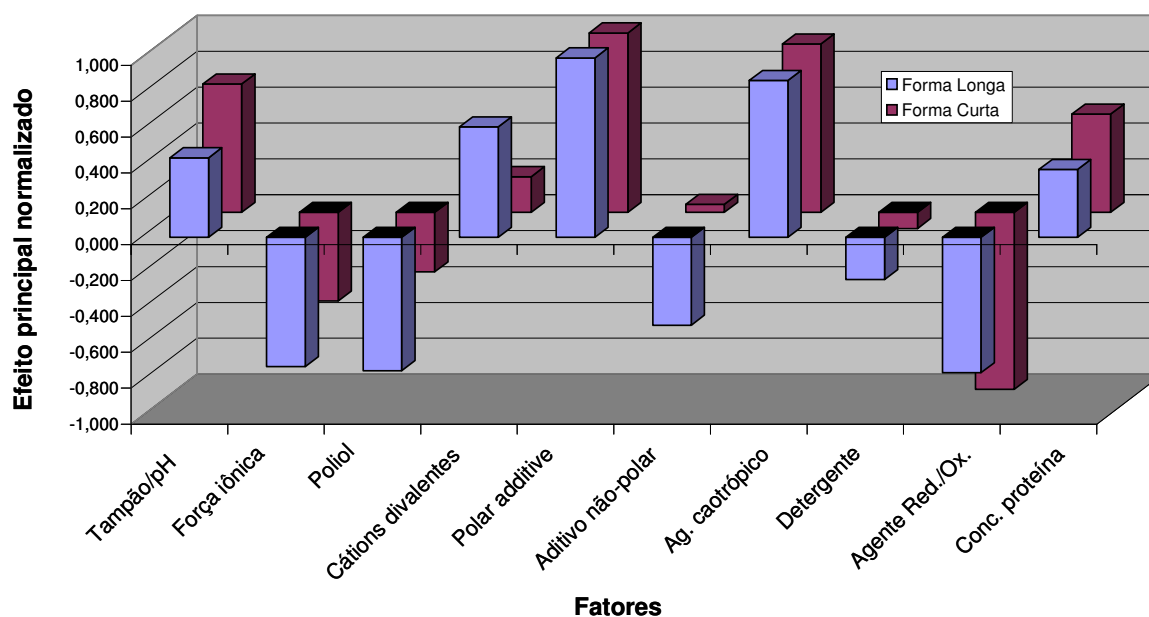


Fig. 48: Gráfico dos efeitos principais normalizados para cada um dos dez fatores avaliados no desenho fatorial fracionado 2^{10-6} com resolução III.

4.5.6 Obtenção de anti-soros policlonais contra frações semi-purificadas de SmCD1 desnaturada

Os antígenos utilizados na imunização de coelhos consistiram nas isoformas de pro-SmCD1 fracionadas em coluna de Ni-IDA. Algumas dificuldades foram encontradas na preparação dos soros policlonais. A primeira delas provém da dificuldade de manipulação de uma proteína insolúvel em água, pois isto reduz as possibilidades de métodos de purificação. Por exemplo, uma proteína solubilizada em 4M de guanidina-HCl não pode ser cromatografada em uma resina de troca iônica. Assim, optou-se por obter um antígeno mais puro recortando-se as bandas das proteínas do gel de poliacrilamida, que foram em seguida maceradas, resuspendidas na emulsão de adjuvante de Freund em PBS e inoculadas por via sub dérmica no dorso do animal. Este procedimento rendeu soros com título muito baixo (entre 1:50 e 1:100), julgando-se pelos resultados de *Immunoblot* usando substrato cromogênico da peroxidase. Títulos tão baixos não seriam úteis devido a mais uma dificuldade: o soro pré-imune dos coelhos apresenta reação cruzada com as proteínas recombinantes até uma diluição de 1:50 nos imunoblots realizados.

Desta forma, optou-se por fazer a imunização com os antígenos semi-purificados, inoculando-se diretamente a fração ligada à Ni-IDA. Como acreditava-se que 4M de guanidina inoculado subdérmico teria um forte efeito necrosante, as proteínas foram primeiramente precipitadas com TCA, o pellet lavado, seco e

finalmente ressuspensão em PBS. O antígeno era então obtido pela emulsão com adjuvante completo ou incompleto de Freund (vide material e métodos). O título dos soros foi acompanhado por imunoblots, obtendo-se após a 4ª inoculação soros com título entre 1:250 e 1:500, contra as proteínas recombinantes (**Fig. 49**).

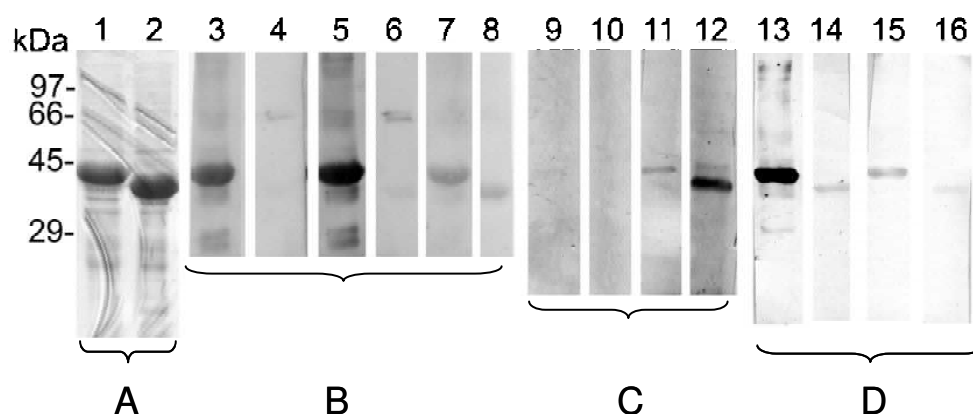


Fig. 49: Análise da produção de soros policlonais em coelhos imunizados com as formas longa e curta de proSmCD1. A- Análise dos antígenos por SDS-PAGE (12 %, redutora revelada por Coomassie G-250): 1- forma longa (FL); 2- forma curta (FC). B, C e D- respectivamente, *immunoblots* utilizando soros da 2ª, 3ª e 4ª inoculação contra material em A transferido para membranas de NC (FL em 3, 5, 7, 9, 11, 13 e 15; FC no restante): 3- anti-FL 1:100; 4- anti-FC 1:100; 5- anti-FL 1:50; 6- anti-FC 1:50; 7 e 8- soro pré-imune 1:50; 9 e 10- soro pré-imune 1:100; 11- anti-FL 1:250; 12- anti-FC 1:100; 13- anti-FL 1:250; 14- anti-FC 1:250; 15- anti-FL 1:500; 16 anti-FC 1:500.

4.6 Domínio COOH-terminal da forma longa de SmCD1 produzido em *E. coli*

4.6.1 Construção pBuff5/pep_C-term e teste de indução em BL21(DE3)

O cDNA codificando a extensão C-terminal de 45 resíduos, específica da forma longa de SmCD1, foi subclonado em vetor de expressão denominado pBUf5, previamente construído a partir de pET28a, conforme esquematizado na **Figura 50B**. Em pBUf5 o cDNA codificando 116 resíduos do C-terminal da proteína D do capsídio do fago λ é encontrado entre os sítios Nco I e Bgl II, na mesma pauta de leitura que o segmento contendo a etiqueta de 6x(His) no N-terminal. Este fragmento da Proteína D corresponde a um domínio bastante compacto constituído basicamente de fitas- β . Ainda em pBUf5, entre os sítios Bgl II e Bam HI, é encontrado o sítio de reconhecimento pela protease TEV codificando a seqüência “ENLYFQGS” (sítio de clivagem sublinhado). A construção resultante da subclonagem de PepC-term entre os sítios Bam HI e Hind III codifica uma proteína de fusão com 179 resíduos (19,1 kDa).

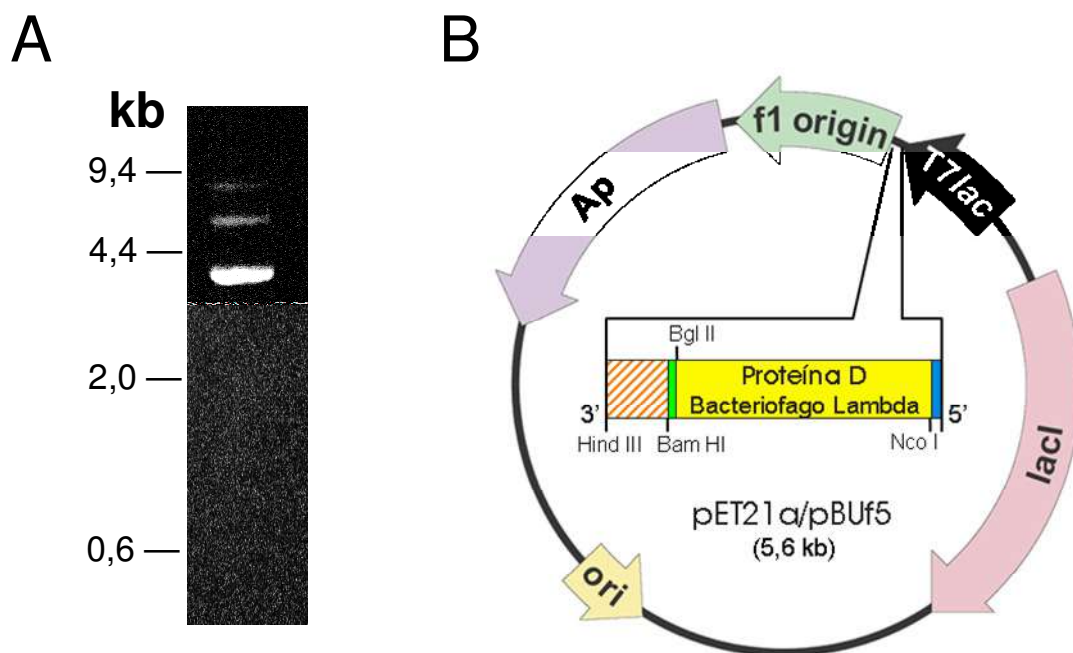


Fig. 50: Representação esquemática da construção do vetor de expressão pBUf5/PepC-term derivado de pET21a. A- Análise em gel de agarose (0,8%) do plasmídeo. B- Esquema da construção destacando os elementos de pET21a (sobre o círculo), como o gene do repressor lac (*lacI*), o promotor da T7 polimerase e operon lac (*T7lac*), além do gene da β -lactamase que codifica resistência à ampicilina (*Ap*). Dentro do círculo estão os elementos específicos de pBUf5, clonados na mesma pauta de leitura que o 6x(His) de pET21a (azul), incluindo o sítio da protease TEV (verde). O cDNA codificando os 45 resíduos da extensão C-terminal da forma longa de SmCD1 foi clonado entre os sítios de Bam HI e Hind III (rajado laranja).

O vetor pBUf5/PepC-term foi usado na transformação de células competentes de *E. coli* BL21(DE3). Um dos transformantes selecionados em placa de cultura contendo 100 mg/mL de ampicilina e confirmado por PCR de colônia, foi cultivado em tubo de ensaio a 37° C. Esta cultura foi usada como pré-inóculo no teste da indução da produção da fusão PtnD/pepC-term por 1 mM de IPTG a 37 °C. A análise por SDS-PAGE das frações solúveis e dos pellets (corpos de inclusão) de alíquotas recolhidas ao longo da indução é mostrada na **Figura 51**.

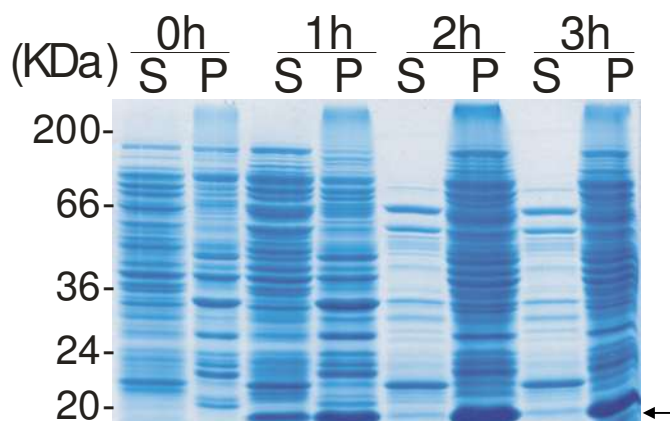


Fig. 51: Análise por SDS-PAGE (12%, redutora) das frações do extrato solúvel (S) e de corpos de inclusão (P) do teste de indução da expressão de PtnD_pepC-term por 1 mM de IPTG a 37 °C. A seta indica a banda de 19,1 kDa esperada para a proteína recombinante. Revelação por Coomassie brilliant blue G-25.

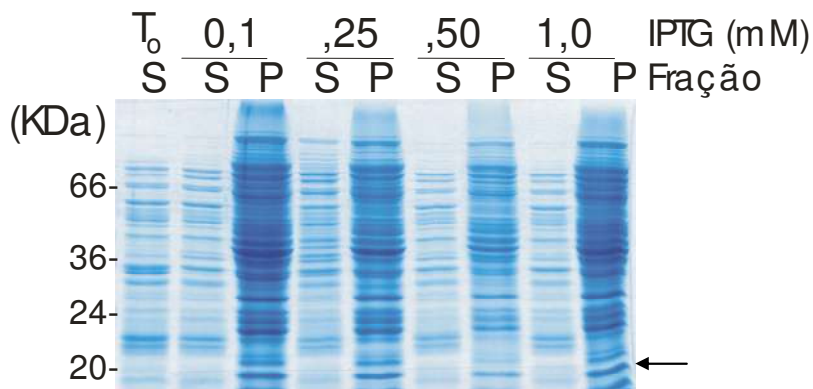
Os resultados do teste de indução mostram que a expressão da fusão PtnD_pepC-term é detectada em SDS-PAGE revelada por coomassie somente nas amostras onde houve adição de IPTG ao meio de cultura. Com 1h de indução observou-se o máximo de produção de PtnD_pepC-term na forma solúvel. Nos tempos de 2 e 3h de indução houve o esperado aumento na quantidade de proteína recombinante acumulada como corpos de inclusão, mas observou-se também uma inesperada redução na quantidade presente na fração solúvel. Esta observação é coerente com agregação causada pelo acúmulo da proteína superexpressa no citoplasma bacteriano.

4.6.2 Otimização da superprodução solúvel da fusão contendo pep_C-term em *E. coli* BL21(DE3)

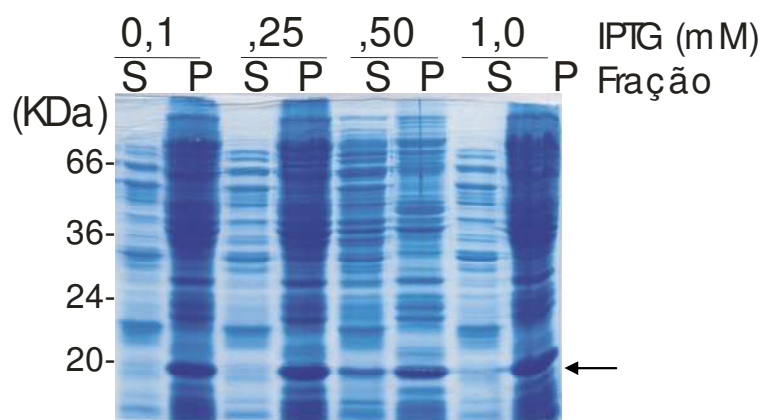
Os resultados obtidos com o teste de indução por IPTG a 37 °C apontaram a agregação da proteína recombinante como um problema em potencial para a sua superprodução na forma solúvel. Assim, procurou-se otimizar a expressão de PtnD_pepC-term na forma solúvel variando-se a temperatura e a concentração de agente indutor. Estes parâmetros controlam, respectivamente, a velocidade do metabolismo basal de *E. coli* e o nível de inibição do repressor lac (e conseqüentemente do nível de transcrição do gene clonado). Em relação a este último parâmetro, deve-se ressaltar que o vetor pET21a contendo o operador lac não é a melhor escolha para ajustar a expressão de uma proteína recombinante em função da concentração de agente indutor. Este sistema foi elaborado pensando-se apenas em ter uma forte repressão da expressão basal

(sem agente indutor) do gene clonado. A análise das alíquotas de 3h de indução no teste variando-se temperatura e concentração de IPTG é encontrada na **Figura 52**.

A



B



C

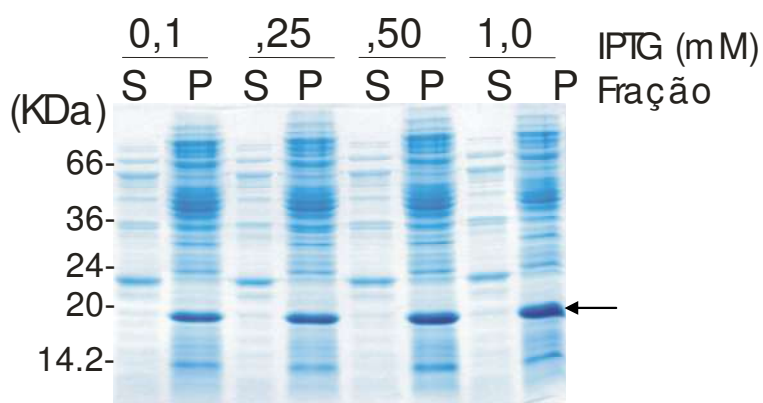


Fig. 52: Análise por SDS-PAGE (15%, redutora) das frações de 3h de indução da expressão de PtnD_pepC-term em *E. coli* BL21(DE3). A indução foi realizada nas temperaturas de 25 °C (A), 30 °C (B) e 37 °C (C) em diferentes concentrações de IPTG (0,1 – 1,0 mM). S = fração solúvel. P = corpos de inclusão. Revelação por Coomassie brilliant blue G-25.

Observou-se que a 25 °C, o nível de expressão da proteína recombinante é demasiadamente baixo para que a espécie de 19,1 kDa seja idubtavelmente detectada na fração solúvel. Na temperatura mais elevada testada (37 °C) também não foi possível detectar PtnD_pepC-term nas frações solúveis. Nestas temperaturas, assim como na temperatura intermediária de 30 °C, não há uma correlação clara entre a concentração de IPTG e a quantidade de proteína encontrada como corpos de inclusão. Este resultado não chega a ser uma surpresa, considerando o mecanismo de controle da transcrição para um gene clonado no sistema de expressão empregado. Entretanto, a 30 °C, observa-se uma destacada banda com cerca de 19 kDa na fração solúvel da cultura adicionada de 0,5 mM de IPTG. Pode-se supor que nesta combinação de temperatura e concentração de agente indutor tem-se uma relação ótima que maximiza a quantidade de proteína expressa, enquanto minimiza a agregação protéica que leva à formação de corpos de inclusão.

4.6.3 Desenvolvimento de uma estratégia de purificação de pepC-term

Conforme demonstrado na **Figura 4.50**, a fusão PtnD/pepC-term conta com uma etiqueta de 6x(His) no N-terminal com o intuito de viabilizar um método de purificação expediente por IMAC. Objetivando determinar as condições ideais para purificação de PtnD/pepC-term, uma nova expressão foi realizada a 30 °C utilizando 0,5 mM de IPTG na indução. Uma alíquota do extrato solúvel foi incubada sob agitação suave com a resina Ni-NTA superflow por 12-14h a 4°C. Tampões com concentrações crescentes de imidazol (5 – 200 mM) foram passados pela resina e as frações eluídas analisadas por SDS-PAGE (**Fig. 53**). A análise deste gel revelou a presença de uma banda contaminante com aproximadamente 16 kDa e mesma intensidade que a banda de 19,1 kDa esperada para PtnD/pepC-term. Outras espécies menores e menos intensas podem ser observadas em 14-15 kDa. Estas bandas co-eluem com PtnD/pepC-term a partir da concentração de 10 mM de imidazol nos tampões (**Fig. 53**, 5 e 6). Mais ainda, pode-se perceber a presença de contaminantes na faixa de 66 kDa que eluem até 50 mM de imidazol (**Fig. 53**, 6-11). Nesta última concentração de imidazol também elui um contaminante de 23-24 kDa (**Fig. 53**, 10 e 11). Além disso, PtnD/pepC-term, juntamente com a espécie de ~16 kDa ainda eluem da coluna com 200 mM de imidazol nos tampões (**Fig. 53**, 12 e 13).

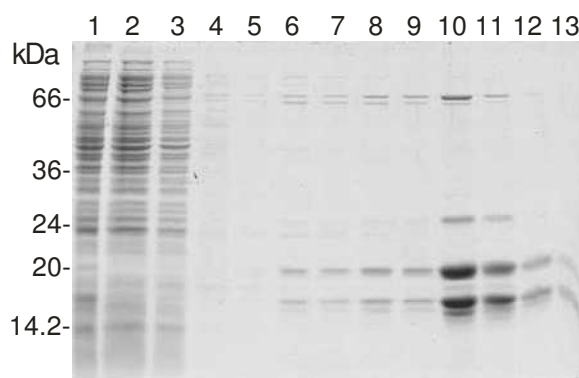


Fig. 53: Análise por SDS-PAGE (15%, redutora) das frações da cromatografia do extrato solúvel de *E. coli* BL21(DE3) expressando PtnD/pepC-term em resina Ni-NTA superflow. 1- Extrato solúvel aplicado; 2- não-ligado; 3-4= lavado com 5 mM imidazol; 5-6= eluído com 10 mM imidazol; 7-8= lavado com 15 mM imidazol; 9-11= eluído com 50 mM imidazol; 12-13 eluído com 200 mM imidazol. Revelação por Coomassie brilliant blue G-25.

Algumas informações para a otimização da purificação da proteína recombinantes puderam ser retiradas deste experimento. As condições do tampão utilizado na IMAC de PtnD/pepC-term podem ser consideradas bastante estridentes, com 300 mM de NaCl, 5% de glicerol e 5 mM imidazol. Assim, não foi surpreendente que a proteína recombinante tenha sido eluída da resina em concentrações de imidazol tão baixas quanto 10 mM. Desta forma, a concentração máxima de imidazol com que as proteínas ligadas inespecificamente à resina podem ser lavadas é 5 mM. Por outro lado, com 50 mM de imidazol, a maior parte de PtnD/pepC-term ligada pôde ser eluída da coluna. Uma eluição mais eficiente poderia ser conseguida com 100 mM de imidazol.

O comportamento observado para as espécies na faixa 14-16 kDa, co-eluído com PtnD/pepC-term em toda a faixa de concentração de imidazol, não é esperado para contaminantes ordinários. Além disso, a recuperação destas espécies, principalmente a de ~16 kDa, em tampões com concentrações superiores a 50 mM de imidazol sugere uma forte interação com o metal de transição imobilizado na resina (Ni^{2+}), tal como é proporcionada pela cauda de 6x(His) propositalmente inserida em proteínas recombinantes. Uma hipótese resultante desta análise foi que estas espécies pudessem ser sub-produtos de degradação proteolítica de PtnD/pepC-term. Mais ainda, os dados eram compatíveis com clivagens próximas ao C-terminal da proteína de fusão, pois o N-terminal contendo a cauda de 6x(His) teria que estar preservado.

A estratégia original de obtenção de pepC-term na forma pura consistia em purificar a proteína de fusão por IMAC e liberar o peptídeo da fusão através de

clivagem com a protease TEV, que em seguida seria obtido na fração não-ligada de uma 2ª IMAC. Esta estratégia baseava-se no pressuposto que os possíveis contaminantes ligados na 1ª IMAC não interfeririam com a liberação e purificação de pepC-term, uma vez que estes, juntamente com a protease TEV (que também contém uma cauda 6x(His)) também ligariam na 2ª IMAC, deixando pepC-term praticamente puro na fração não-ligada. Além disso, contando que PtnD/pepC-term seria obtida com elevado nível de pureza na 1ª IMAC e com a alta especificidade da protease TEV, a geração de peptídeos indesejados a partir da digestão de contaminantes por esta protease seria minimizada. Entretanto, com a hipótese de degradação proteolítica de pepC-term em mente, problemas em potencial eram vislumbrados para a purificação de pepC-term através desta estratégia. A fim de verificar se as clivagens na proteína de fusão ocorriam no segmento contendo pepC-term foi realizado um teste de clivagem com protease TEV. As frações eluídas de Ni-NTA superflow foram reunidas, alíquotas incubadas a 30 °C por cerca de 16h com quantidades crescentes de protease TEV e as amostras analisadas por SDS-PAGE (**Fig. 54**).

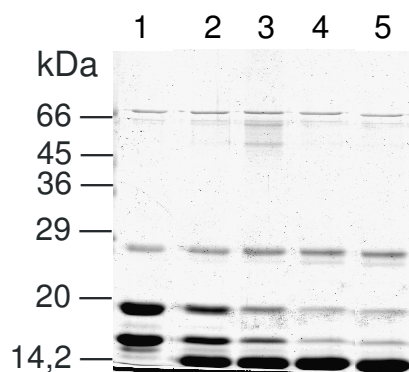


Fig. 54: Análise por SDS-PAGE (12%, redutora) das amostras do teste de clivagem com protease TEV. 1- controle do material ligado à resina Ni-NTA superflow incubada na ausência da protease TEV; 2 a 5- material em 1 incubado com 0,1, 0,2, 0,5 e 1,0 unidade de protease TEV. Revelação por Coomassie brilliant blue G-25.

Como pode ser percebido na **Figura 54**, PtnD/pepC-term juntamente com as espécies migrando entre 15-17 kDa foram todas convertidas em uma espécie de 14 kDa, a qual é compatível com a massa do fragmento da Proteína D do fago λ . Este resultado corrobora a hipótese de que as espécies menores que 19,1 kDa ligadas especificamente à resina Ni-NTA superflow são sub-produtos de degradação proteolítica de PtnD/pepC-term. Além disso, a ausência de bandas inferiores à 14 kDa indica que estes sub-produtos são todos gerados por clivagens na região da proteína de fusão correspondendo à pepC-term.

Estes resultados levaram à uma re-análise das amostras do teste de indução que revelou a presença do sub-produto majoritário da proteólise de PtnD/pepC-term (16,1 kDa) em todos os extratos das culturas induzidas em 25-37 °C e com concentrações de IPTG variando de 0,1-1,0 mM. A experiência adquirida com preparações sucessivas de extratos de culturas a 30 °C e induzidas com 0,5 mM de IPTG, mostrou que a proteólise da proteína de fusão ocorria independentemente da presença de inibidores de protease (1 mM PMSF e 1 mM EDTA) nos tampões e podia ser minimizada reduzindo o tempo entre a preparação do extrato e a cromatografia em Ni-NTA superflow. Estas observações indicam que tal degradação proteolítica de PtnD/pepC-term ocorre durante a expressão e que a(s) protease(s) responsável(eis) pode(m) ser eliminada(s) pela IMAC. A extensa degradação no segmento correspondendo à pepC-term também sugere que este peptídeo apresentava-se em sua maior parte desestruturado nas condições do citossol de *E. coli*.

Com a presença de contaminantes tão relacionados estruturalmente à proteína recombinante de interesse, uma nova estratégia de purificação teria que ser vislumbrada. Uma dentre duas alternativas teria que ser escolhida: purificar a proteína de fusão ou purificar pepC-term dentre os peptídeos liberados pela digestão com protease TEV. Optou-se originalmente pela primeira, em virtude da maior disponibilidade de métodos para purificar proteínas do que peptídeos. Apesar da sabida diferença de massa entre PtnD/pepC-term e seus sub-produtos de proteólise, cromatografia de exclusão molecular não seria uma estratégia eficiente devido à pequena diferença entre os pesos moleculares observados em SDS-PAGE. Assim, optou-se por cromatografia de troca iônica como método de purificação de PtnD/pepC-term a partir do material eluído de Ni-NTA superflow. Uma tentativa inicial foi feita com a coluna catiônica High-Q acoplada a um sistema HPLC. A análise das frações coletadas ao longo da cromatografia revelou que PtnD/pepC-term, assim como seus sub-produtos proteolíticos foram incapazes de seu ligar ao grupo quaternário de amônio da resina (**Fig. 55**). Apenas o contaminante de 23-24 kDa pôde ser purificado com este método. Estes resultados não eram esperados, uma vez que o pI predito pela seqüência da proteína recombinante (pI = 6,92) indica que a proteína estaria com carga negativa no tampão da cromatografia. De fato, 4 resíduos ácidos (Asp + Glu) estariam ionizados em pH 7,5.

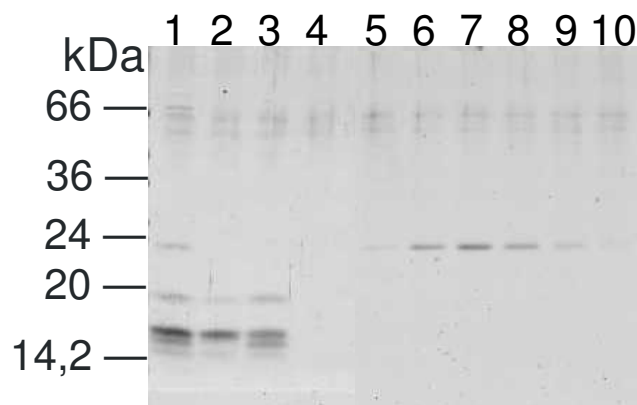


Fig. 55: Análise por SDS-PAGE (15%, redutora) das frações da cromatografia de PtnD/pepC-term na coluna Econopac High-Q. 1- aplicado; 2- não-ligado em Tris-HCl pH 7,5 (tampão A); 3 e 4-lavado em tampão A; 5 a 10- frações 23-28 eluídas durante o gradiente de 50-100% de B (tampão A + 1 M de NaCl). Revelação por Coomassie brilliant blue G-25.

Como PtnD/pepC-term e seus derivados de degradação proteolítica não foram capazes de interagir com a resina catiônica, optou-se por reunir o material não ligado, dialisar contra tampão 50 mM citrato de sódio pH 4,9 (tampão A) e submeter à cromatografia na coluna Resource S acoplada a um sistema HPLC. Conforme esperado após os resultados com a resina High-Q, a proteína de fusão e seus subprodutos foram eluídas durante o gradiente salino com mais de 500 mM de NaCl, indicando uma forte interação destas proteínas com o grupamento aniônico da coluna Resource S (**Fig. 56**). Baseando-se neste resultado, planejou-se nova purificação do extrato solúvel de *E. coli* BL21(DE3) expressando PtnD/pepC-term, usando tampão de eluição da resina Ni-NTA superflow compatível com a ligação na coluna Resource S. O tampão em pH 4,9 e contendo 300 mM de NaCl, além de 100 mM imidazol não interferiria com a eluição — pelo contrário — e permitiria a cromatografia de troca iônica sem a necessidade de um passo adicional de diálise. Entretanto, não foi isso que foi observado na análise das frações eluídas da coluna Resource S (**Fig. 56**). Inesperadamente, ptnD/pepC-term e seus subprodutos de proteólise foram coletados nas frações do lavado da coluna.

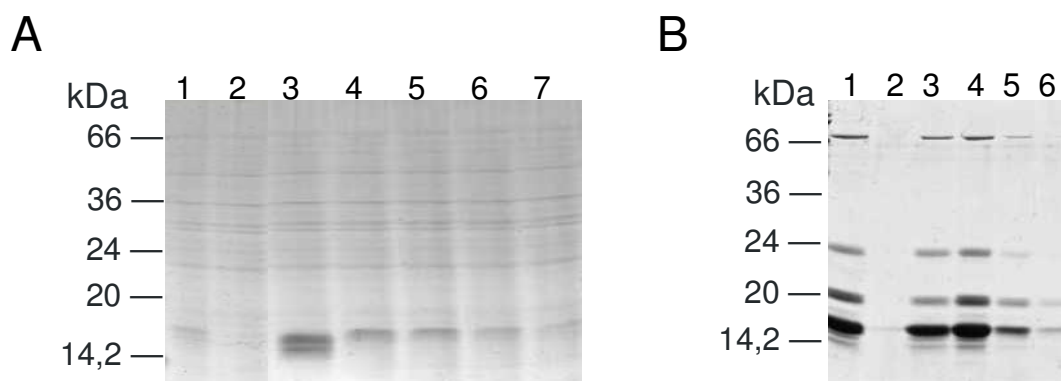


Fig. 56: Análise por SDS-PAGE (12%, redutora) das amostras de PtnD/pepC-term cromatografadas em coluna Resource S acoplada a sistema de HPLC no 1° teste (A) e 2° teste (B). (A) 1-aplicado em tampão A (50 mM citrato de sódio pH 4,9); 2- não-ligado; 3 a 7- frações 7-12 eluídas durante o gradiente de 50-100 B (tampão A + 1M de NaCl). (B) 1- aplicado em tampão A + 300 mM NaCl + 100 mM imidazol; 2- não-ligado; 3 a 6- amostras do lavado. Revelação por Coomassie brilliant blue G-25.

O insucesso com a troca iônica levou a uma mudança de estratégia. Nova batelada da proteína de fusão foi preparada e após diálise contra tampão Tris-HCl pH 8 para retirada do imidazol, o peptídeo pepC-term foi liberado pela digestão com protease TEV. A amostra foi submetida à nova IMAC em resina Ni-NTA superflow e conforme demonstrado na **Figura 57**, a fração não-ligada contendo pepC-term foi obtida praticamente livre da proteína D e de outros contaminantes de mais alto peso molecular.

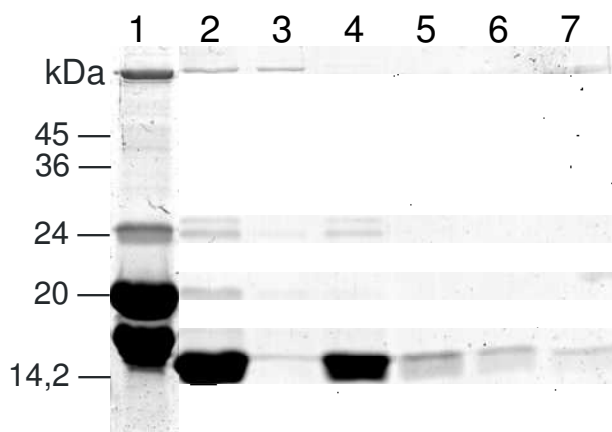


Fig. 57: Análise por SDS-PAGE (15%, redutora) das frações da 2ª IMAC para isolamento de pepC-term a partir da proteína de fusão digerida com a protease TEV. 1- preparação de PtnD/pepC-term antes da digestão com protease TEV; 2- Amostra digerida, aplicada à coluna Ni-NTA superflow; 3- Não-ligado; 4 a 7- frações (1 mL) eluídas com 200 mM de imidazol. Revelação por Coomassie brilliant blue G-25.

A fim de confirmar a presença de pepC-term no não-ligado da 2ª IMAC, esta amostra foi dializada contra tampão volátil (20 mM $\text{NH}_4(\text{HCO}_3)$ pH 8) em uma membrana de 3 kDa de limite de poro. A seguir a amostra foi dividida em duas

alíquotas: a primeira foi evaporada até a secura e a segunda submetida à hidólise com tripsina para então serem analisadas por espectrometria de massas MALDI-TOF (**Fig. 58 e 59**). O espectro da amostra intacta na **Figura 58** evidencia um íon de $m/z = 5149,3$ que poderia corresponder a pepC-term (massa média predita pela seqüência de 5145,8) com alguma modificação em cadeias laterais de resíduos presentes em sua composição como arginina ou metionina. Entretanto, além da abundância de contaminantes possivelmente minoritários em quase toda faixa do espectro, um íon base com $m/z = 2112,9$ chamou atenção pela intensidade.

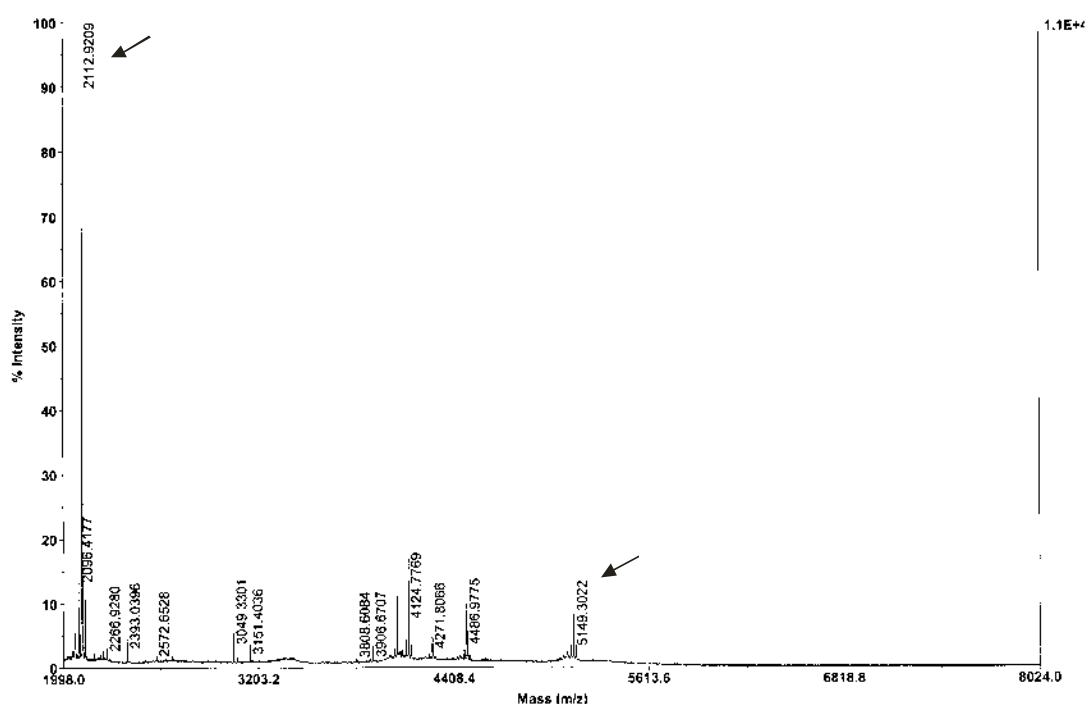


Fig. 58: Análise por espectrometria de massas (MALDI-TOF) da fração não-ligada na 2^a IMAC. Os íons relevantes estão assinalados com setas.

A presença de pepC-term na amostra do não-ligado da 2^a IMAC foi confirmada pela presença dos peptídeos de massa esperada após hidrólise desta amostra com tripsina (**Fig. 59**). A identificação dos íons com m/z 2394.2 e 991.5 compatíveis com os peptídeos trípticos “LFPAQSVPQAASET PNGVFAFSK” e “GSHPDSVHR” ($\pm 1 m/z$), corresponde a uma cobertura de 61% da seqüência de pepC-term.

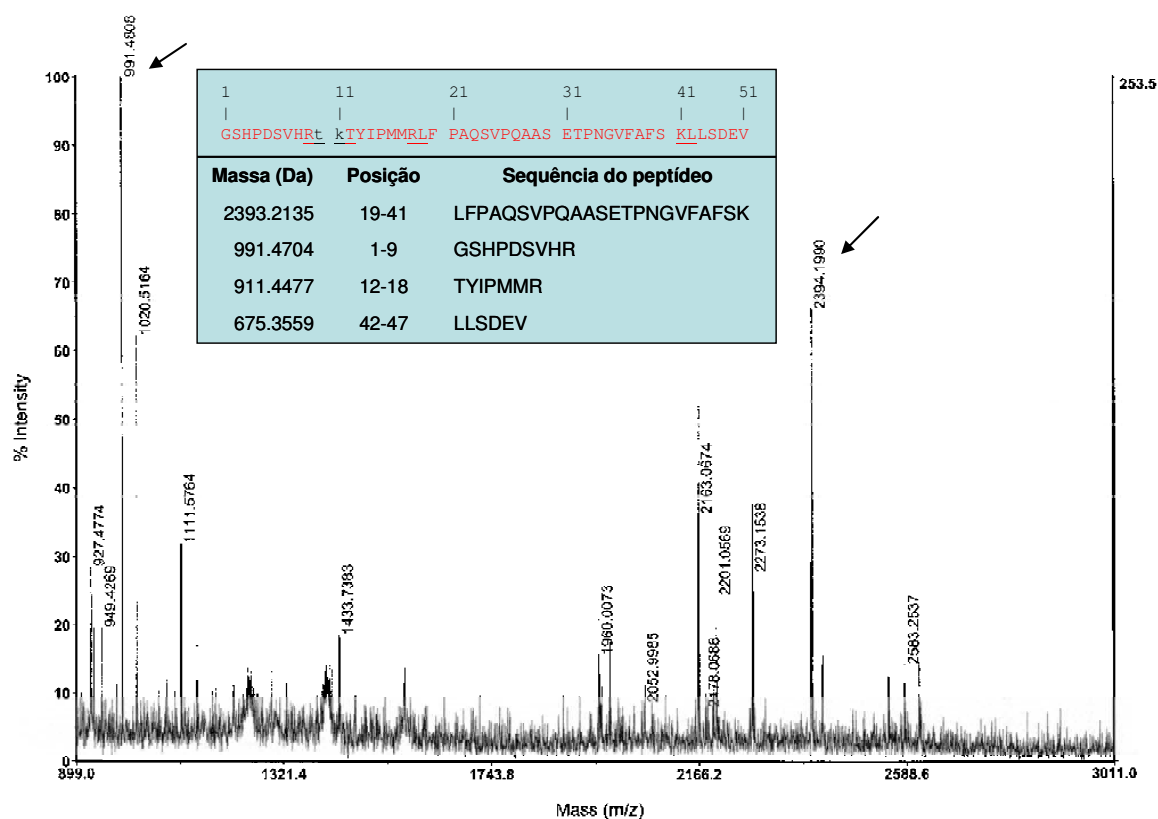


Fig. 59: Análise por espectrometria de massas (MALDI-TOF) da fração não-ligada na 2ª IMAC submetida à hidrólise com tripsina. Os íons relevantes estão assinalados com setas. O quadro em azul mostra a massa e composição dos peptídeos trípticos liberados, conforme previsto a partir da seqüência pepC-term (na parte superior). Os sítios de clivagem aparecem sublinhados, a porção da seqüência representada pelos peptídeos trípticos > 500 Da aparece em letras vermelhas e maiúsculas.

O resultado da análise por espectrometria de massas deixou clara a necessidade de purificação de pepC-term liberado da proteína de fusão. Como primeira alternativa para a purificação, utilizou-se cromatografia de fase reversa (C18) com uma coluna Shimpack VP-ODS acoplada a um sistema de HPLC. A amostra foi passada em um filtro centricon-10 para eliminação de possíveis contaminantes de alto peso residuais e aplicada à coluna. Um gradiente de 0-80% de B (AcCN + 0,1% TFA) foi aplicado, obtendo-se o perfil de eluição (220 nm) mostrado na **Figura 60**. Todos os componentes da amostra são eluídos antes de retenção equivalente a um volume de coluna, indicando que praticamente não houve interação com a matriz hidrofóbica. Apesar deste comportamento frente à fase reversa não ter sido antecipado, ele pode ser explicado pelo alto conteúdo de grupos passíveis de protonação em 0,1% TFA encontrados na seqüência de pepC-term (4 Arg/Lys + 2 His). Entretanto, tentativa de cromatografia na ausência de TFA não mudou este perfil (dados não mostrados).

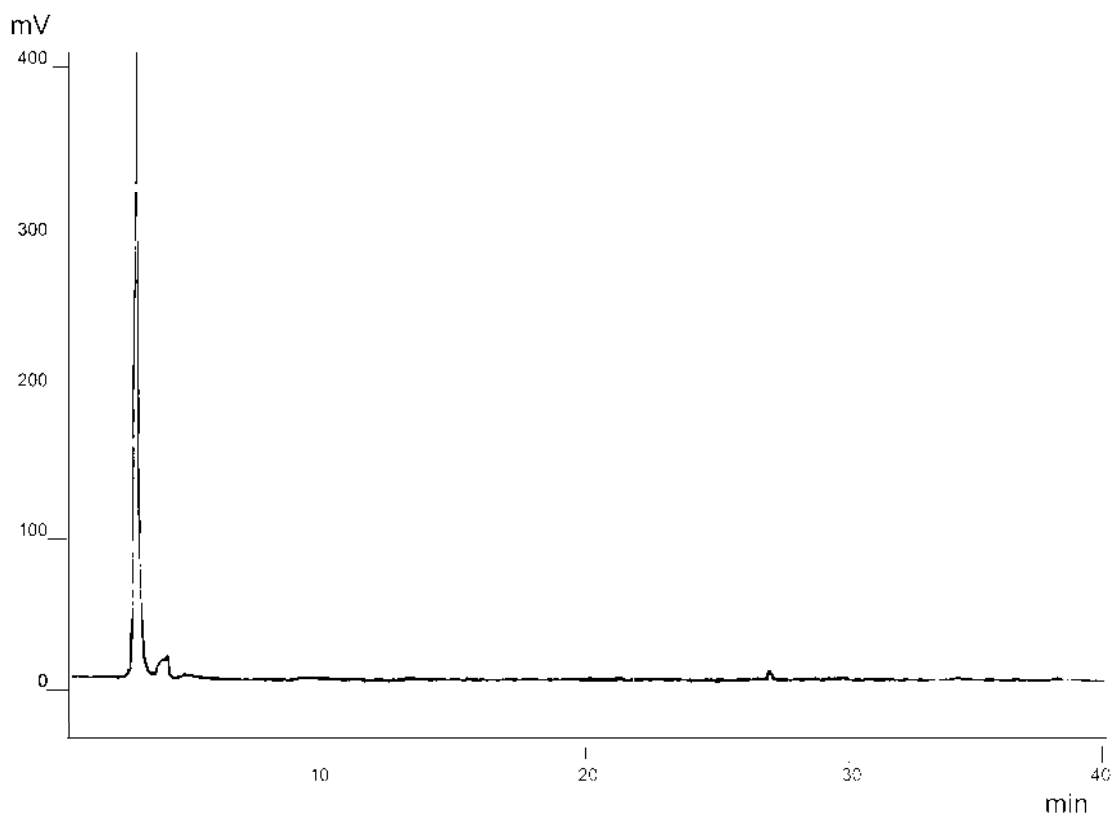


Fig. 60: Cromatografia em fase reversa (coluna C18 VP-ODS) da amostra de pepC-term (não-ligado da 2^a IMAC). A cromatografia foi desenvolvida em 0,1% TFA com um gradiente de 0-80% de ACN de 5-40min e acompanhada pela D.O.²²⁰.

Como segunda e última alternativa disponível para a purificação de pepC-term, empregou-se a cromatografia de exclusão molecular (SEC) em coluna Superdex Peptide HR 10/30. Um teste preliminar foi realizado onde a SEC foi desenvolvida no tampão recomendado pelo fornecedor da coluna (Tris-HCl pH 7.5 + 200 mM NaCl), fornecendo resultados promissores (dados não mostrados). Uma nova batelada de proteína recombinante foi preparada e pepC-term liberado da proteína de fusão como descrito acima. A seguir, tendo em vista as possíveis complicações relacionadas à troca de tampão e concentração de um peptídeo incapaz de ligar em resina C18, como pepC-term, optou-se por realizar a SEC em tampão volátil (150 mM (NH₄)HCO₃ pH 8), obtendo-se o perfil mostrado na **Figura 61**. Cinco picos distintos puderam ser coletados manualmente, todos com absorção a 280 nm, compatível com a presença de um resíduo de tirosina, como encontrado em pepC-term.

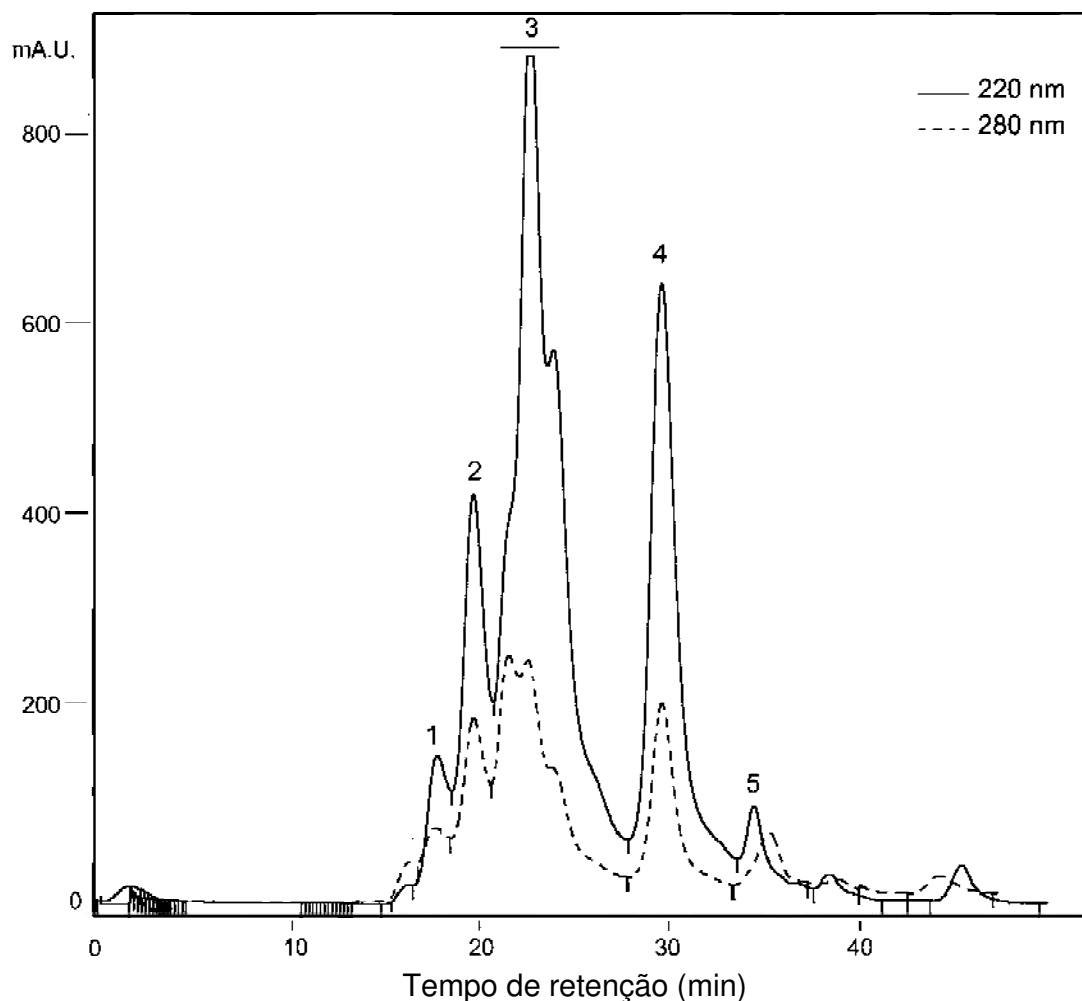


Fig. 61: Cromatografia de exclusão molecular (coluna Superdex Peptide HR 10/30) da preparação de pepC-term. A ordem dos picos coletados para análise posterior está indicada no cromatograma.

A construção de uma curva de calibração (não mostrada) usando como padrões citocromo C (14 kDa), aprotinina (6,5 kDa) e vitamina B12 (1,3 kDa) indicou que a espécie correspondendo ao pico 4 seria compatível com a massa de 5,1 kDa esperada para pepC-term. Assim, o pico 5 poderia corresponder a um dos subprodutos de degradação proteolítica de pepC-term, enquanto que os picos 1, 2 e 3 conteriam contaminantes maiores que 14 kDa (acima do limite de exclusão de 20 kDa da coluna). Esta interpretação era coerente com os dados disponíveis, restando apenas uma confirmação por espectrometria de massas dos picos, como descrito no item a seguir.

4.6.4 Caracterização de um subproduto de degradação proteolítica de pep_C-term por espectrometria de massas

Tendo em vista a massa predita pelo tempo de retenção na SEC, o pico 4 foi selecionado para análise acreditando-se se tratar de pepC-term com 5,1 kDa. O espectro

de MALDI-TOF desta amostra logo desmentiu este pensamento, mostrando um pico base de $m/z = 2113,1$ (**Fig. 62**). Uma espécie com m/z similar foi anteriormente detectada como o íon mais abundante no espectro da fração não-ligada na 2ª IMAC (**Fig. 58**). Mais ainda, um peptídeo com esta massa seria compatível com a diferença nas massas do subproduto de proteólise mais abundante de pepC-term (~16 kDa, observado na SDS-PAGE da fração ligada à 1ª IMAC) em relação à Proteína D (14 kDa). Com o intuito de verificar se esta espécie realmente correspondia a um fragmento de pepC-term e de caracterizar a sua estrutura, nova hidrólise trípica foi realizada. O espectro de MALDI-TOF na **Figura 63A** mostra os íons mais intensos com $m/z = 991,48$ e $m/z = 911,45$. Conforme indicado no espectro detalhado da região com $m/z < 1000$ (**Fig. 63B**), estes valores correspondem às massas de dois dos peptídeos trípticos liberados a partir de pepC-term, com um erro inferior a 0,01.

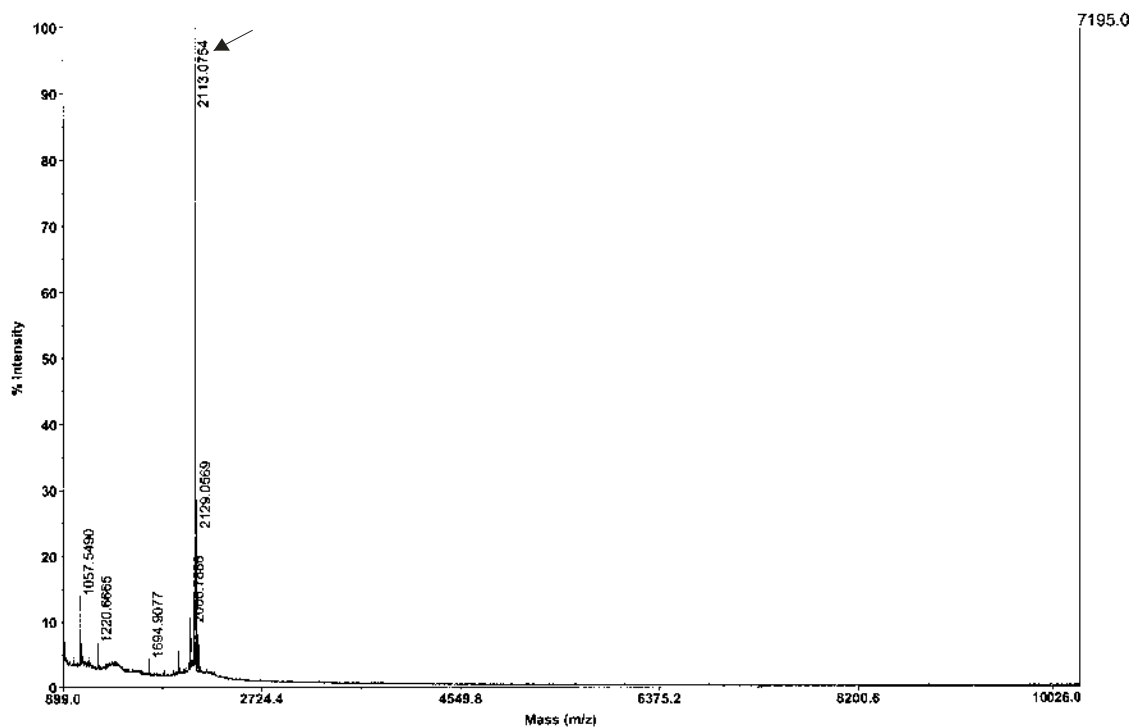
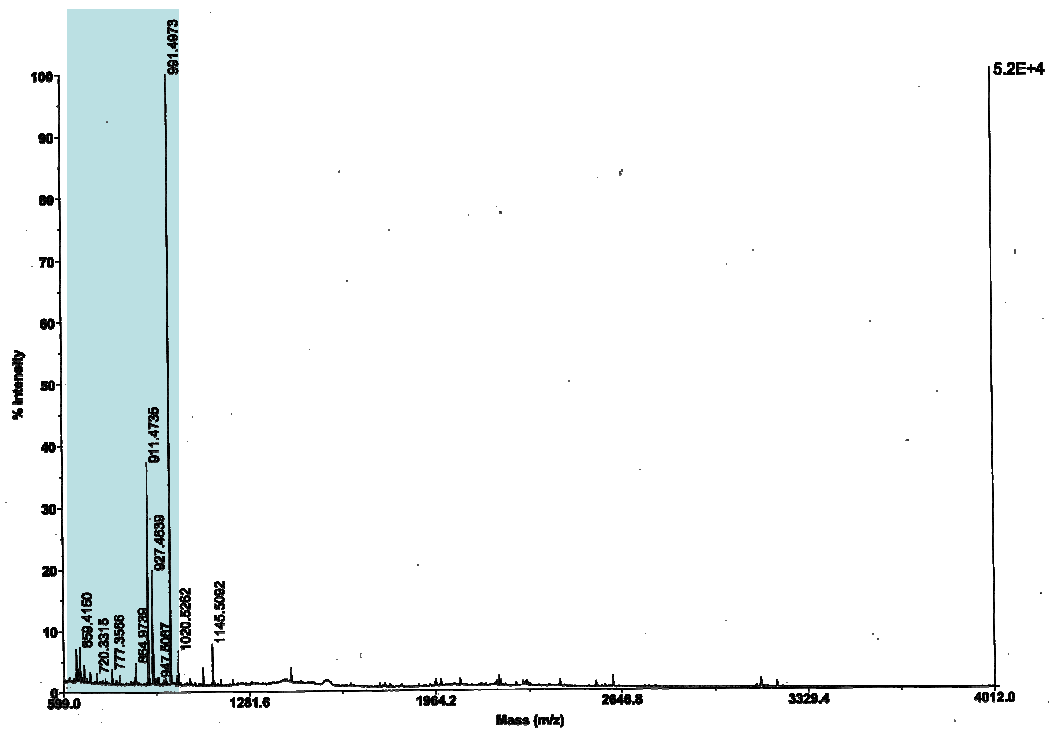


Fig. 62: Análise por espectrometria de massas (MALDI-TOF) do pico 4 da SEC. O íon base está indicado com a seta.

A



B

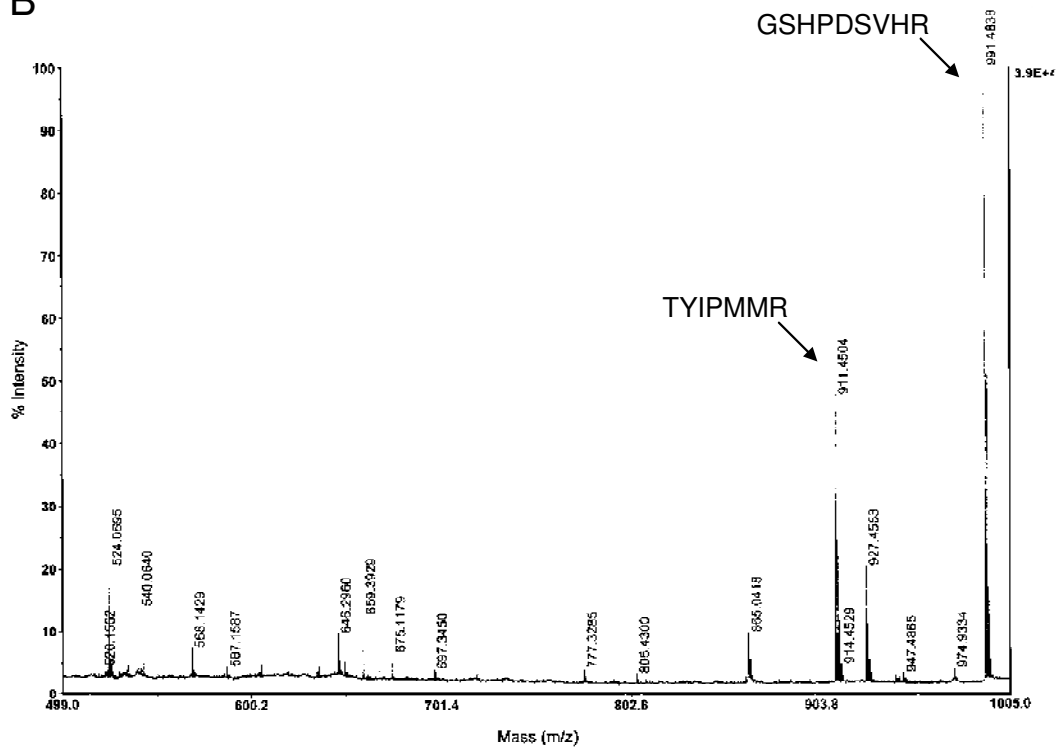


Fig. 63: Análise por espectrometria de massas (MALDI-TOF) do pico 4 digerido com tripsina. A faixa de massa (m/z) hachurada em A foi focalizada no espectro em B. Os íons com m/z correspondendo às massas dos peptídeos trópticos de pepC-term estão indicados.

Os peptídeos GSHPDSVHR e TYIPMMR correspondem aos fragmentos 1-9 e 12-18 de pepC-term. Este resultado levou a uma procura na seqüência de pepC-term por fragmentos que pudessem corresponder à massa de 2113 Da encontrada para o peptídeo presente no pico 4 da SEC. Para isto utilizou-se o programa FindPept que considera possíveis modificações químicas (oxidação, hidrólise, etc.) e pós-traducionais (e.g., glicosilação). A saída deste programa retornou o fragmento N-terminal 1-18, GSHPDSVHRTKTYIPMMR, com massa $[(M+H)^+] = 2113,44$. Desta forma, a espécie isolada do pico 4 da SEC foi identificada como um subproduto proteolítico de pepC-term, correspondendo aos primeiros 18 resíduos a contar do N-terminal (18N_pepC-term). Este resultado é condizente com os dados bioquímicos mostrados anteriormente, incluindo a presença do resíduo de tirosina, responsável pela absorção da radiação à 280 nm ($\epsilon = 1280 \text{ M}^{-1} \cdot \text{cm}^{-1}$) e elevada hidrofobicidade ($pI = 10$), responsável pela pobre interação com a resina C18.

4.6.5 Estudos estruturais em pepC-term

4.6.5.a Métodos teóricos

Conforme exposto na seção introdutória deste capítulo, há um grande interesse na determinação da estrutura tridimensional da extensão C-terminal encontrada na forma longa de SmCD1. Na ausência de estruturas homólogas que possam servir como moldes em exercícios de modelagem por homologia, duas estratégias foram empregadas na predição da conformação espacial de pepC-term: modelagem *ab initio* e reconhecimento do enovelamento. Dentro da primeira estratégia, empregou-se o servidor SCRATCH para obter um modelo de baixa resolução do tipo “colar de pérolas”, onde somente a posição dos carbonos- α é predita (**Fig. 64**). Não há como validar este modelo de “baixa resolução”, senão contrapondo-o a predições realizadas por outros métodos, conforme exposto adiante.

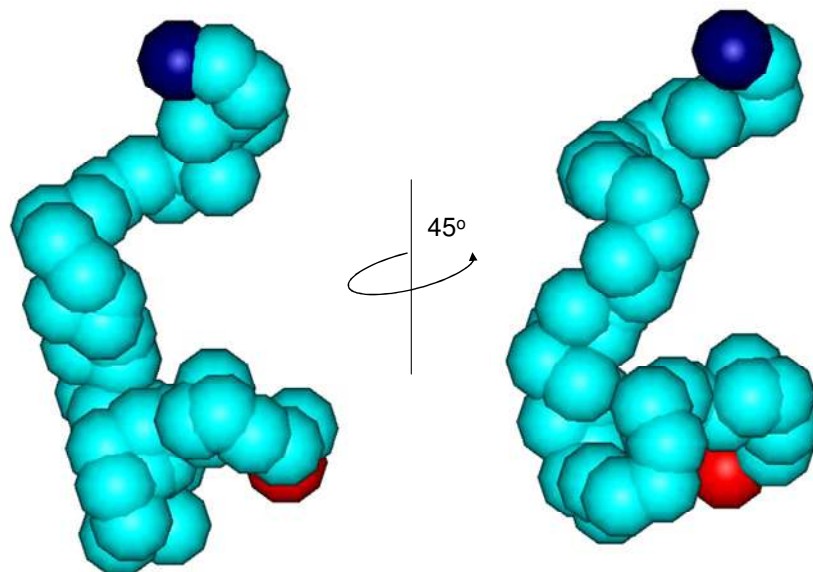


Fig. 64: Modelo ab initio de pepC-term predito pelo programa SCRATCH. Para orientação, o N-terminal é marcado em azul escuro e C-terminal em vermelho.

Um componente importante de ambas estratégias de predição mencionadas anteriormente, é a predição da estrutura secundária. Assim, uma bateria de programas disponíveis (6 ao todo) na forma de servidores WWW foi empregada e uma predição “consenso” da estrutura secundária de pepC-term gerada manualmente pela inspeção da concordância entre as predições individuais (**Fig. 65**). Esta predição parsimoniosa da estrutura secundária de pepC-term indica um conteúdo de pelo menos 14% de fitas- β e 14% de α -hélices. Os elementos de estrutura secundária aparecem concentrados no N- e C-terminal de pepC-term em associações do tipo fita-alça-hélice. Deve-se notar entretanto, que apesar de muitos programas atuais de predição de estrutura secundária conseguirem níveis de acerto na faixa de 60-80% dos resíduos analisados, a confiança nas predições de pepC-term são grandemente prejudicadas pela ausência de homologia clara com outras seqüências em bancos de dados. Isto porque programas como J-Pred, PHD e PSI-Pred têm como componente fundamental de seus algoritmos a construção de perfis de seqüência como os gerados pelo algoritmo de PSI-BLAST. De fato, uma busca nos bancos de dados de proteína tr-EMBL e SwissProt foi realizada com este último programa e não foi capaz de identificar homólogos distantes de pepC-term.

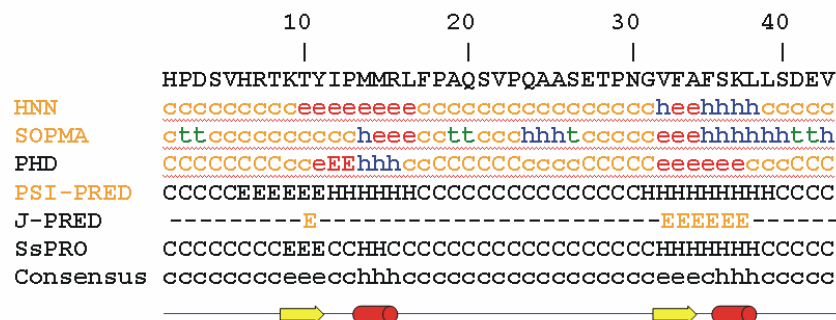


Fig. 65: Predição de estrutura secundária de pepC-term. Legenda: c = aleatória, e = fita- β , h = α -hélice, t = alça- β . Na linha consenso, a seta indica fita- β e o cilindro α -hélice.

Programas de predição *ab initio* como SCRATCH ainda possuem estatus de “em desenvolvimento”. Assim, a metodologia de “reconhecimento de enovelamento”, melhor estabelecida para a predição de estruturas de proteínas-alvo com similaridade a proteínas de estrutura conhecida na faixa de 20-30%, foi empregada. Dois servidores foram utilizados: 3D-PSSM e GenTHREADER. Em ambos os servidores, possíveis moldes foram encontrados na faixa de 15-25% de identidade com pepC-term, porém todos com significado estatístico duvidoso. O modelo mais “significativo” retornado por cada programa, foi avaliado. GenTHREADER forneceu um modelo baseado no segmento 70-110 de uma proteína identificada como “alérgeno de poeira caseira” (código PDB 1A9V). A estrutura desta proteína, resolvida por RMN, consiste num barril de folhas- β , contradizendo as predições de estrutura secundária para pepC-term. Além disso, uma inspeção mais cuidadosa da estrutura revelou que as fitas- β preditas por este modelo para pepC-term não fazem contato entre si, mas sim com segmentos distais da proteína. Assim, as interações (principalmente ligações de hidrogênio) estabilizando esta estrutura na proteína original provavelmente estariam ausentes na estrutura de pepC-term. Por estes motivos o modelo predito pelo programa GenTHREADER foi descartado.

Diferentemente do programa GenTHREADER, o resultado do programa 3D-PSSM foi condizente com as predições feitas anteriormente por outros métodos para a estrutura de pepC-term. O modelo mais “significativo” retornado pelo programa baseava-se numa estrutura de apenas 47 resíduos, correspondendo a um “dedo de zinco clássico” (código PDB 1NCS). Entretanto, em função da baixa similaridade na seqüência (15% de identidade) somente o fragmento 19-46 de pepC-term pôde ser modelado. O modelo de baixa resolução obtido é ilustrado na **Figura 66**. Fica evidente

por este modelo a presença de dois segmentos em α -hélice e um “grampo- β ” na parte central da estrutura.

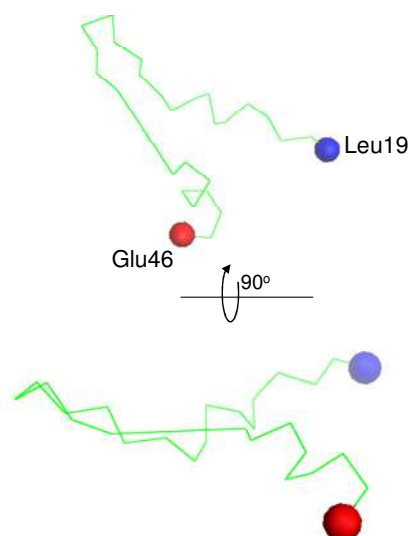


Fig. 66: Modelo do segmento 19-46 de pepC-term predito pelo programa 3D-PSSM. O N-terminal está assinalado em azul e o C-terminal em vermelho.

A comparação do modelo *ab initio* da **Figura 64** e o modelo predito pela técnica de por “reconhecimento de enovelamento” da **Figura 66**, revela alguma semelhança entre as estruturas. Este tipo de concordância entre métodos baseados em diferentes premissas teóricas, é um bom indicativo da validade dos modelos gerados.

4.6.5.b Avaliação do grau de estruturação de 18N_pepC-term em solução por CD

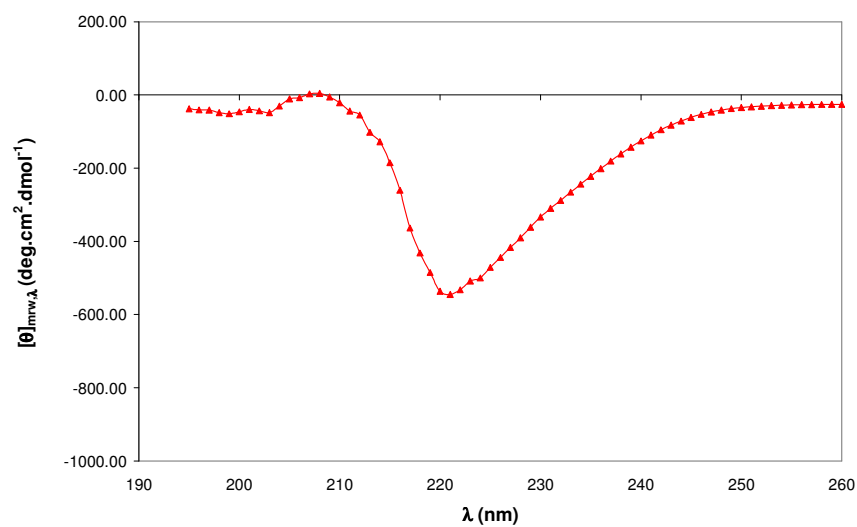
Os resultados dos métodos computacionais relatados acima descrevem diferentes níveis de estruturação de pepC-term e fornecem indícios da estrutura que poderia ser assumida por este segmento de SmCD1. Um dos objetivos do presente trabalho era expressar a extensão C-terminal da forma longa de SmCD1 a fim de permitir estudos estruturais. Dificuldades encontradas em meio a este processo inviabilizaram a obtenção deste peptídeo de 47 resíduos na forma pura e em quantidades suficientes para estudos por CD ou RMN. Entretanto, estes requisitos puderam ser obtidos com 18N_pepC-term, um fragmento de 18 resíduos do N-terminal de pepC-term. Em função da completa falta de dados estruturais sobre este segmento, uma caracterização inicial de sua estrutura em solução foi realizada por CD.

A estrutura de 18N_pepC-term foi primeiramente avaliada em água pura, fornecendo o espectro de CD da **Figura 67A**. É possível perceber claramente a banda

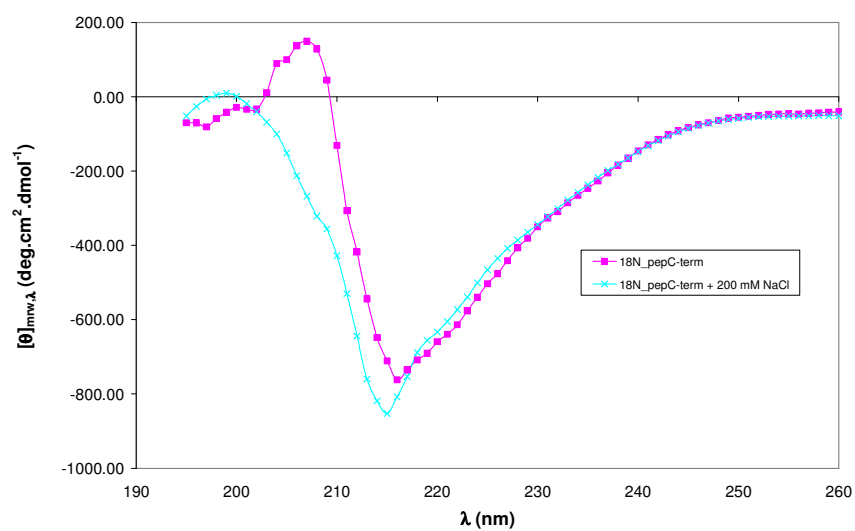
de absorção em 220 nm, característica da cadeia polipeptídica na conformação de α -hélice. A deconvolução do espectro forneceu os seguintes percentuais de estrutura secundária: 7% de α -hélice, 47% de fita- β e 18% de alça- β . Dispondo deste resultado indicando um razoável grau de estruturação do peptídeo, resolveu-se avaliar a conformação deste em pH 3,5. Isto é relevante, pois 18N_pepC-term e seu precursor pepC-term são parte integrante de uma enzima ativa neste pH. O espectro de CD obtido é mostrado pela curva em magenta na **Figura 67B**. Interessantemente, pode-se perceber uma elevação no máximo de $[\theta]_{\text{mrw}}$ indicando uma maior estruturação do peptídeo. Além disso, pode-se perceber um deslocamento do pico de absorção de 220 nm para 215-217 nm. Isto pode indicar um aumento na fração da estrutura em conformação de fita- β . O efeito da adição de 200 mM de NaCl foi testada na estrutura de 18N_pepC-term em pH 3,5. A curva azul clara da **Figura 67B**, mostra um sinal de CD ainda mais intenso para o peptídeo. Também pode-se perceber deslocamento adicional da banda de maior absorção para o UV distante (210 nm). Concentrações baixas de sal costumam ter um efeito estabilizante na estrutura de muitas proteínas. Esta estabilização é em grande parte ocasionada pelo aumento da força iônica, o que torna as interações eletrostáticas menos intensas entre diferentes partes da estrutura protéica.

Os resultados em água pura e tampão pH 3,5 indicaram que a estrutura de 18N_pepC-term é sujeita à uma forte influência do meio. Assim, uma nova série de experimentos foi feita em meio tamponado em pH 7,5, estudando-se novamente a influência da força iônica e pela primeira vez o efeito de um detergente. Os espectros de CD obtidos estão ilustrados na **Figura 67C**. Na ausência de sal (curva em magenta), o espectro assemelha-se ao obtido anteriormente em pH 3,5, porém com uma $[\theta]_{\text{mrw}}$ máxima ligeiramente menor e uma absorção mais evidente em 220 nm. A adição de 200 mM de NaCl provocou o deslocamento visto anteriormente para o UV distante (curva em amarelo) e uma alteração significativa na forma da curva. A adição do detergente aniônico SDS na concentração de 0,5% (acima da CMC) não resultou na perda de estrutura esperada (curva em azul claro). Houve alteração considerável do espectro, provocando um alargamento da banda de absorção máxima em 210-220 nm.

A



B



C

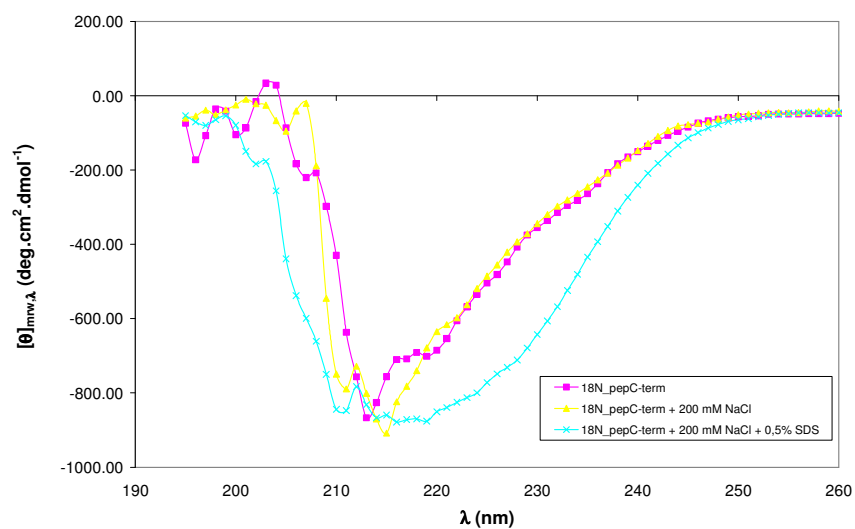


Fig. 67: Espectros de CD de 18N_pepC-term realizados em água pura (A), tampão NaOAc pH 3,5 (B) e tampão fosfato pH 7,5 (C).

4.6.6 Título e especificidade do anti-soro policlonal contra 18N_pep_C-term

Com a disponibilidade de 18N_pepC-term em quantidades apreciáveis e na forma solúvel, procurou-se obter um soro policlonal contra este peptídeo que pudesse distinguir a forma longa da forma curta de SmCD1. Para isto uma amostra purificada de 18N_pepC-term foi inoculada em coelhos e o título e especificidade do soro destes animais avaliado por *Immunoblot* (**Fig. 68**). Um controle com soro anti-pro-SmCD1.2 dil. 1:100 foi realizado, mostrando a forte reação esperada contra ambas as formas de proSmCD1 e nenhuma reação contra PtnD/pepC-term (**Fig. 68**, painel A). O soro anti-18N_pepC-term reagente fortemente contra PtnD/pepC-term e seu subproduto de proteólise em 16,1 kDa no título de 1:100 (**Fig. 68**, painel B). Nesta diluição do soro, não houve reação alguma contra o componente Proteína D da proteína de fusão e fraca reação contra a forma longa e, inesperadamente, contra a forma curta de pro-SmCD1. Um controle com soro pré-imune mostrou que tal nível de reação é inteiramente inespecífico (painel C da **Fig. 68**). Por fim, no painel D da **Figura 68** é mostrado o reconhecimento específico de pepC-term fusionado à Proteína D, através da reação com soro anti-pro-SmCD1.1, mas não contra anti-pro-SmCD1.2 (compare com painel A).

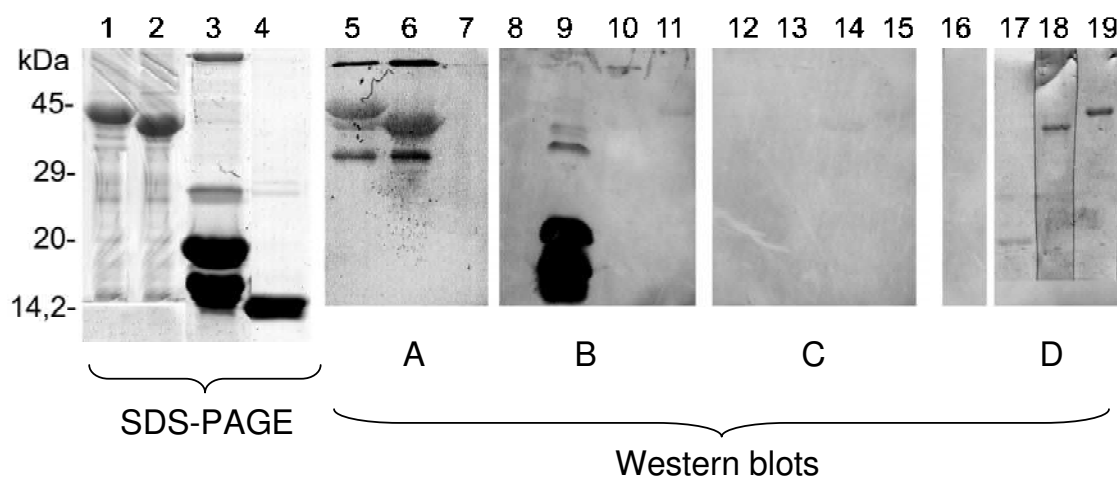


Fig. 68: Determinação do título e especificidade do soro anti-18N_pepC-term por *Immunoblots*. As proteínas transferidas para as membranas de NC são mostradas no painel à esquerda: 1- proSmCD1 forma longa (fitas 5, 11, 15 e 19); 2- proSmCD1 forma curta (fitas 6, 10, 14 e 18); 3- PtnD_pepC-term (fitas 7, 9, 13 e 17); 4- Proteína D livre (fitas 8, 12 e 16). A- soro anti-proSmCD1.2 1:100, B- soro anti-18NpepC-term 1:100, C- soro pré-imune 1:100, D- soro anti-proSmCD1.1 1:250.

Com o intuito de estimar a composição do epítipo B reconhecido em 18N_pepC-term e melhor conhecer a imunogenicidade da extensão C-terminal da

forma longa de SmCD1, a seqüência de pepC-term foi analisada quanto à hidrofiliçidade (escala de Kyte-Doolittle), probabilidade de superfície, flexibilidade e índice antigênico de Jameson-Wolf. O resultado desta análise é apresentado na **Figura 69**, com as seqüências representando potenciais epítomos B em destaque. Este resultado é coerente com a imunogenicidade do peptídeo 18N_pepC-term, indicando que esta pode estar associada a um segmento tão pequeno quanto 6 resíduos de aminoácido.

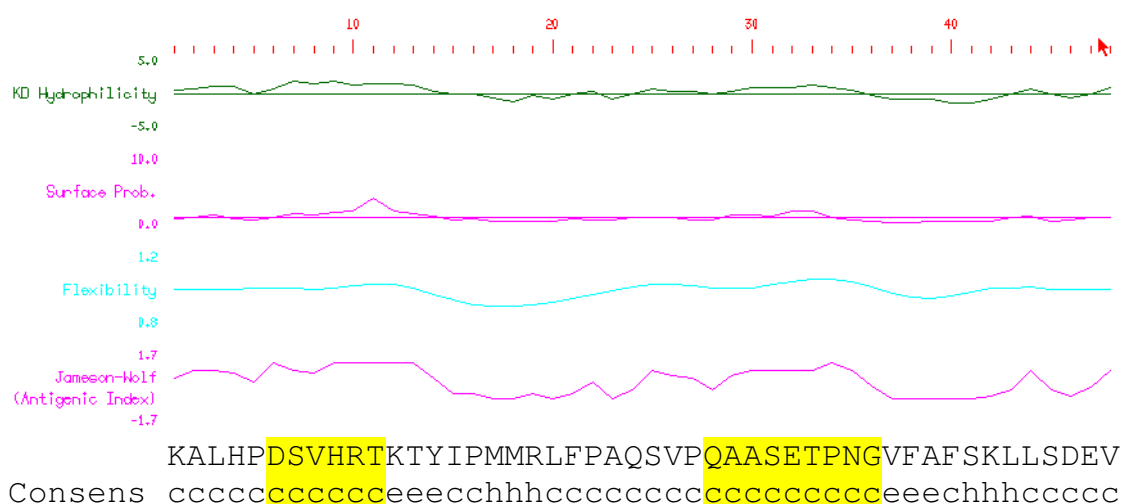


Fig. 69: Análise do potencial imunogênico da extensão C-terminal da forma longa de SmCD1. De cima para baixo estão representadas as escalas de: hidrofiliçidade de Kyte-Doolittle, probabilidade de superfície, flexibilidade e índice antigênico de Jameson-Wolf. No painel inferior é mostrada a seqüência analisada e predição de estrutura secundária: c = aleatório, e = fita- β e h = α -hélice.

4.7 Clonagem das formas longa e curta de SmCD1 em vetor de expressão em *P. pastoris*

Originalmente, a clonagem do cDNA das formas longa e curta de pro-SmCD1 no vetor pPIC9k seria realizada facilmente a partir da ligação dos insertos liberados a partir do vetor de passagem pGEM-t nos sítios Eco RI (5') e Not I (3') de pPIC9k. Entretanto, isto só foi possível para a forma curta (**Fig. 70**). A identificação de uma mutação no clone pGEM-t/pro-SmCD1.1 provocou um desvio na estratégia de clonagem da forma longa da enzima. Como a correção do cDNA da forma longa foi feita diretamente no vetor de expressão pET28a, não estava disponível um clone em pGEM-t com a seqüência correta da forma longa. Por outro lado, a clonagem não poderia ser feita a partir de pET28a, pois o inserto de SmCD1.1 neste vetor não possuía o sítio de Eco RI em 5'. Desta forma, a clonagem em pPIC9k seguiu o plano original para a forma curta, mas teve que passar por uma etapa adicional onde três fragmentos

foram preparados para a ligação: fragmento de 800 pb Eco RI/Acc I de pGEM-t/pro-SmCD1.2, fragmento de 400 pb Acc I/Not I de pET28a/pro-SmCD1.1-corrigida e pPIC9k Eco RI/Not I (**Fig. 70**).

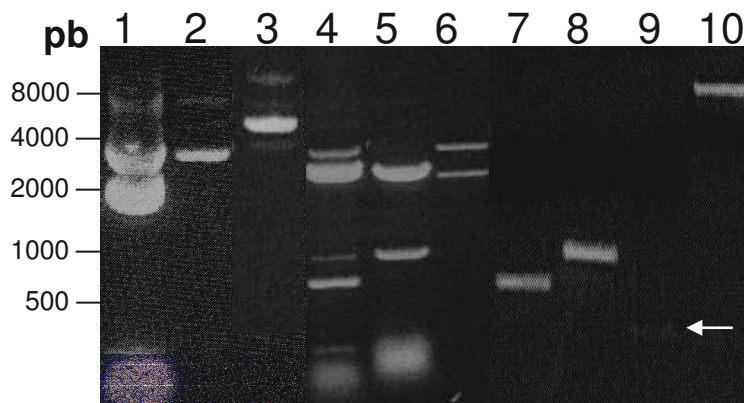


Fig. 70: Análise em gel de agarose (0,8%) dos fragmentos lineares gerados para a clonagem de proSmCD1.1 e proSmCD1.2 em pPIC9k. 1- pGEM-t/pro-SmCD1.2; 2- pET28a/pro-SmCD1.1-corr; 3- pPIC9k; 4- digestão de pGEM-t/pro-SmCD1.2 com Eco RI e Acc I; 5- digestão de pGEM-t/pro-SmCD1.2 com Eco RI e Not I; 6- digestão de pET28a/pro-SmCD1.1-corr com Acc I e Not I; 7- fragmento de 0,8 kb purificado; 8- fragmento de 1,1 kb purificado; 9- fragmento de 400 pb purificado; 10- pPIC9k digerido com Eco RI e Not I, purificado.

Os fragmentos descritos acima foram ligados na reação catalisada pela T4 DNA ligase e o produto utilizado na transformação de células competentes da cepa DH5 α de *E. coli*. Ao redor de 300 colônias foram obtidas a partir de transformação, cerca de 10 vezes mais que o controle negativo, onde o inserto foi suprimido da reação de ligação. Oito colônias de cada placa de transformação com o produto de ligação da forma curta (L1-1 a L1-8) e da forma longa (L3-1 a L3-2) foram cultivadas e os plasmídeos isolados. Os clones positivos foram então identificados por amplificação do inserto por PCR com os *primers* universais AOX1 5' e AOX1 3' (**Fig. 71**). Na **Figura 71** pode-se observar o produto de 492 pb que é amplificado a partir de um vetor pPIC9k vazio. As demais evidenciam os produtos de 1,6 e 1,7 kb para os insertos de pro-SmCD1.2 e pro-SmCD1.1 clonados em pPIC9k.

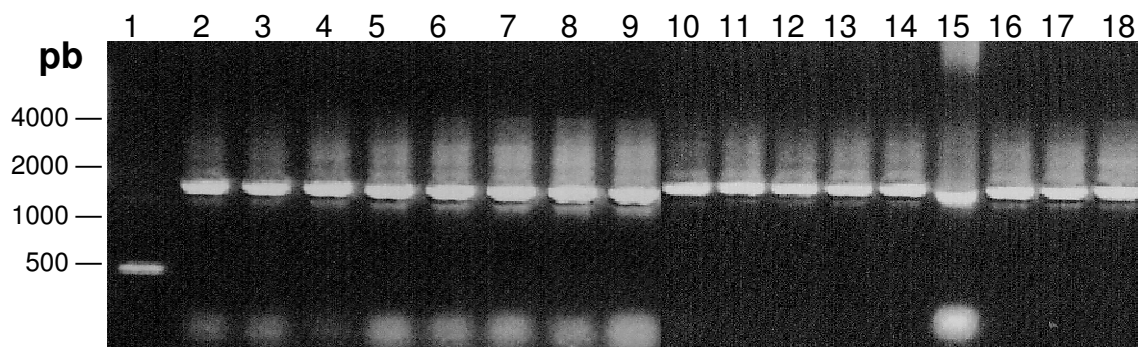


Fig. 71: Análise em gel de agarose (0,8%) das amostras de PCR para detecção da presença dos insertos correspondendo às formas longa e curta de proSmCD1 em pPIC9k. 1- controle com pPIC9k vazio; 2 a 9- clones da transformação com pPIC9k/pro-SmCD1.2; 10 a 18- clones da transformação com pPIC9k/pro-SmCD1.2.

Os clones L1-5 da forma curta e L3-2 da forma longa foram selecionados para análise da integridade do cassete de expressão. A **Figura 72** mostra os insertos de 1,1 e 1,2 kb liberados dos clones L1-5 e L3-2, respectivamente, mediante digestão com Eco RI e Not I. Estes clones foram então submetidos ao sequenciamento com *primers* AOX1, obtendo-se a confirmação que os genes foram clonados na mesma pauta de leitura que a seqüência sinal do fator- α de levedura.

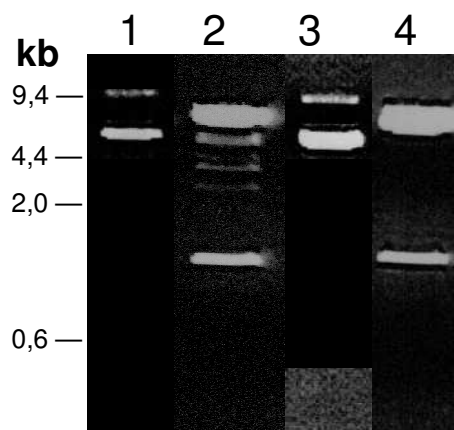


Fig. 72: Análise em gel de agarose (0,8%) dos fragmentos gerados pelas enzimas Eco RI e Not I para confirmação da clonagem de pro-SmCD1.1 e pro-SmCD1.2 em pPIC9k. 1 e 3- pPIC9k/pro-SmCD1.2 e pPIC9k/pro-SmCD1.1 isolados dos clones clones L1-5 e L3-2, respectivamente; 2 e 4- insertos liberados de pPIC9k/pro-SmCD1.2 e pPIC9k/pro-SmCD1.1, respectivamente.

Com os cassetes de expressão devidamente confirmados, culturas dos clones L1-5 e L3-2 foram preparadas e os plasmídeos purificados na escala de “midiprepmin para obter 30-50 μ g de DNA plasmidial. Seguindo o protocolo padrão de clonagem em *P. pastoris*, estes plasmídios seriam linearizados para estimular a recombinação no genoma da levedura. Por questões práticas, optou-se por usar a enzima Sal I na linearização dos

plasmídeos, conforme demonstrado na **Figura 73**. A linearização com Sal I direciona a inserção do cassete de expressão do gene interesse junto ao gene *his4* (não-funcional) no genoma de *P. pastoris*, através de um evento de recombinação simples. Em *P. pastoris* GSS15, este tipo de integração resulta exclusivamente no fenótipo Mut⁺ (rápido crescimento em metanol).

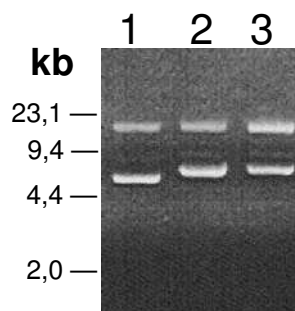


Fig. 73: Análise em gel de agarose (0,6%) das amostras dos vetores pPIC9k (1), pPIC9k/proSmCD1.1 (2) e pPIC9k/proSmCD1.2 (3), submetidas a linearização com Sal I.

Cerca de 10 µg dos plasmídeos linearizados com Sal I, incluindo pPIC9k vazio (como controle negativo), foram usados na transformação de células competentes da cepa GS115 de *P. pastoris*. As células foram plaqueadas para seleção auxotrófica em meio deficiente em histidina (meio MD ágar). Após 3 dias de cultivo a 30° C, colônias translúcidas puderam ser visualizadas. A eficiência da transformação foi calculada em 5×10^3 UFC/µg DNA para a construção da forma longa e 3×10^3 UFC/µg DNA para a construção da forma curta. Tendo em vista a eficiência máxima teórica de 10^4 - 10^3 UFC/µg DNA, as eficiências obtidas podem ser consideradas regulares. Oito colônias da transformação com a construção para cada forma da enzima, além de uma colônia da transformação com pPIC9k vazio, foram repicadas para MD ágar e então submetidas para PCR de colônia (**Fig. 74**). Todos os clones da forma curta, com exceção de um, apresentaram as bandas esperadas de 2,2 kb referente ao gene *AOX1* selvagem e de 1,6 kb referente ao cassete de expressão de pro-SmCD1.2. O clone L1-5_0G7 é mutante natural onde o fenótipo His⁻ foi revertido para His⁺. Os integrantes de proSmCD1.1 foram confirmados pela amplificação do fragmento de 1,2 kb com os *primers* específicos.

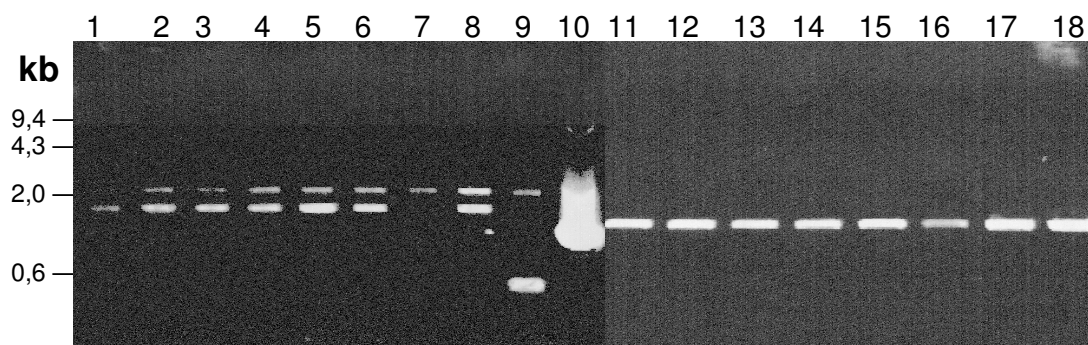


Fig. 74: Análise em gel de agarose (0,8%) das amostras de PCR de colônia para confirmação dos integrantes His⁺. 1 a 8- clones da forma curta (L1-5_0G1 a L1-5_0G8) amplificados com *primers* AOX1; 9- clone de pPIC9k vazio amplificado com *primers* AOX1; 10- controle do PCR com DNA plasmidial de L1-5 de *primers* específicos para o inserto; 11 a 18- clones da forma longa (L3-2_0G1 a L3-2_0G8) amplificados com *primers* específicos para o inserto.

4.7.1 Seleção por geneticina de clones múltipla cópia de *P. pastoris*/pro-SmCD1

Para seleção com geneticina, o restante dos integrantes de *P. pastoris* transformadas com pPIC9k/proSmCD1.1 (L3-2) e pPIC9k/proSmCD1.2 (L1-5) foi ressuspensão em água deionizada estéril numa concentração de $1-9 \times 10^9$ UFC/mL. Cerca de 10^6 células foram inoculadas em placas YPD-ágar com concentrações de geneticina variando de 0,25 até 4,0 mg/mL (0,25G até 4G). Colônias foram obtidas em quantidades decrescentes desde as placas de menor concentração até as de maior concentração de geneticina. Somente na placa G2 (2 mg/mL de geneticina) de L1-5 não foi possível obter colônias. Oito colônias foram repicadas das placas 1G e 4G de L1-5 e das placas 2G e 4G de L3-2 em meio MD-ágar para confirmação do fenótipo His⁺. Estes integrantes (L1-5_1G1 – L1-5_1G8; L1-5_4G1 – L1-5_4G8; L3-2_2G1 – L3-2_2G8; L3-2_4G1 – L3-2_4G8) tiveram a presença dos cassetes de expressão das formas longa e curta confirmados por PCR de colônia com *primers* AOX1 (**Fig. 75**). Os fragmentos amplificados de 1,6 kb e 1,7 kb foram detectados para todos os clones pesquisados, exceto L1-5_4G8. Entretanto a ausência da banda amplificada de 2,2 kb para o gene AOX1 selvagem sugere problemas na PCR desta amostra particular. Para a maior partes das demais amostras, uma tênue banda em 2,2 kb pode ser detectada.

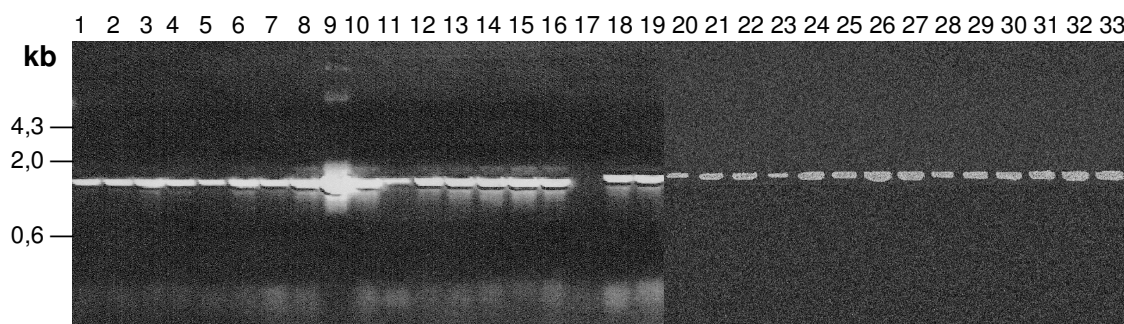


Fig. 75: Análise em gel de agarose (0,8%) das amostras de PCR de colônia (*primers* AOX1) para confirmação dos integrantes His⁺ selecionados em geneticina 1 mg/mL (1G), 2 mg/mL (2G) e 4 mg/mL (4G). 1 a 8- clones da forma curta (L1-5_1G1 a L1-5_1G8); 9- controle do PCR com DNA plasmidial de L1-5; 10 a 17- clones da forma curta (L1-5_4G1 a L1-5_4G8); 18 a 25- clones da forma longa (L3-2_2G1 a L3-2_2G8); 26 a 33- clones da forma longa (L3-2_4G1 a L3-2_4G8).

4.7.2 Testes de expressão de pro-SmCD1 secretada no meio de cultura de *P. pastoris* GS115

Com o intuito de identificar clones de *P. pastoris* capazes de secretar as isoformas de pro-SmCD1 no meio de cultura, um total de 17 clones previamente confirmados por PCR foram cultivados para expressão em pequena escala (25 mL) em meio mínimo. Os clones selecionados para o teste foram: um controle negativo com pPic9k vazio; 8 clones de pro-SmCD1.2 (L1-5): 0G1, 1G1, 4G1-6 e 8 clones de pro-SmCD1.1 (L3-2): 0G2, 2G1, 4G1-6. As células foram cultivadas em meio BMG, utilizando glicerol como única fonte de carbono, até uma $D.O.^{600} = 5$, quando foram coletadas e ressuspensas em meio mínimo até uma $D.O.^{600} = 1,0$. A síntese das isoformas de pro-SmCD1 foi então induzida pela adição de 0,5% de metanol ao meio de cultura. A velocidade de crescimento das células a 30 °C foi verificada pela $D.O.^{600}$ das culturas com 12 e 24h de incubação: com 24h, a $D.O.^{600}$ das culturas variava entre 1,5-2,0 U.A. Mesmo assumindo uma fase lag tão extensa quanto 12h, a densidade das culturas deveria ser aproximadamente 4 U.A., considerando um tempo de duplicação de *P. pastoris* Mut⁺ de 4-6h. Leituras posteriores da densidade ótica das culturas evidenciaram um crescimento com velocidade abaixo da esperada para células com fenótipo Mut⁺. Medidas do pH das culturas revelaram uma acidificação progressiva do meio a partir de 24h (pH 5,5) chegando até a pH 3,5 com 72h de cultura.

Mesmo com o crescimento deficiente das culturas, alíquotas dos sobrenadantes foram retiradas com 12, 24, 48 e 72h e analisadas por SDS-PAGE. Nestes geles, procurava-se por bandas de 47 e 42 kDa ou ainda de 41,5 e 36,5 kDa, correspondendo

à pro-SmCD1.1 e pro-SmCD1.2, caso as proenzimas secretadas no meio ácido das culturas tivessem sofrido processo de auto-ativação, comum às proteases do clã da pepsina. Conforme exemplificado pelos geles da forma curta de SmCD1 revelados pela prata na **Figura 76**, as bandas esperadas não puderam ser evidenciadas pela comparação das amostras de 12h e 48h com as amostras do tempo zero de indução.

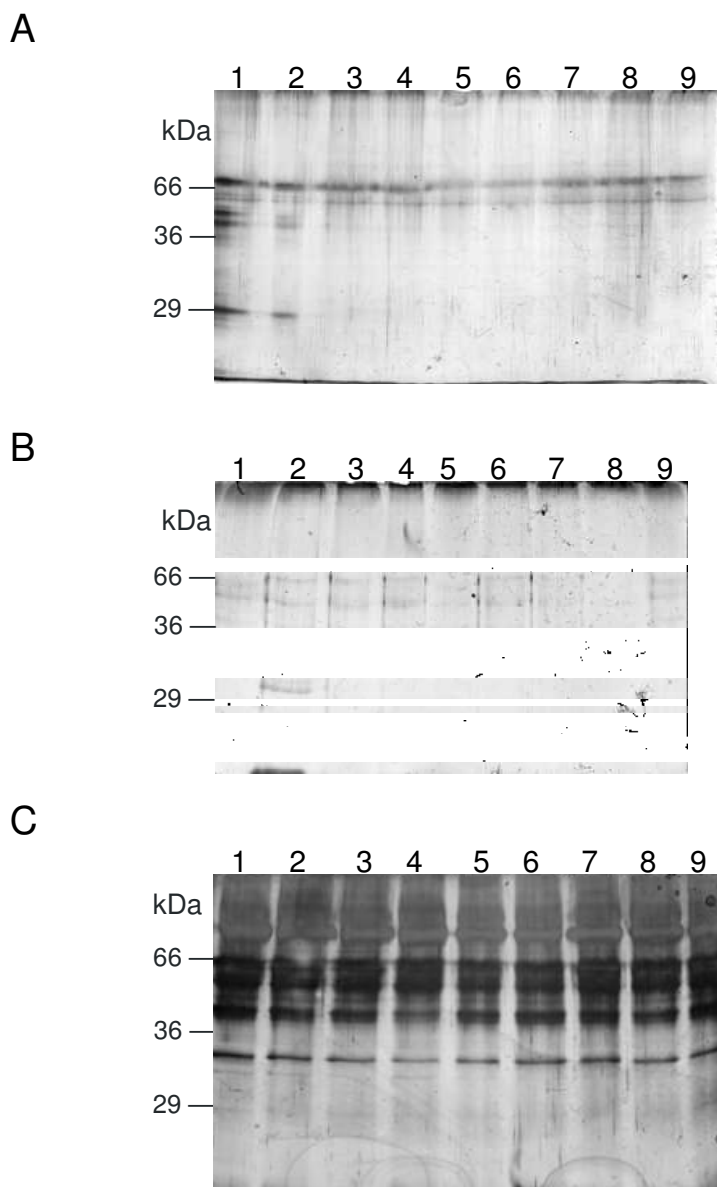


Fig. 76: Análise por SDS-PAGE (12%, redutora) das amostras de sobrenadantes de culturas de *P. pastoris* induzidas com 0,5% de metanol por 0h (A), 12h (B) e 48h (C) para expressão da forma curta de pro-SmCD1. Vinte microlitros dos sobrenadantes de culturas foram aplicados. 1- clone transformado com pPIC9k vazio; 2- clone L3-2_0G2; 3- clone L3-2_2G1; 4 a 9- clones L3-2_4G1 a L3-2_4G6. Revelação por impregnação pela prata.

Os sobrenadantes das culturas com 72h de indução foram concentrados cerca de 20 vezes e as proteínas separadas em geles de poliacrilamida (SDS-PAGE) e reveladas

com o corante coomassie blue. Como pode ser percebido na **Figura 77**, muitas das proteínas presentes no sobrenadante das culturas de *P. pastoris* puderam ser detectadas, mas nenhuma correspondeu à forma curta (**Fig. 77A**) ou à forma longa de SmCD1 (**Fig. 77B**). Estes resultados poderiam refletir um baixo nível de expressão das proteínas recombinantes (abaixo de 0,1 g/L de cultura). Assim, tentou-se detectar as proteínas recombinantes no sobrenadante das culturas de 72h de indução através de *immunoblots*. Conforme evidenciado na **Figura 77C e 77D**, não foi possível detectar (revelação por substrato cromogênico da peroxidase) nenhuma das formas de processamento de SmCD1 curta ou longa.

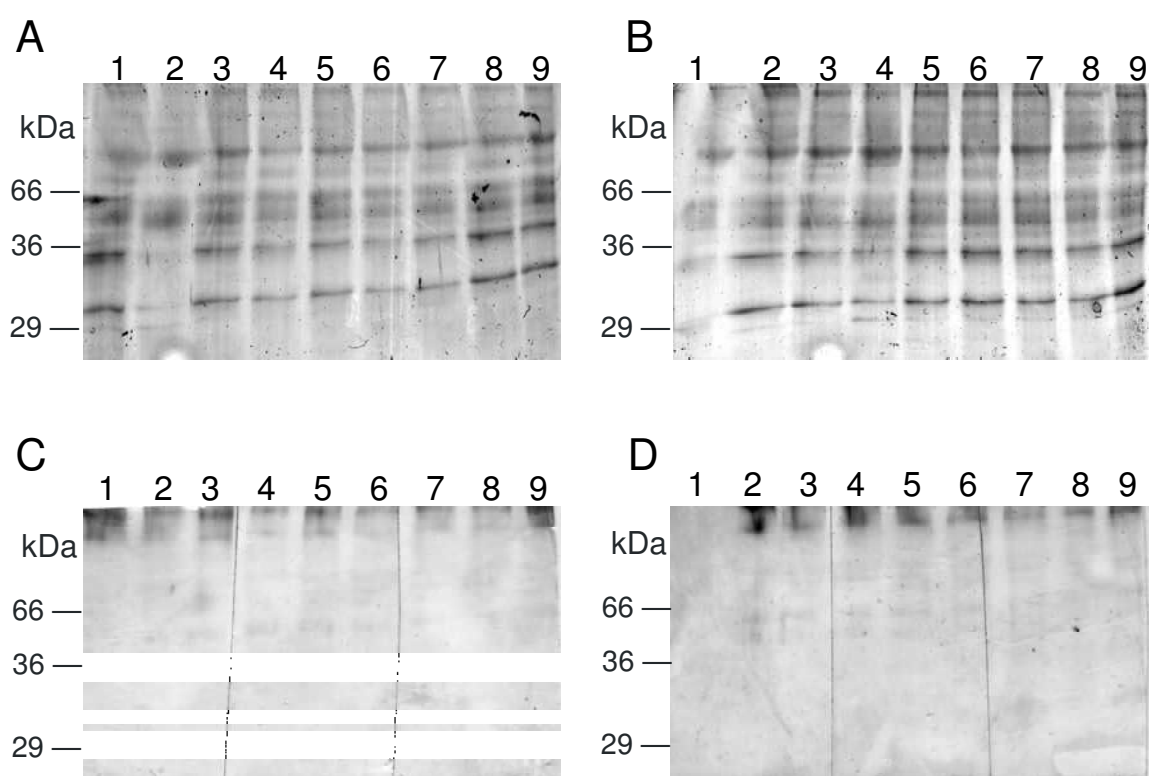


Fig. 77: Análise por SDS-PAGE (12%, redutora) de amostras de sobrenadantes de culturas de *P. pastoris* induzidas com 0,5% de metanol por 72h para expressão da formas curta e longa de proSmCD1. Geles de amostras (24 μ L) da forma curta e da forma longa foram revelados por impregnação pela prata (A-curta e B-longa) ou submetidos à análise por *immunoblot* (C-curta e D-longa). A e C: 1- clone transformado com pPIC9k vazio; 2- clone L1-5_0G1; 3- clone L1-5_1G1; 4 a 9- clones L1-5_4G1 a L1-5_4G6. obrenadantes de culturas (24 μ L). B e D: 1- clone transformado com pPIC9k vazio; 2- clone L3-2_0G2; 3- clone L3-2_2G1; 4 a 9- clones L3-2_4G1 a L3-2_4G6.

Estes resultados sugeriam a ausência de qualquer nível de expressão das formas longa e curta de SmCD1 secretadas no sobrenadante das culturas de *P. pastoris*. Com o intuito de confirmar estes resultados, um número menor de clones foi avaliado de cada forma de SmCD1 em culturas numa escala 8 vezes maior que o teste anterior. Os clones

utilizados no teste foram 0G3 e 4G8 da forma curta (L1-5) e 0G2 e 4G7 da forma longa (L3-2). O meio de cultura e as condições de indução foram mantidas, exceto pela D.O.⁶⁰⁰ inicial que passou a ser 2,0 U.A. Novamente, acompanhou-se a densidade das culturas ao longo da indução, desta vez obtendo-se a curva de crescimento esperada (**Fig. 78**). O tempo de duplicação na fase logarítmica de crescimento (24-72h) foi de 3h, com as culturas atingindo um máximo de densidade entre 72 e 96h. O pH das culturas também foi monitorado, confirmando a acidificação do meio observada no primeiro teste a partir de 24h de indução. Amostras 10x concentradas de alíquotas retiradas com 12 a 120h de indução foram comparadas às amostras das culturas pré-induzidas em SDS-PAGE revelada por prata. Entretanto, apesar da curva de crescimento adequada, não foi possível detectar nenhuma das possíveis formas de processamento (peptídeo-sinal+proenzima, proenzima e enzima) de SmCD1 (dados não-mostrados).

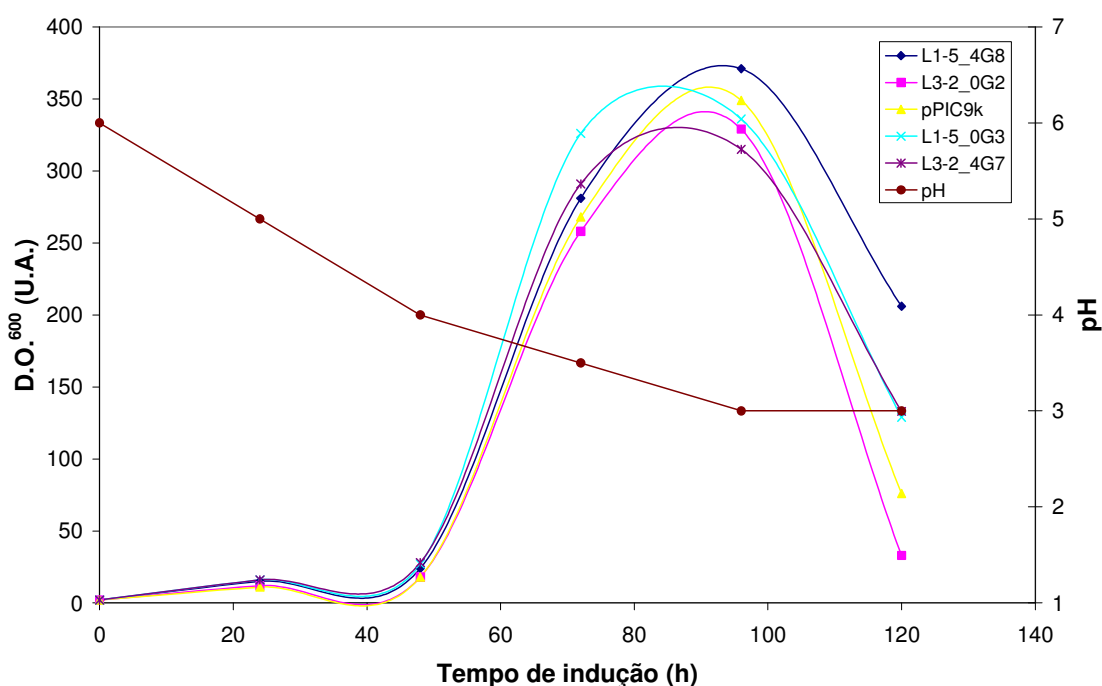


Fig. 78: Curvas de crescimento de culturas de *P. pastoris* GS115 induzidas com 0,5% de metanol em meio mínimo (BM) para expressão das formas longa e curta de pro-SmCD1.

Tendo em vista a confirmação genética da integração do cassete de expressão das formas longa e curta de pro-SmCD1 no genoma de todos os clones testados, a incapacidade de detectar qualquer forma das proteínas recombinantes nos sobrenadantes de cultura foi inesperada. Estes resultados sugerem que as proteínas recombinantes não

foram sintetizadas, enoveladas corretamente ou efetivamente sinalizadas para a via de secreção no meio extracelular. No caso das duas últimas hipóteses serem verdadeiras, as proteínas recombinantes poderiam ser detectadas no extrato total das células de *P. pastoris*. Este não foi o caso, conforme evidenciado no gel de poliacrilamida da **Figura 79**. Nenhum acúmulo apreciável de proteína no citoplasma de amostras induzidas por 24h dos clones expressando as formas longa ou curta de SmCD1 foi detectado em relação aos controles com 0h de indução ao clone obtido da integração de pPIC9k vazio. Estes resultados sugerem que os genes das isoformas de SmCD1 não puderam ser corretamente traduzidos por *P. pastoris*.

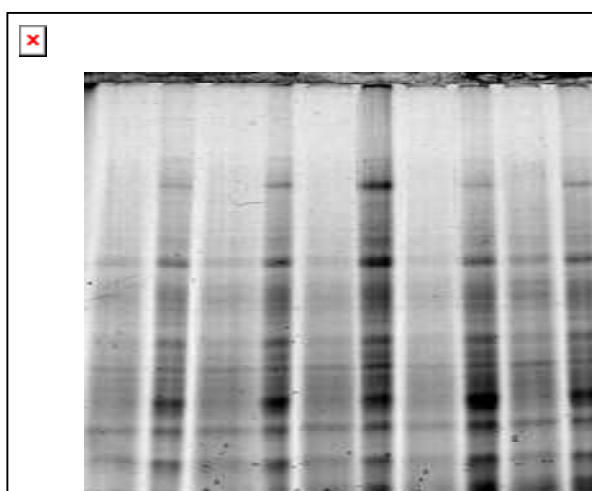


Fig. 79: Análise por SDS-PAGE (12%, redutora) de amostras das frações celulares de clones de *P. pastoris* GS115 expressando as formas longa e curta de SmCD1. Amostras ímpares- 0h de indução e amostras pares- 24h de indução; 1 e 2- pPIC9k vazio (controle negativo); 3 e 4- pPIC9k/proSmCD1.2 (clone L1-5_0G3); 5 e 6- pPIC9k/proSmCD1.2 (clone L1-5_4G8); 7 e 8- pPIC9k/proSmCD1.1 (clone L3-2_0G2); 9 e 10- pPIC9k/proSmCD1.1 (clone L3-2_4G7). Revelação por Coomassie brilliant blue G-25.

4.8 Detecção de isoformas de SmCD1 em extratos de vermes adultos por immunoblot

Experimentos de *immunoblot* foram realizados com o intuito de verificar a utilidade dos soros policlonais obtidos contra as forma longa e curta de SmCD1 e contra os 18 primeiros aminoácidos da extensão C-terminal (18N_pepC-term) que as diferencia (**Fig. 80**). Os três anti-soros foram testados na diluição de 1:1.000 contra o extrato solúvel de vermes adultos de *S. mansoni* (EAV) e isoformas de pro-SmCD1 purificadas por IMAC em Ni-IDA (painel A da **Fig. 80**). As reações foram reveladas usando substrato quimioluminescente da peroxidase conjugada ao anticorpo secundário (1:20.000). Primeiramente, devido à alta sensibilidade do sistema de detecção

empregado pode-se perceber claramente a reação dos soros policlonais anti-pro-SmCD1.1 e anti-pro-SmCD1.2 contra contaminantes presentes nas amostras destas proteínas que foram transferidas para NC e que também estavam presentes nas preparações utilizadas na imunização dos coelhos. Acredita-se que estas reações secundárias ao antígeno majoritário possam ser eliminadas usando-se uma diluição maior do soro. Por outro lado, na diluição empregada para o anticorpo primário, não se observou a reação cruzada com as isoformas de SmCD1 que normalmente é detectada com diluições baixas do soro pré-imune de coelho (**Fig. 80B**). Isto é importante, pois evidencia a especificidade das reações detectadas contra EAV. Inesperadamente, não pôde ser detectada reação do soro anti-18N_pepC-term contra a forma longa de SmCD1. Também não houve reação deste soro contra EAV.

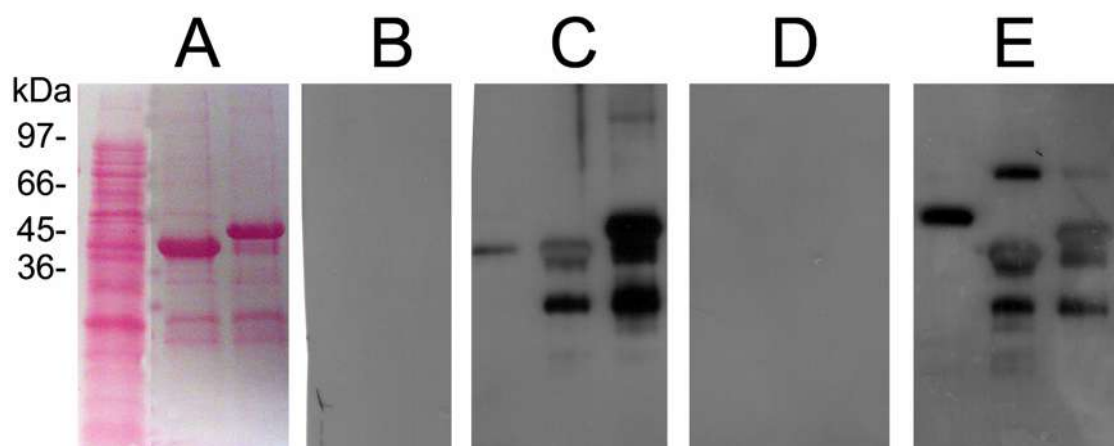


Fig. 80: *Immunoblots* para análise da reatividade dos soros anti-proSmCD1.1, anti-proSmCD1.2 e anti-18N_pepC-term. A- revelação por Ponceau S dos antígenos transferidos para membrana de NC: esquerda - extrato aquoso de vermes adultos de *S. mansoni* (EAV); centro - proSmCD1.2; direita - proSmCD1.1. Réplicas da membrana em A foram reveladas com soros: pré-imune (B), anti-proSmCD1.1 (C), anti-18N_pepC-term (D) e anti-proSmCD1.2 (E).

Os *immunoblots* mostrados na **Figura 80** revelam propriedades curiosas dos soros policlonais obtidos contra as isoformas de proSmCD1 produzidas em *E. coli*. O soro anti-proSmCD1.1 reconhece uma banda em EAV co-migrando com o zimogênio da forma curta, mas não detecta nenhuma banda condizente com a massa de 36 kDa esperada para a forma curta da enzima madura (**Fig. 4.80**). Este resultado poderia ser facilmente interpretado em termos de eventos de auto-ativação e modificação pós-traducional (processamento proteolítico e glicosilação) ocorrendo com a forma longa, se não fosse pela ausência de reação da mesma banda com o soro anti-pro-SmCD1.2. Pelo contrário, o painel D da **Figura 80** mostra que este último soro reage única e fortemente

com uma banda de cerca de 50-55 kDa no EAV. A ausência de bandas reconhecidas pelo mesmo soro co-migrando com esta espécie nas preparações de proSmCD1.1 e proSmCD1.2, descarta uma reação secundária como a possível responsável por esta observação. Contudo, uma análise mais atenta da reação de EAV com o soro anti-proSmCD1.1 sugere um fraco reconhecimento de uma banda entre 50-55 kDa.

5 DISCUSSÃO

Nesta seção, os resultados acima relatados são discutidos em 4 grandes tópicos, divididos por áreas de interesse. A discussão específica sobre a modelagem da especificidade de substrato de SmCD1.2 é consubstanciada pelo artigo encontrado no ANEXO 1. Um aprofundamento desta discussão é apresentado na seção 5.1, à luz dos resultados com inibidores de HIV-1 protease na quimioterapia experimental da esquistossomose mansônica em camundongos e dos dados recentemente publicados por outros.

5.1 *Especificidade e Perspectiva da inibição de SmCD1 como estratégia esquistossomicida*

5.1.1 **Especificidade de substrato de SmCD1.2 analisada por modelagem molecular**

A determinação dos sítios de clivagem na Hb humana reconhecidos por SmCD1 (BRINDLEY *et al.*, 2001) abriu caminho para que estudos comparativos entre aspartil hemoglobinas de esquistossomos e de outros organismos pudessem ser realizados. Brinkworth *et al.* (2001) modelaram as estruturas das plasmepsinas de *P. falciparum* e *P. vivax*, além das aspartil proteases de esquistossomos e *Ancylostoma caninum*. A composição dos subsítios de especificidade foi relatada para todas as enzimas, enquanto que modelos de complexos enzima-substrato foram analisados para a enzima de *S. japonicum* e catepsina D humana (hCD) e os resultados extrapolados para as demais enzimas. Estas comparações revelaram que dentre os 11 resíduos diferentes entre as enzimas de esquistossomos e hCD que são alocados na fenda de ligação ao substrato, 5 também são distintos entre as enzimas das duas espécies de esquistossomos. A maior parte destas mutações é conservativa exceto pela posição 77, onde His em SmCD1.2 foi substituída por Lys em SjCD1.2. His⁷⁷ é situada na base do *flap* e curiosamente é conservada entre a enzima de *S. mansoni* e hCD.

Brindley *et al.* (2001) observaram que entre os 17 sítios de clivagem na Hb humana reconhecidos pelas aspartil proteases de esquistossomo, 12 resíduos em P1 e 6 resíduos em P1' eram Leu ou Phe. Além do mais, 2 dos 3 sítios preferencialmente clivados tinham um resíduo Leu nas posições P1 ou P1'. Esta preferência marcante deve ser resultado de um número maior de contatos favoráveis entre os grupos do substrato e os átomos dos resíduos compondo os subsítios. Esta hipótese serviu de motivação para que dois complexos entre SmCD1 ou hCD e análogos de estado de transição de peptídeos representando as posições P4-P4' nos sítios de clivagem α L105-L106 (SHCL*LVTL) e β L31-L32 (LGRL*LVVY) das cadeias α e β da Hb humana, respectivamente, fossem modelados (vide ANEXO 1). Isto permitiu o mapeamento dos resíduos compondo os subsítios S4-S4' de SmCD1 aliado a uma comparação com as estruturas nos subsítios de hCD. As observações mais importantes provenientes desta análise são resumidas nos parágrafos a seguir.

Os subsítios vizinhos S2 e S4 são pobremente distinguidos na superfície molecular da enzima de *S. mansoni*, não apresentando uma separação definida. Esta situação é atenuada na enzima humana devido à mutação G247Q, onde a cadeia lateral de Gln ajuda a melhor delimitar estes subsítios. Em comum com hCD, o subsítio S4 em SmCD1.2 normalmente acomoda cadeias laterais curtas, contudo resíduos menores e mais hidrofóbicos são favorecidas por esta última em função da mutação de Leu para Met²²⁴ na enzima de *S. mansoni*. O subsítio S2 em SmCD1.2 é hidrofóbico como em hCD, porém é mais amplo devido a duas mutações A226V e L297M da enzima do parasita para a enzima humana. Os subsítios S4 e S2 também podem ser considerados hidrofóbicos, mas são capazes de interagir com as cadeias laterais de P4 e P2 através de ligações de hidrogênio, devido à presença do grupamento hidroxílico de um resíduo de Thr ou Ser. Em ambas as enzimas, o potencial eletrostático na fenda de ligação ao substrato é dominado pela carga negativa dos resíduos de aspartato no sítio catalítico. Entretanto, pôde-se observar uma região neutra na superfície de SmCD1.2 próxima ao subsítio S4 que não está presente em hCD. O desenho de um inibidor que possa alcançar esta região da enzima de *S. mansoni*, pode usufruir de interações estabilizantes que maximizem a afinidade pela enzima do parasito, obtendo assim seletividade frente à enzima humana.

Os resíduos compondo o subsítio S3 de SmCD1.2 e hCD são idênticos, exceto pela mutação conservativa Thr⁸⁰ para Ser no *flap* da enzima humana. Alterações no padrão de ligações de hidrogênio entre o flap e o substrato ligado não foram observadas

em função desta mutação. Entretanto, a inspeção das conformações dos resíduos delineando o subsítio S3 nestas enzimas revelou que a cadeia lateral de Gln¹⁴ é melhor estabilizada em SmCD1.2 numa orientação que abre espaço para a acomodação de cadeias laterais mais volumosas na posição P3 dos substratos comparado à hCD. Isto é coerente com a análise da composição dos substratos reconhecidos pelas duas enzimas. Entretanto, a estabilidade desta conformação de Gln¹⁴ em S3 de SmCD1.2 precisa ser confirmada por simulações de dinâmica molecular com representação explícita do solvente para que maiores conclusões possam ser tiradas. Do lado C-terminal da ligação lábil, o subsítio S2' é estritamente conservado entre as duas enzimas. Entretanto, diferenças marcantes são observadas nas estruturas compondo os subsítios S3' e S4' em SmCD1.2 e hCD. Uma alça na superfície de SmCD1.2 (resíduos 298-303), é dois resíduos mais curta em relação à enzima humana e mais importante, é composta principalmente por resíduos básicos, diferente de hCD cuja alça é rica em Pro. Esta “alça básica” em SmCD1.2 dobra-se sobre a fenda de ligação ao substrato, impondo restrições estéricas aos resíduos em P3' (principalmente) e P4' de seus substratos, que não ocorrem na ligação à hCD. Esta diferença marcante entre as duas enzimas pode ser alvo de exploração no desenho de inibidores seletivos para SmCD1.2.

Os subsítios S1 e S1' estão entre as bolsas de especificidade melhor definidas em aspartil proteases e apresentaram-se extremamente conservados entre SmCD1.2 e hCD. De acordo com a análise descrita no ANEXO 1, entre os 12 resíduos que compõem essas bolsas, apenas 1 sofreu mutação entre a enzima humana e de *S. mansoni* e mesmo assim trata-se de uma mutação conservativa onde Leu²⁹⁹ em SmCD1.2 foi substituído por Ile em hCD. Diferentemente, Brinkworth *et al.* (2001) encontraram em sua análise um resíduo adicional compondo o subsítio S1, Val¹¹⁴, que na enzima humana é substituído por Thr. De qualquer forma, especialmente, há uma excelente sobreposição das estruturas nas regiões compreendendo os subsítios S1 e S1' nas enzimas SmCD1 e hCD. Entretanto, entre os diversos sítios de clivagem na Hb humana que são reconhecidos pela enzima do parasito, mas não por hCD, destacam-se os que diferem principalmente quanto à identidade do resíduo em P1', e.g. α Leu²⁹-Glu e α Phe³⁶-Pro (**Tabela 1**). Por outro lado, o sítio β Glu⁷-Lys é composto por dois resíduos carregados em P1 e P1', sendo reconhecido somente por hCD. Estas diferenças na especificidade podem ser exploradas no desenho de inibidores seletivos para SmCD1.

Uma propriedade marcante da especificidade de substrato das enzimas catépsina D-símile de esquistossomos é a capacidade de clivar o sítio $\alpha\text{Phe}^{36}\text{-Pro}$ na Hb humana. [Brinkworth *et al.* \(2001\)](#) interpretaram esta capacidade exclusiva em termos de contatos hidrofóbicos entre Phe^{36} no substrato e Val^{114} e Thr^{80} em SjCD1.2, que não poderiam ser proporcionados por hCD. Os autores descreveram que a posição dos átomos da cadeia principal do substrato $\text{FLSF}^{36}\text{-PTTK}$ diferiam consideravelmente das posições observadas para os outros substratos de SjCD1.2 modelados devido a restrições conformacionais impostas por Pro em P1min. Assim, [Brinkworth *et al.*](#) justificaram que quando Pro encontra-se em P1min de um substrato a conformação do resíduo em P1 é afetada de forma que SjCD1.2 pode acomodá-las, mas hCD não. Classicamente, substratos contendo Phe-Pro em P1-P1' são eficientemente hidrolisados somente por aspartil proteinases retrovirais ([DUNN *et al.*, 1994](#); [DUNN & HUNG, 2000](#)). Esta especificidade exclusiva das aspartil proteinases retrovirais foi extensivamente explorada como primeira abordagem para alcançar seletividade nos inibidores de HIV-1 protease frente às enzimas ortólogas humanas ([BEYER & DUNN, 1998](#)).

5.1.2 Desempenho de inibidores comerciais de HIV-1 protease na quimioterapia experimental da esquistossomose mansônica

A similaridade inesperada na especificidade de substratos observada para SmCD1 e a proteinase de HIV-1 foi a principal motivação para a avaliação de inibidores comerciais da aspartil protease deste vírus na quimioterapia experimental da esquistossomose mansônica descrita no presente trabalho. Inibidores comerciais foram escolhidos, pois apesar de bem sucedida em encontrar protótipos peptídicos, a estratégia de desenho de inibidores de protease baseados somente na especificidade de substrato não foi a mais eficaz na proposição de fármacos eficientes, principalmente devido à baixa biodisponibilidade dos compostos peptidomiméticos projetados. Assim, ao selecionar fármacos já usados na clínica, o tratamento dos camundongos infectados poderia ser feito por via oral com garantia que níveis séricos relevantes seriam alcançados pelos compostos testados. As doses e a forma de tratamento de cada um dos inibidores testados foram eleitas em função dos dados disponíveis na literatura.

Quando Indinavir é administrado na dose de 20 mg/kg em ratos, na forma de solução oral em 0,05M de ácido cítrico, alcança-se 23% de biodisponibilidade, com o pico na concentração plasmática de 2,8 μM em 30min ([VACCA *et al.*, 1994](#)). A dose

em ratos é cerca de 3-4 vezes maior que a dose praticada em humanos. O valor de K_i contra HIV-1 protease é 0,34 nM e a concentração necessária para alcançar 95% da replicação viral em cultura de células (IC95) é 12-100 nM. Por outro lado, Indinavir é incapaz de inibir a catepsina D humana, pepsina suína, quimosina bovina e renina humana em concentrações inferiores a 10 μ M. Na clínica, Indinavir é usado com duas doses diárias de 8 em 8h. Nos experimentos aqui relatados com camundongos, este inibidor foi administrado em duas doses diárias de 20 ou 100 mg/kg durante cinco dias, sem efeitos significativos observados no número e distribuição de vermes de *S. mansoni*, assim como na oviposição.

A biodisponibilidade oral de Ritonavir administrado a ratos na dose de 10 mg/kg é de 78%, atingindo o pico na concentração plasmática de 2,6 μ M em 2h (KEMPF *et al.*, 1995). Ritonavir liga-se muito forte e especificamente à HIV-1 protease, apresentando um valor de K_i de 15 pM contra esta enzima, comparado aos valores na faixa de 8-20 nM para proteases homólogas humanas. Contudo, a IC50 em cultura de células fica na faixa usual de 9-10 nM que é encontrada para outros inibidores. Na clínica, Ritonavir é empregado em duas doses diárias de 600 mg. Neste trabalho, Ritonavir foi avaliado na dose de 100 mg/kg, duas vezes ao dia, em cinco dias de tratamento de camundongos infectados com *S. mansoni*. Esta dose é 10 vezes superior à dose necessária para alcançar concentrações plasmáticas em ratos que seriam suficientes até para inibir proteases eucarióticas. Contudo, não foi detectada nenhuma atividade esquistossomicida neste fármaco.

Nelfinavir apresenta biodisponibilidade de 52% em ratos quando administrado por via oral na dose de 50 mg/kg (SHETTY *et al.*, 1996). Na clínica, a posologia indicada de Nelfinavir é 2 doses diárias de 1250 mg. No tratamento experimental da esquistossomose relatado neste trabalho, foram utilizadas doses de 50 e 250 mg/kg. A avaliação do número de vermes e demais parâmetros estipulados para verificar o resultado do tratamento, mostrou que Nelfinavir foi inteiramente ineficaz, em ambas as doses testadas. Para Amprenavir, dados de biodisponibilidade e farmacocinética em ratos não estavam disponíveis. A IC90 em cultura de células é 40 nM, com K_i de 0,60 nM sobre a protease de HIV-1 (KIM *et al.*, 1995). Para tratamento da esquistossomose em camundongos foi estipulada uma dose oral de 45 mg/kg, 30 vezes maior que a dose praticada na clínica com humanos. Assim como nos demais fármacos, nenhum impacto

na infecção por *S. mansoni* foi conseguido com este fármaco. Diferentemente, o fármaco de escolha para o tratamento da esquistossomose, PZQ, foi capaz de reduzir consideravelmente o número de vermes (**Fig. 26**) e a oviposição, validando os resultados dos experimentos realizados com os inibidores de HIV-1 protease.

Como os experimentos foram realizados *in vivo*, uma interpretação direta baseada na estrutura dos possíveis complexos formados entre estes compostos e SmCD1 pode ser prejudicada por diversos fatores, de forma que a falta de afinidade destes fármacos por SmCD1 pode não ter sido a única responsável pelos resultados negativos. Por exemplo, dificuldades no acesso dos compostos ao seu local de ação, a saber, as células da gastroderme de *S. mansoni*, podem ter afetado o desempenho destes fármacos. Entretanto, sabendo-se que esquistossomos estão frequentemente ingerindo o sangue do hospedeiro, isto não teria sido um problema majoritário. Além do mais, as doses empregadas permitiriam que os fármacos alcançassem concentrações plasmáticas acima das concentrações efetivas contra a HIV-1 protease e até mesmo contra aspartil proteases eucarióticas. Isto leva a crer que as concentrações plasmáticas obtidas foram insuficientes para inibir a degradação de Hb pelos esquistossomos que é mediada por SmCD1. Outra hipótese é que mesmo tendo conseguido que SmCD1 fosse inibida, outras aspartil proteases (ver discussão na seção seguinte) ou enzimas da família das cisteinil proteases possam ter mantido o catabolismo da Hb funcional nos vermes de *S. mansoni*. O primeiro passo para avaliar estas questões será a realização de ensaios diretos destes inibidores sobre SmCD1.

5.1.3 Identificação de novos inibidores

As dificuldades relatadas neste trabalho em obter formas funcionais de SmCD1.1 e SmCD1.2 recombinantes, expressas em *E. coli* ou *P. pastoris*, inviabilizaram os testes de atividade dos inibidores de HIV-1 protease sobre as enzimas de *S. mansoni*. Entretanto, [Caffrey et al. \(2005\)](#) recentemente relataram estudos de relação estrutura-atividade (SAR) com inibidores derivados do cerne de estatina usando extratos brutos dos vermes adultos de *S. japonicum*, tratados com E-64, como fonte de aspartil proteases envolvidas na degradação da Hb. Como justificativa para a viabilidade desta abordagem, os autores argumentaram que mais de 98% dos ESTs presentes no banco de transcritos de *S. japonicum* provinham de SjCD1.2, sendo portanto esta enzima virtualmente responsável por toda a atividade hemoglobinolítica exibida pelos

extratos tratados com E-64, um inibidor irreversível de cisteinil proteases. De fato, os autores identificaram uma importante restrição estérica no modelo por homologia de SjCD1.2 usado na interpretação de SAR que poderia explicar a falta de atividade inibitória de um dos inibidores possuindo uma ponte intramolecular entre os resíduos em P1 e P3. Através do modelo, os autores identificaram M224 e F120 como os resíduos responsáveis pelo estreitamento na fenda de ligação ao substrato de SjCD1.2 que impedia a ligação do inibidor restrito conformacionalmente em P1-P3, um dos sete derivados de estatina testados. Apoiando esta hipótese, os autores observaram que duas outras aspartil proteases com resíduos diferentes nas posições 224 e 120, incluindo a catepsina D bovina (bCD), eram inibidas por este composto. Um papel para o resíduo M224 havia sido previamente proposto na redução do espaço disponível para a ligação nos subsítios S4-S2 na borda C-terminal da fenda de ligação ao substrato, através da análise do modelo por homologia de SmCD1.2, a qual é consubstanciada neste trabalho pelo ANEXO 1.

Caffrey *et al.* (2005) avaliaram a atividade de uma segunda geração de 4 derivados de estatina desenhados levando em consideração a barreira estérica na fenda de ligação ao substrato de SjCD1.2 mencionada acima e outras propriedades geométricas e físico-químicas observadas em complexos de derivados de estatina com aspartil proteases. Entretanto, a seletividade por SjCD1.2 sobre bCD que os autores esperavam obter não foi conseguida. Os inibidores apresentaram atividade (IC₅₀) na faixa de 15-140 nM para SjCD1.2 e 15-100 nM para bCD. O melhor deles foi 5 vezes mais ativo sobre a protease do parasito do que a protease bovina, enquanto que outros dois foram 2-3 vezes mais ativos contra a enzima mamífera. Curiosamente o inibidor mais seletivo para SjCD1.2 identificado por Caffrey *et al.* apresenta em P1min um biciclo nitrogenado que poderia atuar mimetizando um resíduo de prolina, conforme observado no sítio de clivagem α Phe⁴⁰-Leu da Hb humana. A tímida seletividade conseguida por esta segunda geração de inibidores poder ter sido simplesmente fruto de uma estratégia de desenho mal-sucedida, mas pode ter sido resultante também de interferências de outras aspartil proteases ignoradas atuando nos extratos. Certamente, em vista a análise da família de SmAPs apresentada no presente trabalho, a interpretação de dados de SAR com extratos de *S. mansoni* seria mais complicada.

5.2 Complexidade na família de SmAPs

Parte dos resultados expostos neste trabalho apontam para a existência de outras aspartil proteases em *S. mansoni* diferentes da enzima catepsina D-símile primeiramente clonada por Wong *et al.* (1997). Conforme discutido nesta e nas seções subseqüentes, estas enzimas mais provavelmente constituem uma família de aspartil proteases em *S. mansoni*, batizada aqui de SmAPs. Famílias de genes de aspartil proteases também já foram identificadas em outros agentes parasitários de importância médica para o homem como *Onchocerca volvulus* (JOLODAR & MILLER, 1998) e *Plasmodium sp.* (JEAN *et al.*, 2001), além de nematódeos de vida livre como *C. elegans* (GEIER *et al.*, 1999).

5.2.1 Cromatografia em pepstatina-agarose como estratégia de isolamento de aspartil proteases em *S. mansoni* e outros parasitos

A primeira evidência sustentando a existência da família de SmAPs foi fornecida pelo complexo mapa bidimensional (2D) exibido pela eletroforese da fração do extrato aquoso de vermes adultos (EAV) ligado à coluna de pepstatina-agarose (Fig. 25). Apesar de pepstatina ser um inibidor específico de enzimas da classe das aspartil proteases, poder-se-ia questionar a identidade das espécies reveladas no gel 2D como sendo realmente aspartil proteases em função do elevado número de proteínas ligadas. Entretanto, atividade catepsina D-símile foi detectada na fração ligada à pepstatina agarose usando-se o substrato cromogênico Phe-Ala-Ala-Phe(4-NO₂)-Phe-Val-Leu(4-piridilmetil) éster. Isto sugere que ao menos parte das proteínas observadas no mapa eletroforético 2D sejam realmente aspartil proteases com especificidade de substrato similar à catepsina D de mamíferos. Uma identificação definitiva dos *spots* do gel 2D por espectrometria de massas MALDI-TOF não foi conseguida, possivelmente devido à baixa quantidade de proteína nas amostras e presença de interferentes (supressores de ionização) nas amostras. Por outro lado, é bem provável que a maior parte dos *spots* com massa inferior à 36 kDa sejam fragmentos de proteólise de uma ou mais aspartil proteases ligadas à coluna.

A cromatografia de afinidade em pepstatina-agarose foi utilizada com sucesso na purificação de SjCD1.1 e SmCD1.1 recombinantes secretadas no sobrenadante de cultura de células de inseto (BRINDLEY *et al.* 2001; VERITY *et al.*, 2001). Entretanto, nenhum relato detalhado da aplicação desta cromatografia de afinidade para a purificação de SmAPs a partir de esquistossomos foi descrito. O grupo de I.M. Cesari

do *Centro de Microbiologia y Biología Celular* (Venezuela) relatou a utilização da coluna de afinidade para isolar uma proteína de 45 kDa, apresentando reação sorológica cruzada contra catepsina D bovina (CESARI *et al.*, 1998; VALDIVIESO *et al.*, 2003). A diferença entre os resultados obtidos por estes autores e os dados de purificação em pepstatina-agarose aqui relatados não é facilmente racionalizada. Entretanto, pode-se supor que diferenças nas cepas de *S. mansoni* estudadas, nas formas de preparação de extrato e nas condições de purificação podem ter sido responsáveis por Cesari *et al.* terem supostamente isolado uma espécie majoritária ao invés da diversidade de espécies encontradas neste trabalho. As condições de ligação na coluna foram similares nos dois trabalhos: pH 3,5-3,8 e 1M de NaCl no tampão de equilíbrio da coluna. Desconsiderando as diferenças entre as cepas, os resultados obtidos nos dois trabalhos devem ser fruto da composição das amostras submetidas à cromatografia, haja visto que a interação de múltiplas proteínas com o ligante na coluna de afinidade depende do número de sítios disponíveis na matriz e das constantes de dissociação para cada um dos pares proteína-ligante interagindo.

O perfil de ligação das proteínas de EAV na coluna de pepstatina-agarose obtido neste trabalho é mais compatível com o observado para extratos de *C. elegans* (GEIER *et al.*, 1999). Seis bandas na faixa de 35-52 kDa foram relatadas em SDS-PAGE do material ligado à coluna, resultando num padrão de 19 spots no gel 2D. A disponibilidade do genoma completo deste organismo facilitou a identificação das proteínas ligadas à coluna de afinidade. A partir da análise dos *spots* no gel 2D, seis aspartil proteases foram identificadas, enquanto que outras seis presentes no banco de dados do genoma de *C. elegans* não foram detectadas, totalizando 12 proteases. Sete destas enzimas parecem pertencer a uma mesma família gênica, a julgar pela similaridade nas seqüências.

5.2.2 Espécies detectadas por *immunoblots*

Cesari *et al.* (1998) descreveram a presença de contaminantes de baixo peso molecular na fração ligada à pepstatina-agarose, mas nenhuma análise dos componentes foi mostrada, assim como também não foi descrita a atividade específica da preparação. Com uma Mr de 45 kDa, a espécie relatada por estes autores poderia corresponder tanto à proenzima da forma curta de SmCD1 quanto à enzima madura com a extensão C-terminal. Verity *et al.* (1999) relataram uma única espécie de 47 kDa no extrato de

vermes adultos de *S. japonicum* reativa ao soro de coelho imunizado com SjCD1.1. Estes resultados depõem contra a existência de outras aspartil proteases e até mesmo contra a existência das duas formas das enzimas catepsina D-símile de esquistossomos sugeridas por Wong *et al.* (1997). Os resultados aqui mostrados mostram que em *immunoblot* de extratos de vermes adultos de *S. mansoni* revelados com soros anti-SmCD1.1 e anti-SmCD1.2, duas bandas majoritárias são detectadas com Mr 41 e 47 kDa, respectivamente (**Fig. 80**). Uma terceira banda com pouco mais de 60 kDa também pode ser percebida, porém com muito menor intensidade que estas últimas, na revelação com soro anti-SmCD1.2.

Estas discrepâncias nos resultados de *immunoblot* para as formas longa e curta de SmCD1 devem resultar de diferenças estruturais entre as proteínas recombinantes utilizadas na produção dos anti-soros, mais especificamente, devido à extensão C-terminal presente em SmCD1.1, mas ausente em SmCD1.2. Infelizmente, problemas de origem ainda desconhecida inviabilizaram a comparação destes resultados com *immunoblots* preparados com soro contra a extensão C-terminal. Este soro, direcionado contra os 18 primeiros aminoácidos no N-terminal da extensão C-terminal (18N_pepC-term) mostrou-se altamente reativo contra pepC-term fusionado à proteína D, mas foi incapaz de reagir contra pepC-term presente em SmCD1.1 ou contra esta proteína supostamente presente nos extratos de *S. mansoni* (**Fig. 80**).

5.2.3 Análise dos transcritos de SmAPs

A disponibilidade de números crescentes de seqüências nucleotídicas de *S. mansoni*, provenientes de projetos transcriptoma e genoma deste helminto, criou a oportunidade de identificação de novos genes por ferramentas computacionais. Assim, com o intuito de verificar a existência de outros membros da família de SmAPs, foi feita uma busca por similaridade no banco de seqüências de *S. mansoni* com o programa BLAST, utilizando a seqüência de SmCD1.1. A análise dos TCs, *contigs* e *singlets* resultantes desta busca revelou a existência de 6 outros transcritos pertencentes à família das aspartil proteases claramente distintos de SmCD1, denominados SmCD2 a SmCD7 (**Tabela 8**). Este resultado é compatível com a complexidade do gel 2D da fração do extrato aquoso de *S. mansoni* ligada à pepstatina-agarose. Pode-se inferir a partir do número de ESTs que SmCD3 e SmCD7 representam os genes de SmAP mais abundantes entre os novos transcritos identificados no SmGI. É interessante notar que,

segundo esta análise, SmCD3 e SmCD7 somados correspondem a 18-20% do nível de expressão inferido para SmCD1. Esta diferença no nível de expressão pode estar relacionada à função desempenhada pelos produtos destes genes no parasito. Algumas das regiões alinhadas na **Figura 28** mostram mutações que podem ter um impacto na especificidade de substrato dessas proteases. Por exemplo, a posição 77 (numeração de SmCD1.1 madura) é distinta em SmCD1 (His), SmCD3 (Gln), SmCD4 (Arg) e SmCD5 (Ser) e a posição 80 é conservada como uma Thr em SmCD1 e SmCD3, mas é constituída por um Asp em SmCD3 e SmCD4. Ambas as posições são situadas no *flap* e podem influenciar a afinidade de ligação destas enzimas aos seus substratos. A incomum mutação G35E em SmCD4 chama a atenção, mas precisa ser confirmada neste transcrito que é constituído por um único EST.

Deve-se notar que a futura disponibilização de outras seqüências completas de mRNA pode revelar que alguns dos transcritos aqui identificados por análise computacional compõem, na verdade, um mesmo gene de SmAP. Por exemplo, este é o caso de SmCD2 e SmCD4 que compartilham a seqüência PEQLIN nas posições 54-59 (**Fig. 28**), de qualquer combinação de SmCD2, SmCD3 ou SmCD4 com SmCD6 ou SmCD7, ou ainda da combinação de SmCD5 com SmCD7. Uma vez que estes transcritos não apresentam uma região de sobreposição que permita diferenciá-los, é possível que alguns representem regiões separadas de um mesmo gene.

5.3 Estudos de renaturação de SmCD1 produzida a partir de corpos de inclusão em E. coli

5.3.1 Estratégias de renaturação de aspartil proteases

Ambas as formas, longa e curta, de proSmCD1 foram clonadas como proteínas de fusão com uma seqüência de 6 resíduos de histidina (6xHis), separada por um sítio de clivagem reconhecido pelo fator Xa, no N-terminal das proenzimas (**Fig. 34**). Apesar da presença de alguns códons raros de Arg no cDNA codificando proSmCD1.1 e proSmCD1.2, as proteínas foram abundantemente expressas por *E. coli* BL21(DE3) através da indução com IPTG (**Fig. 38** e **Tabela 9**). Entretanto, as proteínas foram inteiramente expressas na forma de corpos de inclusão (CIs), independente da temperatura e tempo de indução (**Fig. 38**). Este resultado era esperado, uma vez que todas as aspartil proteases expressas até o momento em sistemas heterólogos de origem

procariótica o foram como corpos de inclusão, entre elas: pepsinogênio (LIN *et al.*, 1989), procatepsina E (HILL *et al.*, 1993), proquimosina (YONEZAWA *et al.*, 1993), aspergillopepsinogênio (INOUE *et al.*, 1996), procatepsina D (BEYER & DUNN, 1996) e proplasmepsinas I e II (MOON *et al.*, 1997). Por isto mesmo, as etiquetas de histidina foram inseridas no N-terminal de proSmCD1.1 e proSmCD1.2, pois estas permitem a purificação por cromatografia de afinidade com metais de transição, como na resina Ni-ácido nitriloacético (Ni-NTA), em condições desnaturantes, muito útil para purificação de proteínas expressas como corpos de inclusão (insolúveis). Outra característica é que as proteínas são normalmente clonadas nos vetores de expressão incluindo a região *pro*. Isto ocorre porque se acredita que o propeptídeo seja importante para o enovelamento das proteínas durante a etapa de renaturação e pode ser facilmente eliminado pelo processo de auto-ativação da enzima que é disparado em pH ácido (BEYER & DUNN, 1996; KHAN *et al.*, 1999; BRYAN, 2002). Assim, optou-se pela expressão dos zimogênios das formas longa e curta de SmCD1, ao invés da expressão das formas maduras.

O renovelamento de corpos de inclusão (CIs) não é um processo trivial, frequentemente requerendo processos empíricos exaustivos. O grau de sucesso das tentativas de renaturação de aspartil proteases expressas em *E. coli* é específico de cada proteína, apresentando variações significativas mesmo em proteínas com alta similaridade na seqüência de aminoácidos. Lin *et al.*, (1989) relataram um protocolo bastante inusitado para a renaturação do pepsinogênio produzido como CIs em *E. coli*. Os autores submeteram 1 L da suspensão de CIs (produzidos a partir de 4 L de cultura) a incubação a 4°C por 3 noites em tampão pH 7,2 na presença de 100 mM 2-ME e 150 mM NaCl. A cada noite, o pellet contendo o pepsinogênio desnaturado era novamente ressuspenso e o procedimento repetido. Depois disso, o pellet foi finalmente solubilizado em 60 mL de 6 M de uréia contendo 100 mM de 2-ME e novamente submetido à agitação por 16-18h a 4 °C. A renaturação foi efetuada por rápida diluição em 4L de Tris-HCl pH 8,0. Yonezawa *et al.* (1993) investigaram a renaturação de proquimosina por dois métodos de diálise distintos. Os CIs foram solubilizados em 8 M de uréia pH 7,5 e depois adicionados de 500 mM de NaCl e o pH ajustado para 10,5. Para a renaturação por diálise em um “único passo” 80 mL dos CIs solubilizados foram primeiramente dializados por 4h/25 °C em tampão pH 10,5 e subsequentemente dializados 16-18h contra tampão pH 7,5. Na renaturação por diálise passo-a-passo, a

solução de CIs em 8M de uréia foi submetida a etapas curtas de diálise (4h/25 °C) contra tampões pH 10,5 com concentrações decrescentes de uréia (4, 2, 1 e 0 M), seguido de um passo final de diálise 16-18h a 4° C contra tampão pH 7,5.

Hill *et al.* (1993) induziram a expressão de procatestina E em 8 L de cultura de *E. coli* BL21(DE3) pLys e os CIs obtidos foram solubilizados em 6M de uréia, contendo 1 mM de glicina, 1 mM de EDTA e 50 mM 2-ME. Os autores relataram a obtenção de 3 mg da protease ativa após renaturação por diluição 40 vezes. Maiores detalhes, como a velocidade de diluição e concentração inicial de proteína não foram relatados. Em outro trabalho, Hill *et al.* (1993) descreveram a renaturação de uma construção de proplasmepsina II contendo apenas os 48 últimos resíduos da região pro. Desta vez, antes de solubilizar os CIs (a partir de 1 L de cultura), estes foram incubados sob agitação por 4h em 200 mL de tampão pH 11 na presença de 50 mM de 2-ME. A renaturação foi conduzida diluindo-se 200 vezes os CIs em tampão pH 8,5 e mantendo-se a agitação por 24h. Protocolo semelhante de renaturação foi aplicado à proplasmepsina I, mas a proteína solúvel obtida foi incapaz de se auto-ativar em pH 4,7 (MOON *et al.*, 1997). Os autores prepararam um mutante substituindo um resíduo de aminoácido na região pro (V110P) que era capaz de auto-ativar, porém extensiva agregação durante a etapa de renaturação diminuía consideravelmente o rendimento da preparação. Assim, uma modificação foi introduzida no protocolo de renaturação, na qual os agregados obtidos durante a diluição eram removidos por filtração.

Inoue *et al.* (1996) solubilizaram aspergillopepsinogênio a partir de CIs (provenientes de 1 L de cultura) utilizando 1 L de 8 M de uréia e 100 mM 2-ME. A renaturação foi conduzida por duas diálises contra 10 L de tampão pH 8,0, primeiramente na presença de 100 mM 2-ME e depois sem o agente redutor no tampão, obtendo 19% de rendimento. Até o momento, nossos estudos de expressão em BL21(DE3) mostraram que ambas as formas de SmCD1 são altamente expressas, porém exclusivamente como corpos de inclusão, nas temperaturas de 37 °C e 25 °C, induzidas com 1 mmol/L de IPTG. Beyer & Dunn (1996) otimizaram a renaturação de pro-hCD a partir de CIs solubilizados a 2 mg/mL em 8M de uréia e 50 mM 2-ME em pH 10,7. A renaturação foi conduzida diluindo-se 100 vezes, gota a gota, em água. O pH foi ajustado para 8,7 e a solução de renaturação submetida à agitação por 4h/25 °C O par glutatona reduzida e oxidada (GSH:GSSG 1:0,1 mM) e cisteína 1 mM foram adicionados e a emergência de atividade proteolítica na solução verificada a cada 4h. Os

autores relataram que somente após 8h de incubação nas condições acima descritas houve aparecimento da atividade enzimática e que esta foi máxima somente após o 7º dia. Ao final da renaturação, o pH foi reduzido para 3,7 com adição de 10% v/v de 1M NaOOCH (pH 3,5) + 4 M NaCl e 1% do detergente Brij-35, e a mistura aplicada diretamente em coluna de pepstatina-agarose a 4 °C. Após extensiva lavagem da coluna, a amostra foi eluída em tampão pH 8,5 contendo 0,4 M NaCl e 0,1% Brij-35.

As variações nos protocolos acima descritos para renaturação de diferentes aspartil proteases foi feita com o intuito de demonstrar a diversidade de métodos empregados por diferentes grupos ou até por um mesmo grupo em diferentes proteínas. Muitos dos procedimentos adotados são de cunho prático, outros são frutos do conhecimento acumulado sobre o renovamento de proteínas recombinantes (LILIE *et al.*, 1998; TSUMOTO *et al.*, 2003; BANEYX & MUJACIC, 2004), mas em cada protocolo, muitos detalhes essenciais para o sucesso da renaturação permanecem sem uma razão clara. Assim, pode-se dizer que a renaturação de proteínas a partir de corpos de inclusão continua a ser uma arte. Entretanto, a análise das condições de solubilização e renaturação dos CIs do pequeno grupo de aspartil proteases acima descrito permite apreender algumas lições: (i) todos os CIs foram solúveis em 6-8M de uréia na presença de agente redutor; (ii) em alguns casos, é praticada uma incubação com agente redutor em pH alcalino previamente à solubilização com agente desnaturante; (iii) o método empregado para retirada do agente desnaturante pode ser tanto diluição como renaturação e (iv) o tempo necessário para o renovamento varia de 4 h a alguns dias. Os resultados com proSmCD1.1 e proSmCD1.2 descritos no presente trabalho mostram que algumas das características dos CIs obtidos destas duas proteínas diferem das propriedades comumente observadas para as formas insolúveis de aspartil proteases expressas em *E. coli*.

5.3.2 Solubilização de CIs de SmCD1 por uréia e guanidina-HCl

Há dois aspectos importantes na recuperação de proteínas ativas a partir de CIs, a saber, solubilização e renaturação. A solubilização precisa resultar na dispersão monomolecular e na minimização de interações intra- e intermoleculares não-nativas (TSUMOTO *et al.*, 2003). Mesmo na presença de agente redutor, não foi possível solubilizar eficientemente os CIs de ambas as formas de proSmCD1 em 6-8 M de uréia. Assim, resolveu-se fazer um estudo sistemático das condições de solubilização de

proSmCD1.1 e proSmCD1.2. O pH alcalino mostrou-se importante, proporcionando um ganho de 50 pontos percentuais na quantidade de proteína solubilizada em 8M de uréia (**Fig. 40**). Este resultado é coerente com a presença de 3 pontes dissulfeto presentes nas proteínas recombinantes, uma vez que em pH alcalino é acelerado o processo de quebra e formação destas ligações. Entretanto, o melhor agente solubilizante encontrado foi guanidina-HCl, que na concentração de 6 M proporciona solubilização 20% maior que a condição padrão com 1% de SDS (**Fig. 41**). Com exceção de sarcosil nas concentrações de 1,0-1,5% para SmCD1.2, que resultou num ganho de 50% na solubilização, a presença de detergentes não-iônicos e NaCl teve influência apenas regular na solubilização dos CIs (**Fig. 42**).

É interessante analisar porque guanidina-HCl é capaz de solubilizar as formas longa e curta de SmCD1 tão melhor do que uréia. Historicamente, esses compostos com ação desnaturante sobre proteínas ficaram conhecidos como caotropsos ou agentes caotrópicos devido à desorganização que eles proporcionam na estrutura da água líquida, o que estaria estreitamente vinculado ao seu mecanismo de desnaturação ([KAUSSMAN, 1959](#)). Ao alterar a organização da água, o efeito hidrofóbico, que já foi considerado um dos mais importantes fatores no enovelamento de proteínas, seria atenuado, causando exposição do interior hidrofóbico das proteínas e conseqüente perda de estrutura ([BLOKZIJL & ENGBERTS, 1993](#)). Atualmente o papel da uréia, e em menor extensão de guanidina-HCl, na desnaturação protéica é ativamente discutido. Muitos trabalhos mostraram que além do efeito indireto sobre a estrutura da proteína, estes agentes desnaturantes podem ter um efeito direto, atuando através de interações com a cadeia principal da proteína e cadeias laterais expostas na superfície ([VANZI *et al.*, 1998](#) e referências citadas).

As simulações da estrutura de soluções de uréia em água usando o algoritmo Monte Carlo (MC) por [Vanzi *et al.* \(1998\)](#) e por dinâmica molecular (DM) por [Wallqvist *et al.* \(1998\)](#) concordam que as funções de distribuição radial da água não se alteram na presença de uréia, o que levou [Wallqvist *et al.*](#) a sugerirem que este desnaturante não atuasse através da desestruturação da água. Os resultados de [Vanzi *et al.*](#) discordam desta visão ao mostrar que os ângulos e comprimentos das ligações de hidrogênio efetuadas pela água, na presença de uréia ou guanidina-HCl, na primeira camada de solvatação são significativamente distintos dos encontrados no seio da água líquida. Entretanto, os resultados de [Wallqvist *et al.*](#), são mais realistas, uma vez que

provém de simulações soluções a 6M de uréia. Atualmente, consenso parece existir a respeito da idéia de que o mecanismo de desnaturação protéica mediado por soluções concentradas de uréia (6-8 M) envolve a solvatação de grupos polares na cadeia principal e de cadeias laterais polares na superfície e regiões expostas da proteína (CASTRUOVO *et al.*, 1999; TIMASHEFF & XIE, 2003; MODIG *et al.*, 2003; BENNION & DAGGETT, 2004). Estas interações favoráveis competem e acabam por romper as ligações de hidrogênio intramoleculares na proteína, abrindo caminho para que moléculas de água permeiem a estrutura globular que incha e expõe resíduos hidrofóbicos na superfície

O íon guanidínio é isoeletrônico com a uréia, mas difere pela substituição de um átomo de oxigênio por um grupo amino, o que lhe confere uma carga formal positiva. As simulações computacionais por Vanzi *et al.* sugeriram que o átomo de carbono deste íon apresenta algum caráter apolar, assemelhando-se ao átomo de carbono da uréia. Além disso, os autores predisseram uma variação pequena, porém negativa, na capacidade calorífica (à pressão constante, ΔC_h^p) para a hidratação deste íon que é característica de solutos polares, diferente da uréia que se comporta anormalmente como um soluto apolar, com $\Delta C_h^p > 0$. Estudos que diferenciem o mecanismo de desnaturação protéica do íon guanidínio e da uréia ainda são escassos. Entretanto, intuitivamente é possível supor que este mecanismo esteja relacionado a diferenças na interação com grupamentos carregados nas proteínas. De acordo com esta idéia, Monera *et al.* (1994) demonstraram que a desnaturação mediada por uréia difere marcadamente entre 4 peptídeos de estrutura helicoidal variando quanto ao número de atrações e repulsões eletrostáticas propositalmente inseridas em suas estruturas. Ureia foi mais eficiente em peptídeos contendo interações eletrostáticas repulsivas, enquanto que o número de atrações e repulsões nos peptídeos foi indiferente para a capacidade de desnaturação da guanidina-HCl.

5.3.3 Planejamento estatístico de experimentos, aplicado ao renovelamento de proteínas

A presença de espécies monodispersas após a solubilização dos CIs é indicativo de bons rendimentos de renaturação, pois a presença de microagregados solúveis pode prejudicar o renovelamento da proteína ao semear a formação de agregados maiores e acelerar a precipitação (LILIE *et al.*, 1998; TSUMOTO *et al.*, 2003). As condições de

solubilização selecionadas para as formas longa e curta de proSmCD1 foram tais que ambas as proteínas apresentaram-se mais de 90% no estado monodisperso, conforme demonstrado por SEC em 4M de guanidina-HCl (**Fig. 44**). Deve-se notar que este comportamento foi observado mesmo na ausência de agente redutor no tampão de solubilização, indicando que pontes dissulfeto intermoleculares não estavam envolvidas na estrutura dos CIs de SmCD1.1 e SmCD1.2. Entretanto, as tentativas iniciais de renaturação de SmCD1.1 e SmCD1.2, aplicando os conceitos aprendidos com experiências de relativo sucesso em outras aspartil proteases, fracassaram. Somente agregados insolúveis resultaram da retirada do agente desnaturante, seja por diálise seja por diluição. Diversos fatores podem influenciar o rendimento de um processo de renaturação de proteína na forma de CIs, entre eles: concentração da proteína, forma de renaturação (diálise, diluição, em coluna, etc.) e presença de aditivos. Para alguns destes fatores, existem interpretações bem aceitas de como eles podem influenciar o rendimento da renaturação (TSUMOTO *et al.*, 2003), mas para muitos outros simplesmente não há.

O número de variáveis que podem ter a influência testada sobre um experimento de renaturação pode chegar a 15 ou 20, dependendo da complexidade da proteína, incluindo: pH, força iônica, concentração protéica, método de renaturação, tempo, temperatura, etc. Muitos destes fatores podem ser irrelevantes, enquanto que alguns poucos podem realmente impactar no resultado observado. Tradicionalmente, as variáveis de um processo são examinadas uma a uma, através de alterações sistemáticas seguidas da avaliação do rendimento do processo. Claramente, esta é uma abordagem ineficaz para processos influenciado por múltiplos fatores, assim como a renaturação de proteínas. O planejamento estatístico de experimentos (*DoE*) é um procedimento eficiente pelo qual ensaios podem ser desenhados de forma que os dados possam ser analisados para se tirar conclusões objetivas e válidas (TYE, 2004). Desenhos fatoriais fracionados (FFD) constituem uma forma de *DoE* que permite determinar eficientemente que fatores são importantes no rendimento de um determinado processo, com um número mínimo de ensaios. Outra vantagem da abordagem de FFDs para o estudo de condições de renaturação é que estes podem ser prontamente aplicados na forma de “experimentos de triagem” (“*Screening*”), podendo ser aplicados a diferentes proteínas, sem a exigência de um conhecimento prévio aprofundando sobre suas características e sem inserir tendências no planejamento e análise do experimento.

A primeira tentativa de estabelecer um método geral de triagem de condições de renaturação de proteínas expressas como CIs foi relatada por [Hoffman *et al.* \(1995\)](#). Estes autores empregaram uma matriz esparsa de 50 condições originalmente desenhadas para realizar a triagem de condições de cristalização, ou seja, condições desenhadas para favorecer a redução na solubilidade de proteínas em tampões aquosos, favorecendo a cristalização. Mesmo assim, os autores relataram rendimentos variando de 9-333% relativos às condições originais de renaturação de mais de dez proteínas diferentes. O emprego de DoE na forma de desenhos fatoriais fracionados foi primeiramente relatado por [Chen & Gouaux \(1997\)](#), que mostraram que um desenho de resolução III para avaliar 12 fatores, em dois níveis cada, em 16 tampões, foi capaz de efetivamente encontrar múltiplas condições de renaturação do domínio de ligação do receptor de glutamato (GluR2), com apenas 1/256 do desenho fatorial completo. Os autores encontraram que uma dessas condições provou ser mais econômica e eficiente comparada ao procedimento de renaturação original desta proteína. Mais tarde, o mesmo grupo demonstrou a aplicabilidade geral do método ao empregá-lo na identificação de condições favoráveis à renaturação de 3 outras proteínas: receptor de cainato, lisozima e anidrase carbônica ([ARMSTRONG *et al.*, 1999](#)). [Tobbell *et al.* \(2002\)](#) relataram a aplicação de métodos de DoE em estudos comparativos de renaturação de catepsina S e seu zimogênio procatepsina S. Inicialmente, os autores empregaram um FFD de 8 fatores, em dois níveis, em 16 condições, representando 1/16 do fatorial completo, para identificar condições favoráveis a renaturação das proteínas. Muitos dos fatores empregados faziam parte do desenho original de Chen & Gouaux, incluindo: concentração de proteína, aditivo polar (L-arg), detergente não-iônico (Brij-35), pH, reagentes redox (DTT / GSH:GSSG), força iônica e aditivo não-polar. Os autores investigaram ainda o efeito da temperatura, como oitavo fator. Nas condições testadas, poucas amostras puderam ser renaturadas, levando Tobbell *et al.* a analisarem os oito fatores citados anteriormente em dois FFDs de 4 fatores, em 3 níveis cada, com 9 experimentos cada (1/9 do fatorial completo). A partir daí, 5 fatores principais identificados foram reavaliados em um novo FFD em dois níveis de cada fator com 16 experimentos. Finalmente, os autores otimizaram as condições de renaturação de Catepsina S avaliando os 2 fatores principais que surgiram das triagens acima (pH alcalino e 0,5 M de L-Arg) em um FFD de 5 experimentos, com cada fator em dois níveis e um ponto central. Tobbell *et al.* também demonstraram que no caso desta

cisteinil protease, a presença da região pro foi prejudicial à obtenção de catepsina S ativa.

5.3.4 Viabilidade de produção de isoformas de SmCD1 funcionais a partir da renaturação de corpos de inclusão produzidos em *E. coli*

Neste trabalho, um FFD com resolução III foi aplicado, pela primeira vez, ao estudo da influência de 10 fatores, em dois níveis cada, na renaturação de duas aspartil proteases, as forma longa e curta de proSmCD1. Os fatores avaliados (**Tabela 4**) nos 16 tampões de renaturação empregados, representando 1/64 do fatorial completo (**Tabela 3**), foram inspirados nas condições pesquisadas originalmente por [Chen & Gouaux](#) e representam variáveis cuja influência no processo de renaturação de proteínas foi previamente demonstrada ([LILIE *et al.*, 1998](#); [TSUMOTO *et al.*, 2003](#)). Diferentemente das experiências anteriores da aplicação de FFDs para a identificação de condições favoráveis à renaturação protéica, nenhuma das condições aqui testadas foi capaz de fornecer SmCD1.1 ou SmCD1.2 funcionais. Entretanto, formas solúveis destas proteínas foram obtidas em todas as condições experimentais. Isto indica que a ausência de atividade enzimática nestas amostras pode tanto estar relacionada ao aprisionamento das formas solúveis em uma conformação diferente da nativa, ou simplesmente à uma dificuldade de auto-ativação das enzimas nas condições experimentais avaliadas. Esta última hipótese foi verificada com plasmepsina II, exigindo a mutação de um resíduo crítico na região pro da enzima para permitir sua auto-ativação ([MOON *et al.*, 1997](#)). Não possível verificar as propriedades estruturais destas formas solúveis de proSmCD1.1 ou proSmCD1.2, que pudessem indicar quais das duas possibilidades acima seria mais plausível, devido às quantidades insuficientes das proteínas obtidas.

De qualquer forma, na esperança de que as formas solúveis obtidas pudessem representar um primeiro passo na obtenção das formas longa e curta de SmCD1 enzimaticamente ativas, análises estatísticas foram aplicadas ao rendimento da renaturação medido pela concentração protéica (D.O.²⁸⁰) das amostras solúveis (após centrifugação a 20.000 g). As análises empregadas envolveram a avaliação dos chamados perfis de predição, obtidos a partir do modelo estatístico gerado (**Fig. 47**) e dos fatores principais (**Fig. 48**), como feito anteriormente por [Armstrong *et al.* \(1998\)](#). A partir destas análises, identificou-se os fatores “agente caotrópico”, “aditivo polar” e “agente redox ativo” como os mais importantes na obtenção das formas solúveis de

SmCD1.1 e SmCD1.2. Como mencionado no início desta discussão, a renaturação de proteínas permanece uma arte. Assim, através da aplicação de outras condições experimentais para a renaturação, como por exemplo, o uso de oxidação pelo ar controlada (MENZELLA *et al.*, 2002), aliado aos fatores já identificados favorecendo a obtenção de forma solúveis destas proteínas, pode-se finalmente obter formas ativas de SmCD1.1 e SmCD1.2. Alternativamente, o papel da região pro na capacidade de autoativação das formas solúveis obtidas pode ser investigado por mutação sítio-dirigida ou ainda, outros vetores e sistemas de expressão procarióticos podem ser investigados, como p.ex. fusões para tentar aumentar a solubilidade da proteína expressa (SACHDEV & CHIRGWIN, 1999).

5.4 ***Estudos de expressão de SmCD1 em P. pastoris***

5.4.1 **Expressão de proteínas heterólogas secretadas**

No curso deste trabalho, com o intuito de se alcançar o objetivo original de obter as formas longa e curta de SmCD1 funcionais, decidiu-se por investir o tempo restante e esforço na migração para sistema de expressão eucariótico, ao invés de continuar as tentativas de renaturação a partir de CIs produzidos em *E. coli*. O sistema de expressão heterólogo escolhido foi a levedura *Pichia pastoris* e as razões para isso foram as seguintes propriedades conhecidas deste sistema (DALY & HEARN, 2005; MACAULEY-PATRICK *et al.*, 2005): (i) simplicidade das técnicas necessárias para a manipulação molecular genética (ii) produção de proteínas recombinantes em elevadas quantidades, seja intracelular ou extracelularmente; (iii) a capacidade de promover modificação pós-traducionais típicas de eucariotos, tais como glicosilação, formação de ligações dissulfeto e processamento proteolítico; e (iv) disponibilidade do sistema na forma de um kit comercial. Desde os primeiros relatos de utilização desse sistema a partir de 1994 até a última revisão extensiva da literatura, cerca de 100 proteínas haviam sido expressas com sucesso neste organismo (ROSENFELD, 1999; CEREGHINO & CREGG, 2000). Até a redação deste trabalho, a julgar pela ocorrência das palavras-chaves “express*” e “pichia” na base de dados do ISI *Web of Science*, este número teria quadruplicado, mostrando a aplicabilidade deste sistema de expressão. Entretanto, deste número, cerca de apenas 10 proteínas constituíam aspartil proteases, incluindo: catepsina E gástrica humana (YAMADA *et al.*, 1994), uma enzima do fungo *Mucor pusillus* (MONTESINO *et al.*, 1998; 1999), ciprosina da planta *Cynara cardunculus*

(WHITE *et al.*, 1999), PEP2 do fungo *Aspergillus fumigatus* (REICHARD *et al.*, 2000), SAP1 da levedura *Candida albicans* (BEGGAH *et al.*, 2000), uma enzima do fungo *Rhizopus microsporus* (SCHOEN *et al.*, 2002) e pepsinogênio suíno (YOSHIMASU *et al.*, 2002; 2003). Mais marcante é o fato de que até o momento, somente um gene de esquistossomos foi expresso com sucesso em *P. pastoris*, SmAE (uma asparaginil endopeptidase) de *S. mansoni* (CAFFREY *et al.*, 2000).

Uma das grandes vantagens do sistema de expressão em *P. pastoris* é a possibilidade de secreção da proteína heteróloga no sobrenadante de cultura em meio químicamente definido (meio mínimo), aliada à capacidade da levedura utilizar metanol como única fonte de carbono. Ao endereçar o gene expresso para o meio extracelular, a etapa de purificação subsequente é em muito facilitada, pois a proteína recombinante é usualmente encontrada como o componente protéico majoritário no sobrenadante de uma cultura feita em meio mínimo (DALY & HEARN, 2005; MACAULEY-PATRICK *et al.*, 2005). Ao submeter um gene integrado em *P. pastoris* ao controle do promotor da enzima álcool oxidase (AOX1) é possível induzir a expressão do produto deste gene pela transferência das células da levedura para meio de cultura contendo metanol como única fonte de carbono no meio. Duas vantagens podem ser citadas ao se usar metanol como agente indutor da expressão de genes em *P. pastoris*: primeiramente, como muitos microorganismos são incapazes de sobreviver utilizando metanol como única fonte de carbono, o meio de cultura empregado na expressão pode ser considerado seletivo; segundo, ao fornecer um meio de controle da síntese do gene, pode evitar que este se acumule em quantidades suficientes para ter um efeito tóxico sobre a levedura. Todos os genes de aspartil proteases de que se tem conhecimento, além do gene SmAE, foram clonados em vetores de integração em *P. pastoris*, e.g. pHIL-S1, pPIC9(k) e pPICZ α , desenhados para resultar tirar proveito das propriedades acima descritas. O vetor pHIL-S1, mais antigo, utiliza a seqüência sinal da fosfatase ácida (PHO1) para secreção no meio de cultura, enquanto que os demais vetores utilizam a seqüência sinal do fator- α de acasalamento (ROSENFELD, 1999; CEREGHINO & CREGG, 2000).

Diversos fatores podem influenciar a produção de uma proteína recombinante por *P. pastoris* (SREEKRISHMA *et al.*, 1997): características intrínsecas da proteína (estabilidade, estrutura, modificação pós-traducional), número de cópias do cassete de expressão, sítio e modo de integração do cassete de expressão, contexto do códon de

iniciação da tradução (ATG), composição A+T do cDNA, blocos de transcrição e de tradução, natureza do sinal de secreção, atividade proteolítica endógena, cepa hospedeira, meio e condições de crescimento, e parâmetros de fermentação. Tendo em vista as vantagens discutidas no parágrafo anterior e o histórico de sucesso na expressão de enzimas homólogas e de proteínas de *S. mansoni*, optou-se por clonar os genes proSmCD1.1 e proSmCD1.2 no vetor pPIC9k (**Fig. 72**). Os plasmídeos foram digeridos com Sal I para a integração por recombinação simples no sítio *HIS4* do genoma de *P. pastoris* GS115 (**Fig. 73**), fornecendo clones com o fenótipo Mut⁺/His⁺, permitindo o crescimento rápido na presença de metanol e seleção com histidina. Conforme verificado pelo sequenciamento dos vetores de integração e PCR de colônia, os integrantes selecionados apresentaram um cassete de expressão constituído pelo promotor AOX1 (elevado nível de transcrição induzido por metanol), seguido da seqüência sinal do fator- α (para secreção no meio extracelular) e finalmente pelas seqüências de SmCD1.1 ou SmCD1.2 (**Fig. 74**).

Os diferentes níveis de resistência ao antibiótico geneticina foram utilizados para a seleção de clones das formas longa e curta de proSmCD1 carreando múltiplas cópias do cassete de expressão e, portanto com o potencial de expressar maiores quantidades das proteínas recombinantes. Os testes de expressão foram realizados com culturas induzidas por 0,5% de metanol em meio mínimo e a presença das proteínas recombinantes pesquisadas por SDS-PAGE revelada por impregnação pela prata (**Fig. 76 e 77**) e *immunoblots* com o anticorpo anti-proSmCD1.1 (dados não mostrados).

5.4.2 Viabilidade de produção das isoformas de SmCD1

As análises dos sobrenadantes e da fração celular das culturas de *P. pastoris* (**Fig. 78**) indicaram que as isoformas de proSmCD1 não puderam ser sintetizadas nas condições experimentais testadas. A seguir serão analisados alguns dos fatores que poderiam ter resultado na incapacidade de *P. pastoris* expressar esses genes. Primeiramente, é possível que diferenças entre as preferências de códons exibidas pela levedura e por *S. mansoni* tenham inviabilizado a expressão dos genes de interesse. A análise da freqüência de utilização dos 428 códons presentes na seqüência da forma longa de proSmCD1 por *P. pastoris* de fato revelou um total de 25 códons utilizados em menos de 10% dos genes conhecidos desta levedura. Estes códons raros aparecem por

toda a extensão da sequência e dividem-se entre códons codificando Pro (6 CCG), Gly (GGG), Ser (2 TCG e 3 AGC), Leu (4 CTC), Ala (3 GCG) e Arg (1 CGC e 1 CGA).

O significado desta análise pode ser avaliado comparando-se com a frequência de aparecimento de códons raros em outros genes expressos com sucesso em *P. pastoris*. Dois genes foram escolhidos para comparação, por razões óbvias, o da asparaginil endopeptidase de *S. mansoni* (CAFFREY *et al.*, 2000) e outro funcionalmente relacionado expressando o pepsinogênio suíno (YOSHIMASU *et al.*, 2002). Um total de 17 códons raros foram encontrados ao longo dos 409 códons de proSmAE: 3 códons de Gly (GGG), 4 códons de Arg (1 CGC e 3 CGA), 5 códons de Ser (3 TCG e 2 AGC) e outros 5 entre códons de Pro (2 CCG), Leu (2 CTC) e Ala (GCG). Pode-se perceber que todos os códons raros presentes na sequência de SmCD1.1 estão presentes em SmAE, porém em menor quantidades. Uma vez que está última consegue ser expressa em *P. pastoris*, esta análise parcial poderia indicar que o maior número de códons raros na sequência de SmCD1.1 poderia ser responsável pela impossibilidade de detectar este produto no sobrenadante de cultura ou no citoplasma das leveduras. Entretanto, pepsinogênio suíno com 30 códons raros de *P. pastoris* em sua sequência pode é expresso em quantidades superiores a 30 mg/L de cultura (YOSHIMASU *et al.*, 2002), um nível de expressão que seria facilmente detectado em nossas análises se fosse o caso de SmCD1.1 ou SmCD1.2. Na sequência codificando os 385 resíduos do pepsinogênio suíno são encontrados 20 códons raros AGC codificando serina, seguido de 6 códons CTC codificando leucina, 2 códons de prolina (CCG) e um de arginina (CGC).

A segunda hipótese a ser analisada diz respeito a uma possível terminação prematura da transcrição do gene através do reconhecimento de um sítio de terminação/poliadenilação na sequência de proSmCD1 pela maquinaria de molecular empregada por *P. pastoris*. Este fenômeno está usualmente associado a um alto conteúdo de A+T na sequência nucleotídica (e.g., WITHERS-MARTINEZ *et al.*, 1999; OUTCHKOUROV *et al.*, 2002; WOO *et al.*, 2002). Sreekrisma *et al.* (1997) propuseram que o conteúdo de AT de gene para ser eficientemente expresso em *P. pastoris* deve pairar entre 30 e 55%. O conteúdo de AT do genoma de *S. mansoni* é sabidamente elevado (66%), assim a terminação prematura da transcrição dos genes de SmCD1.1 e SmCD1.2 pode ter sido responsável pela incapacidade de detectar as proteínas recombinantes no sobrenadante de cultura ou no extrato celular. Para melhor

considerar esta possibilidade, é relevante novamente comparar os conteúdos AT de outros genes expressos com sucesso em *P. pastoris*. Os conteúdos de AT dos genes de SmCD1.1, pepsinogênio suíno e SmAE são 57,9%, 41,9% e 65,7%, respectivamente. Isto indica que uma análise superficial do conteúdo AT, não pode explicar o rendimento na expressão destes genes em *P. pastoris*.

A regulação da transcrição gênica é um dos processos moleculares mais intrincados de que se tem conhecimento, envolvendo uma diversidade de sinais imputados não só pela identidade de algumas seqüências curtas específicas no DNA (sinais), mas também pela posição espacial destas seqüências e pelos fatores envolvidos no reconhecimento destas. A clivagem do mRNA precursor após a transcrição é uma etapa crucial de sua maturação. Seguindo-se à clivagem, a maior parte dos mRNAs eucarióticos adquire uma calda poli(A) em sua extremidade 3' (revisado por [ZHAO *et al.*, 1999](#)). A poliadenilação acarreta na terminação da transcrição e transporte do mRNA maduro do núcleo para o citoplasma. A presença da calda poli(A) também favorece a tradução e a estabilidade do mRNA mensageiro. Conforme mencionado anteriormente, sinais presentes na seqüência nucleotídica é que codificam em última instância o sítio de clivagem e poliadenilação. Somente no final da década de 90, os sinais responsáveis pelo direcionamento do processamento da extremidade 3' de genes de leveduras passaram a ser melhores conhecidos ([GUO & SHERMAN, 1996](#); [VAN HELDEN *et al.*, 2000](#)). Dentro da região codificante ou ORF, 3 elementos são distinguidos: EE ou elemento de eficiência (TATATA e seqüências relacionadas TATTTA, TATGTA e TTTTATA), PE ou elemento de posicionamento (AATAAA ou AAAAAA) e o sítio de adição da calda poli(A) propriamente dita (YA_n, onde Y é uma base pirimidínica). PE é normalmente encontrado 10-30 nucleotídeos acima (5') ao sítio de poliadenilação, enquanto que a posição de EE varia com a eficiência de transcrição do gene ([ZHAO *et al.*, 1999](#)). Abaixo do códon de terminação, outros elementos ricos em AT foram detectados através de métodos estatísticos, sugerindo uma possível ligação com o processamento 3' do mRNA ([VAN HELDEN *et al.*, 2000](#)).

O insucesso na expressão dos genes de SmCD1 pode estar relacionado à eventual presença de algum destes elementos regulatórios em levedura, em meio à seqüência codificando estas proteínas. De fato, a seqüência mais comum de um PE é encontrada na posição 948 do cDNA de SmCD1.1 (contado a partir da 1ª base da preproenzima) e na posição esperada precedendo o sítio de poliadenilação de mamíferos

no pepsinogênio suíno, mas não no cDNA de SmAE. Contudo, uma seqüência típica de EE não é encontrada acima deste PE no cDNA de SmCD1.1. Assim, não foi possível identificar um motivo claro pelo qual uma suposta terminação prematura da transcrição em *P. pastoris* dos genes das formas longa e curta de SmCD1 teria ocorrido, em comparação com outros dois genes relacionados funcional e filogeneticamente. O fato é que a otimização total da seqüência nucleotídica para reduzir o conteúdo AT e alterar a utilização de códons tem sido o único expediente pelo qual se tem conseguido reverter a situação em que um gene não é expresso numa situação em que elevados níveis de expressão em *P. pastoris* são observados (WITHERS-MARTINEZ *et al.*, 1999; OUTCHKOUROV *et al.*, 2002; WOO *et al.*, 2002). Particularmente, o estudo sistemático de Woo *et al.* mostrou que a mudança da seqüência sinal para secreção no meio de cultura, número de cópias do cassete de expressão e fenótipo Mut^s/Mut⁺ não foram capazes de resultar na expressão de uma proteína recombinante previamente demonstrada com incapaz de ser expressa em *P. pastoris*.

5.5 A extensão C-terminal de SmCD1.1

5.5.1 Função e papel biológico

A elucidação do papel do domínio C-terminal na função da aspartil protease catepsina D-símile de *S. mansoni* constituía um dos objetivos originais deste trabalho. Para isto, tanto o cDNA da suposta forma curta quanto o cDNA da forma longa de SmCD1 foram clonados por uma estratégia de RT-PCR. Deve-se ressaltar que a importância da clonagem destas “duas formas” de SmCD1 para estudo da estrutura e função do domínio C-terminal independe da real existência destas formas *in vivo*. Esta última questão foi considerada apenas colateral no presente trabalho. Contudo, em função da carência de dados disponíveis que comprovem a ocorrência da forma curta de SmCD1 no parasito, alguns dos resultados aqui apresentados fornecem subsídios para aprofundar a discussão desta questão. Por exemplo, os experimentos de RT-PCR para amplificação dos cDNAs correspondentes às duas formas de SmCD1, indicam que mRNAs com diferentes abundâncias serviram de moldes nestas reações, tendo em vista a intensidade nitidamente maior do cDNA obtido da amplificação com *primers* para SmCD1.2 (Fig. 31). Esta observação sugere que genes distintos codifiquem as formas longa e curta de SmCD1 em *S. mansoni*. Na hipótese de existir um molde de mRNA para a forma curta e outro molde de mRNA para a forma longa, ambos na mesma

quantidade, seria de se esperar uma intensidade duas vezes maior para o produto amplificado por RT-PCR de SmCD1.2. A inspeção visual do gel de agarose na **Figura 31** permite se afirmar que isto não só é verdade como também a intensidade do cDNA amplificado de SmCD1.2 é certamente muitas vezes superior ao de SmCD1.1

No último ano, [Morales et al. \(2004\)](#) descreveram a organização genômica de SmCD1.1 como sendo composta por seis *introns* em meio a sete *exons*, totalizando 13 kb de comprimento. Os *introns* variam de 30 a 5025 bp de tamanho e revelaram a presença de elementos genéticos móveis, como retrotransposons. Usando como sonda o cDNA de SmCD1.1, estes autores identificaram oito clones positivos numa biblioteca de BACs, representando uma cobertura de 8x do genoma de *S. mansoni*. Baseados nestes resultados, os autores concluíram que o gene de SmCD1.1 deve existir em cópia única no genoma desta parasito. Entretanto, isto não foi confirmado por sequenciamento, pois somente um dos clones positivos foi caracterizado pelos autores. O sequenciamento de fragmentos Hind III deste clone revelou uma seqüência contígua de 2470 pb acima do códon de iniciação e 3954 pb abaixo do códon de parada. A seqüência genômica do *exon 1* revelou ter 2 aminoácidos a mais no N-terminal (peptídeo sinal) que a preproenzima deduzida anteriormente por [Wong et al. \(1997\)](#). Curiosamente, mais uma vez a questão da extensão C-terminal de SmCD1.1 foi negligenciada pelos autores.

As diferentes abundâncias dos transcritos das formas curta e longa de SmCD1 observadas no presente trabalho poderiam ser resultantes de regulação ao nível da transcrição e não de diferenças no número de cópias de genes distintos para cada forma da enzima, uma vez que até o momento apenas um único gene foi identificado codificando SmCD1 ([MORALES et al. 2004](#)). De fato, a estrutura genômica relatada para SmCD1 mostra que a região do gene codificando a extensão C-terminal não fica isolada num exon separado, como podia se esperar no caso do transcrito da forma curta ser resultante de *splicing* alternativo. Na verdade, a estrutura do gene de SmCD1 não forneceu indício algum de como as supostas duas formas de SmCD1 poderiam ser obtidas a partir do mesmo gene. A análise computacional das abundâncias de ESTs relacionados estritamente à forma longa (alinhando com a extensão C-terminal) sugere que o suposto mRNA da forma curta é pelo menos tão abundante quanto o da forma longa (**Fig. 29**). Esta análise indica uma ligação genética entre as duas formas, coerente com a presença de um único gene codificando apenas a forma longa de SmCD. Desta

forma, os dados discutidos até aqui fornecem maior suporte à noção de que a única forma de SmCD1 existente em *S. mansoni* seja a forma incluindo a extensão C-terminal, SmCD1.1.

Por outro lado, permanecem conflitantes os resultados de RT-PCR sugerindo que a mensagem da forma curta é mais abundante que a da forma longa. Uma possível explicação para esta discrepância poderia vir do anelamento inespecífico dos *primers* desenhados para amplificar SmCD1.2 com o mRNA de outros genes parálogos de SmAPs. Os resultados da análise *in silico* da família de SmAPs indicam que pelo menos outros 4 mRNAs poderiam ser amplificados pelos *primers* desenhados para clonar o gene da forma curta de SmCD1. Entretanto, o sequenciamento de alguns dos poucos clones de SmCD1.2 em pGEM-t mostrou, para todos, 100% de identidade com a seqüência de conhecida de SmCD1.2.

A análise da inter-relação desses transcritos e SmCD1 revelou que nenhum se alinha com a extensão C-terminal que define a forma longa de SmCD1 (**Fig. 28**). Este fato pode ser apenas decorrente das seqüências incompletas dos demais genes, mas é possível especular que a presença do domínio C-terminal esteja vinculada a uma função diferenciada de SmCD1 associada a uma localização intracelular específica determinada pelo seu peptídeo sinal, uma vez que, até o momento, este também não foi compartilhado por nenhuma outra SmAP encontrada nos bancos de dados. Interessantemente, em aspartil proteases de plantas, um segmento de cerca de 100 resíduos, específico das enzimas destes organismos, parece estar envolvido no endereçamento para vacúolos de armazenamento e não na regulação da atividade destas enzimas (WHITE *et al.*, 1999; SHAAF *et al.*, 2004). A obtenção da seqüência completa de outros genes da família SmAP e subsequente clonagem para estudos funcionais irá ajudar a esclarecer estas e outras questões que permanecem sobre a função desta extensão C-terminal.

5.5.2 Expressão em *E. coli* e análises estruturais

Modelos teóricos foram propostos para o domínio principal de SmCD1, conforme descrito no presente trabalho (vide ANEXO 1) e também por outros (BRINKWORTH *et al.*, 2001). Entretanto, a extensão C-terminal de SmCD1.1 (pepC-term) não pôde ser alvo dos mesmos estudos de modelagem por homologia, em virtude de não apresentar similaridade significativa com a seqüência de nenhuma proteína

disponível nos bancos de dados públicos. No presente trabalho foram relatadas análises computacionais para a predição da estrutura secundária e do enovelamento geral deste pequeno domínio. A análise da seqüência indicou que pepC-term tem cerca de 30-40% de estrutura secundária composta por duas α -hélices intercaladas à duas fitas- β curtas (**Fig. 65**). Dois modelos 3D foram gerados para os 43 resíduos de pepC-term, um por método *ab initio* (**Fig. 64**) e outro utilizando técnicas reconhecimento de enovelamento (**Fig. 67**). Ambas as predições sugerem uma estrutura para pepC-term dobrada, assemelhando-se a um grampo de cabelo torcido. O resultado destas predições associado à importância inerente de se determinar a estrutura de pepC-term por métodos físicos, levou ao início de um programa de estudos estruturais sobre a extensão C-terminal de SmCD1.1. Para este fim, um sistema de expressão deste peptídeo na forma de uma proteína de fusão de 20 kDa foi desenvolvido.

O cDNA correspondendo à pepC-term foi clonado no vetor pBUF5 (**Fig. 50**) e as condições otimizadas de expressão forneceram quantidades apreciáveis da proteína de fusão na fração solúvel dos extratos de *E. coli* BL21(DE3) (~ 30 mg/L de cultura; **Fig. 52**). Estudos por espectrometria de massas demonstraram que a ligação 18-19 de pepC-term era clivada por uma protease endógena de *E. coli*, durante a expressão (**Fig. 58 e 63**). A estratégia de purificação desenvolvida, tendo como etapa crucial a cromatografia de exclusão molecular do produto de digestão da proteína de fusão com a protease TEV (**Fig. 61**), mostrou-se apropriada para a purificação do fragmento N-terminal de 18 resíduos de pepC-term (18N_pepC-term). Estudos de CD com 18N_pepC-term indicaram que o peptídeo apresenta estrutura secundária em água deionizada e que o grau de estruturação é aumentado em soluções tamponadas em pH ácido (3,5) e próximo da neutralidade (7,5) (**Fig. 67**). Além disso, pela análise das curvas de CD nos dois valores de pH, pode-se inferir que a estrutura é distinta em cada condição. A estrutura de pepC-term também mostrou-se sensível à força iônica da solução.

A falta de similaridade com outras proteínas depositadas em bancos de dados (atualmente, mais de um milhão de seqüências) desperta especial interesse no domínio C-terminal de SmCD1. Primeiramente, a elucidação final da estrutura dos pouco menos de 50 resíduos compondo este domínio pode revelar uma conformação tridimensional única, já que não foi possível detectar semelhança significativa com nenhuma das mais de 20.000 estruturas protéicas já determinadas. Assim, a caracterização estrutural do domínio C-terminal de SmCD1 irá acrescentar um novo tipo de enovelamento ao

repertório já existente, ajudando assim a desvendar a estrutura de outras proteínas já sequenciadas, mas que não tiveram a estrutura 3D resolvida. Além disso, a disponibilidade de informações estruturais irá auxiliar na caracterização do papel deste domínio na função das formas longa e curta de SmCD1, um promissor alvo para o desenho de novos fármacos esquistossomicidas.

5.5.3 Potencial de exploração no diagnóstico da esquistossomose

O diagnóstico coprológico pela técnica de Kato-Katz (KATZ *et al.*, 1972) ainda é o método de escolha para o diagnóstico da esquistossomose mansônica. Entretanto, este método é pouco eficiente no diagnóstico de indivíduos com reduzida carga de vermes, tal como é freqüente em regiões não endêmicas. Por este motivo se persegue o desenvolvimento de um método diagnóstico mais sensível, utilizando técnicas sorológicas (DOENHOFF *et al.*, 2004) ou moleculares (RABELLO *et al.*, 2002). Em relação às técnicas imunológicas, têm-se utilizado desde frações brutas dos vermes em *immunoblots* (SULAHIAN *et al.*, 2005) ou ELISA (DE OLIVEIRA *et al.*, 2005) até anticorpos monoclonais ou antígenos puros em técnicas rápidas utilizando os chamados “dipsticks” (BOSOMPEM *et al.*, 2004) ou “reagent strips” (VAN DAM *et al.*, 2004).

O caráter único do domínio C-terminal de SmCD1, mesmo quando comparado às seqüências depositadas de outras parasitas próximos na escala evolutiva, motiva o estudo do potencial desta região no diagnóstico diferencial e na detecção prematura de anticorpos no soro de pacientes infectados com *S. mansoni*. Neste trabalho, foi relatada a produção deste peptídeo recombinante em *E. coli* e a efetiva purificação de um fragmento de 18 resíduos do N-terminal. O fragmento foi capaz de induzir anticorpos em coelhos (Fig. 68) indicando que pelo menos um epitopo B está presente neste segmento, conforme predito pela análise da seqüência com o algoritmo de Kyte-Doolittle (Fig. 69). Experimentos futuros utilizando pepC-term ou seu fragmento N-terminal de 18 resíduos imobilizados em placas de ELISA poderão verificar se anticorpos circulantes contra estes peptídeos estão presentes no soro de pacientes crônicos ou infectados na fase aguda da esquistossomose mansônica. Caso a presença destes anticorpos seja verificada, um novo método diagnóstico potencialmente específico para esta parasitose pode vir a ser desenvolvido baseado neste peptídeo.

6 CONCLUSÕES

- Estudos de modelagem molecular de SmCD1.2 evidenciaram que as preferências de substrato observadas experimentalmente para SmCD1.1 e hCD são dependentes de interações com os resíduos que compõem os subsítios de especificidade destas enzimas. As diferenças estruturais na fenda de ligação ao substrato de SmCD1 e hCD são significativas e podem ser exploradas no desenho de inibidores seletivos para a protease de esquistossomos;
- Estudos *in vivo* com inibidores comerciais de HIV-1 protease mostraram que estes são incapazes de reduzir a carga parasitária de camundongos infectados com *S. mansoni*;
- Cromatografia de afinidade em pepstatina-agarose não é capaz de ligar seletivamente SmCD1 e suas isoformas e revela a verdadeira complexidade das aspartil proteases de *S. mansoni* (SmAPs);
- Confirmando a análise da fração ligada à pepstatina-agarose, até 6 genes de SmAPs podem ser identificados em bancos de seqüências transcritas de diferentes estágios do crescimento de *S. mansoni*. Os genes de SmAPs denominados sequencialmente de SmCD2 até SmCD7 são significativamente menos representados nos bancos de seqüências transcritas de *S. mansoni*, podendo estar associados a funções diferenciadas no parasita. Entre os novos genes de SmAPs identificados *in silico*, SmCD3 e SmCD7 são os mais abundantes, com 20% do nível de expressão de SmCD1;
- Até o momento, nenhum outro gene de SmAPs parcialmente caracterizado compartilha a extensão C-terminal que define a chamada forma longa de SmCD1 (SmCD1.1). Os dados disponíveis não permitem concluir se existe uma forma curta de SmCD1 (SmCD1.2), como uma espécie distinta de SmCD1.1;
- Conforme esperado para uma aspartil proteinase, os estudos com as formas longa e curta de SmCD1 mostraram que estas são exclusivamente expressas como corpos de inclusão, no sistema de expressão em *E. coli* empregado neste

trabalho. Através das análises também ficou demonstrado que os corpos de inclusão de SmCD1 só podem ser eficientemente solubilizados em valores de pH na faixa de 6-8 em 5-6 M de guanidina-HCl, enquanto que 8 M de uréia na presença de agentes redutores pode ser uma alternativa em valores de pH > 8;

- A aplicação da técnica de planejamento estatístico de experimentos, conhecida como desenho fatorial fracionado, foi capaz de identificar condições favoráveis à obtenção de formas solúveis, porém inativas, de SmCD1.1 e SmCD1.2 após o processo de renaturação. Dentre os 10 fatores analisados, aqueles que se mostraram mais importantes nos tampões de renaturação de ambas as formas de SmCD1 foram: 1 M de guanidina-HCl, 500 mM de L-Arg e o par redox 1 mM GSH : 0,1 mM GSSG;
- O peptídeo correspondendo à extensão C-terminal de SmCD1.1 foi abundantemente expresso na forma de uma proteína de fusão com a proteína D de fago 1, em *E. coli*. Entretanto, um evento proteolítico provavelmente ocorrendo durante a expressão foi responsável pela diminuição do rendimento da preparação em até 50%. Uma estratégia de purificação eficiente foi desenvolvida para a purificação do fragmento N-terminal de 18 resíduos (18N_pepC-term) resultante desta clivagem proteolítica;
- Análises de CD mostraram que o peptídeo 18N_pepC-term apresenta-se estruturado em solução, sendo sensível a alterações no pH e na força iônica do meio. Resultados teóricos e experimentais sugeriram um grau de estrutura secundária de até 40% neste peptídeo. Estudos de modelagem computacional sugeriram uma estrutura em forma de grampo de cabelo torcido para o segmento completo da extensão C-terminal de SmCD1.1;
- Os produtos dos genes das formas longa e curta de SmCD1 não puderam ser detectados nos sobrenadantes nem na fração celular de clones de *P. pastoris* GSS15 previamente confirmados quanto à integridade do cassete de expressão. Análises preliminares sugeriram que esta observação está possivelmente associada à uma terminação prematura da transcrição ou a problemas na tradução, devido ao alto conteúdo AT das seqüências nucleotídicas destes genes;

- Soros policlonais construídos contra as proteínas recombinantes expressas em *E. coli* mostraram-se específicos contra os antígenos empregados na imunização de coelhos. O soros contra as formas longa e curta de SmCD1 reconheceram bandas majoritárias com cerca de 41 e 47 kDa no extrato solúvel de vermes adultos de *S. mansoni*. O soro contra o peptídeo 18N_pepC-term, por sua vez, não foi capaz de reagir contra a forma longa recombinante, nem contra espécie alguma presente nos extratos do verme.

Capítulo 2

Serininil proteases ativas sobre a coagulação sanguínea

1 INTRODUÇÃO

1.1 *Estrutura e função de serinil proteases da família da quimiotripsina*

1.1.1 Aspectos gerais e enovelamento

Enzimas proteolíticas dependentes de um resíduo de serina para a atividade catalítica são amplamente disseminadas e abundantes (RAWLINGS & BARRET, 1994). De acordo com as informações reunidas no banco de dados MEROPS, serinil peptidases são encontradas em vírus, bactérias, fungos, protozoários, plantas e animais (RAWLINGS *et al.*, 2004). Este grupo de enzimas inclui exo-, endo-, oligo- e omegapeptidases. Originalmente, Rawlings & Barret (1994) vislumbraram 4 origens evolutivas distintas (formando 4 diferentes clãs) para serinil peptidases com base nas estruturas tridimensionais das enzimas quimiotripsina (clã SA), subtilisina (clã SB), carboxipeptidase C (clã SC) e D-Ala-D-Ala peptidase A de *E. coli* (clã SE).

Atualmente, o clã da quimiotripsina é classificado no MEROPS como um subclã [PA(S)] do clã PA, composto por 2 tipos catalíticos distintos (serinil e cisteinil peptidases). Os membros deste clã compartilham um enovelamento comum e acredita-se que os subclãs tenham divergido de um ancestral comum bastante antigo (RAWLINGS *et al.*, 2004). Todos os membros do subclã PA(S) empregam dois componentes no mecanismo catalítico: a tríade His⁵⁷, Asp¹⁰², Ser¹⁹⁵ e o chamado sítio de ligação ao oxianion. A tríade catalítica localiza-se no centro da molécula e é parte de uma extensiva rede de ligações de hidrogênio. O sítio de ligação ao oxianion é formado pelos grupos NH da cadeia principal de Gly¹⁹³ e Ser¹⁹⁵. Estes átomos formam uma cavidade com carga positiva que estabiliza o intermediário tetraédrico negativamente carregado que se forma após ataque da carbonila da ligação lábil (vide próximo item).

Dezoito famílias distintas são agrupadas no clã PA(S), entre elas a família S1, contendo a própria quimiotripsina A bovina, a protease prototípica do clã PA. Esta família é a maior de todas as famílias de peptidases contidas no MEROPS, contendo até o momento 2666 seqüências protéicas divididas em 6 subfamílias. Todas as peptidases caracterizadas da família S1 são endopeptidases (proteínases). Há três tipos principais de “atividades” ou “especificidades”: “tripsina-símile”, onde há clivagem de substratos peptídicos após um resíduo de Arg ou Lys em P1; “quimiotripsina-símile”, onde a

clivagem ocorre após aminoácidos hidrofóbicos (Phe/Tyr/Trp) em P1 e “elastase-símile” com a clivagem ocorrendo após uma Ala/Val em P1 do substrato (RAWLINGS & BARRET, 1994). A especificidade de substrato na família S1 depende primordialmente do resíduo ocupando a posição P1. A maior parte das peptidases desta família entra na via de secreção e possui um peptídeo sinal N-terminal. Estas são sintetizadas como precursores com uma extensão N-terminal (propeptídeo) que é clivada para formar a enzima ativa. Quatro segmentos formam o chamado domínio de ativação em enzimas quimiotripsina-símile: do N-terminal ao resíduo 19, além dos resíduos 142-152, 184-193 e 26-223. O domínio de ativação contém o subsítio S1 e o sítio de ligação ao oxianion e apresenta-se deformado no zimogênio inativo. Entretanto, a ativação nem sempre requer que o propeptídeo seja removido (RAWLINGS & BARRET, 1994). Nos fatores da coagulação sanguínea, o propeptídeo permanece ligado por uma ligação dissulfeto à cadeia pesada que é gerada após a clivagem. A clivagem do precursor é essencial para permitir um rearranjo estrutural no novo N-terminal exposto (usualmente Ile¹⁶ ou Val¹⁶), que se liga ao aspartato próximo à Ser¹⁹⁵ catalítica no motivo G¹⁹³DSGG¹⁹⁷ (vide Fig. 81). O N-terminal varia em tamanho, desde apenas 2 resíduos na catepsina G até segmentos mais longos, contendo centenas de resíduos.

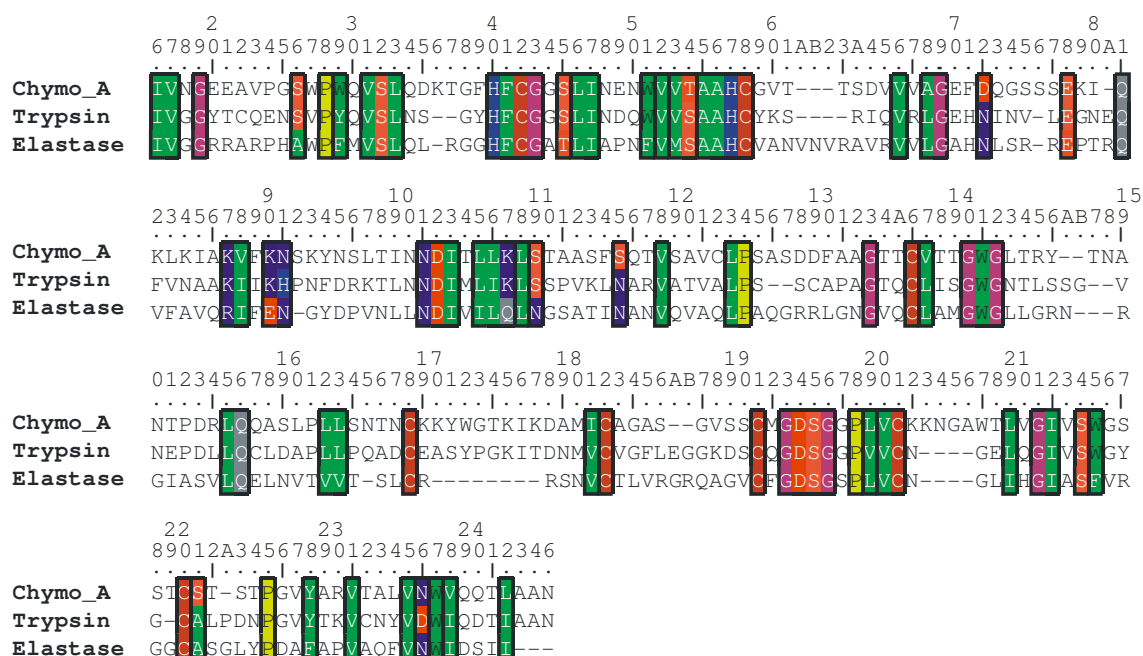


Fig. 81: Alinhamento estrutural de serinil proteases pertencentes à família S1A. Chymo_A = alfa-quimiotripsina bovina (PDB 5CHA), Trypsin = tripsina aniônica de rato II (PDB 1ANE) e Elastase humana de neutrófilo (PDB 1H1B). O alinhamento é numerado de acordo com a sequência do quimiotripsinogênio. As posições do alinhamento com aminoácidos conservados nas três proteínas são coloridas de acordo com a matriz de escore BLOSUM62 (HENIKOFF, 1992).

Serinil proteases da família S1A contém dois domínios de estrutura similar, os quais se acreditam terem surgido por duplicação de gene e divergência (RAWLINGS & BARRET, 1994). Os domínios empacotam-se assimetricamente, com o sítio catalítico disposto entre eles e cada domínio contém uma folha- β antiparalela, enovelada em um barril- β (Fig. 82). Lesk & Fordham (1996) analisaram as estruturas de 13 enzimas da família S1A e identificaram regiões do enovelamento que foram preservadas através da evolução, tanto internas a cada domínio como fazendo parte da interface. Na maior parte das enzimas de mamíferos, um curto segmento do N-terminal forma parte do domínio C-terminal (domínio 2) e um curto segmento do C-terminal forma parte do domínio N-terminal (domínio 1); isto é, cada domínio contém duas seções contíguas e a cadeia (na estrutura intacta imatura) cruza entre os domínios três vezes. A trombina é uma exceção; uma longa alça de um domínio estende-se no outro domínio. Os domínios da trombina, cada um possui três seções separadas da cadeia polipeptídica. Desta forma, a cadeia polipeptídica da proteinase imatura intacta cruza entre os domínios 4 vezes.

Cada domínio contém duas repetições de um motivo consistindo de três fitas- β antiparalelas conectadas por dois grampos- β . O primeiro grampo é dobrado em um ângulo de aproximadamente 90° . Os dois motivos são conectados por ligações de hidrogênio, cada um contendo dois grampos- β . Os primeiros grampos- β de cada motivo formam ligações de hidrogênio entre si enquanto que os segundos grampos- β de cada motivo o fazem entre si, fechando assim o barril- β . Os domínios catalíticos das proteinases diferem entre si nos comprimentos das fitas das folhas- β compondo as partes comuns do enovelamento e no comprimento e conformações das regiões conectando as fitas (i.e., as alças). A estrutura da interface é caracterizada por uma região central conservada próxima à tríade catalítica, cercada por regiões adicionais que variam entre as proteinases. Os resíduos da região central provêm de seis fitas, três de cada domínio, incluindo as fitas no domínio contendo os resíduos catalíticos His e Asp e a fita no domínio 2 contendo o resíduo de Ser catalítico.

A bolsa de especificidade primária de serinil proteases ou subsítio S1 é responsável por receber a cadeia lateral na posição P1, sendo adjacente ao sítio ativo. As laterais desta bolsa são formadas por duas alças conectando fitas sucessivas da folha- β do 2º domínio (4ª e 5ª fitas- alça 1 e 5ª e 6ª fitas- alça 2, vide Fig. 82D). As alças 1 e 2 também são designadas por “alça-187” e “alça-220s, respectivamente, indicando as suas

posições na estrutura. A base da bolsa é ocupada pela posição 189 (Ser em quimi tripsina, Asp em tripsina e Gly na elastase, vide [Fig. 81](#)), sendo a contribuição do resíduo ocupando esta posição fundamental para a especificidade da proteinase ([PERONA & CRAIK, 1995](#)). Os resíduos nas posições 216 e 226 também influenciam na forma da cavidade S1 e contribuem para a especificidade das enzimas, principalmente na elastase, onde resíduos diferentes de Gly nestas posições restringem o tamanho da cadeia lateral podendo ocupar o subsítio.

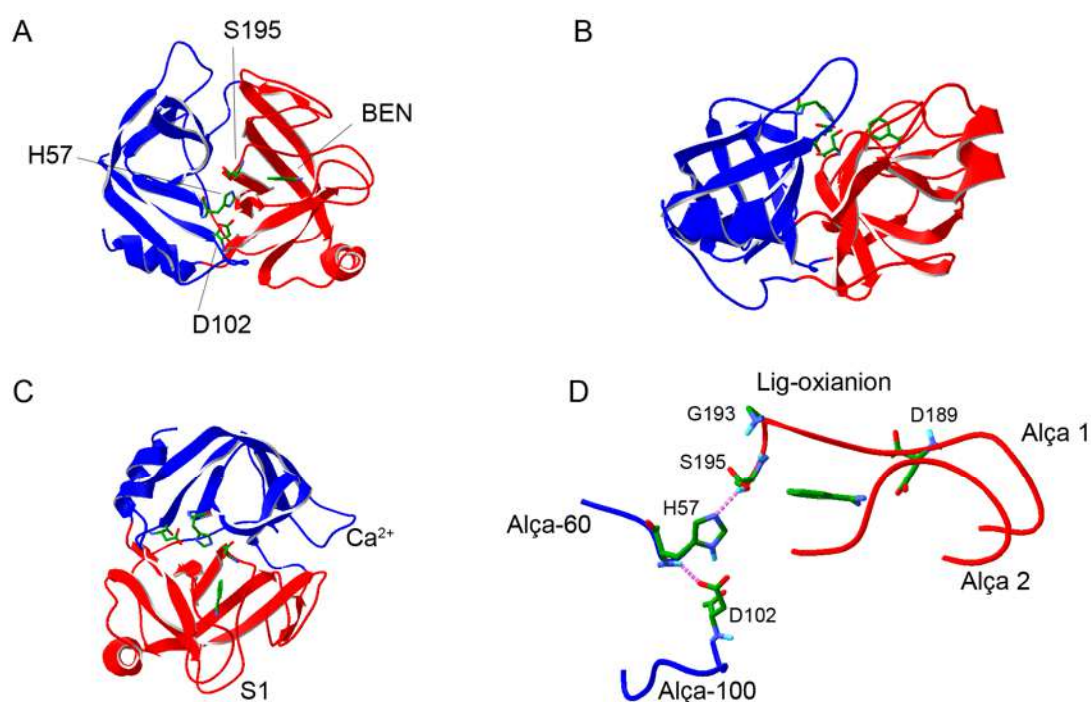


Fig. 82: Enovelamento de enzimas da família da quimi tripsina. A estrutura da tripsina aniônica de rato ligada à benzamidina (PDB 1ANE) é mostrada com os elementos de estrutura secundária (*ribbons*) coloridos em azul no domínio N-terminal e em vermelho no C-terminal. A- vista de cima da fenda catalítica destacando os resíduos catalíticos e o inibidor ligado. B- vista ao longo da fenda de ligação ao substrato obtida girando-se 90° a estrutura em A (notar disposição perpendicular dos barris- β). C- vista semelhante a A, seguindo a “orientação de Bode” ([BODE & SCHWAGER, 1975](#)), com o subsítio S1 ao sul, o sítio de ligação ao cálcio ao nordeste e a fenda de ligação do substrato correndo de oeste para leste. D- detalhamento das alças compondo o sítio catalítico e subsítio S1 (alça 1 = 185-188 e alça 2 = 217-225), mostrando as ligações de hidrogênio (pontilhado) da tríade catalítica.

[Krem et al. \(1999\)](#) analisaram a seqüência dos domínios catalíticos de 90 membros da família S1A e concluíram que o domínio C-terminal tem papel fundamental na definição da função da serinil protease. Esta porção do domínio catalítico constitui a maior parte da superfície acessível dos subsítios S1-S3. Em particular, os resíduos ocupando as últimas 50 posições do domínio C-terminal (187-

237) são críticos para determinar a função de enzimas da família S1A (ROSE & DI CERA, 2002), incluindo: o resíduo 189 importante para a especificidade no subsítio S1; o resíduo 193 do sítio de estabilização do oxianion; o resíduo catalítico Ser¹⁹⁵; os resíduos 216, 225 e 226, que controlam o acesso e tem papel crucial na arquitetura do subsítio S1; e o Trp²³⁷ contido na hélice C-terminal que ajuda a estabilizar a alça-100 contendo o resíduo catalítico Asp¹⁰². Entretanto, deve-se ressaltar que alguns resíduos importantes para a função de proteases da família da quimiotripsina estão localizados no domínio N-terminal, como os resíduos catalíticos His⁵⁷ e Asp¹⁰², e os resíduos compondo a alça-60, os quais restringem o acesso aos subsítios S2 e S3 em serinil proteínases da coagulação e do sistema complemento.

1.1.2 Mecanismo catalítico e cinética enzimática

A Fig. 83 mostra o mecanismo comumente aceito para a etapa de acilação da hidrólise da ligação peptídica, catalisada por serinil proteases da família S1A (revisado por HEDSTROM, 2002). Após a formação do complexo de Michaelis-Menten entre enzima e substrato, o resíduo de Ser¹⁹⁵ no sítio catalítico ataca a carbonila da ligação peptídica entre os resíduos P1 e P1'. O ataque nucleofílico é assistido por His⁵⁷, atuando como base geral, fornecendo o intermediário tetraédrico. O íon His-H⁺ resultante é estabilizado por ressonância dentro do anel imidazólico e pela ligação de hidrogênio com Asp¹⁰². O oxianion do intermediário tetraédrico é estabilizado pela interação com os grupos NH de Ser¹⁹⁵ e Gly¹⁹³. O arranjo de ligações de hidrogênio presente na tríade catalítica de serinil proteases, permitindo a estabilização da carga que se forma no estado de transição do ataque nucleofílico à ligação peptídica, ficou conhecido como “sistema de relé de carga”. Na etapa subsequente do mecanismo catalítico, o intermediário tetraédrico entra em colapso e expulsa o grupo de saída contendo os resíduos P' do substrato. Isto é facilitado por His-H⁺ que atua como um ácido geral para protonar a amina do grupo de saída, culminando na formação do intermediário acil-enzima.

A etapa de desacilação da reação essencialmente repete a seqüência acima: uma molécula de água ataca a acil-enzima, assistida por His⁵⁷, fornecendo um segundo intermediário tetraédrico. Este último se decompõe na Ser¹⁹⁵ livre e no ácido carboxílico liberado. A existência do intermediário acil-enzima é bem documentada por estudos cristalográficos, mas os intermediários tetraédricos foram inferidos de estudos em

solução (HEDSTROM, 2002). Acredita-se que os estados de transição para as reações de acilação e desacilação assemelham-se aos intermediários tetraédricos altamente energéticos. Desta forma, os termos “estado de transição” e “intermediário tetraédrico” são frequentemente utilizados indiscriminadamente na literatura.

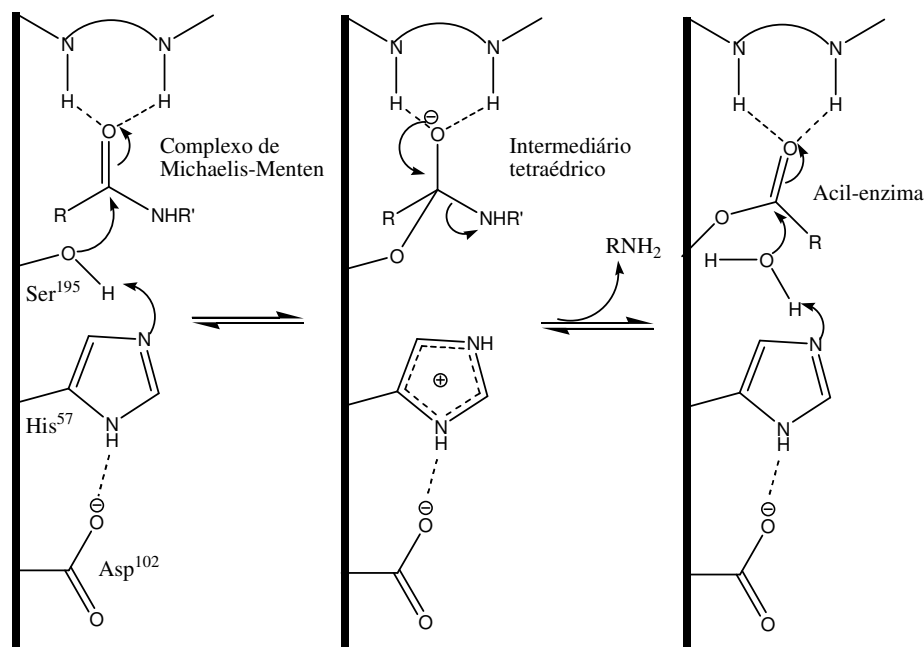
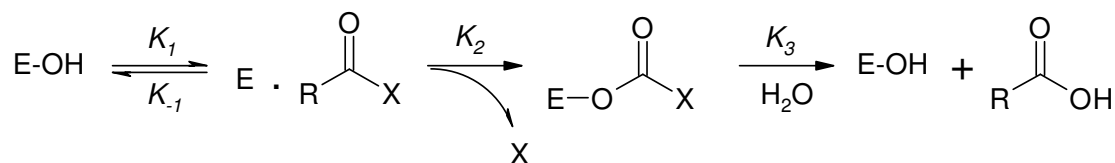


Fig. 83: Mecanismo catalítico em serinil proteases da família da quimiotripsina.

Geralmente considera-se que a reação catalisada por serinil proteases possui um mecanismo cinético de três etapas (Fig. 84): (i) formação do complexo enzima-substrato (ES; $K_s = k_{-1}/k_1$); (ii) acilação da serina do sítio ativo (k_2) e (iii) hidrólise do intermediário acil-enzima (k_3). Deve-se notar que alguns pontos deste mecanismo são aproximações ainda sujeitas ao debate. Primeiramente, k_2 e k_3 são constantes de velocidade compostas que incluem uma transformação química e uma etapa de dissociação do produto. É geralmente assumido que esta última é rápida relativamente à transformação química, baseado em resultados de efeitos isotópicos do solvente, mas estes experimentos não constituem prova cabal de que uma etapa é limitante num mecanismo cinético (HEDSTROM, 2002). Os parâmetros de Michaelis-Menten k_{cat} , K_m e k_{cat}/K_m são composições de constantes cinéticas, conforme mostrado na Fig. 84. Valores crescentes de k_2 irão aumentar K_m até $k_2 \geq k_3$, porém o aumento de k_2 acima deste limite irá decrescer o valor de K_m . Desta forma, pode ser extremamente difícil interpretar K_m em relação à afinidade de substrato. Além disso, as expressões na Fig. 84 são frequentemente simplificadas assumindo que $k_{-1} \gg k_2$. Esta premissa é

provavelmente válida para substratos pouco específicos, do contrário o processo é controlado por difusão e $k_2 \geq k_{-1}$ (HEDSTROM, 2002). Outra generalização sobre este mecanismo que é feita indiscriminadamente é que “a acilação é a etapa limitante da hidrólise de amidas ($k_3 \gg k_2$), enquanto que a desacilação é a etapa determinante da velocidade de hidrólise de ésteres por serinil proteases ($k_2 \gg k_3$)”. Entretanto, diversas evidências experimentais sugerem maior cautela e cada substrato deve ser analisado especificamente (HEDSTROM, 2002).



$$k_{\text{cat}} = k_2 \cdot k_3 / (k_2 + k_3)$$

$$K_m = K_3(k_{-1} + k_2) / k_1(k_2 + k_3)$$

$$k_{\text{cat}} / K_m = k_1 \cdot k_2 / (k_1 + k_2)$$

$$\text{Se } k_{-1} \gg k_2,$$

$$K_m = k_{-1} \cdot k_3 / k_1(k_2 + k_3) = K_s \cdot k_3 / (k_2 + k_3)$$

$$k_{\text{cat}} / K_m = k_1 \cdot k_2 / k_{-1} = k_2 / K_s$$

$$\text{Se } k_3 \gg k_2,$$

$$K_m = (k_{-1} + k_2) / k_1$$

Fig. 84: Mecanismo cinético da hidrólise da ligação peptídica catalisada por serinil proteases da família S1A.

O conhecimento detalhado do mecanismo cinético de hidrólise por serinil proteases permitiu que as hipóteses referentes ao mecanismo estrutural pudessem ser reavaliados. O resíduo Ser²¹⁴ foi anteriormente implicado na estabilização de Asp¹⁰² através da geração de um ambiente polar mais favorável na forma de ligações de hidrogênio ancorando 2 moléculas do solvente (HEDSTROM, 2002). Estudos de mutação sítio dirigida não sustentaram esta hipótese, mostrando que a substituição de Ser²¹⁴ por Ala afeta, na verdade, positivamente a catálise por tripsina e trombina (KREM *et al.*, 2002a). De fato, nenhum dos demais mutantes de trombina preparados (Ser²¹⁴ → Ala, Thr, Cys, Asp, Glu e Lys) apresentou comprometimento significativo da função catalítica. No entanto, as mutações menos conservativas resultaram em grandes aumentos em K_m devido às menores velocidades de difusão do substrato no sítio ativo. Assim, o papel de Ser²¹⁴ parece ser o de promover a formação produtiva do complexo enzima-substrato. Mais recentemente, o papel de Gly¹⁹³ no sítio de ligação ao oxianion

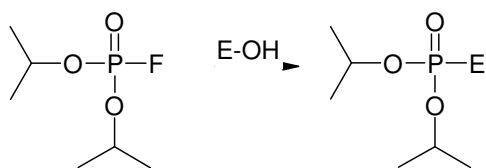
foi reavaliado pela combinação de estudos cinéticos detalhados e mutação sítio-dirigida (BOBOFCHAK *et al.*, 2005). Os dados relatados neste estudo revelaram que Gly¹⁹³ contribui praticamente de maneira idêntica para a estabilidade do estado fundamental e de transição durante o ciclo catalítico. Isto põe em questão o consenso atual sobre o papel do sítio de ligação ao oxianion na catálise efetuada por serinil proteases.

1.1.3 Mecanismos de inativação

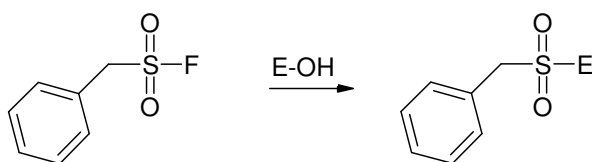
1.1.3.a Inibidores covalentes sintéticos

Inibidores de serinil proteases podem ser divididos em duas grandes classes: irreversíveis (ou covalentes) e reversíveis. Inicialmente, inibidores covalentes como DFP (diisopropilfluorofosfato) e TPCK (tosil-L-fenilalanil clorometilcetona) foram usados para revelar as primeiras pistas sobre o mecanismo catalítico de serinil proteases. DFP inativa enzimas da família da quimiotripsina ao modificar especificamente a Ser¹⁹⁵, enquanto que TPCK reage com His⁵⁷ da tríade catalítica (Fig. 85). Outros compostos eletrofílicos (aldeídos peptídicos, trifluorometilcetonas e compostos similares) formam adutos tetraédricos com serinil proteases ao serem atacados por Ser¹⁹⁵ (Fig. 85). A especificidade destes reagentes é adaptada a uma protease em especial ao combinar os grupos reativos com as porções peptídicas apropriadas (POWERS *et al.*, 2002). Clorometilcetonas como TPCK e TLCK (derivado de lisina) são potentes inibidores de enzimas empregando um nucleófilo ativado no mecanismo catalítico, tal como cisteinil e serinil proteases. O grupo reativo de clorometilcetonas inativa serinil proteases ao formar um hemiacetal com Ser¹⁹⁵ e subsequentemente alquilar His⁵⁷ (Fig. 85C). A reatividade de alguns compostos atacados por serinil proteases foi modificada para fornecer intermediários acil-enzima mais estáveis. Por exemplo, as velocidades de desacilação decrescem pela presença de substituintes eletrodoadores ou pela presença de heteroátomos adjacentes à carbonila de ésteres (HEDSTROM, 2002). Esta menor reatividade é usualmente contrabalanceada pela introdução de melhores grupos de saída, e.g., p-nitrofenila (Fig. 85F). De fato, guanidinobenzoato de p-nitrofenila (NPGb) é muito utilizado como titulante de sítio-ativo de serinil proteases tripsina-símile devido sua alta reatividade contra estas enzimas (KNIGHT, 1995).

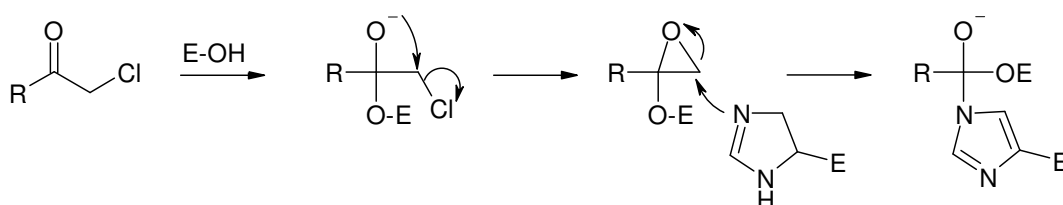
A. Reação com DFP



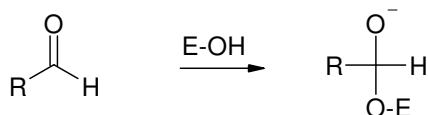
B. Reação com PMSF



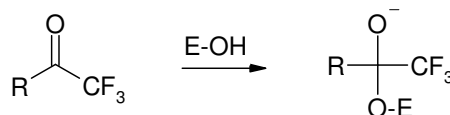
C. Reação com clorometilcetonas



D. Reação com aldeídos



E. Reação contra trifluorometilcetonas



F. Reação com guanidinobenzoato de p-nitrofenila (pNP)

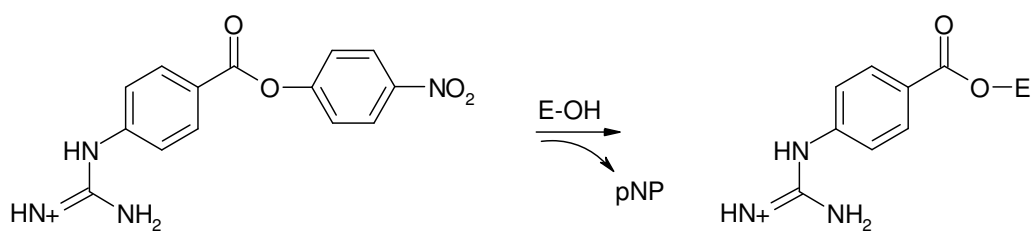


Fig. 85: Mecanismos de inativação de serinil proteases da família da quimiotripsina por modificação covalente com compostos sintéticos. Nos esquemas, “E” denota um grupo químico da enzima.

1.1.3.b Inibidores covalentes naturais: SERPINS

Uma classe de inibidores protéicos de serinil proteases, conhecidos por SERPINS, também atua modificando covalentemente o resíduo de Ser¹⁹⁵ (GETTINS, 2002). Estes inibidores atuam como verdadeiras “ratoeiras moleculares”, sofrendo uma

mudança conformacional ao serem atacados pela protease que culmina na deformação do sítio ativo da enzima e conseqüentemente na sua inativação. A chave para a base estrutural do mecanismo de ação de SERPINS reside numa alça conectando duas folhas- β na molécula. Esta última, conhecida como alça do sítio reativo (ASR), consiste num segmento flexível de 17 resíduos, que se situa na superfície da conformação nativa de SERPINS. A **Fig. 86** ilustra as principais etapas do mecanismo suicida adotado por SERPINS. Após a formação do complexo não-covalente (EI), o mecanismo pode ser visto como uma via bifurcada, onde o intermediário acilado ($EI^\#$) pode resultar no complexo inibido (EI^+) ou no inibidor clivado (I^*) e na enzima livre (E), através da via produtiva. Ao sofrer a clivagem, a ASR em SERPINS sofre rearranjo conformacional e se insere na folha- β na base da molécula do inibidor. Acredita-se que o ganho de energia resultante dessa nova conformação é canalizado para o aprisionamento da proteinase numa conformação inativa (**GETTINS, 2002**). A via inibitória é cineticamente favorecida ($k_4 < k_5$), pois as mudanças conformacionais impostas à protease desaceleram em até 100 milhões de vezes a velocidade de desacilação (k_5).

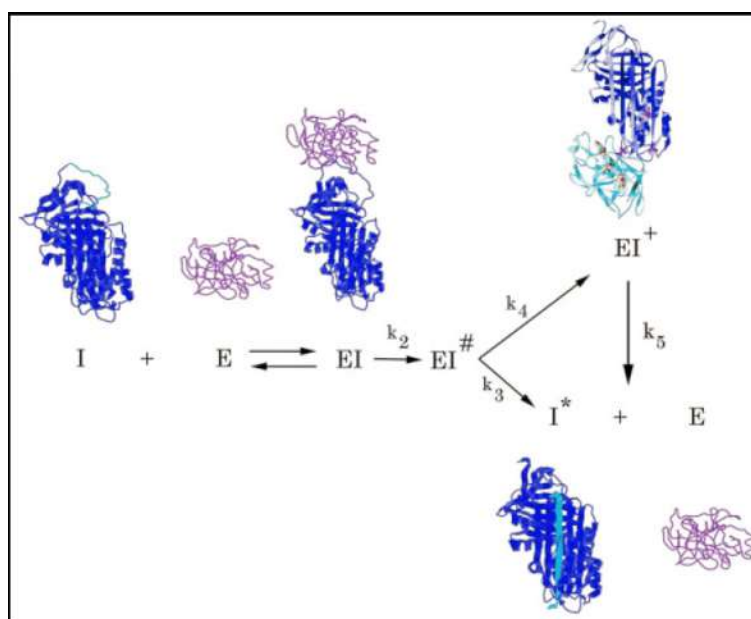


Fig. 86: Mecanismo cinético e aspectos estruturais da inibição por SERPINS. O inibidor é mostrado em azul escuro com a alça do sítio reativo (ASR) em azul claro. A proteinase é mostrada em lilás. A clivagem da ligação lábil em SERPIN proporciona uma modificação conformacional na ASR que se insere na folha- β na base da molécula. Quando $k_4 < k_3$, a proteinase acilada é comprimida contra a base de SERPIN provocando uma deformação que inativa a enzima. Do contrário, a via não-inibitória (ou produtiva) é seguida e há regeneração da proteinase nativa e produção de SERPIN clivado. Adaptado de **Silverman et al. (2001)**. Para maiores detalhes vide texto.

1.1.3.c Inibidores não-covalentes protéicos

A segunda grande classe de inibidores de serinil proteases atua associando-se não-covalentemente à enzima, através de mecanismos competitivos (bloqueio do sítio ativo) ou não-competitivos (modulação alostérica). Pequenas moléculas sintéticas normalmente atuam competindo pelo sítio de ligação ao substrato, enquanto que macromoléculas naturais podem também atuar em sítios distantes do sítio catalítico (BODE & HUBER, 2000). Os inibidores protéicos de serinil proteases constituem uma classe extremamente diversificada com pelo menos 18 famílias distintas, incluindo os tipos Kazal, Kunitz e Bowman-Birk (LASKOWSKI & QASIM, 2000). Estes se ligam à proteinase através de uma alça exposta à superfície que faz ligações de hidrogênio complementares à fenda de ligação ao substrato (BODE & HUBER, 1992).

Os resíduos P1-P3 do inibidor interagem analogamente a um substrato, formando uma folha- β com os resíduos 214-216 da enzima (Fig. 87). Diferenças na composição da alça canônica fornecem especificidade para as diferentes enzimas da família S1A (HELLAND *et al.*, 2003). Apesar dos enovelamentos distintos adotados nas diferentes famílias de inibidores, a conformação com que essa alça liga-se à proteinase é conservada, sendo, portanto conhecida como “conformação canônica” (APOSTOLUK & OTLEWSKI, 1998). Entre 50-70% da energia de ligação à proteinase provém da interação do resíduo ocupando a posição P1 do inibidor com o subsítio complementar. Os resíduos distais ao segmento P5-P3' do inibidor fazem contatos secundários com a proteinase que podem auxiliar na estabilização do complexo enzima-inibidor (BODE & HUBER, 2000). Os inibidores canônicos naturais de serinil proteases servem como um paradigma para o reconhecimento substrato-protease, uma vez que suas alças expostas na superfície ligam-se às fendas catalíticas de maneira similar aos substratos naturais.

Os valores de K_i de inibidores protéicos naturais correlacionam-se muito bem com os valores de k_{cat}/K_m para os substratos análogos, indicando que as alças canônicas ligam-se às serinil proteases mimetizando o estado de transição da reação de hidrólise da ligação peptídica (HEDSTROM *et al.*, 1994a; LASKOWSKI & QASIM, 2000). Em vista disso, tem sido proposto que a conformação canônica pode ser uma propriedade destes inibidores que se assemelham ao estado de transição de substratos protéicos, representando uma conformação de alta energia que ativa a ligação lábil para a

hidrólise. Entretanto, está ainda é uma questão em aberto, uma vez que [Coombs *et al.* \(1999\)](#) demonstraram que duas ligações de hidrogênio envolvidas no modo de ligação canônico podem ser perturbadas sem prejudicar a afinidade de ligação aos inibidores. Estes autores propuseram que o modo de ligação de inibidores protéicos naturais de serinil proteases representa uma interação similar ao modo de ligação do substrato, porém não-produtivo.

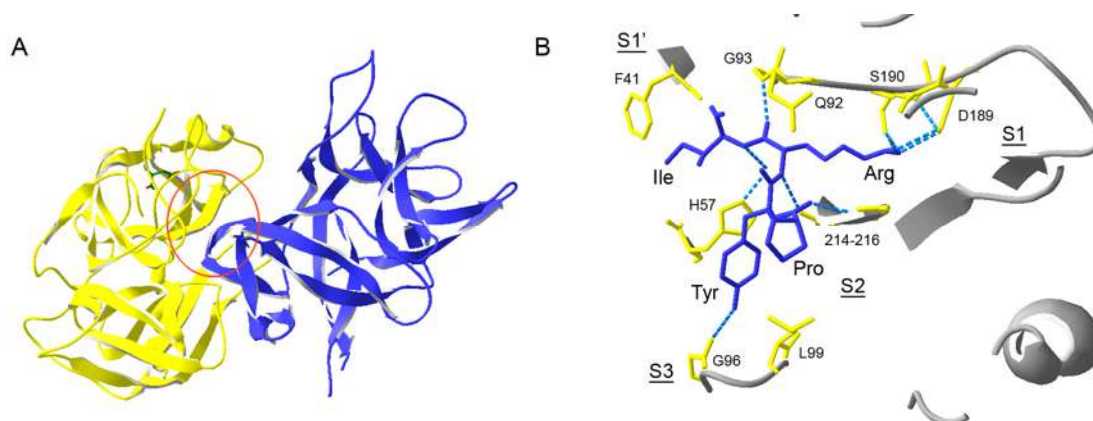


Fig. 87: Modo de ligação canônico de inibidores protéicos reversíveis de serinil proteases conforme ilustrado pelo complexo entre tripsina suína e SBTI (PDB 1AVW). A- orientação de SBTI (azul) ligado através da alça canônica à tripsina. B- detalhamento da região em destaque em A, mostrando os resíduos de tripsina fazendo contato com P1-P3 de SBTI.

1.1.3.d Inibidores não-covalentes sintéticos: afinidade de ligação ao subsítio S1

Enzimas com atividade tripsina-símile estão envolvidas nos mais diversos processos biológicos, incluindo coagulação sanguínea, imunidade celular e humoral, remodelagem tecidual, fertilização, fibrinólise e embriogênese ([KREM & DI CERA, 2002b](#)). Assim, a identificação de moléculas capazes de modular reversivelmente a atividade destas enzimas apresenta grande interesse. Em um trabalho seminal, [Mares-Guia & Shaw \(1965\)](#) propuseram que derivados alifáticos e aromáticos de amidinas e guanidinas atuam como inibidores competitivos da tripsina ao mimetizarem a cadeia lateral da arginina, presente na posição P1 de substratos. O modo de interação destes inibidores foi interpretado em termos de interações com um sítio aniônico e com um sítio hidrofóbico na enzima. Trabalhos subsequentes ressaltaram a característica lipofílica do sítio de interação, baseados na natureza hidrofóbica requerida para os derivados de amidina e guanidina mais potentes ([MARES-GUIA & FIGUEIREDO, 1970](#)).

A disponibilidade da estrutura cristalográfica da tripsina permitiu que as bases estruturais para a interação de amidinas e guanidinas fossem reveladas (BODE & SCHWAGER, 1975). A molécula de benzamidina encaixa-se no subsítio S1 da protease fazendo uma ponte salina com Asp¹⁸⁹ através do grupo amidínio. De posse das novas informações estruturais, Mares-Guia *et al.* (1977) conduziram um estudo de QSAR que racionalizou a atividade de derivados de benzamidina substituídos na posição 4 em termos do caráter eletrônico dos substituintes, empregando as constantes sigma (σ) de Hammet (1935). Estes autores apontaram a importância de substituintes doadores de elétrons (ERS, sigla inglesa) na posição 4 do anel aromático de forma a maximizar a potencia inibitória através da formação de uma interação dipolar adicional com o átomo O γ do resíduo catalítico Ser¹⁹⁵ (Fig. 88).

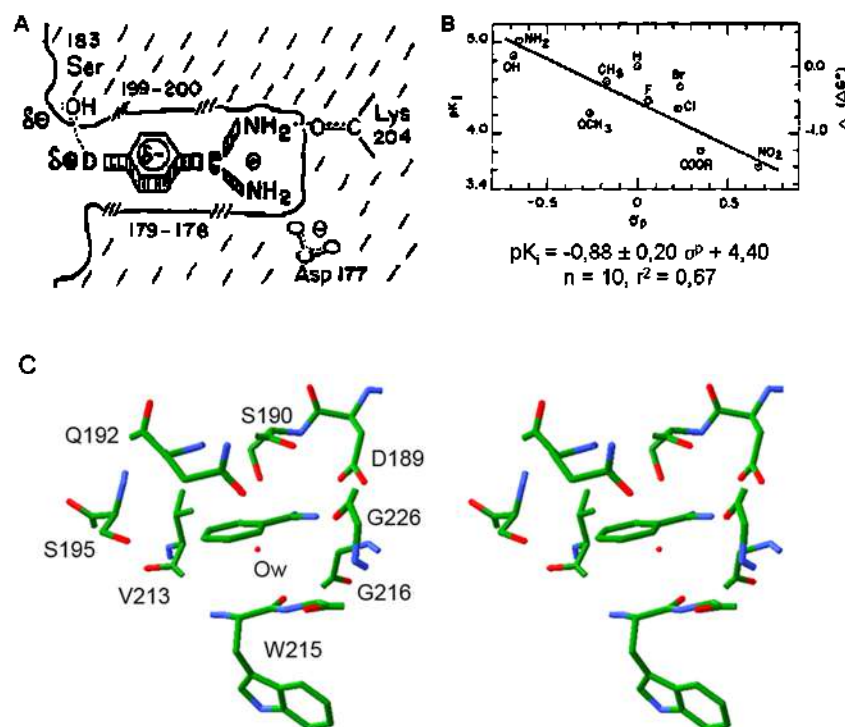


Fig. 88: Sítio de ligação e modo de interação de derivados de benzamidina com enzimas tripsina-símile. A- esquema do modelo de interação dipolar original proposto por Mares-guia e colaboradores. B- correlação entre a constante σ de Hammet e potencia (pK_i) para derivados substituídos na posição *para* de benzamidina (MARES-GUIA *et al.*, 1977). C- vista estereoscópica dos resíduos do subsítio S1 fazendo contato com benzamidina no complexo com tripsina (PDB 1ANE).

O modelo da interação dipolar de Mares-Guia perdurou por muitos anos como um importante paradigma na ligação de benzamidinas a proteases tripsina-símile. Na década de 80, o grupo de Corwin Hansch (RECONTINI *et al.*, 1986) propôs um modelo

de QSAR onde termos hidrofóbicos e estéricos foram adicionados de modo a melhorar a correlação com os dados de inibição: $pK_i = -0,45 \pm 0,45 \sigma + 0,41 \pm 0,11 \text{LogP} - 1,07 \pm 0,30$ ($n = 8$, $r^2 = 0,955$, $s = 0,088$). Interessantemente, estes autores propuseram outro papel para o caráter necessariamente eletrofílico do ERS. Este seria o de restabelecer a hidrofobicidade do anel aromático, diminuída pelo caráter retirador de elétrons do grupamento benzamidina protonado.

Publicações recentes, em geral, apóiam a importância da hidrofobicidade e de interações não-polares em inibidores direcionados ao subsítio S1 de enzimas tripsina-símile (ORTIZ *et al.*, 1995; ESSEX *et al.*, 1997; TALLHOUT & ENGBERTS, 2001; DULLWEBER *et al.*, 2001; GUIMARÃES & DE ALENCASTRO, 2002). Este é um fato curioso, uma vez que o subsítio S1 provê muitos grupos aceptores e doadores capazes de participar de interações direcionais mais específicas com grupamentos correspondentes nas estruturas dos inibidores (vide Fig. 88C). O grupo de Jorgensen conduziu uma simulação de Monte Carlo a fim de computar as afinidades de ligação de benzamidinas substituídas na posição 4 pela tripsina, a qual ajudou a esclarecer esta aparente contradição (ESSEX *et al.*, 1997). A simulação revelou que ao se trocar o substituinte 4-amino pelo grupo metil ou pelos átomos de cloro ou hidrogênio não se introduz ganho líquido algum no número de ligações de hidrogênio formadas entre grupos funcionais do inibidor e grupos na superfície de S1. Isto ocorre porque o número de potenciais ligações de hidrogênio perdidas ao se passar os inibidores da fase aquosa para a bolsa S1 é constante entre os inibidores estudados. Assim, o número total de ligações de hidrogênio envolvendo átomos da proteína no sítio ativo é praticamente independente do inibidor, portanto ligações de hidrogênio não podem ser invocadas para racionalizar a tendência nas afinidades de ligação.

Baseados puramente em princípios do contínuo eletrostático, alguns autores (ESSEX *et al.*, 1997; TALHOULT & ENGBERTS, 2001) sugeriram que a substituição por ERS decresce o caráter dipolar de benzamidinas 4-substituídas, diminuindo por sua vez a penalidade na energia livre ao se dessolvatar estes inibidores mediante associação com a enzima. Entretanto, quando alguma medida do caráter ERS dos substituintes, como as constantes sigma, são grafadas contra pK_i somente uma fraca correlação é obtida (MARES-GUIA *et al.*, 1977; MAGALHÃES *et al.*, 1997; TALLHOUT & ENGBERTS, 2001). Desta forma, independentemente da interpretação empregada,

parece que as propriedades eletrônicas do substituinte em derivados de benzamidina não podem ser utilizadas para racionalizar as afinidades de ligação com enzimas tripsina-símile.

Atualmente, é possível questionar se modelos combinando constantes sigma e descritores de hidrofobicidade, tais como LogP (RECONTINI *et al.*, 1986), somente melhoraram a correlação com os dados de atividade devido a colinearidade dos descritores empregados. De fato, tem sido discutido que como existe apenas uma fraca relação entre LogP e atividade inibitória, não se deveria racionalizar as afinidades de ligação de inibidores direcionados ao S1 em termos de sua hidrofobicidade (MAGALHÃES *et al.*, 1997; TALHOUT & ENGBERTS, 2001). No entanto, tem-se reunido evidências da importância do efeito hidrofóbico (SHARP & MADAM, 1997) na ligação de inibidores direcionados ao S1 em muitos estudos termodinâmicos (ROGANA *et al.*, 1989; DULLWEBER *et al.*, 2001).

É interessante notar que a inclusão da contribuição eletrostática para a energia de interação intermolecular não é estritamente requerida para descrever a maior parte da variância nos dados de inibição. Isto é curioso, pois seria esperado que a necessidade de um grupamento carregado positivamente, como os grupos amidínio ou guanidínio dos inibidores, deveria ser prontamente traduzida numa importante contribuição eletrostática para as afinidades de ligação. A análise de Checa *et al.* (1997) com respeito à inibição da tripsina por flavonóides ajudou a esclarecer esta aparente contradição. Estes autores concluíram que as interações de vdW seriam dominantes na descrição das afinidades de ligação devido à pequena variação no termo eletrostático, o qual é largamente independente do solvente.

1.1.4 Origens da especificidade de substrato

1.1.4.a Bases cinéticas e energéticas

A especificidade de uma enzima por um determinado substrato é medida pelo parâmetro k_{cat}/K_m . Em serinil proteases, a interação P1/S1 governa a especificidade pelo substrato. Conforme ilustrado pela tripsina, há uma preferência de 10^5 por substratos contendo uma arginina ou lisina na posição P1, em relação ao próximo resíduo mais favorecido (CZAPINSKA & OTLEWSKI, 1999). Isto resulta principalmente de um aumento em k_{cat} ($\sim 10^4$ vezes) e não de um decréscimo em K_m (~ 10 vezes). Da mesma forma, interações favoráveis adicionais nos subsítios S2-Sn e S2'-Sn' são refletidas por

aumentos significativos em k_{cat} . É importante mencionar que o subsítio S1 e os subsítios mais remotos atuam em conjunto para determinar a especificidade de substrato. Isto se torna evidente ao se notar que o valor de k_{cat}/K_m só aumenta com a extensão da cadeia do substrato de enzimas tripsina-símile quando há a interação com Arg ou Lys em S1, mas não com Phe (HEDSTROM, 2002).

O parâmetro k_{cat}/K_m depende somente das etapas de ligação (K_s) e acilação (k_2 ; Fig. 84). Como a especificidade de serinil proteases é ditada por k_{cat} , a etapa de acilação deve ser a mais importante para a preferência de substratos. Ou seja, quando dois substratos competem pelo sítio catalítico de serinil proteases, a etapa de desacilação é irrelevante, pois a competição termina quando o primeiro substrato é acilado. De fato, os valores de k_2 variam por um fator de 1000x entre substratos bons e ruins de serinil proteases, enquanto que K_s varia apenas 10-100x (HEDSTROM, 2002). Entretanto, deve-se notar que se o passo de desacilação for muito lento (k_3 pequeno, acil-enzima muito estável), o ciclo catalítico como um todo será lento. Outra consequência das relações algébricas das quais o parâmetro k_{cat}/K_m é derivado, é que a medida que k_2 cresce, k_1 torna-se insignificante ($k_{cat}/K_m = k_1$) e conseqüentemente a reação torna-se controlada por difusão ($k_1 \sim 10^8 \text{ M}^{-1} \cdot \text{s}^{-1}$). No entanto, tem-se relatado valores de k_1 10^6 - $10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$ nas reações catalisadas por serinil proteases, sugerindo que a formação do complexo ES é mais complicada do que a simples colisão entre os reagentes. Etapas de isomerização e rearranjos conformacionais devem-se ser responsáveis pela redução na velocidade da reação. A especificidade da hidrólise de ésteres por serinil proteases pode ser dominada pela etapa de associação com o substrato, principalmente quando bons grupos de saída estão envolvidos, como é o caso de pNP (HEDSTROM, 2002). Ou seja, uma vez ligado na fenda catalítica, a reatividade intrínseca de um éster assegura a acilação e o progresso através do ciclo catalítico. A consequência disto é que a utilização de peptidil ésteres para avaliar a especificidade pode subestimar o potencial de uma serinil protease em discriminar dois substratos.

Uma nova abordagem para o estudo da especificidade de serinil proteases e processos de reconhecimento molecular em geral — termodinâmica sítio-específica — foi desenvolvida pelo grupo de Enrico Di Cera na Universidade de Washington (revisado por VINDIGNI, 1999; DI CERA, 1998). Basicamente, esta abordagem fornece informações sobre a contribuição energética de cada subsítio de especificidade e sobre o acoplamento entre eles. Isto é conseguido para os subsítios S_n , perturbando-se a

estrutura do substrato XaaXaaXaa-pNA (onde pNA é o grupo p-nitroanilida que permite a quantificação espectrofotométrica do substrato clivado) ao se introduzir um conjunto restrito de modificações nas posições P1, P2 e P3. Os substratos funcionam como um conjunto de sondas específicas que podem relatar as propriedades energéticas dos subsítios S_n da enzima do estado de transição. Um painel mínimo de 8 substratos derivados de FPR-pNA foi testado contra tripsina e as enzimas trombina (selvagem e outros 3 mutantes R221aA, K224A e R221aA/K224A), plasmina e tPA (selvagem e mutante Y99L), e as constantes de especificidade (K_s) determinadas (VINDIGNI *et al.*, 1997). O ativador de plasminogênio, t-PA selvagem, mostrou uma preferência Arg/Lys em P1 incomum, de cerca de 70x em favor de Arg. Há uma perda energética de 1 kcal/mol a mais em t-PA do que na trombina, na alteração de Arg para Lys no substrato. Diferentemente, plasmina mostrou uma leve preferência (50%) por Lys em P1. A enzima t-PA é praticamente indiferente à substituição de Pro por Gly na posição P2 dos substratos, enquanto que a mesma mutação compromete severamente a especificidade das enzimas homólogas trombina, tripsina e plasmina.

Os autores racionalizaram a maior especificidade de t-PA por Arg em P1 em termos da presença incomum entre serinil proteases de His¹⁸⁸, próximo à Asp¹⁸⁹. Por outro lado, foi demonstrado posteriormente que a posição 190 em serinil proteases pode influenciar a preferência por Arg ou Lys em P1 dos substratos (SICHLER *et al.*, 2002). Tripsina e plasmina apresentam Ser¹⁹⁰ enquanto que t-PA e trombina têm Ala¹⁹⁰. Em relação ao subsítio S2, acredita-se que Tyr⁹⁹ em t-PA seja responsável pela incapacidade desta enzima de discriminar Pro ou Gly em P2 dos substratos, uma vez que a substituição Y99L (Leu ocupa a posição 99 em tripsina e trombina) restabelece parcialmente esta propriedade. Outra importante observação decorrente dos estudos de Vindigni *et al.* é que as perturbações introduzidas nas posições P1-P3 são simplesmente aditivas nos substratos de tripsina, tPA e plasmina, indicando que os subsítios S1-S3 destas enzimas não estão acoplados energeticamente. O mesmo não foi observado para trombina, na qual o acesso mais restrito à fenda catalítica provavelmente resulta em um estado de transição mais tencionado e, portanto, dependente de interações ótimas nos subsítios S1-S3.

1.1.4.b Bases estruturais

Diferentemente da cinética da reação catalisada por serinil proteases, o conhecimento sobre os mecanismos estruturais responsáveis pela especificidade de substrato nessas enzimas ainda encontra-se em um nível introdutório. Entretanto, os esforços para se determinar as bases estruturais do reconhecimento de substrato não foram poucos e envolveram não somente experimentos de cristalografia, mas também modelagem molecular, cinética enzimática, mutação sítio dirigida e métodos biofísicos variados (revisado por [PERONA & CRAIK, 1995](#); [CZAPINSKA & OTLEWSKI, 1999](#); [VINDIGNI, 1999](#); [HEDSTROM, 2002](#)). Logicamente, um dos primeiros alvos de estudo foi a posição 189, localizada no fundo do subsítio S1 e ocupada pelo resíduo de ácido aspártico em enzimas tripsina-símile. Mutações introduzidas nesta posição diminuem em 10^5 vezes a eficiência catalítica contra substratos contendo Arg/Lys na posição P1, a qual pode ser restabelecida pela re-introdução de Asp ou Glu nas posições 189/190 ([PERONA & CRAIK, 1995](#)). Interessantemente, a atividade catalítica do mutante D189S da tripsina é aumentada em 300 vezes pela presença de altas concentrações de acetato de sódio, indicando que a imobilização do íon acetato pode, em parte, restabelecer o potencial eletrostático negativo do subsítio S1 apesar do alto custo entrópico envolvido ([CZAPINSKA & OTLEWSKI, 1999](#)).

[Perona et al. \(1993\)](#) mostraram que a carga negativa representada pelo Asp¹⁸⁹ pode ser re-allocada no subsítio S1 (D189G/G226D) sem comprometer significativamente a eficiência catalítica, porém causando uma mudança na preferência pelo resíduo em P1 por 30-40 vezes em favor de Lys. A carga negativa do aspartato introduzido na posição 226 é parcialmente compensada por interações polares com S190 e Y228, explicando a menor afinidade (maior K_m) da enzima mutante por substratos contendo Lys/Arg em P1 em termos de uma diminuição na energia de interação eletrostática. Os dados estruturais indicam que a preferência por Lys em P1 decorre de uma maior plasticidade da cavidade S1 ([CZAPINSKA & OTLEWSKI, 1999](#)). Esta última pode remodelar-se na enzima mutante para melhor acomodar a cadeia mais curta de Lys (em relação à Arg) e promover contatos diretos entre o grupo amônio e a carga negativa do íon aspartato na posição 226. Estas observações sugerem porque o resíduo de Asp é conservado na posição 189 em serinil proteases tripsina-símile. Nesta posição, o aspartato não pode fazer interações com outros grupos químicos no subsítio S1.

Conforme introduzido na seção 1.1.1 acima, a especificidade do subsítio S1 de serinil proteases é usualmente atribuída a um conjunto limitado de elementos estruturais referido em livros-texto de bioquímica como o paradigma tripsina-quimiotripsina-elastase. Trocar os elementos de um arcabouço protéico para outro deveria ser suficiente para transferir a especificidade entre as enzimas, mas as mutações óbvias nesse propósito (e.g., D189S) simplesmente falharam (GRAF *et al.*, 1988). Tripsina adquire nova especificidade somente quando as seguintes substituições mais abrangentes visando a introdução de um subsítio S1 quimiotripsina-símile são realizadas: I138T, D189S, Q192M, inserção Δ T219 e substituição das alças 187 (resíduos 185-188) e 220 (resíduos 221-224). A protease mutante obtida, Tr \rightarrow Ch[S1+L1+L2], é 100 vezes mais ativa que o mutante D189S e possui 0,5% da atividade da quimiotripsina (HEDSTROM *et al.*, 1992). A substituição Y172W aumenta a atividade deste mutante para até 15% da atividade da quimiotripsina (HEDSTROM *et al.*, 1994b). Compatível com o papel crucial da etapa de acilação na determinação da especificidade de substrato em serinil proteases, a introdução de atividade quimiotripsina-símile na tripsina é quase inteiramente resultado de um aumento de 10^4 em k_2 . A desacilação é praticamente inalterada nas proteases mutantes, enquanto que a formação de ES também é comprometida (por um fator de até 10x). Os resultados destes estudos demonstraram claramente que a discriminação do resíduo em P1 dos substratos de serinil proteases da família S1A depende da integridade do subsítio S1 como um todo, mas também de interações remotas que mantêm o substrato corretamente alinhado para o ataque da ligação peptídica.

Outro aspecto importante dos determinantes estruturais da especificidade de serinil proteases no subsítio S1 é que a especificidade da reação de hidrólise de ésteres pode ser facilmente alterada, em contraste com a atividade amidásica. Hung & Hedstrom (1998) introduziram alterações adicionais no mutante de tripsina Tr \rightarrow Ch[S1+L1+L2] para refletir os elementos estruturais que possivelmente estão envolvidos na especificidade elastase-símile. O mutante Tr \rightarrow El-2 foi desenhado para imitar o subsítio estericamente restrito da elastase com as mutações G216V e G226T, além das modificações necessárias nas alças 187 e 220. Tr \rightarrow El-2 apresentou-se como uma esterase ainda mais específica que a própria elastase para substratos contendo Ala em P1. Entretanto, este mutante é desprovido de atividade amidásica, o que leva a

acreditar que a facilidade com que a atividade esterásica é manipulada se deve à maior reatividade intrínseca de ésteres (HEDSTROM, 2002).

Mutações reversas, no sentido de introduzir a especificidade da tripsina por P1 Arg/Lys na quimiotripsina, revelaram outro aspecto peculiar dos determinantes estruturais da especificidade de substrato em enzimas da família S1A (VENEKEI *et al.*, 1996, JELINEK *et al.*, 2004). Foi demonstrado que a simples mutação S189D é capaz de introduzir atividade tripsina-símile no mutante A226G da quimiotripsina B de rato (equivalente à quimiotripsina A, onde Gly ocupa naturalmente a posição 226). O mutante S189D/A226G apresenta especificidade 100x maior que o quimiotripsinogênio selvagem para substratos com P1=Lys e apresenta a atividade quimiotripsina-símile prejudicada por um fator de 10^3 . Ainda, a atividade tripsina-símile deste mutante é cerca de 10^3 - 10^4 x menor que a tripsina selvagem. A maior rigidez do subsítio S1 na quimiotripsina A deve fornecer um arcabouço mais estável, onde mutações podem ser inseridas alterando a polaridade do sítio sem comprometer a sua estabilidade. O mesmo parece ocorrer em serinil proteases primitivas, como, SGT, a enzima tripsina-símile de *Streptomyces griseus*. Em SGT, a mutação T190P reduz em apenas 25% a atividade catalítica da enzima, enquanto que alterna sua especificidade de Arg para Lys em P1 (PAGE *et al.*, 2003).

Estruturas cristalográficas dos mutantes S189D de quimiotripsina B e D189S de tripsina de rato foram comparadas com as estruturas das enzimas selvagens correspondentes a fim de elucidar as razões estruturais para o fracasso na instalação das atividades tripsina- e quimiotripsina-símile, respectivamente (SZABÓ *et al.*, 2003). A ligação de substrato nesses mutantes é prejudicada por deslocamentos substanciais na ponte dissulfeto C191-C220 e nos segmentos 185-195 e 217-224. Foi observado que as mutações introduzidas conferem alterações estruturais no domínio de ativação que tornam as enzimas recombinantes semelhantes aos zimogênios inativos. Os resultados discutidos nos parágrafos acima são compatíveis com a visão de que as diferentes especificidades esculpidas durante a evolução do subsítio S1 no enovelamento de enzimas da família S1A foram resultantes de diferentes soluções para problemas estruturais semelhantes. Assim, cada protease da família da quimiotripsina pode exibir um conjunto específico de resíduos para determinar cada tipo de atividade enzimática e/ou especificidade de substrato. De fato, Pál *et al.* (2004) relataram a obtenção de uma tripsina mutante onde uma rota alternativa para a instalação de atividade quimiotripsina-

símile foi vislumbrada. Esta consistiu na alteração da tripsina de rato através da substituição D189S e da deleção Δ D223. A protease mutante obtida teve a eficiência catalítica reduzida em muitas ordens de magnitude, mas apresentou extraordinária especificidade por resíduos de Tyr em P1.

Diferentemente do subsítio S1, as bolsas S1' e S2' podem ser facilmente redesenhadas para exibir diferentes especificidades (WILLETT *et al.*, 1996; KURTH *et al.*, 1997; KURTH *et al.*, 1998). Isto pode ser interpretado em termos das diferentes topologias dos subsítios S1, S1' e S2'. O subsítio S1 é uma cavidade profunda que veste todo o resíduo em P1. Diferentemente, os subsítios S1' e S2' apresentam-se com bolsas rasas na superfície de enzimas da família da quimiotripsina. Além disso, o subsítio S1 compartilha uma das laterais de sua arquitetura com o domínio de ativação. Estas características são refletidas na estrutura intrinsecamente instável do subsítio S1, em oposição aos subsítios Sn' adjacentes que são mais resistentes à mutações (HEDSTROM, 2002).

Diferentemente das proteases digestivas, serinil proteases envolvidas em processos fisiológicos requerendo o processamento de apenas um ou poucos substratos, evoluíram para apresentar modificações em alças na superfície que introduzem novas restrições à especificidade de substrato (PERONA & CRAIK, 1997; ROSE & DICERA, 2002). Tais determinantes funcionais estão mapeados na estrutura da tripsina aniônica de rato na **Fig. 89**. Sete alças separadas (A-E, 2 e 3) foram implicadas em sítios de contato distais com inibidores e/ou substratos em diferentes enzimas. A alça C é invariavelmente posicionada para contatar diretamente o substrato estendido do lado N-terminal da ligação lábil, enquanto que as alças A, B, D e E interagem com o lado do grupo de saída. As alças 2 e 3 modulam tanto a especificidade do subsítio S1 como a preferência de substrato dos sítios Sn. Até onde é sabido, a alça 1 determina apenas a especificidade no sítio S1. A alça C com a seqüência K-R-R-K, por exemplo, está envolvida no reconhecimento de enteroquinase com seu substrato fisiológico, tripsinogênio. Outras modificações na estrutura da enteroquinase ocorrem nas alças B e 2, todas possivelmente envolvidas na evolução da especificidade de substrato nesta enzima (PERONA & CRAIK, 1997).

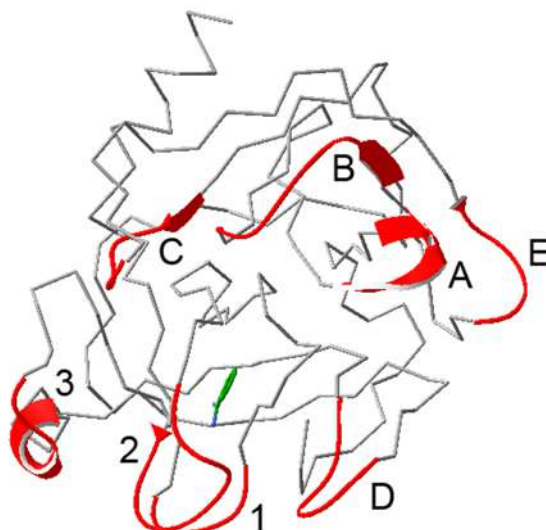


Fig. 89: Localização das oito alças na superfície responsáveis pela determinação da especificidade de substrato em enzimas da família da quimiotripsina. A estrutura da tripsina aniônica de rato representa o arcabouço protéico conservado na família S1A (segmentos tubulares em cinza). Os determinantes funcionais estão destacados em vermelho e representados como fitas (*ribbons*). A orientação do inibidor benzamidina (representado em palitos) marca a localização do subsítio S1. Alça A: resíduos 34-41; Alça B: resíduos 56-64; Alça C: resíduos 97-103; Alça D: resíduos 143-149; Alça E: resíduos 74-80; Alça 1: resíduos 185-188; Alça 2: resíduos 217-225; Alça 3: resíduos 169-174.

1.2 *Serininil proteases ativas sobre a coagulação sanguínea*

1.2.1 A coagulação sanguínea

Dois grupos de bioquímicos trabalhando independentemente na década de 60 propuseram um modelo de cascata da coagulação que consistia em uma série de passos nos quais a ativação de cada fator da coagulação leva a ativação de outro, resultando finalmente numa aceleração da formação de trombina (DAVIE, 2003). Cada fator da coagulação existe como um zimogênio ou co-fator inativo na circulação sanguínea que é ativado ou convertido em sua forma ativa por proteólise limitada. Os fatores procoagulantes (trombina, FVIIa, FIXa, FXa, FXIa e FXIIa) e enzimas envolvidas em atividades anticoagulantes (aPC, plasmina e ativadores de plasminogênio) são todas serininil proteases da família S1A. Atualmente, sabe-se que a coagulação sanguínea em mamíferos possui um componente celular (responsável pela hemóstase primária) e um protéico (revisado mais recentemente por MANN, 2003; HOFFMAN, 2004). Este último, consistindo na cascata da coagulação, é iniciado e regulado por uma complexa rede de interações regulada por ciclos de retro-alimentação negativos e positivos.

A exposição do fator VII (FVII) ao fator tecidual (TF) na superfície de células extravasculares (via extrínseca) resulta na ativação de FIX e FX, pelo complexo TF-FVIIa (**Fig. 90**). Na ausência de seu co-fator ativado, FVa, FXa gera somente quantidades traço de trombina. Apesar de ser insuficiente para iniciar polimerização significativa da fibrina, as quantidades ínfimas de trombina formada neste estágio de iniciação da coagulação são capazes de retro-ativar FV, FVIII e XI por proteólise limitada. Isto leva à geração “explosiva” de trombina pela ação da protrombinase (FVa-FXa) durante a etapa de propagação, que em última análise leva à formação do coágulo de fibrina. Este processo de propagação ocorre independentemente do complexo TF-FVIIa, o qual é rapidamente inativado por TFPI, o inibidor da via do TF (**BROZE-JR, 1995**), e é mantido pelo complexo tenase (VIIIa-IXa) formado na via intrínseca.

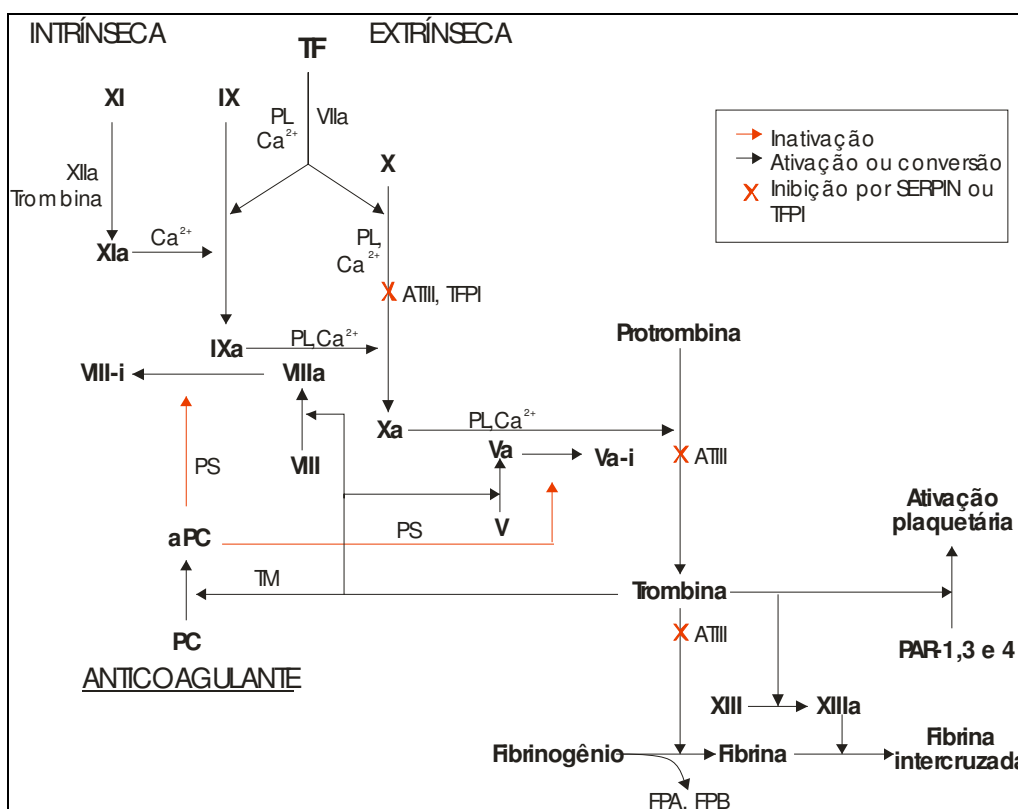


Fig. 90: Cascata da coagulação. A via extrínseca inicia a coagulação em condições fisiológicas e consiste no fator tecidual (TF) e no fator VII (FVII). Na via intrínseca, responsável pela manutenção da coagulação, FXa é gerado pelo complexo tenase (FIXa-VIIIa). Para maiores detalhes vide texto.

A trombina, formada pela ativação da protrombina (PT) ao final da cascata, converte o fibrinogênio solúvel no plasma em um polímero insolúvel (fibrina) pela remoção dos fibrinopeptídeos A e B (FPA e FPB). Contribuindo para a formação de um coágulo insolúvel, trombina ativa FXIII, uma transglutaminase que entrecruza as

cadeias de fibrina (HERRICK *et al.*, 1999). Finalmente, trombina atua como um dos mais potentes ativadores plaquetários conhecidos. Isto ocorre através de proteólise limitada de PARs (receptores ativados por protease) uma classe específica de receptores acoplados à proteína G, presentes na superfície de plaquetas (BRASS, 2003). Dos 4 subtipos de PAR conhecidos, trombina é capaz de ativar 3 (PAR-1, PAR-3 e PAR-4).

A restrição da formação do coágulo ao local da lesão é conseguida por diversos mecanismos. Primeiramente, as reações descritas acima para os complexos TF-FVIIa, tenase e protrombinase somente ocorrem eficientemente na presença de fosfolipídios (PL) e Ca^{2+} . O componente PL é oriundo da membrana celular de células sub-endoteliais (TF-FVIIa) ou plaquetas ativadas, que por sua vez secretam Ca^{2+} , entre outros efetores (Fig. 91). Plaquetas são capturadas no local da lesão através da interação com o fator de von Willebrand (vWF), uma proteína plasmática imensa, que serve como ponte entre o colágeno exposto no tecido subendotelial e as glicoproteínas (GPIb) na membrana de plaquetas (WU & THIAGARAJAN, 1996). Isto efetivamente restringe a polimerização da fibrina às vizinhas do endotélio lesado e do tampão plaquetário formado. A trombina também ativa proteolicamente uma via anticoagulante ao se ligar à trombosmodulina (TM), uma proteína transmembrana presente nas células do endotélio (WU & THIAGARAJAN, 1996). No complexo com TM, trombina não mais cliva eficientemente os substratos procoagulantes, incluindo FV, FVIII e fibrinogênio. O substrato do complexo trombina-TM é a proteína C, cuja forma ativada (aPC) associa-se ao co-fator proteína S (PS) e rapidamente inativa FVa e FVIIIa por clivagem proteolítica adicional (ESMON, 2000). Regulação negativa adicional da cascata da coagulação ocorre através da ação de anti-trombina III (ATIII), uma SERPIN que inibe a atividade de FIXa, FXa e trombina (JANCIAUSKIENE, 2001).

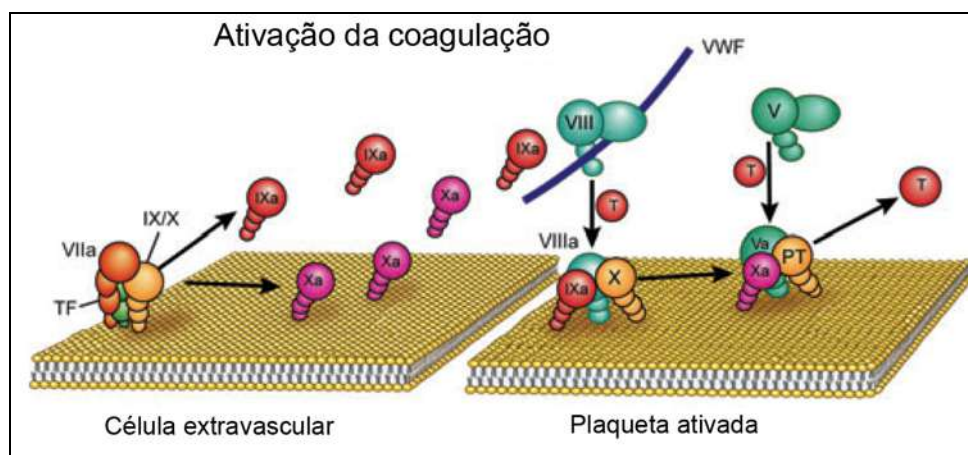


Fig. 91: Localização celular de processos envolvidos na ativação da coagulação. O fator tecidual (TF, em verde) é exposto na superfície de células extravasculares. Os fatores IXa e Xa podem difundir-se pela bicamada fosfolipídica na forma do complexo com TF-FVIIa ou podem difundir-se pelo plasma até encontrarem com seus co-fatores, FVIII e FV, respectivamente. FVIII circula no plasma ligado ao fator de von Willebrand (VWF). A trombina (T), formada a partir da ativação de traços de protrombina (PT) no início da coagulação, retro-alimenta a cascata ao ativar os fatores VIII e V, que irão compor, respectivamente, os complexos tenase e protrombinase na superfície de plaquetas. Adaptado de [Dahlbäck \(2005\)](#).

1.2.2 Estrutura e função de serinil proteases da cascata da coagulação

Serinil proteases envolvidas na cascata da coagulação (SPCs) estão presentes em todos os vertebrados e algumas são compartilhadas por invertebrados superiores ([AIRD, 2003](#)). As funções especializadas destas enzimas foram adquiridas há mais de 450 milhões de anos, quando cordados divergiram dos invertebrados ([DAVIDSON *et al.*, 2003](#)). Num processo relativamente rápido em termos evolutivos (~50 milhões de anos), SPCs tiveram suas funções esculpidas por processos de duplicação e divergência gênica. Quatro zimogênios principais (FVII, FIX, FX e PC) evoluíram a partir de reduplicação e diversificação funcional de uma pro-enzima ancestral com a estrutura [Gla-EGF1]-[EGF2]-[SP]. Gla-EGF1 é o primeiro domínio similar ao fator de crescimento epidérmico (EGF) contendo um segmento N-terminal de resíduos de ácido glutâmico γ -carboxilados (Gla), EGF2 é o segundo domínio EGF e SP é o domínio catalítico de serinil proteases. Este zimogênio ancestral e todas as SPCs dele originadas dependem de vitamina K para a γ -carboxilação dos resíduos de ácido glutâmico. PT também é o zimogênio de uma SP dependente de vitamina K, mas possui dois domínios “kringle” no lugar dos domínios EGF. Entretanto, a protease madura, trombina, não mais apresenta estes domínios, diferente das demais SPCs, que os mantém ligados através de interações não-covalentes e uma ponte dissulfeto com a chamada cadeia leve

(contendo Gla-EGF1-EGF2). Foi proposto que a PT primitiva também possuía domínios EGF, mas estes teriam sido substituídos durante a duplicação gênica e embaralhamento de exons (DAVIDSON *et al.*, 2003). Análise da seqüência de aminoácidos e utilização de códons por resíduos atrelados à função catalítica de serinil proteases diferencia trombina de FVII, FIX, FX e PC, sugerindo que a primeira foi a SPC ancestral (KREM & DI CERA, 2002b).

Em SPCs, as alças c36, c60, c70, c90, c140 e c170 na superfície delineiam a fenda de ligação ao substrato. A posição das alças é conservada no envelhecimento quimiotripsina-símile, mas a composição de aminoácidos é variável (Fig. 92). SPCs clivam ligações peptídicas em resíduos básicos (principalmente Arg, mas também Lys) como a tripsina, mas adquiriram um elevado grau de especificidade e clivam somente um número limitado de ligações.

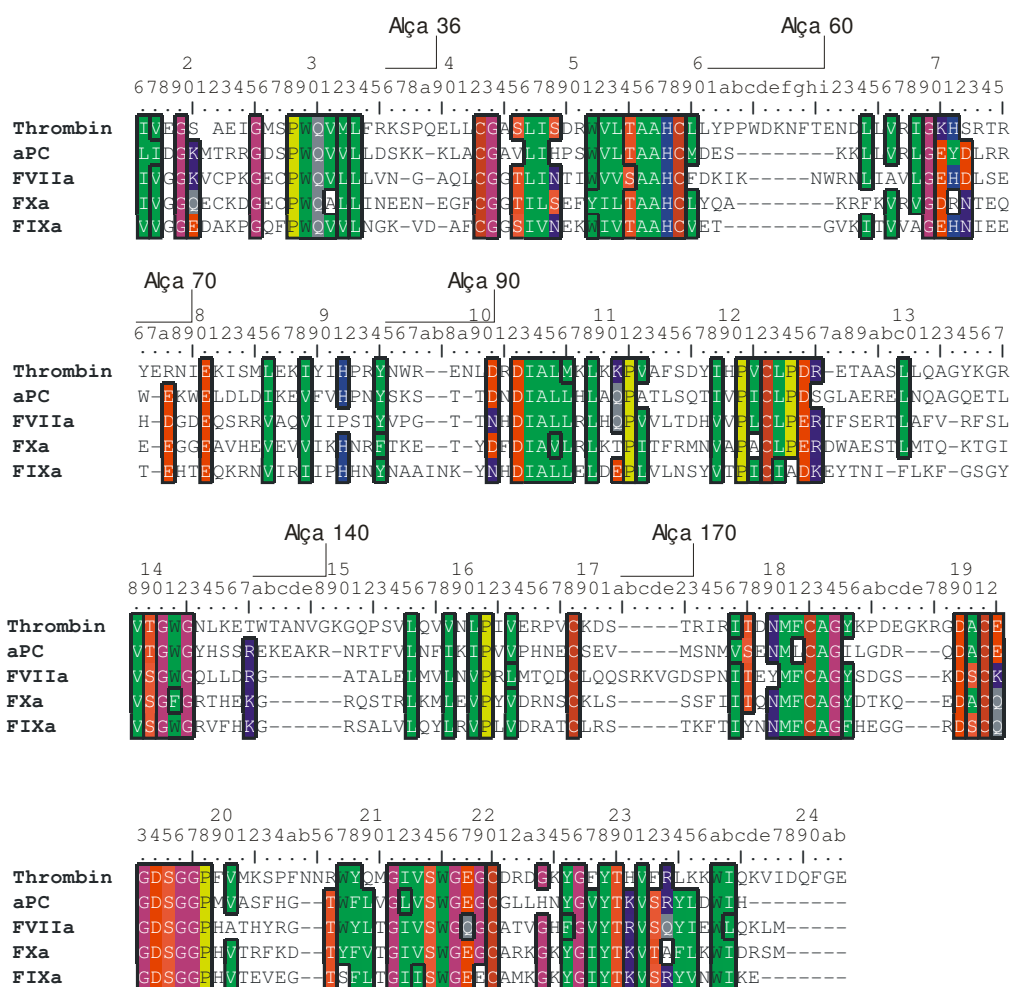


Fig. 92: Alinhamento estrutural das principais serinil proteases da cascata da coagulação humana. As posições do alinhamento com aminoácidos conservados em pelos menos 4 das cinco proteínas são coloridas de acordo com a matriz de escore BLOSUM62 (HENIKOFF, 1992).

1.2.2.a Fator VII

O FVII tem importante papel na iniciação da coagulação. Esta glicoproteína circula no plasma como um zimogênio de cadeia única com 50 kDa e 406 resíduos de aminoácidos. A deficiência de FVII é uma doença autossômica recessiva rara, com um amplo espectro de fenótipos hemorrágicos de baixa correlação com a atividade FVIIa residual (PEYVANDI *et al.*, 2001; PINOTTI *et al.*, 2002; GOMEZ *et al.*, 2004). A maior parte de FVII está presente no sangue na forma do zimogênio inativo, com as quantidades traço de FVIIa presentes, provavelmente surgindo da ação de FXa e FIXa (revisado por EIGENBROT, 2002). Assim como nos zimogênios de outras SPs da família S1A, o evento chave formal na ativação de FVII é a clivagem entre cArg¹⁵-cIle¹⁶ (onde “c” designa o sistema de numeração do quimiotripsinogênio). Entretanto, para ativação completa de FVIIa a associação com TF é requerida. Desta forma, FVIIa existe em um equilíbrio entre as formas parcialmente e totalmente ativadas, o qual é alterado em direção desta última pela associação com TF. FVIIa é sujeito a inúmeras influências alostéricas, a maior parte das quais se originam da ligação de 8 íons Ca²⁺ no domínio Gla-EGF1. Entretanto, pouco se sabe como estes eventos alteram a atividade amidase no sítio catalítico distante.

Um problema de certa forma mais tratável é o estudo de influências alostéricas originando-se de exossítios dentro do próprio domínio catalítico. Os detalhes moleculares da influência alostérica de TF e a relação estrutura-função de FVIIa foram estudados através da determinação de diversas estruturas cristalográficas, incluindo: FVIIa ligado a 1,2dansil-EGR-CMK (KEMBALL-COOK *et al.*, 1999), FVIIa em complexo com TF solúvel (BANNER *et al.*, 1996), FVIIa no complexo ternário com TFs e BPTI (ZHANG *et al.*, 1999), o zimogênio FVII ligado a peptídeo no exossítio A (EIGENBROT *et al.*, 2001), zimogênio FVII ligado a TF solúvel (SICHLER *et al.*, 2002) e mais recentemente FVIIa com o inibidores peptídeomiméticos (KADONO *et al.*, 2004; OLIVERO *et al.*, 2005). Adicionalmente, há uma ampla literatura sobre os efeitos funcionais de mutações sítio-dirigida de FVIIa e suas relações com propriedades alostéricas.

Usualmente considera-se que o domínio SP de FVIIa é provido de três sítios alostéricos: a região de ligação à TF (alça-c170 e hélice α -1), a fenda catalítica e o exossítio de ligação a substratos macromoleculares (Fig. 93). Este último consiste

fundamentalmente na região entre o sítio de ligação ao cálcio e a região de inserção N-terminal. Um evento de ligação em qualquer um desses três sítios induz mudanças nos outros dois. A ligação de TF aumenta a atividade de FVIIa tanto sobre substratos macromoleculares como sobre substratos pequenos cromogênicos. Este aumento na atividade também depende da presença de fosfolipídios por um fator de 10^4 . Desta forma, o efeito ativador da formação do complexo “tenase extrínseco” (FVIIa-TF-PL- Ca^{2+}) pode ser dividido num efeito puramente alostérico (conforme refletido no aumento de 50 vezes na atividade amidásica de FVIIa-TF solúvel) e outros efeitos derivados simplesmente da co-localização de TF, FVIIa e o substrato macromolecular na membrana, e de contatos diretos entre TF e o substrato.

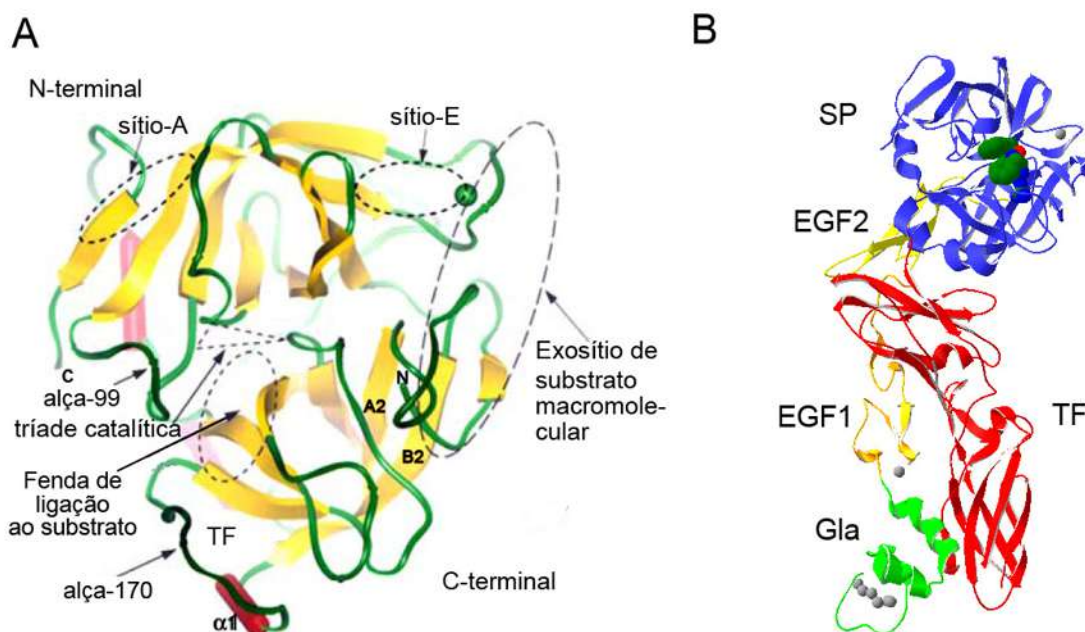


Fig. 93: Sítios alostéricos no domínio catalítico de FVIIa (A) e estrutura de domínios em FVIIa em complexo com fator tecidual (B). Em (A) o sítio de ligação ao Ca^{2+} na alça c70-c80 é representado pela esfera em verde. As alças c99 e c170 (resíduos c168-c182) compõem os subsítios S2 e S3. Os sítios A (próximo ao C-terminal) e E (entre o domínio de ativação e o sítio de ligação ao Ca^{2+}) interagem com peptídeos em solução e podem estar envolvidos na interação com ligantes fisiológicos. As fitas- β A2 (c134-c140) e B2 (c160-c153) no C-terminal estão envolvidas na ativação do zimogênio e podem transmitir alterações originadas pela ligação de TF (alça-c170 e hélice α 1) até o sítio de ligação a macromoléculas e ao sítio ativo.

As alças c99 e c170 delinham a extremidade N-terminal da fenda de ligação ao substrato, interagindo com as posições P2 e P3 dos substratos (**Fig. 93**). Moléculas de FVIIa recombinante onde a composição destas alças foi alterada para refletir a composição dos segmentos homólogos na tripsina mostraram-se ~20 vezes mais ativas

que FVIIa selvagem contra substratos peptídicos e macromoleculares (SOEJIMA *et al.*, 2002). A alça-c99 contribui para a afinidade de ligação ao substrato, enquanto que as interações com a alça-c170 melhor posicionam este para a catálise. Comparações entre as estruturas do zimogênio de FVII, FVIIi (inibido) e FVIIa no complexo com TF revelou que além das alterações conformacionais típicas no domínio de ativação, mudanças importantes na região de ligação a TF também podem ser observadas. Estas consistem num deslocamento inesperado de três resíduos na fita- β B2 (c160-c158) relativa à fita A2, alterando o registro das ligações de hidrogênio nesta folha- β no domínio C-terminal (Fig. 93).

Estudos recentes tentaram impor um caráter de protease-ativa (como quando ligada a TF) em FVIIa ao introduzir ligações dissulfeto conectando as fitas A2 e B2 no intuito de imobilizá-las na conformação encontrada no complexo FVIIa-TF ativo (OLSEN *et al.*, 2004; MAUN *et al.*, 2005). Olsen *et al.* alteraram o padrão de ligações dissulfeto em FVIIa de cC22:cC27 para cC22:cC157 ao introduzir as mutações cC27V/cV157C, mas não conseguiram aumento na atividade de FVIIa em comparação à enzima ativa selvagem. Algumas das mutações introduzidas por Maun *et al.* (e.g., cC136:cC160 e cC138:cC160) foram mais eficientes, conseguindo níveis de atividade amidásica em FVIIa mutante, sobre certos substratos peptídicos, comparáveis a FVIIa selvagem no complexo com TF. Entretanto, a atividade contra substratos macromoleculares não foi instituída nestes mutantes, indicando que as mutações introduzidas podem ter afetado negativamente a conformação de alguns epitopos funcionais em FVIIa.

1.2.2.b Fator IX

FIX tem papel essencial na coagulação, conforme demonstrado pela tendência ao sangramento associada à deficiência congênita deste fator (hemofilia B). Além disso, estudos mostram que níveis elevados de FIX estão associados com o aumento no risco de tromboembolismo venoso e trombose arterial coronária aguda (LOWE, 2001). Mais ainda, estudos recentes sugerem que quantidades ínfimas de FIXa circulando no plasma sejam responsáveis pela geração de trombina na ausência de TF (BUTENAS *et al.*, 2004). A proteína é sintetizada nos hepatócitos e sofre extensa modificação pós-traducional, resultando num zimogênio glicosilado (17% de carboidratos) com 415 resíduos e 57 kDa (revisado por SCHMIDT & BAJAJ, 2003). A análise da seqüência de

aminoácidos revela que a estrutura de domínios de FIX é semelhante à FVII, contendo um curto segmento hidrofóbico (resíduos 41-46 na numeração do zimogênio de FIX) entre a região Gla e EGF1, um segmento de ligação (resíduos 84-87) entre EGF1 e EGF2, e o peptídeo de ativação (resíduos 146-180) entre EGF2 e o domínio SP. Espacialmente, as estruturas também são bastante semelhantes, conforme evidenciado pela cristalografia de FIXa suíno (Gla-EGF1-EGF2-SP; BRANDSTETTER *et al.*, 1995) e humano (EGF2-SP; HOPFNER *et al.*, 1999). Assim como FVII, FIX possui um único sítio de alta afinidade de ligação ao cálcio no domínio catalítico (alça c70-c80), envolvido na modulação das funções desempenhadas por esta enzima (SCHMIDT & BAJAJ, 2003).

Diferentemente de FVII, a ativação de FIX envolve a clivagem de duas ligações nas posições Arg¹⁴⁵-Ala¹⁴⁶ e Arg¹⁸⁰-Val¹⁸¹ (ou cArg¹⁶-cVal¹⁷) para remoção do peptídeo de ativação. A cadeia leve (resíduos 1-145) permanece ligada à cadeia pesada (resíduos) através de uma ponte dissulfeto. Importaneamente, estudos de mutação sítio dirigida revelaram que além dos sítios de reconhecimento primários das ligações onde os eventos proteolíticos ocorrem (Arg¹⁴⁵ e Arg¹⁸⁰), pelo menos dois exossítios em FIX — os domínios Gla e EGF1 — estão envolvidos no complexo ternário TF-FVIIa-FIX (NDONWI *et al.*, 2005). Adicionalmente, experimentos de mutação sítio-dirigida na chamada alça de autólise (resíduos 143-154) de FIXa demonstraram que cArg¹⁴³ e cLys¹⁴⁷ são sítios de reconhecimento para FX, independentemente de FVIIIa, enquanto que cArg¹⁵⁰ é um sítio de interação com a SERPIN ATIII na conformação ativada por heparina (YANG *et al.*, 2003). A alça de autólise tem importante papel na determinação da especificidade por substratos ou inibidores em trombina e aPC.

Durante a etapa de propagação da coagulação, a ativação de FX pelo complexo tenase intrínseco (FIXa-FVIIIa) é a etapa limitante da geração de trombina. Resumidamente, FVIIIa é um heterotrímero composto de três cadeias protéicas associadas não covalentemente: A1, A2 e A3-C1-C2. Esta última é conhecida como a “cadeia leve” e associa-se com a subunidade A1 numa interação dependente de íon metálico divalente. Muitos grupos tem se dedicado à investigação do modo de ligação de FIXa ao seu co-fator FVIII (MATHUR *et al.*, 1997; KOLKMAN *et al.*, 1999; MATHUR & BAJAJ, 1999). Dados estruturais e funcionais foram usados para gerar modelos aproximados da interação de FIXa com FVIII (BAJAJ *et al.*, 2001; STOILOVA-McPHIE *et al.*, 2002). Nestes modelos, o domínio SP e possivelmente o

domínio EGF2 de FIXa ligam-se a FVIIIa, através dos domínios A2 e A3. Acredita-se que EGF1 exerça um papel estrutural, ao posicionar os domínios EGF2 e SP para a ligação a FVIII, enquanto que nenhum papel para o domínio Gla está definido. Mais recentemente, os dados experimentais sobre a interação FVIIIa/FIXa foram reavaliados e um total de 10 modelos foram gerados em concordância com esses dados (AUTIN *et al.*, 2005). Cinco conformações do modelo de FIXa humano (sem o domínio Gla) no complexo com o modelo A1-A2-A3 de FVIIIa foram gerados. Para cada um destas, 2 modelos alternativos foram construídos ao adicionar os domínios C1-C2. A Fig. 94 mostra os sítios de interação entre FVIIIa e FIXa esquematicamente mapeados nas estruturas 3D.

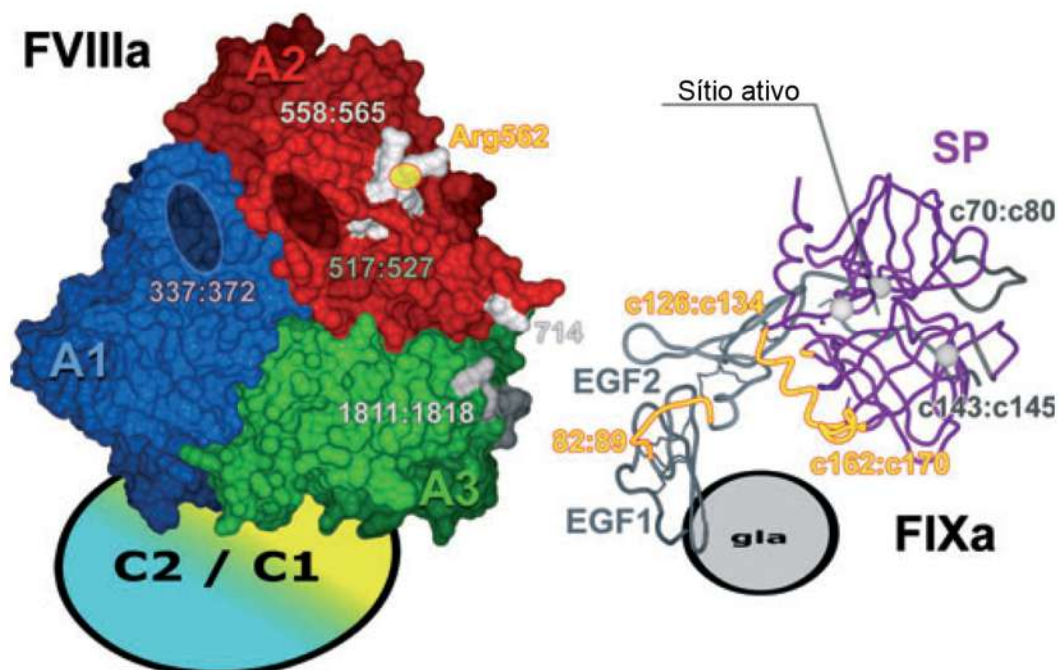


Fig. 94: Sítios de interação entre FIXa e FVIIIa mapeados nas estruturas 3D. Em FVIIIa (superfície molecular), os sítios de interação com FIXa visíveis aparecem em branco; Arg⁵⁶² é o sítio de clivagem para aPC; e as regiões ovais escurecidas foram omitidas nos modelos construídos por Autin *et al.* (2005). Em FIXa (em fitas), as esferas em branco demarcam a tríade catalítica; as regiões em amarelo são sítio de interação com FVIIIa; e as regiões em cinza no domínio SP não devem interagir diretamente com FVIIIa. Adaptado de Autin *et al.* (2005).

Heparina é um conhecido agente antitrombótico, mas os mecanismos moleculares precisos pelos quais este oligossacarídeo heterogêneo atua ainda são desconhecidos. A ativação conformacional de ATIII por um pentassacarídeo derivado da heparina estimula a inibição de ambos FIXa e FXa através de interações mediadas por um exossítio. No entanto, a formação dos complexos enzimáticos ligados aos

fosfolipídios da membrana resulta em proteção parcial da ação da heparina. Por outro lado, oligossacarídeos da heparina inibem a ativação de FX pelo complexo tenase intrínseco através da ligação direta ao um exossítio em FIXa que inviabiliza a interação crítica com o co-fator VIIIa (SHEEHAN *et al.*, 2003). Mais recentemente, Yuan *et al.* (2005) expressaram mutantes de FIXa onde resíduos básicos (cH92, cR170, cR233 e cK241) numa região homóloga aos exossítio de ligação à heparina na trombina e FXa foram substituídos por Ala. O mutante cR233A apresentou marcante redução na afinidade por heparina, enquanto que as reduções observadas para cH92A, cR170A e cK241A foram moderadas. Estes dados sugerem que o principal sítio de ligação de FIXa à heparina seja restrito à região entre a α -hélice c165-c170 e a hélice C-terminal c235-c241. Mais ainda, ensaios de ligação com FVIIIa mostraram que o exossítio de ligação à heparina sobrepõe-se extensivamente com um importante sítio de interação (pontuado nos resíduos cR170 e cR233) com o domínio A2 deste co-fator.

1.2.2.c Fator X

FXa ocupa posição única na cascata da coagulação como a primeira enzima na via comum de formação da trombina. A deficiência de FX é a mais severa das doenças hemorrágicas congênitas raras (UPRICHARD & PERRY, 2002). FX é sintetizado no fígado como uma glicoproteína de 488 resíduos que após remoção dos 40 resíduos da região pré-pro N-terminal é submetida a extensa modificação pós-traducional antes de ser secretada, incluindo: glicosilação (Asn¹⁸¹ e Asn¹⁹¹), γ -carboxilação (primeiros 11 resíduos de ácido glutâmico), β -hidroxilação (Asp⁶³) e remoção do tripeptídeo (Arg-Lys-Arg) entre Arg¹³⁹ e Ser¹⁴³. A proteína madura resultante é um zimogênio de serinil protease que após remoção do peptídeo de ativação de 52 resíduos (resíduos 143-194), fornece a protease madura (FXa) com uma cadeia pesada de 253 resíduos (o domínio SP, 195-448) e uma cadeia leve de 139 resíduos (domínios Gla-EGF1-EGF2) ligadas pela ponte dissulfeto C132-C302. FXa é eficientemente gerado no plasma pela associação em complexos quaternários com FVIIa-TF-PL/Ca²⁺ ou FIXa-FVIIIa-PL/Ca²⁺.

O principal substrato fisiológico de FXa é a protrombina (PT), a qual tem a conversão em trombina acelerada em 300 mil vezes pelo complexo protrombinase (FXa-FVa-PL/Ca²⁺). FXa pode ativar ainda os fatores V, VII e VIII. FXa livre no plasma é inativado por ATIII ou pode formar complexo 1:1 com TFPI, o qual pode

ligar-se a FVIIa-TF, formando um complexo quaternário inativo. Neste último, TFPI atua ligando-se primeiramente ao sítio ativo de FXa através do 2º domínio Kunitz e tornando subsequentemente o complexo completamente inativo pela ligação do 2º domínio Kunitz no sítio ativo de FVIIa (BROZE-JR, 1995). Em geral, FXa compartilha muitas das propriedades estruturais de FIX, incluindo uma alça de ativação proteolítica inicial, um exossítio de interação com co-fator de estrutura A1/A2/A3-C1-C2 (FV) requerida para ativação completa da protease e um exossítio de interação com heparina (PADMANABHAN *et al.*, 1993).

Uma característica fundamental de SPCs é a regulação alostérica da atividade enzimática através de interações com exossítios no domínio catalítico ou na cadeia leve. A **Fig. 95** mostra a localização de alguns destes sítios na superfície de FXa. O aumento geral na eficiência de ativação de PT pelo complexo protrombinase em relação a FXa isolado resulta de uma redução em 100 vezes de K_m e um aumento de 1000 vezes em k_{cat} . Estudos funcionais com mutantes de FXa (RUDOLPH *et al.*, 2001) evidenciaram que resíduos básicos nas α -hélices c125-c130, c165-c170 (hélice intermediária) e na ponta da fita- β c226-c230 formam um sítio de reconhecimento de FVa na região sudeste da molécula de FXa. Mais recentemente, o resíduo cK186 no segmento c185-c188 da alça 1 compoendo o subsítio S1 também foi implicado na interação com FVa (REZAIIE & KITTUR, 2004). Resíduos carregados no segmento c82-116, principalmente K96 na alça c90 e K109 na fita- β 103-109 nas proximidades do subsítio S2, interagem diretamente com protrombina e estão energeticamente ligados ao sítio de interação com FVa (MANITHODY *et al.*, 2005).

Glicosaminoglicanos polissulfatados derivados da heparina atuam como co-fatores, acelerando a inativação de FXa (assim como de FIXa) pela SERPIN circulante ATIII por um fator de 4-5 ordens de magnitude. Resíduos básicos na alça de autólise afetam a interação de FXa com ATIII ativada por heparina por um fator de 10 (MANITHODY *et al.*, 2002). Dados estruturais indicam que sete dos onze resíduos básicos do exossítio de ligação à heparina de trombina foram conservados em posições equivalentes na superfície de FXa. Destes, três se sobrepõem ao sítio de interação com FVa — cK169, cR165 e cR125 (REZAIIE, 2000). Contudo, o sítio primário de ligação à heparina parece estar concentrado na α -hélice C-terminal c235-c241.

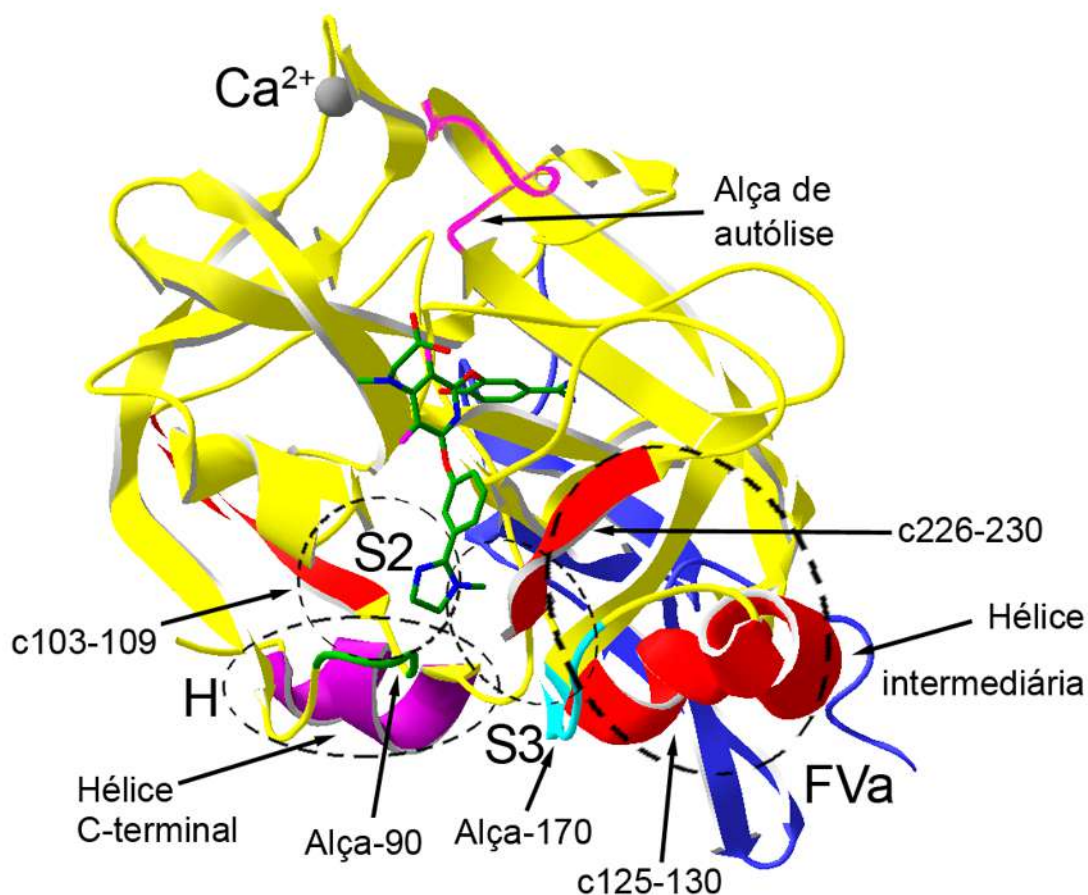


Fig. 95: Determinantes de especificidade e sítios alostéricos em FXa. A estrutura do domínio catalítico é mostrada em amarelo com um inibidor (palitos) ligado à fenda do substrato e com o domínio EGF2 aparecendo em azul na parte posterior da molécula. Os exossítios de ligação a FVa e heparina (H) estão destacados por elipses tracejadas. Os subsítios S2 e S3 também aparecem demarcados. O sítio de ligação ao cálcio está indicado pela esfera de vdW do íon metálico em cinza. Regiões energeticamente ligadas ao sítio de ligação ao FVa aparecem em vermelho. Os demais segmentos comentados no texto estão indicados por setas e codificados por cores.

FXa possui 12-sítios de ligação ao Ca^{2+} , dois dos quais, um situado no domínio EGF1 da cadeia leve e outro no domínio catalítico, são de alta afinidade (SABHARWAL *et al.*, 1997). O sítio de ligação ao Ca^{2+} do domínio SP na estrutura cristalográfica original de FXa determinada por Padmanabhan *et al.* (1993) apresentava-se desordenado. Baseado na estrutura do sítio homólogo na tripsina, foi proposto que em FXa o sítio de ligação de alta afinidade ao Ca^{2+} no domínio catalítico fosse composto pelos grupos carboxilato de cE70 e cE80, e pelos oxigênios carbonílicos dos resíduos nas posições c72 e c75 (SABHARWAL *et al.*, 1997). O papel de cE70 havia sido confirmado anteriormente por mutação sítio dirigida (REZAIIE & ESMON, 1994). Os outros dois ligantes necessários para compor o sítio de coordenação octaédrico

seriam ocupados por moléculas de água do solvente. Diferentemente de trombina e aPC, a alça c70-c80 de FXa possui uma polaridade revertida com somente uma carga positiva e cinco resíduos carregados negativamente. Nessas outras duas proteases, os resíduos básicos da alça c70-c80 contribuem para a interação com seus co-fatores. Um papel protetor para a alça 70-80 em FXa foi proposto, uma vez que esta é capaz de estabilizar a alça de autólise contra clivagem proteolítica e garantir máxima apresentação da atividade biológica, incluindo a interação com FVa (SABHARWAL *et al.*, 1997). Mais recentemente, mutações X→Ala (X = cR71, cE74, cE76 e cE77) demonstraram um papel nas propriedades do zimogênio de FX, ao influenciarem a ativação deste por FIXa no complexo tenase intrínseco, independentemente da interação com FVIIIa, mas não pelo complexo extrínseco (CHEN *et al.*, 2004).

FXa cliva apenas duas ligações em seu substrato, protrombina, tendo (Glu/Asp)-Gly-Arg-(Thr/Ile) como resíduos P3-P2-P1-P1'. As estruturas cristalográficas de FXa e trombina indicaram que o resíduo ocupando a posição c192 (Gln e Glu, respectivamente) poderia exercer importante função na especificidade de substrato destas enzimas devido à sua localização na base do subsítio S1 e elevada flexibilidade. De fato, o fator Xa mutante Q192E teve a interação com FVa prejudicada na formação do complexo protrombinase. Outro resíduo variante ocupa a posição c99, com Tyr em FXa, Thr em aPC e Leu na trombina (Fig. 92). Acredita-se que o resíduo na posição c99, juntamente com os resíduos em c174 e c215, contribua para a interação com o substrato em S2. O mutante Y99T/Q192E de FXa foi preparado e mostrou-se resistente a ATIII, mas somente na ausência de heparina (REZAIE, 1996). O mutante também se mostrou específico para substratos de aPC e incapaz de ativar protrombina, apesar de também não poder inativar FVa. Posteriormente, com importância para o desenho de fármacos, foi observado que Y99 constitui um sítio de ligação às aminas quaternárias em FXa (MONNAIE *et al.*, 2000).

Painéis de substratos fluorogênicos representando as posições P3-P3' do substrato foram testados contra FXa, revelando uma inesperada baixa seletividade na presença ou na ausência de FVa (BIANCHINI *et al.*, 2002). Assim, estes estudos revelaram que com relação à fenda catalítica, FXa assemelha-se mais com uma enzima de baixa especificidade como tripsina do que com trombina. Mais ainda, a impressionante seletividade do complexo protrombinase deve ser originada da interação dos domínios acessórios, o co-fator FVa ou de exossítios em FXa com o substrato

protrombina. Uma abordagem alternativa foi empregada pelo grupo de Milton Stubbs para elucidar os determinantes estruturais na fenda catalítica de FXa para a especificidade de substrato (REYDA *et al.*, 2003). Mutantes de tripsina de rato foram construídos contendo elementos da alça c90 sozinha (X99rT) ou contendo elementos da alça c90, da alça c170 e o resíduo c190 [X(99/175/190)rT] de FXa. Uma característica proeminente da fenda de ligação ao substrato em FXa é a presença de uma “caixa aromática” na vizinhança do sítio ativo, formada pelas cadeias laterais de cTrp²¹⁵, cTyr⁹⁹ e cPhe¹⁷⁴, dispostas perpendicularmente entre si. Num primeiro estágio para introduzir a caixa aromática de FXa na tripsina os autores introduziram as mutações E97K e Y99L obtendo o mutante X99rT. Este último foi cristalizado com um inibidor específico de FXa, mostrando importantes rearranjos estruturais que culminaram na formação da caixa aromática. Em X(99/175/190)rT, a introdução dos resíduos cS172, cS173, cS174 e cS175 da alça c170 e do resíduo c190 (S→A) resultou numa inesperada reorientação da “hélice intermediária” (resíduos 165-170) acompanhada de isomerização da ponte dissulfeto C168-C182 e perda de acessibilidade ao solvente de Phe¹⁷⁴. Estes rearranjos conformacionais podem ser resgatados gerando uma conformação típica de FXa mediante adição do inibidor específico desta enzima. Estes resultados, apoiados por estudos complementares de mutação sítio-dirigida na posição 174 (RAUH *et al.*, 2004), revelaram uma plasticidade da hélice 165-170, implicada na interação de FXa com o co-fator FVa.

1.2.2.d Proteína C

A importância de aPC na regulação da coagulação sanguínea pode ser ilustrada pela observação que a deficiência heterozigótica de PC está associada a um alto risco de trombose venosa e sua deficiência homozigótica resulta na doença fatal conhecida como *púrpura fulminans* (revisado por REZAIIE, 2003). PC circula no sangue como uma glicoproteína heterodimérica de 62 kDa e compartilha a estrutura de domínios com outras enzimas da coagulação dependentes de vitamina K, incluindo o domínio Gla N-terminal e os dois domínios EGF da cadeia leve e o domínio SP na cadeia pesada. O promotor fisiológico da conversão PC→aPC é complexo trombina-trombomodulina (complexo de ativação) na superfície endotelial. Esta ativação é aumentada em 20x mediante ligação ao receptor de PC na célula endotelial (EPCR) através do domínio Gla. Durante a conversão de PC em sua forma ativa, um peptídeo curto de 12 resíduos é removido do N-terminal da cadeia pesada. A proteinase ativa resultante cliva, tornando

inativos os co-fatores Va e VIIIa. aPC precisa clivar pelo menos duas ligações P1-Arg nas posições 306 (na subunidade A1) e 506 (na subunidade A2) de FVa para torná-lo completamente inativo. A clivagem da ligação 506 é rápida e independente de co-fatores, enquanto que a clivagem da ligação 306 necessária para completa inativação de FVa é acelerada pelo co-fator de aPC, proteína S. Além da inativação dos fatores Va e VIIIa, aPC também é capaz de clivar PAR-1 quando em complexo com EPCR, uma atividade que parece estar envolvida em sua atividade apoptótica (ESMON, 2003). A atividade proteolítica de aPC no plasma é regulada pelas SERPINS: inibidor de PC (PCI), inibidor de ativador de plasminogênio-I (PAI-I) e inibidor- α_1 de proteinase (α_1 -PI). A velocidade de inativação de aPC por estas SERPINS é bastante lenta (na ordem de 10^1 - 10^3 $M^{-1}.s^{-1}$), mas na presença de heparina a inibição por PCI é acelerada 1000x.

Desde a elucidação da estrutura cristalográfica de des-Gla aPC (MATHER *et al.*, 1996), dados funcionais e de mutagênese indicaram que resíduos básicos de três alças expostas na superfície constituem um exossítio de ligação a ânions que interage com substratos macromoleculares e inibidores de aPC e na interação do zimogênio com o complexo de ativação (KNOBE *et al.*, 1999; REZAIIE, 2003): alça 36 (cLys³⁷-cLys³⁹), alça-60 (cLys⁶², cLys⁶³) e a alça 70 (cArg⁷⁴, cArg⁷⁵ e cLys⁷⁸) (Fig. 96).

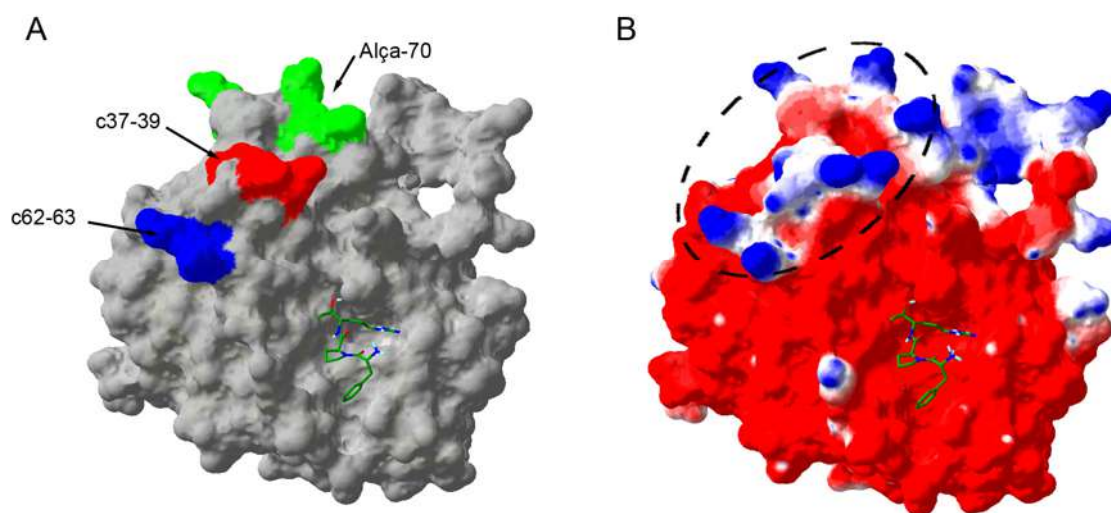


Fig. 96: Exossítio de ligação a ânions em aPC. A- localização dos resíduos básicos em aPC envolvidos na interação com co-fatores e inibidores. B- Potencial eletrostático mapeado na superfície de aPC (vermelho – negativo; branco – neutro; azul – positivo). Note a região de potencial positivo no exossítio de ligação a ânions. Para propósitos de orientação, um inibidor direcionado ao sítio ativo é mostrado (em palitos).

Resíduos na alça 36 e 70 são críticos na determinação da especificidade de clivagem da ligação lábil na subunidade A2 de FVa. Mais recentemente, foi

demonstrado por mutação sítio-dirigida que os mesmos resíduos estão envolvidos na determinação da especificidade de ambas as subunidades A1 e A2 do fator VIIIa (MANITHODY *et al.*, 2003). Na alça 70 também está situado o sítio de ligação ao Ca^{2+} (REZAIIE *et al.*, 1994), o qual foi demonstrado ser necessário para o aumento da especificidade de ativação de PC pelo complexo trombina-trombomodulina através numa mudança conformacional no peptídeo de ativação (YANG *et al.*, 2004) e estar energeticamente ligado ao subsítio S1 (SCHMIDT *et al.*, 2002). Mais recentemente, um epitopo de interação de PC com trombomodulina no complexo de ativação foi determinado, envolvendo os resíduos K38, R74 e R75 (GALE & GRIFFIN, 2004). Isto mostra que os epitopos de interação de aPC com os co-fatores e de PC com trombomodulina são amplamente sobreponíveis.

1.2.2.e Trombina e a modulação alostérica de SPCs pelo íon Na^+

Protrombina é certamente a mais fascinante das SPCs. Trombina é a enzima que promove a reação final na cascata da coagulação, convertendo o fibrinogênio solúvel no coágulo de fibrina. Trombina cliva a molécula solúvel de fibrinogênio, com estrutura $(\text{A}\alpha, \text{B}\beta, \gamma)_2$, liberando 4 fibrinopeptídeos (dois FpA e dois FpB). Isto produz o monômero de fibrina $(\alpha, \beta, \gamma)_2$, cuja polimerização espontânea forma o coágulo, frouxo, insolúvel nascente. O mecanismo aceito de liberação de fibrinopeptídeos (AROSIO *et al.*, 2000) estabelece que FpA é liberado primeiro do fibrinogênio levando à formação de monômeros de fibrina I. Estes monômeros agregam para formar protofibrilas de fibrina I, a partir das quais FpB é liberado para dar origem a protofibrilas de fibrina II, a qual forma o arcabouço do coágulo de fibrina. Desta forma, os epitopos de ligação ao fibrinogênio e fibrina à trombina podem ser mapeados de maneira direta, ao se estudar a liberação de fibrinopeptídeos.

Somente 4 das 376 potenciais ligações Arg/Lys-Xaa no fibrinogênio são clivadas pela trombina, destacando a elevada especificidade desta interação. A forma ativa da enzima consiste apenas no domínio catalítico, diferentemente das demais SPCs que apresentam domínios acessórios na cadeia leve. Isso faz com que a complexidade de funções cravejadas na seqüência desta serinil protease seja ainda maior, i.e., todos os exossítios contendo os epitopos responsáveis pela interação com os diversos co-fatores, substratos e inibidores estão montados no domínio catalítico da protease. Como conciliar a presença destes sítios de modulação (muitos alostéricos) com a função

catalítica num mesmo enovelamento deve ter sido um dos casos de evolução molecular mais formidáveis exercidos pela natureza.

A protrombina (PT) humana é uma glicoproteína composta por uma cadeia de 579 resíduos de aminoácidos. A conversão de PT na forma ativa α -trombina pelo complexo protrombinase ocorre inicialmente pela clivagem em Arg³²⁰ (ou cArg¹⁵) dando origem ao produto intermediário meizotrombina. Isto é seguido pela clivagem em Arg²⁷¹ para produzir α -trombina, a sofrer clivagem autocatalítica em Arg²⁸⁴ para fornecer des(Thr²⁷²-Arg²⁸⁴)- α -trombina. A forma estável de α -trombina é então composta pelos resíduos originais da protrombina que vão de Thr²⁸⁵ a Arg³²⁰, correspondendo à cadeia A (ou leve), e por Ile³²¹ até Glu⁵⁷⁹ na cadeia B (ou pesada). A cadeia A, ligada por uma ponte dissulfeto, é parte integral da estrutura da enzima, mas não está envolvida na catálise. Trombina apresenta atividades procoagulantes e anticoagulantes. As atividades favoráveis à coagulação são: conversão de fibrinogênio à fibrina, ativação dos fatores V, VIII, XI e XIII (usando fibrina como co-fator) e ativação de PARs 1 e 4 (empregando GpIb α como co-fator, uma glicoproteína da superfície de plaquetas). Por outro lado, quando a trombina se liga à trombosmodulina a enzima passa por uma incrível transformação, adquirindo funções anticoagulantes, i.e.: ativação de proteína C e do inibidor de fibrinólise (TAFI), um zimogênio de carboxipeptidase B circulante no plasma cuja função é clivar resíduos de Lys no C-terminal de fibrina parcialmente digerida, contribuindo para a diminuição da incorporação e ativação do plasminogênio (LEUNG & HALL, 2000). A ligação à TM aumenta em 1000x a especificidade da trombina por PC e suprime a atividade sobre o fibrinogênio.

Um passo crucial para o entendimento da função da α -trombina foi dado com a elucidação de sua estrutura cristalográfica no final da década de 80 (BODE *et al.*, 1989). A estrutura foi resolvida numa resolução de 1.9 Å graças a progressos nos métodos cristalográficos e à observação de que cristais das formas inibidas com clorometil cetonas (FPRCK ou PPACK no caso da trombina) eram estáveis e difratavam até altas resoluções. As conseqüências estruturais das importantes inserções em alças na superfície do enovelamento quimiotripsina-símile foram reveladas (vide Fig. 92). As mais destacadas para a interação com substratos macromoleculares são a alça c60 e a inserção de 5 resíduos (cA149a-K149e) na alça de autólise (conhecida também por alça- γ). Pouco tempo depois, as estruturas dos complexos com hirudina, fibrinopeptídeo A,

peptídeos do receptor de trombina assim como fragmentos de protrombina foram determinados (revisado por [STUBBS & BODE, 1995](#)). Estes trabalhos, juntamente com estudos bioquímicos revelaram a presença de outras duas regiões importantes para o controle da função da trombina, i.e., o sítio de ligação ao Na^+ , um importante modulador alostérico, e um segundo exossítio de ligação à ânions (ABE-II), responsável pela ligação à heparina e glicosaminoglicanos (revisado mais recentemente por [DICERA, 2003](#); [HUNTINGTON, 2005](#)).

A heparina funciona como um adaptador para a atividade ótima do inibidor ATIII, acelerando a velocidade de inativação da trombina por um fator de 10^3 . Os determinantes estruturais envolvidos nas funções moleculares da trombina estão mapeados na **Fig. 97**. Os resíduos cY60A-cP60B-cP60C-cW60D restringem consideravelmente o acesso à fenda de ligação ao substrato, particularmente em S2, e sugeriu-se que estes estivessem implicados na pobre ligação de trombina à inibidores protéicos naturais de serinil proteases. Foi demonstrado também que o exossítio de ligação a ânions (ABE-I), envolvido na interação com fibrinogênio e como o inibidor hirudina, estivesse localizada na região rica em Arg, ao redor da alça c70. Diversos estudos revelaram ainda um papel para a região hidrofóbica formada por cI174, cW215, cL99, cH57, cY60A e cW60D na determinação da potência de inibidores contendo resíduos aromáticos em P3.

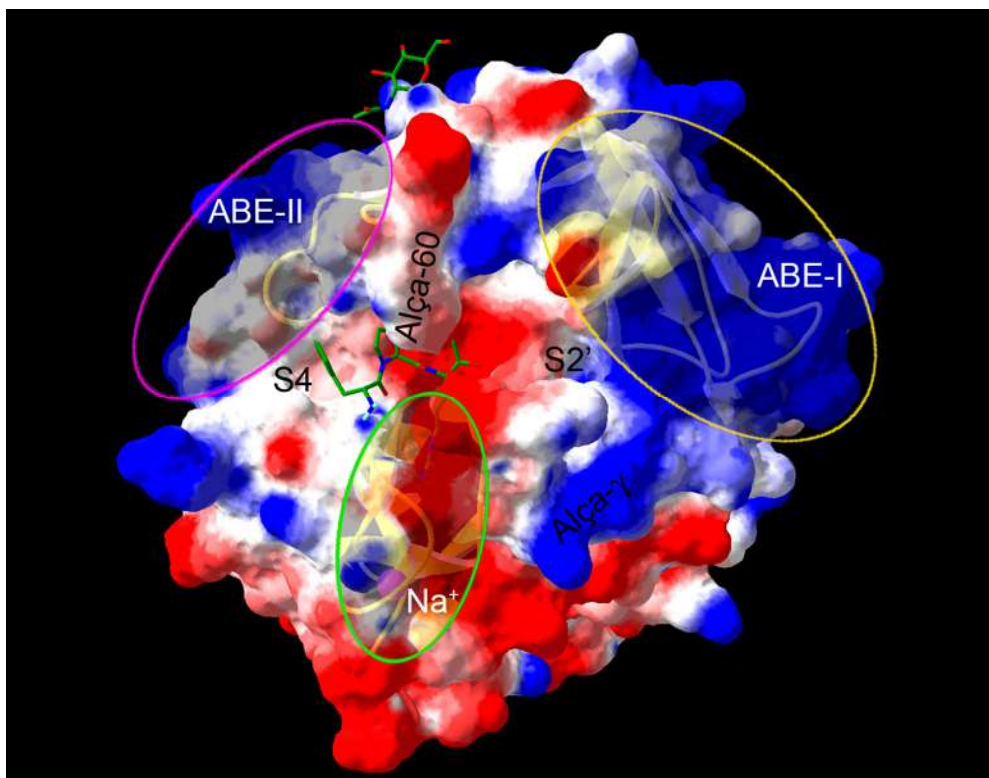


Fig. 97: Determinantes estruturais da especificidade de substrato na α -trombina. A superfície da cadeia B da enzima é mostrada com o potencial eletrostático mapeado (vermelho – negativo; branco – neutro; azul – positivo), com o inibidor PPACK ligado ao sítio ativo (palitos) e com o resíduo de N-acetilglicosamina em cAsn⁶⁰ (palitos). Os sítios alostéricos — ABE-I (exosítio de ligação a anions-I), ABE-II e de ligação ao Na⁺ são delimitados pelas áreas ovais onde a superfície foi feita transparente para destacar as alças e elementos de estrutura secundária responsáveis pela estrutura do sítio. O íon Na⁺ é representado como uma esfera de vdW em magenta. As regiões da superfície correspondendo às alças 60 e γ (ou alça de autólise, resíduos 140-150), importantes na restrição do acesso à fenda de ligação ao substrato (S4-S2'), também são destacadas.

A ligação de Na⁺ à trombina promove a ligação do substrato (aumento em k_1 , diminuição em K_m), juntamente com sua hidrólise (maior k_2 , maior k_{cat}) e é altamente dependente da temperatura, com uma grande mudança negativa na capacidade calorífica (revisado por [DI CERA, 2004](#)). O sítio de ligação ao Na⁺ da trombina é estrategicamente posicionado proximamente ao subsítio S1, entre as alças 1 e 2. Os sítios de ligação ao Na⁺ em FXa ([ZHANG & TULINSKY, 1997](#); [REZAIIE & HE, 2000](#)), aPC ([HE & REZAIIE, 1999](#)), FIXa ([SCHMIDT *et al.*, 2005](#)) e supostamente em FVIIa estão similarmente localizados e arranjados estruturalmente. Neste último, a presença do sítio de ligação ao sódio é predita pela ocorrência de um resíduo de Phe (ou Tyr em outras SPCs) e não Pro na posição c225 ([DANG & DI CERA, 1996](#)). As restrições conformacionais impostas por cP225 impedem a correta orientação do importante sítio de coordenação na posição 224. Na trombina, o Na⁺ ligado é

coordenado octaedricamente por dois átomos de oxigênio da carbonila da cadeia principal de cArg^{221a} e cLys²²⁴ e quatro moléculas de água ancoradas às cadeias laterais e cAsp¹⁸⁹, cAsp²²¹ e átomos da cadeia principal de cGly²²³ e cTyr^{184a}. Em termos gerais, nas demais SPCs, duas destas moléculas de água na esfera de coordenação são substituídas por grupamentos carbonila da cadeia principal dos resíduos nas posições c184a e c185.

Estudos de mutagênese, cinéticos e biofísicos acompanhados de cuidadosos experimentos de cristalografia na presença e ausência de Na⁺, mostraram que o aumento da competência catalítica na trombina tem natureza alostérica (DI CERA, 2004). A ligação do íon Na⁺ é convertida nesta maior aptidão catalítica através de um rearranjo na rede de ligações de hidrogênio do canal aquoso que liga o sítio de coordenação do íon monovalente até a base do subsítio S1 e ao resíduo de cSer¹⁹⁵ catalítico. É motivo de notável disputa na literatura que mudanças conformacionais muito tênues nos resíduos cArg¹⁸⁷, cAsp¹⁸⁹, cGlu¹⁹², cSer¹⁹⁵, cAsp²²¹ e cAsp²²², conforme observadas nas estruturas cristalográficas, tem sido responsabilizadas pelas marcantes alterações nas propriedades catalíticas da trombina (PINEDA *et al.*, 2002; HUNTINGTON, 2003; CARTER *et al.*, 2004; PINEDA *et al.*, 2004).

A K_d para ligação de Na⁺ na trombina é 110 nM, o que em condições fisiológicas de pH, temperatura e concentração de Na⁺ (140 mM) resulta em 60% da enzima na forma ligada ao sódio. A importância fisiológica do equilíbrio entre as formas rápida (ligada ao sódio) e lenta da trombina decorre da observação de que a forma livre de sódio assume papel anticoagulante, enquanto que a forma rápida tem natureza procoagulante. Em condições saturantes de Na⁺, trombina cliva os substratos fibrinogênio e PAR-1, com constantes de especificidade (K_s) 40-60 vezes maior que na ausência deste íon (onde a forma lenta predomina). A forma livre de sódio, por outro lado, mantém sua atividade normal contra PC, mas é incapaz de promover as clivagens de fibrinogênio e PAR-1 a velocidades fisiologicamente aceitáveis. O papel fisiológico do controle alostérico mediado por Na⁺ em trombina e em outras SPCs é enfatizado por resultados mostrando uma ligação energética entre o sítio de ligação deste íon e os ABEs, assim como o sítio de ligação ao Ca²⁺ (exceto trombina, onde a substituição cLys70 torna o sítio não-funcional), importantes na interação com co-fatores e inibidores estendidos (UNDERWOOD *et al.*, 2000; CAMIRE, 2002; SCHMIDT *et al.*, 2002; SCHMIDT *et al.*, 2005; MENGWASSER *et al.*, 2005).

1.2.3 Serinil proteases de venenos de serpentes (SVSP) ativas sobre a coagulação

1.2.3.a Componentes do veneno interagindo com o sistema hemostático

Existem mais de 200 espécies de serpentes venenosas, classificadas em diversas famílias, incluindo Elapidae, Viperidae, Crotalidae, Hydrophidae e Colubidae. O veneno é composto por uma rica mistura de proteínas e peptídeos (50-60 componentes), secretada e armazenada em glândulas (MARKLAND, 1998). A função biológica do veneno é tanto de defesa para a serpente como de rapidamente imobilizar, matar e pré-digerir a presa que servirá de alimento. Muitos componentes do veneno são comuns e relacionados dentro de cada família de serpentes. Por exemplo, neurotoxinas são típicas do veneno de Hydrophidae e Elapidae, enquanto que componentes ativos sobre a hemóstase são geralmente encontrados em serpentes Viperidae, Crotalidae e Elapidae (BRAUD *et al.*, 2000a). Por outro lado, variações quantitativas e/ou qualitativas na composição do veneno de uma espécie podem ocorrer em função de mudanças climáticas (sazonais) e geográficas (FRANCISCHETTI *et al.*, 1998; MUNEKIYO & MACKESSEY, 1998; SERRANO *et al.*, 2005).

Os venenos de serpentes contém uma ampla gama de componentes que afetam a hemóstase (Fig. 98). Muitos são enzimas tais como nucleotidases, fosfolipases A₂ (PLA₂), metalo e serinil proteases, enquanto que outros não possuem atividade catalítica, tais como desintegrinas e lectinas tipo-C (revisado por MARKLAND, 1998; BRAUD *et al.*, 2000a; LU *et al.*, 2005). Nucleotidases degradam o ADP, atuando como potentes inibidores da agregação plaquetária. Fosfolipases podem atuar como anticoagulantes ao degradarem PLs da membrana de plaquetas que atuam como cofatores dos complexos tenase e protrombinase, ou como procoagulantes ao liberarem fatores estimulantes da agregação plaquetária (e.g., ácido araquidônico) a partir da hidrólise de PLs. Desintegrinas são pequenas proteínas contendo o motivo RGD que atuam competindo especificamente com integrinas endógenas, tal como o receptor de plaquetas GPIIb/IIIa. Por outro lado, lectinas tipo-C ligam-se seletivamente à porção carboidrato de glicoproteínas na superfície de plaquetas ou de fatores da coagulação.

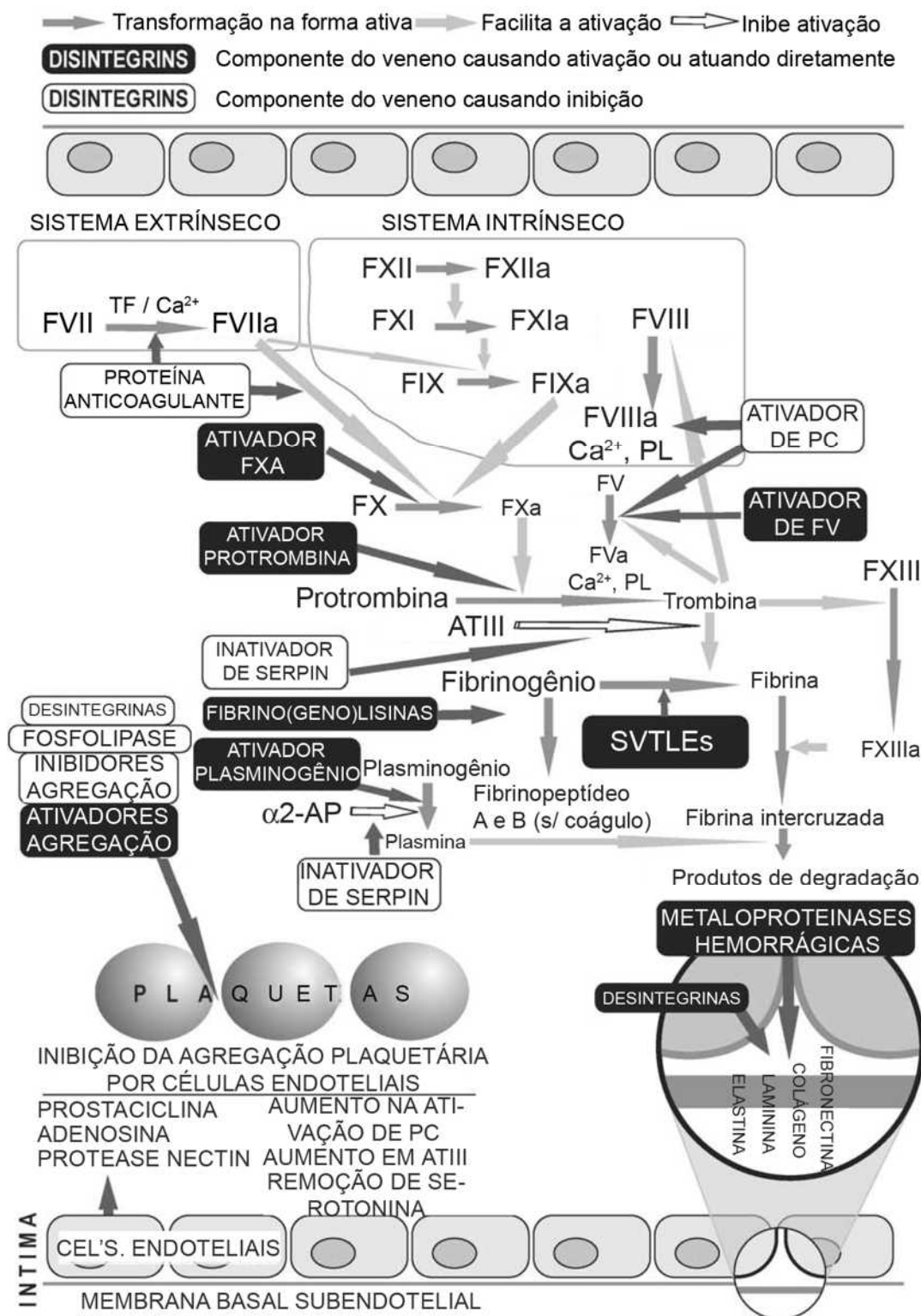


Fig. 98: Componentes do veneno de serpentes afetando a hemóstase. Adaptado de White (2005).

Atividades proteolíticas em venenos de serpentes puderam ser atribuídas, até o momento, à metalo e serinil proteases (revisado por MATSUI *et al.*, 2000; MARKLAND-JR & SWENSON, 2005). Estas enzimas interferem com o sistema hemostático da presa através de atividades fibrinogenolíticas, fibrinolíticas ou exibindo

ações mais específicas mimetizando fatores da coagulação, tais como ativadores de FV, FX e protrombina. Em particular, serinil proteases foram caracterizadas exibindo atividade calicreína-símile (cininogenases), trombina-símile (coagulam o fibrinogênio) ou ainda ativadora de plasminogênio. Metaloproteases, particularmente, foram encontradas apresentando atividade hemorrágica associada à degradação da matriz extracelular de vasos sanguíneos ou à clivagem de ligações específicas em proteínas da membrana basal do endotélio.

1.2.3.b SVSPs e enzimas trombina-símile do veneno de serpentes (SVTLEs)

SVSPs estão entre as enzimas atuando sobre o sistema hemostático melhor caracterizadas (revisado recentemente por [SERRANO & MAROUN, 2005](#)). Estas enzimas são bastante abundantes nos venenos de membros das famílias Viperidae e Crotalidae, podendo chegar a 20% do conteúdo protéico total ([BRAUD *et al.*, 2000a](#)), mas também estão presentes em Elapidae e Colubridae. SVSPs não são letais por si só, mas contribuem para o efeito tóxico quando associadas à outras proteínas do veneno. Todas as que foram seqüenciadas até o momento mostraram pertencer à família da quimiotripsina, sendo compostas por cerca de 240 resíduos de aminoácidos e contendo 12 cisteínas em 6 pontes dissulfeto intramoleculares. Apresentam graus de similaridade variando na faixa de 30-40% (identidade) com outras enzimas da família S1A, enquanto que dentro do próprio grupo, esta similaridade sobe para 50-60%, mesmo entre enzimas de diferentes espécies e/ou gêneros, apresentando funções distintas. O alinhamento das seqüências destas proteinases com os domínios catalíticos de serinil proteases de mamíferos mostra a presença de regiões variáveis restritas em SVSPs ([Fig. 99](#)). A presença de Asp na posição 189 denota que estas enzimas apresentam especificidade primária de substrato similar à tripsina. Apesar do elevado grau de similaridade na seqüência, SVSPs são bastante específicas contra um dado substrato macromolecular. Isto contrasta com a atividade menos específica da tripsina. O alinhamento da [Fig. 99](#) também revela que 5 das pontes dissulfeto em SVSPs são conservadas na tripsina, enquanto que o par dissulfeto restante envolvendo Cys⁹¹-Cys^{245e} é exclusivo deste grupo de enzimas.

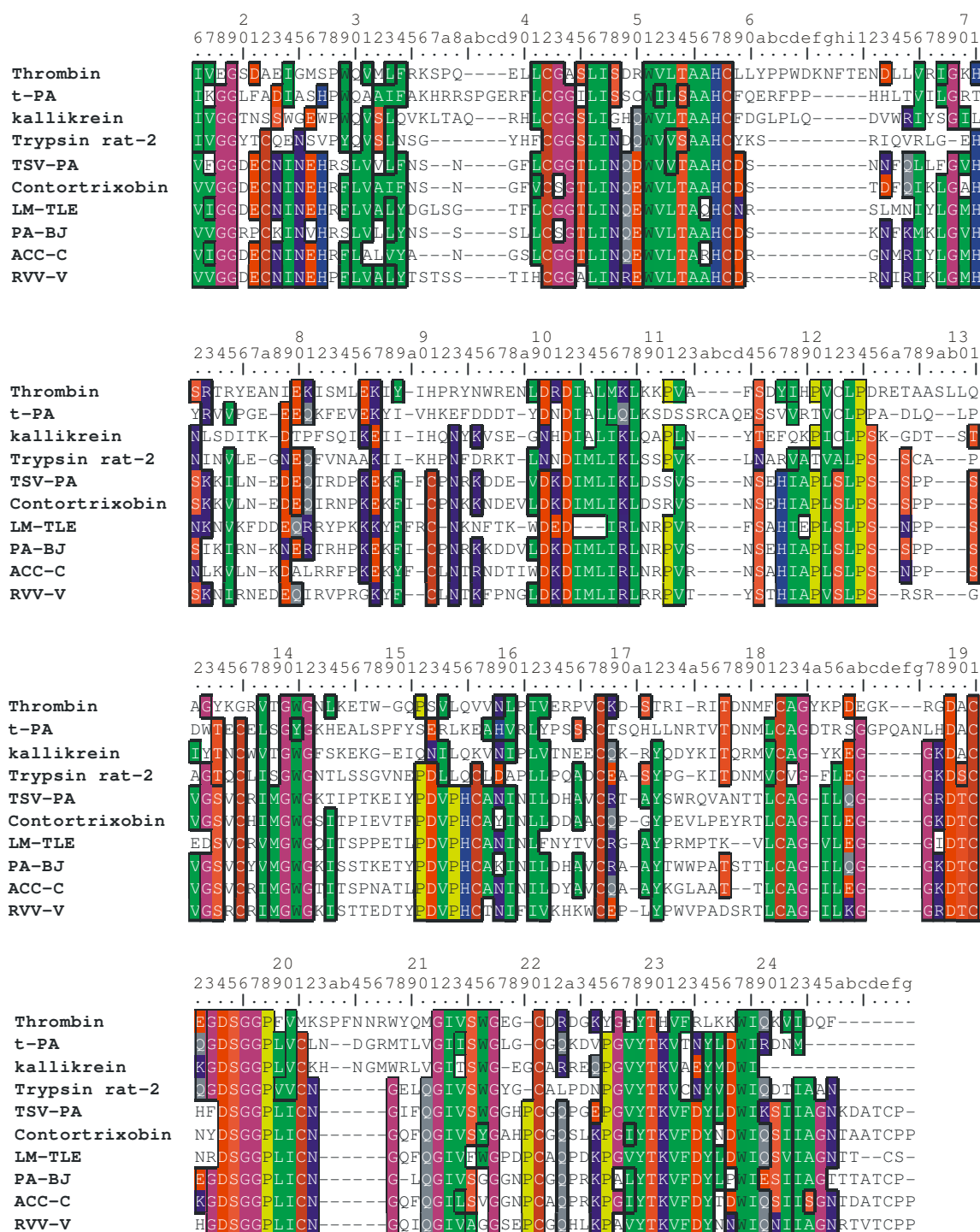


Fig. 99: Alinhamento entre SVSPs e serinil proteases de mamíferos selecionadas. TSV-PA = ativador de plasminogênio do veneno de *Trimeresurus stejnegeri* (PDB 1BQY); Contortrixobin = SVSP com atividades coagulante (libera fpB) e ativadora de FV e FXIII do veneno de *Agkistrodon contortrix* (SwissProt P82981); LM-TLE = enzima trombina-símile do veneno de *Lachesis muta* (SwissProt P33589); PA-BJ = ativador de PAR1 e PAR4 do veneno de *Bothrops jararaca* (SwissProt P81824); ACC-C = ativador de proteína C do veneno de *A. contortrix* (SwissProt P09872). RVV-V = ativador alfa de fator V do veneno de *Vipera russelli* (SwissProt P18964). As posições com 60% de conservação foram coloridas de acordo com a matriz de escore BLOSUM62 (HENIKOFF, 1992).

SVSPs apresentam graus variáveis de glicosilação, sendo mais comum a adição de carboidratos em resíduos de Asn (SERRANO *et al.*, 2005). Uma característica marcante de um subgrupo de SVSPs distinguido pela atividade fibrinogenolítica (inclui enzimas trombina-símile) é a resistência à inativação por extremos de calor e pH. Em algumas enzimas fibrinogenolíticas, a termorresistência está associada à presença de um alto conteúdo de carboidratos (>30% p/p) (SWENSON *et al.*, 2005). Uma característica mais geral de SVSPs é a purificação relativamente simples destas enzimas a partir do veneno. A maior parte dos procedimentos inicia por uma separação cromatográfica baseada em peso molecular, seguido de troca iônica ou afinidade (benzamidina- ou arginina-sefarose) (SERRANO *et al.*, 2005). As enzimas isoladas apresentam valores de pI variando desde a faixa ácida até valores de pH alcalinos, devido à variações na seqüência de aminoácidos e composição de carboidratos. A purificação é mais comumente acompanhada por ensaios com substratos cromogênicos derivados da arginina tendo uma ligação éster (e.g., BAME, TAME ou BAEE) ou amida (e.g., BApNA) lábil. As enzimas purificadas são então caracterizadas como pertencentes à classe das serinil proteases ao serem submetidas a ensaios de inibição com um painel de reagentes diagnósticos, como DFP, PMSF, TLCK/TPCK, benzamidina, etc. Uma etapa importante da caracterização da função biológica de SVSPs consiste na realização de ensaios enzimáticos contra substratos macromoleculares de importância na hemóstase, tais como, fibrinogênio, cininogênio, fator V, proteína C, receptores presentes na superfície de plaquetas, etc.

A maior parte das SVSPs caracterizadas até o momento são capazes de coagular o fibrinogênio (SERRANO & MAROUN, 2005). Diferentemente de enzimas fibrino(geno)líticas ordinárias, que também incluem metaloproteínases, muitas SVSPs clivam apenas o fibrinogênio intacto, liberando um ou ambos os fibrinopeptídeos A e B. Uma vez que estas proteases mimetizam esta atividade procoagulante da trombina mamífera, estas tem sido denominadas “enzimas trombina-símile” (SVTLEs). No entanto, as atividades de SVTLEs sobre o fibrinogênio não são exatamente iguais à da trombina mamífera, já que muitas são incapazes de ativar FXIII ou geram fibrinopeptídeos com cinética diferente da trombina mamífera (quando não liberam apenas um destes), produzindo coágulos que são prontamente removidos da circulação (revisado por CASTRO *et al.*, 2004). De fato, o termo SVTLE não é apropriado para descrever este importante subgrupo de SVSPs, pois estas enzimas não são

verdadeiramente equivalentes à trombina de mamíferos. Primeiramente, para qualquer SVSP, não há maior similaridade na seqüência com uma enzima funcionalmente relacionada do que com qualquer outra SVSP. Por exemplo, SVTLEs apresentam 30-40% de identidade com a trombina humana, mas possuem mais de 50% de identidade com uma SVSPs não funcionalmente relacionada, como TSV-PA. Segundo, SVTLEs não correspondem às enzimas responsáveis pela coagulação do fibrinogênio em serpentes, sendo incapazes de coagular o plasma retirado de outras serpentes ou mesmo da própria serpente que produziu o veneno (JOSEPH *et al.*, 2002). Por fim, outras atividades específicas desempenhadas pela trombina de mamíferos são mimetizadas por SVSPs, tais como ativação de FV e proteína C e estimulação da agregação plaquetária, e nem por isso estas enzimas são denominadas “trombina-símile”. Tendo feito estas ressalvas, o termo SVTLE será mantido neste trabalho por razões históricas e devido à ampla disseminação desta denominação para SVSPs capazes de coagular o fibrinogênio.

1.2.3.c As SVTLEs de surucucu (*Lachesis muta*): Lmm-TLE e Lmr-47

As serpentes da espécie *L. muta*, conhecida no Brasil como surucucu, atingem até 3m de comprimento. São distribuídas em florestas tropicais desde o sul da Nicarágua até a Costa Rica e Panamá na América Central (*L. m. stenophrys*, *L. m. melanocephala*), na região equatorial da Amazônia (*L. m. muta*) e na floresta Atlântica Brasileira, desde o estado do Rio de Janeiro até a Paraíba (*L. m. rhombeata*). Os venenos das serpentes da família Viperidae destacam-se por suas atividades hemorrágicas, coagulantes e proteolíticas. De fato, o acidente ofídico com surucucu apresenta semelhanças com o causado por jararacas (*Bothrops sp.*), apesar do soro antibotrópico não ser eficaz na neutralização destas atividades no envenenamento por *Lachesis muta* (JORGE *et al.*, 1997). Serinil proteases têm sido isoladas do veneno destas serpentes, muitas das quais com atividade coagulante (ARAGON-ORTIZ, 1986; ARAGON-ORTIZ, 1988; DA SILVA *et al.*, 1989; YARLEQUE *et al.*, 1989; SILVEIRA *et al.*, 1989; ARAGON-ORTIZ & GUBENSEK, 1993; AGUIAR *et al.*, 1996).

A melhor caracterizada destas enzimas é uma SVTLE originalmente isolada do veneno de *L. m. muta* através de gel filtração seguida de uma etapa de afinidade utilizando agmatina como ligante, uma amina gerada a partir da descarboxilação da L-Arginina (SILVEIRA *et al.*, 1989). Esta SVTLE, aqui denominada Lmm-TLE, migra

como uma banda difusa de 41-47 kDa em SDS-PAGE e apresenta pI variando de 3,1 a 5,0 para a enzima nativa e pI 8,5 para a enzima desialilada. Foi relatado que Lmm-TLE cliva exclusivamente a cadeia A α do fibrinogênio, sendo incapaz de ativar FXIII. Lmm-TLE, assim como a serinil proteinase de 60 kDa isolada do veneno de *L. m. muta* por Da Silva *et al.* (1989), induz episódios de giro ao redor do eixo longitudinal de cobaias (atividade giroxina-símile). A sequência de aminoácidos completa da glicoproteína foi posteriormente determinada por degradação de Edman, revelando 4 sítios potenciais de N-glicosilação em 228 resíduos (MAGALHÃES *et al.*, 2003). Foi demonstrado que os 25 resíduos N-terminais são idênticos ao N-terminal da giroxina de 60 kDa previamente isolada do mesmo veneno.

Uma SVTLE com 45-47 kDa foi isolada do veneno da subespécie nativa da floresta Atlântica Brasileira, *L. m. rhombeata* (Lmr-TLE ou Lmr-47; AGUIAR *et al.*, 1996), mostrando características semelhantes às SVTLEs purificadas a partir do veneno de *L. m. muta* da região da Amazônia Peruana (P_Lmm-TLE; YARLEQUE *et al.*, 1989) e Brasileira (B_Lmm-TLE; SILVEIRA *et al.*, 1989). Lmr-TLE libera preferencialmente fpA a partir do fibrinogênio, também possui pI na faixa ácida (3,1-5,0) e apresenta os 30 resíduos de aminoácidos N-terminais idênticos à B_Lmm-TLE. Como P_Lmm-TLE, a enzima da subespécie da floresta Atlântica foi relatada com 13% de carboidratos. Outras importantes propriedades de Lmr-TLE também foram reveladas através de estudos cinéticos com substratos peptídicos cromogênicos e ensaios de susceptibilidade a inibidores específicos. A atividade esterásica de Lmr-TLE sobre TAME (N-tosil-L-Arg metil éster) é 3 ordens de magnitude maior do que a atividade amidásica sobre BapNA (N-benzoil-DL-Arg-*p*-nitroanilida), fato que pode surgir exclusivamente da maior reatividade natural de ésteres. Mais informativa é a observação de que a constante de especificidade (K_s) de Lmm-TLE é praticamente inalterada frente aos substratos estendidos NBz-PFR-pNA e NBz-VGR-pNA, desenhados para ocupar os subsítios S2 (P2=Phe/Gly) e S3 (P3=Pro/Val). Por outro lado, Lmr-TLE foi eficientemente inibida pelo inibidor irreversível de tripsina antipaína {[*(S)*]-1-carboxi-2-fenil]-carbamoil-Arg-Val-arginal} e menos pronunciadamente por leupeptina (acetil-leucil-leucil-arginal). Estes inibidores apresentam Val e Leu em P2 e Arg e Leu em P3, respectivamente, indicando que a baixa seletividade pelos substratos cromogênicos estendidos testados pode ter sido resultante de uma alta especificidade para resíduos alifáticos ramificados na bolsa S2 e por resíduos volumosos em S3.

Mais recentemente, Magalhães *et al.* (2003) relataram uma extensa comparação das propriedades enzimáticas e bioquímicas das SVTLEs purificadas dos venenos de *L. m. muta* Peruana e Brasileira. Os autores concluíram que as duas enzimas são praticamente idênticas, podendo representar uma única enzima diferindo pelo grau de glicosilação. Tanto B_Lmm-TLE como P_Lmm-TLE apresentaram PM de 27 kDa após todo o carboidrato ser retirado pela enzima PNGase F. Este resultado (42% de carboidratos totais) contraria os relatados anteriormente para P_Lmm-TLE (YARLEQUE *et al.*, 1989), indicando que a análise química pode subestimar o conteúdo de carboidratos nestas enzimas. Mais importante, os autores também reviram a atuação das enzimas sobre o fibrinogênio, concluindo desta vez que estas não liberam especificamente, mas sim, preferencialmente fpA e que liberam um peptídeo de 4,5 kDa oriundo da clivagem em Arg⁴²-Ala da cadeia B β . Estes resultados dão suporte à hipótese de que Lmr-TLE e as variantes de Lmm-TLE tratam-se da mesma enzima, uma vez que a primeira também libera preferencialmente fpA do fibrinogênio (AGUIAR *et al.*, 1996).

Assim, semelhanças nas propriedades bioquímicas e estruturais (parciais) de Lmr-TLE e B_Lmm-TLE permitem que estas enzimas sejam consideradas idênticas, como de fato será feito no presente trabalho. Suporte adicional a esta hipótese provém do fato que a própria separação da espécie *L. muta* em suas 4 subespécies é atualmente alvo de discussão de taxonomistas (ZAMUDIO & GREENE, 1997). Baseados em análise comportamental, morfológica e molecular (filogenia do DNA mitocondrial) de espécimes do gênero *Lachesis*, foi concluído que três espécies distintas devem ser reconhecidas, *L. stenophrys*, *L. melanocephala* e *L. muta*. O estudo concluiu também que a distinção entre as subespécies Sul-americanas carece de nitidez, provavelmente devido a um fluxo genético que teria ocorrido quando estas repartiam o mesmo habitat no Pleistoceno.

1.2.3.d Relações estrutura-função em SVSPs

A primeira estrutura 3D determinada de uma SVSP foi a do ativador de plasminogênio do veneno de *Trimeresurus stejnegeri* (TSV-PA; PARRY *et al.*, 1998). A estrutura do aduto com EGR-cmk foi determinada com 2,5 Å de resolução e mostrou o padrão de seis ligações dissulfeto e a extensão C-terminal única de SVSPs (Fig. 100). Mutações nas alças 30, 70, 90 e 140 que controlam o acesso à fenda de ligação ao

substrato em serinil proteases quimiotripsina-símile e na alça 2 suportando o subsítio S1, formam a base para a especificidade de substrato de TSV-PA. Mesmo antes da estrutura 3D experimental de TSV-PA estar disponível, modelos por homologia indicaram possíveis determinantes da especificidade de substrato desta enzima. [Zhang et al. \(1997\)](#), construíram mutantes onde resíduos da alça 90 foram substituídos por seus equivalentes em batroxobin (uma SVTLE) e tPA, descobrindo um importante papel de Asp⁹⁷ no reconhecimento molecular do plasminogênio.

Com a disponibilidade da estrutura 3D de TSV-PA, estudos de mutação sítio-dirigida foram planejados para verificar hipóteses sobre o papel de algumas propriedades estruturais na determinação de aspectos funcionais desta enzima. Em um destes trabalhos, as posições 192 e 193 foram alvo de mutações visando esclarecer a origem da especificidade de substrato de TSV-PA ([BRAUD et al., 2000b](#)). A posição 192 é importante para a especificidade de substrato em trombina, aPC e outras SPCs, enquanto que a posição 193 é ocupada por Gly em todas as SPCs e em batroxobin. A presença de um resíduo volumoso em 193 é responsável pela resistência do mesotripsinogênio a SBTI e poderia ser responsável pelo mesmo comportamento exibido por TSV-PA. Os resíduos de TSV-PA (His¹⁹² e Phe¹⁹³) foram substituídos pelos resíduos equivalentes em tPA (Gln¹⁹² e Gly¹⁹³) e batroxobin (Gly¹⁹² e Gly¹⁹³). Os resultados mostraram que a posição 193 tem importante papel na determinação da especificidade de substratos macromoleculares e susceptibilidade a inibidores protéicos, restringindo o acesso ao sítio S2'. De fato, o mutante F193G apresentou atividade cerca de 10x maior contra plasminogênio e tornou-se sensível à inibição por SBTI.

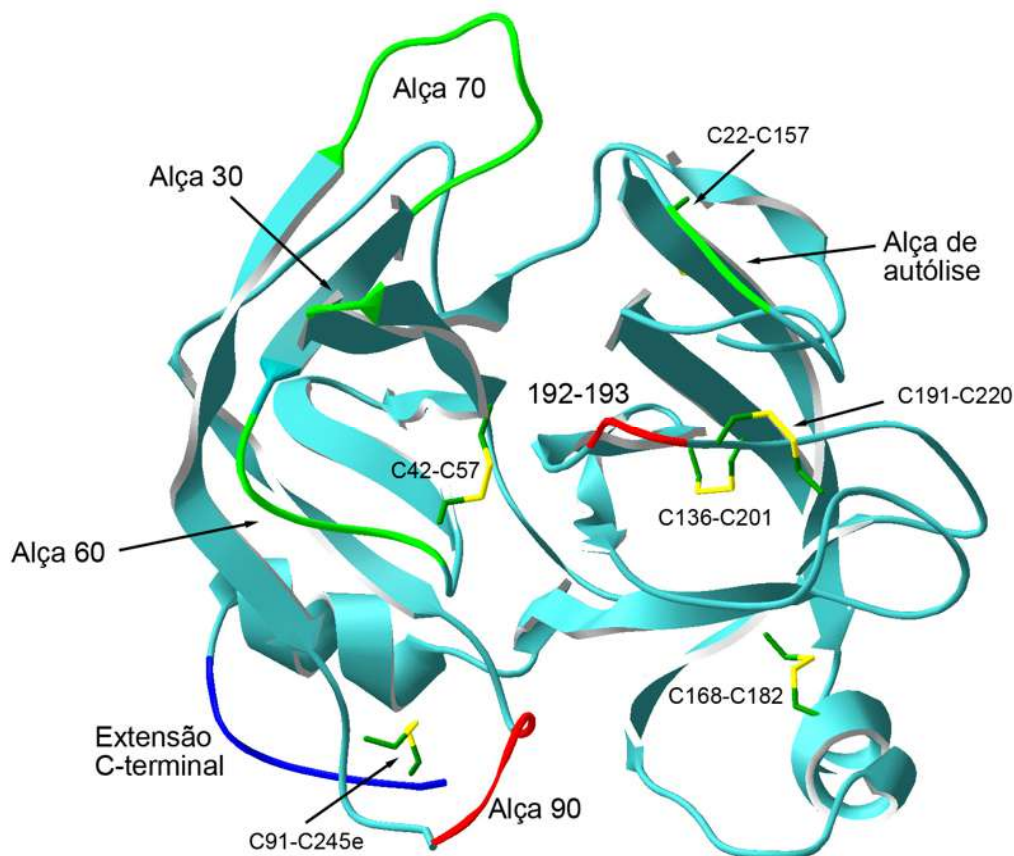


Fig. 100: Estrutura de uma serinil protease de veneno de serpentes. A estrutura de TSV-PA foi determinada com 2,5Å de resolução (PDB 1BQY; PARRY *et al.*, 1998). Note as pontes dissulfeto (palitos), a extensão C-terminal de SVSP (azul) e os determinantes de especificidade com evidência experimental (vermelho) ou preditos por homologia com serinil proteases da coagulação (verde).

Mesmo antes da elucidação estrutural de TSV-PA, modelos de SVSPs foram construídos com base na homologia com serinil proteases tripsina-símile. Um modelo simplificado, focando o subsítio S1, foi construído para Lmm-TLE num estudo que os autores denominaram de “análise topográfica da estrutura do sítio-ativo” (MAGALHÃES *et al.*, 1993). A potência dos inibidores competitivos do tipo amidina e guanidina (2/3 dos quais eram derivados simples de benzamidina) foi interpretada em termos do paradigma da interação dipolar de Mares-Guia e um modelo de regressão linear, utilizando constantes sigma de Hammett, capaz de explicar cerca de 60% dos dados experimentais, foi gerado. Posteriormente, um modelo completo de Lmm-TLE foi construído com base na homologia com tripsina e trombina, e usado na interpretação de dados de inibição com inibidores macromoleculares e da interação com fpA (CASTRO *et al.*, 2001). Interessantemente, o modelo foi gerado sem considerar a estrutura de TSV-PA, já disponível na época. Uma bolsa hidrofóbica composta por

Phe⁹⁵, Trp⁹⁹, Tyr¹⁷², Phe²¹⁴ e Trp²¹⁵ mostrou-se importante na interação com Phe⁸ e Leu⁹ de fpA. Os autores implicaram os resíduos Phe⁹⁵ e Trp⁹⁹ na restrição do acesso de BPTI ao subsítio S2 da enzima. Adicionalmente, foi ressaltada a presença do resíduo volumoso Arg¹⁹³ em Lmm-TLE, o qual analogamente à Phe¹⁹³ em TSV-PA poderia bloquear o acesso ao subsítio S2'. Ainda neste estudo, inibidores de trombina dependentes da interação em ABE-I, botrojaracin, hirugen e botroalternin, foram demonstrados serem ineficazes contra Lmm-TLE. Os autores sugeriram que um exosítio distinto do apresentado na trombina de mamíferos poderia ser responsável por estes resultados.

A elucidação da estrutura de TSV-PA consistiu num importante passo para o estudo da relação estrutura-função de SVSPs. Mais recentemente, as estruturas cristalinas de dois variantes estritamente relacionados de uma outra SVSP, idêntica ao ativador de plasminogênio de *Agkistrodon acutus* (DAV-PA), foram determinadas com 2,0 e 2,1 Å de resolução (ZHU *et al.*, 2005). Estas estruturas viabilizam a construção de modelos por homologia mais precisos para os membros restantes da família de SVSPs e portanto contribuem para um melhor entendimento da função destas enzimas. Entretanto, a determinação das estruturas de membros apresentando funções biológicas distintas, tais como SVTLEs, ativadores de FV e PC, irá certamente acelerar e incentivar mais estudos visando esclarecer a relação estrutura-função neste interessante grupo de serinil proteases ativas sobre a coagulação. Isto não deve estar muito longe de ocorrer, uma vez que a cristalização de uma SVTLE já foi relatada (BORTOLETO *et al.*, 2002).

1.2.4 Serinil proteases como fonte de novas abordagens para o tratamento e diagnóstico de distúrbios da coagulação

1.2.4.a Importância das patologias afetando os fatores da coagulação

Em última análise, a hemóstase consiste no balanço entre sangramento e trombose. As desordens mais comuns e graves do sistema hemostático advêm de hemofilias ou como conseqüência de complicações após agressões externas, desde os efeitos graduais e cumulativos de hábitos de vida arriscados (fumo, dieta rica em gorduras, etc.) até efeitos adversos de medicamentos e acidentes com animais peçonhentos, como o envenenamento ofídico (PATRUSHEV, 2002; KHRENOV *et al.*, 2002; LEE & LIP, 2003; WHITE, 2005; DAHLBÄCK, 2005). As formas mais comuns

de hemofilia são devidas a diferentes tipos de defeitos genéticos (deleções, inserções, inversões ou mutações pontuais) resultando em deficiência hereditária de FVIII (hemofilia A) ou FIX (hemofilia B; [PATRUSHEV, 2002](#)). Os genes de FVIII e FIX localizam-se no cromossomo X, em contraste com os fatores de coagulação restantes que estão localizados em cromossomos autossômicos. Isto explica porque, salvo raras exceções, somente machos são afetados por hemofilia A ou B, e porque hemofilias devido a outros fatores da coagulação são herdadas como tratos recessivos em ambos os sexos. A prevalência de hemofilia A na população masculina é aproximadamente 1/5000 e de hemofilia B é 1/30.000. Em cerca de 30% dos hemofílicos, a doença surge de novas mutações, explicando porque frequentemente esta acomete garotos nascidos em famílias sem histórico prévio de hemofilia. Hemofilias A ou B ocorrem em formas brandas (5-30%), moderadas (1-5%) e severas (<1%) dependendo da concentração plasmática dos fatores VIII e IX, respectivamente. O mesmo tipo de correlação não é clara para deficiências nos demais fatores da coagulação. A hemofilia severa caracteriza-se por episódios aparentemente espontâneos de sangramento afetando juntas, músculos, órgãos internos e o cérebro.

Sangramentos patológicos também podem ocorrer devido ao desenvolvimento de auto-anticorpos contra fatores da coagulação (hemofilias adquiridas), os mais comuns sendo direcionados contra FVIII ([LOLLAR, 2004](#); [LOLLAR *et al.*, 2005](#)). Estas condições, as quais podem ser muito severas e ocasionalmente fatais, afetam principalmente pessoas idosas através de um mecanismo molecular ainda pouco entendido. Outras desordens do sistema hemostático surgem da deficiência de vitamina K, a qual é necessária para a biossíntese de muitos fatores da coagulação. Isto pode ocorrer por mal absorção desta vitamina lipossolúvel ou pelo consumo excessivo de antagonistas da vitamina K, tal como a Varfarina usada na terapia anticoagulante.

Um trombo pode impedir parcial ou totalmente os vasos sanguíneos, potencialmente privando tecidos de oxigênio e suprimento nutricional. Alternativamente, um trombo pode se soltar dando origem a um êmbolo que viaja com a corrente sanguínea, podendo entupir outro vaso. A trombose pode se manifestar de muitas formas, incluindo trombose venosa profunda e embolismo pulmonar, conhecidas coletivamente como tromboembolismo venoso (VTE), trombótico ou embólico e infarto do miocárdio ([GUSTAFSSON *et al.*, 2004](#)). Há três principais grupos de risco: hipercoagulabilidade sanguínea (como resultado de genótipo, câncer, mudanças

hormonais, cirurgia, trauma, etc.); estase do fluxo sanguíneo (causado por imobilidade, fibrilação arterial, etc.); e dano à parede dos vasos (tal como na aterosclerose). Estima-se que a incidência de VTE em caucasianos exceda 1 caso a cada 1.000 pessoas; 1,3-4,1 pessoas em 1.000 vivenciam um infarto; e ~5 em 1.000 sofrem um infarto do miocárdio anualmente (HEIT, 2003).

Aterosclerose é uma doença reconhecida como uma das principais causas de morte em países industrializados, principalmente devido a fatores de risco associados ao estilo de vida (LEE & LIP, 2003). Uma forma comum de aterosclerose caracteriza-se pelo rápido acúmulo focalizado de lipídeos com o desenvolvimento de um amplo depósito de gordura subendotelial recoberta por uma fina capa, mecanicamente frágil. Frequentemente, há pequena intrusão no lúmen do vaso, ocasionando um considerável remodelamento na parede do vaso para acomodar o acúmulo de lipídeos. Com a ruptura da capa, as células e fatores solúveis da coagulação são expostos a esse grande depósito de substâncias procoagulantes. A formação de trombo oclusivo no sítio de uma lesão é a maior causa de complicações de aterosclerose, tais como derrame e infarto do miocárdio, os quais são geralmente fatais (KHRENOV *et al.*, 2002). Em alguns casos, o trombo pode ser reabsorvido e a placa de ateroma pode evoluir para uma placa fibrosa avançada (algumas vezes calcificada) levando a estenose significativa e sintomas de angina pectoris estável.

Na aterosclerose, a inter-relação dos fatores inerentes à coagulação sanguínea, inflamação e sistema imune pode levar à rápida formação de um trombo massivo no sítio da lesão (TRACY, 2003). Devido à ampla disponibilidade de TF, a atividade de ambas as vias extrínsecas e intrínsecas estão aumentadas no sítio da lesão, a qual é acompanhada por um decréscimo na atividade de mecanismos anticoagulantes. Em relação a este último, a modulação negativa do EPCR e da expressão de trombomodulina na placa pode contribuir para o potencial protrombótico desta. Os mecanismos moleculares envolvidos no elevado potencial protrombótico da placa de ateroma ainda são pouco conhecidos, mas parecem envolver o depósito de lipoproteínas de baixa densidade oxidadas e corpos apoptóticos que favorece o acúmulo de fatores de coagulação nas superfícies fosfolipídicas no local da lesão (KHRENOV *et al.*, 2002).

A interferência com aspectos do sistema hemostático humano é um tema comum entre venenos de serpentes, englobando todas quatro principais famílias de serpentes

peçonhentas. Isto é refletido na importância clínica da coagulopatia que se segue ao acidente ofídico ao redor do globo: mais de 2,5 milhões de pessoas são picadas por serpentes anualmente, das quais cerca de 4% morrem de acidentes graves (WHITE, 2005). Deve-se notar que coagulopatias nem sempre constituem o efeito principal do veneno responsável pela morbidade ou mortalidade, mas podem atuar de modo sinérgico com outras atividades do veneno para o detrimento da saúde humana. O impacto de acidentes ofídicos é maior em áreas rurais dos trópicos, mas a invasão de áreas urbanas por serpentes venenosas é um problema crescente em países em desenvolvimento. O espectro de coagulopatias induzidas pelo envenenamento ofídico pode ser resumido em 4 quadros principais: (i) coagulabilidade sanguínea reduzida; (ii) franco sangramento devido ruptura de vaso sanguíneo; (iii) efeitos secundários do sangramento, tais como choque hipovolêmico e dano aos órgãos; e (iv) trombose patológica direta, tal como VTE. Coagulopatia pode se desenvolver rápido, mas nem sempre, com alguns casos tendo coagulação normal nas primeiras duas horas após a picada, mas apresentando severa coagulopatia acompanhada de desfibrinação após 4-8h. Além do sangramento local, pode haver sangramento gastrointestinal, com hematêmese e melena, sangramento cutâneo, causando equimoses, hematúria e menos comumente menorragia e sangramento intracraniano (WHITE, 2005).

Como pode ser notado pela descrição acima, a maior parte dos problemas clínicos do envenenamento ofídico está associado à hemorragia e incoagulabilidade do sangue. Entretanto, *in vitro*, muitos componentes do veneno podem atuar como procoagulantes, tais como SVTLEs e ativadores de FV, FX e PT (Fig. 98). A explicação para esta aparente discrepância é que *in vivo* estes fatores atuam consumindo os fatores da coagulação, tornando o sangue incoagulável. Deve-se ressaltar que alguns venenos contêm verdadeiros componentes anticoagulantes, tal como enzimas fibrinolíticas e ativadores de plasminogênio ou PC (Fig. 98).

1.2.4.b Aplicações práticas de SVSPs na hemóstase

Muitos dos componentes de venenos de serpentes têm sido empregados no diagnóstico e tratamento de desordens da hemóstase. De fato, o estudo destas toxinas deu importantes contribuições para a pesquisa básica da coagulação, conforme revisado recentemente (MARSH & WILLIAMS, 2005). SVTLEs são usadas para quantificar o fibrinogênio ou seus produtos de degradação no plasma e na detecção de disfunções

desta macromolécula. Muitas SVTLEs não ativam FXIII e processam o fibrinogênio diferentemente da trombina humana, resultando na formação de coágulos de fibrina friáveis e facilmente removíveis da circulação. Isto forma a base para a desfibrinação terapêutica realizada pelas SVTLEs batroxobin de *Bothrops atrox* (Reptilase®, Pentapharm, Basel, Suíça) e Ancrod de *Callosellasma rhodostoma* (Knoll, Ludwigshafen, Alemanha). Muitas SVTLEs não são inibidas por heparina nem reagem com SERPINS plasmáticas (CASTRO *et al.*, 2004; SERRANO & MAROUN, 2005). Assim, SVTLEs podem ser úteis na desfibrinação de amostras visando a quantificação de ATIII ou outros fatores hemostáticos, mesmo em amostras heparinizadas. Reptilase é utilizada como reagente alternativo ao ensaio do tempo de trombina, como um teste rápido do fibrinogênio em amostras contendo heparina.

Das quatro classes de ativadores de PT encontrados em venenos de serpentes, os tipos C e D são serinil proteases, sendo este último representado por homólogos estruturais de FXa (SERRANO & MAROUN, 2005). Estas enzimas têm sido utilizadas em estudos clínicos de variedades descarboxiladas de PT ocorrendo na deficiência de vitamina K por doença hepática ou tratamento com anticoagulantes, em deficiências de PT e no diagnóstico de coagulação intravascular disseminada (MARSH & WILLIAMS, 2005). Ativadores de PT ainda têm sido usados como reagentes para a pesquisa, na geração de meizotrombina e outras formas inativas de trombina, assim como na ativação de PT recombinante. O ativador de FV do veneno de *Vipera russelli* (RVV-V) difere da trombina na medida em que esta cliva a glicoproteína plasmática após Arg¹⁵⁴⁵, enquanto que a enzima humana cliva FV em dois outros pontos (Arg⁷⁰⁹ e Arg¹⁰¹⁸; LU *et al.*, 2005). RVV-V tem recebido aplicação na pesquisa para o mapeamento dos eventos moleculares operando na ativação de FV, devido à sua ação seletiva (MARSH & WILLIAMS, 2005). Ensaio funcionais de PC são problemáticos uma vez que a ativação por trombina pode ser incompleta e esta última pode interferir em ensaios cromogênicos. Estes problemas são contornados pelo emprego de ACC-C, um potente ativador de PC do veneno de *Agkistrodon contortrix* (Protac®, Pentapharm). Além disso, Protac tem sido extremamente útil na pesquisa da via anticoagulante da PC, em ensaios de Proteína S e na ativação de PC recombinantes (MARSH & WILLIAMS, 2005).

1.2.4.c Potencial de serinil proteases no desenvolvimento de terapias para doenças do sistema hemostático

Serinil proteases atuam em ambos os lados do fiel da balança representada pela hemóstase. Tanto o sangramento descontrolado como as doenças relacionadas à trombose representam graves ameaças à saúde. Enzimas selecionadas da família da quimiotripsina atuam em diferentes pontos do sistema hemostático e são sujeitas a um conjunto de mecanismos moleculares de controle na modulação de suas atividades procoagulantes e anticoagulantes.

Deficiências nos fatores da coagulação em pacientes sofrendo de hemofilia (hereditária ou adquirida) e quadros hemorrágicos decorrentes de trauma físico ou cirúrgico são tradicionalmente tratados por transfusão de sangue, ou seus derivados, como plasma congelado fresco (FFP), hemácias, plaquetas e fatores da coagulação purificados do plasma (HOLCOMB *et al.*, 2004; HESS & HIIPPALA, 2005). Desde os anos 70, concentrados de complexo protrombina ativado (aPCC) ou não-ativado (PCC) tem sido empregados no tratamento de complicações de hemofilias A e B, doença hepática e reversão de terapia antitrombótica (YOLKEN & HILGARTNER, 1978; LUSHER *et al.*, 1984; YASAKA *et al.*, 2005). O desenvolvimento de alo-anticorpos inibitórios contra fatores VIII e IX em pacientes com formas severas de hemofilia A e B, respectivamente, tratados com reposição dos fatores purificados do plasma, constituem a principal complicação no tratamento destas doenças (GREEN, 1999). Nestes casos, PCCs e aPCCs tais como FEIBATM (Baxter Health Care Corp., Glendale, EUA), AUTOPLEXTM (Nabi, Boca Raton, EUA) e OCTAPLEXTM (Laboratórios Clausen, Uruguai) representam a rota alternativa de tratamento (LEISSINGER, 1999). No entanto, um excesso de episódios trombóticos tem sido observados nos pacientes submetidos à terapia com PCCs ou aPCCs, incluindo coagulação intravascular disseminada (DIC), VTE, MI (LUBETSKY *et al.*, 2004). PCCs contém protrombina, FIX e FX, com quantidades variáveis de FVII. Estudos têm verificado que os episódios trombóticos estão relacionados a níveis variáveis de FIXa nestas preparações (LOWE, 2001).

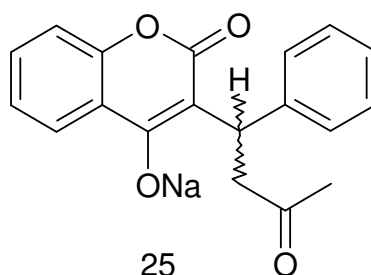
Entretanto, com o aparecimento do HIV na década de 80, o ressurgimento da hepatite C e mais recentemente a descoberta dos prions, a preocupação com o risco de transmissão de infecções pelo tratamento a base de medicamentos preparados a partir do

sangue humano aumentou consideravelmente. Assim, a busca por estratégias que reduzam ou eliminem a transfusão de componentes do sangue para indivíduos com distúrbios hemorrágicos torna-se extremamente desejável uma vez que estas têm o potencial de reduzir a morbidade, mortalidade e custo associados (LOWE, 2001; KESSLER, 2004; HESS & HIIPALA, 2005). Neste sentido, duas SPC têm sido alvo do desenvolvimento de proteínas recombinantes terapêuticas: rFVIIa (KESSLER, 2004) e rFIXa (LOWE, 2001). rFVIIa é sem dúvida o mais amplamente estudado e tem sido considerado como um coagulante de amplo espectro. O concentrado de rFVIIa é inicialmente produzido na forma do zimogênio em células baby do rim de hamster, transfectadas com o gene humano (HEDNER, 2004). A proteína é secretada no meio de cultura e ativada autoproteoliticamente, gerando a forma de duas cadeias que é purificada por cromatografia de troca iônica. Vírus de envelope lipídico são destruídos por tratamento com Triton X-100. O produto final é estruturalmente idêntico ao fator purificado do plasma humano, exceto por diferenças minoritárias no conteúdo de ácido siálico e fucose e pela presença de um resíduo de Gla com níveis variáveis de γ -carboxilação. rFVIIa é logicamente livre de vírus oriundos do sangue e não induz uma resposta de alo-anticorpos em indivíduos desprovidos de FVII plasmático. Além disso, rFVIIa possui potencial trombogênico equivalente a aPCCs. Por todas estas propriedades desejáveis, o uso de rFVIIa tem sido ampliado do alvo original de seu desenvolvimento (reversão do sangramento em hemofilias congênitas A e B, complicadas pelo desenvolvimento de alo-anticorpos inibitórios) para o tratamento de hemofilias adquiridas (contra FVIII e FIX), reversão de tratamento anticoagulante com antagonistas de vitamina K e inibidores diretos e indiretos de trombina (KESSLER, 2004).

Pelo que foi exposto acima fica evidente que o futuro da terapia anti-hemorrágica está no desenvolvimento de biofármacos recombinantes. De fato, a produção de fatores da coagulação recombinantes para o controle de distúrbios hemorrágicos ocorre há mais de uma década (GILL, 1993). Mais recentemente, SPCs entraram na alça de mira dos projetos visando obter biofármacos ativos sobre o sistema hemostático mais eficientes através do emprego da engenharia de proteínas (GRUBER *et al.*, 2002; CURTIS *et al.*, 2005). O desenvolvimento de fatores da coagulação mutantes com propriedades projetadas para aumentar a segurança e eficácia de seu uso é extremamente desejável. Uma alternativa para alcançar este objetivo pode estar no uso

de SVSPs. Muitas destas enzimas representam versões funcionalmente simplificadas de SPCs humanas, onde apenas uma ou poucas de suas funções bioquímicas estão presentes. Devido à dificuldades na humanização de proteínas de outras fontes e o conseqüente risco de desenvolvimento de reações imunológicas indesejáveis, uma alternativa viável seria introduzir as modificações necessárias nas SPCs humanas, inspiradas pela estrutura de SVSPs, para criar enzimas recombinantes com atividades planejadas. Para que isto seja possível num futuro próximo, se faz cada vez mais necessário o estudo da relação estrutura-função em ambos os grupos de serinil proteases, de veneno de serpentes e humanas.

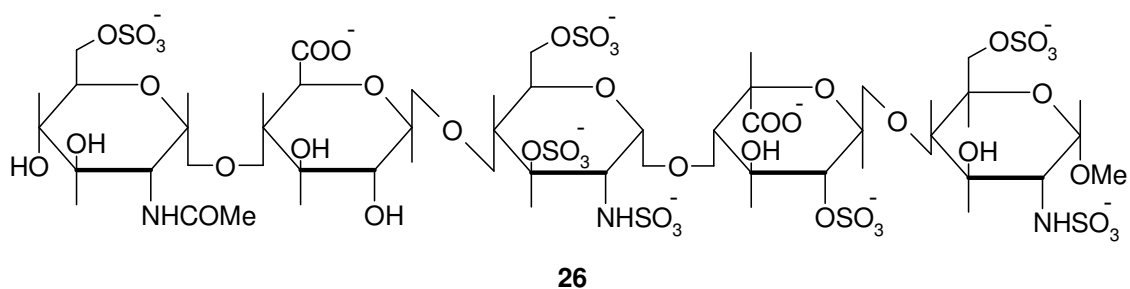
O outro lado da moeda na patofisiologia da hemóstase é representado pelas desordens tromboembólicas. As terapias atuais inibem a coagulação de diferentes formas (MAJERUS *et al.*, 2001). Heparinas e cumarinas pertencem a um grupo de anticoagulantes atuando indiretamente sobre a trombina e outros fatores da coagulação. As cumarinas inibem a síntese dos fatores dependentes de vitamina K. Varfarina (**25**) e outras cumarinas são os únicos anticoagulantes que podem ser dados por via oral, mas suas concentrações plasmáticas e efeitos anticoagulantes podem ser erráticos, como resultado de suas freqüentes interações fármaco-fármaco e fármaco-alimento; significando que pacientes precisam evitar medicação concomitante e manter uma dieta bem controlada. Além disso, como a janela terapêutica é estreita, varfarina requer o monitoramento contínuo do seu efeito anticoagulante para evitar a superdose e um conseqüente risco de sangramento severo, ou a subdose, resultando em falha da prevenção da trombose.



Heparinas atuam indiretamente na inibição da trombina ao induzirem mudanças conformacionais no inibidor endógeno ATIII e fazerem uma ponte entre este e a trombina através do ABE-II da enzima (Fig. 101). A interação com ATIII leva a um aumento da velocidade de inativação da trombina, FXa e outros fatores sensíveis à

SERPIN. Entretanto, uma desvantagem desta classe de anticoagulantes é que o complexo heparina:ATIII não pode atuar no complexo ternário heparina:trombina:fibrina que se forma normalmente no coágulo (**Fig. 101**). Como heparina se liga fortemente à fibrina e à trombina, é difícil para a heparina no complexo com ATIII deslocar o oligossacarídeo do complexo ternário e efetivamente aumentar a velocidade de inativação da protease pela SERPIN.

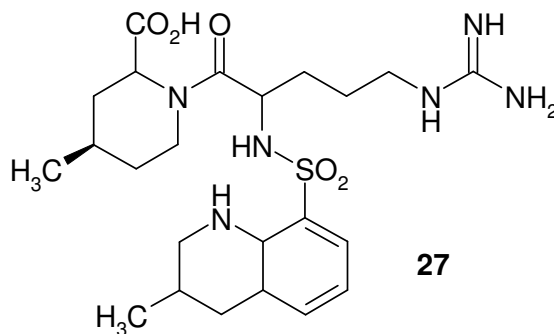
Heparina não-fracionada tem um efeito variável devido à interação com proteínas endógenas de forma que esta também requer monitoramento ([MAJERUS *et al.*, 2001](#)). Além disso, pacientes podem desenvolver uma reação imune conhecida como trombocitopenia induzida por heparina (HIT). Um avanço clínico foi conseguido com o uso de frações de baixo peso molecular da heparina (LMWH). Estas possuem massa molecular entre 4-5 kDa, comparado com a faixa de 3-30 kDa (média 15 kDa) para heparina não-fracionada. Mais recentemente, o núcleo ativo da estrutura da heparina passou a ser preparado sinteticamente (Fondaparinux, 1.728 Da; [BAUER *et al.*, 2003](#)). Fondaparinux (**26**) é na verdade um inibidor indireto de FXa, porque este é muito curto para fazer uma ponte entre ATIII e trombina (**Fig. 101**). Diferentemente de LMWHs, fondaparinux aumenta a velocidade de inativação de FXa por ATIII, bloqueando a geração de trombina, mas sem efeito sobre a inativação da trombina ([WEITZ & BATES, 2005](#)). LMWH e fondaparinux produzem respostas mais previsíveis e um ajuste da dose é possível. A frequência de HIT também é considerada baixa. No entanto, o elevado tamanho molecular ainda requer administração subcutânea, o que tende a inviabilizar o uso em longo prazo.



Hirudina, o inibidor direto de trombina (DTI) prototípico, foi originalmente isolado da sanguessuga medicinal, *Hirudo medicinalis* (revisado por [WARKENTIN, 2004](#)). Uma vantagem de DTIs é que estes podem inibir tanto trombina livre como a enzima ligada à fibrina (**Fig. 101**). Hirudina é um polipeptídeo de 65 aminoácidos (~ 7

kDa) que se liga à trombina com alta afinidade ($K_i \sim 10^{-13}$ M) bloqueando tanto o sítio ativo (através do sítio D) como ABE-I, formando um complexo estequiométrico 1:1. Hirudina recombinante (r-Hirudina ou Lepirudina), não possui um grupo sulfato em Tyr⁶³ e desta forma tem sido denominada dessulfatohirudina (desirudin). Hirudina nativa e desirudin precisam ser administradas por via parenteral. Bivalirudina (Hirulog), um oligopeptídeo de 20 aminoácidos, modelado com base na hirudina, consiste num domínio N-terminal D-Phe-Arg-Pro que interage com o sítio-ativo da trombina, um espaçador de glicinas, e um análogo dodecapeptídico do C-terminal de hirudina que se liga ao ABE-I da trombina. Hirulog também requer administração parenteral, com uma vida plasmática de 25min (contra 1 h de hirudina nativa ou recombinante).

Durante a década de 90, diversos DTIs foram sintetizados e testados quanto à sua eficácia *in vivo*, e o composto Argatroban, 508,7 Da (**27**), surgiu como a primeira molécula pequena a ser aprovada para uso na clínica (revisado por [KATHIRESAN *et al.*, 2002](#)). DTIs univalentes (direcionados somente ao sítio ativo) tem a vantagem de inibir mais eficientemente a trombina ligada à fibrina, pois não têm de competir pelo ABE-I como análogos de hirudina ([Fig. 101](#)). Entretanto, Argatroban necessita ser administrado por via parenteral devido à sua baixa biodisponibilidade oral.



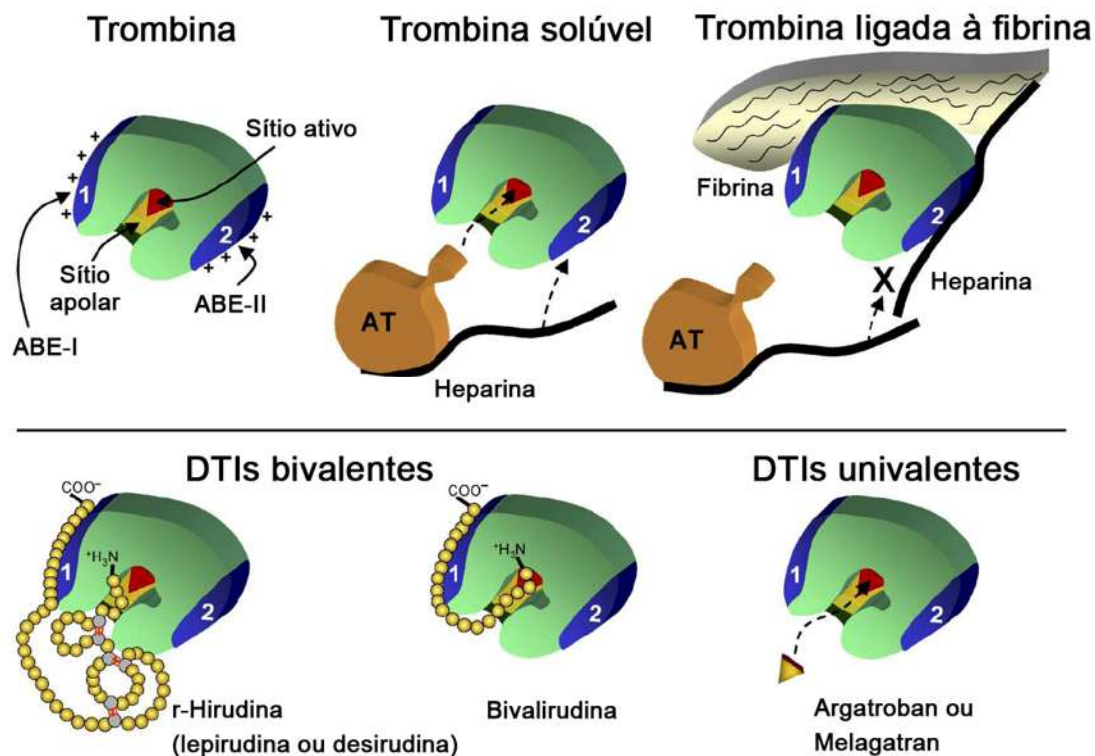
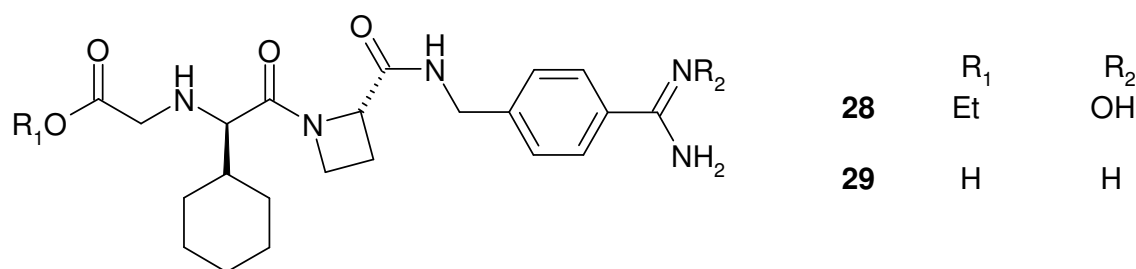


Fig. 101: Modo de ação de inibidores indiretos e diretos de trombina. Vide texto para maiores detalhes. Adaptado de [Warkentin \(2004\)](#).

É possível perceber, pelo exposto acima, que há uma forte necessidade médica de novos anticoagulantes orais, bem tolerados, com uma ampla janela terapêutica. Isto combinado com propriedades favoráveis de absorção, distribuição e eliminação que permitam uma proteção previsível sem a necessidade de monitoramento representa o eldorado do desenvolvimento de fármacos anticoagulantes. Estes fármacos são atualmente agrupados em duas classes: DTI orais e inibidores orais de FXa. Deve-se mencionar que diferentemente de inibidores indiretos, não há antídotos para sangramentos por problemas de superdosagem com DTIs ([KAM *et al.*, 2005](#)). Entretanto, acredita-se que devido à meia-vida plasmática curta destes compostos isso não seja um problema. Após 50 anos de pesquisa para o desenvolvimento de anticoagulantes diretos, Ximelagatran, o primeiro DTI oral, teve o uso em humanos aprovado na Europa, apesar de ainda restrito à profilaxia de curto prazo de VTE seguindo cirurgia ortopédica ([GUSTAFSSON *et al.*, 2004](#)). Ximelagatran (**28**) é o resultado de um programa de 15 anos de pesquisa na Astra-Zeneca buscando a descoberta de inibidores de trombina análogos da interação de fibrinogênio com os subsítios S1, S2 e com o sítio hidrofóbico da trombina. Ximelagatran é na verdade um pró-fármaco, sendo Melagatran (**29**) o metabólito ativo, um inibidor mais potente,

porém menos biodisponível, previamente desenvolvido na Astra-Zeneca. É importante mencionar que foi detectado que Ximelagatran produz alterações significativas nos níveis de alanina transaminase hepática em 8% dos pacientes tratados por mais de um mês (WEITZ & BATES, 2005). Apesar das alterações serem reversíveis, o significado a longo prazo destas alterações ainda é desconhecido, motivo pelo qual Ximelagatran não pôde ser aprovado para uso nos EUA. Atualmente, há outros DTIs e inibidores de FXa orais em fase II de testes clínicos, onde o problema com a enzima hepática está sendo investigado (WEITZ & BATES, 2005).



Mais recentemente, o conceito de “moléculas pequenas moduladoras da atividade da trombina” surgiu como resultado dos estudos demonstrando que trombina é uma enzima regulada por mecanismos alostéricos, podendo exibir atividades procoagulantes e anticoagulantes fisiologicamente significativas (revisado por HALL *et al.*, 1997). Segundo este conceito, um anticoagulante ideal seria um agente que inibisse especificamente a atividade procoagulante da trombina, sem risco de sangramento, exibindo ainda um amplo perfil farmacológico, boa disponibilidade oral e meia-vida relativamente longa. Alternativamente, dado o caráter alostérico da trombina, compostos sintéticos que pudessem modular suas propriedades procoagulantes e anticoagulantes poderia ser explorado para acelerar a via endógena da proteína C. Como prova do conceito, Berg *et al.* (1996) descobriram, em meio a mais de 8.000 compostos sintéticos testados, um composto que induzia um aumento de 10x na geração de aPC e uma redução de 95% na atividade coagulante da trombina, representando um aumento de 200x na especificidade de substrato. Os autores demonstraram que este efeito era independente da modulação alostérica por trombomodulina e íons Na⁺, e propuseram que seria resultante de uma mudança conformacional induzida no sítio S3.

Vale mencionar que nenhum avanço no desenvolvimento de moduladores de trombina foi relatado deste o trabalho seminal de Berg *et al.* Entretanto, o conceito foi absorvido pelo restante da comunidade científica e muitos DTIs foram testados quanto

aos seus efeitos sobre a geração de aPC e fpA pela trombina. De fato, em um estudo recente, [Linder *et al.* \(2003\)](#) mostraram que diversos DTIs, incluindo hirudina, melagatran e PPACK, inibiram as atividades procoagulantes e anticoagulantes da trombina similarmente, exceto hirugen (um inibidor de ABE-I, baseado no domínio C-terminal da hirudina). Os autores associaram esta falta de seletividade à possível causa da falha de DTIs em testes clínicos do tratamento da doença coronária arterial instável. Analisando de forma mais ampla, o paradigma de desenho de inibidores de SPCs precisa mudar para refletir a natureza alostérica destas enzimas, a fim de produzir fármacos anticoagulantes mais eficientes. Isto também ressalta a importância da contínua caracterização dos mecanismos estruturais responsáveis pelas diferentes funções desempenhadas por SPCs.

2 OBJETIVOS

O objetivo geral deste trabalho foi estudar a estrutura e função de serinil proteases ativas sobre a coagulação sanguínea empregando técnicas bioquímicas, biofísicas e computacionais. Os objetivos específicos estão listados a seguir:

- Purificar serinil proteases tripsina-símile potencialmente ativas sobre a hemóstase a partir dos venenos das serpentes *Lachesis muta rhombeata* e *Bothrops jararacussu*, usando matrizes de afinidade cromatográficas;
- Caracterizar a seletividade de ligação de SVSPs por matrizes cromatográficas baseadas em benzamidina por abordagens bioquímicas e modelagem molecular;
- Caracterizar BJ-48, a serinil protease purificada do veneno de *B. jararacussu*, quanto às suas propriedades bioquímicas;
- Estudar a especificidade de substrato e função de Lmr-47 e BJ-48 através da determinação de constantes cinéticas contra substratos e inibidores de importância biológica;
- Caracterizar a estrutura em solução de Lmr-47 e BJ-48 por métodos espectroscópicos (fluorescência e dicroísmo circular) e por espalhamento de raios-X a baixos ângulos (SAXS);
- Estudar o papel da porção carboidrato na estrutura e função de Lmr-47 e BJ-48;
- Estudar a afinidade de ligação de inibidores direcionados ao subsítio S1 de Lmm-47, outras SVTLEs e enzimas tripsina-símile através de modelagem molecular;
- Caracterizar o canal aquoso abrigando o sítio de ligação ao íon Na^+ em serinil proteases da coagulação humana e estudar o mecanismo de regulação alostérica da trombina por este íon, usando modelagem molecular.

3 MATERIAIS E MÉTODOS

Esta seção é consubstanciada pelos procedimentos descritos nos artigos e manuscritos reproduzidos nos ANEXOS 2-6. A modelagem molecular e os cálculos aplicados ao estudo da afinidade de ligação ao subsítio S1 de enzimas tripsina-símile estão descritos no ANEXO 2. A metodologia empregada no estudo da interação das SVSPs com matrizes baseadas em benzamidina é encontrada no ANEXO 3. Os métodos computacionais utilizados no estudo do canal aquoso e na alosteria mediada por Na⁺ de enzimas da coagulação estão relatados no ANEXO 4. Os procedimentos experimentais da caracterização bioquímica e estudo inicial da estrutura em solução de BJ-48 estão contidos no ANEXO 5. Por fim, os experimentos de SAXS empregadas na caracterização adicional da estrutura de BJ-48 estão descritos no ANEXO 6.

3.1 Reagentes

Os venenos de *Lachesis muta rhombeata* e *Bothrops jararacussu* foram obtidos de indivíduos mantidos no serpentário do Instituto Vital Brazil (Niterói, RJ), liofilizado e mantido a 4 °C. Os demais reagentes utilizados foram do mais alto grau analítico. O peptídeo PAR1 foi obtido de Biomolecules Midwest (St. Louis, EUA). PPACK (H-D-Phe-Pro-Arg-chlorometilcetona) foi adquirido de Calbiochem (Merck KGaA, Darmstadt, Alemanha). Fibrinogênio humano, Proteína C, trombomodulina de pulmão de coelho e anti-trombina III foram do mais alto grau de pureza, sendo conseguidos junto a Hematologic Technologies, Inc. (Essex Junction, EUA)

3.2 Procedimentos gerais

3.2.1 Eletroforese em gel de poliacrilamida-dodecil sulfato de sódio (SDS-PAGE)

Para as análises por SDS-PAGE utilizou-se uma adaptação do sistema de tampões de Laemmli (1970) e o sistema de eletroforese em gel vertical Mini-protean II (Bio-Rad Laboratories, Califórnia, EUA). Resumidamente, alíquotas de amostras de proteínas eram diluídas em tampão de aplicação de amostra (80 mM Tris-HCl pH 6,8 contendo 2 % SDS, 12 % glicerol e 0,015 % azul de bromofenol) em condições redutora (5% β-mercaptoetanol adicionado ao tampão de amostra) ou não-redutora. As corridas eram rotineiramente feitas aplicando-se uma diferença de potencial constante de 120 V.

As proteínas eram reveladas com o corante Coomassie Brilliant Blue G-250 ou com um kit comercial de revelação por impregnação com prata (Bio-Rad Laboratories).

3.2.2 Estimativa do conteúdo protéico

O conteúdo protéico das amostras foi rotineiramente estimado pelos métodos de [Lowry \(1951\)](#) ou de [Bradford \(1976\)](#), conforme implementado nos kits “DC Protein Assay” e “Protein Assay” (Bio-Rad Laboratories), respectivamente, ou apenas monitorado pela D.O.²⁸⁰, dependendo da agilidade requerida ou da compatibilidade com o tampão em que a amostra encontrava-se dissolvida.

3.2.3 Sequenciamento N-terminal por degradação de Edman automatizada

O sequenciamento da região N-terminal de proteínas foi realizado num sequenciador de fase gasosa PSQ-1 (Schimadzu, Kyoto, Japão), através do método de microssequenciamento das proteínas imobilizadas em um suporte de PVDF ([DE-SIMONE et al., 1994](#)).

3.2.4 Bioinformática

Buscas em bancos de dados foram realizadas usando programas da família BLAST ([ALTSCHUL et al., 1990](#)), implementados no NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>) ou ExPASy Proteomics Server (<http://ca.expasy.org/>). Alinhamentos múltiplos foram executados usando o programa ClustalX para Windows ([TOMPSON et al., 1997](#)). Alinhamentos de seqüências foram manipulados no programa BIOEDIT ([TOM HALL, Ibis Pharmaceuticals](#)). Ferramentas de análise proteômica usadas foram as disponíveis no ExPASy. Predição de sítios de glicosilação foram realizadas nos servidores WWW NetNGlyc 1.0 e NetOGlyc 3.5 ([GUPTA et al., 1999](#)). Manipulação e análise de estruturas de proteínas foi conduzida nos programas WebLab Viewer Lite ([Accelrys, San Diego, CA](#)) e SwissPDB Viewer ([SCHWEDE et al., 2003](#)).

3.3 Obtenção de Lmr-47 e BJ-48

3.3.1 Cromatografia de afinidade

Cerca de 500 mg dos venenos de *L. muta rhombeata* ou de *B. jararacussu* foram solubilizados em 6 mL de tampão 50 mM Tris-HCl pH 7,5 + 0,5 M NaCl (tampão de

ligação). O material foi centrifugado a 20.000g / 30min / 4 °C e o sobrenadante filtrado a 0,22 µm antes de ser aplicado a uma coluna empacotada com 3 mL de resina benzamidina-sepharose 6B (Amersham Biosciences, New Jersey, EUA) no caso do veneno de *L. m. rhombeata* e p-aminobenzamidina-agarose (Sigma-Aldrich, St. Louis, EUA) para o veneno de *B. jararacussu*, previamente equilibrada com tampão de ligação. O material não-ligado foi coletado a um fluxo de 1 mL/min e a coluna foi lavada com tampão de ligação até que a D.O.²⁸⁰ do efluente chegasse a zero. As proteínas ligadas especificamente à resina foram eluídas a 1 mL/min em frações de 1 mL com 1 mM de HCl + 0,5 M NaCl (pH 3,0). A D.O.²⁸⁰ das frações foi determinada, os tubos correspondendo ao pico foram reunidos. As amostras da cromatografia foram posteriormente analisadas por SDS-PAGE 12%.

3.3.2 Gel filtração

A cromatografia líquida de alta eficiência (HPLC) do material eluído da coluna de benzamidina-sepharose 6B (frações reunidas 12-16) foi realizada em uma coluna Shinpack Diol-150 (Shimadzu), previamente equilibrada em tampão 50 mM fosfato de sódio pH 7,2. A cromatografia foi desenvolvida sob fluxo de 1 mL/min a 25 °C e o efluente da coluna monitorado a 280 nm.

3.3.3 Ensaios de atividade

Os ensaios foram conduzidos utilizando 1,25 µM de DL-BApNA como substrato num volume final de 200 µL em placas de 96 poços. Sucintamente, 40 µL de solução do substrato a 6,25 µM era misturada a 100 µL de solução do tampão 2x (100 mM Tris-HCl pH 7,5) e 40 µL de água deionizada ou solução de inibidor. As reações eram disparadas pela adição de 20 µL de solução de enzima (1-2 µg) e a velocidade de hidrólise do substrato a 37° C acompanhada a 410 nm por 15min em um espectrofotômetro SPECTRAMax PLUS³⁸⁴ (Molecular Devices, Sunnyvale, EUA). Os ensaios eram rotineiramente realizados em triplicata e brancos eram preparados omitindo-se a adição de enzima.

3.4 **Efeito da temperatura na atividade de Lmr-47**

3.4.1 **Temperatura ótima de atividade amidásica**

A temperatura ótima de hidrólise de DL-BApNA foi determinada em ensaios como os descritos no item 3.3.3, porém variando a temperatura de 25 a 45° C. Para temperaturas superiores a 45 °C, os ensaios foram realizados em volume 10 vezes menor em microtubos incubados em banho-maria aquecidos até 80 °C. A cada 5min de incubação, por cerca de 20-25min, alíquotas de 2 µL eram retiradas e lidas em um espectrofotômetro ND-1000 (Nanodrop Technologies, Wilmington, EUA). Os ensaios foram conduzidos em triplicata e brancos eram preparados omitindo-se a adição de enzima. Os resultados foram expressos em unidades de atividade específica (1 u = mM substrato hidrolisado.min⁻¹.mg⁻¹ de proteína) usando o coeficiente de extinção molar para o grupo pNA livre $\epsilon = 8.800 \text{ M}^{-1}.\text{cm}^{-1}$.

3.4.2 **Estabilidade térmica**

Para a pesquisa da estabilidade térmica das enzimas, estas eram submetidas a 1h de aquecimento a 65 °C para então terem a atividade amidásica determinada como descrito no item 2.3.

3.5 **Caracterização da estrutura em solução de LMR-47**

3.5.1 **Espectropolarimetria de dicroísmo circular (CD)**

Para este experimento, utilizou-se espectropolarímetro J-720 (JASCO Inc., Easton, EUA). A partir de amostra concentrada (4 mg/mL) de LMR-47 em 1 mM Tris-HCl obteve soluções 1 mg/mL em 50 mM de tampão fosfato nos valores de pH 6,0 e 7,5. As medidas foram realizadas a 25 °C em cubetas de 0,2 cm de caminho ótico e os dados de dicroísmo coletados na faixa espectral de 260 – 195 nm com resolução de 0,5 – 1,0 nm. Os espectros obtidos são resultantes de 3 acumulações e subtração do espectro do respectivo tampão. O sinal de CD em mdeg foi convertido em elipsidade ($\theta_{mrw,\lambda}$) ou absortividade molar residual média ($\Delta\epsilon$) para usar programas de predição de estrutura secundária. A conversão em $\theta_{mrw,\lambda}$ (deg.cm².dmol⁻¹) foi feita a partir da fórmula $\theta_{mrw,\lambda} = \frac{mrw * CD}{10 * L * c}$, onde *mrw* é o “peso médio do resíduo” = MM/(NR-1), com MM = massa molecular (g/mol) e NR = número de resíduos; CD é o sinal de CD (mdeg), *L* é o caminho ótico (cm) e *c* a concentração (mg/mL). Para obter os dados em $\Delta\epsilon$ (M⁻¹.cm⁻¹)

usou-se a relação $\theta_{mrw,\lambda} = 3298 * \Delta\epsilon$. A porcentagem de estrutura secundária do tipo α -hélice, folha- β , grampos- β e estrutura randômica foi calculada com o auxílio do programa CDNN Deconvolution (Versão 2– www.bioinformatik.biochemtech.uni-halle.dee/cdnn).

3.5.2 Espectroscopia de Emissão da Fluorescência Intrínseca

As medidas foram realizadas num espectrofluorímetro FP-6300 (Jasco Inc.) em cubeta de quartzo de 1 x 1 cm e a 25 °C. As amostras de LMR-47 (1 μ M) foram preparadas em diferentes valores de pH, em soluções a 50 mM dos tampões acetato de sódio (pH 5,0 e 5,5), fosfato de sódio (pH 6,0 e 7,0) e Tris-HCl (pH 7,5 e 8,0). Os espectros de emissão de fluorescência foram adquiridos no intervalo espectral de 300 a 400 nm (em intervalos de 0,5 nm), após excitação da proteína a 280 nm. O espectro de emissão de fluorescência do tampão (3 leituras cumulativas) foi subtraído do espectro de emissão fluorescência da proteína (3 leituras cumulativas). Os dados foram analisados no programa Origin® v.7.0 (OriginLab Corp., Northampton, EUA) e três parâmetros de fluorescência foram considerados: intensidade de fluorescência total (IF^t), pela aplicação da integral da área espectral na faixa de 320 a 380 nm; centro de massa espectral (CME), $CME = (\sum \lambda * IF_{\lambda}) / \sum IF_{\lambda}$, onde IF_{λ} é a intensidade de fluorescência no comprimento de onda λ ; os comprimentos de onda correspondendo aos dois máximos da intensidade de fluorescência (λ_{max}^I e λ_{max}^{II}), calculados pelo ajuste do espectro a duas curvas Gaussianas.

3.5.3 Espalhamento de raios-X a baixos ângulos (SAXS)

3.5.3.a Preparo da amostra

Uma amostra de LMR-47, purificada como descrito nos itens 2.3 e 2.4 acima, foi concentrada até 4 mg/mL em tampão 50 mM fosfato de sódio pH 6,0 usando filtro Centricon-10 (Millipore, Billerica, EUA) e mantida a 4 °C. Imediatamente antes à realização das medidas, as amostras eram centrifugadas a 20.000g / 10min / 4 °C e o sobrenadante utilizado para a coleta dos dados.

3.5.3.b Configuração experimental, coleta e tratamento preliminar de dados

Os experimentos foram realizados na linha de SAXS do Laboratório Nacional de Luz Síncrotron ([KELLERMANN et al., 1997](#)). Duas distâncias amostra-detector foram utilizadas: 849,5 e 430,1 nm. A amostra foi aplicada em um porta-amostras de mica. O

comprimento de onda das análises foi 1,488 Å. A proteína foi exposta à radiação por 4 ciclos de 10min, enquanto que o tampão teve os dados de espalhamento coletados por apenas 1 ciclo. Para determinação da massa molecular foi utilizado uma amostra a 5 mg/mL de ovalbumina (47 kDa) como padrão secundário. O tratamento preliminar dos dados experimentais foi realizado com auxílio do programa TRAT1D 1.3 (OLIVEIRA, 2003). Este tratamento consiste em corrigir a curva de espalhamento quanto à absorção, à intensidade do feixe incidente, ao tempo de aquisição e ao espalhamento parasita. Este último representa o espalhamento adicional causado pelos elementos ópticos inseridos no caminho do feixe após a amostra (fendas, janelas, etc.). Além dos efeitos já descritos, o programa também leva em conta que detectores à gás não têm a mesma resposta em todas as posições e como tal deve-se efetuar uma correção adicional que consiste em dividir a intensidade por uma curva característica da resposta do detector (curva de homogeneidade). Esta curva é determinada a partir da emissão homogênea de uma fonte de ^{55}Fe e fornecida pelo pessoal da linha de SAXS junto com o conjunto de dados experimentais.

O programa TRAT1D oferece como saída os dados corrigidos em função do vetor de espalhamento q e também a propagação de erro para a curva experimental. O módulo do vetor de espalhamento é definido por: $q = \frac{4\pi \cdot \sin\theta}{\lambda}$, onde θ é a metade do ângulo de espalhamento e λ é o comprimento de onda da radiação utilizada. Para determinação do ângulo de espalhamento 2θ usa-se a seguinte relação: $2\theta = \arctang [(C - C_z) \cdot K / D]$, onde C é um canal qualquer do detector e C_z (canal zero) corresponde à posição de incidência do feixe direto no detector, D é a distância amostra-detetor e K é a distância entre dois canais consecutivos (relação mm/canal).

3.5.3.c Análise dos dados

Os dados experimentais tratados pelo programa TRAT1D para os espalhamentos à longa e curta distância foram reunidos e em seguida manualmente truncados até o valor de q onde a intensidade de espalhamento confundia-se com o erro experimental. Estes dados de espalhamento “editados” foram fornecidos ao programa Gnom 4.5 (SVERGUN *et al.*, 1988; SVERGUN, 1992). Este programa busca uma função de distribuição de pares $[p(r)]$ no chamado “espaço real” que se relaciona às intensidades experimentais pela transformada indireta de Fourier (SVERGUN, 1991). Os dados foram processados através de ciclos iterativos de ajuste da função $p(r)$, onde valores de

D_{\max} (diâmetro máximo da partícula espalhadora) eram fornecidos a cada ciclo. A correção de distorções experimentais (*desmearing*) para o detector uni-dimensional empregado na linha de SAXS do LNLS, foi realizada em termos da divergência no feixe, através do fornecimento do parâmetro correspondente à projeção da altura do feixe no plano do detector (AH), calculado através da equação: $AH = \frac{2\pi f}{\lambda D}$, onde f é a fenda do detector (8 mm), λ é o comprimento de onda da radiação (1,488 Å) e D é a distância amostra-detector. AH precisa ser fornecido com a mesma unidade do vetor de espalhamento (neste caso, Å⁻¹). Além de maximizar o índice de qualidade do ajuste fornecido pelo programa Gnom, buscou-se obter uma $p(r)$ ajustada onde se tinha um decaimento a zero suave e assintótico ao eixo das abscissas.

A partir dos dados ajustados também se gerou uma curva $I.q^2$ x q , onde I é a intensidade e q o vetor de espalhamento. Esta curva é conhecida como o gráfico de Kratky e serve à avaliação da flexibilidade da partícula espalhadora (KRATKY, 1963). O grau de dispersão da proteína foi avaliado comparando-se a massa molecular de LMR-47 conhecida por espectrometria de massas com a estimativa feita a partir dos dados de SAXS. Esta última é obtida pela relação: $M_w a = \frac{I_a / C_a}{I_p / C_p} M_w p$, onde os índices ‘a’ e ‘p’ significam ‘amostra’ e ‘padrão’, respectivamente; M_w é a massa molecular, I é $I_{(0)}$ ou a intensidade extrapolada para $q = 0$ a partir da curva de espalhamento ajustada pelo programa Gnom e C é a concentração da proteína.

3.5.3.d Modelagem *ab initio*

Os dados ajustados pelo programa Gnom foram fornecidos ao programa GASBOR (SVERGUN *et al.*, 2001) para o cálculo *ab initio* de modelos de baixa resolução da estrutura de LMR-47. O programa Gasbor representa a estrutura da proteína através de um conjunto de “resíduos anônimos” (dummy residues, DR) compatível com um polímero de aminoácidos (i.e., cada DR tem pelo menos um vizinho a uma distância de 3,8 Å). A posição espacial destes resíduos tenta aproximar a posição dos carbonos alfa na estrutura da proteína, mas a seqüência com que estes aparecem ordenados no modelo não tem relação com a seqüência primária da proteína. Assim, para o cálculo dos modelos *ab initio*, considerou-se como 228 o número de resíduos de aminoácidos de LMR-47 (obtido pelo número de aminoácidos na seqüência

madura de Lmm-TLE; MAGALHÃES *et al.*, 1993) e os demais parâmetros do programa foram *default*. Um total de 10 modelos independentes foi obtido e um modelo médio foi calculado com o programa DamAver do pacote ATSAS 2.1 (VOLKOVA & SVERGUN, 2003).

3.6 Caracterização bioquímica adicional de LMR-47 e BJ-48

Estes experimentos foram realizados em colaboração com o grupo do Dr. Enrico Di Cera (Washington University School of Medicine, St. Louis, EUA).

3.6.1 Especificidade do subsítio S1 e concentração do sítio-ativo

Os valores de K_m e k_{cat} para a hidrólise de H-D-Phe-Pro-Xaa-p-nitroanilida (Xaa = Arg, Lys, Phe ou Tyr; Midwest Biotech, Indianápolis, EUA) foram determinados usando curvas de progresso monitoradas a 405 nm analisadas por integração direta e correção para a inibição do produto (KREM & DI CERA, 2003). Nestes substratos tripeptídicos, o resíduo P3 é um D-aminoácido e atua como um resíduo P4. Os ensaios foram realizados a 25 °C no tampão 5 mM Tris-HCl pH 8,0 + 200 mM NaCl + 1 % PEG. No caso de FPR-pNa, também foi realizado um controle com cloreto de colina para manter a força iônica e verificar o efeito do íon Na^+ . A concentração de enzima foi tipicamente 1-2 nM e a concentração de substrato variou de 5-100 μ M, dependendo do resíduo em P1. As enzimas Lmr-47 e BJ-48 tiveram a concentração do sítio-ativo determinada por titulação com o inibidor PPACK, uma vez que foi previamente relatado que Lmm-TLE é fracamente inibida por NPGB e SBTI, titulantes de sítio-ativo de tripsina (MAGALHÃES *et al.*, 2003). Em resumo, quantidades variáveis de PPACK foram incubadas com uma quantidade fixa das proteases por aproximadamente 1 hora. A velocidade inicial (V_o) da reação de hidrólise de H-D-FPR-pNA foi então grafada contra a quantidade de inibidor ([I]) para fornecer uma linha reta cuja inclinação e intercepto fornecem a concentração do sítio-ativo (KNIGHT, 1995).

3.6.2 Interação com moléculas ativas sobre o sistema hemostático

Os ensaios foram realizados a 37° C no tampão Tris-HCl 5 mM pH 8,0 + 1% PEG + 145 mM NaCl. O ensaio com fibrinogênio também foi realizado com cloreto de colina para manter a força iônica e verificar o efeito do íon Na^+ .

3.6.2.a Hirudina

A constante de dissociação em equilíbrio (K_i) para hirudina foi determinada pela análise das curvas de progresso da hidrólise de H-d-FPR-pNA a 405 nm, como função da concentração de inibidor, disparado pela adição de trombina (STONE & HOFSTEENGE, 1986; AYALA & DI CERA, 1994). A ligação de hirudina às enzimas foi medida nas mesmas condições que nos ensaios com H-d-FPR-pNA sozinho (vide item 3.6.1 acima). Seis curvas de progresso foram medidas em concentrações de hirudina variando de na faixa de 0-1000 nM. A concentração de enzima foi tipicamente 2 nM e a concentração de substrato foi 50 μ M.

3.6.2.b Proteína C

A ativação de PC foi determinada usando curvas de progresso da hidrólise de H-D-Asp-Arg-Arg-p-nitroanilida (MidWest Biotech) para o valor de $s = k_{cat}/K_m$ (XU *et al.*, 2004). Este substrato cromogênico é altamente seletivo para proteína C ativada (DANG & DI CERA, 1997), e sua hidrólise foi empregada para monitorar quantitativamente a conversão de PC na enzima ativa mediada pelas SVTLEs. Experimentos controle para a hidrólise de H-D-Asp-Arg-Arg-p-nitroanilida pelas SVTLEs foram conduzidos e confirmaram a ausência de qualquer atividade significativa em todas as condições testadas. As condições experimentais foram 5 mM Tris-HCl pH 7,4; 0,1% PEG; 5 mM CaCl_2 ; a 37 °C. As SVTLEs foram usadas nas concentrações de 0,2 a 10 nM. A concentração de PC foi fixada em 300 nM, bem abaixo do valor de K_m estimado (GALVIN *et al.*, 1987), para assegurar determinações acuradas de k_{cat}/K_m pela análise das curvas de progresso (DANG *et al.*, 1995; DANG *et al.*, 1997). Os experimentos de ativação de PC foram realizados na presença de 10 nM trombomodulina.

3.6.2.c PAR1

A interação das SVSPs com PAR-1 foi estudada usando a porção extracelular do receptor abrangendo os resíduos 33-62 (TR³³⁻⁶²). O fragmento tem a sequência A³³TNATLDPRSFLLRNPNDKYEPFWEDEEKN⁶² e, como referência, é clivado pela trombina humana entre R41 e S42. As constantes cinéticas para a hidrólise do peptídeo de TR³³⁻⁶² por Lmr-47 e BJ-48 foram determinadas a partir de curvas de progresso, conforme descrito anteriormente (AROSIO *et al.*, 2000; AYALA *et al.*, 2001). A concentração de TR³³⁻⁶² usada no ensaio foi 3 μ M e de enzima foi 5 nM.

3.6.2.d FpA

A cinética de liberação de FpA a partir do fibrinogênio humano foi determinada conforme descrito por [Vindigni & Di Cera \(1996\)](#). A concentração de enzima ativa variou de 1 a 2 nM. Fibrinogênio foi usado em concentrações de 0,22 μM , bem abaixo do valor estimado de K_m . A reação, disparada pela adição das SVTLEs à solução de fibrinogênio, foi paralisada em diferentes intervalos de tempo com 3M de ácido perclórico. Os peptídeos foram separados por cromatografia de fase reversa, quantificados e os valores de k_{cat}/K_m calculados a partir das curvas de progresso conforme descrito ([VINDIGNI & DI CERA, 1996](#)).

3.6.2.e ATIII

A constante cinética de segunda ordem aparente (k_{on}) para inibição das SVTLEs por ATIII foi determinada a partir da análise de curvas de progresso da hidrólise de H-D-FPR-pNA na presença de concentrações variáveis do inibidor (0-100 nM), 50 μM do substrato FPR-pNA e 2 nM de enzima, conforme [Olson *et al.* \(1993\)](#).

3.7 Caracterização dos carboidratos de LMR-47 e BJ-48

3.7.1 Teste de ligação em Con A

Concanavalina A (Con A) é uma lectina capaz de ligar moléculas contendo α -D-manopiranosil, α -D-glucopiranosil e outros grupamentos com estereoquímica similar ([BREWER & BHATTACHARYYA, 1988](#)). Para o teste, 250 μL de resina Con A-sepharose 4B (Amersham Biosciences) foram empacotados em mini-coluna de 1 mL (Bio-Rad) por centrifugação a 500g / 30s / 25 °C. A coluna foi equilibrada por centrifugação com 5 adições de 1 mL de tampão 50 mM NaOAc pH 6,0 + 0,5 M NaCl (tampão 1). Cerca de 700 μg de cada uma das enzimas puras foram aplicados a colunas independentes em 250 μL de tampão 1. Cada coluna foi acoplada a um tubo de 1,5 ml e o material não-ligado foi coletado por centrifugação. Em seguida cada coluna foi lavada com 3 aplicações de 250 μL de tampão 1. Adicionou-se 250 μL de tampão 1 + 500 mM de α -metilmanosídeo ou maltose (α -manosilglucopiranosídeo) e após 5min de incubação eluiu-se com auxílio de centrifugação. Outras 3 frações de 250 μL foram eluídas desta forma. As amostras foram analisadas por SDS-PAGE 12%.

3.7.2 Determinação enzimática do conteúdo e tipo de carboidratos em Lmr-47

O conteúdo total de carboidratos foi estimado pelo aumento na mobilidade eletroforética da glicoproteína submetida à deglicosilação enzimática em condições desnaturantes com o kit E-DEGLY (Sigma, St. Louis, MI). Seguiu-se as instruções do fabricante do kit, exceto pelo tempo de incubação que foi aumentado de 3 para 18h. Os tipos de sacarídeos presentes na glicoproteína foram pesquisados pela adição seqüencial de glicosidases específicas. Carboidratos N-ligados foram retirados pela peptidil N-glicosidase F (PNGase F) fornecida com o kit. A presença de carboidratos O-ligados foi testada pela ação combinada da O-glicosidase com pelo menos uma das seguintes enzimas: α -2(3,6,8,9)-neuraminidase, β (1-4)-galactosidase ou β -N-acetilglicosaminidase.

3.7.3 Deglicosilação enzimática em condições nativas

3.7.3.a Estudos preliminares

Para o ensaio piloto de deglicosilação de BJ-48 e LMR-47 com PNGase F em condições nativas, avaliou-se a eficiência da reação em pH 6,0 e em pH 8,5. Dois controles foram realizados. Um consistiu em deglicosilar a enzima em condições desnaturantes (controle positivo) e o outro em substituir o volume de PNGase F adicionado à reação em pH 8,5 por volume equivalente de água (controle negativo). Para as reações em condições nativas, 20 μ g de cada glicoproteína foram misturados a 4 unidades de PNGase F (P0704S, New England Biolabs, Ipswich, EUA) em 30 μ L de tampão fosfato de sódio 100 mM pH 6,0 ou Tris-HCl 100 mM pH 8,5. Para a reação em condições desnaturantes, a mesma quantidade das glicoproteínas foi diluída em 20 μ L de “tampão desnaturante” (0,5% SDS + 1% 2-ME). Ferveu-se a 100 °C por 10min. Após resfriamento dos tubos até a 25 °C, adicionou-se 3 μ L de 0,5M tampão fosfato de sódio pH 7,5, 3 μ L de 10% NP-40, 4 unidades de PNGase F e água deionizada q.s.p. 30 μ L. As reações, incluindo o controle negativo, foram incubados a 37 °C. Alíquotas foram analisadas por SDS-PAGE 12% com 18 e 48h de incubação.

3.7.3.b Deglicosilação “semi-preparativa”

Aproximadamente 1 mg de cada glicoproteína foi incubado a 37 °C em 500 μ L de tampão 50 mM fosfato de sódio pH 7,0 com 9 unidades de PNGase F (P2619,

Sigma), na presença de 1 mM *p*-NH₂-benzamidina. Após 72h de incubação, alíquotas das reações foram analisadas por SDS-PAGE 12%. Adicionou-se então 9 unidades de α -2(3,6,8,9)-neuraminidase (E-DEGLY kit, Sigma), suplementando a reação com *p*-NH₂-benzamidina suficiente para uma concentração de 1 mM. As reações foram então incubadas por mais 72h a 37 °C, quando alíquotas foram analisadas por SDS-PAGE 12%. As reações foram então novamente suplementadas com 250 unidades de PNGase F, 1 mM *p*-NH₂-benzamidina e incubadas por mais 48h a 37 °C. Após análise de alíquotas das reações por SDS-PAGE 12%, decidiu-se adicionar mais 200 unidades de PNGase F e 1 mM *p*-NH₂-benzamidina. A reação foi incubada por mais 48h nas mesmas condições. Os tubos foram então centrifugados (20.000g, 10min, 4 °C) e pellet e sobrenadantes analisados por SDS-PAGE 12%.

3.7.3.c Deglicosilação preparativa

Para BJ-48, 2,4 mg da glicoproteína foram incubados com 480 unidades de PNGase F em 1,8 mL de tampão 100 mM Tris-HCl pH 8,5 contendo 1 mM de *p*-NH₂-benzamidina. Para LMR-47, 5 mg da glicoproteína foram incubados com 1000 unidades de PNGase F em 1,8 mL de 100 mM tampão fosfato de sódio pH 6,0 contendo 1 mM de *p*-NH₂-benzamidina. Após 48h de incubação a 37 °C, alíquotas foram analisadas por SDS-PAGE 12% e suplementou-se as reações com 1000 unidades de PNGase F e 1 mM de *p*-NH₂-benzamidina. Após 72h de incubação nas mesmas condições e nova análise por SDS-PAGE 12%, adicionou-se CHAPS a 1% e mais 1000 unidades de PNGase F às reações. Incubou-se por mais 48h a 37 °C e novas alíquotas foram retiradas para análise eletroforética. Neste ponto, 7 e 14 unidades de α -2(3,6,8,9)-neuraminidase foram adicionados às reações de BJ-48 e LMR-47, respectivamente, além de 1 mM *p*-NH₂-benzamidina. Após 36h de incubação a 37 °C, amostras das reações foram analisadas por SDS-PAGE 12%. Finalmente, mais 1000 unidades de PNGase F foram adicionados e as reações incubadas por 60h a 37 °C.

3.7.4 Análise química dos carboidratos liberados

As amostras de proteínas deglicosiladas em condições nativas conforme o item 3.7.3.c tiveram o volume elevado para 8 mL com tampão 50 mM NaOAc pH 6,0. A seguir as amostras foram concentradas com filtros centricon-10 até redução do volume para 500–1000 μ L. O material filtrado contendo os carboidratos liberados na deglicosilação foram analisados conforme os itens 3.7.3 a e b. O material concentrado

retido nos filtros, contendo as proteínas deglicosiladas, foi fracionado conforme o item 3.7.4.

3.7.4.a Carboidratos sialilados

O conteúdo de ácido siálico foi determinado de acordo com (Dionex technical note 41; Dionex Inc., Sunnyvale, EUA). Sucintamente, alíquotas das amostras de carboidrato foram dissolvidas em ácido acético 1M e aquecidos a 80 °C por 2h. O ácido D-glicurônico foi adicionado como um padrão interno e as amostras analisadas em um sistema cromatográfico Dionex Biol C, equipado com uma coluna PA-10. As corridas cromatográficas foram realizadas a um fluxo de 1 mL/min, aplicando-se um gradiente de 7-30% de B (1M NaOAc em 100 mM NaOH) em A (100 mM NaOH).

3.7.4.b Carboidratos neutros

Alíquotas dos carboidratos foram secas em Speedvac (Thermo Savant, Somerset, EUA), dissolvidas em 200 µL of ácido acético 2M e aquecidas a 100 °C por 4h. As amostras hidrolisadas foram novamente secas e então dissolvidas em água para serem injetadas em um sistema Dionex Biol C, como descrito acima. A cromatografia foi desenvolvida usando 200 mM NaOH e água deionizada como fases móveis A e B, respectivamente. A eluição foi efetuada aplicando-se 90% A por 20min. Fetuína foi usada como controle da quantificação de açúcares e 2-deoxi-D-glicose foi empregada como padrão interno.

3.7.5 Fracionamento do produto de deglicosilação nativa

3.7.5.a Teste com cromatografia de afinidade

As amostras de proteína da deglicosilação nativa foram submetidas à cromatografia de afinidade em coluna de Con A-sepharose 4B, semelhantemente ao descrito no item 3.7.1. Desta vez, o volume de resina usado foi de 1 mL e o eluente adotado foi 500 mM de maltose em 50 mM NaOAc pH 6,0. Frações de 500 µL foram coletadas do não-ligado (2 frações), lavado (5 frações) e eluído (10 frações). As amostras foram analisadas por SDS-PAGE.

3.7.5.b Teste com cromatografia de exclusão molecular (SEC)

Alíquotas das amostras de deglicosilação nativa (5 – 20 µL) foram aplicadas em um sistema HPLC equipado com coluna (24 mL de leito) Superdex 75 HR 10/30

(Amersham Biosciences). A cromatografia foi desenvolvida com fluxo de 0,5 mL/min em tampão 50 mM Tris-HCl pH 7,5 + 200 mM NaCl. A D.O. do efluente da coluna foi monitorada a 220 e 280 nm.

3.7.5.c Fracionamento por cromatografia de interação hidrofóbica (HIC)

Amostras da reação de deglicosilação nativa das glicoproteínas tiveram o tampão trocado para 50 mM Tris-HCl pH 7,5 + 2M Na₂SO₄ com auxílio de filtros centricon-10. As amostras (500 – 1000 µL) foram então aplicadas em um sistema HPLC equipado com coluna Phenyl 5PW (TosoHas, Montgomeryville, EUA) de 5 mL de leito. A cromatografia foi desenvolvida a 25 °C sob fluxo de 0,5 mL/min, usando 50 mM Tris-HCl pH 7,5 + 2M Na₂SO₄ como tampão A e 50 mM Tris-HCl pH 7,5 como tampão B. Após aplicação da amostra, a coluna foi lavada por 20min com 100% A. Em seguida, a concentração de tampão A em B foi reduzida a 20% em 5min e assim mantida por mais 5min. A seguir, um gradiente de 20-0% de A em B foi realizado em 40min. A partir daí, aplicou-se 0% de A em B (100% B) por mais 20min, quando se iniciou o reequilíbrio do sistema. Durante a cromatografia, a D.O. do efluente da coluna foi monitorada a 220 e 280 nm. Os picos observados no cromatograma foram coletados manualmente.

3.7.6 Caracterização das frações deglicosiladas de BJ-48

3.7.6.a Estabilidade térmica

Amostras das frações obtidas por HIC e da enzima nativa foram submetidas a 1h de aquecimento a 55 °C ou a 65 °C e então tiveram a atividade específica sobre BApNA determinada a 37° C, conforme o item 3.3.3. Controles foram realizados onde a etapa de aquecimento foi omitida (100% de atividade).

3.7.6.b Inibição por SBTI e Infestina

As frações deglicosiladas foram testadas quanto a inibição da atividade amidásica a 37 °C sobre BApNA por SBTI (50 µg/mL) e Infestina (0,5 nM), conforme o item 3.3.3. Os dados de inibição foram expressos em termos da atividade residual comparado a um controle onde a adição do inibidor foi omitida (100% de atividade).

4 Resultados

Esta seção é consubstanciada pelos resultados relatados nos artigos e manuscritos submetidos para publicação constantes dos ANEXOS 2-6. O modelo por homologia de Lmm-TLE, a qual foi assumida no presente trabalho como idêntica a Lmr-47, a enzima correspondente no veneno de *L. m. rhombeata*, é relatado no ANEXO 2. A análise dos inibidores amidínicos e guanidínicos ancorados ao subsítio S1 de Lmm-TLE e o modelo empírico gerado no estudo da afinidade de ligação ao subsítio S1 de enzimas tripsina-símile também são descritos no ANEXO 2.

O rendimento da purificação de BJ-48 e Lmr-47, as propriedades bioquímicas da interação das SVSPs com as matrizes baseadas em benzamidina e o estudo de modelagem molecular visando a interpretação estrutural das seletividades cromatográficas de Lmr-47 são relatados no ANEXO 3. É relevante destacar que a serinil protease purificada do veneno de *L. m. rhombeata* por cromatografia de afinidade mostrou-se indistinguível de Lmr-47 por critérios bioquímicos, incluindo a seqüência dos 15 primeiros resíduos do N-terminal (vide ANEXO 3).

A caracterização da forma do canal aquoso e o mecanismo proposto para a alosteria mediada por Na^+ em serinil proteases da coagulação humana estão descritos no ANEXO 4. Os resultados da caracterização bioquímica e estudo inicial da estrutura em solução de BJ-48 estão contidos no ANEXO 5. Finalmente, os dados de SAXS obtidos para BJ-48 visando a caracterização da estabilidade térmica, efeito do pH e da ligação de benzamidina na estrutura em solução da enzima estão descritos no ANEXO 6.

4.1 **Determinação da temperatura ótima de atividade amidásica de Lmr-47**

A velocidade de hidrólise de BApNA por Lmr-47 foi medida em temperaturas variando de 25 a 65 °C. A **Fig. 102:** mostra o perfil de variação da atividade amidásica de Lmr-47 com a temperatura. A enzima mostrou temperatura ótima de atividade em 44 °C, exibindo cerca de 70% da atividade máxima na temperatura fisiológica de 37 °C e apenas 20% desta a temperatura ambiente (25 °C). Lmr-47 mostrou elevada estabilidade térmica, mantendo 60% da atividade máxima em temperaturas na faixa de 50-60 °C. A atividade amidásica da enzima somente foi inteiramente abolida em temperaturas superiores a 80 °C (dados não mostrados).

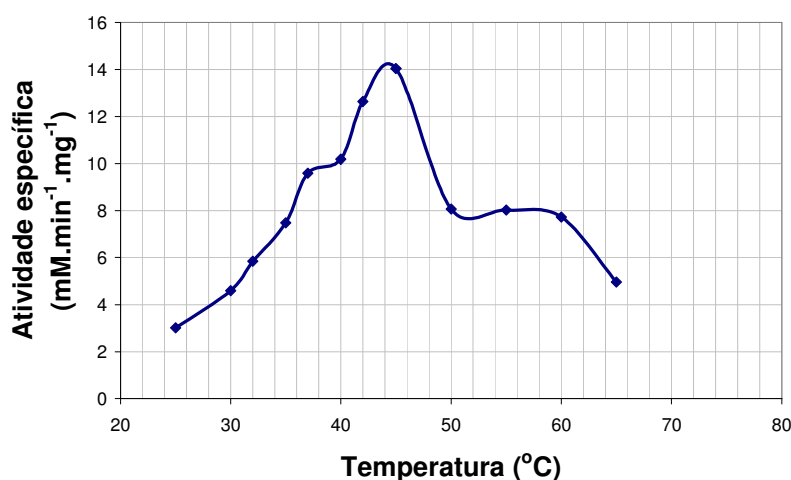


Fig. 102: Perfil de variação da atividade amidásica de Lmr-47 com a temperatura. BApNA foi utilizado como substrato. Cada ponto representa a média de três replicatas.

4.2 Especificidade do subsítio S1 de BJ-48 e Lmr-47

Com o intuito de melhor caracterizar a preferência por substratos no subsítio S1 das enzimas BJ-48 e Lmr-47 um painel de derivados cromogênicos de tripeptídeos variando na posição P1 foi testado e os valores de K_s obtidos estão listados na **Tabela 10**. Somente os valores de K_s são relatados, pois os valores de K_m foram muito superiores (mM) à faixa de concentração de substrato, [S], empregada nos ensaios (μM). Nestas condições, a velocidade de formação do produto segue uma exponencial simples, cuja cinética depende apenas da concentração de enzima e de K_s . De fato, para a maior parte das enzimas sob condições fisiológicas, o valor de K_s é o único que importa, pois normalmente o substrato está presente numa concentração tal que $[S] \ll K_m$.

Tabela 10: Constantes de especificidade (K_s) das enzimas Lmr-47 e BJ-48 contra tripeptídeos variando em P1.

Tripeptídeos	K_s ($\mu\text{M}^{-1}.\text{s}^{-1}$)	
	Lmr-47	BJ-48
D-FPR-pNA	8,5 (9,8)	2,9 (3,1)
D-FPK-pNA	$1,2 \times 10^{-5}$	$1,2 \times 10^{-4}$
D-FPF-pNA	$2,3 \times 10^{-3}$	n.c.
D-FPY-pNA	n.c.	n.c.

n.c. = não cliva. Os valores entre parênteses foram medidos na ausência do íon Na^+ no tampão.

Conforme esperado para enzimas tripsina-símile, ambas as SVSPs foram incapazes de clivar substratos contendo um resíduo de Tyr em P1 e apresentaram valores de K_s na ordem de $10^6 \text{ M}^{-1}.\text{s}^{-1}$ para o tripeptídeo com Arg em P1. O restante dos

dados constantes da **Tabela 10** são bastante inusitados para enzimas tripsina-símile. As enzimas compartilham uma preferência marcante (10^4 – 10^5 vezes) por Arg contra Lys na posição P1 dos substratos, com Lmr-47 mostrando-se 10 vezes mais exigente que BJ-48. Por outro lado, Lmr-47 é capaz de clivar FPF, enquanto que BJ-48 não. Curiosamente, Lmr-47 é 100 vezes mais específica para FPF, um substrato típico de quimiotripsina, em relação a FPK. Não obstante, a presença grupo hidroxila na posição 4 do anel benzênico da fenilalanina, torna o substrato FPY resistente à clivagem por Lmr-47. Estes resultados caracterizam um perfil de especificidade do subsítio S1 único em Lmr-47. Mais ainda, os dados sugerem que o caráter lipofílico e o volume do resíduo em P1 sejam importantes requisitos para alcançar alta afinidade pelo subsítio S1 de Lmr-47. Em BJ-48, um perfil mais semelhante com outras enzimas tripsina-símile é revelado, uma vez que a carga positiva no resíduo em P1 é essencial para a ligação e clivagem do substrato por esta enzima.

O efeito do íon Na^+ também foi testado na hidrólise do substrato FPR (**Tabela 10**). Conforme esperado para SVSPs, as enzimas mostraram-se não-responsivas à presença deste íon nos tampões de atividade. Lmr-47, assim como as demais SVSPs seqüenciadas até o momento, possuem um resíduo de Pro na posição 225, de onde se infere que estas são incapazes de manter um sítio de ligação ao Na^+ como o encontrado na trombina de animais vertebrados. BJ-48 também não teve a atividade incrementada pela presença do íon Na^+ , indicando que esta enzima também possui Pro²²⁵.

4.3 Interação com moléculas ativas sobre o sistema hemostático

Os papéis biológicos das enzimas Lmr-47 e BJ-48 foram estudados através da determinação de constantes cinéticas contra diferentes moléculas ativas sobre o sistema hemostático (**Tabela 11**). Conforme esperado para estas SVTLEs, a cadeia A α do fibrinogênio foi eficientemente hidrolisada para liberar FpA. Pela primeira vez, a atividade destas enzimas foi testada contra Proteína C, um substrato anticoagulante, e PAR1, um receptor ativado por proteases, mediando importantes ações das plaquetas na coagulação e outros processos de importância fisiológica, como a inflamação. Diferentemente da trombina humana, o peptídeo derivado de PAR1 não foi clivado pelas enzimas de veneno. Isso indica uma importante diferença no mecanismo de ação

destas enzimas em relação à trombina. Lmr-47 também foi capaz de clivar PC na presença do co-fator trombomodulina, um substrato anticoagulante, porém com especificidade 10^5 vezes menor do que o fibrinogênio humano. Pesquisou-se ainda o efeito do íon Na^+ na interação com FpA. Diferentemente dos substratos tripeptídicos direcionados ao sítio catalítico e interagindo com os subsítios S1-S3, a especificidade da reação de liberação de FpA por ambas as SVTLEs aumentou em pouco mais de 2 vezes na presença do cátion monovalente (145 mM). Tendo em vista que a força iônica e o contra-íon (Cl^-) foram mantidos nos ensaios, este efeito pode ser efetivamente atribuído ao íon Na^+ .

Tabela 11: Constantes cinéticas das enzimas Lmr-47 e BJ-48 contra moléculas ativas sobre o sistema hemostático.

Ligantes	Lmr-47	BJ-48
Substratos	K_s ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	
FpA	1,8 (0,8)	2,1 (1,0)
Proteína C ^a	$3,4 \times 10^{-5}$	n.d.
PAR1	n.c.	n.c.
Inibidores	K_i ou K_{on}	
Hirudina	n.l.	n.l.
ATIII ^b	n.l.	n.l.

n.c. = não cliva. n.d. = não determinado. n.l. = não liga. Os valores entre parênteses foram medidos na ausência do íon Na^+ no tampão. ^a Na presença de 5 mM Ca^{2+} e 10 nM trombomodulina. ^b Na ausência de heparina.

As interações das SVSPs com duas outras moléculas ativas sobre fatores da coagulação sanguínea, hirudina e anti-trombina (na ausência de heparina) foram pesquisadas (**Tabela 11**). Lmr-47 e BJ-48 foram incapazes de ligar estes inibidores, inviabilizando a determinação das constantes cinéticas. Estes resultados apontam diferenças nas estruturas destas SVSPs em relação à trombina e outros fatores da coagulação, nas regiões equivalentes aos exosítios de ligação a ânions, ABE-I e ABE-II, implicados na interação com hirudina e ATIII, respectivamente, com as enzimas de mamíferos.

4.4 Caracterização da fração carboidratos em Lmr-47 e BJ-48

4.4.1 Interação com Con A

As enzimas Lmr-47 e BJ-48 foram submetidas a um teste de ligação com a lectina Con A, através de um ensaio cromatográfico. A análise por SDS-PAGE das

frações cromatográficas mostrou que ambas as SVSPs são capazes de se ligar à Con A immobilizada ao suporte inerte, conforme ilustrado para BJ-48 na **Fig. 103**. A eluição foi realizada com sucesso através de soluções 500 mM de maltose ou α -metilmanosídeo, demonstrando a natureza específica da interação com a resina cromatográfica. Este resultado mostra que as enzimas são glicoproteínas contendo grupos piranosídicos unidos por ligações α -glicosídicas na porção de carboidratos de suas estruturas. Mais ainda, a interação com Con A fornece mais uma opção de purificação destas enzimas por cromatografia de afinidade, a qual pode ser muito útil na separação de formas nativas e deglicosiladas de Lmr-47 e BJ-48.

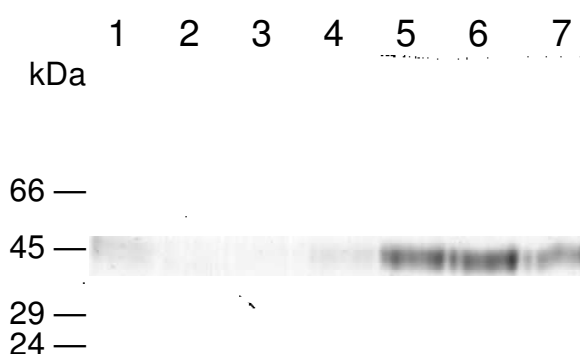


Fig. 103: Análise por SDS-PAGE 12% redutora de amostras do teste de ligação de BJ-48 em coluna Con A. A ligação e lavagem foi feita em tampão 50 mM NaOAc pH 6,0 + 0,5 M NaCl (tampão 1). A eluição foi feita em tampão 1 + 500 mM de maltose. Canaletas: 1-aplicado; 2- não-ligado; 3- lavado; 4 a 7- frações eluídas 1 a 4. O gel foi revelado por coomassie G-250.

4.4.2 Análise enzimática da porção carboidrato em Lmr-47

Com o intuito de caracterizar enzimaticamente a quantidade e a composição de carboidratos em Lmr-47, amostras da glicoproteína foram submetidas à ação de glicosidases com diferentes especificidades e então analisadas por SDS-PAGE. Conforme demonstrado na **Fig. 104**, a Mr de Lmr-47 foi reduzida em 2 kDa (4% p/p), mediante ação da α -2(3,6,8,9)neuraminidase, uma sialidase capaz de liberar resíduos de ácido siálico (ácido N-acetilneuramínico) unidos por virtualmente todo tipo de ligação glicosídica aos demais resíduos de carboidrato da glicoproteína. A ação concomitante da enzima O-glicosidase (endo- α -N-acetilgalactosaminidase) não produziu aumento adicional na mobilidade eletroforética da amostra tratada de Lmr-47. Isto indica a ausência de carboidratos O-ligados com cernes Gal- β (1-3)-GalNAc simples mono-, di- ou Trisialilados em Lmr-47. Entretanto, esta observação por si só não exclui a

possibilidade de resíduos de Ser/Thr apresentando este cerne com outros tipos de modificação.

O tratamento de Lmr-47 com a enzima PNGase F resultou num decréscimo da Mr da glicoproteína para 28-26 kDa, indicando a presença de 36-40% de carboidratos N-ligados em Lmr-47. Tendo em vista a massa molecular de 25,6 kDa prevista pela seqüência de Lmm-TLE, é possível concluir que praticamente todo o carboidrato presente em Lmr-47 encontra-se ligado à resíduos de Asn. O programa NetNGlyc aponta 4 sítios potenciais de N-glicosilação na seqüência de Lmm-TLE: N59, N95, N164 e N245.

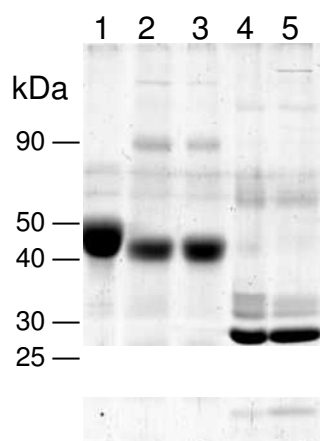


Fig. 104: Análise por SDS-PAGE 12% redutora de amostras da deglicosilação enzimática de Lmr-47 em condições desnaturantes. Amostras de Lmr-47 foram submetidas à ação de glicosidases de diferentes especificidades. Ao término de 18h de incubação a 37 °C, as reações foram paralisadas por fervura em tampão de amostra da SDS-PAGE. Canaletas: 1- controle com nenhuma glicosidase adicionada à glicoproteína; 2- Lmr-47 + α -2(3,6,8,9)neuraminidase; 3- Lmr-47 + neuraminidase + O-glicosidase; 4- Lmr-47 + PNGase F; 5- Lmr-47 + PNGase F + O-glicosidase + neuraminidase + β (1-4)-galactosidase + β -N-acetilglicosaminidase. O gel foi revelado por coomassie G-250.

A conclusão de que Lmr-47 contém apenas carboidratos N-ligados foi apoiada pelo ensaio mostrando que o tratamento concomitante desta glicoproteína com todas as enzimas mencionadas anteriormente mais as enzimas β (1-4)-galactosidase e β -N-acetilglicosaminidase, não resultou em aumento adicional da mobilidade eletroforética em comparação com a amostra tratada com PNGase F somente (**Fig. 104**). Este resultado também demonstra que a estrutura de carboidratos de Lmr-47 é livre de núcleos Gal- β (1-3)-GalNAc- α -O-Ser/Thr contendo a modificação comum NeuAc- α (2-3)Gal- β (1-4)GlcNAc- β (1-6) no resíduo de GalNAc.

4.5 Análise da estrutura em solução de Lmr-47

4.5.1 CD

A estrutura secundária de Lmr-47 em soluções aquosas tamponadas em pH 6,0 e pH 7,5, foi avaliada pelos espectros de CD na região do UV distante (190-260 nm). O pH 6,0 corresponde ao pH do veneno de *L. muta rhombeata* in natura, enquanto que em pH 7,5 a enzima exibe seu máximo de atividade amidásica. Os espectros sobrepostos na **Fig. 105** mostram que a estrutura secundária da enzima é mantida nestes dois valores de pH. Os espectros obtidos foram característicos de uma proteína $\alpha+\beta$, com uma forte banda negativa de absorção na faixa de 215 a 208 nm. A deconvolução do espectro de CD de Lmr-47 no programa CDNN mostrou que a enzima é composta por cerca de 40% de folhas- β , 30% de estrutura aleatória, 20% de grampos- β e 10% de α -hélices. Esta composição é característica de membros da família da quimiotripsina, tais como tripsina e elastase.

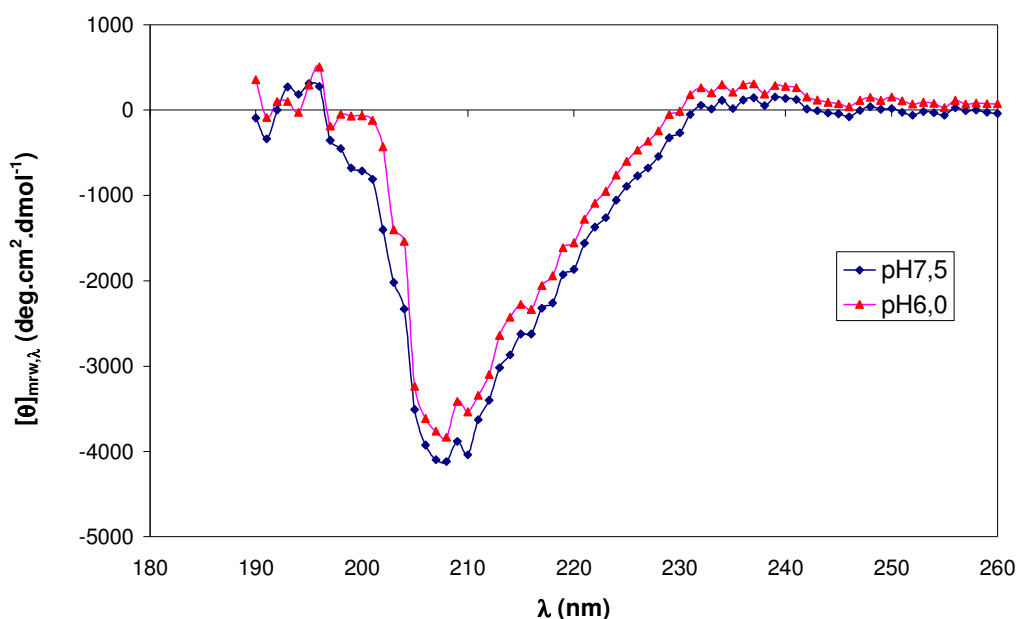
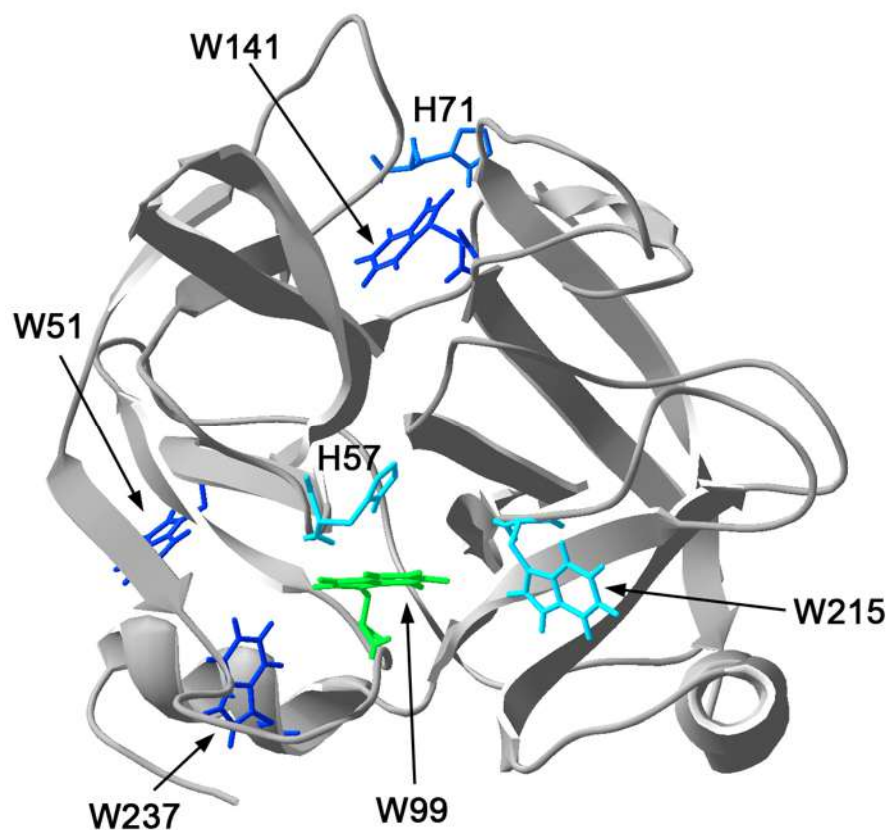


Fig. 105: Espectros de CD de Lmr-47 em soluções aquosas tamponadas em pH 6,0 e 7,5.

4.5.2 Fluorescência

Espectros de fluorescência intrínseca de Lmr-47 foram medidos em soluções aquosas tamponadas em valores de pH variando de 5,0 a 8,0 (**Fig. 106**). Independente do valor de pH, dois comprimentos de onda de fluorescência máxima puderam ser identificados, $\lambda'_{\max} = 230$ nm e $\lambda''_{\max} = 238-239$ nm, compatíveis com a emissão de resíduos de Trp na proteína. A existência destes dois valores de λ_{\max} na fluorescência intrínseca de Lmr-47 está de acordo com a distribuição

dos 5 resíduos de Trp na estrutura da enzima, conforme inferido pelo modelo por homologia (Fig.



107

Fig. 107). λ''_{\max} , mais deslocado para o vermelho corresponderia aos dois resíduos de Trp próximos à superfície da molécula (W99 e W215) enquanto que os três outros (W51, W141 e W237), mais protegidos do solvente no interior da enzima, seriam responsáveis por λ'_{\max} . O CME é pouco afetado pela variação do pH, sofrendo um deslocamento para o vermelho (maiores valores de λ) inferior a 2 nm. Isto indica que a variação do pH não provoca alterações conformacionais em Lmr-47 que resultariam numa maior ou menor exposição dos resíduos de Trp ao solvente.

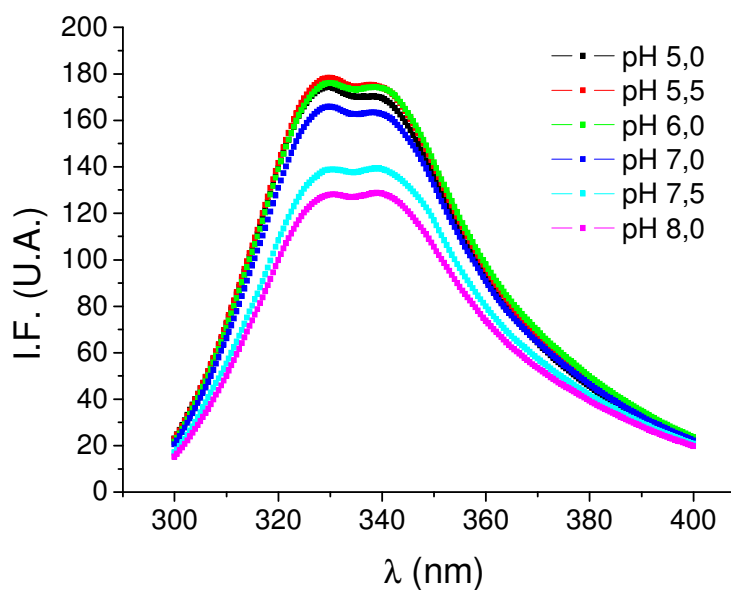


Fig. 106: Espectros de fluorescência intrínseca de Lmr-47 em soluções aquosas tamponadas em pH variando de 5,0 a 8,0.

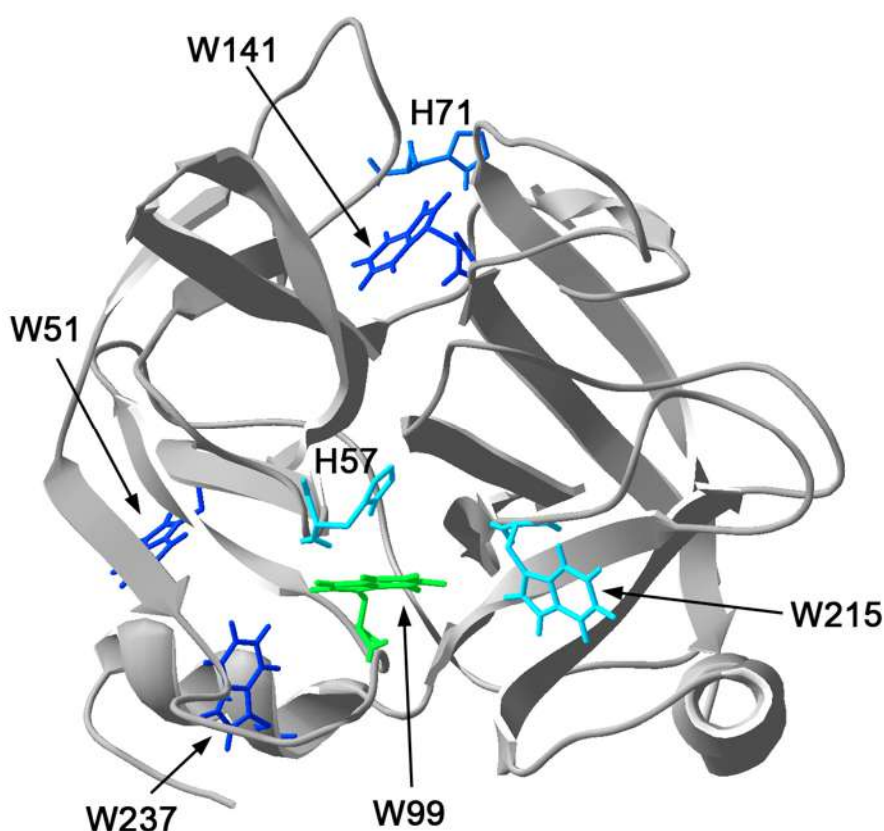
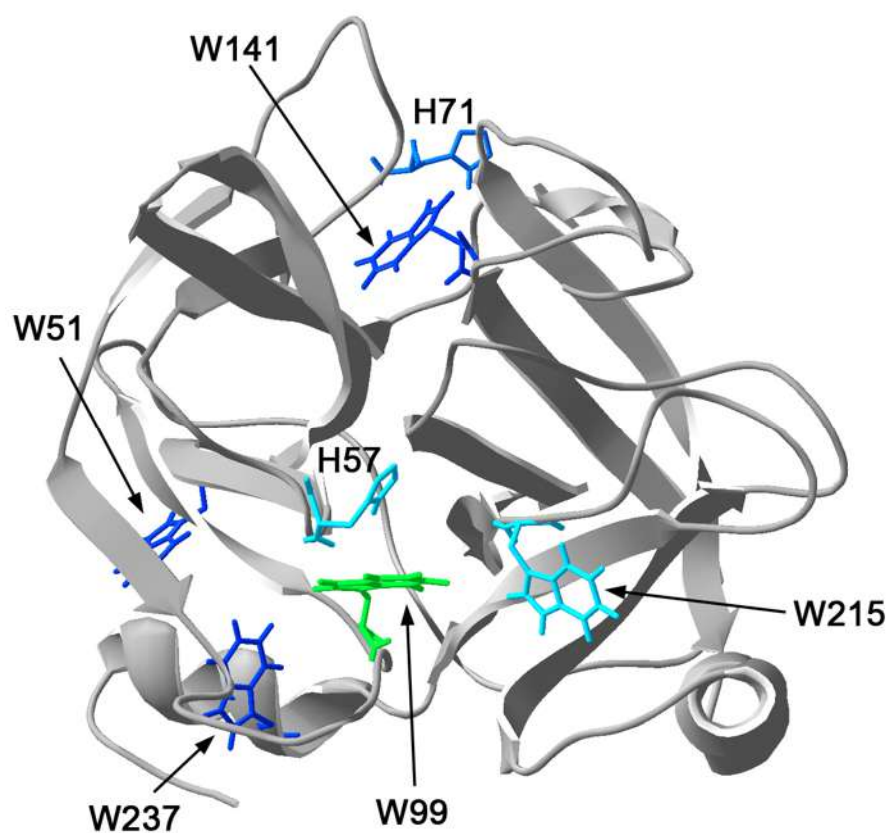


Fig. 107: Localização espacial de resíduos potencialmente envolvidos nas propriedades de fluorescência de Lmr-47. O modelo teórico de Lmr-47 é mostrado com os resíduos identificados e coloridos por acessibilidade ao solvente (vermelho – totalmente exposto, azul – totalmente protegido). Note a distribuição dos resíduos de Trp em um grupo pouco acessível ao solvente (W141, W51 e W237) e outro parcialmente acessível (W99 e W215). Dois resíduos de His em distância de contato ($< 4,0 \text{ \AA}$) com W99 (His⁵⁷) e

W141 (His⁷¹) são destacados. Estes podem estar envolvidos na supressão da fluorescência de Lmr-47 que é observada com o aumento do pH.

Entretanto, uma clara tendência à diminuição na I.F. com o aumento do pH pôde ser observada nos espectros de Lmr-47. Esta tendência é mais claramente percebida ao se grafar a I.F. total sob os espectros da **Fig. 106** (obtida pela integração entre 310 e 370 nm) em função do pH (**Fig. 108**). É possível notar que a transição estrutural relacionada ao decréscimo da fluorescência ocorre ao redor de pH 7,0. Uma vez que o CME não varia significativamente com o pH, o decréscimo na I.F. não pode ser associado a uma possível maior exposição ao solvente e conseqüente supressão da emissão do Trp pelo oxigênio dissolvido na água, mediante aumento do pH. Uma proposta compatível com os dados experimentais seria uma em que a emissão de parte dos resíduos de Trp fosse suprimida por grupos próximos na estrutura da proteína, sofrendo ionização mediante aumento do pH de 5,0 para 8,0. Resíduos de His sofrem ionização nesta faixa de pH e são bons candidatos para exercer esse papel em Lmr-47. De fato, dois resíduos de His são encontrados próximos a resíduos de Trp em cada um dos dois grupos de resíduos de Trp na molécula de Lmr-47 (**Fig.**



107

Fig. 107). His⁵⁷, próxima a Trp⁹⁹, integra a tríade catalítica e reconhecidamente possui pKa próximo à 7,0 em todas as serinil proteinases quimiotripsina-símile, coincidente com o pH de transição na estrutura de Lmr-47, detectado pela redução na I.F. (**Fig. 108**). O outro resíduo, His⁷¹, situa-se próximo a Trp¹⁴¹ na “alça de autólise”. No mecanismo aqui proposto, a energia absorvida durante a excitação de Trp⁹⁹ e Trp¹⁴¹ pode ser transferida por ressonância (RET) mediante contato com os resíduos His⁵⁷ e His⁷¹, respectivamente. O aumento do pH favorece a ionização destes últimos, que por

sua podem absorver por RET a energia de excitação dos resíduos de Trp próximos e suprimir a fluorescência destes.

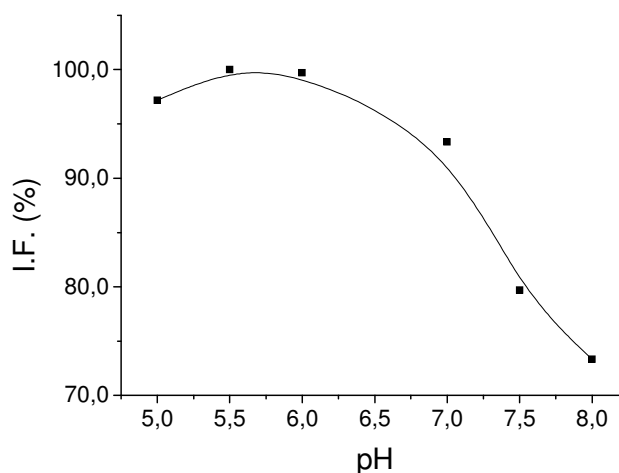


Fig. 108: Efeito do pH na intensidade de fluorescência de Lmr-47. A intensidade de fluorescência em cada valor de pH corresponde ao valor da integral no intervalo 310-370 nm e é expressa em unidades percentuais em relação à intensidade máxima. Uma transição em I.F. de Lmr-47 pode ser percebida em pH 7,0.

4.5.3 SAXS

Dados de SAXS foram coletados para Lmr-47 na concentração de 4 mg/mL e pH 6,0. Este valor de pH foi selecionado para evitar a autoproteólise que poderia ocorrer numa solução tamponada no pH ótimo de atividade da enzima (pH 8,0). A **Fig. 109** mostra o gráfico da intensidade em função do vetor de espalhamento [$I(q)$, **Fig. 109A**], a função de distribuição de distâncias [$p(r)$, **Fig. 109B**] e o gráfico de Kratky ($I \cdot q^2 \times q$; **Fig. 109C**) para as medidas realizadas a 10 °C. A curva $I(q)$ exibida na **Fig. 109A** é resultante de experimentos realizados em duas distâncias amostra-detector distintas, visando-se obter dados de boa qualidade estatística em um grande intervalo de q . A função $p(r)$ foi calculada a partir da transformada indireta de Fourier da função $I(q)$. Para isto usou-se o programa Gnom, onde valores de D_{\max} foram fornecidos num processo iterativo até obter um ajuste ótimo dos dados com $D_{\max} = 102 \text{ \AA}$. O valor de $R_g = 34,5 \pm 0,6$ obtido a partir da função $p(r)$ foi compatível com o valor fornecido pelo gráfico de Guinier ($\ln I \times q^2$; não mostrado). Este último mostrou-se linear na faixa de $q \cdot R_g < 1,3$, indicando que Lmr-47 apresentava-se monodispersa.

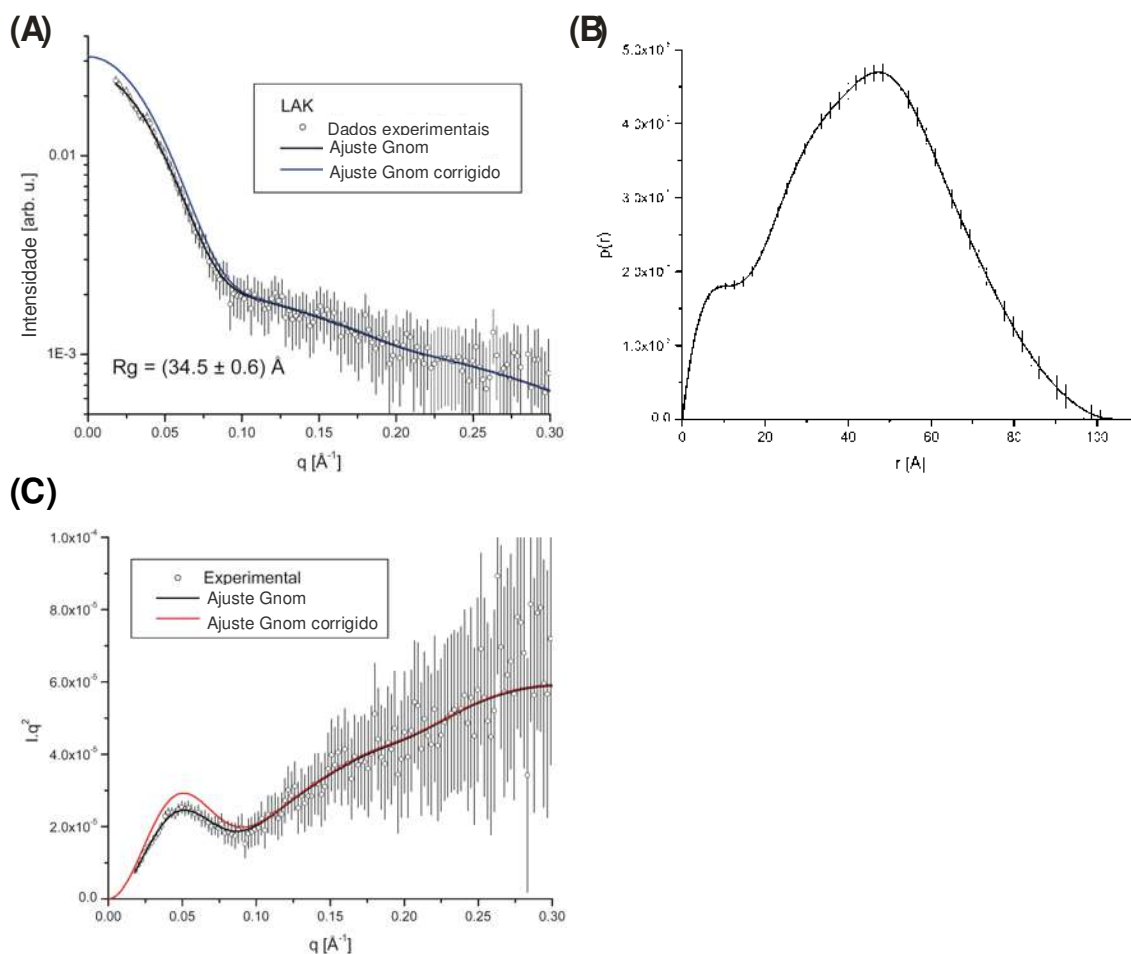


Fig. 109: Análise de dados de SAXS para Lmr-47 em pH 6,0 e 10 °C. Os dados de espalhamento foram tratados com o programa TRAT1D e as intensidades corrigidas para distorções por efeito de pequenas distâncias no programa Gnom. (A) curva $I(q)$, com o valor de R_g deduzido da análise. (B) função de distribuição de distâncias, $p(r)$. (C) Gráfico de Kratky.

O ajuste da função $p(r)$ aos dados experimentais foi excelente para valores de $0,10 \text{ \AA} < q > 0,30 \text{ \AA}$ e satisfatório para $q < 0,10 \text{ \AA}$, conforme percebido pelo gráfico da curva $I(q)$ ajustada pelo Gnom (Fig. 109A). O P.M. de $43 \pm 2 \text{ kDa}$ obtido em $I(0)$ é concordante com o valor de 45 kDa obtido pela análise por SDS-PAGE. Isto corrobora a qualidade do ajuste e também indica que Lmr-47 encontrava-se monomérica nas condições experimentais. O elevado valor de R_g obtido para Lmr-47 e os aspectos da função $p(r)$ (Fig. 109B) indicam que esta proteína não apresenta uma forma globular em solução. Uma proteína globular com cerca de $41\text{--}45 \text{ kDa}$ deveria apresentar R_g em torno de 20 \AA , enquanto que o ombro na função $p(r)$ indica que a proteína apresenta uma forma “oca” em solução, como um “U”, por exemplo. Por fim, o gráfico de Kratky (Fig. 109C) é compatível com esta conclusão, pois indica que a proteína possui uma

conformação flexível. Proteínas globulares compactas apresentam gráfico de Kratky com forma de sino.

O modelo *ab initio* da estrutura em solução de Lmr-47 em pH 6,0, apresentado na **Fig. 110A** foi gerado pelo programa DamAver a partir da média de dez modelos independentes gerados pelo programa Gasbor a partir dos dados experimentais de SAXS. Cada modelo gerado no programa Gasbor foi construído a partir de um número de “dummy residues” (DRs), estimado pelo peso molecular da proteína (cada resíduo ~ 110 Da), os quais foram configurados no espaço de forma a obter o melhor ajuste possível da função $p(r)$ do modelo com a função $p(r)$ experimental. A qualidade deste ajuste pode ser inferida pela satisfatória acomodação da curva $I(q)$, predita pelo modelo médio final, aos dados de espalhamento experimental (**Fig. 110B**). O modelo de Lmr-47 apresenta a forma de “Y” contorcido, com um dos segmentos do “Y” mais curto que os demais (**Fig. 110A**). Assim, a forma da partícula espalhadora predita pelo modelo *ab initio* é compatível com a função $p(r)$ e o gráfico de Kratky calculados a partir dos dados de espalhamento experimental, que indicam uma forma não-globular e flexível para a partícula.

A forma de Lmr-47 em pH 6,0 determinada por SAXS é marcadamente distinta da forma globular compacta do modelo teórico da enzima obtido por comparação com estruturas de proteínas homólogas cristalizadas em pH 7,5-8,0. De acordo com os dados de CD e fluorescência de Lmr-47 (vide seções 4.5.1 e 4.5.2 acima), a diferença no pH em que os modelos foram construídos não justificaria tamanha discrepância entre eles. Entretanto, uma diferença fundamental entre os dois modelos é o modelo com resolução atômica, obtido por modelagem comparativa, representa uma partícula de aproximadamente 25 kDa, enquanto que o modelo determinado por SAXS representa uma partícula de 45-47 kDa. A diferença na massa dos modelos provém de cadeias de carboidrato em Lmr-47, totalizando 40% do peso molecular total da enzima, distribuídas e 4 potenciais sítios de N-glicosilação (vide seção 4.4.2). A sobreposição destes dois modelos na **Fig. 111** ilustra bem a diferença em suas dimensões e formas. Tendo em vista que de cada sítio de glicosilação em Lmr-47 “brota”, em média, uma cadeia ramificada e flexível contendo 20-25 unidades piranosídicas, é razoável propor que as propriedades de SAXS da enzima nativa sejam dominadas pelos carboidratos na superfície da enzima.

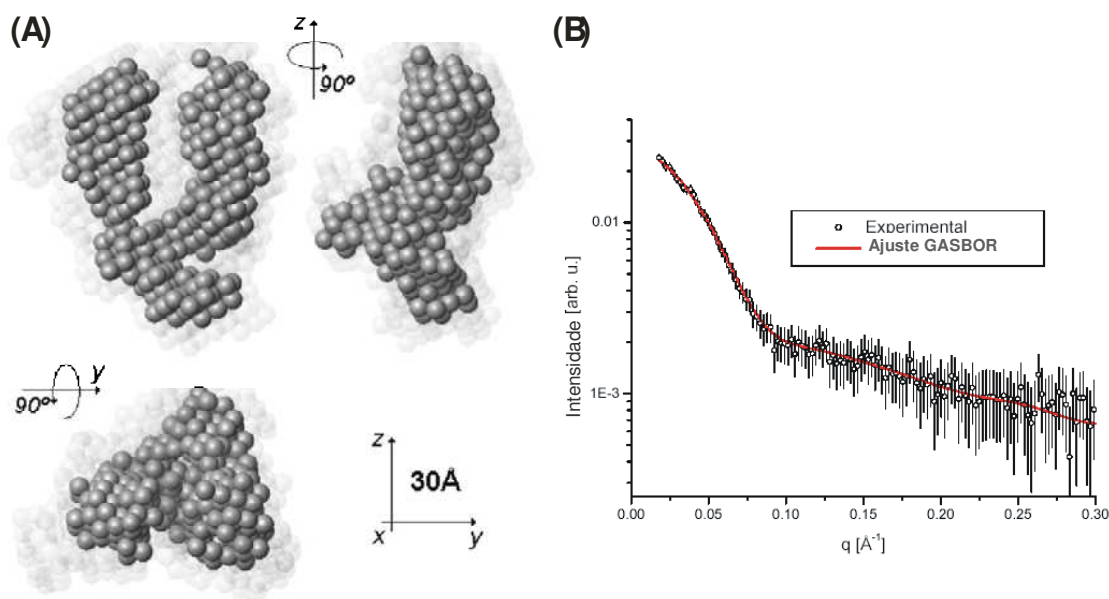


Fig. 110: Modelo *ab initio* de Lmr-47 calculado a partir dos dados experimentais de SAXS. (A) Modelo final (esferas sólidas) obtido da media de 10 modelos independentes gerados pelo programa Gasbor. As esferas translúcidas representam o volume total ocupado pela sobreposição dos 10 modelos. (B) Ajuste da função $I(q)$ predita pelo modelo sobre os dados de espalhamento experimentais.

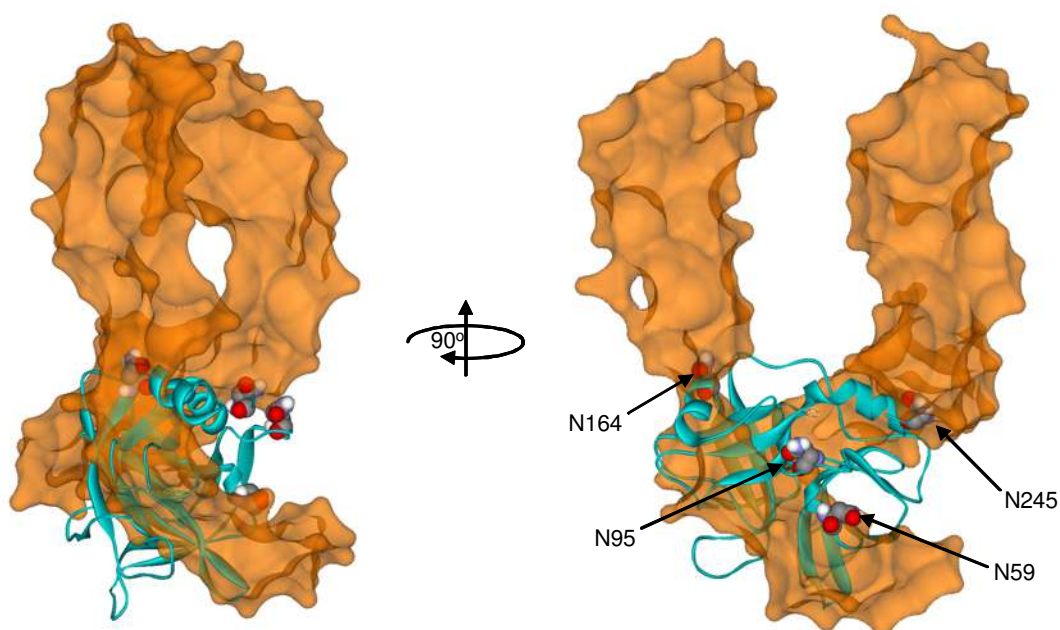


Fig. 111: Comparação entre o envelope molecular de Lmr-47 obtido por SAXS e o modelo teórico da cadeia polipeptídica da enzima. O enovelamento globular da cadeia polipeptídica, conforme previsto pelo modelo de resolução atômica, é representado em fitas sólidas em azul claro. O envelope molecular previsto pelo modelo *ab initio* de baixa resolução é representado pela superfície laranja translúcida. Os resíduos de Asn envolvidos nos potenciais sítio de N-glicosilação de Lmr-47 são mostrados como esferas de vdW.

4.6 Análise do papel do carboidrato na estrutura e função de Lmr-47 e BJ-48

4.6.1 Deglicosilação enzimática de BJ-48 e Lmr-47 em condições não-desnaturantes

Uma maneira de avaliar o papel da porção de carboidratos na estrutura e função de BJ-48 e Lmr-47 é realizar estudos comparativos entre as formas nativas (glicosiladas) destas proteases e formas artificialmente produzidas pela deglicosilação parcial e/ou completa das enzimas. Para isto, a deglicosilação em condições não-desnaturantes é requisito crucial. Considerando que, essencialmente, todo carboidrato presente nas enzimas BJ-48 e Lmr-47 está ligado a resíduos de Asn, procurou-se estabelecer condições apropriadas para a ação eficiente da enzima PNGase F sobre a estrutura nativa das glicoproteínas. Reconhecidamente, o principal empecilho para a deglicosilação nativa de glicoproteínas efetuada pela enzima PNGase F é o bloqueio estérico da ligação entre GlcNAc-Asn na base da cadeia de carboidrato ligada à proteína. Assim, pesquisou-se o desempenho da enzima PNGase F na deglicosilação de Lmr-47 e BJ-48 nativas em meios tamponados em pH 6,0 e 8,5, onde as glicoproteínas poderiam apresentar conformações distintas. Em uma destas conformações, o acesso da glicosidase ao sítio de N-glicosilação poderia ser mais facilitado do que na outra. Os resultados destes experimentos foram analisados por SDS-PAGE e são mostrados na **Fig. 112**. Em geral, conforme esperado, a deglicosilação de ambas, Lmr-47 e BJ-48, em condições nativas é muito menos eficiente comparada à deglicosilação das enzimas previamente desnaturadas. É interessante notar que mesmo em condições desnaturantes, a deglicosilação de Lmr-47 e BJ-48 não é completa após 18h, sendo evidente a permanência de formas parcialmente deglicosiladas. Com 18h de incubação a 37 °C, o desempenho da enzima PNGase F na deglicosilação de BJ-48 não-desnaturada em pH 8,5 já é marcadamente superior à deglicosilação em pH 6,0. O mesmo não ocorre com Lmr-47, onde a deglicosilação em pH 6,0 é apenas ligeiramente superior em eficiência ao pH 8,5.

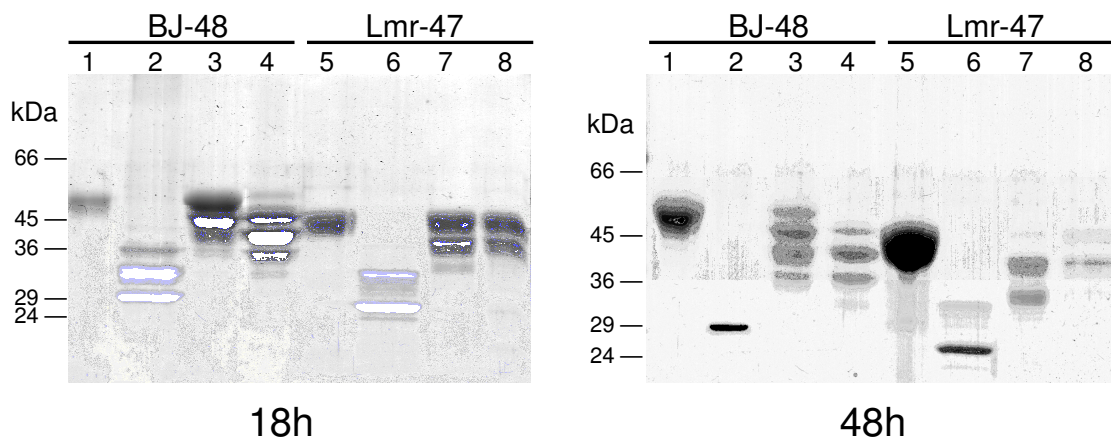


Fig. 112: Análise por SDS-PAGE 12% redutora de amostras do ensaio piloto de deglicosilação de BJ-48 e Lmr-47 com PNGase F. Amostras foram analisadas após 18 e 48h de incubação a 37° C. Canaletas: 1- controle sem adição de PNGase F à glicoproteína; 2- controle em condições desnaturantes; 3- deglicosilação nativa em pH 6,0; 4- deglicosilação nativa em pH 8,5. Os geles foram revelados por coomassie G-250.

É interessante notar que mesmo após 48h de incubação a 37° C em meios tamponados fora do pH de atividade recomendado (pH 7,0-7,5), a enzima PNGase F permanece ativa, conforme verificado pela deglicosilação virtualmente completa de BJ-48 e Lmr-47 em condições desnaturantes. Após 48h de incubação, o progresso da deglicosilação de BJ-48 nativa manteve-se superior em pH 8,5 comparado ao pH 6,0. Mais destacadamente, com o tempo prolongado de incubação, o pH 6,0 mostrou-se mais apropriado para a deglicosilação nativa de Lmr-47, pois permitiu um maior avanço da reação catalisada pela PNGase F. A deglicosilação mais favorável de Lmr-47 em meio ligeiramente ácido é interessante, pois em pH 6,0 a enzima apresenta apenas 60% da sua atividade amidásica máxima, dificultando a degradação auto-proteolítica da enzima.

De fato, uma importante observação originando destes ensaios preparatórios é que tanto BJ-48 como Lmr-47 permanecem estáveis à autoproteólise durante a deglicosilação, apesar da incubação prolongadas em condições favoráveis a este processo, como é o caso de BJ-48. Isto pode ser verificado pela ausência de bandas inferiores à 26-28 kDa, conforme esperado para fragmentos proteolíticos destas proteínas. Por outro lado, a ausência destas bandas de autoproteólise sugere que as bandas abaixo de 52 e 47 kDa, observadas durante a deglicosilação de BJ-48 e Lmr-47, respectivamente, referem-se exclusivamente às formas parcialmente deglicosiladas destas enzimas. A quantidade aparentemente elevada de bandas observadas está relacionada ao número de sítios de N-glicosilação ocupados nestas enzimas. Por exemplo, em Lmr-47 4 sítios em potencial estão presentes na seqüência primária.

Considerando todos os sítios como efetivamente ocupados, o total de formas possíveis desde a totalmente deglicosilada até a completamente deglicosilada soma 12 formas distintas.

Os resultados anteriores sugeriram que o caminho para a deglicosilação completa de BJ-48 e Lmr-47 estaria na aplicação de tempos prolongados de incubação, em meio tamponado no pH apropriado e na adição de quantidades extra de PNGase F para repor a fração da enzima que se perde naturalmente durante o processo. Entretanto, uma deglicosilação em escala “semi-preparativa” (1 mg) destas glicoproteínas foi primeiramente tentada em meio tamponado em pH 7,0, conforme sugerido pelo fornecedor da enzima PNGase F. Para evitar quaisquer problemas com autoproteólise, adições periódicas às reações de deglicosilação do inibidor reversível p-aminobenzamidina foram feitas. Com 72h de incubação, o progresso das reações em pH 7,0 ficou abaixo do esperado. A partir deste ponto procedeu-se a adição de neuraminidase para remover resíduos de ácido siálico que eventualmente estivessem contribuindo para um impedimento estérico da ação da PNGase F sobre os sítios de glicosilação. De fato, a análise das amostras após 72h de incubação confirmou as expectativas, apesar das formas totalmente deglicosiladas não terem sido detectadas (dados não mostrados). No entanto, com adições subsequentes de PNGase F em mais de 96h adicionais de incubação houve aparecimento de material floculado esbranquiçado nos tubos de ensaio de ambas as glicoproteínas, BJ-48 e Lmr-47, sugerindo a formação de proteína insolúvel. O material foi centrifugado e a análise de precipitado e sobrenadante por SDS-PAGE revelou a presença de quase todas as glicoproteínas na forma insolúvel (**Fig. 113A**). Mais ainda, a análise do precipitado revelou a presença de bandas referentes a formas parcialmente deglicosiladas diversas, bandas compatíveis com as formas totalmente deglicosiladas, além de bandas possivelmente relacionadas a produtos de autoproteólise de BJ-48 e Lmr-47. Esta análise indica que com o avanço da deglicosilação houve perda de estabilidade, desnaturação e agregação das glicoproteínas. Com a desnaturação, as enzimas não mais se mostraram resistentes à auto-degradação proteolítica (intermolecular).

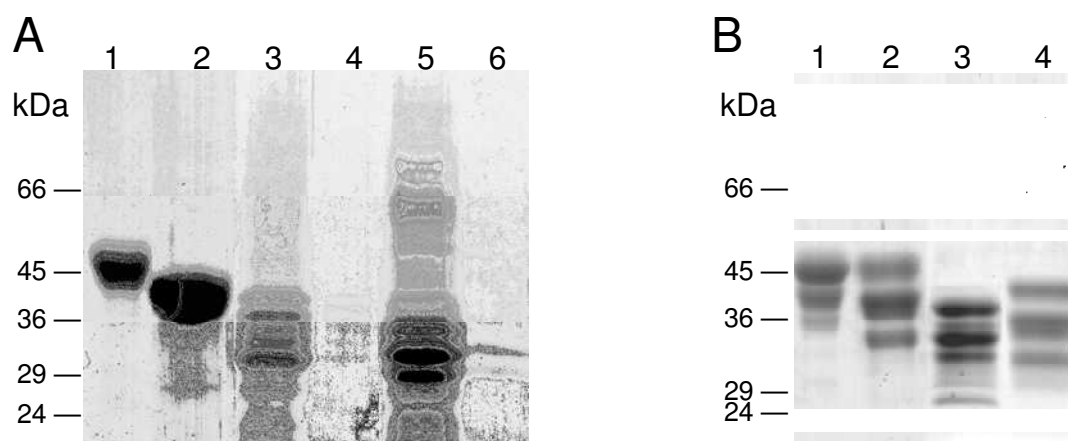


Fig. 113: Análise por SDS-PAGE 12% redutora de amostras da deglicosilação de BJ-48 e Lmr-47 nativas. A- amostras da reação em escala “semi-preparativa” sem adição de detergente ao tampão de reação. Canaletas: 1 e 2- controles de BJ-48 e Lmr-47, respectivamente, sem adição de PNGase F; 3/5- pellets e 4/6 sobrenadantes de amostras de BJ-48 e Lmr-47, respectivamente, obtidas após 240h de reação. B- amostras da reação em escala “preparativa” na presença de 1% CHAPS. Canaletas: 1 e 2- BJ-48 e Lmr-47 com 120h de incubação; 3 e 4- BJ-48 e Lmr-47, 36h após a adição sequencial de neuraminidase. Os gels foram revelados por coomassie G-250.

Em um segundo procedimento de deglicosilação, agora em escala “preparativa” (2,4-5,0 mg), adotou-se os valores de pH ótimos para o tratamento das enzimas BJ-48 e Lmr-47 previamente estabelecidos pelo ensaio piloto. Com 120h de incubação a 37 °C, a deglicosilação mostrou-se apenas parcial e com eficiência abaixo da esperada em comparação com os resultados do ensaio piloto (Fig. 113B). Neste ponto, antes de adições subseqüentes de PNGase F e neuraminidase, adicionou-se 1% de CHAPS às reações com o objetivo de estabilizar as glicoproteínas. Isto seria conseguido através da ação detergente deste aditivo sobre porções hidrofóbicas expostas nas superfícies a medida que as glicoproteínas fossem perdendo suas cadeias de carboidrato.

Aparentemente, a adição do detergente teve influência positiva sobre a estabilidade das proteínas uma vez que após 48h de incubação, nenhum precipitado foi observado. Entretanto, as amostras das glicoproteínas não mostraram evolução significativa na deglicosilação por PNGase F neste período. Assim, com o tempo total de reação já ultrapassando 160h de incubação, adicionou-se neuraminidase às amostras. Conforme esperado, a análise por SDS-PAGE de amostras de 36h após adição da neuraminidase evidenciou progresso significativo da deglicosilação das glicoproteínas desialiladas (Fig. 113B). Não houve formação de precipitado, indicando que a presença do detergente CHAPS teve a ação positiva esperada sobre as formas deglicosiladas de BJ-48 e Lmr-47. Na amostra de BJ-48, pôde-se inclusive detectar uma banda

compatível com a forma totalmente deglicosilada da enzima. Para Lmr-47, três bandas principais foram detectadas na faixa de 32-40 kDa, mostrando que uma forma totalmente deglicosilada não foi obtida. A análise das reações nas 60h subseqüentes, mesmo após suplementação com PNGase F adicional, não revelou mudança no produto da deglicosilação das enzimas.

4.6.2 Análise química dos carboidratos liberados

Amostras intactas (não-hidrolisadas) dos carboidratos liberados da deglicosilação nativa de Lmr-47 e BJ-48 foram analisadas quanto ao conteúdo de oligossacarídeos mono-, di- e trisialilados. Para BJ-48, todos os oligossacarídeos liberados da deglicosilação apresentaram-se desialilados, enquanto que oligossacarídeos monosialilados foram encontrados na amostra proveniente de Lmr-47. Isto mostra que em Lmr-47, a eficiência do tratamento com neuraminidase não foi tão alta como em BJ-48, com algumas cadeias de carboidrato permanecendo sialiladas. Este resultado é compatível com a maior eficiência observada na deglicosilação de BJ-48 em relação à Lmr-47, uma vez que a presença de resíduos de ácido siálico nestas glicoproteínas parece dificultar a ação da enzima PNGase F. A razão molar de resíduos de ácido siálico em BJ-48 e Lmr-47 foi determinada como 12:1 e 20:1, respectivamente, sugerindo uma estrutura triantenada complexa para os carboidratos em BJ-48 e possivelmente uma estrutura tetraantenada com um resíduo adicional de ácido siálico ligado ao cerne para os carboidratos N-ligados complexos em Lmr-47.

Os cromatogramas da análise de monossacarídeos neutros na fração de carboidratos liberados de Lmr-47 e BJ-48 no sistema Dionex Biol C são mostrados na **Fig. 114**. Resíduos de fucose (Fuc), galactosamina (GalNAc na amostra original), glicosamina (GlcNAc na amostra original), galactose (Gal) e manose (Man) foram encontrados nos carboidratos liberados de ambas as SVSPs. Esta composição, exceto por Fuc, a qual é mais rara, é compatível com a estrutura típica de carboidratos N-ligados em eucariotos. Semi-quantitativamente, há uma nítida distinção entre as duas enzimas em relação ao conteúdo de GalNAc. A razão molar de monossacarídeos neutros determinada na fração de carboidratos liberados de BJ-48 foi Fuc(2):GalNAc(4):GlcNAc(5):Gal(2):Man(3), a qual é compatível com uma estrutura do tipo tri- ou tetraantenada complexa de glicanos N-ligados, com o resíduo de Fuc mais comumente encontrado conectado ao resíduo de GlcNAc ligado diretamente à

proteína através de ligação glicosídica $\alpha(1-6)$. A estrutura do carboidrato em Lmr-47 é semelhante, exceto pela menor quantidade de GalNAc (1).

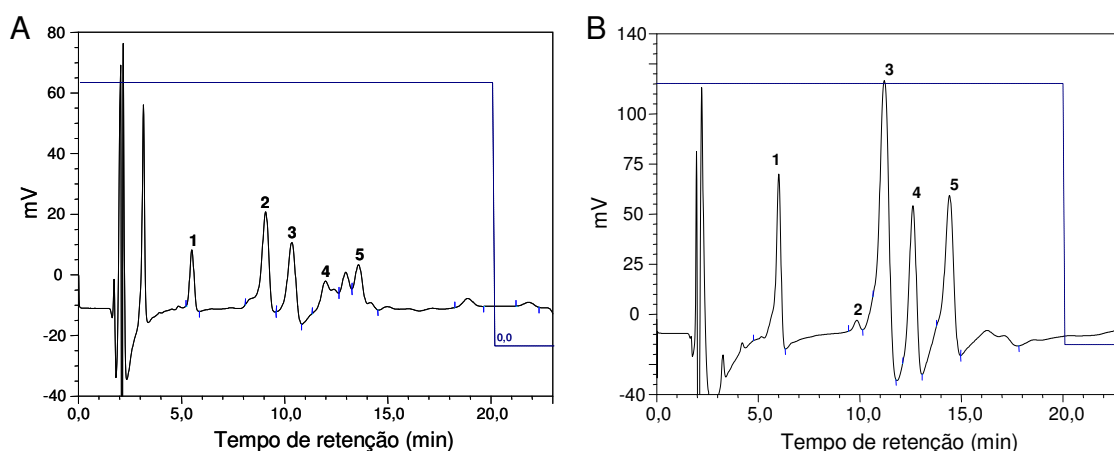


Fig. 114: Análise cromatográfica de monossacarídeos neutros liberados pela deglicosilação não-desnaturante de BJ-48 (A) e Lmr-47 (B) no sistema Dionex Biol C. Picos: 1- fucose; 2- galactosamina; 3- glicosamina; 4- galactose; 5- manose.

4.6.3 Resolução das formas parcialmente deglicosiladas de Lmr-47 e BJ-48 por HIC

O passo seguinte crucial para a determinação da função dos carboidratos em Lmr-47 e BJ-48 é desenvolver uma metodologia adequada para a separação das formas com diferentes conteúdos de carboidrato oriundas da deglicosilação em condições nativas. Como primeira opção para a separação, tentou-se a coluna de afinidade Con A, entretanto este suporte cromatográfico mostrou-se ineficaz devido à baixa resolução (dados não mostrados). Alternativamente, empregou-se cromatografia de exclusão molecular (SEC) na tentativa de separar os produtos de deglicosilação pela diferença em suas massas moleculares. Novamente, a técnica cromatográfica não apresentou resolução suficiente para separação das diferentes formas deglicosiladas (dados não mostrados).

Um terceiro método de separação, baseado em interação hidrofóbica, foi testado na tentativa de efetuar o fracionamento do produto de deglicosilação nativa de BJ-48 e Lmr-47. A [Fig. 115](#) mostra os perfis cromatográficos das amostras submetidas à coluna Phenyl 5PW, juntamente com os perfis eletroforéticos das frações coletadas.

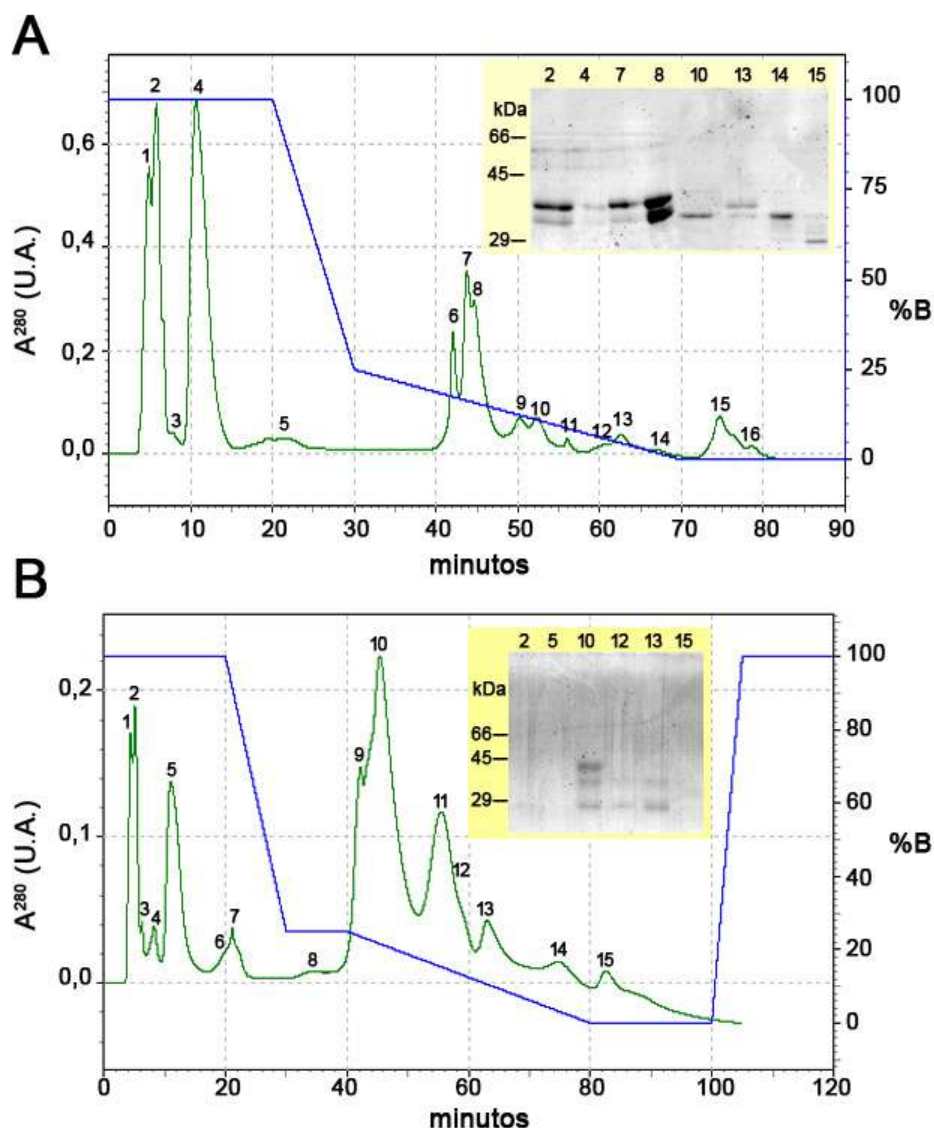


Fig. 115: Análise das amostras da deglicosilação nativa de BJ-48 (A) e Lmr-47 (B) fracionadas por HIC na coluna Phenyl 5PW. Os cromatogramas são apresentados com as frações coletadas conforme a numeração. A linha azul representa o gradiente de tampão B contendo 2M de Na₂SO₄. No destaque em amarelo são mostradas as análises por SDS-PAGE 12% redutora (revelação por coomassie G-250) de frações selecionadas da HIC.

Com as amostras de ambas glicoproteínas há uma nítida tendência de eluição das formas mais glicosiladas a frente das amostras menos glicosiladas, especialmente durante o gradiente de 25-0% de B (tampão com 2M de sulfato de sódio). O comportamento cromatográfico das formas com diferentes níveis de glicosilação é compatível com um papel dos carboidratos na blindagem do contato com o solvente de regiões hidrofóbicas na superfície destas glicoproteínas. De fato, formas mais glicosiladas interagem fracamente com o grupo fenil da coluna e são eluídas mesmo em altas concentrações de Na₂SO₄, onde a interação hidrofóbica é máxima. Com a

diminuição da concentração salina durante o gradiente de 100-0% do tampão B, a atividade aquosa é restituída, enfraquecendo a interação hidrofóbica e permitindo que componentes mais hidrofóbicos sejam eluídos da coluna.

Por exemplo, a fração 15 da cromatografia do produto de deglicosilação de BJ-48 (BJ-F15) é constituída predominantemente pela forma totalmente deglicosilada, sendo eluída somente depois que o gradiente atinge 0% de B (Fig. 115A). O oposto ocorre com frações de maior conteúdo de glicanos (maior Mr na análise eletroforética), como as frações 2 a 8 de BJ-48, as quais passam direto pela coluna ou são predominantemente eluídas no início do gradiente (Fig. 115A). Com algumas frações intermediárias, aparentes exceções a esta regra são encontradas, como por exemplo, nas frações 10 (BJ-F10) e 13 (BJ-F13), onde a primeira é predominantemente composta por uma espécie de menor Mr (menos glicosilada) que a segunda (Fig. 115A). Entretanto, este comportamento pode ser prontamente reconciliado com a hipótese proposta no parágrafo anterior se for imaginado que diferentes sítios de glicosilação nestas enzimas protegem do solvente regiões de diferentes níveis de hidrofobicidade em suas superfícies moleculares. Desta forma, o carboidrato adicional em BJ-F13 pode ser menos útil na blindagem de uma determinada região hidrofóbica nesta molécula do que a(s) cadeia(s) de carboidrato(s) em BJ-F10.

4.6.4 Estabilidade térmica

O efeito do carboidrato na estabilidade térmica de BJ-48 foi pesquisado medindo-se a redução na atividade amidásica específica de amostras desta enzima com diferentes conteúdos de glicanos, mediante aquecimento por 1h nas temperaturas de 55 e 65° C. Estes resultados estão listados na Tabela 12 e mostram claramente que os carboidratos em BJ-48 estão envolvidos na proteção desta enzima à desnaturação térmica, conforme medido pela perda de atividade amidásica.

Tabela 12: Avaliação do efeito do carboidrato na redução da atividade amidásica de BJ-48 mediante tratamento térmico.

Amostra	Atividade específica residual (%) [*]	
	55 °C	65 °C
Fração 15	49 ± 9	51 ± 7
Fração 10	58 ± 6	58,7 ± 4
Fração 13	70 ± 2	60,6 ± 5
BJ-48	84 ± 3	79,3 ± 5

^{*} Os resultados representam média de 3 replicatas ± erro padrão.

Após tratamento a 55 °C, ou 65 °C a enzima nativa, contendo cerca de 40% de carboidrato, retém ao redor de 80% de sua atividade amidásica. Diferentemente, BJ-F15, representando predominantemente a forma deglicosilada da enzima, sofre um decréscimo 30% maior na atividade amidásica mediante tratamento térmico idêntico. As espécies BJ-F10 e BJ-F13 com níveis intermediários de glicosilação apresentam reduções aproximadamente proporcionais na estabilidade térmica.

4.6.5 Resistência a inibidores protéicos

A resistência a Infestina e SBTI, dois inibidores protéicos aos quais BJ-48 é naturalmente refratária foi testada nas frações da deglicosilação desta enzima. A inibição foi medida em termos da redução na atividade amidásica específica das amostras na presença do inibidor. Os resultados apresentados na **Tabela 13** mostram que a resistência à Infestina é indiferente à presença de carboidrato enquanto que a inibição por SBTI é afetada pelo carboidrato em BJ-48. SBTI provoca uma redução leve, de apenas 20%, na atividade da enzima de BJ-48 nativa comparada à completa abolição da atividade da tripsina bovina, nas condições experimentais, enquanto que as formas deglicosiladas de BJ-48 apresentam reduções de 50-65% em suas atividades. As espécies BJ-F15 e BJ-F10, apresentam níveis de inibição por SBTI, que dentro do erro experimental, podem ser consideradas equivalentes. BJ-F13, com um conteúdo um pouco maior de carboidrato é ligeiramente menos susceptível à ação de SBTI. Estes resultados indicam que uma ou mais cadeias de carboidrato em BJ-48 contribuem para a baixa atividade de SBTI sobre a enzima, mais provavelmente através de um bloqueio estérico de parte da fenda de ligação ao substrato.

Tabela 13: Avaliação do efeito do carboidrato na resistência de BJ-48 à inibição da atividade amidásica por inibidores protéicos.

Amostra	Atividade específica residual (%) [*]	
	Infestina	SBTI
Tripsina	N.D.	0,0 ± 0,2
Fração 10	100 ± 5	35 ± 3
Fração 15	100 ± 5	44 ± 8
Fração 13	100 ± 5	48 ± 1
BJ-48	100 ± 5	77 ± 2

^{*} Os resultados representam a média de 3 replicatas ± erro padrão. N.D. = não determinado.

5 DISCUSSÃO

5.1 Especificidade das SVSPs BJ-48 e Lmr-47

5.1.1 Subsítio S1

5.1.1.a BJ-48 e Lmr-47 são enzimas específicas para Arg em P1

Os estudos cinéticos com Lmr-47 e BJ-48 mostraram que estas enzimas apresentam preferência por Arg sobre Lys em P1 de substratos tripeptídicos na ordem de 10^6 e 10^4 , respectivamente (**Tabela 10**). Esta extraordinária especificidade é absolutamente extraordinário para enzimas da família da quimiotripsina, sendo somente equiparada à preferência de 10^5 de enzimas tripsina-símile por Arg/Lys em relação ao segundo resíduo favorecido na posição P1 de substratos (**CZAPINSKA & OTLEWSKI, 1999**). A discriminação entre Arg e Lys em P1 é ainda mais excepcional. Por exemplo, tripsina e trombina preferem Arg sobre Lys em P1 de tripeptídeos *p*-nitroanilida com a seqüência D-FPX, como os usados neste trabalho, por um fator de cerca de apenas 10 vezes (**VINDIGNI, 1999**). Gabonase, uma TLE do veneno de *Bitis gabonica* cliva com baixíssima eficiência ($K_s \sim 10^2 \text{ M}^{-1} \cdot \text{s}^{-1}$) substratos com a estrutura N α -Tos-Gly-Pro-Xaa, porém é praticamente incapaz de discriminar Arg de Lys em P1 (**PIRCKLE et al., 1986**). A mesma falta de seletividade foi observada com PA-BJ, uma TLE ativa sobre a agregação plaquetária isolada do veneno de *B. jararaca*, usando substratos do tipo D-Val-Leu-Xaa (**SERRANO et al., 1995**).

É importante mencionar que a preferência relativa de uma serinil proteinase pelo resíduo em P1 pode variar, dependendo dos resíduos ocupando P2 e P4. Conforme mostrado nos estudos enzimáticos no ANEXO 5, trombina é ainda mais seletiva (~20 vezes) por Arg em relação a Lys na posição P1 de substratos do tipo N α -Tos-Gly-Pro-Xaa. Esta seletividade dobrada por Arg sobre Lys em P1 resulta de uma redução desproporcional na eficiência de clivagem destes substratos pela trombina. O mesmo ocorre com BJ-48 que simplesmente não cliva após Lys destes substratos enquanto cliva o derivado com Arg em P1 com eficiência 5000 vezes menor. Em Lmm-TLE, a qual se considera neste trabalho como essencialmente idêntica à Lmr-47, a preferência por Arg sobre Lys em P1 é reduzida a um fator de 100 vezes quando os parâmetros cinéticos são determinados sobre D-Val-Leu-Xaa-pNA (**MAGALHÃES et al., 2003**), um substrato típico de caliceínas. Estas observações estão de acordo com o paradigma corrente estabelecendo que os subsítios distais S4-S2 contribuem para a especificidade primária

ao registrar o substrato para o ataque nucleofílico eficiente no sítio ativo (PERONA & CRAIK, 1995; HEDSTROM, 2002).

Algumas enzimas tripsina-símile são capazes de discriminar Arg de Lys em P1, entretanto nenhuma se compara a BJ-48 ou Lmr-47. Por exemplo, tPA é 70 vezes mais específica na clivagem de substratos contendo Arg ao invés de Lys em P1 enquanto que plasmina é apenas 50% mais eficiente contra Lys sobre Arg em P1 (VINDIGNI, 1999). Ativadores de plasminogênio de veneno de serpentes também apresentam preferência na ordem de 10^2 por Arg sobre Lys em P1 (ZHANG *et al.*, 1995; ZHU *et al.*, 2005). Calicreínas da glândula submaxilar de camundongo são usualmente destacadas por suas seletividades em favor de Arg sobre Lys em P1 de substratos cromogênicos curtos. Entretanto, esta preferência é de apenas 40-80 vezes (BLABER *et al.*, 1989).

5.1.1.b Evolução da especificidade de substrato em SVSPs: similaridades com calicreínas teciduais

A avaliação da especificidade do subsítio S1 das SVSPs Lmr-47 e BJ-48 revelou ainda uma importante diferença entre as duas enzimas (Tabela 10). BJ-48 atua como uma enzima tripsina-símile clássica, sendo incapaz de clivar substratos sem uma carga positiva em P1. Por outro lado, Lmr-47 é não somente capaz de fazê-lo, como é mais específica para o substrato com Phe em P1 do que para o análogo com Lys nesta posição. Curiosamente, Lmr-47 foi incapaz de clivar D-FPY-pNA. Como o grupo hidroxil na posição 4 da fenila de Tyr pode prejudicar tanto a eficiência catalítica de Lmr-47 sobre este substrato é ainda desconhecido e a questão merece estudos adicionais. A especificidade de Lmr-47 é possivelmente remanescente das propriedades de um ancestral comum com as calicreínas teciduais.

Calicreínas constituem um grupo heterogêneo de serinil proteases pertencentes ao clã da quimiotripsina, caracterizadas por uma especificidade geral tripsina-símile, porém apresentando um espectro de substratos macromoleculares bem definido (revisado por YOUSEF & DIAMANDIS, 2003). Calicreínas são usualmente divididas em calicreínas plasmáticas e teciduais, diferindo quanto à massa molecular, especificidade de substrato, características imunológicas, estrutura gênica e no tipo de cinina liberada. Em humanos, a calicreína plasmática é codificada por um único gene (KLKB1) composto por 15 exons localizado no lócus 4q35. A enzima é expressa exclusivamente no fígado e tem participação na cascata da coagulação e na regulação de

processos inflamatórios e no tônus muscular, estas últimas ações através da liberação de bradicinina a partir do cininogênio de alto peso molecular.

A calicreína tecidual libera o peptídeo vasoativo e espasmogênico Lys-bradicinina, também conhecido como calidina, pela proteólise altamente específica dos cininogênios de alto e baixo peso molecular. Em função disto, calicreína tecidual é um importante regulador da pressão e fluxo sanguíneo. A proteína foi encontrada até o momento em seis diferentes ordens de mamíferos: Primatas, Rodentia, Carnivora, Proboscidea, Perissodactyla e Artiodactyla (YOUSEF & DIAMANDIS, 2003). Mais recentemente, o conceito de “família multigênica de calicreínas” foi introduzido, no qual o termo genérico “calicreína tecidual” não é restrito à descrição de enzimas que liberam peptídeos bioativos (cininas) de moléculas precursoras (cininogênios). O termo agora é usado para descrever um grupo de enzimas com estrutura gênica e protéica, que também partilham similaridade na seqüência e colocalizam-se no mesmo locus cromossômico (DIAMANDIS *et al.*, 2004). O uso do termo “calicreína” não implica necessariamente que qualquer um dos membros desta família (exceto hK1 em humanos e mGK6 em camundongos) possui atividade cininogenase. Estas enzimas foram reunidas no mesmo grupo baseado em semelhanças estruturais e localização cromossômica.

Lmr-47 e calicreínas teciduais compartilham a capacidade de clivar preferencialmente ligações Arg-Xaa em substratos sintéticos pequenos. Além disso, Lmr-47 é capaz de clivar peptídeos contendo um aminoácido hidrofóbico em P1, uma característica empregada por calicreínas teciduais para liberar calidina através da hidrólise de ligações Met-Lys ou Leu-Lys de cininogênios. Em vista dessa semelhança na especificidade, é possível que o cininogênio seja um substrato para Lmr-47, uma atividade que não foi testada até o momento nesta enzima (AGUIAR *et al.*, 1996). De fato, atividade cininogenase foi demonstrada em outras SVTLEs de diferentes espécies, tais como crotalase de *Crotalus durissus* (MARKLAND *et al.*, 1982), Lmr-32 de *L. m. rhombeata* (DE-SIMONE *et al.*, 1997), KN-BJ 1 e 2 de *B. jararaca* (SERRANO *et al.*, 1998), e duas SVSPs de *Trimeresurus mucrosquamatus* (HUNG & CHIOU, 2001). Baseado na organização do gene de batroxobin envolvendo 5 exons e 4 íntrons, semelhantemente aos genes de calicreínas teciduais, foi proposto que estas últimas e SVSPs são homólogos distantes pertencentes a uma mesma família de enzimas glandulares (ITOH *et al.*, 1988). Os dados gerados neste trabalho mostrando a rara

semelhança na especificidade de substrato de Lmr-47 e calicreínas teciduais fornece evidência adicional sobre a origem evolutiva comum destas enzimas. Isto sugere que SVSPs divergiram de calicreínas glandulares e que suas funções mudaram durante a evolução. De fato, a comparação dos cDNAs codificando SVSPs mostrou que as regiões codificando a proteína são muito mais diversificadas que as regiões não-codificantes, e que substituições nucleotídicas não-sinonímicas são pelo menos tão freqüentes quanto, senão mais comuns, que substituições sinonímicas nas regiões codificando a proteínas (DESHIMARU *et al.*, 1996). Estas propriedades incomuns indicam que SVSPs evoluíram rapidamente e adquiriram funções fisiológicas diversas.

5.1.1.c Bases estruturais para a especificidade por Arg na família da quimiotripsina

Proteases Arg/Lys-específicas pertencentes à família da quimiotripsina tem sido isoladas de diversas fontes, de eubactérias até humanos. Entretanto, nenhuma explicação estrutural satisfatória foi fornecida até o momento para explicar a capacidade destas enzimas de discriminarem Arg de Lys na posição P1 de seus substratos.

Achromobacter protease I (aPRO) é uma lisil endopeptidase com estrutura 3D semelhante à tripsina bovina (subfamília S1D), apesar de compartilhar apenas 20% de identidade na seqüência (TSUNASAWA *et al.*, 1989; NORIOKA *et al.*, 1994). Foi proposto que a estrita especificidade por Lys em P1 resulta de uma organização incomum da bolsa de especificidade primária, onde dois resíduos ácidos E189 e D226 estão presentes (revisado por CZAPINSKA & OTLEWSKI, 1999). A estrutura cristalográfica do complexo covalente com Tos-Lys-CMK (PDB 1ARC) revela a ponte salina entre Lys do inibidor e D226. Como a cadeia lateral deste último é localizada a uma distância 1,5 Å menor que D189 na tripsina, foi colocado que o subsítio S1 em aPRO seria mais raso e, portanto, incapaz de acomodar a cadeia lateral mais longa de Arg em P1. Entretanto, esta proposta terá que ser revista, pois posteriormente uma endopeptidase com 30% de identidade com aPRO e especificidade para Arg em P1 foi clonada de *Lysobacter enzymogenes* exibindo os mesmos resíduos ácidos nas posições 189 e 226 (WRIGHT *et al.*, 1998). Conforme discutido na seção introdutória deste trabalho, a enzima da subfamília S1A, plasmina, também é capaz de discriminar Lys de Arg na posição P1 (VINDIGNI, 1997). Contudo, uma análise estrutural mais aprofundada ainda não foi realizada com esta enzima.

Entre as enzimas tripsina-símiles capazes de favorecer Arg sobre Lys na posição P1 dos substratos hidrolisados, um papel na definição desta especificidade tem sido atribuído ao resíduo ocupando a posição 190 na alça 1 compondo o subsítio S1 (CZAPINSKA & OTLEWSKI, 1999). Proteases como trombina, tPA e FXa apresentam a substituição S190A em relação à tripsina, a qual foi proposta ser responsável por tornar o subsítio S1 mais hidrofóbico e amplo nestas enzimas. Com isto, a cadeia lateral curta de Lys em P1 não teria um par para fazer ligação de hidrogênio e, portanto não se acomodaria tão bem ao subsítio S1 de proteases apresentando Ala¹⁹⁰. Entretanto, esta proposta pode ser questionada a partir dos resultados relatados no presente trabalho, mostrando que Lmr-47 com sua extrema especificidade por Arg em P1, apresenta Thr na posição 190, assim como a vasta maior parte das SVSPs seqüenciadas até o momento. O resíduo Thr¹⁹⁰ não só tornaria o subsítio S1 menos amplo como também poderia satisfazer os requerimentos de ligação de hidrogênio por Lys em P1. Desta forma, claramente o resíduo na posição 190 não pode ser responsabilizado pelo favorecimento de Arg em P1. Uma proposta mais coerente, porém ainda carecendo de uma fundamentação estrutural mais sólida e abrangente, foi feita por Vindigni (1997) implicando His¹⁸⁸, único em tPA entre diversas seqüências de enzimas tripsina-símile, no favorecimento de tPA por Arg sobre Lys em P1.

Os estudos de modelagem molecular apresentados no ANEXO 2 deste documento sugerem que a forma e a hidrofobicidade estão entre os fatores mais importantes na maximização da afinidade de ligação de pequenas moléculas direcionadas ao subsítio S1 de enzimas tripsina-símile. Assim, modificações estruturais que tornem o subsítio S1 mais apropriado ao encaixe ótimo da cadeia lateral de Arg e ao mesmo tempo maximizem a sua lipofilicidade podem ser a chave para a melhor compreensão da especificidade na bolsa de especificidade primária de enzimas tripsina-símiles. Análises comparativas envolvendo o modelo por homologia de Lmr-47 (Lmm-TLE), construído neste trabalho (vide ANEXO 2), as estruturas cristalográficas de calicreínas teciduais (BODE *et al.*, 1983; BAX *et al.*, 1997; TIMM, 1997) e mais recentemente as estruturas de diferentes ativadores de plasminogênio de venenos de serpentes (PARRY *et al.*, 1998; ZHU *et al.*, 2005) serão importantes para alcançar este objetivo.

5.1.1.d Uma possível aplicação biotecnológica para BJ-48 e Lmr-47?

Uma das enzimas arg-específicas encontradas na glândula submaxilar de camundongos é mGK9 (BLABER *et al.*, 1987), originalmente identificada como uma proteína ligante do fator de crescimento epidérmico ou EGF-BP (TAYLOR *et al.*, 1974). Esta proteinase é encontrada numa preparação comercial conhecida como Arg-C (Sigma código P6056), a qual é muito utilizada na proteólise limitada de proteínas visando o sequenciamento ou mapeamento peptídico, pois cliva preferencialmente ligações Arg-Xaa. Entretanto, tem sido demonstrado que Arg-C pode clivar proteínas em ligações Lys-Arg e Lys-Lys (BERTRAND *et al.*, 1989) e mesmo não clivar todas as ligações Arg-Xaa em um substrato protéico. Tendo em vista a excepcional especificidade do subsítio de BJ-48 e Lmr-47 é interessante considerar o potencial biotecnológico destas SVSPs como enzimas Arg-específicas para aplicações em estudos proteômicos. É provável que este uso seja dificultado pela elevada especificidade que as enzimas também apresentam contra substratos macromoleculares (vide abaixo). Contudo, futuramente, os elementos estruturais responsáveis por esta restrição no reconhecimento molecular poderão ser removidos através da clonagem dos genes destas SVSPs e técnicas de mutação sítio-dirigida.

5.1.1.e Cromatografia de afinidade baseada em benzamidina: uma técnica para a purificação seletiva de amidases Arg-específicas de venenos de serpentes?

Um aspecto interessante das SVSPs investigadas neste trabalho é que estas compartilham muitas propriedades, enzimáticas e estruturais (vide tópicos abaixo e ANEXOS). Outra característica em comum entre estas enzimas está no método de purificação pelo qual estas foram isoladas de seus venenos de origem, o qual consistiu fundamentalmente de uma única etapa de cromatografia de afinidade com derivados de benzamidina. Lmr-47 e BJ-48 interagem especificamente com matrizes cromatográficas diferindo na estrutura dos espaçadores inseridos no processo de acoplamento, mais provavelmente através de sítios hidrofóbicos distais à bolsa de especificidade primária (vide ANEXO 3). A partir desta constatação é possível questionar se o método de purificação tem algum papel na seleção de SVSPs de gêneros distintos, porém apresentando características comuns tão diversificadas como: espectros de atividade sobre ligantes macromoleculares, grau e tipo de glicosilação e estrutura em solução. Considerando que a característica comum nestas enzimas que mais chama a atenção é a

extraordinária especificidade por Arg sobre Lys em P1, é possível reduzir a questão postulada anteriormente para a indagação constante no título desta seção.

O primeiro passo para se formular uma resposta satisfatória é verificar a aplicabilidade do método como uma técnica de purificação de SVSPs específicas. Realmente, é impressionante o desempenho desta estratégia de purificação em selecionar apenas uma dentre muitas SVSPs encontradas nos venenos de serpentes nos quais a técnica já foi empregada. Por exemplo, além de BJ-48 no veneno de *B. jararacussu* e Lmr-47 no veneno de *L. m. rhombeata*, relatadas neste trabalho, batroxobin do veneno de *B. atrox* (HOLLEMAN & WEISS, 1976) e jararassin-I do veneno de *B. jararaca* (VIEIRA *et al.*, 2004), já foram purificadas em uma única etapa utilizando resinas cromatográficas baseadas em benzamidina como ligante. Em todos estes venenos, SVSPs distintas das que foram purificadas por este método foram descritas.

Por exemplo, ao menos três SVSPs no veneno de *B. atrox* [trombocitin (NIEWIAROWSKI *et al.*, 1979) e MSP 1 e 2 (SERRANO *et al.*, 1993)]; e oito no veneno de *B. jararaca* [BPA (MANDELBAUM & HENRIQUES, 1964), botrombin (NISHIDA *et al.*, 1994), PA-BJ (SERRANO *et al.*, 1995), KN-BJ 1 e 2 (SERRANO *et al.*, 1998), TL-BJ (SERRANO *et al.*, 2000), HS114 e HS120 (SAGUCHI *et al.*, 2005)] são conhecidas e estas enzimas não foram retidas nas resinas. Isto indica que as SVSPs ligadas são efetivamente selecionadas a partir de uma mistura de serinil proteases potencialmente capazes de também interagir com a benzamidina. De fato, algumas SVSPs dos venenos mencionados anteriormente podem ser purificadas por afinidade em resina de benzamidina após uma etapa cromatográfica preliminar, normalmente de troca iônica, que tem o efeito de retirar do veneno a enzima que se liga preferencialmente à coluna. Este é o caso de botrombin (WATANABE *et al.*, 2002) e jararacussin-I do veneno de *B. jararacussu* (BORTOLETO *et al.*, 2002).

O passo seguinte na formulação da resposta à pergunta no título é verificar se estas SVSPs são efetivamente enzimas Arg-específicas. De fato, BJ-48 e Lmr-47, conforme demonstrado neste trabalho, e batroxobin (BRAUD *et al.*, 2000b; YOU *et al.*, 2004) são capazes de distinguir Arg de Lys em P1. Entretanto, o número de enzimas purificadas e caracterizadas quanto à especificidade no subsítio S1 é ainda limitado, não permitindo uma resposta definitiva. De qualquer forma, a origem da possível

seletividade destas resinas por SVSPs Arg-específicas deve ser uma questão cinética. Isto porque não se espera que os valores de K_i destas enzimas contra um determinado derivado de benzamidina (vide ANEXO 2) variem numa faixa ampla o bastante para justificar valores de K_d suficientemente distintos contra o ligante imobilizado na matriz que poderiam responder por tamanha seletividade. Assim, é mais provável que a seleção ocorra por conta de um processo cinético onde a enzima com o subsítio S1 previamente melhor moldado para acomodar o grupo benzamidínico vença a competição para ligação na resina, ao dispensar um passo extra de rearranjo conformacional para receber o ligante.

5.1.2 Interação com ligantes estendidos: Ações fisiológicas de Lmr-47 e BJ-48

Conforme demonstrado anteriormente para Lmr-47 (AGUIAR *et al.*, 1996) e neste trabalho para BJ-48 (vide ANEXO 5), a cadeia A α do fibrinogênio humano é preferencialmente hidrolisada por estas enzimas gerando um coágulo de fibrina. Em BJ-48, foi demonstrado que a hidrólise é específica liberando fibrinopeptídeo A (FpA) e em menor proporção FpB. Com o intuito de aprofundar e estender o conhecimento sobre as funções *in vivo* de Lmr-47 e BJ-48, um painel de ligantes ativos sobre o sistema hemostático capazes de interagir com trombina de mamíferos foi testado (Tabela 11). Entre os três substratos testados, somente fibrinogênio foi clivado eficientemente, conforme acompanhado pela liberação de FpA, com especificidade cerca de 10 vezes inferior à trombina humana nas mesmas condições experimentais (VINDIGNI & DI CERA, 1996). A especificidade da reação de liberação de FpA com ambas as SVTLEs na presença de Na⁺ dobrou, um efeito insignificante frente ao aumento de 10 vezes que foi relatado para a trombina humana (VINDIGNI & DI CERA, 1996). De acordo com o paradigma estabelecido (vide ANEXO 4), este resultado era esperado, pois como é sabido para Lmr-47 (MAGALHÃES *et al.*, 1993) e conforme é esperado para BJ-48 e outras SVSPs, um resíduo de prolina ocupando a posição 225 nestas enzimas inviabiliza a sustentação de um sítio de ligação ao Na⁺ funcionalmente similar à trombina.

Apenas Lmr-47 foi capaz de hidrolisar PC na presença de TM, porém com especificidade 5 ordens de magnitude inferior à trombina humana (DANG & DI CERA, 1997). Desta forma, é possível antecipar que as SVSPs investigadas no presente trabalho sejam incapazes de ativar mecanismos anticoagulantes *in vivo* através da via da

PC. Lmr-47 e BJ-48 foram incapazes de hidrolisar PAR1, indicando que uma ativação de plaquetas, análoga à exercida pela trombina (BRASS, 2003), não é efetuada por estas enzimas *in vivo*. O sítios de hidrólise na cadeia A α do fibrinogênio, em PC e em PAR1 reconhecidos pela trombina apresentam as seguintes seqüências FLAEGGGVR↓GPRVVERH (BLOMBÄCK, 1969), NQGDQVDPR↓LIDGKMTRR (FOSTER & DAVIE, 1984), ATNATLDPR↓SFLLRNPND (VU *et al.*, 1991). A estrutura da trombina em complexo com um fragmento de FpA mostra que a partir do resíduo em P2 até a extremidade N-terminal deste substrato, uma volta para dentro do solvente é responsável por re-alocar o resíduo de Phe em P9 no lugar do resíduo que ocuparia a posição P4 (MARTIN *et al.*, 1992). O resíduo nesta posição interage com uma bolsa hidrofóbica na superfície da trombina. Na região equivalente em Lmr-47, destaca-se o resíduo Trp⁹⁹, o qual se propôs anteriormente, por modelagem molecular, ter um papel fundamental no reconhecimento de FpA por esta enzima (CASTRO *et al.*, 2001).

Os resíduos em P1 e P2 dos sítios de hidrólise em PAR1 e PC são os mesmos presentes no substrato tripeptídico D-FPR, o qual é clivado eficientemente por ambas as SVSPs (Tabela 10). Por outro lado, os resíduos de Leu ou Val em P4 destes substratos não deve representar um empecilho à hidrólise por BJ-48 ou Lmr-47, uma vez que os grupos volumosos e hidrofóbicos D-Phe (Tabela 10), N α -Tos (ANEXO 5), N α -Bz (AGUIAR *et al.*, 1996) e D-Val (MAGALHÃES *et al.*, 2003) são bem aceitos nesta posição. Um resíduo ácido em P3, como o que ocorre no sítio de hidrólise em PAR1 e PC pela trombina, nunca foi examinado nos substratos testados contra BJ-48 ou Lmr-47. Um rearranjo conformacional em Glu¹⁹² necessário à hidrólise eficiente destes substratos tem sido responsabilizado pelo intercâmbio entre as formas procoagulantes e anticoagulantes da trombina (revisado por DI CERA, 2004). O resíduo homólogo em Lmr-47 é Asn, mas o resíduo ocupando esta posição tem se mostrado extremamente variável em SVSPs. Resíduos hidrofóbicos como Phe e Leu, Gly, pequeno e conformacionalmente livre, e Pro, volumoso e conformacionalmente restrito, são encontrados na posição P3 de substratos clivados com níveis variados de eficiência (ANEXO 5; AGUIAR *et al.*, 1996; MAGALHÃES *et al.*, 2003). Em resumo, os dados disponíveis até o momento sobre a especificidade de substrato de Lmr-47 e BJ-48 ainda são limitados, dificultando uma interpretação mais detalhada das bases estruturais para o reconhecimento molecular na fenda de ligação ao substrato destas enzimas.

Não somente a fenda de ligação ao substrato está envolvida na interação com FpA e PAR1 na trombina (HUNTINGTON, 2005) e portanto, é interessante investigar se os exosítios implicados nestas interações na trombina também estão presentes nas SVSPs estudadas neste trabalho. Primeiramente, não foi possível detectar a ligação de hirudina à Lmr-47 ou BJ-48 (Tabela 11), indicando que os ABE-I's em trombina e nestas SVSPs apresentam composição distinta. Um total de 15 resíduos foram implicados por estudos de mutação sítio-dirigida na interação com fibrinogênio, dentre os quais K36, R67, H71, R73, Y76, R77a, K81, K109 e K110 pertencem a ABE-I (TSIANG *et al.*, 1995; HALL *et al.*, 2001). De fato, a maior parte dos resíduos nestas posições não está presente em Lmr-47 (vide ANEXO 2). A hidrólise de PAR1 pela trombina envolve a interação da seqüência ácida DKYEPFWEDEE no receptor com ABE-I, de maneira análoga à região C-terminal da hirudina (LIU *et al.*, 1991; VU *et al.*, 1991; AYALA *et al.*, 2001). Além disso, a interação de PAR1 tem como co-fator a glicoproteína integral da membrana de plaquetas, GpIb α , a qual se liga a ABE-II (LI *et al.*, 2001; DE CRISTOFARO *et al.*, 2001). Assim, em relação ao ABE-I de Lmr-47, era de se esperar a ausência de atividade desta enzima sobre PAR1.

Trombomodulina (TM) é outro ligante interagindo fortemente com a de trombina, cujo principal efeito é aumentar em três ordens de magnitude a hidrólise de PC pela trombina. TM se liga através dos domínios EGF5 e EGF6 ao ABE-I na enzima, através de uma ponte salina entre Lys¹¹⁰ na trombina e Asp⁴⁶¹ em TM e contatos hidrofóbicos distribuídos numa superfície de 900Å² (FUENTES-PRIOR *et al.*, 2000). A julgar pela virtual falta de atividade de ambas Lmr-47 e BJ-48 sobre PC na presença de TM é sensato propor que estas SVSPs são incapazes de interagir produtivamente com TM. A inibição de trombina por ATIII é acelerada em 1000 vezes pela interação do co-fator heparina (ou outros glicosaminoglicanos) com ABE-II, porém acredita-se que a reação não-catalisada tenha importância fisiológica (OLSON *et al.*, 1993). Desta forma, a ação de ATIII sobre as SVSPs alvo deste trabalho foi inicialmente avaliada na ausência do co-fator heparina, mostrando que nestas condições não há interação apreciável (Tabela 11). Este resultado tem grande importância fisiológica, pois mostra que Lmr-47 e BJ-48 são resistentes a ATIII, a qual é o principal inibidor circulante da ação da trombina. Trabalhos de modelagem molecular subsequentes utilizando o modelo de Lmr-47 poderão auxiliar na interpretação estrutural dos mecanismos de reconhecimento molecular empregados por esta SVTLE.

5.2 Estrutura em solução de Lmr-47

A estrutura em solução de Lmr-47 foi investigada por CD, fluorescência e SAXS. Os espectros de CD no UV-distante (190-260 nm) obtidos nos valores de pH 6,0 e 7,5 (**Fig. 105**) são compatíveis com a composição de estrutura secundária rica em folhas- β e com um elevado conteúdo de alças- β , esperada para membros da família da quimiotripsina (**LESK & FORDHAM, 1996**). Estes são nitidamente distintos do apresentado por BJ-48 em pH 8,0, o qual mostra um alto conteúdo de estruturas de conformação aleatória (vide ANEXO 5). Os espectros de CD Lmr-47 também se mostraram inalterados frente à variação de pH, indicando que a estrutura secundária de Lmr-47 é mantida mediante acidificação do pH ótimo para a catálise até pH 6,0. Estes resultados foram corroborados pela medida dos espectros de fluorescência intrínseca na faixa de pH variando de 5,0 até 8,0 (**Fig. 106**). Uma vez que cinco resíduos de triptofano estão distribuídos pela estrutura da enzima (**Fig. 107**) é provável que qualquer mudança conformacional que viesse a ser induzida pela variação de pH teria sido detectada pelo deslocamento do CME (ou nos valores de λ_{max}) de emissão de fluorescência por estes resíduos. A transição na I.F. que é observada na variação de pH (**Fig. 108**) é compatível com processos de RET (**LAKOWICS, 1999**) ocorrendo em virtude da ionização de resíduos de His em contato com parte dos resíduos de triptofano responsáveis pela fluorescência (**Fig. 107**).

Os dados de SAXS em pH 6,0, por outro lado, referem-se a uma partícula de dimensões maiores que a esperada para uma proteína globular de 45 kDa e que ao mesmo tempo apresenta regiões flexíveis em sua estrutura (**Fig. 109**). A forma média de Lmr-47 em solução assemelha-se, na maior parte do tempo, com uma estrutura estendida em forma de “Y” (**Fig. 110**). Este resultado é compatível com a análise por SAXS de BJ-48 no mesmo pH (ANEXO 6), mostrando que a proteína possui um formato de “U” de dimensões ($D_{\text{max}} = 110\text{\AA}$) e flexibilidade comparáveis a Lmr-47. Entretanto, as semelhanças estruturais entre Lmr-47 e BJ-48 cessam quando se analisa o efeito do pH sobre a estrutura desta última (ANEXO 6). Em pH 8,0, a estrutura de BJ-48 é mais rígida e apresenta forma prolata. Assim, diferente do que se conclui da análise por fluorescência de Lmr-47, BJ-48 parece sofrer uma mudança significativa na forma de sua molécula em soluções tamponadas em pH 6,0 ou 8,0.

Tendo em vista a contribuição inquestionável da extensiva glicosilação de BJ-48 para as suas propriedades em solução é possível propor que estes efeitos da variação de pH também sejam resultantes de alterações na flexibilidade das cadeias de carboidrato. Estudos utilizando técnicas de baixa resolução, como SAXS, são particularmente interessantes nesta fase ainda inicial em que se encontra a caracterização estrutural de SVSPs. A discussão acima mostra que há muito mais entre o enovelamento globular esperado para SVSPs como membros da família da quimiotripsina e as formas que estas enzimas realmente adotam em solução. Este é particularmente o caso de enzimas extensamente glicosiladas como BJ-48 e Lmr-47. Mais ainda, com estas glicoproteínas a técnica de SAXS pode ser a única disponível para o estudo das estruturas nativas, pois técnicas de alta resolução como cristalografia de raios-X e ressonância magnética nuclear podem não ser viáveis. A primeira devido à dificuldades inerentes à cristalização de proteínas contendo longas cadeias de carboidrato altamente flexíveis (WORMALD *et al.*, 2002; CHERNOV, 2003) e a segunda devido ao tamanho das moléculas acima dos limites práticos da técnica (WURTHRICH, 1990; SATTLER *et al.*, 1999).

5.3 Papel do carboidrato na estrutura e função de Lmr-47 e BJ-48

Conforme mostrado pela análise enzimática de Lmr-47 (seção 4.4.2) e BJ-48 (ANEXO 5), estas glicoproteínas são compostas por cerca de 38-42% de carboidratos, exclusivamente ligados através de ligações N-glicosídicas. Estas enzimas estão entre as SVSPs de mais alto nível de glicosilação, sendo superadas por poucos membros do grupo, como no caso excepcional de BPA, com 15% de O-glicanos e 47% de N-glicanos, totalizando 62% de carboidratos (MURAYAMA *et al.*, 2003). A análise da composição dos carboidratos mostrou estruturas semelhantes em Lmr-47 (seções 4.4.2 e 4.6.2) e BJ-48 (seção 4.6.2 e ANEXO 5), típicas de N-glicanos complexos (SPIRO, 2002), contendo entre 3 a 5 resíduos de ácido siálico ligados a um cerne fucosilado. Este último é composto por 4 resíduos de GlcNAc, 5 resíduos de GalNAc e 3 resíduos de Man em BJ-48, diferindo de Lmr-47 apenas pelo conteúdo significativamente menor de GalNAc (1 resíduo por cerne).

Uma análise detalhada da estrutura de carboidratos em batroxobin foi conduzida anteriormente e revelou que ambos os sítios de glicosilação na proteína são ocupados

por carboidratos heterogêneos compostos predominantemente de glicanos diantennados do tipo complexo, parcialmente incompletos, além de espécies do tipo híbrido (LOCHNIT & GEYER, 1995). Acredita-se que em Lmr-47, os 4 sítios potenciais de N-glicosilação (N59, N95, N164 e N245) estejam ocupados, baseado nos ciclos em branco obtidos durante a degradação de Edman e pela análise de aminoácidos que indicou resíduos Asx (MAGALHÃES *et al.*, 1993). Em concordância geral com a estrutura revelada neste trabalho para Lmr-47 e BJ-48, foi demonstrado que a maior parte dos carboidratos em batroxobin são fucosilados em C6 do resíduo de GlcNAc mais interno do cerne enquanto que parte dos resíduos de GlcNAc de antenas do tipo complexa são fucosilados em C3. Mais ainda, de acordo com o alto conteúdo de GalNAc em BJ-48, foi relatado que a maior parte dos resíduos de Gal em batroxobin são substituídos por GalNAc.

Como os estudos de SAXS deixaram claro, a contribuição dos carboidratos para a estrutura de Lmr-47 e BJ-48 e, conseqüentemente, para a função destas enzimas, não pode ser ignorado. Desta forma, ensaios foram conduzidos almejando estabelecer uma metodologia eficiente para a deglicosilação destas SVSPs em condições não-desnaturantes utilizando a enzima PNGase F. A deglicosilação foi somente conseguida em sua completude em condições desnaturantes, indicando que os sítios de N-glicosilação nestas SVSPs são relativamente inacessíveis à glicosidase. Dificuldades semelhantes foram recentemente relatadas na deglicosilação de BPA nativa (MURAYAMA *et al.*, 2003). Contudo, em Lmr-47 e BJ-48 este acesso pode ser grandemente facilitado pela remoção de resíduos de ácido siálico da estrutura do tipo complexa, mais provavelmente tri ou tetraantennada (Fig. 112).

Durante os experimentos de deglicosilação nativa também ficou claro o papel da porção carboidrato em Lmr-47 e BJ-48 na manutenção da solubilidade destas proteínas em concentrações variando de 1 a 3 mg/mL, pois em ensaios preliminares sem adição de CHAPS 1% foi visto que o progresso da deglicosilação por PNGase F induzia a formação de agregados insolúveis e sujeitos à autoproteólise (Fig. 113). O mesmo foi observado por Murayama *et al.* com BPA, sugerindo que os sítios de glicosilação nestas SVSPs são implicados na proteção da agregação por interação hidrofóbica. De fato, uma análise estatística da ocupação de sítios de N-glicosilação em proteínas de estrutura 3D conhecida mostrou que a maior parte destes sítios ocorre em regiões hidrofóbicas parcialmente acessíveis ao solvente (PRETRESCU *et al.*, 2004). Isto é compatível com

o reconhecimento cada vez maior de um papel da N-glicosilação na estabilização tanto de cadeias polipeptídicas nascentes no retículo endoplasmático rugoso como de proteínas enoveladas em sua conformação definitiva (REUTER & GABIUS, 1999; IMPERIALI & O'CONNOR, 1999; ROTH, 2002; SPIRO, 2002). Finalmente, conforme esperado, uma separação cromatográfica satisfatória das formas parcialmente deglicosiladas de BJ-48 e Lmr-47 foi conseguida utilizando interação hidrofóbica como princípio físico-químico (Fig. 115).

Com o intuito de melhor caracterizar a função dos carboidratos em SVSPs, as propriedades de duas frações parcialmente deglicosiladas e outra totalmente deglicosiladas de BJ-48 foram analisadas quanto à estabilidade térmica e susceptibilidade de inibição por SBTI, um inibidor contra o qual BJ-48 foi demonstrada ser resistente (ANEXO 5). Assim como demonstrado neste trabalho para BJ-48 (ANEXOS 5 e 6) e Lmr-47 (Fig. 102), muitas outras SVSPs extensivamente glicosiladas, tais como BPA (MURAYAMA *et al.*, 2003) e BJ-FCE (ZAGANELLI *et al.*, 1996) apresentam máximo de atividade amidásica em temperaturas superiores a 45°C e possuem elevada resistência à desnaturação pelo calor. Ao passo que as seis pontes dissulfeto conservadas em SVSPs (SERRANO & MAROUN, 2005) podem ter algum papel nesta estabilidade, este não deve ser crucial, uma vez que outras enzimas tripsina-símile como a tripsina bovina e caliceínas teciduais também apresentam padrão semelhante de resíduos de cistina e não apresentam a mesma estabilidade térmica (SIPOS & MERCKEL, 1970). Os dados obtidos neste trabalho para formas de BJ-48 com diferentes níveis de glicosilação (Tabela 12) mostram, até onde se sabe, pela primeira vez, que a porção de carboidratos está efetivamente associada à estabilidade térmica em uma SVSP. Anteriormente, Murayama *et al.* mostraram através de espectrofluorimetria e ensaios cinéticos que em BPA a glicosilação afeta a estabilidade geral da estrutura e a resistência a extremos de pH. Esta diminuição na estabilidade também está provavelmente associada à redução na eficiência de hidrólise de substratos cromogênicos curtos observada por estes autores, conforme refletido na redução em k_{cat} .

Neste trabalho, uma segunda função foi demonstrada para os carboidratos em BJ-48. Esta consiste no bloqueio do acesso de SBTI, um inibidor protéico, ao sítio ativo da enzima, conforme demonstrado pelo aumento na susceptibilidade de inibição aparentemente proporcional ao grau de deglicosilação da enzima (Tabela 13). Estes

resultados confirmam e estendem as recentes análises estruturais de [Zhu et al. \(2005\)](#) com o ativador de plasminogênio do veneno de *A. actus* (AaV-SP), que sugeriram um papel para o sítio de N-glicosilação em Asn³⁵ no bloqueio da ligação de SBTI a esta SVSP. O resíduo de GlcNAc fucosilado em C6 ligado à Asn³⁵, projeta-se na região entre as alças 30 e 60, fazendo ligações de hidrogênio com His⁵⁷ na tríade catalítica e Asp⁵⁹. Além disso, parte do subsítio S1' é efetivamente ocupado pelo resíduo de fucose. Este mecanismo é fundamentalmente diferentemente do que é empregado por TSV-PA, cuja resistência à SBTI está associada à presença do resíduo volumoso de fenilalanina na posição 192 ([BRAUD et al., 1992](#)), uma vez que AaV-SP apresenta um resíduo de glicina nesta posição. Os dados aqui relatados para BJ-48 demonstram que o mecanismo de resistência ao SBTI proposto por [Zhu et al.](#), meramente baseado na análise estrutural, é efetivamente viável em SVSPs e indica que BJ-48 guarda importante semelhança estrutural com AaV-SP. Esta informação será muito útil na construção de futuros modelos teóricos, assim que a seqüência completa de BJ-48 estiver disponível.

A discussão acima mostra que a glicosilação em SVSPs pode estar implicada não somente na estabilidade estrutural intrínseca ou na estabilidade mostrada frente à agressores ambientais como temperatura e extremos de pH, mas também pode controlar o acesso de substratos e inibidores de importância fisiológica ao sítio ativo destas enzimas. A continuidade destes estudos poderá revelar ainda outros papéis biológicos para os carboidratos em SVSPs relevantes para o envenenamento, como a estabilidade plasmática e propriedades antigênicas, estas últimas já demonstradas para outras glicoproteínas ([LISOWSKA, 2002](#)). Por fim, a dependência da N-glicosilação para a atividade proteolítica de enzimas de diferentes classes vem sendo demonstrada em trabalhos recentes ([YOSHIMAZU et al., 2004](#); [BARINKA et al., 2004](#)), sugerindo que a análise de carboidratos precisa ser adicionada como uma nova dimensão no estudo da estrutura e função de proteases.

6 CONCLUSÕES

- O modelo teórico da estrutura 3D de Lmr-47 (ou Lmm-TLE) mostrou estereoquímica satisfatória e outras propriedades estruturais condizentes com o conhecimento disponível sobre a estrutura e função de SVTLEs, evidenciando a qualidade do modelo construído. De acordo com os estudos de modelagem molecular, as posições 190, 192, 213 e 218 são candidatas a um papel no controle da afinidade de ligação ao subsítio S1 em SVTLEs. As análises também mostraram que esta afinidade é ditada pela lipofilicidade e forma da bolsa de especificidade primária em enzimas tripsina-simile, em oposição à idéia de que interações eletrostáticas e polares direcionais seriam responsáveis por este processo.
- Cromatografia de afinidade em resinas a base de derivados de benzamidina é um método eficiente para a purificação, em uma única etapa, de fibrinogenases de venenos de serpentes. A estrutura de diferentes espaçadores, usados na imobilização de p-aminobenzamidina ao suporte inerte, é responsável pela seletividade das resinas por SVSPs de diferentes venenos. A caracterização cinética das SVSPs retidas nestas colunas indica que estas podem ser pertencentes a um grupo específico de serinil proteases específicas para arginina em P1.
- A enzima de 48-52 kDa purificada por cromatografia de afinidade a partir do veneno de *B. jararacussu*, BJ-48, mostrou ser uma serinil protease com atividade coagulante (SVTLE), capaz de liberar FpA mais rapidamente que FpB do fibrinogênio humano. Assim como demonstrado para Lmr-47, BJ-48 é altamente glicosilada, onde carboidratos N-ligados respondem por 40% da massa molecular.
- BJ-48 e Lmr-47 apresentam extraordinária especificidade por Arg sobre Lys em P1 de substratos sintéticos tripeptídicos. Esta especificidade pode ser remanescente de uma origem evolutiva comum com arginil amidases glandulares. A interação destas SVSPs com diversos componentes ativos sobre o sistema hemostático mostrou um espectro de interações moleculares bastante seletivo, onde somente o fibrinogênio é eficientemente hidrolisado entre o

conjunto de substratos testados incluindo PAR1 e PC. BJ-48 e Lmr-47 também são resistentes a inibição por ATIII, uma observação de evidente importância fisiológica para a ação tóxica destas enzimas. Esta seletividade no reconhecimento molecular provavelmente advém não somente de interações improdutivas com resíduos na fenda de ligação ao substrato em BJ-48 e Lmr-47, mas também de estruturas alternativas de exosítios nestas enzimas.

- A estrutura de carboidratos de BJ-48 e Lmr-47 é semelhante à determinada para batroxobin, sendo composta por cadeias complexas tri ou tetraantenas ramificadas e fucosiladas. Estes carboidratos são importantes para a manutenção da solubilidade das enzimas ao impedirem a agregação por contatos hidrofóbicos, assim como estão implicados na elevada estabilidade térmica de BJ-48 e possivelmente de outras SVSPs. A presença de carboidratos N-ligados em BJ-48 também é responsável pela resistência a SBTI e possivelmente outros inibidores protéicos de serinil proteinases.
- As diferentes técnicas biofísicas de baixa resolução forneceram um quadro em que Lmr-47 aparece como uma proteína com estrutura secundária típica de outros membros da família da quimiotripsina, relativamente resistente à variações no pH, cujas dimensões e forma são significativamente afetadas pelas cadeias de carboidrato flexíveis em sua superfície. Estudos de SAXS em BJ-48 indicam que a forma da proteína se altera mediante mudança do pH de 6,0 para 8,0. Ao comparar os dados estruturais para BJ-48 e Lmr-47, duas SVSPs compartilhando diversas propriedades bioquímicas, é possível propor que estas alterações sejam devidas não à cadeia polipeptídica, mas sim à estrutura de carboidratos nestas enzimas.
- Estudos de modelagem molecular sobre o canal de sódio em serinil proteases da coagulação revelaram importantes diferenças topográficas que tornam trombina única entre as enzimas sensíveis à regulação alostérica pelo íon Na^+ . Os resultados da comparação das superfícies moleculares nas formas lenta e rápida da trombina mostram que alterações conformacionais em F34 e W141, até então desapercibidas pelas análises cristalográficas, podem transmitir o evento de ligação de Na^+ pode afetar interações com substratos fisiológicos em ABE-I de trombina, mais de 20 Å distantes do sítio de ligação ao cátion monovalente.

REFERÊNCIAS

- ADAMS, M.D.; DUBNICK, M.; KERLAVAGE, A.R. *et al.* Sequence identification of 2,375 human brain genes. *Nature*, v. 355, p. 632-634 (1992).
- AGUIAR, A.S.; ALVES, C.R.; MELGAREJO, A. *et al.* Purification and partial characterization of a thrombin-like/gyroxin enzyme from bushmaster (*Lachesis muta rhombeata*) venom. *Toxicon*, v. 34, p. 555-565 (1996).
- ALTSCHUL, S.F.; GISH, W.; MILLER, W. *et al.* Basic local alignment search tool. *J. Mol. Biol.*, v. 215, p. 403-410 (1990).
- ANDREEVA, N.S.; RUMSH, I.D. Analysis of crystal structures of aspartic proteinases: on the role of amino acid residues adjacent to the catalytic site of pepsin-like enzymes. *Protein Sci.*, v. 10, p. 2439 (2001).
- APOSTOLUK, W.; OTLEWSKI, J. Variability of the Canonical Loop Conformations in Serine Proteinases Inhibitors and Other Proteins. *PROTEINS: Struct. Funct. Gen.*, v. 32, p. 459-474 (1998).
- ARAGON-ORTIZ, F. Isolation and partial characterization of *Lachesis muta melanocephala* coagulant proteinase: biochemical parameters of the venom. *Rev. Biol. Trop.*, v. 36, p. 387-392 (1988).
- ARAGON-ORTIZ, F. Purification and properties of a coagulant proteinase isolated from bushmaster (*Lachesis muta*) venom (Serpentes: Viperidae). *Rev. Biol. Trop.*, v. 34, p. 55-58 (1986).
- ARAGON-ORTIZ, F.; GUBENSEK, F. A thrombin-like enzyme from bushmaster (*Lachesis muta stenophrys*) venom. *Toxicon*, v. 31, p. 1435-43 (1993).
- ARMSTRONG, N.; DE LENCASTRE, A.; GOUAUX, E. A new protein folding screen: application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase. *Prot. Sci.*, v. 8, p. 1475-1483 (1999).
- AROSIO, D.; AYALA, Y.M.; DI CERA, E. Mutation of W215 Compromises Thrombin Cleavage of Fibrinogen, but Not of PAR-1 or Protein C. *Biochemistry*, v. 39, p. 8095-8101 (2000).
- AUTIN, L.; MITEVA, M.A.; LEE, W.H. *et al.* Molecular models of the procoagulant factor VIIIa-Factor IXa complex. *J. Thromb. Haemost.*, v. 3, p. 2044-2056 (2005).

AYALA, Y.H.; DI CERA, E. Molecular recognition by thrombin - role of the slow-fast transition, site-specific ion-binding energetics and thermodynamic mapping of structural components. *J. Mol. Biol.*, v. 235, p. 733-746 (1994).

AYALA, Y.M.; CANTWELL, A.M.; ROSE T. *et al.* Molecular Mapping of Thrombin-Receptor Interactions. *PROTEINS: Struct. Funct. Genet.*, v. 45, p. 107-116 (2001).

AZUMA, T.; PALS, G.; MOHANDAS, T.K. *et al.* Human gastric cathepsin E: predicted sequence, localization to chromosome 1 and sequence homology with other aspartic proteinase. *J. Biol. Chem.*, v. 264, p. 16748-16753 (1989).

BAJAJ, S.P.; SCHMIDT, A.E.; MATHUR, A. *et al.* Factor IXa:factor VIIIa interaction. Helix 330-338 of factor IXa interacts with residues 558-565 and spatially adjacent regions of the A2 subunit of factor VIIIa. *J. Biol. Chem.*, v. 276, p. 16302-16309 (2001).

BANEYX, F.; MUJACIC, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.*, v. 22, p. 1399-1408 (2004).

BANNER, D.W.; D'ARCY, A.; CHENE, C. *et al.* The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature*, v. 380, p. 41-46 (1996).

BARINKA, C.; SACHA, P.; SKLENAR, J. *et al.* Identification of the N-glycosylation sites on glutamate carboxypeptidase II necessary for proteolytic activity. *Prot. Sci.*, v. 13, p. 1627-1635 (2004).

BAUER, K.A. New Pentasaccharides for Prophylaxis of Deep Vein Thrombosis. *Pharmacology. CHEST*, v. 124, p. 124-130 (2003).

BAX, B.; BLUNDELL, T.L.; MURRAY-RUST, J. *et al.* Structure of mouse 7S NGF: a complex of nerve growth factor with four binding proteins. *Structure*, v. 5, p. 1275-1285 (1997)

BECKER, M.M.; HARROP, S.A.; DALTON, J.P. *et al.* Cloning and characterization of *Schistosoma japonicum* aspartic proteinase involved in hemoglobin digestion. *J. Biol. Chem.*, v. 270, p. 24496-24501(1995).

BEGGAH, S.; LECHENNE, B.; REICHARD, U. *et al.* Intra- and intermolecular events direct the propeptide-mediated maturation of the *Candida albicans* secreted aspartic proteinase Sap1p. *Microbiology-UK*, v. 146, p. 2765-2773 (2000).

- BENNION, B.J.; DAGGETT, V. Counteraction of urea-induced protein denaturation by trimethylamine N-oxide: A chemical chaperone at atomic resolution. *Proc. Natl. Acad. Sci. USA*, v. 101, p. 6433–6438 (2004).
- BENSON, D.A.; KARSCH-MIZRACHI, I.; LIPMAN, D.J. *et al.* Genbank. *Nucleic Acids Res.*, v. 33, p. D34-38 (2005).
- BERG, D.T.; WILEY, M.R.; GRINNELL, B.W. Enhanced protein C activation and inhibition of fibrinogen cleavage by a thrombin modulator. *Science*, v. 273, p. 1389-1391 (1996).
- BERMAN, H.M.; WESTBROOK, J.; FENG, Z. *et al.* The Protein Data Bank. *Nucleic Acids Res.*, v. 28, p. 235-242 (2000).
- BERRY, C.; DAME, J.B.; DUNN, B.M. *et al.* Aspartic proteinases from the human malarial parasite *Plasmodium falciparum*. *Biochem. Soc. Trans.*, v. 22, p. 305S (1994).
- BERTRAND, R.; DERANCOURT, J.; KASSAB, R. Selective cleavage at lysine of the 50 kDa-20 kDa connector loop segment of skeletal myosin S-1 by endoproteinase Arg-C. *FEBS Lett.*, v. 246, p. 171-176 (1989).
- BEYER, B.M.; DUNN, B.M. Prime region subsite specificity characterization of human cathepsin D: the dominant role of position 128. *Protein Sci.*, v. 7, p. 88-95 (1998).
- BEYER, B.M.; DUNN, B.M. Self-activation of recombinant human lysosomal procathepsin D at a newly engineered cleavage junction, "short" pseudocathepsin D. *J. Biol. Chem.*, v. 271, p. 15590-15596 (1996).
- BIANCHINI, E.P.; LOUVAIN, V.B.; MARQUE, P-E. *et al.* Mapping of the catalytic groove preferences of factor Xa reveals an inadequate selectivity for its macromolecule substrates. *J. Biol. Chem.*, v. 277, p. 20527-20534 (2002).
- BLABER, M.; ISACKSON, P.J.; MARSTERS-JR, J.C. *et al.* Substrate Specificities of Growth Factor Associated Kallikreins of the Mouse Submandibular Gland. *Biochemistry*, v. 28, p. 7813-7819 (1989).
- BLABER, M.; ISACKSON, P.J.; BRADSHAW, R.A. A Complete cDNA Sequence for the Major Epidermal Growth Factor Binding Protein in the Male Mouse Submandibular Gland. *Biochemistry*, v. 26, p. 6742-6749 (1987).

- BLOKZIJL, W.; ENGBERTS, J.B.F.N. Hydrophobic effects - opinions and facts. *Ang. Chem.-Int. Ed. Eng.*, v. 32, p. 1545-1579 (1993).
- BLOMBÄCK, G. (1969) The N-terminal disulfide knot of human fibrinogen. *Br. J. Haematol.*, v. 17, p. 145-151 (1969).
- BLUNDELL, T.; SIBANDA, L.B.; PEARL, L. Three-dimensional structure, specificity and catalytic mechanism of renin. *Nature*, v. 304, p. 273-275 (1983).
- BOBOFCHACK, K. M.; PINEDA, A. O.; MATHEWS, S. *et al.* Energetic and structural consequences of perturbing Gly-193 in the oxyanion hole of serine proteases. *J. Biol. Chem.*, v. 280, p. 25644-25650 (2005).
- BODE, W.; HUBER, R. Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.*, v. 204, p. 433-451 (1992).
- BODE, W.; CHEN, Z.; BARTELS, K. *et al.* Refined 2Å X-ray crystal structure of porcine pancreatic kallikrein A, a specific trypsin-like serine proteinase: crystallization, structure determination, crystallographic refinement, structure and its comparison with bovine trypsin. *J. Mol. Biol.*, v. 164, p. 237-282 (1983).
- BODE, W.; HUBER, R. Structural basis of the endoproteinase-protein inhibitor interaction. *Biochim. Biophys. Acta*, v. 1477, p. 241-252 (2000).
- BODE, W.; MAYR, I.; BAUMANN, U. *et al.* The refined 1.9 Å crystal structure of human α -thrombin: interaction with D-Phe-Pro-Argchloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.*, v. 8, p. 3467-3475 (1989).
- BODE, W.; SCHWAGER, P. Refined crystal-structure of bovine beta-trypsin at 1.8Å resolution .2. Crystallographic refinement, calcium-binding site, benzamidine binding-site and active-site at pH 7.0. *J. Mol. Biol.*, v. 98, p. 693-717 (1975).
- BOGITSH, B.J. observations on digestion in schistosomes or blood and guts. *Trans. Am. Microsc. Soc.*, v. 108, p. 1-5 (1989).
- BOGITSH, B.J., KIRSCHNER, K.F., ROTMANS, J.P. *Schistosoma japonicum*: immunoinhibitory studies on hemoglobin using heterologous antiserum to bovine cathepsin D. *J. Parasitol.*, v. 78, p. 454-459 (1992).

- BOGITSH, B.J.; KIRSCHNER, K.F. *Schistosoma japonicum*: ultrastructural localization of a hemoglobinase using mercury labeled pepstatin. *Exp. Parasitol.*, v. 62, p. 207-215 (1986).
- BORTOLETO, R.K.; MURAKAMI, M.T.; WATANABE, L. *et al.* Purification, characterization and crystallization of Jararacussin-I, a fibrinogen-clotting enzyme isolated from the venom of *Bothrops jararacussu*. *Toxicon*, v. 40, p. 1307-1312 (2002).
- BOSOMPEM, K.M.; OWUSU, O.; OKANLA, E.O.; *et al.* Applicability of a monoclonal antibody-based dipstick in diagnosis of urinary schistosomiasis in the Central Region of Ghana. *Trop. Med. Int. Health*, v. 9, p. 991-996 (2004).
- BRADFORD, M. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* v. 72, p. 248-252 (1976).
- BRADY, C.P.; BRINKWORTH, R.I.; DALTON, J.P. *et al.* Molecular modeling and substrate specificity of discrete cruzipain-like and cathepsin L-like cysteine proteinases of the human blood fluke *Schistosoma mansoni*. *Arch. Biochem. Biophys.*, v. 380, p. 46-55 (2000).
- BRASS, L.F. Thrombin and platelet activation. *CHEST*, v. 124, p. 18S-25S (2003).
- BRAUD, S.; BON, C.; WISNER, A. Snake venom proteins acting on hemostasis. *Biochimie*, v. 82, p. 851-859 (2000a).
- BRAUD, S.; PARRY, M.A.; MAROUN, R. *et al.* The contribution of residues 192 and 193 to the specificity of snake venom serine proteinases. *J. Biol. Chem.*, v. 275, p. 1823-1828 (2000b).
- BRETNER, M. Existing and future therapeutic options for hepatitis C virus infection. *Acta Biochim. Pol.*, v. 52, p. 57-70 (2005).
- BREWER, C.F.; BHATTACHARYYA, L. Concanavalin-A interactions with asparagine-linked glycopeptides - the mechanisms of binding of oligomannose, bisected hybrid, and complex type carbohydrates. *Glycoconj. J.*, v. 5, p. 159-173 (1988).
- BRIK, A.; WONG, C-H. HIV-1 protease: mechanism and drug discovery. *Org. Biomol. Chem.*, v. 1, p. 5-14 (2003).

BRINDLEY, P.J. Drug resistance to schistosomicides and other anthelmintics of medical significance. *Acta Trop.*, v. 56, p. 213-231 (1994).

BRINDLEY, P.J.; KALINNA, B.H.; WONG, J.Y.M. *et al.* Proteolysis of human hemoglobin by schistosome cathepsin D. *Mol. Biochem. Parasitol.*, v. 112, p. 103-112 (2001).

BRINDLEY, P.J.; KALINNA, B.H.; DALTON, J.P. *et al.* Proteolytic degradation of host hemoglobin by schistosomes. *Mol. Biochem. Parasitol.*, v. 89, p. 1-9 (1997).

BRINKWORTH, R.I.; PROCIV, P.; LOUKAS, A. *et al.* Hemoglobin-degrading, aspartic proteases of blood-feeding parasites: substrate specificity revealed by homology models. *J. Biol. Chem.*, v. 276, p. 38844-38851 (2001).

BROZE-JR, G.J. Tissue factor pathway inhibitor and the revised theory of coagulation. *Annu. Rev. Med.*, v. 46, p. 103-112 (1995).

BRYAN, P.N. Prodomains and Protein folding catalysis. *Chem. Rev.*, v. 102, p. 4805-4815 (2002).

BUTENAS, S.; ORFEO, T.; GISSEL, M.T. *et al.* The significance of circulating factor IXa in blood. *J. Biol. Chem.*, v. 279, p. 22875-22882 (2004).

BUTLER, R.; MICHEL, A.; KUNZ, W. *et al.* Sequence of *Schistosoma mansoni* cathepsin C and its structural comparison with papain and cathepsins B and L of the parasite. *Prot. Pept. Lett.*, v. 2; p. 313-320 (1995).

CAFFREY, C.R.; MATHIEU, M.A.; GAFFNEY, A.M. *et al.* Identification of a cDNA encoding an active asparagynil endopeptidase of *Schistosoma mansoni* and its expression in *Pichia pastoris*. *FEBS Lett.*, v. 466, p. 244-248 (2000).

CAFFREY, C.R.; MCKERROW, J.H.; SALTER, J.P. *et al.* Blood 'n' guts: an update on schistosome digestive peptidases. *Trends Parasitol.*, v. 20, p. 241-248 (2004).

CAMIRE, R.M. Prothrombinase assembly and S1 site occupation restore the catalytic activity of FXa impaired by mutation at the sodium-binding site. *J. Biol. Chem.*, v. 277, p. 37863-37870 (2002).

CARROLL, C.D.; JOHNSON, T.O.; TAO, S. *et al.* Evaluation of a structure-based statine cyclic diamino amide encoded combinatorial library against plasmepsin II and cathepsin D. *Bioorg. Med. Chem. Lett.* v. 8, p. 3203-3206 (1998).

CARTER, W.J.; MYLES, T. GIBBS, C.S. *et al.* Crystal structure of anticoagulant thrombin variant E217K provides insights into thrombin allostery. *J. Biol. Chem.*, v. 279, p. 26387-26394 (2004).

CASTRO, H.C.; SILVA, D.M.; CRAIK, C. *et al.* Structural features of a snake venom thrombin-like enzyme: thrombin and trypsin on a single catalytic platform? *Biochim. Biophys. Acta*, v. 1547, p. 183-195 (2001).

CASTRO, H.C.; ZINGALLI, R.B.; ALBUQUERQUE, M.G. *et al.* Snake venom thrombin-like enzymes: from reptilase to now. *Cell. Mol. Life. Sci.*, v. 61, p. 843-856 (2004).

CASTRONUOVO, G.; ELIA, V.; POSTIGLIONE, C. *et al.* Interactions of aminoacids in concentrated aqueous solutions of urea or ethanol. Implications for the mechanism of protein denaturation. *Thermochim. Acta*, v. 339, p. 11-19 (1999).

CEREGHINO, J.L.; CREGG, J.M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.*, v. 24, p. 45-66 (2000).

CESARI, I.M.; VALDIVIESO, E.; SCHRÉVEL, J. Biochemical characterization of cathepsin D from adult *Schistosoma mansoni* worms. *Mem. Inst. Oswaldo Cruz*, v. 93, p. S165-168 (1998).

CHAPPELL, C.L.; DRESDEN, M.H. *Schistosoma mansoni*: protease activity of "hemoglobinase" from the digestive tract of adult worms. *Exp. Parasitol.*, v. 61, p. 160-167 (1986).

CHECA, A.; ORTIZ, A.R.; PASCUAL-TEREZA, B. *et al.* Assessment of Solvation Effects on Calculated Binding Affinity Differences: Trypsin Inhibition by Flavonoids as a Model System for Congeneric Series. *J. Med. Chem.*, v. 40, p. 4136-4145 (1997).

CHEEVER, A.W.; WELLER, T.H. Observations on the growth and nutritional requirements of *Schistosoma mansoni* *in vitro*. *Am. J. Hyg.*, v. 68, p. 322-329 (1958).

CHEN, K.C.S.; TANG, J. Amino acid sequence around the epoxide-reactive residues in pepsin. *J. Biol. Chem.*, v. 247, p. 2566-2574 (1972).

CHEN, L.; MANITHODY, C.; YANG, L. *et al.* Zymogenic and enzymatic properties of the 70-80 loop mutants of factor X/Xa. *Prot. Sci.*, v. 13, p. 431-442 (2004).

- CHERNOV, A.A. Protein crystals and their growth. *J. Struct. Biol.*, v. 142, p. 3–21 (2003).
- CHITSULO, L.; ENGELS, D.; MONTRESOR, A. *et al.* The global status of schistosomiasis and its control. *Acta Trop.*, v. 77, p. 41-51 (2000).
- CHITSULO, L.; LOVERDE, P.; ENGELS, D. Schistosomiasis. *Nat. Rev. Microbiol.*, v. 2, p. 12-3 (2004).
- CHOMCZYNSKI, P.; SACCHI, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, v. 162, p. 156-159 (1987).
- CIOLI, D. Chemotherapy of schistosomiasis: an update. *Parasitol. Today*, v. 14, p. 418-422 (1998).
- CIOLI, D. Praziquantel: is there real resistance and are there alternatives? *Curr. Opin. Infect. Dis.*, v. 13, p. 659-663 (2000).
- CIOLI, D.; PICA-MATTOCCIA, L.; ARCHER, S. Antischistosomal drugs: Past, present... and future? *Pharmac. Ther.*, v. 68, p. 35-85 (1995).
- COATES, L.; ERSKINE, P.T.; WOOD, S.P. *et al.* A neutron Laue diffraction study of endothiapepsin: implications for the aspartic proteinase mechanism. *Biochemistry*, v. 40, p. 13159 (2001).
- COMBET, C.; BLANCHET, C.; GEOURJON, C. *et al.* NPS@: Network Protein Sequence Analysis. *Trends Biochem. Sci.*, v. 25, p. 147-150 (2000).
- COOMBS, G. S.; RAO, M. S.; OLSON, A. J. *et al.* Revisiting catalysis by chymotrypsin family serine proteases using peptide substrates and inhibitors with unnatural main chains. *J. Biol. Chem.*, v. 274, p. 24074-24079 (1999).
- CUFF, J.A.; CLAMP, M.E.; SIDDIQUI, A.S. *et al.* Jpred: A Consensus Secondary Structure Prediction Server, *Bioinformatics*, v. 14, p. 892-893 (1998).
- CURTIS, L.D.; BROWN, A.; COMER, M.B. *et al.* Pharmacokinetics and pharmacodynamics of BB-10153, a thrombin-activatable plasminogen, in healthy volunteers. *J. Thromb. Haemost.*, v. 3, p. 1180-1186 (2005).

- CURWEN, R.S.; ASHTON, P.D.; JOHNSTON, D.A. *et al.* The *Schistosoma mansoni* soluble proteome: a comparison across four life-cycle stages. *Mol. Biochem. Parasitol.*, v. 138, p. 57-66 (2004).
- CZAPINSKA, H.; OTLEWSKI, J. Structural and energetic determinants of the S1-site specificity in serine proteases. *Eur. J. Biochem.*, v. 260, p. 571-595 (1999).
- DA SILVA, N.J.; AIRD, S.D.; SEEBART, C. *et al.* A gyroxin analog from the venom of the bushmaster (*Lachesis muta muta*). *Toxicon*, v. 27, p.763-771 (1989).
- DAHLBÄCK, B. Blood coagulation and its regulation by anticoagulant pathways: genetic pathogenesis of bleeding and thrombotic diseases. *J. Int. Med.*, v. 257, p. 209-223 (2005).
- DALTON, J.P.; CLOUGH, K.A.; JONES, M.K. *et al.* Characterization of the cathepsin-like cysteine proteinases of *Schistosoma mansoni*. *Infect. Immun.*, v. 64, p. 1328 (1996).
- DALTON, J.P.; HOLA-JAMRISKA, L.; BRINDLEY, P.J. Asparaginyl endopeptidase activity in adult *Schistosoma mansoni*. *Parasitology*, v. 111, p. 575-580 (1995a).
- DALTON, J.P.; SMITH, A.M.; CLOUGH, K.A. *et al.* Digestion of haemoglobin by schistosomes: 35 years on. *Parasitol. Today*, v. 11, 299-303 (1995b).
- DALY, R.; HEARN, M.T.W. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J. Mol. Recognit.*, v. 18, p. 119-138 (2005).
- DANG, Q.D.; DI CERA, E. Residue 225 determines the Na⁺-induced allosteric regulation of catalytic activity in serine proteases. *Proc. Natl. Acad. USA*, v. 93, p. 10653-10656 (1996).
- DANG, Q.D.; GUINTO, E.R.; DI CERA, E. Rational engineering of activity and specificity in a serine protease. *Nat. Biotech.*, v. 15, p. 146-149 (1997).
- DANG, Q.D.; DI CERA, E. Chromogenic substrates selective for activated protein C. *Blood*, v. 89, p. 2220-2222 (1997).
- DANG, Q.D.; VINDIGNI, A.; DI CERA, E. An allosteric switch controls the procoagulant and anticoagulant activities of thrombin. *Proc. Natl. Acad. Sci. USA*, v. 92, p. 5977-5981 (1995).

- DAVIDSON, C.J.; TUDDENHAM, E.G.; McVEY, J.H. 450 million years of hemostasis. *J. Thromb. Haemost.*, v. 1, p. 1487-1494 (2003).
- DAVIE, E.W. A brief historical review of the waterfall/cascade of blood coagulation. *J. Biol. Chem.*, v. 278, p. 50819-50832 (2003).
- DAVIES, D.R. The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Chem.*, v. 19, p. 189-215 (1990).
- DE CRISTOFARO, R.; DE CANDIA, E.; LANDOLFI, R. *et al.* Structural and functional mapping of the thrombin domain involved in the binding to the platelet glycoprotein Ib. *Biochemistry*, v. 40, p. 13268-13273 (2001).
- DeMARTINO, G.N.; SLAUGHTER C.A. The proteasome, a novel protease regulated by multiple mechanisms. *J. Biol. Chem.*, v. 274, p. 22123-22126 (1999).
- DE OLIVEIRA, E.J.; KANAMURA, H.Y.; LIMA, D.M.C. Efficacy of an enzyme-linked immunosorbent assay as a diagnostic tool for schistosomiasis mansoni in individuals with low worm burden. *Mem. Inst. Oswaldo Cruz*, v. 100, p. 421-425 (2005).
- DESHIMARU, M.; OGAWA, T.; NAKASHIMA, K.I. *et al.* Accelerated evolution of Crotalinae snake venom gland serine proteases. *FEBS Lett.*, v. 397, p. 83-88 (1996).
- DE SIMONE, S.G.; AGUIAR, A.S.; GIMENEZ, A.R. *et al.* Purification, properties, and N-terminal amino acid sequence of a kallikrein-like enzyme from the venom of *Lachesis muta rhombeata* (bushmaster). *J. Prot. Chem.*, v. 16, p. 809-818 (1997).
- DE SIMONE, S.G.; SANTOS, R.; ARAUJO, M.F. *et al.* Preparative isolation of the lectin jacalin by anion-exchange high performance liquid chromatography. *J. Chromatogr. A*, v. 688, p. 356-362 (1994).
- DHANARAJ, V.; DEALWIS, C.G.; FRAZAO, C. *et al.* X-ray analysis of peptide-inhibitor complexes define the structural basis of specificity for human and mouse renins. *Nature*, v. 357, p. 466-472 (1992).
- DIAMANDIS, E.P.; YOUSEF, G.M.; OLSSON, A.Y. An update on human and mouse glandular kallikreins. *Clinical Biochemistry*, v. 37, p. 258-260 (2004).
- DI CERA, E. Site-Specific Thermodynamics: Understanding Cooperativity in Molecular Recognition. *Chem. Rev.*, v. 98, p. 1563-1591 (1998).

- DI CERA, E. Thrombin: a paradigm for enzyme allosterically activated by monovalent cations. *C. R. Biologies*, v. 327, p. 1065-1076 (2004).
- DIAS-NETO, E.; CORREA, R.G.; VERJOVSKI-ALMEIDA, S. *et al.*, Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc. Natl. Acad. Sci. USA*, v. 97, p. 3491-3496 (2000).
- DOENHOFF, M.J.; CHIODINI, P.L.; HAMILTON, J.V. Specific and sensitive diagnosis of schistosome infection: can it be done with antibodies? *Trends Parasitol.*, v. 20, p. 35-39 (2004).
- DOENHOFF, M.J.; KIMANI, G.; CIOLI, D. Praziquantel and the control of schistosomiasis. *Parasitol. Today*, v. 16, p. 364-366 (2000).
- DRESDEN, M.H.; DEELER, A.M. *Schistosoma mansoni*: thiol protease properties of adult worm "hemoglobinase". *Exp. Parasitol.*, v. 14, p. 68-73 (1979).
- DRESDEN, M.H.; RUTLEDGE, M.L.; CHAPPELL, C.L. Properties of the acidic thiol protease from *Schistosoma mansoni* adults. *Mol. Biochem. Parasitol.*, v. 4, p. 61-66 (1981).
- DULLWEBER, F., STUBBS, M. T., MUSIL, D. *et al.* Factorising ligand affinity: a combined thermodynamic and crystallographic study of trypsin and thrombin inhibition. *J. Mol. Biol.* 313, 593-614 (2001).
- DUNN, B.M. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem. Rev.*, v. 102, p. 4431-4458 (2002).
- DUNN, B.M.; GUSTCHINA, A.; WLODAWER, A. *et al.* Subsite preferences of retroviral proteinases. *Methods Enzymol.*, v. 241, p. 254-278 (1994).
- DUNN, B.M.; HUNG, S-H. The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases. *Biochim. Biophys. Acta*, v. 1477, p. 231-240 (2000).
- EHRMANN, M.; CLAUSEN, T. Proteolysis as a regulatory mechanism. *Annu. Rev. Genet.*, v. 38, p. 709-24.
- EIGENBROT, C. Structure, function, and activation of coagulation factor VII. *Curr. Prot. Pep. Sci.*, v. 3, p. 287-299 (2002).

- EIGENBROT, C.; KIRCHHOFER, D.; DENNIS, M.S. *et al.* The factor VII zymogen structure reveals registration of b strands during activation. *Structure*, v. 9, p. 627-636 (2001).
- EL-SAYED, N.M.A.; BARTHOLOMEU, D.; IVENS, A. Advances in schistosome genomics. *Trends Parasitol.*, v. 20, p. 154-157 (2004).
- ESMON, C.T. Regulation of blood coagulation. *Biochim. Biophys. Acta*, v. 1477, p. 349-360 (2000).
- ESMON, C.T. The protein C pathway. *CHEST*, v. 124, p. 26S-32S (2003).
- ESSEX, J.W.; SEVERANCE, D.L.; TIRADO-RIVES, J. *et al.* Monte Carlo Simulations for Proteins: Binding Affinities for Trypsin-Benzamidine Complexes via Free-Energy Perturbations. *J. Phys. Chem. B*, v. 101, p. 9663-9669 (1997).
- FAUST, P.L.; KORNFELD, S.; CHIRGWIN, J.M. Cloning and sequence analysis of cDNA for human cathepsin D. *Proc. Natl. Acad. Sci. USA*, v. 82, 4910-4914.
- FELDMEIER, H. *Diagnosis*. In Jordan, P.; Webbe, G.; Sturrok, R.F. *Human schistosomiasis*, Wallingford: CAB International, 1993, p. 271-303.
- FENWICK, A.; SAVIOLI, L.; ENGELS, D. Drugs for control of parasitic diseases: current status and development in schistosomiasis. *Trends Parasitol.*, v. 19, p. 509-515 (2003).
- FISCHER, D.; BARRET, C.; BRYSON, K. *et al.* Critical Assessment of Fully Automated Structure Prediction Methods. *Proteins: Struct. Funct. Genet.*, Suppl 3, p. 209-217 (1999).
- FODA, H.D.; ZUCKER, S. Matrix metalloproteinases in cancer invasion, metastasis and angiogenesis. *Drug Discov. Today*, v. 6, p. 478-482 (2001).
- FOLTMANN, B.; PEDERSEN, V.B.; JACOBSEN, H. *et al.* The primary structure of calf chymosin. *J. Biol. Chem.*, v. 254, p. 8447-8456 (1977).
- FORTENBERRY, S.C.; CHIRGWIN, J.M. The propeptide is nonessential for the expression of human cathepsin D. *J. Biol. Chem.*, v. 270, p. 9778 (1995).
- FOSTER, D.; DAVIE, E.W. Characterization of a cDNA coding for human protein C. *Proc. Natl. Acad. Sci. USA*, v. 81, p. 4766-4770 (1984).

FRANCIS, S.E.; BANERJEE, R.; GOLDBERG, D.E. Biosynthesis and maturation of the malaria hemoglobins plasmeprins I and II. *J. Biol. Chem.*, v. 272, p. 14961-14968 (1997).

FRANCISCHETTI, I.M.B.; CASTRO, H.C.; ZINGALI, R.B. *et al.* Bothrops sp. snake venoms: comparison of some biochemical and physicochemical properties and interference in platelet function. *Comp. Biochem. Physiol.*, v. 119C, p. 21-29 (1998).

FRANCO, G.R.; ADAMS, M.D.; SOARES, M.B. *et al.* Identification of new *Schistosoma mansoni* genes by the EST strategy using a directional cDNA library. *Gene*, v. 152, p. 141-147 (1995).

FRANCO, G.R.; VALADÃO, A.F.; AZEVEDO, V. *et al.* The Schistosoma gene discovery program: state of the art. *Int. J. Parasitol.*, v. 30, p. 453-463 (2000).

FRUITIER, I.; GARREAU, I.; PIOT, J-M. Cathepsin D is a good candidate for the specific release of a stable hemorphin from hemoglobin in vivo: VV-hemorphin-7. *Biochem. Biophys. Res. Commun.*, v. 246, p. 719-724 (1998).

FUENTES-PRIOR, P.; IWANAGA, Y.; HUBER, R. *et al.* Structural basis for the anticoagulant activity of the thrombin-thrombomodulin complex. *Nature*, v. 404, p. 518-525 (2000).

FUHRMANN, M.; HAUSHERR, A.; FERBITZ, L. *et al.* Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Mol. Biol.*, v. 55, p. 869-881 (2004).

FUKUDA, R.; UMEBAYASHI, K. HOVIUSHI, H. *et al.* Degradation of *Rhizopus niveus* aspartic proteinase-I with mutated prosequences occurs in the endoplasmic reticulum of *Saccharomyces cerevisiae*. *J. Biol. Chem.*, v. 271, p. 14252 (1996).

REUTER G. & GABIUS H-J. Eukaryotic glycosylation: whim of nature or multipurpose tool? *Cell. Mol. Life Sci.*, v. 55, p. 368-422 (1999).

GALE, A.J.; GRIFFIN, J.H. Characterization of a thrombomodulin binding site on protein C and its comparison to an activated protein C binding site for factor Va. *Proteins: Struct. Funct. Bioinfo.*, v. 54, p. 433-441 (2004).

GALLEGO, S.G.; SLADE, R.W.; BRINDLEY, P.J. A cDNA encoding a pepsinogen-like, aspartic protease from the human roundworm parasite *Strongyloides stercoralis*. *Acta Tropica*, v. 71, p. 17-26 (1998).

GALVIN, J.B.; KUROSAWA, S.; MOORE, K. *et al.* Reconstitution of rabbit thrombomodulin into phospholipid-vesicles. *J. Biol. Chem.*, v. 262, p. 2199-2205 (1987).

GEHLHAAR, D.K.; VERKHIVKER, G.M.; REJTO, P.A. *et al.* Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.*, v. 2, p. 317-24 (1995).

GEIER, G.; BANAJ, H-J.; HEID, H. *et al.* Aspartyl proteases in *Caenorhabditis elegans*: Isolation, identification and characterization by a combined use of affinity chromatography, two-dimensional gel electrophoresis, microsequencing and databank analysis. *Eur. J. Biochem.*, v. 264, p. 872-879 (1999).

GEOURJON, C.; DELEAGE, G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci.*, v. 11, p. 681-684 (1995).

GETTINS, P.G.W. Serpin Structure, Mechanism, and Function. *Chem. Rev.*, v. 102, p. 4751-4803 (2002).

GILL, J.C. Therapy of factor VIII deficiency. *Semin. Thromb. Hemost.*, v. 19, p. 1-12 (1993).

GOMEZ, K.; LAFFAN, M.A.; KEMBALL-COOK, G. *et al.* Two novel mutations in severe factor VII deficiency. *British J. Haemat.*, v. 126, p. 105-110 (2004).

GOTZ, B.; KLINKERT, M-Q. Expression and partial characterization of a cathepsin B-like enzyme (Sm31) and a proposed "haemoglobinase" (Sm32) from *Schistosoma mansoni*. *Biochem. J.*, v. 290, p. 801-806 (1993).

GRAF, L.; JANCISO, A.; SZILAGYI, L. *et al.* Electrostatic complementarity within the substrate-binding pocket of trypsin. *Proc. Natl. Acad. Sci. USA*, v. 85, p. 4961-4965 (1988).

GREEN, D. Complications associated with the treatment of haemophiliacs with inhibitors. *Haemophilia*, v. 5, S11-S17 (1999).

GROLL, M.; BOCHTLER, M.; BRANDSTETTER, H. *et al.* Molecular machines for protein degradation. *Chembiochem*, v. 6, p. 222-256 (2005).

GRUBER, A.; CANTWELL, A.M.; DI CERA, E. *et al.* The thrombin mutant W215A/E217A shows safe and potent anticoagulant and antithrombotic effects in vivo. *J. Biol. Chem.*, v. 277, p. 27581-27584 (2002).

GUIMARÃES, C.R.W.; DE ALENCASTRO, R.B. Thermodynamic Analysis of Thrombin Inhibition by Benzamidine and p-Methylbenzamidine via Free-Energy Perturbations: Inspection of Intraperturbed-Group Contributions Using the Finite Difference Thermodynamic Integration (FDTI) Algorithm. *J. Phys. Chem. B*, v. 106, p. 466-476 (2002).

GUPTA, R.; JUNG, E.; GOOLEY, A.A. *et al.* Scanning the available *Dictyostelium discoideum* proteome for O-linked GlcNAc glycosylation sites using neural networks. *Glycobiology*, v. 9, p. 1009-1022 (1999).

GUSTAFSSON, D.; BYLUND, R.; ANTONSSON, T. *et al.* A new oral anticoagulant: the 50-year challenge. *Nat. Rev. Drug. Discov.*, v. 3, p. 649-659 (2004).

GUYATT, H.L.; BROOKER, S.; KIHAMIA, C.M. *et al.* Evaluation of efficacy of school-based anthelmintic treatments against anaemia in children in the United Republic of Tanzania. *Bull. WHO*, v. 79, p. 695-703 (2001).

HALL, S.W.; GIBBS, C.S.; LEUNG, L.L.K. Strategies for development of novel antithrombotics: modulating thrombin's procoagulant and anticoagulant properties. *Cell. Mol. Life Sci.*, v. 53, p. 731-736 (1997).

HALL, S.W.; GIBBS, C.S.; LEUNG, L.L. Identification of critical residues on thrombin mediating its interaction with fibrin. *Thromb. Haemost.*, v. 86, p. 1466-1476 (2001).

HAMMETT, L.P. Some Relations between Reaction Rates and Equilibrium Constants. *Chem. Rev.*, v. 17, p. 125-136 (1935).

HANNA, G.P.; SMALLING, R.W. Therapy with thrombolytic agents in coronary artery disease. *Curr. Opin. Cardiol.*, v. 13, p. 267-273 (1998).

HARROP, A.S.; PROCIV, P.; BRINDLEY, P.J. Acasp, a gene encoding a cathepsin D-like aspartic protease from the hookworm *Ancylostoma caninum*. *Biochem. Biophys. Res. Commun.*, v. 227, p. 294-302 (1996).

HE, X.; REZAIIE, A.R. Identification and characterization of the sodium-binding site of activated protein C. *J. Biol. Chem.*, v. 274, p. 4970-4976 (1999).

- HEDNER, U. Recombinant FVIIa. *Vox Sang.*, v. 87, p. S25-S28 (2004).
- HEDSTROM, L. Serine protease mechanism and specificity. *Chem. Rev.*, v. 102. p. 4501-4523 (2002).
- HEDSTROM, L.; FARR-JONES, S.; KETTNER, C. A. *et al.* Converting trypsin to chymotrypsin: ground-state binding does not determine substrate specificity. *Biochemistry*, v. 33, p. 8764-8769 (1994a).
- HEDSTROM, L.; PERONA, J. J.; RUTTER, W. J. Converting trypsin to chymotrypsin: residue 172 is a substrate specificity determinant. *Biochemistry*, v. 33, p. 8757-8763 (1994b).
- HEDSTROM, L.; SZILAGYI, L.; RUTTER, W. J. Converting trypsin to chymotrypsin: the role of surface loops. *Science*, v. 255, p. 1249-1253 (1992).
- HEIT, J.A. The potential role of direct thrombin inhibitors in the prevention and treatment of venous thromboembolism. *CHEST*, v. 124, p. 40S-48S (2003).
- HELLAND, R.; CZAPINSKA, H.; LEIROS, I. *et al.* Structural Consequences of Accommodation of Four Non-cognate Amino Acid Residues in the S1 Pocket of Bovine Trypsin and Chymotrypsin. *J. Mol. Biol.*, v. 333, p. 845–861 (2003).
- HENIKOFF, J.G. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA*, v. 89, p. 10915–10919 (1992).
- HERRICK, S.; BLANC-BRUDE, O.; GRAY, A. *et al.* Fibrinogen, *Int. J. Biochem. Cell Biol.*, v. 31, p. 741-746 (1999).
- HESS, J.R.; HIIPPALA, S. Optimizing the use of blood products in trauma care. *Crit. Care*, v. 9, S10-S14 (2005).
- HILL, J.; MONTGOMERY, D.S.; KAY, J. Human cathepsin E produced in *E. coli*. *FEBS Lett.*, v. 326, p. 101-104 (1993).
- HOBART, P.M.; FOGLIANO, M.; O'CONNOR B.A. *et al.* Human rennin gene: structure and sequence analysis. *Proc. Natl. Acad. Sci. USA*, v. 81, p. 5026-5030 (1984).
- HOEGL, L.; OLLERT, M.; KORTING, H.C. The role of *Candida albicans* secreted aspartic proteinase in the development of candidoses. *J. Mol. Med.*, v. 74, p. 135-142 (1996).

- HOFFMAN, G.; GARRISON, T.R.; DOHLMAN, H.G. Analysis of RGS proteins in *Saccharomyces cerevisiae*. *Meth. Enzymol.*, v. 344, p. 617-631 (2002b).
- HOFFMAN, M. Remodeling the blood coagulation cascade. *J. Thromb. Thrombol.*, v. 16, p. 17-20 (2003).
- HOFFMANN, K.F.; JOHNSTON, D.A.; DUNNE, D.W. Identification of *Schistosoma mansoni* gender-associated gene transcripts by cDNA microarray profiling. *Genome Biol.*, v. 3, p. 1-12 (2002a).
- HOLCOMB, J.B. Methods for improved hemorrhage control. *Crit. Care*, v. 8, p. S57-S60 (2004).
- HOLLEMAN, W.H.; WEISS, L.J. The thrombin-like enzyme from *Bothrops atrox* snake venom. Properties of the enzyme purified by the affinity chromatography on p-aminobenzamidine-substituted agarose. *J. Biol. Chem.*, v. 251, p. 1663-1669 (1976)
- HOPFNER, K-P.; LANG, A.; KARCHER, A. *et al.* Coagulation factor IXa: the relaxed conformation of Tyr99 blocks substrate binding. *Structure*, v. 7, p. 989-996 (1999).
- HORTELANO, G.; OFOSU, F.A. Therapeutic approaches for haemophilia. *Exp. Opin. Therap. Pat.*, v. 10, p. 929-938 (2000).
- HUNG, S-W.; HEDSTROM, L. Converting trypsin to elastase: substitution of the S1 site and adjacent loops reconstitutes esterase specificity but not amidase activity. *Prot. Eng.*, v. 11, p. 669-673 (1998).
- HUNG, C.C.; CHIOU, S.H. Fibrinogenolytic proteases isolated from the snake venom of Taiwan habu: Serine proteases with kallikrein-like and angiotensin-degrading activities. *Biochem. Biophys. Res. Commun.*, v. 281, p. 1012-1018 (2001).
- HUNTINGTON, J.A. Molecular recognition mechanisms of thrombin. *J. Thromb. Haemost.*, v. 3, p. 1861-1872 (2005).
- HUNTINGTON, J.A.; ESMON, C.T. The molecular basis of thrombin allostery revealed by a 1.8Å structure of the "slow" form. *Structure*, v. 11, p. 469-479 (2003).
- IMPERIALI, B.; O'CONNOR, S.E. Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Curr. Opin. Chem. Biol.*, v. 3, p. 643-649 (1999).
- INOUE, H.; HAYASHI, T.; HUANG, X-P. *et al.* Heterologous expression and site-directed mutagenesis studies on the activation mechanism and the roles of the basic

residues in the prosegment of aspergillopepsinogen I. *Eur. J. Biochem.*, v. 237, p. 719-725 (1996).

ITOH, N.; TANAKAG, N.; FUNAKOSHI, I. *et al.* Organization of the gene for batroxobin, a thrombin-like snake venom enzyme. Homology with the trypsin/kallikrein gene family. *J. Biol. Chem.*, v. 263, p. 7628-7631 (1988).

JAMES, M.N.; SIELECKI, A.R. Stereochemical analysis of peptide bond hydrolysis catalyzed by the aspartic proteinase penicillopepsin. *Biochemistry*, v. 24, p. 3701-3713 (1985).

JAMES, M.N.; SIELECKI, A.R.; HAYAKAWA, K. *et al.* Crystallographic analysis of transition state mimics bound to penicillopepsin: difluorostatine- and difluorostatone-containing peptides. *Biochemistry*, v. 31, p. 3872 (1992)

JANCIAUSKIENE, S. Conformational properties of serine proteinase inhibitors (serpins) confer multiple pathophysiological roles. *Biochim. Biophys. Acta*, v. 1535, p. 221-235 (2001).

JEAN, L.; LONG, M.; YOUNG, J. *et al.* Aspartyl proteinase genes from apicomplexan parasites: evidence for evolution of the gene structure. *Trends Parasitol.*, v. 17, p. 491-498 (2001).

JEDINAK, A.; MALIAR, T. Inhibitors of proteases as anticancer drugs – Minireview. *Neoplasma*, v. 52, p. 185-192 (2005).

JELINEK, B.; ANTAL, J.; VENEKEI, I. *et al.* Ala226 to Gly and Ser189 to Asp mutations convert rat chymotrypsin B to a trypsin-like protease. *Prot. Eng. Des. Sel.*, v. 17, p. 127-131 (2004).

JOLODAR, A.; MILLER, D.J. Identification of a novel family of non-lysosomal aspartic proteases in nematodes. *Biochim. Biophys. Acta*, v. 1382, p. 13-16 (1998).

JONES D.T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.*, v. 292, p. 195-202 (1999).

JONES, D.T. GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. *J Mol Biol.*, v. 287, p. 797-815 (1999).

- JORGE, M.T.; SANO-MARTINS, I.S.; TOMY, S.C. *et al.* Snake bite by the bushmaster (*Lachesis muta*) in Brazil: case report and review of the literature. *Toxicon*, v. 35, p. 545-554 (1997).
- JOSEPH, J.S.; CHUNG, M.C.M; MIRTSCHEIN, P.J. *et al.*, Effect of snake venom procoagulants on snake plasma: implications for the coagulation cascade of snakes. *Toxicon*, v. 40, p. 175-183 (2002).
- KADONO, S.; SAKAMOTO, A.; KIKUCHI, Y. *et al.* Crystal structure of human factor VIIa/tissue factor in complex with peptide mimetic inhibitor. *Biochem. Biophys. Res. Commun.*, v. 324, p. 1227-1233 (2004).
- KALDOR, S.W.; KALISH, V.J.; DAVIES-II, J.F. *et al.* Viracept (Nelfinavir Mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease, *J. Med. Chem.*, v. 40, p. 3979 (1997).
- KAM, P.C.A.; KAUR, N.; THONG, C.L. Direct thrombin inhibitors: pharmacology and clinical relevance. *Anaesthesia*, v. 60, p. 565–574 (2005).
- KASSCHAU, M.R.; ROBINSON, D.C.; DRESDEN, M.H. *Schistosoma mansoni*: activation of hemolytic activity in homogenates from live adult worms. *Exp. Parasitol.*, v. 62, p. 442-449 (1986).
- KATHIRESAN, S.; SHIOMURA, J.; JANG, I-K. Argatroban. *J. Thromb. Thrombol.*, v. 13, p. 41–47 (2002).
- KATZ, N.K.; CHAVES, A.; PELLEGRINO, J. A simple device for quantitative stool thick-smear in schistosomiasis mansoni. *Rev. Inst. Med. São Paulo*, v. 14, p. 397-400 (1972).
- KAUZMANN, W. Some factors in the interpretation of protein denaturation. *Adv. Prot. Chem.*, v. 14, p. 1-63 (1959).
- KAY, J.; DUNN, B.M. Substrate specificity and inhibitors of aspartic proteinases. *Scand. J. Clin. Lab. Invest.*, v. 52, p. S23-S30 (1992).
- KELLERMANN, G.; VICENTIN, F.; TAMURA, E. *et al.* The Small-Angle X-ray Scattering Beamline of the Brazilian Synchrotron Light Laboratory. *J. Appl. Crystallogr.*, v. 30, p. 880–883 (1997).

KEMBALL-COOK, G.; JOHNSON, J.D.; TUDDENHAM, E.G.D. *et al.* Crystal structure of active site-inhibited human coagulation factor VIIa (des-Gla). *J. Struct. Biol.*, v. 127, p. 213-223 (1999).

KEMPF, D.J.; MARSH, K.C.; DENISSEN, J.F. *et al.* ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proc. Natl. Acad. Sci. USA.* v. 92, p. 2484-2488 (1995).

KESAVULU, M.M.; PRAKASHA-GOWDA, A.S.; RAMYA, T.N. *et al.* Plasmepsin inhibitors: design, synthesis, inhibitory studies and crystal structure analysis. *J. Pept. Res.*, v. 66, p. 211-219 (2005).

KESSLER, C.M. Antidotes to haemorrhage: recombinant factor VIIa. *Best Pract. Res. Clin. Haematol.*, v. 17, p. 183-197 (2004).

KHAN, A.R.; BERNSTEIN, N.K.; BERGMANN, E.M.; JAMES, M.N.G. Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. *Proc. Natl. Acad. Sci. USA*, v. 96, p. 10968-75 (1999).

KHRENOV, A.V.; ANANYEVA, N.M.; GRIFFIN, J.H. *et al.* Coagulation pathways in atherothrombosis. *Trends Cardiovasc. Med.*, v. 12, p. 317-324 (2002).

KIM, E.E.; BAKER, C.T.; DWYER, M.D. *et al.* Crystal-structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. *J. Am. Chem. Soc.*, v. 117, p. 1181-1182 (1995).

KNIGHT, C.G. Active-site titration of peptidases. *Meth. Enzymol.*, v. 248, p. 85-101 (1995).

KNOBE, K.E.; BERNTSDOTTER, A.; SHEN, L. *et al.* Probing the activation of protein C by the thrombin-thrombomodulin complex using structural analysis, site-directed mutagenesis and computer modeling. *Proteins: Struct. Funct. Genet.*, v. 35, p. 218-234 (1999).

KOBAYASHI, M.; FUJIWARA, Y.; GODA, M. *et al.* Identification of active sites in amidase: Evolutionary relationship between amide bond- and peptide bond-cleaving enzymes. *Proc. Natl. Acad. USA*, v. 94, p. 11986-11991 (1997).

KRATKY O. X-ray small angle scattering with substances of biological interest in diluted solutions. *Progr. Biophys. Mol. Biol.*, v. 13, p. 105 (1963).

KREM, M. M.; DI CERA, E. Evolution of enzyme cascades from embryonic development to blood coagulation. *TRENDS Biochem. Sci.*, v. 27, p. 67-74 (2002b).

KREM, M. M.; PRASAD, S.; DI CERA, E. Ser²¹⁴ is crucial for substrate binding to serine proteases. *J. Biol. Chem.*, v. 277, p. 40260-40264 (2002a).

KREM, M.M.; ROSE, T.; DI CERA, E. The C-terminal sequence encodes function in serine proteases. *J. Biol. Chem.*, v. 274, p. 28063-28066 (1999).

KURTH, T.; GRAHN, S.; THORMAN, M.; ULLMANN, D. *et al.* Engineering the S1' subsite of trypsin: design of a protease which cleaves between dibasic residues. *Biochemistry*, v. 37, p. 11434-11440 (1998).

KURTH, T.; ULLMANN, D.; JAKUBKE, H-D. *et al.* Converting trypsin to chymotrypsin: structural determinants of S1' specificity. *Biochemistry*, v. 36, p. 10098-10104 (1997).

KUSEL, J.; HAGAN, P. Praziquantel – Its use, cost and possible development of resistance. *Parasitol. Today*, v. 15, p. 352-354 (1999).

LAEMMLI, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, v. 227, p. 680-685 (1970).

LAKOWICS, J.R. *Protein fluorescence. In Principles of fluorescence spectroscopy*. 2nd ed. New York: Kluwer Academic, 1999, p. 446-487.

LASKOWSKI-JR, M.; QASIM, M. A. What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? *Biochim. Biophys. Acta*, v. 1477, p. 324-337 (2000).

LAWRENCE, J.D. The ingestion of red blood cells by *Schistosoma mansoni*. *J. Parasitol.*, v. 59, p. 60-63 (1973).

LEE, K.W.; LIP, G.Y.H. Effects of lifestyle on hemostasis, fibrinolysis, and platelet reactivity. A systematic review. *Arch. Int. Med.*, v. 163, p. 2368-2392 (2003).

LEISSINGER, C.A. Use of prothrombin complex concentrates and activated prothrombin complex concentrates as prophylactic therapy in haemophilia patients with inhibitors. *Haemophilia*, v. 5, p. S25-S32 (1999).

LESK, A.M.; FORDHAM, W.D. Conservation and variability in the structure of serine proteinases of the chymotrypsin family. *J. Mol. Biol.*, v. 258, p. 501-537 (1996).

- LEUNG, L.L.K.; HALL, S.W. Dissociation of thrombin's substrate interactions using site-directed mutagenesis. *Trends Cardiovasc. Med.*, v. 10, p. 89-92 (2000).
- LI, C.Q.; VINDIGNI, A.; SADLER, J.E. *et al.* Platelet glycoprotein Ib alpha binds to thrombin anion-binding exosite II inducing allosteric changes in the activity of thrombin. *J. Biol. Chem.*, v. 276, p. 6161-6168 (2001).
- LILIE, H.; SCHWARZ, E.; RUDOLPH, R. Advances in refolding of proteins produced in *E. coli*. *Curr. Opin. Biotechnol.*, v. 9, p. 497-501 (1998).
- LIN C.; FRISHMAN, M.D. Renin inhibition: a novel therapy for cardiovascular disease. *Am. Heart J.*, v. 131, p. 1024-34 (1996).
- LIN, X.; WONG, R.N.S.; TANG, J. Synthesis, purification, and active site mutagenesis of recombinant porcine pepsinogen. *J. Biol. Chem.*, v. 294, p. 4482-4489 (1989).
- LINDER, R.; FREBELIUS, S.; GRIP, L. *et al.* The influence of direct and antithrombin-dependent thrombin inhibitors on the procoagulant and anticoagulant effects of thrombin. *Thromb. Res.*, v. 110, p. 221-226 (2003).
- LISOWSKA, E. The role of glycosylation in protein antigenic properties. *Cell. Mol. Life Sci.*, v. 59, p. 445-455 (2002).
- LIU, L.W.; VU, T.K.; ESMON, C.T. *et al.* The region of the thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. *J. Biol. Chem.*, v. 266, p. 16977-16980 (1991).
- LOCHNIT G.; GEYER, R. Carbohydrate structure analysis of batroxobin, a thrombin-like serine protease from *Bothrops moojeni* venom. *Eur. J. Biochem.*, v. 228, p. 805-816 (1995).
- LOETSCHER, H.; NIEDERHAUSER, O.; KEMP, J. *et al.* Is caspase-3 inhibition a valid therapeutic strategy in cerebral ischemia?, *Drug Discov. Today*, v. 6, p. 671-80 (2001).
- LOWE, G.D.O. Factor IX and thrombosis. *Brit. J. Haematol.*, v. 115, p. 507-513 (2001).
- LOLLAR, P. Pathogenic antibodies to coagulation factors. Part one: Factor VIII and Factor IX. *J. Thromb. Haemost.*, v. 2, p. 1082-1095 (2004).

- LOLLAR, P. Pathogenic antibodies to coagulation factors. Part II. Fibrinogen, prothrombin, thrombin, factor V, factor XI, factor XII, factor XIII, the protein C system and von Willebrand factor. *J. Thromb. Haemost.*, v. 3, p. 1385–1391 (2005).
- LONGBOTTOM, D.; REDMOND, D.L.; RUSSELL, M. *et al.* *Mol. Biochem. Parasitol.*, v. 88, p. 63-72 (1997).
- LOVERDE, P.T.; HIRAI, H.; MERRICK, J.M. *et al.* Schistosoma mansoni genome project: an update. *Parasitol. Int.*, v. 53, p. 183-192 (2004).
- LOWE, G.D.O. Factor IX and thrombosis. *British J. Haemat.*, v. 115, p. 507-513 (2001).
- LOWRY, O.H.; ROSEBROUGH, N.J.; FARR A.L. *et al.* Protein measurement with the folin phenol reagent. *J. Biol. Chem.*, v. 193, p. 265-275 (1951).
- LU, Q.; CLEMETSON, J.M.; CLEMETSON, K.J. Snake venoms and hemostasis. *J. Thromb. Haemost.*, v. 3, p. 1791-1799 (2005).
- LUBETSKY, A.; HOFFMAN, R.; ZIMLICHMAN, R. *et al.* Efficacy and safety of a prothrombin complex concentrate (Octaplex) for rapid reversal of oral anticoagulation. *Thromb. Res.*, v. 113, p. 371-378 (2004).
- LUSHER, J.M. Controlled clinical trials with prothrombin complex concentrates. *Prog. Clin. Biol. Res.*, v. 150, p. 277-290 (1984).
- MACAULEY-PATRICK, S.; FAZENDA, M.L.; McNEIL, B. *et al.* Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, v. 22, p. 249-270 (2005).
- MAGALHÃES, A.; DA-FONSECA, B.C.; DINIZ, C.R. *et al.* The complete amino acid sequence of a thrombin-like enzyme/gyroxin analogue from venom of the bushmaster snake (*Lachesis muta muta*). *FEBS Lett.* v. 329, p. 116-120 (1993).
- MAGALHÃES, A.; FERREIRA, R.N.; RICHARDSON, M. *et al.* Coagulant thrombin-like enzymes from the venoms of Brazilian and Peruvian bushmaster (*Lachesis muta muta*) snakes. *Comp. Biochem. Physiol. B*, v. 136, p. 255-266 (2003).
- MAGALHÃES, A.; MONTEIRO, M.R.; MAGALHÃES, H.P.B. *et al.* Thrombin-like enzyme from *Lachesis muta muta* venom: isolation and topographical analysis of its

active site structure by means of the binding of amidines and guanidines as competitive inhibitors. *Toxicon*, v. 35, p. 1549-1559 (1997).

MAJERUS, P.W.; TOLLEFSEN, D.M. *Anticoagulant, thrombolytic, and antiplatelet drugs*. In GILMAN, A.G.; HARDMAN, J.G.; LIMBIRD, L.E. (Eds) Goodman & Gilman's The pharmacological basis of therapeutics. 10th ed. New York: McGraw-Hill, 2001, p. 1519-1538.

MANDELBAUM, F.R.; HENRIQUES, O.B. Purification and properties of Bothrops protease A. *Arch. Biochem. Biophys.*, v. 104, p. 369-374 (1964).

MANITHODY, C.; FAY, P.J.; REZAIE, A.R. Exosite-dependent regulation of factor VIIIa by activated protein C. *Blood*, v. 101, p. 4802-4807 (2003).

MANITHODY, C.; REZAIE, A.R. Functional mapping of charged residues of the 82-116 sequence in factor Xa: evidence that lysine 96 is a factor Va independent recognition site for prothrombin in the prothrombinase complex. *Biochemistry*, v. 44, p. 10063-10070 (2005).

MANITHODY, C.; YANG, L.; REZAIE, A.R. Role of basic residues of the autolysis loop in the catalytic function of factor Xa. *Biochemistry*, v. 41, p. 6780-6788 (2002).

MANN, K.G. Thrombin formation. *CHEST*, v. 124, p. 4S-10S (2003).

MARES-GUIA, M.; FIGUEIREDO, A.F.S. Thermodynamics of the hydrophobic interaction in the active center of trypsin. Investigation with amidines and guanidines. *Biochemistry*, v. 9, p. 3223-3227 (1970).

MARES-GUIA, M.; NELSON, D.L.; ROGANA, E. Electronic effects in the interaction of para-substituted benzamidines with trypsin: the involvement of the π -electronic density at the central atom of the substituent in binding. *J. Am. Chem. Soc.*, v. 99, p. 2331-2336 (1977).

MARES-GUIA, M.; SHAW, E. Studies on the Active Center of Trypsin. The binding of amidines and guanidines as models of the substrate side chain. *J. Biol. Chem.*, v. 240, p. 1579-1585 (1965).

MARKLAND, F.S. Snake venoms and the hemostatic system. *Toxicon*, v. 36, p. 1749-1800 (1998).

- MARKLAND, F.S.; KETTNER, C.; SCHIFFMAN, S. *et al.* Kallikrein-like activity of crotalase, a snake-venom enzyme that clots fibrinogen. *Proc. Natl. Acad. Sci. U.S.A.*, v. 79, p. 1688-1692 (1982).
- MARSH, N.; WILLIAMS, V. Practical applications of snake venom toxins in haemostasis. *Toxicon*, v. 45, p. 1171-1181 (2005).
- MARTIN, U.; KAUFMANN, B.; NEUGEBAUER, G. Current clinical use of reteplase for thrombolysis - A pharmacokinetic-pharmacodynamic perspective. *Clin. Pharmacok.*, v. 36, p. 265-276 (1999).
- MARTIN, P.D.; ROBERTSON, W.; TURK, D. *et al.* The structure of residues 7-16 of the A α chain of human fibrinogen bound to bovine thrombin at 2.3Å resolution. *J. Biol. Chem.*, v. 267, p. 7911-7920 (1992).
- MATHER, T.; OGANESSYAN, V.; HOF, P. *et al.* The 2.8 Å crystal structure of Gla-domainless activated protein C. *EMBO J.*, v. 15, p. 6822-6831 (1996).
- MATHIEU, M.A.; BOGYO, M.; CAFFREY, C.R. *et al.* Substrate specificity of schistosoma versus human legumain determined by P1-P3 peptide libraries. *Mol. Biochem. Parasitol.*, v. 121, p. 99-105 (2002).
- MATHUR, A.; BAJAJ, S.P. Protease and EGF1 domains of factor IXa play distinct roles in binding to factor VIIIa: importance of helix 330 (helix 162 in chymotrypsin) of protease domain of factor IXa in its interaction with factor VIIIa. *J. Biol. Chem.*, v. 274, p. 18477-18486 (1999).
- MATHUR, A.; ZHONG, D.; SABHARWAL, A.K. *et al.* Interaction of factor IXa with factor VIIIa: Effects of protease domain Ca²⁺ binding site, proteolysis in the autolysis loop, phospholipids, and factor X. *J. Biol. Chem.*, v. 272, p. 23418-23426 (1997).
- MATSUI, T.; FUJIMURA, Y.; TITANI, K. Snake venom proteases affecting haemostasis and thrombosis. *Biochim. Biophys. Acta.*, v. 1477, p. 146-156 (2000).
- MAUN, H.R.; EIGENBROT, C.; RAAB, H. *et al.* Disulfide locked variants of factor VIIa with a restricted b-strand conformation have enhanced enzymatic activity. *Prot. Sci.*, v. 14, p. 1171-1180 (2005).
- McGUFFIN, L.J.; BRYSON, K.; JONES, D.T. The PSIPRED protein structure prediction server. *Bioinformatics*, v. 16, p. 404-405 (2000).

McGUFFIN, L.J.; JONES, D.T. Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics*, v. 19, p. 874-881 (2003).

MCKERROW, J.H.; ENGEL, J.C.; CAFFREY, C.R. Cysteine protease inhibitors as chemotherapy for parasitic infections. *Bioorg. Med. Chem.* v. 7, p. 639-44 (1999).

McTIGUE, M.A.; WILLIAMS, D.R.; TAINER, J.A. Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase from *Schistosoma japonica* and its complex with the leading antischistosomal drug praziquantel. *J. Mol. Biol.*, v. 246, p. 21-27 (1995).

MEHLIN, C. Structure-based drug discovery for *Plasmodium falciparum*. *Comb. Chem. High Throughput Screen.*, v. 8, p. 5-14 (2005).

MENGWASSER, K.E.; BUSH, L.A.; SHIH, P. *et al.* Hirudin binding reveals key determinants of thrombin allostery. *J. Biol. Chem.*, v. 280, p. 26997-27003 (2005).

MERRICK, J.M.; OSMAN, A.; TSAI, J. The *Schistosoma mansoni* gene index: gene discovery and biology by reconstruction and analysis of expressed gene sequences. *J. Parasitol.*, v. 89, p. 261-169 (2003).

MEZELLA, H.G.; GRAMAJO, H.C.; CECCARELLI, E.A. High recovery of prochymosin from inclusion bodies using controlled air oxidation. *Prot. Exp. Purif.*, v. 25, p. 248-255 (2002).

MICHEL, A.; GHONEIM, H.; RESTO, M. Sequence, characterization and localization of a cysteine proteinase cathepsin L in *Schistosoma mansoni*. *Mol. Biochem. Parasitol.*, v. 73, p. 7-18 (1995).

MODIG, K.; KURIAN, E.; PRENDERGAST, F.G. *et al.* Water and urea interactions with the native and unfolded forms of a β -barrel protein. *Prot. Sci.*, v. 12, p. 2768-2781 (2003).

MONERA, O.D.; KAY, C.M.; HODGES, R.S. Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Prot. Sci.*, v. 3, p. 1984-1991 (1994).

MONNAIE, D.; AROSIO, D.; GRIFFON, N. *et al.* Identification of a binding site for quaternary amines in factor Xa. *Biochemistry*, v. 39, p. 5349-5354 (2000).

- MONTESINO, R.; GARCIA, R.; QUINTERO, O. *et al.* Variation in N-linked oligosaccharide structures on heterologous proteins secreted by the methylotrophic yeast *Pichia pastoris*. *Prot. Expr. Purif.*, v. 14, p. 197-207 (1998).
- MONTESINO, R.; NIMTZ, M.; QUINTERO, O. *et al.* Characterization of the oligosaccharides assembled on the *Pichia pastoris*-expressed recombinant aspartic protease. *Glycobiol.*, v. 9, p. 1037-1043 (1999).
- MOON, R.P.; TYAS, L.; CERTA, U. *et al.* Expression and characterization of plasmepsin I from *Plasmodium falciparum*. *Eur. J. Biochem.*, v. 244, p. 552-560 (1997).
- MORALES, M.E.; KALINNA, B.H.; HEYERS, O. *et al.* Genomic organization of the *Schistosoma mansoni* aspartic protease gene, a platyhelminth orthologue of mammalian lysosomal cathepsin D. *Gene*, v. 338, p. 99-109 (2004).
- MOSS, M.L.; WHITE, J.M.; LAMBERT, M.H. *et al.* TACE and other ADAM proteases as targets for drug discovery, *Drug Discov. Today*, v. 6, p. 417-426 (2001).
- MUNEKIYO, S.M.; MACKESSY, S.P. Effects of temperature and storage conditions on the electrophoretic, toxic and enzymatic stability of venom components. *Comp. Biochem. Physiol.*, v. 119B, p. 119-127 (1998).
- MURAYAMA, N.; SAGUCHI, K.; MENTELE, R. *et al.* The unusual high molecular mass of *Bothrops* protease A, a trypsin-like serine peptidase from the venom of *Bothrops jararaca*, is due to its high carbohydrate content. *Biochim. Biophys. Acta*, v. 1652, p. 1-6 (2003).
- NAVIA, M.A.; FITZGERALD, P.M.D.; McKEEVER, B.M. *et al.* 3-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature*, v. 337, p. 615-620 (1989).
- NDONWI, M.; BROZE-JR, G.; BAJAJ, S.P. The first epidermal growth factor-like domains of factor Xa and factor IXa are important for the activation of the factor VII-tissue factor complex. *J. Thromb. Haemost.*, v. 3, p. 112-118 (2005).
- NEURATH, H. Proteolytic Enzymes, past and future. *Proc. Natl. Acad. Sci. USA*, v. 96, p. 10962-10963 (1999).
- NIEWIAROWSKI, S.; KIRBY, E.P.; BRUDZYNSKI, T.M. *et al.* Thrombocytin, a serine protease from *Bothrops atrox* venom. 2. Interaction with platelets and Plasma-clotting factors. *Biochemistry*, v. 18, p. 3570-3577 (1979).

- NISHIDA, S.; FUJIRAMA, Y.; MIURA, S. *et al.* Purification and characterization of bothrombin, a fibrinogen-clotting serine protease from the venom of *Bothrops jararaca*. *Biochemistry*, v. 33, p. 1843-1849 (1994).
- NORTHROP, D.B. Follow the protons: a low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Acc. Chem. Res.*, v. 34, p. 790-797 (2001).
- NOVAK, J.F.; TRNKA, F. Proenzyme therapy of cancer. *Anticancer Res.*, v. 25, p. 1157-1177 (2005).
- O'FARRELL, P.H. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.*, v. 250, p. 4007-4021 (1975).
- OLIVEIRA, C.L.P.. "TRAT1D - Computer Program for SAXS Data Treatment", LNLS technical Manual MT 01/2003 (2003).
- OLIVEIRA, G.; RODRIGUES, N.B.; ROMANHA, A.J. *et al.* Genome and genomics of schistosomes. *Can. J. Zool.*, v. 82, p. 375-390 (2004).
- OLIVERO, A.G.; EIGENBROT, C.; GOLDSMITH, R. *et al.* A selective, slow binding inhibitor of factor VIIa binds to a nonstandard active site conformation and attenuates thrombin formation in vivo. *J. Biol. Chem.*, v. 280, p. 9160-9169 (2005).
- OLSEN, O.H.; NIELSEN, P.F.; PERSSON, E. Prevention of β strand movement into a zymogen-like position does not confer higher activity to coagulation factor VIIa. *Biochemistry*, v. 43, p. 14096-14103 (2004).
- OLSON, S.T.; BJORK, I.; SHORE, J.D. Kinetic characterization of heparin-catalyzed and uncatalyzed inhibition of blood-coagulation proteinases by antithrombin. *Meth. Enzymol.*, v. 222, p. 525-559 (1993).
- ORTIZ, A.R.; PISABARRO, M.T.; GAGO, F. *et al.* Prediction of Drug Binding Affinities by Comparative Binding Energy Analysis. *J. Med. Chem.*, v. 38, p. 2681-2691 (1995).
- OUELLETTE, T.; DESTRAU, S.; OUELLETTE, T.; *et al.* Production and purification of refolded recombinant human IL-7 from inclusion bodies. *Prot. Express. Purif.*, v. 30, p. 156-166 (2003).
- OUTCHKOUROV, N.S.; STIEKEMA, W.J.; JONGSMA, M.A. Optimization of the expression of equistatin in *Pichia pastoris*. *Prot. Exp. Purif.*, v. 24, p. 18-24 (2002).

- PADMANABHAN, K.; PADMANABHAN, K.P.; TULINSKY, A. *et al.* Structure of human des(1-45) factor Xa at 2.2Å resolution. *J. Mol. Biol.*, v. 232, p. 947-966 (1993).
- PAGE, M. J.; WONG, S-L.; HEWITT, J. *et al.* Engineering the primary specificity of *Streptomyces griseus* trypsin. *Biochemistry*, v. 42, p. 9060-9066 (2003).
- PAPASSOTIROPOULOS, A.; BAGLI, M.; KURZ, A. *et al.* A genetic variation of cathepsin D is a major risk factor for Alzheimer's disease. *Ann. Neurol.*, v. 47, p. 399-403 (2000).
- PARK, H.; SUH, J.; LEE, S. *Ab initio* on the catalytic mechanism of aspartic proteinases: nucleophilic versus general acid/general base mechanism. *J. Am. Chem. Soc.*, v. 122, p. 3901-3908 (2000).
- PARRY, M.A.A.; JACOB, U.; HUBER, R. *et al.* The crystal structure of the novel snake venom plasminogen activator TSV-PA: a prototype structure for snake venom serine proteinases. *Structure*, v. 6, p. 1195-1206 (1998).
- PATRUSHEV, L.I. Genetic mechanisms of hereditary hemostasis disorders. *Biochemistry (Moscow)*, v. 67, p. 33-46 (2002).
- PEARL, L.H.; TAYLOR, W.R. A structural model for the retroviral proteases. *Nature*, v. 329, p. 351-354 (1987).
- PEARL, L.; BLUNDELL, T. The active site of aspartic proteinases. *FEBS Lett.*, v. 174, p. 96-101, (1984).
- PERONA, J. J.; TSU, C. A.; McGRATH, M. E. *et al.* Relocating a negative charge in the binding pocket of trypsin. *J. Mol. Biol.*, v. 230, p. 934-949 (1993).
- PERONA, J.J.; CRAIK, C.S. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.*, v. 272, p. 29987-29990 (1997).
- PERONA, J.J.; CRAIK, C.S. Structural basis of substrate specificity in the serine proteases. *Prot. Sci.*, v. 4, p. 337-360 (1995).
- PETRESCU, A-J.; MILAC, A-L.; PETRESCU, S.M. *et al.* Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*, v. 14, p. 103-114 (2004).
- PEYVANDI, F.; CAREW, J.A.; PERRY, D.J. *et al.* Abnormal secretion and function of recombinant human factor VII as the result of modification to a calcium binding site caused by a 15-base pair insertion in the F7 gene. *Blood*, v. 97, p. 960-965 (2001).

- PIANA, S.; SEBASTIANI, D.; CARLONI, P. *Ab initio* molecular dynamics-based assignment of the protonation state of pepstatin A/HIV-1 protease cleavage site. *J. Am. Chem. Soc.*, v. 123, p. 8730-8737 (2001).
- PIMENTA, D.C.; OLIVEIRA, A.; JULIANO, M.A. *et al.* Substrate specificity of human cathepsin D using internally quenched fluorescent peptides derived from reactive site loop of kallistatin. *Biochim. Biophys. Acta*, v. 1544, p. 113-122 (2001).
- PINEDA, A.O.; CARRELL, C.J.; BUSH, L.A. *et al.* Molecular dissection of Na⁺ binding to thrombin. *J. Biol. Chem.*, v. 279, p. 31842-31853 (2004).
- PINEDA, A.O.; SAVVIDES, S.N.; WAKSMAN, G. *et al.* Crystal structure of the anticoagulant slow form of thrombin. *J. Biol. Chem.*, v. 277, p. 40177-40180 (2002).
- PINOTTI, M.; ETRO, D.; BINDINI, D. *et al.* Residual factor VII activity and different hemorrhagic phenotypes in CRM⁺ factor VII deficiencies (Gly331Ser and Gly283Ser). *Blood*, v. 99, p. 1495-1497 (2002).
- PIRKLE, H.; THEODOR, I.; MIYADA, D. *et al.* Thrombin-like enzyme from the venom of *Bitis gabonica*. Purification, properties and coagulant actions. *J. Biol. Chem.*, v. 261, p. 8830-8835 (1986).
- POLLASTRI, G.; BALDI, P. Prediction of Contact Maps by Recurrent Neural Network Architectures and Hidden Context Propagation from All Four Cardinal Corners. *Bioinformatics*, v. 18, p. S62-S70 (2002).
- POLLASTRI, G.; PRZYBYLSKI, D.; ROST, B. *et al.* Improving the Prediction of Protein Secondary Structure in Three and Eight Classes Using Recurrent Neural Networks and Profiles, *Proteins*, v. 47, p. 228-235 (2002).
- POWERS, J. C.; ASGIAN, J. L.; EKICI, O. D. *et al.* Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev.*, v. 102, p. 4639-4750.
- PROSDOCIMI, F.; FARIA-CAMPOS, A.C.; PEIXOTO, F.C. *et al.* Clustering of *Schistosoma mansoni* mRNA sequences and analysis of the most transcribed genes: implications in metabolism and biology of different developmental stages. *Mem. Inst. Oswaldo Cruz*, v. 97(S1), p. 61-69 (2002).
- RABELLO, A.; PONTES, L.A.; DIAS-NETO, E. Recent advances in the diagnosis of *Schistosoma* infection: the detection of parasite DNA. *Mem. Inst. Oswaldo Cruz*, v. 97, p. S171-172 (2002).

RAGAGOPOLAN, T.G.; STEIN, W.H.; MOORE, S. The inactivation of pepsin by diazoacetyl-norleucine methyl ester. *J. Biol. Chem.*, v. 241, p. 4295-4297 (1966).

RANG, H.P.; DALE, M.M.; RITTER, J.M. *Farmacologia*. 4a ed. Rio de Janeiro: Guanabara-Koogan, 2001, 703p.

RATHORE, D.; MCCUTCHAN, T.F.; SULLIVAN, M. *et al.* Antimalarial drugs: current status and new developments. *Exp. Opin. Invest. Drugs*, v. 14, p. 871-883 (2005).

RAUH, D.; KLEBE, G.; STUBBS, M.T. Understanding protein-ligand interactions: the price of protein flexibility. *J. Mol. Biol.*, v. 335, p. 1325-1341 (2004).

RAWLINGS N.D.; BARRET A.J. Families of serine peptidases. *Meth. Enzymol.*, v. 244, p. 19-61 (1994).

RAWLINGS, N.D.; BARRET, A.J. Evolutionary families of peptidases. *Biochem. J.*, v. 290, p. 205-218 (1993).

RAWLINGS, N.D.; BARRET, A.J. MEROPS: The Peptidase database. *Nucleic Acids Res.*, v. 28, p. 323-325 (2000).

RAWLINGS, N.D.; TOLLE, D.P.; BARRETT, A.J. MEROPS: the peptidase database. *Nucleic Acids Res.*, v. 32, p. D160-D164 (2004).

RECANTINI, M., KLEIN, T., YANG, C.-Z. *et al.* Quantitative structure-activity relationships and molecular graphics in ligand receptor interactions: amidine inhibition of trypsin. *Mol. Pharmacol.*, v. 29, p. 436-440 (1986).

REDMAN, C.A.; ROBERTSON, A.; FALLON, P.G. *et al.* Praziquantel: An urgent and exciting challenge. *Parasitol. Today*, v. 12, p. 14-20 (1996).

REICHARD, U.; COLE, G.T.; RUCHEL, R.; *et al.* Molecular cloning and targeted deletion of PEP2 which encodes a novel aspartic proteinase from *Aspergillus fumigatus* *Int. J. Med. Microbiol.*, v. 290, p. 85-96 (2000).

REICHARD, U.; EIFERT, H.; RÜCHEL, R. Purification and characterization of an extracellular aspartic proteinase from *Aspergillus fumigatus*. *J. Med. Vet. Mycol.*, v. 32, p. 427-436 (1994).

REY, L. *Parasitologia*, 2^a ed., Rio de Janeiro: Guanabara-Koogan, 1991, 731p.

- REYDA, S.; SOHN, C.; KLEBE, G. *et al.* Reconstructing the binding site of factor Xa in trypsin reveals ligand-induced structural plasticity. *J. Mol. Biol.*, v. 325, p. 963-977 (2003).
- REZAIE, A.R. Exosite-dependent regulation of the protein C anticoagulation pathway. *Trends Cardiovasc. Med.*, v. 13, p. 8-15 (2003).
- REZAIE, A.R. Identification of basic residues in the heparin-binding exosite of factor Xa critical for heparin and factor Va binding. *J. Biol. Chem.*, v. 275, p. 3320-3327 (2000).
- REZAIE, A.R. Role of residue 99 at the S2 subsite of factor Xa and activated protein C in enzyme specificity. *J. Biol. Chem.*, v. 271, p. 23807-23814 (1996).
- REZAIE, A.R.; ESMON, C.T. Asp-70 to Lys mutant of factor Xa lacks the high affinity Ca^{2+} binding site yet retains function. *J. Biol. Chem.*, v. 269, p. 21495-21499 (1994).
- REZAIE, A.R.; HE, X. Sodium binding site of factor Xa: role of sodium in the prothrombinase complex. *Biochemistry*, v. 39, 1817-1825 (2000).
- REZAIE, A.R.; KITTUR, F.S. The critical role of the 185-189 loop in the factor Xa interaction with Na^+ and factor V in the prothrombinase complex. *J. Biol. Chem.*, v. 279, p. 48262-48269 (2004).
- REZAIE, A.R.; MATHER, T.; SUSSMAN, F. *et al.* Mutation of Glu-80 \rightarrow Lys results in a protein C mutant that no longer requires Ca^{2+} for rapid activation by the thrombin-thrombomodulin complex. *J. Biol. Chem.*, v. 269, p. 3151-3154 (1994).
- RICH, D.H.; BERNATOWICS, M.S.; AGARWAL, N.S. Inhibition of aspartic proteases by pepstatin and 3-methylstatine derivatives of pepstatin - evidence for collected-substrate enzyme-inhibition. *Biochemistry*, v. 24, p. 3165 (1985).
- RICHTER, C.; TANAKA, T.; KOSEKI, T. *et al.* Contribution of a prosegment lysine residue to the function and structure of porcine pepsinogen A and its active form pepsin A. *Eur. J. Biochem.*, v. 261, p. 746 (1999).
- ROGANA, E.; PENHA-SILVA, N.; MARES-GUIA, M. The substituent effect on complex-formation between alpha-trypsin and para-substituted benzamidinium ions - a thermodynamic study. *Brazilian J. Med. Biol. Res.*, v. 22, p. 1177-1190 (1989).
- ROSE T.; DI CERA, E. Substrate recognition drives the evolution of serine proteases. *J. Biol. Chem.*, v. 277, p. 19243-19246 (2002).

- ROSENFELD, S.A. Use of *Pichia pastoris* for expression of recombinant proteins. *Meth. Enzymol.*, v. 306, p. 154-169 (1999).
- ROST, B.; SANDER, C. Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.*, v. 232, p. 584-99 (1993).
- ROTH, J. Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem. Rev.*, v. 102, p. 285-303 (2002).
- RUPPEL, A.; BRETERNITZ, U.; BURGER, R. Diagnostic Mr 31.000 *Schistosoma mansoni* proteins: requirement of infection, but not immunization, and use of the 'miniblot' technique for the production of monoclonal antibodies. *J. Helminthol.*, v. 61, p. 95-101 (1987).
- SACHDEV, D.; CHIRGWIN, J.M. Properties of soluble fusions between mammalian aspartic proteinases and bacterial maltose-binding protein. *J. Prot. Chem.*, v. 18, p. 127-136 (1999).
- SAGUCHI, K-I.; HAGIWARA-SAGUCHI, Y.; MURAYAMA, N. *et al.* Molecular cloning of serine proteinases from *Bothrops jararaca* venom gland. *Toxicon*, v. 46, p. 72-83 (2005).
- SAMBROOK, J.; FRITSCH, D.F.; MANIATIS, T. *Molecular Cloning: A laboratory manual*. 2nd Edition. New York: Cold Spring Harbor Press, 1989, p. 9.16-9.23.
- SATTLER, M.; SCHLEUCHER, J.; GRIESINGER, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nuc. Magnet. Reson. Spectr.*, v. 34, p. 93-158 (1999).
- SCHAAF, A.; RESKI, R.; DECKER, E.L. A novel aspartic proteinase is targeted to the secretory pathway and to the vacuole in the moss *Physcomitrella patens*. *Eur. J. Cell. Biol.*, v. 83, p. 145-152 (2004).
- SCHECHTER I & BERGER A On size of active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.*, v. 27, p. 157-162 (1967).
- SCHIMMOLLER, F.; HIGAKI, J.N.; CORDELL, B. Amyloid forming proteases: Therapeutic targets for Alzheimer's disease. *Curr. Pharm. Des.*, v. 8, p. 2521-2531 (2002).

SCHMIDT, A.E.; PADMANABHAN, K.; UNDERWOOD, M.C. *et al.* Thermodynamic linkage between the S1 site, the Na⁺ binding site, and the Ca²⁺ site in the protease Domain of human activated protein C (APC). Sodium ion in the APC crystal structure is coordinated to four carbonyl groups from two separate loops. *J. Biol. Chem.*, v. 277, p. 28987-28995 (2002).

SCHMIDT, A.E.; STEWART, J.E.; MATHUR, A. *et al.* Na⁺ site in blood coagulation factor IXa: effect on catalysis and factor VIIIa binding. *J. Mol. Biol.*, v. 350, p. 78-91 (2005).

SCHOEN, C.; REICHARD, U.; MONOD, M.; *et al.* Molecular cloning of an extracellular aspartic proteinase from *Rhizopus microsporus* and evidence for its expression during infection. *Med. Mycol.*, v. 40, p. 61-71 (2002).

SCHWEDE, T.; KOPP, J.; GUEX, N. *et al.* SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* v. 31, p. 3381-3385 (2003).

SERRANO, S.M.T.; HAGIWARA, Y.; MURAYAMA, N. *et al.* Purification and characterization of a kinin-releasing and fibrinogen-clotting serine proteinase (KN-BJ) from the venom of *Bothrops jararaca*, and molecular cloning and sequence analysis of its cDNA. *Eur. J. Biochem.*, v. 251, p. 845-853 (1998).

SERRANO, S.M.T.; MAROUN, R.C. Snake venom serine proteases: sequence homology vs. substrate specificity, a paradox to be solved. *Toxicon*, v. 45, p. 1115-1132 (2005).

SERRANO, S.M.T.; MATOS, M.F.C.; MANDELBAUM, F.R. *et al.* Basic proteinases from *Bothrops moojeni* (caissaca) venom—I. Isolation and activity of two serine proteinases, MSP 1 AND MSP 2, on synthetic substrates and on platelet aggregation. *Toxicon*, v. 31, p. 471-481 (1993).

SERRANO, S.M.T.; MENTELE, R.; SAMPAIO, C.A.M. *et al.* Purification, characterization and amino acid sequence of a serine proteinase, PA-BJ, with platelet-aggregating activity from the venom of *Bothrops jararaca*. *Biochemistry*, v. 34, p. 7186-7193 (1995).

SERRANO, S.M.T.; SAMPAIO, C.A.M.; MENTELE, R. *et al.* A novel fibrinogen-clotting enzyme, TL-BJ, from the venom of the snake *Bothrops jararaca*: purification and characterization. *Thromb. Haemost.*, v. 83, p. 438-444 (2000).

SERRANO, S.M.T.; SHANNON, J.D.; WANG, D. *et al.* A multifaceted analysis of viperid snake venoms by two-dimensional gel electrophoresis: an approach to understanding venom proteomics. *Proteomics*, v. 5, p. 501-510 (2005).

SHARP, K.A.; MADAM, B. Hydrophobic effect, water structure, and heat capacity changes. *J. Phys. Chem. B*, v. 101, p. 4343-4348 (1997).

SHEEHAN, J.P.; KOBBERVIG, C.E.; KIRKPATRICK, H.M. Heparin inhibits the intrinsic tenase complex by interacting with an exosite on factor IXa. *Biochemistry*, v. 42, p. 11316-11326 (2003).

SHETTY, B.V.; KOSA, M.B.; KHALIL, D.A. *et al.* Preclinical pharmacokinetics and distribution to tissue of AG1343, an inhibitor of human immunodeficiency virus type 1 protease. *Antimicrob. Agents Chemother.* v. 40, p. 110-114 (1996).

SHEVCHENKO, A.; WILM, M.; VORM, O. *et al.* Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical Chemistry*, v. 68, p. 850-858 (1996).

SIBLEY, C.H.; HUNT, S.Y. Drug resistance in parasites: can we stay ahead of the evolutionary curve? *Trends Parasitol.*, v. 19, p. 532-537 (2003).

SICHLER, K.; BANNER, D.W.; D'ARCY A. *et al.* Crystal structures of uninhibited factor VIIa link its cofactor and substrate-assisted activation to specific interactions. *J. Mol. Biol.*, v. 322, p. 591-603 (2002).

SICHLER, K.; HOPFNER, K-P.; KOPETZKI, E. *et al.* The influence of residue 190 in the S1 subsite of trypsin-like serine proteases on substrate selectivity is universally conserved. *FEBS Lett.*, v. 530, p. 220-224 (2002).

SIELECKI, A.R.; HAYAKAWA, K.; FUJINAGA, M. *et al.* Structure of recombinant human renin, a target for cardiovascular-active drugs, at 2.5 Å resolution. *Science*, v. 243, p. 1346-1350 (1989).

SILVA-JR, F.P.; VEYL, F.Z.; CLOS, J. *et al.* Molecular Modeling Approaches for Determining Gene Function: Application to a Putative Poly-A Binding Protein from *Leishmania amazonensis* (LaPABP). *Mem. Inst. Oswaldo Cruz*, v. 97, p. 335-341 (2002).

- SILVEIRA, A.M.; MAGALHÃES, A.; DINIZ, C.R. *et al.* Purification and properties of the thrombin-like enzyme from the venom of *Lachesis muta muta*. *Int. J. Biochem.*, v. 21, p. 863-871 (1989).
- SILVERMAN, G.A.; BIRD, P.I.; CARREL R.W. *et al.* The Serpins Are an Expanding
- SIMPSON, A.J.G.; SHER, A.; McCUTCHAN, T.F. Isolation of DNA, its size, bases and repetitive sequences. *Mol. Biochem. Parasitol.*, v. 6, p. 125-137 (1982).
- SINHA, S.; ANDERSON, J.P.; BARBOUR, R. *et al.* Purification and cloning of amyloid precursor protein beta-secretase from human brain. *Nature*, v. 402, p. 537-540 (1999).
- SIPOS, T.; MERCKEL, J.R. An effect of calcium ions on the activity, heat stability, and structure of trypsin. *Biochemistry*, v. 9, p. 2766-2775 (1970).
- SKELLY, P.J.; DA'DARA, A.; HARN, D.A. Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference. *Int. J. Parasitol.*, v. 33, p. 363-369 (2003).
- SMITH, A.M.; DALTON, J.P.; CLOUGH, K.A. *et al.* Adult *Schistosoma mansoni* express cathepsin L proteinase. *Mol. Biochem. Parasitol.*, v. 67, p. 11-19 (1994).
- SOEJIMA, K.; YUGUCHI, M.; MIZUGUCHI, J. *et al.* The 99 and 170 loop-modified factor VIIa mutants show enhanced catalytic activity without tissue factor. *J. Biol. Chem.*, v. 277, p. 49027-49035 (2002).
- SOGAWA, K.; FUJII-KURIYAMA, Y.; MIZUKAMI, Y. *et al.* Primary structure of human pepsinogen gene. *J. Biol. Chem.*, v. 258, p. 5306-5311 (1983).
- SOLARY, E.; EYMIN, B.; DROIN, N. *et al.* Proteases, proteolysis and apoptosis. *Cell. Biol. Toxicol.*, v. 14, p. 121-132 (1998).
- SOLMS, S.J.; GIULIANI, E.A.; GUARE, J.P. *et al.* Design and synthesis of HIV proteinase inhibitors. Variations of the carboxy terminus of the HIV protease inhibitor L-682-679. *J. Med. Chem.*, v. 34, p. 2852-2857 (1991).
- SOUTHAN, C. A genomic perspective on human proteases as drug targets. *Drug Discov. Today*, v. 6, p.681-688 (2001).
- SPIRO, R.G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, v. 12, p. 43R-56R (2002).

SREEKRISHNA, K.; BRANKAMP, R.G.; KROPP, K.E. *et al.* Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene*, v. 190, p. 55-62 (1997).

STOILOVA-McPHIE, S.; VILLOUTREIX, B.O.; MERTENS, K. *et al.* 3-Dimensional structure of membrane-bound coagulation factor VIII: modeling of the factor VIII heterodimer within a 3-dimensional density map derived by electron crystallography. *Blood*, v. 99, 1215-1223 (2002).

STONE, S.R.; HOFSTEENGE, J. Kinetics of the inhibition of thrombin by hirudin. *Biochemistry*, v. 25, p. 4622-4628 (1986).

STUBBS, M.T.; BODE, W. The clot thickens: clues provided by thrombin structure. *Trends. Biochem. Sci.*, v. 20, p. 23-28 (1995).

STURROK, R.F. Schistosomiasis epidemiology and control: how did we get here and where should we go. *Mem. Inst. Oswaldo Cruz*, v. 96, p. S17-S27 (2001).

SUGUNA, K.; BOTT, R.R.; PADLAN, E.A. *et al.* Structure and refinement at 1.8 Å resolution of the aspartic proteinase from *Rhizopus chinensis*. *Proc. Natl. Acad. Sci. USA*, v. 84, p. 7009-7013 (1987).

SULAHIAN, A.; GARIN, Y.J.F.; IZRI, A. *et al.* Development and evaluation of a western blot kit for diagnosis of schistosomiasis. *Clin. Diag. Lab. Immunol.*, v. 12, p. 548-551 (2005).

Superfamily of Structurally Similar but Functionally Diverse Proteins. *J. Biol. Chem.*, v. 276, p. 33293-33296 (2001).

SVERGUN, D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.*, v. 25, 495-503 (1992).

SVERGUN, D.I. Mathematical methods in small-angle scattering data analysis. *J. Appl. Cryst.* v. 24, p. 485-492 (1991).

SVERGUN, D.I.; PETOUKHOV, M.V.; KOCH, M.H.J. Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.*, v. 80, p. 2946-2953 (2001).

SVERGUN, D.I.; SEMENYUK, A.V.; FEIGIN, L.A. *Acta Cryst.* A44, 244-250 (1988).

SWENSON, S.; MARKLAND-Jr., F.S. Snake venom fibrin(ogen)olytic enzymes, *Toxicon*, v. 45, p. 1021-1039 (2005).

SYNTICHAKI, P.; XU, K.; DRISCOLL, M. *et al.* Specific aspartyl and calpain proteases are required for neurodegeneration in *C. elegans*. *Nature*, v. 419, p. 939-944 (2002).

TALHOUT, R.; ENGBERTS, J.B.F.N. Thermodynamic analysis of binding of p-substituted benzamidines to trypsin. *Eur. J. Biochem.*, v. 268, p. 1554-1560 (2001).

TANAKA, T.; YADA, R.Y. N-terminal portion acts as an initiator of the inactivation of pepsin at neutral pH. *Prot. Eng.*, v. 14, p. 669-674 (2001).

TANG, J.; JAMES, M.N.G.; HSU, N. *et al.* Structural evidence for gene duplication in the evolution of the acid proteases. *Nature*, v. 271, p. 618-21 (1978).

TANG, J.; WONG, R.N.S. Evolution in the structure and function of aspartic proteases. *J. Cell. Biochem.*, v. 33, p. 53-63 (1987).

THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F. *et al.* The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, v. 25, p. 4876-4882 (1997).

THOMPSON, W.J.; FITZGERALD, P.M.; HOLLOWAY, M.K. *et al.* Synthesis and antiviral activity of a series of HIV-1 protease inhibitors with functionality tethered to the P1 or P1' phenyl substituents: X-ray crystal structure assisted design. *J. Med. Chem.*, v. 35, p. 1685-701 (1992).

TIMASHEFF, S.N.; XIE, G. Preferential interactions of urea with lysozyme and their linkage to protein denaturation. *Biophys. Chem.*, v. 105, p. 421-448 (2003).

TIMMS, A.R. & BUEDING, E. Studies of a proteolytic enzyme from *Schistosoma mansoni*. *Br. J. Pharmacol.*, v. 14, p. 68-73 (1959).

TIMM, D.E. The crystal structure of the mouse glandular kallikrein-13 (prorenin converting enzyme). *Protein Sci.*, v. 6, p. 1418-1425 (1997).

TOH, H.; ONO, M.; SAIGO, K. *et al.* Retroviral protease-like sequence in the yeast transposon Ty1. *Nature*, v. 315, p. 691-692 (1985).

TRACY, J.W.; WEBSTER, L.T. *Drugs used in the chemotherapy of helminthiasis*. In GILMAN, A.G.; HARDMAN, J.G.; LIMBIRD, L.E. (Eds) Goodman & Gilman's The

pharmacological basis of therapeutics. 10th ed. New York: McGraw-Hill, 2001, p. 1121-1140.

TRACY, R.P. Thrombin, inflammation, and cardiovascular disease. An epidemiologic perspective. *CHEST*, v. 124, p. 49S-57S (2003).

TRAVIS, J.; POTEPA, J. Bacterial proteinases as targets for the development of second-generation antibiotics. *Biochim. Biophys. Acta.* v. 1477, p. 35-50 (2000).

TSIANG, M.; JAIN, A.K.; DUNN, K.E. *et al.* Functional mapping of the surface residues on human thrombin. *J. Biol. Chem.*, v. 270, p. 16854-16863 (1995).

TSUMOTO, K.; EJIMA, D.; KUMAGAI, I. *et al.* Practical considerations in refolding proteins from inclusion bodies. *Prot. Exp. Purif.*, v. 28, p. 1-8 (2003).

TYE, H. Application of statistical 'design of experiments' methods in drug discovery. *Drug. Discov. Today.*, v. 9, p. 485-491 (2004).

TYNDALL, J.D.A.; FAIRLIE, D.P. Conformational homogeneity in molecular recognition by proteolytic enzymes. *J. Mol. Recognit.*, v. 12, p. 363-370 (1999).

TYNDALL, J.D.A.; FAIRLIE, D.P. Macrocycles mimic the extended peptide conformation recognized by aspartic, serine, cysteine and metallo proteases. *Curr. Med. Chem.*, v. 8, p. 893-907 (2001).

UNDERWOOD, M.C.; ZHONG, D.; MATHUR, A. *et al.* Thermodynamic linkage between the S1 site, the Na⁺ site, and the Ca²⁺ site in the protease domain of human coagulation factor Xa. Studies on catalytic efficiency and inhibitor binding. *J. Biol. Chem.*, v. 275, p. 36876-36884 (2000).

UPRICHARD, J.; PERRY, D.J. Factor X deficiency. *Blood*, v. 16, p. 97-110 (2002).

VACCA, J.P. Design of tight-binding human immunodeficiency virus type 1 protease inhibitors. *Meth. Enzymol.*, v. 241, p. 311-334 (1994).

VACCA, J.P.; DORSEY, B.D.; SCHLEIF, W.A. *et al.* L-735,524: an orally bioavailable human immunodeficiency virus type 1 protease inhibitor. *Proc. Natl. Acad. Sci. USA.*, v. 91, p. 4096-4100 (1994).

VAN DAM, G.J.; WICHERS, J.H.; FERREIRA, T.M.F.; *et al.* Diagnosis of schistosomiasis by reagent strip test for detection of circulating cathodic antigen. *J. Clin. Microbiol.*, v. 42, p. 5458-5461 (2004).

VAN DEN HAZEL, H.B.; KILLAND-BRANDT, M.C.; WINTHER, J.R. Random substitution of large parts of the propeptide of yeast proteinase A. *J. Biol. Chem.*, v. 270, p. 8602-8607 (1995).

VAN DEN HAZEL, H.B.; KILLAND-BRANDT, M.C.; WINTHER, J.R. The propeptide is required for in vivo formation of stable active yeast proteinase A and can function even when not covalently linked to the mature region. *J. Biol. Chem.*, v. 268, p. 18002-18007 (1993).

VAN ETTEN, L.; FOLMAN, C.C.; EGGELTE, T.A. *et al.* Rapid diagnosis of schistosomiasis by antigen detection in urine with a reagent strip. *J. Clin. Microbiol.*, v. 32, p. 2404-2406 (1994).

VANZI, F.; MADAN, B.; SHARP, K. Effect of the protein denaturants urea and guanidinium on water structure: A structural and thermodynamic study. *J. Am. Chem. Soc.*, v. 120, p. 10748-10753 (1998).

VEERAPANDIAN, B.; COOPER, J.B.; SALI, A. *et al.* Direct observation by X-ray analysis of the tetrahedral "intermediate" of aspartic proteinases. *Prot. Sci.*, v. 1, p. 322 (1992).

VENEKEI, I.; SZILAGYI, L.; GRÁF, L. *et al.* Attempts to convert chymotrypsin to trypsin. *FEBS Lett.*, v. 379, p. 143-147 (1996).

VERITY, C.K.; LOUKAS, A.; McMANUS, D.P. *et al.* *Schistosoma japonicum* cathepsin D aspartic protease cleaves human IgG and other serum components. *Parasitology*, v. 122, p. 415-421 (2001).

VERITY, C.K.; McMANUS, D.P.; BRINDLEY, P.J. Developmental expression of cathepsin D aspartic protease in *Schistosoma japonicum*. *Int. J. Parasitol.*, v. 29, p. 1819-1824 (1999).

VERITY, C.K.; McMANUS, D.P.; BRINDLEY, P.J. Vaccine efficacy of recombinant cathepsin D aspartic protease from *Schistosoma japonicum*. *Parasite Immunol.*, v. 23, p. 153-162 (2001).

VERJOVSKI-ALMEIDA, S.; DeMARCO, R.; MARTINS, E.A.L. *et al.* Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. v. 35, p. 148-157 (2003).

- VIEIRA, D.F.; WATANABE, L.; SANT'ANA, C.D. *et al.* Purification and Characterization of Jararassin-I, A Thrombin-like Enzyme from *Bothrops jararaca* Snake Venom. *Acta Biochim. Biophys. Sinica*, v. 36, p. 798–802 (2004).
- VINDIGNI, A. Energetic dissection of specificity in serine proteases. *Comb. Chem. High Throughput Screen.*, v. 2, p. 139-153 (1999).
- VINDIGNI, A.; DANG, Q. D.; DI CERA, E. Site-specific dissection of substrate recognition by thrombin. *Nat. Biotechnol.*, v. 15, p. 891-895 (1997).
- VINDIGNI, A.; DI CERA, E. Release of fibrinopeptides by slow and fast forms of thrombin. *Biochemistry*, v. 35, p. 4417-4426 (1996).
- VOLKOVA, V.V.; SVERGUN, D.I. Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Cryst.*, v. 36, p. 860-864 (2003).
- VU, T.K.H.; WHEATON, V.I.; HUNG, D.T. *et al.* Domains defining thrombin receptor interactions. *Nature*, v. 353, p. 674-677 (1991).
- WALLQVIST, A.; COVELL, D.G.; THIRUMALAI, D. Hydrophobic interactions in aqueous urea solutions with implications for the mechanism of protein denaturation. *J. Am. Chem. Soc.*, v. 120, p. 427-428 (1998).
- WARKENTIN T.E. Bivalent direct thrombin inhibitors: hirudin and bivalirudin. *Best Pract. Res. Clin. Haematol.*, v. 17, p. 105-125 (2004).
- WASILEWSKI, M.M., LIM, K.C.; PHILLIPS, J. *et al.* Cysteine protease inhibitors block schistosome hemoglobin degradation *in vitro* and decrease worm burden and egg production *in vivo*. *Mol. Biochem. Parasitol.*, v. 81, p. 179-189 (1996).
- WATANABE, L.; VIEIRA, D.F.; BORTOLETO, R.K. *et al.* Crystallization of bothrombin, a fibrinogen-converting serine protease isolated from the venom of *Bothrops jararaca*. *Acta. Cryst. Sect. D.*, v. 58, p. 1036-1038 (2002).
- WEITZ, J.I.; BATES, S.M. New Anticoagulants. *J. Thromb. Haemost.*, v. 3, p. 1843–1853 (2005).
- WHITE, J. Snake venoms and coagulopathy. *Toxicon*, v. 45, p. 951-967 (2005).
- WHITE, PC; CORDEIRO, MC; ARNOLD, D; *et al.* Processing, activity, and inhibition of recombinant cyprosin, an aspartic proteinase from cardoon (*Cynara cardunculus*). *J. Biol. Chem.*, v. 274, p. 16685-16693 (1999).

WILLETT, W. S.; BRINEN, L. S.; FLETTERICK, R. J. *et al.* Delocalizing Trypsin Specificity with Metal Activation. *Biochemistry*, v. 35, p. 5992-5998 (1996).

WILLIAMS, S.A.; JOHNSTON, D.A. Helminth genome analysis: the current status of the filarial and schistosome genome projects. *Parasitology*, v. 118, p. S19-S38 (1999).

WITHERS-MARTINEZ, C.; CARPENTER, E.P.; HACKETT, F. *et al.* PCR-based gene synthesis as an efficient approach for expression of the A+T-rich malaria genome. *Prot. Eng.*, v. 12, p. 1113-1120 (1999).

WLODAVER, A.; MILLER, M.; JASKOLSKI, M. *et al.* Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science*, v. 245, p. 616-621 (1989).

WONG, J.Y.M.; HARROP, S.A.; DAY, S.R. *et al.* Schistosomes express two forms of cathepsin D. *Biochim. Biophys. Acta.*, v. 1338, p. 156-160 (1997).

WOO, J.H.; LIU, Y.Y.; MATHIAS, A. *et al.* Gene optimization is necessary to express a bivalent anti-human anti-T cell immunotoxin in *Pichia pastoris*. *Prot. Expr. Purif.*, v. 25, p. 270-282 (2002).

WORLD HEALTH ORGANIZATION. Prevention and control of schistosomiasis and soil-transmitted helminthiasis: report of a WHO expert committee. *WHO Tech. Rep. Ser.*, 912, WHO, Geneva.

WORLD HEALTH ORGANIZATION. Schistosomiasis. *WHO Fact Sheet*, v. 115, pp 14 (1996).

WORMALD, M.R.; PETRESCU, A.J.; PAO, Y.L. *et al.* Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modeling. *Chem. Rev.*, v. 102, p. 371-386 (2002).

WU, K.K.; THIAGARAJAN, P. Role of endothelium in thrombosis and hemostasis. *Annu. Rev. Med.*, v. 47, p. 315-331 (1996).

WÜTHRICH, K. Protein Structure Determination in Solution by NMR Spectroscopy. *J. Biol. Chem.* v. 265, p. 22059-22062 (1990).

XU, H.; BUSH, L.A.; PINEDA, A.O. *et al.* Thrombomodulin Changes the Molecular Surface of Interaction and the Rate of Complex Formation between Thrombin and Protein C. *J. Biol. Chem.*, v. 280, p. 7956-7961 (2005).

YAMADA, M.; AZUMA, T.; MATSUBA, T. *et al.* Secretion of human intracellular aspartic proteinase cathepsin-E expressed in the methylotrophic yeast, *Pichia Pastoris* and characterization of produced recombinant cathepsin-E. *Biochim. Biophys. Acta-Prot. Struct. Mol. Enzymol.*, v. 1206, p. 279-285 (1994).

YANG, L.; MANITBODY, C.; OLSON, S.T. *et al.* Contribution of basic residues of the autolysis loop to the substrate and inhibitor specificity of factor IXa. *J. Biol. Chem.*, v. 278, p. 25032-25038 (2003).

YANG, L.; PRASAD, S.; DI CERA, E. *et al.* The conformation of the activation peptide of protein C is influenced by Ca^{2+} and Na^{+} binding. *J. Biol. Chem.*, v. 279, p. 38519-38514 (2004).

YARLEQUE, A.; CAMPOS, S.; ESCOBAR, E. *et al.* Isolation and characterization of a fibrinogen-clotting enzyme from venom of the snake, *Lachesis muta muta* (Peruvian bushmaster). *Toxicon*, v. 27, p. 1189-1197 (1989).

YASAKA, M.; SAKATA, T.; NARITOMI, H. *et al.* Optimal dose of prothrombin complex concentrate for acute reversal of oral anticoagulation. *Thromb. Res.*, v. 115, p. 455-459 (2005).

YASUDA, Y.; KALETA, J.; BROMME, D. The role of cathepsins in osteoporosis and arthritis: Rationale for the design of new therapeutics. *Adv. Drug Deliv. Rev.*, v. 57, p. 973-993 (2005).

YOLKEN, R.H.; HILGARTNER, M.W. Prothrombin complex concentrates. Use in treatment of hemophiliacs with factor VIII inhibitors. *Am. J. Dis. Child.*, v. 132, p. 291-293 (1978).

YONEZAWA, M.; SUZUKI, J.; NISHIYAMA, M. *et al.* Role of the amino-terminal amino acid sequences determining the in vivo refolding process of prochymosin polypeptide. *J. Biotechnol.*, v. 28, p. 85-97 (1993).

YOSHIMASU, M.A.; AHN, J-K.; TANAKA, T. *et al.* Soluble expression and purification of porcine pepsinogen from *Pichia pastoris*. *Prot. Expr. Purif.*, v. 25, p. 229-236 (2002).

YOSHIMASU, M.A.; TANAKA, T.; AHN, J-K. *et al.* Effect of N-linked glycosylation on the aspartic proteinase porcine pepsin expressed from *Pichia pastoris*. *Glycobiol.*, v. 14, p. 417-429 (2004).

- YOU, W-K.; CHOI, W-S.; KOH, Y-S. *et al.* Functional characterization of recombinant batroxobin, a snake venom trombin-like enzyme, expressed from *Pichia pastoris*. *FEBS Lett.*, v. 571, p. 67-73 (2004).
- YOUSEF, G.M.; DIAMANDIS, E.P. An overview of the kallikrein gene families in humans and other species: Emerging candidate tumour markers. *Clinical Biochemistry*, v. 36, p. 443-452 (2003).
- YUAN, Q-P.; WALKE, E.N.; SHEEHAN, J.P. The factor IXa heparin-binding exosite is a cofactor interactive site: mechanism for antithrombin-independent inhibition of intrinsic tenase by heparin. *Biochemistry*, v. 44, p. 3615-3625 (2005).
- ZAGANELLI, G.L.; ZAGANELLI, M.G.M.; MAGALHÃES, A. *et al.* Purification and characterization of a fibrinogen-clotting enzyme from the venom of jararacuçu (*Bothrops jararacussu*). *Toxicon*, v. 34, p. 807-819 (1996).
- ZAMUDIO, K.R.; GREENE, H.W. Phylogeography of the bushmaster (*Lachesis muta*: Viperidae): implications for neotropical biogeography, systematics, and conservation *Biol. J. Linn. Soc.*, v. 62, p. 421-442 (1997).
- ZAVRSKI, I.; JAKOB, C.; SCHMID, P. *et al.* Proteasome: an emerging target for cancer therapy. *Anti-Cancer Drugs*, v. 16, p. 475-481 (2005).
- ZHANG, Y.; WISNER, A.; XIONG, Y. *et al.* A Novel Plasminogen Activator from Snake Venom. Purification, characterization, and molecular cloning. *J. Biol. Chem.*, v. 270, p. 10246-10255 (1995).
- ZHANG, E.; CHARLES, R.; TULINSKY, A. Structure of Extracellular Tissue Factor Complexed with Factor VIIa Inhibited with a BPTI Mutant. *J. Mol. Biol.*, v. 285, p. 2089-2104 (1999).
- ZHANG, E.; TULINSKY, A. The molecular environment of the Na⁺ binding site of thrombin. *Biophys. Chem.*, v. 63, p. 185-200 (1997).
- ZHANG, Y.; WISNER, A.; MAROUN, R.C. *et al.* *Trimeresurus stejnegeri* snake venom plasminogen activator. Site-directed mutagenesis and molecular modeling. *J. Biol. Chem.*, v. 272, p. 20531-20537 (1997).
- ZHU, Z.; LIANG, Z.; ZHANG, T. *et al.* Crystal structure and amidolytic activities of two glycolylated snake venom serine proteinases. *J. Biol. Chem.*, v. 280, p. 10524-10529 (2005).

ZUSSMAN, R.A.; BAUMAN, P.M. Schistosome hemoglobin protease; search for inhibitors. *J. Parasitol.*, v. 57, p. 233-234 (1971).

ZUSSMAN, R.A.; BAUMAN, P.M.; PETRUSKA, J.C. The role of ingested hemoglobin in the nutrition of *Schistosoma mansoni*. *J. Parasitol.*, v. 56, p. 75-9 (1970).

Exploring the subsite specificity of *Schistosoma mansoni* aspartyl hemoglobinase through comparative molecular modelling

F.P. Silva Jr.^a, F. Ribeiro^b, N. Katz^b, S. Giovanni-De-Simone^{a,c,*}

^aLaboratório de Bioquímica de Proteínas e Peptídeos, Departamento de Bioquímica e Biologia Molecular, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro-RJ, Brazil

^bCentro de Pesquisas René Rachou, FIOCRUZ, Belo Horizonte-MG, Brazil

^cDepartamento de Biologia Celular e Molecular, Instituto de Biologia, UFF, Niterói-RJ, Brazil

Received 4 October 2001; revised 21 December 2001; accepted 21 December 2001

First published online 15 February 2002

Edited by Gunnar von Heijne

Abstract Blood flukes of the genus *Schistosoma* currently infect millions of people in tropical and subtropical countries. An enzyme playing a major role in hemoglobin (Hb) degradation by *Schistosoma mansoni* has been cloned and shown to be highly similar to the human cathepsin D aspartyl proteinase, although presenting a distinct substrate specificity from the latter. Investigating the structural features responsible for this difference has a major application in the design of selective anti-schistosomal drugs. In order to achieve this goal a homology model for the *S. mansoni* aspartyl hemoglobinase was constructed and then used to simulate the complexes formed with two transition state analogues of Hb-derived octapeptide substrates. Comparison with human cathepsin D showed that different pocket volumes and surface electrostatic potentials arise from substitutions in residues comprising the S4, S3, S2 and S3' subsites. Since the primary specificity of the *S. mansoni* enzyme resembles that of HIV-1 protease, we have discussed the applicability of current retroviral protease inhibitors as leads for the design of new anti-schistosomal drugs. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Schistosomiasis; Aspartyl hemoglobinase; Specificity; Cathepsin D; *Schistosoma mansoni*

1. Introduction

Schistosomiasis is a major human parasitic disease. An estimate from the World Health Organization in 1993 [1] indicates that the blood flukes of the genus *Schistosoma* may infect over 200 million people in more than 74 countries throughout the world. Schistosome infestation is currently both treatable and preventable. The former is being achieved by oral administration of praziquantel or oxamniquine and the later by health educational programs. Nevertheless, there is no apparent sign of improvement in schistosomiasis morbidity indexes in most developing countries, possibly due to a premature increased tolerance to praziquantel [2]. Thus, seeking novel strategies for schistosomiasis chemotherapy is of major importance.

Schistosomes as any other blood-feeding parasite depend

on hemoglobin (Hb) catabolism for survival. Several already characterized or putative Hb degrading enzymes have been isolated from protozoans and helminths, including: *Plasmodium* sp. [3–6], *Ancylostoma caninum* [7], *Strongyloides stercoralis* [8] and *Haemonchus contortus* [9]. These proteolytic enzymes have been shown to belong either to the cysteinyl or to the aspartyl families of peptidases [10]. In schistosomes, a cathepsin D-like proteinase from both *Schistosoma mansoni* and *Schistosoma japonicum* species has been cloned [11,12]. Human cathepsin D is an aspartyl protease that has been assumed to be involved in the generation of peptides for antigen presentation [13], cancer metastasis [14] as well as in the formation of the plaques encountered in the Alzheimer's disease [15]. Today, a lot of knowledge has been accumulated from structural studies of aspartyl proteinases involved in the pathogeny of important human diseases [16]. Nevertheless, where no experimental information on the structure of the studied protease is available, molecular modelling plays a crucial role on the prediction of structure–function relationships [17–20]. The value of these computational techniques has been proved by the subsequent confirmation of numerous predictions through comparison with the structures resolved by nuclear magnetic resonance techniques or, in a greater extent, by X-ray crystallography [21–23].

Brindley et al. [24] have shown that the *S. mansoni* enzyme presents a discrete substrate preference from its orthologous host enzyme, cleaving human Hb in several distinct sites. Therefore, exploring this difference has a potential application on the design of selective inhibitors of *S. mansoni* cathepsin D-like hemoglobinase, which can represent a novel strategy for the combat of schistosomiasis. Noticeably, Brinkworth et al. [25] have recently focused the structural features responsible for the α -chain Phe36–Pro37 cleavage site on human Hb. This is a very unusual site for an eukaryotic aspartyl proteinase since it was essentially only observed in the aspartyl proteases of retroviral origin [16]. Pursuing the Phe–Pro cleavage site is a very interesting approach for substrate-based inhibitor design because it can furnish selectivity over eukaryotic enzymes. Nevertheless, addressing the general topological, electrostatic and steric features responsible for the discrete substrate preference shown by the schistosome and human enzymes should be of greater interest for broad-spectrum projects in structure-based drug design. Following this approach, a homology model for the *S. mansoni* aspartyl hemoglobinase has been constructed and used to simulate the complexes formed with two transition state analogues of

*Corresponding author. Fax: (55)21-25903495.

E-mail address: dsimone@ioc.fiocruz.br (S. Giovanni-De-Simone).

Abbreviations: 3D, three dimensional

Hb-derived octapeptide substrates (α L105–L106/SHCL* I.VTI. and β L31–L32/I.GRI* I.VVY). The modelled complexes were used to pinpoint the structural discrepancies found in the subsites of the schistosoma and human cathepsin D enzymes that could be correlated to the enzymatic activities and binding affinities reported in the literature for these proteases. In addition, in view of the similarity at the substrate preference level of the aspartyl proteases of schistosoma and HIV-1, we have discussed the feasibility of current HIV-1 protease inhibitors as leads for the design of novel anti-schistosomal drugs.

2. Materials and methods

2.1. Sequence analysis

The sequence of the mature *S. mansoni* cathepsin D-like protease (CatD_Sman), comprising 377 residues, was extracted from the zymogen sequence deposited in TrEMBL protein databank under accession number P91802. Aspartyl proteases found on the complexes of cathepsin D-pepstatin (PDB ID 1LYB, 2.50 Å resolution) [21], mouse renin CH-66 (PDB ID 1SMR, 2.00 Å resolution) [22] and pepsin 3A-phosphonate inhibitor (PDB ID 1QRP, 1.96 Å resolution) [26] were used as templates in the modelling procedure. A reliable alignment between CatD_Sman and its templates was achieved by adjusting the multiple sequence alignments obtained either from the pepsin A1 family in MEROPS database [27] or from ClustalW v. 1.72 [28], in order to maximize the placement of insertions and deletions in loops and to align features conserved in the pepsin family of proteases, such as the three pairs of cysteine residues. This alignment was also used as an input to the program MEGA v. 2.1 [29] to calculate the pairwise distances (using the Poisson correction model and a pairwise deletion of gap containing segments), which were used to generate a phenogram by the neighbor-joining method. The reliability of each internal branch was assessed by the bootstrap method implemented in MEGA.

2.2. Comparative molecular modelling

The sequence of CatD_Sman was folded onto the superposed structures of its templates according to the optimal alignment, obtained as explained above, using the SwissPDB Viewer v. 3.6 program and then submitted in optimize mode to the Swiss-Model automated modelling server [30]. This modelling method has been rigorously assessed by a large-scale initiative called the 3DCrunch project (http://www.expasy.ch/swissmod/SM_3DCrunch.html). Briefly, models whose sequences shared around 50% identity to templates deviate from experimental structures by approximately 1 Å relative mean square deviation in the common core. Suitability of the modelled CatD_Sman three dimensional (3D) structure for this study was assessed with the analytical tools available in SwissPDB Viewer v. 3.6 and with the programs in the PROCHECK suite [31]. Residues corresponding to the carboxy-terminal extension found in *S. mansoni* aspartyl protease were not considered in the modelling because this segment showed to have no counterpart on protein structures determined to date. As no function has been assigned to this region yet and as schistosomes are believed to express both forms of the enzyme, with and without this segment [11], we decided not to address this question here.

2.3. Modelling of active site residues

The geometry of the residues composing the catalytic triad in the active site is highly conserved among the aspartyl proteases [16]. Following this principle, we have applied a manual torsion on the χ^2 angle of D33 and D219 side chains, taking as a guide the catalytic aspartate side chains from human cathepsin D structure [21], in order to guarantee that both groups belong to the plane observed in aspartyl proteases. Finally, 20 cycles of steepest descent and 20 cycles of conjugate gradient energy minimization with the Gromos96 force field [32] implemented in SwissPDB Viewer v. 3.6 were applied to release any problems with stereochemistry that would have been imposed by the manual torsion procedure.

2.4. Substrate modelling

The 3D structure of mouse renin complexed with the decapeptide inhibitor CH-66 was used as the archetype to model the transition

state isoster of the pseudo-substrates LGRLW(CH(OH)CH₂)LVVY and SHCLW(CH(OH)CH₂)LVYL. CH-66 was chosen because: (i) it almost perfectly resembles the transition state intermediate formed during the cleavage of a peptide bond and (ii) it would be large enough to furnish the extended backbone necessary to model an octapeptide. Substrate modelling was achieved by superimposing the CatD_Sman model on the enzyme inhibitor complex of mouse renin (0.70 Å in 1264 backbone atoms) and then transferring the coordinates of the inhibitor onto the model structure. Finally, collisions and improper stereochemistry of the mutated side chains in the inhibitor enzyme complex were released by 20 cycles of steepest descent and 20 cycles of conjugate gradient energy minimization with the Gromos96 force field. Hydrogen bonds and non-bonded contacts between the modelled transition state isosteres of Hb-derived substrates and CatD_Sman model were calculated with HBPLUS [33].

3. Results and discussion

3.1. Overall model structure and validation

The structure of CatD_Sman, as expected, resembles that of most eukaryotic aspartyl proteinases (Fig. 2A). A six-stranded

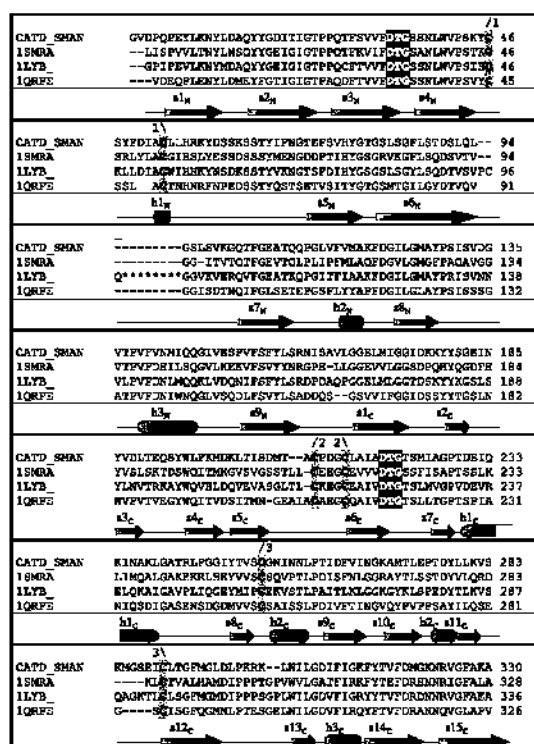


Fig. 1. Multiple sequence alignment used on the construction of *S. mansoni* aspartyl hemoglobinase model. CatD_Sman *S. mansoni* aspartyl hemoglobinase; 1SMRA mouse renin; 1LYB human cathepsin D and 1QRF5 human pepsin 3A. The black arrowhead indicates the end of cathepsin D light chain. Asterisks in 1LYB sequence correspond to residues excised during post-translational processing of human cathepsin D. The three conserved cysteine pairs are highlighted in gray while the two catalytic triads are highlighted in black. In the bottom of each alignment block is represented the secondary structure of CatD_Sman derived from its model (coded to facilitate discussion in the text). Secondary structure elements are represented as cylinders (α -helices) and arrows (β -strands). The first two residues in the CatD_Sman sequence (shown in grayish) were not modelled neither considered in the numeration.



Table 1
Hydrogen bonds between residues mapped in the S4–S4' subsites of the CatD_Sman model and the atoms of the substrates transition state analogues

Enzyme residue and atom	Substrate analogue	Substrate analogue moiety	Distance (Å)	
S223	O γ	α L105 L106	His (P3) N	2.99
	N	β L31 L32	Gly (P3) N	2.92
T80	O γ^1	α L105 L106	His (P3) O	2.90
		β L31 L32	Gly (P3) O	2.80
	N	α L105 L106	Cys (P2) N	2.62
		β L31 L32	Arg (P2) N	2.49
G79	N	α L105 L106	Cys (P2) O	3.31
		β L31 L32	Arg (P2) O	2.90
	O δ^1	α L105 L106	Cys (P2) O	3.08
		β L31 L32	Arg (P2) O	3.13
D33	O δ^2	α L105 L106	Leu (P1') O	3.31
	O δ^2	β L31 L32	Leu (P1') O	3.31
D219	O δ^2	α L105 L106	Leu (P1) O	2.85
		β L31 L32	Leu (P1) O	2.94
G221	O	α L105 L106	Leu (P1) O	2.38
		β L31 L32	Leu (P1) N	3.10
G35	O	α L105 L106	Leu (P1) N	3.07
		β L31 L32	Val (P2') N	2.68
H77	O	β L31 L32	Val (P2') N	2.73
		β L31 L32	Val (P3') N	3.27
K301	O	β L31 L32	Tyr (P4') OH	2.75

for the whole CatD_Sman model were inside the expected values.

3.2. Subsite specificity and comparison with human cathepsin D active site

Brindley et al. [24] have observed that among the 17 Hb sites cleaved by the schistosome aspartyl proteases, 12 P1 residues and six P1' residues were either Leu or Phe. In addition, two out of the three initial cleavage sites had a Leu residue either on P1 or P1' positions. Considering that this markedly preference must be a result of a higher number of favorable contacts between substrate moieties and the active site residues atoms we have modelled two complexes between the CatD_Sman model and the transition state analogues of the peptides representing positions P4–P4' in the cleavage sites α L105–L106 (SHCL*LVTL; Fig. 2B) and β L31–L32 (LGRL*LVVY) from α - and β -chains, respectively, of human

Hb. These substrates could be appropriately modelled from the decapeptide renin inhibitor CH-66, which contains a Leu- ψ (CHOHCH₂)-Leu moiety, since a very tight superposition of the CatD_Sman model and mouse renin could be achieved. This procedure allowed us to define the enzyme residues corresponding to the subsite pockets S4–S4' (Schechter and Berger's nomenclature [34]). For all of the subsites studied, with exception of S2, several hydrogen bonds could be detected between substrate backbone atoms and enzyme moieties (Table 1). Nevertheless, not all hydrogen bonds predicted between the O δ^1 and O δ^2 atoms from the catalytic aspartates and the P1-Leu carbonyl oxygen (as detected in the mouse renin-CH-66 complex [22]) were detected in CatD_Sman and its modelled substrates. This observation is probably explained by the existence of a slight difference in the rotameric position of the aspartate side chains in the crystallized mouse renin structure in comparison to the CatD_Sman model.

Table 2
Comparison between subsite composition of *S. mansoni* and human cathepsin D proteases^a

Subsite	<i>S. mansoni</i>	Human	Subsite	<i>S. mansoni</i>	Human
S4	S223	S235	S4'	K301	P314
	M224	L236		R302	P313
	M295	M307			
	<i>L291</i>	L303			
S3	Q14	Q14	S3'	H77	H77
	T80	S80		R302	P313
	F120	F131		<i>L299</i>	I311
	<i>F115</i>	F126			
S2	T222	T234	S2'	G35	G35
	A226	V238		S36	S36
	L297	M309		H77	H77
	I306	I320		I131	I142
	<i>I217</i>	I229		<i>Y194</i>	Y205
S1	V31	V31	S1'	Y78	Y78
	D33	D33		G79	G79
	Y78	Y78		Y194	Y205
	G79	G79		I217	I229
	I123	I134		D219	D231
	G221	G233		L299	<i>I317</i>

^aHighlighted in bold are residues mutated in *S. mansoni* protease in relation to cathepsin D. Detached in italic are the residues from CatD_Sman not detected in the stipulated contact range of 4.0 Å.

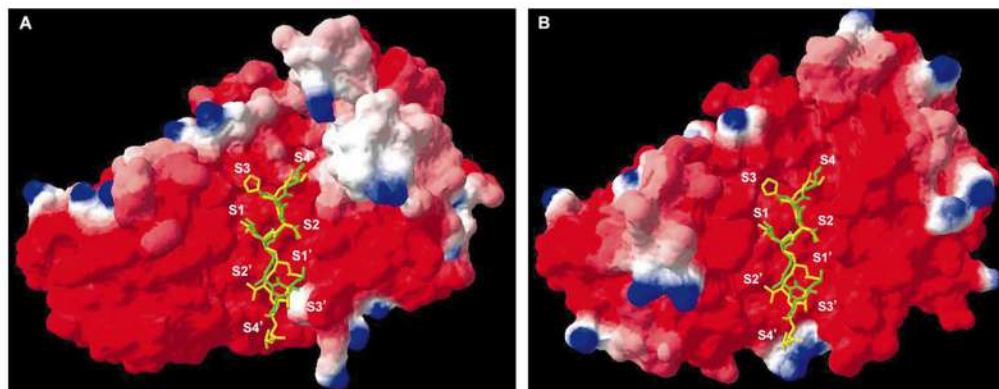


Fig. 4. Comparison between the electrostatic potential at the active site cleft surfaces of the CatD_Sman model (A) and the human cathepsin D structure (B). Subsite pockets are indicated in each structure. The surface corresponding to the flap and the second N-terminal helix are not shown for a better visualization of the substrate binding groove.

The first step in validating a parasitic protease as a target for chemotherapy should be the study on how to selectively block its activity over the host orthologous enzyme. This can be effectively achieved by correlating the differences found in substrate preference of the parasitic and human proteases with the structures in the active site of these enzymes. *S. mansoni* aspartyl hemoglobinase shows only moderated similarity to most human aspartyl proteases (including cathepsin E, renin and pepsin). Nevertheless, it is significantly similar to the lysosomal human enzyme cathepsin D (71% of similarity in 240 residues from the mature enzymes). Indeed it can be assumed from the phenogram shown in Fig. 3 that CatD_Sman is clearly closer to human cathepsin D than to any other host protease. The tree shown in Fig. 3 was rooted with HIV-1 proteinase monomer sequence, since it would better represent the ancestral gene that has given rise to the known eukaryotic aspartyl proteases [35]. This close relationship between the schistosome protease and human cathepsin D led us to focus on the comparison between the residue composition and the electrostatic properties on the surfaces of the subsite pockets from these enzymes. In Table 2 are listed the residues occupying homologous positions in the S4–S4' subsites, according to a structural superimposition of the human and schistosome protease. Differences were encountered in the residue composition as well as in the electrostatic properties on the surfaces of the S4, S3, S2 and S3' pockets.

The S2 and S4 subsites are fundamentally formed by the same molecular framework, where a clear distinction between the individual pockets was difficult to define. These pockets are formed by the β -turn between the $s6_C$ and $s7_C$ β -strands (spanning residues S223–G227, as defined in Fig. 1) and by the $s12_C$ strand (comprising residues C290–L297). Although a similar analysis by Brinkworth et al. [25] has classified this pocket as essentially hydrophobic, we found that this subsite architecture results in a shallow and amphipathic S4 pocket, which is similar to that described for human cathepsin D [21]. Mutation of L236 to M224 in the parasite enzyme provides a hydrophobic environment which can result in a more favorably interaction with a short hydrophobic side chain in opposition to the human enzyme. This could explain why a serine can be found among the residues occupying the P4 position in the sites cleaved in bovine Hb by the human cathepsin D [36]

while this amino acid never occurred in the same position of human Hb sites cleaved by the schistosome enzyme. Important discrepancies between the parasitic and human enzymes could be encountered in the S2 subsite. In the human enzyme, A226 is substituted for V238 in the loop between $s6_C$ and $s7_C$ strands and L297 is substituted for M309 in the tip of the $s12_C$ β -strand. The large diversity in the kind of side chain favored in the large and hydrophobic S2 subsite of human cathepsin D poses further difficulty in interpreting the structure–activity relationships of the same pocket in the schistosome enzyme. Gulnick et al. [37] have observed that the human enzyme more effectively cleaves substrates having sulfur-containing residues such as methionine or cysteine in the P2 position. However, Scarborough and Dunn [38] have shown that the wild-type human enzyme bearing M287 (M309 according to the numeration in Fig. 1) displays a preference for hydrophobic residues such as Leu while moderately accommodating polar side chains such as Glu in the P2 position of

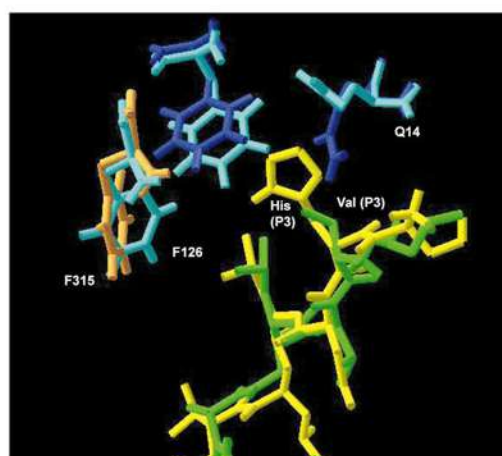


Fig. 5. Superimposition of S3 residues of CatD_Sman (cyan) and human cathepsin D (blue). Q14 assumes different conformations in both enzymes. F115 (orange) in the parasitic enzyme is in a different orientation from that of the correspondent F126 residue in the human protease.

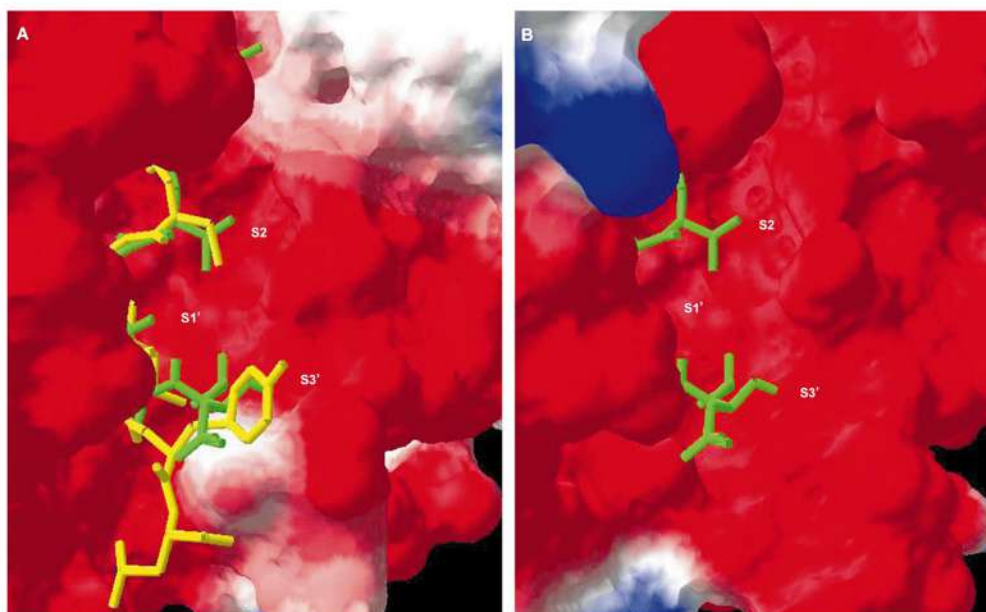


Fig. 6. Comparison of the surfaces at the active site C-terminal boundaries in CatD_Sman model (A) and human cathepsin D structure (B). A: The substrate analogue comprising α L105–L106 cleavage site of human Hb (SHCL ψ [CH(OH)CH₂]LVTL; yellow) is shown superimposed on the statin-based inhibitor of human cathepsin D (Iva-VV-Sta- ψ [CH(OH)CH₂OH]GA-Sta; green [21]). B: Only the statin-based inhibitor is shown docked at the human cathepsin D active site cleft. Iva = isovaline and Sta = statin.

substrates. Thus, large hydrophobic (Met or Leu), medium polar (cysteine) as well as large positively charged (Glu) residues can fit the S2 pocket of human cathepsin D. According to the mutations observed in the S2 pocket of the parasitic enzyme we can only propose that this subsite being even larger would favor the binding of large hydrophobic groups. In addition, examining the surface of both enzymes, two major differences that can affect S4 and S2 subsites are encountered (Fig. 4). First, a protuberance in the C-terminal border is clearly differentiable in the human enzyme, which can reflect in less conformational freedom in the binding of the P4 and P2 residues of the substrate. This difference is caused by the mutation of G247 for Q258 in the human enzyme. Second, there is a considerable hydrophobic patch near the S4 pocket surface in CatD_Sman that is absent from the human enzyme, which may also favor the binding of hydrophobic residues in this subsite. In the flap position above the S3 subsite region, S80 is mutated for T80 in the parasite enzyme. This change seemed to have no major effect on the hydrogen-bonding pattern observed between the flap and the substrate backbone.

Although much information can be obtained directly by comparison of the primary sequence of the two enzymes, certain observations can only be made using 3D models. For instance, as depicted in Fig. 5, residue Q14 in the S3 subsite, in spite of being conserved, assumes different conformations in both enzymes. According to our model, the rotamer for Q14 in the parasitic protease has its side chain pointed outside the pocket, while in the human enzyme it is directed toward the binding cleft. This would enable bulkier side chains, such as that of histidine in the α L105–L106 substrate, occupy the S3 pocket (yellow in Fig. 5) where only a short side chain, such as that of valine in the P3 position of the statin-based

inhibitor of human cathepsin D [21], would fit (green in Fig. 5). Indeed, carefully analysis of the cleavage sites in human Hb suggests that residues with longer side chains, such as arginine or glutamine, are preferred by the parasitic enzyme in the P3 position. On the other hand, it has recently been shown that the human enzyme attacks with greater efficiency fluorogenic substrates bearing the residues leucine, valine or glutamic acid [39] in this position. We also detected that F115 (orange in Fig. 5) in the parasitic enzyme is in a different orientation from that of the homologous F126 residue in the human protease. This possibly accounts for the fact that F115 was not initially detected in the set of residues within 4.0 Å from substrate atoms. Other residues have been substituted between the two enzymes in the loop connecting s7_N and h2_N (F115–I123) making up the framework of the S3 pocket, as noted before by Wong et al. [11], but none of these should affect interactions with P3 side chains of substrates.

None of the residues that are near to substrate atoms differs in S2' subsite of both schistosome and human proteases. However, mutations in residues surrounding I131 in the loop connecting the s8_N strand and the h2_N helix (S130 for R141, D134 for N145 and G135 for N146) would turn the pocket in the mammalian enzyme more hydrophilic than in the parasitic homologue. It has been proposed before that the hydrophobic S2' pocket in human enzyme can accommodate polar or charged residues by enabling the polar group to point out of the cleft and be stabilized by the solvent [40]. Indeed, inspection of the residues favored in P2' position of protein substrates cleaved by human cathepsin D characterizes a preference for short and polar residues.

A special feature of the structures of renin and cathepsin D is the presence of a proline-rich loop (also referred to as the

'rigid loop') connecting the $s11_C$ and $s12_C$ β -strands forming the border of the S3' and S4' subsites. The equivalent structure in the CatD_Sman is formed by a sequence of three basic residues K301, R302 and K303, for this reason being hereafter referred to by us as the 'basic loop'. The proline-rich loop in human cathepsin D is displaced by 2.7 Å away from the center of the binding groove in relation to the basic loop in CatD_Sman. An immediate consequence of this difference can be noted as a protuberance in the surface of the CatD_Sman model, which would leave much less space for accommodating large side chains in these pockets (Fig. 6A). A good example of how this can affect the binding of substrates/inhibitors can be found in the backbone torsion within the statin-based inhibitor of human cathepsin D responsible for making P3' pointing toward S1' in the complex [21]. Comparison of the surfaces of the complexes on the S3' binding pockets of both enzymes (Fig. 6A,B) shows that it would be impossible for an inhibitor of *S. mansoni* aspartyl protease to suffer this backbone torsion because of the steric blockage that is imposed by the basic loop. The complex between CatD_Sman and the modelled substrate (α L105–L106 cleavage site of IIb) shows that the bulk Tyr side chain in P3' would have to be torsioned out of the binding pocket in order to be accommodated in S3' (Fig. 6A). Based on the precedent discussion we propose that although larger side chains can occupy P3' position in CatD_Sman substrates, a favorable fit would only be achieved by short and polar residues. Again, our conclusions disagree with the observations made by Brinkworth et al. [25] that considered the S3' pocket as markedly hydrophobic. Analysis of interactions within the S4' pocket of CatD_Sman was made difficult by the fact that: (i) P4' in modelled substrates was positioned in the limit of the active site N-terminal half or prime region (adopting the nomenclature of Schechter and Berger [34]) and (ii) a direct comparison with the human cathepsin D complexed inhibitor was impossible in view of its particular binding mode as discussed before.

S1 and S1' subsite structures of both schistosome and human enzymes were very similar. They showed an almost perfect superimposition of the strands forming the ψ -like structure in the active site of several aspartyl proteinases and all of the residues within 4.0 Å from the atoms of the modelled substrates were conserved between the two proteases. Nevertheless, several sites in human Hb α -chain are cleaved by the schistosome enzyme but are not recognized by the human protease, based on the orthologous sites in bovine Hb. Among these sites, α Phe36–Pro37 deserves special attention because it represents a very unusual primary specificity for an eukaryotic aspartyl proteinase. Brinkworth et al. [25] have proposed that a rationale for this would be that when Pro is at P1' the backbone atom locations in P2, P1 and P1' are changed, which in human cathepsin D would produce a destabilization because of the lack of important hydrophobic contacts between Phe at P1 and residues in the S1 pocket. By careful inspection of the models generated we have observed that the substitution of Ala preceding D219 in CatD_Sman by Val in the human protease would make the S1' pocket of the later smaller and more hydrophobic. The larger S1' pocket in CatD_Sman would collaborate to a better accommodation of the rigid proline cycle at this position and thus help Phe at P1 to find a more stable conformation in S1.

Cleavage sites bearing Phe–Pro at P1–P1' positions have

only been efficiently hydrolyzed by retroviral aspartyl proteinases [41] and it has been suggested that the selectivity evidenced by HIV-1 protease is the consequence of an extensive network of interactions, not only within the S4–S3' subsites, but also extending beyond this segment [41]. This pattern would indicate that only a more rigorous and systematic study involving substitutions in P4–P4' positions of synthetic substrates will be able to identify any subtle differences existent in S1–S1' subsites from the schistosome and human proteases. The exclusive specificity of retroviral aspartyl proteinases has been extensively explored, as it became the first approach to reach selectivity within the anti-HIV-1 drugs [40]. Although successful in finding selective peptidic leads, this approach has fallen short in designing effective drugs mainly because of the low bioavailability of the designed peptidomimetics compounds [41]. The two compounds designed following this concept that have reached the market are the Hoffmann-La Roche RO318959 (Saquinavir) and the Merck L-735,524 (Indinavir). Most recent approaches that have taken full advantage on the 3D structure of the HIV-1 protease have been more productive in finding anti-viral drugs [41,42]. At this time three drugs were successfully designed in this way: the Vertex VX478 (Amprenavir), the Agouron AG-1343 (Nelfinavir) and the Pharmacia-Upjohn PNU-140690 (Tipranavir). The approach followed in this work, in which the general topological and chemical features responsible for substrate recognition by the schistosome and human orthologous enzyme were compared, would facilitate this type of drug design strategy.

The abnormal similarity with the HIV-1 protease can be explored in the design of inhibitors of the schistosome aspartyl hemoglobinase through the *in vitro* or *in silico* screening of known HIV-1 protease inhibitors. Nevertheless, besides the affinity for substrates bearing a hydrophobic Pro motif as P1–P1' residues, a number of studies have pointed out the essentially hydrophobic S2 and S2' subsites as the most critical in defining effective substrate interactions with the HIV-1 protease [43]. Thus, many substrate-based HIV-1 protease inhibitors reflect these requirements [44]. For example, the *cis*-decahydroisoquinoline moiety found in Saquinavir and Nelfinavir mimics the proline at P1' as well as it provides some interactions at the S3' binding site, while the bulky *t*-butylamide fills S2'. The general hydrophilic character of its S2' binding pocket would lower the binding affinity of such inhibitors to the schistosome enzyme but it would also enhance the selectivity for the parasitic enzyme over human cathepsin D, since this subsite in the human orthologue tends to be even more hydrophilic. As discussed above, S2 subsites on both the parasitic and human enzymes are hydrophobic. However, this pocket seems to be narrower in the schistosome protease, which would favor HIV-1 protease inhibitors with shorter P2 groups.

Acknowledgements: This work has been partially supported by the CNPq, Faperj, Fapemig and FioCruz (Papes). F.P.S.Jr. was a PIBIC CNPq/FioCruz fellowship recipient.

References

- [1] World Health Organization (1993) WHO Technical Report 830.
- [2] Doenhoff, M.J., Kimani, G. and Cioli, D. (2000) *Parasitol. Today* 16, 364–366.
- [3] Berry, C., Humphreys, M.J., Matharu, P., Granger, R., Hor-

- rocks, P., Moon, R.P., Certa, U., Ridley, R.G., Bur, D. and Kay, J. (1999) *FEBS Lett.* 447, 149–154.
- [4] Moon, R.P., Tyas, L., Certa, U., Rupp, K., Bur, D., Jaquet, C., Matile, H., Loetscher, H., Gruninger-Leitch, F., Kay, J., Dunn, B.M., Berry, C. and Ridley, R.G. (1997) *Eur. J. Biochem.* 244, 552–560.
- [5] Silva, A.M., Lee, A.Y., Gulnick, S.V., Majer, P., Collins, J., Bhat, T.N., Collins, P.J., Cachau, R.E., Luker, K.E., Gluzman, I.Y., Francis, S.E., Oksman, A., Goldberg, D.E. and Erickson, J.W. (1996) *Proc. Natl. Acad. Sci. USA* 93, 10034–10039.
- [6] Francis, S.E., Gluzman, I.Y., Oksman, A., Banerjee, D. and Goldberg, D.E. (1996) *Mol. Biochem. Parasitol.* 83, 189–200.
- [7] Harrop, S.A., Prociw, P. and Brindley, P.J. (1996) *Biochem. Biophys. Res. Commun.* 227, 294–302.
- [8] Gallego, S.G., Slade, R.W. and Brindley, P.J. (1998) *Acta Trop.* 71, 17–26.
- [9] Longbottom, D., Redmond, D.L., Russell, M., Liddell, S., Smith, W.D. and Knox, D.P. (1997) *Mol. Biochem. Parasitol.* 88, 63–72.
- [10] Dalton, J.P., Smith, A.M., Clough, K.A. and Brindley, P.J. (1995) *Parasitol. Today* 11, 299–303.
- [11] Wong, J.Y.M., Harrop, S.A., Day, S.R. and Brindley, P.J. (1997) *Biochem. Biophys. Acta* 1338, 156–160.
- [12] Becker, M.M., Harrop, S.A., Dalton, J.P., Kalinna, B.H., McManus, D.P. and Brindley, P.J. (1995) *J. Biol. Chem.* 270, 2496–24501.
- [13] Hewitt, E.W., Treuman, A., Morrice, N., Tatnell, P.J., Kay, J. and Watts, C. (1997) *J. Immunol.* 159, 4693–4699.
- [14] Rochefort, H. and Llaudat-Coopman, E. (1999) *APMIS* 107, 86–95.
- [15] Chevallier, N., Vizzavona, J., Marambaud, P., Baur, C.P., Spillantini, M., Pulerand, P., Martinez, J., Goedert, M., Vincent, J.P. and Checler, U. (1997) *Brain Res.* 750, 11–19.
- [16] Davies, D.R. (1990) *Annu. Rev. Biophys. Chem.* 19, 189–215.
- [17] Chou, K.-C., Tomasselli, A.G. and Henrikson, R.L. (2000) *FEBS Lett.* 470, 249–256.
- [18] Sauder, J.M., Arthur, J.W. and Dunbrack Jr., R.L. (2000) *J. Mol. Biol.* 300, 241–248.
- [19] Guruprasad, K., Tórná kangas, K., Kervinen, J. and Blundell, T.L. (1994) *FEBS Lett.* 352, 131–136.
- [20] Azim, M.K. and Zaidi, Z.H. (1999) *Biochem. Biophys. Res. Commun.* 264, 825–832.
- [21] Baldwin, E.T., Bhat, T.N., Gulnick, S., Hosur, M.V., Sowder, R.C., Cachau, R.E., Collins, J., Silva, A.M. and Erickson, J.W. (1993) *Proc. Natl. Acad. Sci. USA* 90, 6796–6800.
- [22] Dhanaraj, V., Dealwis, C.G., Frazao, C., Badasso, M., Sihanta, B.L., Tickle, I.J., Cooper, J.B., Driessen, H.P.C., Newman, M., Aguilar, C., Wood, S.P., Blundell, T.L., Hobart, P.M., Geoghagan, K.F., Ammirati, M.J., Danley, D.E., O'Connor, B.A. and Hoover, D.J. (1992) *Nature* 357, 466–472.
- [23] Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana, B.K., Baldwin, E., Weber, U.T., Selk, L.M., Clawson, L., Schneider, J. and Kent, S.B.H. (1989) *Science* 245, 616–621.
- [24] Brindley, P.J., Kalinna, B.H., Wong, J.Y.M., Bogitsh, B.J., King, L.T., Smyth, D.J., Verity, C.K., Abbenante, G., Brinkworth, R.L., Fairlie, D.P., Smythe, M.L., Milburn, P.J., Bielefeldt-Olmann, H., Zheng, Y. and McManus, D.P. (2001) *Mol. Biochem. Parasitol.* 112, 103–112.
- [25] Brinkworth, R.L., Prociw, P., Loukas, A. and Brindley, P.J. (2001) *J. Biol. Chem.* 276, 38846–38851.
- [26] Fujinaga, M., Cherney, M.M., Tarasova, N.I., Bartlett, P.A., Hanson, J.E. and James, M.N.G. (2000) *Acta Crystallogr. Sect. D* 56, 272–280.
- [27] Rawlings, N.D. and Barrett, A.J. (2000) *Nucleic Acids Res.* 28, 323–325.
- [28] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) *Nucleic Acids Res.* 22, 4673–4680.
- [29] Kumar, S., Tamura, K., Jakobsen, I.B. and Nei, M. (2001) *Bioinformatics* (submitted).
- [30] Guex, N. and Peitsch, M.C. (1997) *Electrophoresis* 18, 2714–2723.
- [31] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) *J. Appl. Crystallogr.* 26, 283–291.
- [32] Van Gunsteren, W.F. and Mark, A.E. (1992) *Eur. J. Biochem.* 204, 917–961.
- [33] McDonald, I.K. and Thornton, J.M. (1994) *J. Mol. Biol.* 238, 777–793.
- [34] Schechter, I. and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* 27, 157–162.
- [35] Tang, J., James, M.N.G., Hsu, I.N., Jenkins, J.A. and Blundell, T.L. (1978) *Nature* 271, 618–621.
- [36] Frutiger, J., Garrazou, I. and Pict, J.M. (1998) *Biochem. Biophys. Res. Commun.* 246, 719–724.
- [37] Gulnick, S.V., Suvorov, L.I., Majer, P., Collins, J., Kane, B.P., Johnson, D.G. and Erickson, J.W. (1997) *FEBS Lett.* 413, 379–384.
- [38] Scarborough, P.E. and Dunn, B.M. (1994) *Protein Eng.* 7, 495–502.
- [39] Pimenta, D.C., Oliveira, A., Juliano, M.A. and Juliano, L. (2001) *Biochim. Biophys. Acta* 1544, 113–122.
- [40] Beyer, B.M. and Dunn, B.M. (1997) *Protein Sci.* 7, 88–95.
- [41] Darke, P.L., Nutt, R.F., Brady, S.F., Garsky, V.M., Ciccarone, T.M., Leu, C.T., Lumma, P.K., Freidinger, R.M., Veber, D.F. and Sigal, I.S. (1988) *Biochem. Biophys. Res. Commun.* 156, 297–303.
- [42] Tomasselli, A.G. and Henrikson, R.L. (2000) *Biochim. Biophys. Acta* 1477, 189–214.
- [43] Kay, J. and Dunn, B.M. (1992) *Scand. J. Clin. Lab. Invest.* 52 (Suppl. 210), 23–30.
- [44] Vacca, J.P. (1994) *Methods Enzymol.* 211, 311–334.

ANEXO 2



presence of two similar β -barrel domains asymmetrically packed against each other, with the catalytic site lying between them.² The catalytic apparatus is composed from one side by Ser195 from the C-terminal domain and by His57 and Asp102 from the N-terminal domain, on the other side (following chymotrypsinogen numbering system).³ As recently reviewed,⁴ this triad is aligned in such a way that the Ser195 hydroxyl is polarized to attack the carbonyl group of the scissile bond with concomitant capture of the hydroxylic proton by His57, which is stabilized by Asp102, resulting in a tetrahedral intermediate/transition state that is readily stabilized by the backbone N–H groups of Ser195 and Gly193 (forming the so-called 'oxyanion-hole'). As the latter collapses to the acylated enzyme, a proton is transferred from His57 to the leaving group (the polypeptide fragment C-terminal to the labile bond). The enzyme is regenerated by a similar mechanism where a water molecule acts as the nucleophile.

The S1 subsite of the chymotrypsin-like proteases is responsible for the so-called 'primary specificity' of these enzymes [here we follow Schechter and Berger nomenclature,⁵ where P1–P1' represent the residues N- and C-terminal to the scissile bond, respectively; Distal residues are denoted outward; S1, S1', etc. denote the corresponding binding sites]. This cavity is delineated by three surface loops (loop 1: 184–195, loop 2: 213–228, and loop 3: 169–175) in the C-terminal domain. There is a remarkable structural match between this pocket on the enzyme surface and the amino acid occupying the P1 position of its substrates.⁶ Indeed, in the digestive enzyme trypsin,⁷ the negatively charged Asp189, positioned at the base of the S1 pocket, determines the absolute requirement for a positively charged residue such as Arg and Lys in the P1 position of its substrates. The same electrostatic interaction accounts for the competitive inhibitory activity of positively charged ammonium, amidinium, and guanidinium derivatives. Nevertheless, site-directed mutagenesis experiments have shown that the S1 subsite is not the single determinant of substrate specificity in serine proteinases.⁸ Thus, in more specialized enzymes such as thrombin,⁹ binding affinities are defined by the cooperative contribution of more distal sites in the protein surface.

Serine peptidases from some Viperidae venoms are able to convert fibrinogen into fibrin by releasing fibrinopeptides A and/or B. However, in most cases, these fibrin monomers are unable to be cross-linked by Factor XIII, resulting in aberrant clots that are independent of plasmin activity for dissolution.¹⁰ The in vivo consequence of this activity is dependent upon the plasmatic concentration of the enzyme but in ordinary envenomations it usually causes a marked anticoagulant effect. Although thrombin has many activities,⁹ these proteinases have been termed thrombin-like enzymes (TLE) because of their skill to mimic the fibrinogen clotting ability of this enzyme, which is central to the coagulation process in mammals. Moreover, a number of snake venom thrombin-like enzymes (SVTLEs) can mimic other thrombin functions such as activation of coagulation factors,¹¹ even though the mechanisms by

which they work may differ from those of the mammalian enzymes. Having a M_r ranging from 41 to 47 kDa, which is much greater than the predicted molecular weight of 25 kDa (\sim 230 residues), it has been shown that SVTLEs are extensively glycosylated. These proteins possess hydrolytic activity over synthetic substrates (arginyl and lysyl esters/amides), as well as an inhibitory profile against amidines and guanidines similar to trypsin.

The venoms of all *Lachesis muta* (*muta*, *rhombeata*, and *stenophrys*) pit vipers cause considerable hemorrhage and the *L. m. rhombeata* (Lmr-TLE) and *L. m. muta* (Lmm-TLE) subspecies have had, to the present, at least one SVTLE characterized.^{12,13} Protein sequencing of Lmm-TLE has revealed a high degree of sequence similarity (\sim 60%) to other SVTLEs such as flavoxobin, bathroxobin, and ancrod.^{14–17} All these SVTLEs, including Lmm-TLE and Lmr-TLE, are able to release only fibrinopeptide A after cleavage of the Arg16-Gly17 peptide bond in the A α -chain of fibrinogen. Because of this specific feature, ancrod, and batroxobin have been used as defibrinogenating agents for a number of clinical conditions.¹⁰ Although not functionally related, the tissue plasminogen activator-like enzyme from the *Trimeresurus stejnegeri* venom (TSV-PA) is structurally related to Lmm-TLE (54% identity). TSV-PA, in turn, is the only snake venom serine proteinase to have its structure determined to date.¹⁸ Lmm-TLE is also homologous to the mammalian serine proteases trypsin (40% identity) and thrombin (30% identity). Thus, sequence similarity allied to the biochemical data available has allowed Lmm-TLE and other SVTLEs to be classified in the chymotrypsin family. Interestingly, it has been recently proposed, based on a homology model, that Lmm-TLE has hybrid properties unifying the activities and structural features of the mammalian thrombin and trypsin enzymes.¹⁹

In view of the potential biotechnological and clinical application of SVTLEs as well as the importance of the S1 subsite in determining the binding affinities of small synthetic substrates and inhibitors of trypsin-like serine proteases, we have performed a series of computational analyses aiming to point out the dominating contributions to the inhibitory profile of small amidines/guanidines targeting the S1 subsite of a representative SVTLE—Lmm-TLE. In order to accomplish this, we have generated a homology model of this enzyme and then used it in docking simulations of a set of 21 known S1 specific reversible inhibitors of Lmm-TLE²⁰ (Fig. 1). We have mapped the S1 subsite interactions of the docked complexes so as to find, which residues and type of interactions can be important for dictating the distinct properties of this subsite among SVTLEs. The proposed formalism for calculation of binding affinities includes the van der Waals (vdW) component of the intermolecular interaction energy of the docked complexes, the hydration energy (ΔH_{hyd}) of the inhibitors and accessible area buried (AAB) upon formation of the ligand–protein complex. The relative importance of each of these terms has been determined by a statistical regression method. According to the regression model obtained, the vdW as well as the ΔH_{hyd} terms are the most important in



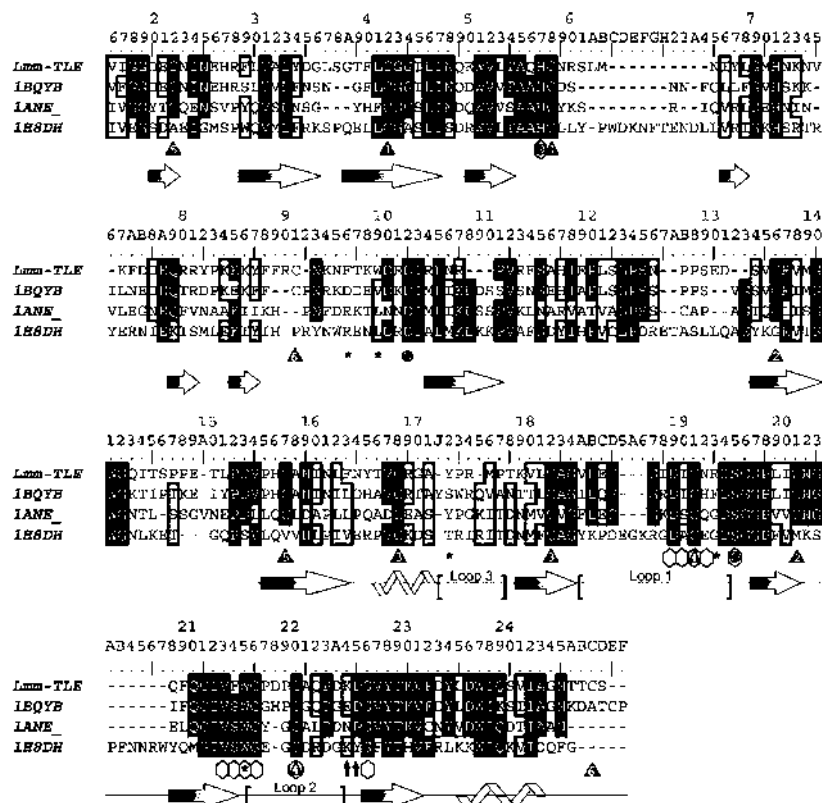


Figure 2. Structural alignment used on homology model construction. Lmm-TLE *Loaches mitta mitta* thrombin-like enzyme;¹⁴ IBQYB *Trimeresurus stejnegeri* plasminogen activator (TSV-PA);¹⁵ LANE_ Anionic rat trypsin;²⁷ IBSDH human α -thrombin.²² The alignment was boxed according to the conservation threshold of 70% and numbered after the chymotrypsinogen sequence.³ After each alignment block is represented the predicted secondary structure of Lmm-TLE: arrow β -strand; helix α -helix; straight line nondefined (loop). Loops 1–3 are delimited by brackets. Residues with a structural, catalytic, substrate recognition, or any particular function in the chymotrypsin family are detached: catalytic triad residues (gray circles); residues lining the S1 subsite walls (beehive-like polygons); Na⁺ binding site (black arrows); aryl binding site (asterisks); conserved cysteines (triangles ∇); and cysteines residues in Lmm-TLE possibly involved in disulfide bridges (triangles \blacktriangle). The black filled triangles mark disulfide bridges predicted by the Lmm-TLE 3D model while gray filled triangles spot these that were not.

In most members of the family, intradomain disulfide bonds help to keep the active conformation of the peptidase. Indeed, the proposed Lmm-TLE model predicted four cysteine pairs: 22–157, 42–58, 136–201, and 168–182. Other four Cys residues present in the Lmm-TLE sequence (191, 220, 91, 245) were not modeled as disulfide-bridged cysteines. Nevertheless, careful inspection of the model reveals that Cys191 and Cys220 are sufficient close to each other to afford a disulfide bond linking them. Loops 1, 2, and 3 as detached in Figure 4, set the limits of the S1 subsite in the C-terminal domain of the enzyme. At the bottom of loop 1 is displayed the Asp189 residue responsible for the primary specificity of this family of serine proteinases (Fig. 4).

2.3. Structural features of the docked complexes between Lmm-TLE and its inhibitors

Inhibitors docked into Lmm-TLE S1 binding pocket are displayed in Figure 1 and listed in Table 1 followed by

the observed inhibitory activities and the decomposition of the energetic scores (intermolecular and intramolecular) given by the dock program. The flexible docking calculations have confirmed the presumed binding mode of these amidine and guanidine derivatives. The conserved interaction observed in all simulated complexes is a face-to-face electrostatic attraction between the positively charged amidine or guanidine moiety and the negatively charged carboxyl group of Asp189 localized at the bottom of the S1 pocket. This salt-bridge governs the orientation of the inhibitors at the enzyme pocket as can be perceived from the superposition of inhibitors 13, 17, and 19 in Figure 5A. A diagram of the extensive nonpolar contacts and hydrogen bonds at the S1 subsite made by the most active compound in the series, 2-nafthamide, is depicted in Figure 5B.

A set of Lmm-TLE residues—Asp189, Thr190, Asn192, Ser195, Phe214, Trp215, Gly216, Cys220, and Gly226 were involved in several contacts with the majority of the inhibitors docked. Among these, only Asp189,

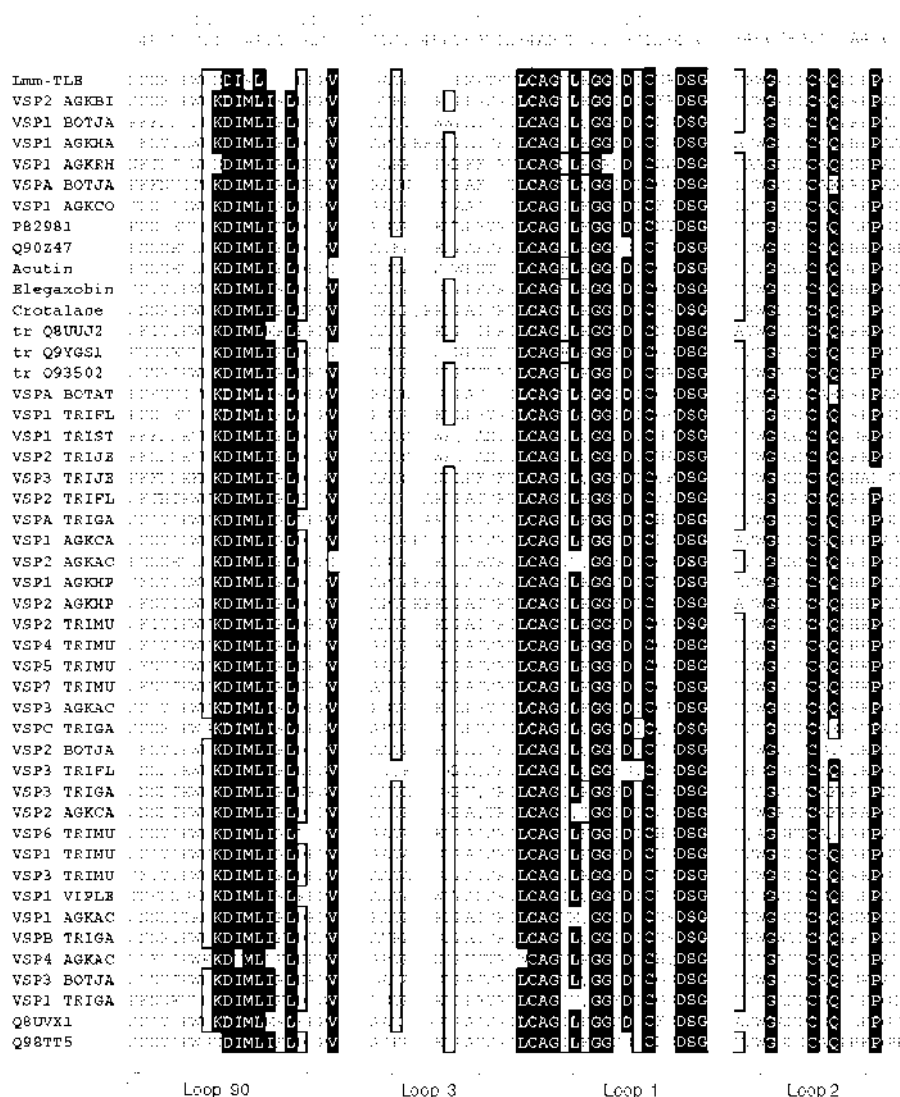


Figure 3. Multiple sequence alignment of snake venom thrombin-like enzymes (SV TLEs). Sequences were retrieved from protein sequence databanks using the sequence retrieval system (SRS at EMBL server, <http://www.embl.ch/>). SV TLEs discussed in the text or that could only be extracted from publications are identified by the ordinary names. The other sequences are denoted by the codenames found in databases. The alignment was boxed according to the conservation threshold of 85% and numbered after the chymotrypsinogen sequence.¹

Thr190, Gly216, and Cys220 were unanimous to all complexes analyzed. Moreover, four extra residues His57, Cys191, Val213, and Asp218 were also encountered making contacts with twelve of the studied inhibitors. Two types of hydrogen bonds were observed in all simulated complexes: the first involving one of the amidine or guanidine nitrogen atoms and the O₇ of Thr190; the second one involving the other amidine or guanidine nitrogen atom and the O₇2 of Asp189. These interactions, combined, form the basic attachment points for these inhibitors at the S1 subsite of Lmm-TLE. Nevertheless, some benzamidine inhibitors (7, 11,

13, and 14) bearing acceptor groups in position 4 of the aromatic ring are able to hydrogen bond to the Ser195 hydroxyl (Fig. 5C). On the other hand, compounds 2 and 19, which carry acceptor groups at the position 4 of the phenyl ring, are able to hydrogen bond to the Asn192 O_{δ1} and to the Asp218 O_{δ1} atoms, respectively. A summarized account of the interactions between inhibitor moieties and enzyme residues is presented in Table 2. No obvious trend has been observed between the extended hydrogen bonding capacity of some inhibitors and the experimental binding affinities of these compounds. The accessible area of Lmm-TLE S1

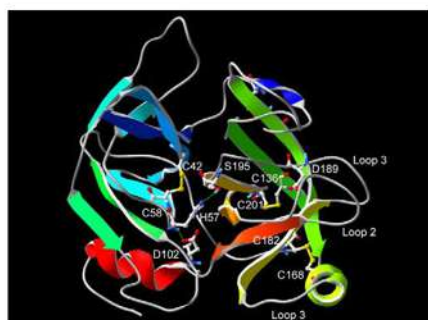


Figure 4. *Lachesis m. muta* thrombin-like enzyme (Lmm-TLE) homology model. Ribbon diagram colored by secondary structure succession from the N-terminal (blue) to the C-terminal (red). The catalytic and Cys residues involved in disulphide bridges are depicted in sticks and colored according to the CPK model. Loops 1–3 composing the S1 subsite are also identified.

subsite buried upon complexation is also listed in Table 2 and is barely proportional to the molecular size of the inhibitors.

2.4. Calculation of binding affinities for the amidine and guanidine inhibitors of Lmm-TLE

In the search for the relative contribution of each component of the formalism proposed to describe binding affinity to Lmm-TLE, we have performed PLS regression (PLSR) over three parameters: E_{vdw} , ΔE_{hyd} and AAB. The PLSR model could be derived from the

first LV (as assigned by crossvalidation), which is able to explain 76% of the variance in activity data (pK_i). This model presented thoroughly statistical meaning as indicated by the accompanying statistics in the MLR-like equation shown below:

$$pK_i = 5.364 \pm 0.701 - 0.100 \pm 0.012E_{vdw} + 0.090 \pm 0.016\Delta E_{hyd} + 0.008 \pm 0.002AAB$$

$$n = 21, \quad s = 0.41, \quad R^2 = 0.76, \quad F_{(1,19)} = 37.2, \\ p_{(1,19)} < 0,001, \quad Q^2 = 0.60, \quad S_{PRESS} = 0.62$$

Judging from the regression coefficients of the equation above, it is clear that the major contributions for the observed binding affinities come from the complementary fit of the inhibitors in the S1 subsite pocket (represented by E_{vdw}) followed by the hydrophobicity of the inhibitors (measured by the positive ΔE_{hyd} values). A negative coefficient for the vdW component occurs because the more negative is this term more favorable are the nonpolar interactions of the inhibitor with the enzyme residues. On the other hand, as a measure of the free energy of solvation in water, the more positive the value of ΔE_{hyd} the less favorable is this process, that is, more hydrophobic is the inhibitor. In turn AAB is a measure of the gain (mainly entropic in nature) in freeing bound water molecules from the interface of the protein–ligand complex upon association.

From crossvalidated standard deviation (S_{PRESS}), which is well below the standard deviation from mean pK_i values (0.96), we see that the model is able to improve

Table 1. Observed pK_i values and calculated energetic descriptors for the S1-directed inhibitors of *Lachesis m. muta* thrombin-like enzyme (Lmm-TLE)

S1-directed inhibitors	pK_i^a	Intermolecular interaction energy ^b			ΔE_{hyd}^c (kcal/mol)
		Total (kcal/mol)	VdW (kcal/mol)	Electrostatic (kcal/mol)	
2-Naphtamidine (1)	5.21	-38.0	-22.0	-12.3	-55.6
4-Aminobenzamidine (2)	4.89	-40.7	-21.2	-12.8	-56.7
4-Methylbenzamidine (3)	4.70	-37.3	-21.3	-12.6	-54.9
3-Methylbenzamidine (4)	4.80	-34.9	-20.0	-11.7	-55.0
4-Bromobenzamidine (5)	4.30	-36.5	-22.1	-12.3	-59.2
Benzamidine (6)	4.27	-35.2	-20.9	-11.8	-56.1
4-Methoxybenzamidine (7)	4.23	-34.2	-18.0	-12.5	-56.1
4-Chlorobenzamidine (8)	4.11	-36.2	-21.6	-12.3	-58.5
4-Fluorobenzamidine (9)	3.95	-37.3	-21.3	-12.6	-57.6
3-Aminobenzamidine (10)	3.80	-37.3	-21.7	-12.4	-58.5
4-Carboxybenzamidine (11)	3.40	-36.5	-22.4	-9.3	-60.7
3-Fluorobenzamidine (12)	3.39	-36.6	-21.1	-13.0	-57.7
4-Carboethoxybenzamidine (13)	3.24	-35.5	-19.1	-12.1	-60.8
4-Nitrobenzamidine (14)	3.12	-37.2	-22.5	-13.0	-66.8
3-Nitrobenzamidine (15)	3.04	-37.2	-21.6	-12.8	-65.5
Acetamidine (16)	1.37	-22.9	-17.2	-9.3	-64.4
Phenylguanidine (17)	3.72	-36.4	-14.5	-11.8	-56.0
Cyclohexylguanidine (18)	3.66	-42.6	-20.8	-10.7	-55.1
4-Aminophenylguanidine (19)	2.79	-38.8	-14.4	-13.2	-57.6
Methylguanidine (20)	2.66	-33.8	-14.6	-12.0	-61.8
Guanidine (21)	2.03	-39.8	-12.2	-13.2	-68.2

^a Experimental values taken from Magalhães et al.²⁰

^b Scores produced by the dock program.⁶⁵

^c Calculated using the SM5.4/A solvation model,⁶⁴ implemented in PC Spartan PRO release 1.05, Wavefunction, Inc.

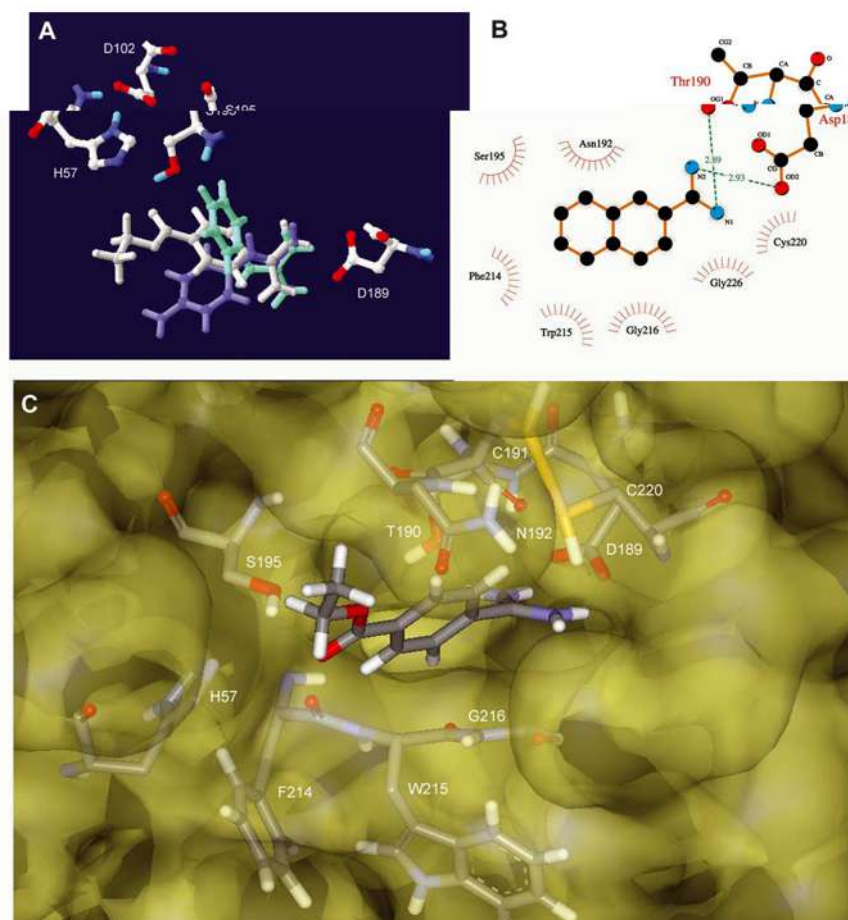


Figure 5. Binding mode of amidine and guanidine inhibitors within the S1 subsite of *Lachesis m. muta* thrombin-like enzyme (Lmm-TLE). (A) Superposed inhibitors **13** (4-carboethoxybenzamidine, gray), **17** (phenylguanidine, green) and **19** (4-aminophenylguanidine, blue) are represented in sticks. (B) Schematic representation of the interactions made by inhibitor **1** (2-naphitamidine); contacting atoms are detached with red marks and hydrogen bonds are indicated in dashed green lines along with donor/acceptor pairs distances. (C) Inhibitor **13** docked at the Lmm-TLE S1 pocket; the solvent accessible surface (1.4 Å probe) was generated transparent in order to make interacting residues visible whereas the hydrogen bonds are indicated by green lines.

our basic predictive ability on the inhibitory activity of the set of compounds analyzed. Other feature of this model is its progress in relation to the published model of Magalhães et al.²⁰ Its now possible to account for the order of binding affinities for the role set of 21 assayed inhibitors in contrast to the previous model, which only included the 14 benzamidines. The later also presented an inferior fit to data (F -ratio = 21.3) and a slightly worse predictive ability ($Q^2 = 0.56$). Analysis of fitted versus experimental pK_i plots shown in Figure 6 detected no severe outliers in the data set.

2.5. Analysis of the relative lipophilicity of S1 subsites of trypsin-like enzymes

As noted above, the regression coefficients of the PLSR equation have pointed out the importance of the

inhibitor hydrophobicity in order to attain strong association to Lmm-TLE. Therefore, it should be reasonable to ask if there is a correlation between the differential binding affinities of small S1-directed inhibitors of trypsin-like serine proteinases and the lipophilicity of this subsite on the enzymes surfaces. As a first step to test this hypothesis, we have determined the lipophilic potential (LP) over the molecular surfaces of the S1 subsite cavities of 20 structures from seven trypsin-like enzymes: trypsin, thrombin, t-PA, plasmin, Factor Xa, the catalytic domain of CIs component of complement and Lmm-TLE (Fig. 7). Subsequently, we devised a scheme to quantify the relative lipophilicity of the S1 subsite that was based on an arbitrary division of the LP scale given by the program MOLCAD (Table 3). HCA was applied as an objective methodology to obtain an independent grouping of the analyzed structures according to the relative lipophilicity values. It can be



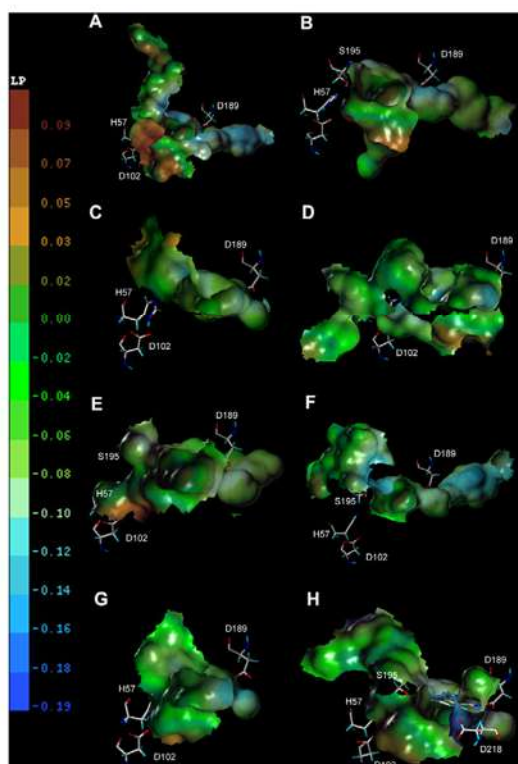


Figure 7. Analysis of lipophilicity potentials over S1 subsite surfaces from 8 trypsin-like enzymes. (A) Human α -thrombin (extracted from the complex with hirugen and *p*-aminophenylpyruvate, PDB code 1AHT); (B) Human tissue-type plasminogen activator (extracted from the complex with Dansyl-Glu-Gly-Arg Chloromethyl Ketone, PDB code 1BDA); (C) Human plasmin (extracted from the ternary complex Microplasmin-Staphylokinase-Microplasmin, PDB code 1BCT); (D) Catalytic domain of human C1s complement component (PDB code 1ELV); (E) Human Factor Xa (PDB code 1HCG); (F) Human α -thrombin 'slow' form (PDB code 1MH0); (G) Bovine β -trypsin (extracted from the benzamidine complex, PDB code 3PTB); (H) *Lachesis m. muta* thrombin-like enzyme (Lmm-TLE, this work) with benzamidine docked. Visible active site residues are labeled in each panel along with Asp189. The lipophilicity scale, shown in the left, ranges from blue (less lipophilic) to brown (more lipophilic).

method has enabled a classification consistent with the following order of relative S1 subsite lipophilicity: thrombin \ll [Plasmin $<$ Lmm-TLE $<$ trypsin] \ll [t-PA $<$ C1s $<$ Factor Xa]. After carefully determining the order of S1 subsite lipophilicity among the trypsin-like enzymes analyzed we could investigate if there was a trend in binding affinities of these proteases to S1-directed inhibitors that would follow this order. Indeed, as the plot in Figure 9 shows, the order of S1 subsite lipophilicity can be used to correctly sort the binding affinities of four trypsin-like enzymes—thrombin, plasmin, Lmm-TLE, and trypsin—to several small-molecule inhibitors. This result reinforces our conclusion that the hydrophobic character of S1-directed inhibitors, in addition to the overall shape complementarity to the cavity and the

deep salt bridge to Asp189, has a major contribution to the observed binding affinities of small S1-directed inhibitors.

3. Discussion

The lack of experimental structural data for the SVTLEs poses a further difficulty in understanding the structural features governing the activity of their S1-directed inhibitors. We have tried to circumvent this problem by generating a homology model of Lmm-TLE. Although most residues with a putative function have also been assigned in the alignment of Figure 2, most of the general features of the Lmm-TLE structure have been discussed before, through the model constructed by Castro et al.¹⁹ Thus, considering that our primary goal in this work was to study the significant aspects of S1 subsite recognition and binding affinities, other features of the Lmm-TLE model will not be discussed in depth. Nevertheless, one point concerning the possible existence of a 'fibrinogen recognition exosite' (FRE) in Lmm-TLE deserves attention. Noteworthy, Castro et al. have hypothesized a FRE homologous to the one found in mammalian thrombin (Arg60, Lys73, Lys76, Arg81, Arg82, Lys85, Lys86, Lys87, Arg110, Arg113), but their experimental data have not supported this proposal. Considering the distribution of basic residues over the Lmm-TLE surface, a FRE analogous to that proposed for crotonase,²¹ formed by a dislocated groove surrounded by Arg80 and Arg113 from one side and Lys85 and Arg107 from the other, should represent an alternative for the FRE location in Lmm-TLE.

The S1 subsite in serine proteases has been the target of a plethora of studies since the identification of the complementary aspect of the interaction of these enzymes with their polypeptide substrates. The availability of the first X-ray crystal structures of serine proteases of pharmaceutical and biological importance has greatly contributed to elucidate the structural features of the complementary fit between side chains of amino acids residues occupying the P1 position of substrates and the structure of this pocket on the enzyme surface.²² These early structural studies revealed that the S1 subsite architecture is composed by three surface loops from the C-terminal domain.² These segments in Lmm-TLE are represented by loops 1 (184–195), 2 (215–225), and 3 (172–178) as detached in the alignments of Figures 2 and 3 as well as in the 3D model of Figure 4. As can be perceived from the secondary structure representation of Lmm-TLE in Figure 4, loops 1 and 2 offer the main organization of the pocket while loop 3 contributes with more distal interactions to extended ligands of this subsite. Analysis of the predicted configuration of the complexes formed by the 21 S1-directed inhibitors of Lmm-TLE shown in Figure 1 and listed in Table 1 has disclosed the set of protein residues essential for binding (Table 2). Position 189 has been by far the most characterized in view of its primary role in the preference displayed by the enzyme for the P1 residue.^{4,6}

Table 3. Analysis of lipophilicity potentials over S1 subsite surfaces from selected trypsin-like enzymes

Enzymes (PDB code) ^a	Area of S1 subsite containing cavity		Lipophilicity potential ^d							
	Total ^b	Core ^c	Low		Medium		High		Total lipophilic ^e	
			(Å ²)	%	(Å ²)	%	(Å ²)	%		
α-Thrombin human (1MHO)	670.2	293.0	43.7	170.2	25.4	461.8	68.9	38.2	5.7	74.6
ε-Thrombin bovine (1ETR)	665.0	303.1	45.6	153.1	23.0	429.9	64.6	82.0	12.3	77.0
α-Thrombin human (1AHT)	916.7	352.3	38.4	197.0	21.5	507.5	55.4	212.2	23.1	78.5
α-Thrombin bovine (1HRT)	662.7	333.3	50.3	132.2	19.9	372.4	56.2	158.1	23.9	80.1
Plasmin human (1BUU)	280.2	232.5	83.0	55.1	19.7	200.8	71.7	24.3	8.7	80.3
β-Trypsin bovine (1BTY)	316.9	265.6	83.8	56.1	17.7	239.8	75.7	20.9	6.6	82.3
Plasmin human (1BLM)	326.6	276.1	84.5	55.4	17.0	271.2	83.0	0.0	0.0	83.0
Lmm-TLE <i>L. m. mata</i>	523.7	316.7	60.5	86.6	16.5	392.9	75.0	44.1	8.4	83.4
γ-Thrombin human (2HNT)	716.8	335.5	46.8	113.8	15.9	511.5	71.4	91.5	12.8	84.1
β-Trypsin bovine (3PTB)	298.7	226.7	75.9	45.4	15.2	227.9	76.3	25.3	8.5	84.8
α-Thrombin human (1AWH)	858.4	378.0	44.0	121.3	14.1	571.2	66.5	165.8	19.3	85.9
Anionic-2-Trypsin rat (1E2S)	432.6	251.4	58.1	53.8	12.4	328.4	75.9	50.4	11.7	87.6
Anionic-2-Trypsin rat (1ANE)	423.5	221.8	52.4	46.8	11.1	341.3	80.6	35.5	8.4	89.0
t-PA human (1BDA)	558.9	325.2	58.2	52.8	9.4	413.2	73.9	92.9	16.6	90.6
t-PA human (1ASH)	523.3	289.2	55.3	42.6	8.1	428.4	81.9	52.3	10.0	91.9
β-Trypsin bovine (1CLO)	310.2	228.7	73.7	23.9	7.7	251.4	81.0	34.9	11.3	92.3
C1s human (1ELV)	580.0	292.8	50.5	41.3	7.1	442.1	76.2	96.3	16.6	92.8
t-PA human (1RTF)	583.9	314.4	53.8	40.8	7.0	475.9	81.5	67.2	11.5	93.0
Anionic-2-Trypsin rat (1DPO)	487.3	258.1	53.0	23.7	4.9	378.2	77.6	85.4	17.5	95.1
Factor Xa human (1ICG)	463.9	338.2	72.9	11.3	2.4	410.5	88.5	42.1	9.1	97.6

^a Species from where proteins were derived follows the enzyme name.

^b Areas calculated from the total cavity detected by the MOLCAD module in Sybyl v.6.8, Tripos Inc.

^c Surface area within 4.0 Å from any of the benzamidine atoms. For the purposes of this analysis, benzamidine was docked into Lmm-TLE and thrombin subsites by superposition of the crystallographic complex with bovine trypsin.

^d Areas under the following lipophilicity ranges: low: -0.19 to -0.10; medium: -0.10 to -0.01 and high: -0.01 to 0.09. Percentages are expressed in relation to total cavity area.

^e Calculated from the sum of areas within medium and high I.P.

It can be perceived from the panels in Figure 5 that once the basic complementary salt-bridge has been established, other residues in loops 1 and 2 may afford additional interactions that could determine the observed binding affinities to a number of small ligands able to fit the S1 subsite. For instance, the structural data listed in Table 2 indicate that Thr190, Asn192, and Ser195 are able to provide such additional interactions in the form of hydrogen bonds or electron pair acceptor-donor interactions. Position 190 has been recognized as an important point of differentiation between enzymes presenting additional hydrogen bonding capability, due to serine or threonine residues in this position, as in urokinase type plasminogen activator (uPA) and factor VIIa, and those presenting an alanine, like in thrombin and tPA.²³ On the other hand, position 192 has recently been questioned, as its role in macromolecular substrate recognition in TSV-PA has not confirmed the results obtained with mammalian coagulant enzymes.¹⁸ Among other SVTLEs, it can be seen from

the alignment in Figure 3 that almost any kind of residue can be found in this position. Nevertheless, within certain groups some conservation may still be encountered: the two database entries for anrod (VSP1_AGKCO, VSP1_AGKRH)¹⁷ present either an asparagine or lysine in this position; Lmm-TLE; batroxobin (VSPA_BOTAT)¹⁶, bothrombin (VSPA_BOTJA),²⁴ and flavoxobin (VSP1_TRIFL)¹⁵ display a smaller glycine residue while calobin (VSP1_AGKCA)²⁵ and bilineobin (VSP2_AGKBI)²⁶ present a basic amino acid in this position.

The analysis of the docked complexes also revealed that Cys191 and Asp218 could also provide further stabilization to inhibitors bearing appropriately positioned polar groups. Surprisingly, position 218 has received little attention besides being located in the entrance of the S1 pocket. This residue is absolutely nonconserved among TSV-PA, rat anionic trypsin and human thrombin (see the alignment in Fig. 2). Considering the

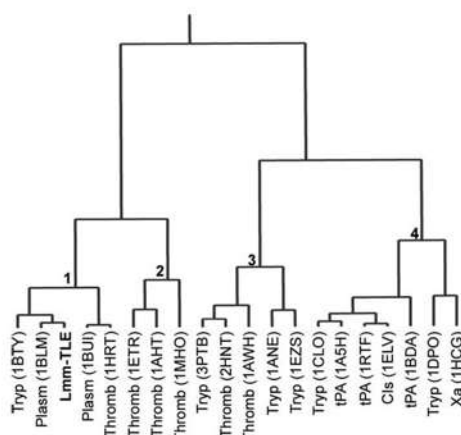


Figure 8. Dendrogram of trypsin-like enzymes characterized by the relative lipophilicity of the S1 subsite. Enzymes names are abbreviated next to the respective branch and are followed by the PDB code of the structure used in the analysis. Significant clusters are identified by the numbers 1–4 next to them. The hierarchical clustering analysis was made using complete linkage.

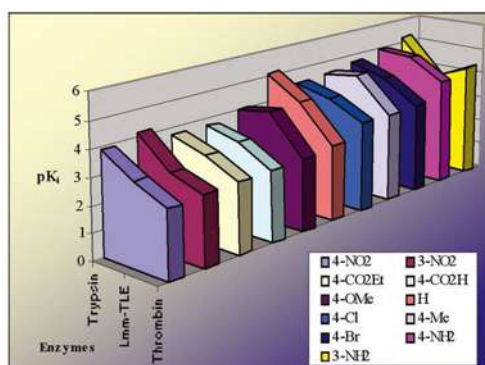


Figure 9. Trend in binding affinities of trypsin, *Lachesis m. muta* thrombin-like enzyme (Lmm-TLE), Plasmin, and thrombin to S1-directed benzamidines. For each inhibitor, the negative logarithm of the experimental inhibition constants (pK_i) are plotted according to the color code exhibited in the inset. Lmm-TLE data was retrieved from Magalhães et al.²⁰ For thrombin, plasmin, and trypsin, data relative to benzamidines and its 3-NO₂, 4-NO₂, 4-OMe, 3-NH₂ derivatives was obtained from Andrews et al.⁷⁹ while data corresponding to the remaining 4-substituted derivatives (Me, Cl, CO₂H, CO₂Et, Br, NH₂) was extracted from Coats et al.⁸⁰

different biological substrates and more specifically the distinct kinetic properties of these enzymes, a relevant role for position 218 could be envisaged for these proteins. However, besides being occupied by polar residues, position 218 displays little conservation between SVTLEs (Fig. 3). For example, calobin, both thrombin, batroxobin, and Lmm-TLE display an aspartate; flavoxobin, like TSV-PA, exhibits a histidine while anocrod and bilineobin present an asparagine and lysine, respectively.

The structural data listed in Table 2 along with the inspection of the binding mode of Lmm-TLE inhibitors

(Fig. 5) also indicate that the Val213 and Cys220 side chains along with the main chain atoms from Phe214, Trp215, Gly216, and Gly226 can supply mainly non-specific (vdW) interactions. Based on a theoretical 3D model of crotalase, Massova et al.²¹ have proposed that the Phe214 side chain, which pointed in the direction of the catalytic His57, would turn this region of the substrate binding-groove narrower than in trypsin and thrombin that have a smaller serine in this position. Supporting this prediction, our Lmm-TLE model displays the same kind of geometry for the Phe214 side chain. Interestingly, this feature seems to be rare among SVTLEs because a serine is conserved in this position for the large majority of the enzymes aligned, as can be observed in the alignment of Figure 3. Position 213, in turn, presumably contributes significantly to the hydrophobicity of the S1 subsite since a nonpolar amino acid (valine, leucine, or alanine) is conserved among all SVTLEs (Fig. 3), trypsin, thrombin, and TSV-PA (Fig. 2). A nonpolar residue in this position should be important for increasing the binding affinities of S1 subsite ligands by increasing the lipophilicity of the S1 subsite.

Although being unable to present sufficient size to achieve high selectivity to close-related serine proteases, ligands large enough to fill just the S1 subsite are still object of much research work because the perfect anchoring in this pocket is the first step to design a potent and selective inhibitor.^{27,28} In this way, the obvious applicability of a tight-binding noncovalent inhibitor has promoted numerous studies aiming to predict the binding affinities of small ligands directed to the S1 subsite of these proteases. In the case of trypsin-like serine proteases, there is a 10⁵-fold preference for substrates containing an arginine or lysine in the P1 position, in relation to the next favored residue.⁶ In a seminal work, Mares-Guia and Shaw have investigated this amazing specificity, using amidines and guanidines as structural analogues of the arginine side chain.²⁹ Following works have suggested the lipophilic feature of the interaction site based on the hydrophobic nature of the most potent amidine and guanidine derivatives as well as synthetic substrates.^{30,31}

The subsequent availability of X-ray crystal structures has greatly influenced these studies.³² For instance, in other pioneering work, Mares-Guia et al.³³ have conducted a QSAR study that has rationalized the activity of 4-substituted benzamidines by pointing out the importance of electron-releasing substituents (ERS) in position 4 of the aromatic ring in order to maximize the inhibitory potency through formation of an additional dipolar interaction with the O_γ atom of the catalytic Ser195 residue. This model has until recently represented a major paradigm in the binding of benzamidines to trypsin-like proteases. Later, the group of Hansch et al.³⁴ has proposed a QSAR model where hydrophobic and steric terms were added in order to improve the correlation to inhibition data. Interestingly, these authors have proposed another role for the required electrophilic character of ERS, which was to re-establish the hydrophobicity of the aromatic ring

diminished by the electron-withdrawing character of the protonated benzamidine moiety.

Recent publications have, in general, supported the importance of the hydrophobicity and nonpolar interactions of S1-directed inhibitors of trypsin-like enzymes.^{35–39} Indeed, the regression coefficients of the PLSR model of Lmm-TLE inhibitors clearly show the importance of the vdW interaction term and hydration energy in order to achieve maximum affinity to the S1 subsite of this SV-TLE. This is a curious observation since the S1 subsite provides many acceptor and donor groups able to participate in more specific, directional interactions with corresponding moieties on the inhibitors structures. The Jorgensen's group has conducted a Monte Carlo simulation, in order to compute the binding affinities of 4-substituted benzamidines to trypsin, which has helped to clarify this issue.³⁶ The simulation revealed that mutation of the 4-amino substituent to methyl, chlorine or hydrogen introduces no net gain in the number of hydrogen bonds formed between inhibitor moieties and S1 surface groups because the number of potential hydrogen bonds lost in passing the inhibitors from the aqueous phase to the S1 pocket is constant among the studied inhibitors. Hence, the total number of hydrogen bonds involving protein atoms in the active site is nearly independent of the inhibitor and hydrogen bonding cannot therefore be invoked to rationalize the trends in binding affinities.

Based on purely continuum electrostatic grounds some authors have suggested that substitution by ERS decreases the dipolar character of the closely related 4-substituted benzamidines, decreasing in turn the free-energy penalty in desolvating these inhibitors upon association with the enzyme.^{36,37} Nevertheless, when some measure of the ERS character of the substituents, like Hammett's sigma constants,⁴⁰ are plotted against pK_i , only a poor correlation is obtained.^{20,35,37} Thus, independent of the interpretation employed, the electronic properties of the substituent in these benzamidines cannot be invoked to rationalize binding affinities in this case. Hence, it is now possible to ask if models combining sigma constants and hydrophobicity descriptors, such as LogP,³⁴ have only improved the correlation to activity data by a misleading collinearity (correlation) of the descriptors employed. In fact, it has been argued that as there is only a poor relationship between LogP and inhibitory activity, one should not rationalize the binding affinities of S1-directed inhibitors in terms of their hydrophobicity.^{20,37} Nevertheless, evidence for the role of a hydrophobic effect⁴¹ in the binding of S1-directed inhibitors has been observed in a number of thermodynamic studies.^{37,38,42} The results of our analysis have also clearly indicated a role for the hydrophobicity of the inhibitors if high binding affinity to the S1 subsite in Lmm-TLE is to be achieved.

Our results also show that when more structural diverse S1-directed inhibitors are considered, the overall shape complementarity of the inhibitors to the S1 cavity becomes important to maximize the binding affinity. This contribution is represented by the vdW component

of the intermolecular interaction energy present in the PLSR model obtained in this work. Results from more accurate calculations with thrombin inhibitors³⁹ also support this finding. It is interesting to note that inclusion of the electrostatic contribution to the intermolecular interaction energy is not strictly required to describe the great majority of the variance in inhibition data. It is curious since it should be expected that the requirement for a positively charged moiety like the amidinium or guanidinium groups would be readily translated into an important electrostatic contribution to binding affinities. The analysis of Checa et al.⁴³ concerning the inhibition of trypsin by flavonoids has helped to shed some light in this apparent contradiction. It was concluded that vdW interactions would be dominant in describing binding affinities because there is little variation in the electrostatic term, which is largely solvent independent. Any attempt to include the electrostatic interaction term has not been successful in improving the statistics of our PLSR model. The above discussion indicates that the hydrophobic character of S1-directed inhibitors, in addition to the overall shape complementarity to the cavity and the deep salt bridge to Asp189, possesses a considerable contribution to the observed differential binding affinities among small S1-directed inhibitors of Lmm-TLE and related serine proteinases.

If such a relationship between strong binding to the S1 subsite and the hydrophobic nature of the inhibitor is as true for other trypsin-like enzymes as it is true for Lmm-TLE, it should be possible to observe a direct correlation between the order in binding affinities of these enzymes to their inhibitors and the lipophilicity of this cavity on their surfaces. A comparison between the overall order in the LP over the surfaces of S1 subsite boundaries in Lmm-TLE, trypsin, plasmin, and thrombin (Table 3) and the trend in binding affinities to S1-directed inhibitors shared by these enzymes (Fig. 9) provides general support for this hypothesis as the order of binding affinities, as expressed by pK_i , among simple 4- and 3-substituted benzamidine inhibitors follows, in most cases, the order of lipophilicity of the S1 subsite from the four proteases: thrombin \ll plasmin $<$ Lmm-TLE $<$ trypsin.

It is important to note that the proposed order of lipophilicity in S1 subsites of thrombin and trypsin is the inverse of the usual order described in the literature.⁴⁴ In view of the importance of our conclusions regarding S1 subsite lipophilicity to a number of scientists working in the fields of thrombin inhibition and haemostasis control, we decided to discuss this issue in depth. At this point, it is wise to note that an inherent handicap of our lipophilicity quantification scheme is that its relative, since the range of the LP scale depends on which structures entered the analysis. Hence, our lipophilicity measure, although simple, is not absolute. For instance, the lower limit of the LP scale applied on our analysis was imposed by the high hydrophilic region on the entrance of the Lmm-TLE S1 subsite that could be attributed to an aspartate residue at position 218 (Fig. 7, panel H). A negatively charged residue at this position is common to many SVTLEs (Fig. 3) and thrombin.

The rationalization for the proposed order of lipophilicity between trypsin and thrombin that is commonly found in the literature has been primarily based on the substitution in position 190 of a serine residue in the former enzyme by an alanine residue in the later. Nevertheless, it became clear from the results of our analysis, which was based on the 3D LP (calculated from atomic hydrophobicity constants⁴⁵), that not only the composition of the loops giving the S1 subsite architecture but also the orientation of the residues present may influence the lipophilicity of this cavity. In fact, other factors may affect the effective lipophilicity of the S1 subsite. For instance, recent crystallographic studies have evidenced that the result of the S190A (trypsin → thrombin) mutation may be counterbalanced by the presence of a conserved water molecule bridging Ala190 to the benzamidinium moiety of inhibitors.²³ Similarly, even more pronounced should be the effect of the presence of a water channel communicating with the S1 subsite, as in fact is found in thrombin. This channel contains a Na⁺ binding site that has been shown to serve as an allosteric switch between the so-called 'slow' and 'fast' forms of thrombin.^{46,47} Notably, for the purposes of our analysis, this channel was considered as an integral part of the S1 subsite of the thrombin structures analyzed, as evidenced in Figure 7 (panels A and F). Indeed, such channel was clearly responsible for the great proportion of low LP over the S1 subsites surfaces of thrombins (Table 3).

Dang and Di Cera⁴⁸ observed that for several serine proteases, the presence of a proline residue at position 225 could be successfully correlated with the inability of the enzyme to be allosterically regulated by Na⁺. In accordance with this proposal, our analytic procedure could detect a channel communicating with the S1 subsite of factor Xa, which, like thrombin, has a Y225. Interestingly, analysis of the single C1s structure available (PDB code 1ELV) could not reveal such channel, although this enzyme also presents a tyrosine residue at position 225 (Fig. 7, panel D). Accordingly, careful inspection of the results from the Na⁺ regulation experiments reported by Dang and Di Cera⁴⁸ reveals that not all serine proteases not carrying a P225 can be as much activated by Na⁺ as thrombin and factor Xa can, suggesting possible drastic structural differences in the Na⁺ binding sites among such enzymes. Conversely, we could detect the presence of a channel starting from the surface opposite to the active site and ending in the S1 subsite of tPA that would not be predicted by the rule of Dang and Di Cera, since this serine protease presents a P225. Whether this channel would be able to shelter a Na⁺ binding site or not needs experimental confirmation.

Another interesting observation from our analysis concerning the structural differences between the fast and slow forms of thrombin was that, in agreement with the literature, we detected a much smaller accessible surface area for the S1 subsite in a putative slow form of thrombin (PDB code 1MHO) compared to the sodium-bound fast form (see Fig. 7, panels A and F). A handful of biochemical experiments collected along the years^{46–51}

have crystallized the idea that the slow form of thrombin would have a much tighter substrate binding cleft in order to account for the drastic decrease in the rate of substrate binding upon the fast → slow transition. In this manner, our result is consistent with the claim of Di Cera and co-workers⁵² that the sodium-free structure reported by them (PDB code 1MHO) would correspond to the slow form of thrombin, although a recent publication disagrees from that.⁵³

It is now clear that electrostatic interactions or differential hydrogen bonding cannot be used to explain the observed binding affinities of simple amidines/guanidines inhibiting trypsin-like enzymes. On the other hand, the combination of all nonpolar contacts that can be made by these compounds in the S1 subsite of trypsin-like enzymes, as an example of the lock-and-key concept, is the primary force contributing to maximization of binding affinities. Destabilization of the inhibitors in the aqueous phase by the hydrophobic effect also makes an important contribution to affinity for the lipophilic S1 subsite. In consideration of the stated above, it is tempting to think that in the course of evolution, SVTLEs, may have been shaped to present the catalytic efficiency of trypsin while achieving the required specificity of thrombin by different strategies, like the maximization of lipophilicity through specific mutations in residues composing the S1 subsite architecture or by employing an alternative FRE.

SVTLEs have been successfully applied in a number of clinical conditions.¹⁰ Their therapeutic action is probably related to the fibrinogen depletion *in vivo*, which results in decreased blood viscosity and better blood flow.¹⁰ However, the prolonged usage of these enzymes is inefficient due to the development of immunological resistance by the patient. Fortunately, replacement by a different SVTLE is able to by-pass this handicap in successive courses of long-term anticoagulant therapy because of the lack of cross-reactivity between two SVTLEs, for example, ancrod, and batroxobin.¹⁰ Related studies with other STVI.Fs should be able to disclose such beneficial lacks of cross-reactivity within these specific venom components. Other use of SVTLEs should be envisaged in the clinical research of anomalies in the coagulation cascade of patients, since these enzymes act in more specific points of this pathway. Thus, considering all possible applications of SVTLEs, the design of potent small-molecule inhibitors aimed to manipulate the activity of these enzymes should be of great relevance. The results from this study should have application in the development of small compounds for controlling the enzymatic activity of SVTLEs in whatever biotechnological or clinical research application that these enzymes can have, in special the prevention of possible hemorrhagic counter-effects of the clinical use of these enzymes and the exploitation as chemical probes for functional-proteomics studies of snake venoms. With these straightforward computational analysis we hope to have contributed to clarify the debate over the structural features responsible for S1 binding affinities in the trypsin family, in special among the structure scarce SVTLEs.

4. Methods

4.1. Sequence analysis and multiple sequence alignment

The sequences of Lmm-TLE and other similar SVTLEs were retrieved from the SWISS-PROT or Tr-EMBL protein data banks using the Expert Protein Analysis System WWW server (ExpASY; <http://www.expasy.ch/>). An initial alignment of these proteinases was obtained in the ClustalW v.1.81 software,⁵⁴ using program's default parameters. This alignment was then inspected and manually corrected. An analysis of the secondary structure potential of Lmm-TLE sequence was also carried out using the PSI-PRED⁵⁵ software.

4.2. Template selection and model building

Templates were retrieved after similarity (BLAST) search with Lmm-TLE sequence against the ExNRL-3D database available through the SWISS-MODEL server (<http://www.expasy.ch/swissmod>).⁵⁶ Among the highest scoring structures, chain B of TSV-PA,¹⁸ rat anionic trypsin bound to benzamidine⁵⁷ and chain H of human α -thrombin in complex with a tripeptide phosphonate inhibitor⁵⁸ (Ex-PDB codes 1BQYB, 1ANE_ and 1H8DH, respectively) were selected as templates for the homology modeling procedure. An initial alignment between Lmm-TLE and its templates was accomplished by combining the information derived from the multiple sequence alignment of SVTLEs described above to the S1A chymotrypsin family alignment retrieved from the MEROPS database of peptidases.¹ Improvement of this alignment was accomplished by maximizing the superposition of secondary structural elements and conserved functional residues within the chymotrypsin family. Following this optimal alignment, the sequence of Lmm-TLE was folded over its templates structures using the SwissPDB Viewer v.7.2b2 software.⁵⁹ Preliminary model building was accomplished by an automated procedure through the use of the optimize mode in SWISS-MODEL server. Several cycles of alignment refinement and model construction were carried out. The modeling progress was accompanied through the use of Ramachandran plots generated with the PROCHECK software.⁶⁰ All operations conducted in SwissPDB Viewer were performed on a Pentium III 800 MHz personal computer.

4.3. Model refinement and validation

All energy minimization operations described in this section were conducted with the GROMOS96 v.43B1 force field⁶¹ implemented in SwissPDB Viewer v.3.7b2 running on a Pentium III 800 MHz personal computer. At first, any missing peptide bonds in the Lmm-TLE model were introduced and locally minimized by 20 steps of steepest descent followed by 100 steps of conjugated gradients minimization methods, until the energy difference between two steps was below 0.01 kJ/mol. Two successive refinements of the Lmm-TLE model were performed using a similar minimization

protocol in which the type of residues constrained during the minimization were varied: initially, only residues out of the most favored regions of the Ramachandran plot (53 total) were allowed to move; then residues with high model B-factor and/or force field energy were included (37 total). The final refined theoretical structure of Lmm-TLE was achieved by an energy minimization with a harmonic constraint over all residues showing a low model B-factor. Refinement progress was accompanied by the distribution of Lmm-TLE residues over the most favored and allowed regions of the Ramachandran plot. The geometry of Asp189 and the residues composing the catalytic triad (His57, Asp102, and Ser195) received special attention due to their importance in this work. When necessary, manual torsions were imposed on the dihedral angles of these residues in order to reproduce the conserved architecture of the S1 subsite and active site of trypsin-like serine proteinases. A thorough stereochemical validation of the final Lmm-TLE model was achieved by submitting the structure to the analytical softwares contained in the PROCHECK and WHATCHECK packages.^{60,62} An extensive functional crossvalidation of the model was also performed by assuring that any biochemical information available to Lmm-TLE was in accordance to the structural features predicted by the proposed 3D model.

4.4. Small compound modeling and calculation of hydration energy

The structures for the set of 21 amidine and guanidine inhibitors (Fig. 1) reported in the work of Magalhães et al.²⁰ were constructed in PC Spartan Pro release 1.05 (Wavefunction, Inc., Irvine, CA, USA; kindly made available by Dr. Ricardo B. de Alencastro and Dr. Magaly G. Albuquerque from Lab. de Modelagem Molecular, DQO/IQ, UFRJ, Brazil) using the semi-empirical AM1 Hamiltonian.⁶³ All amidinium and guanidinium groups were protonated and the carboxylic group of inhibitor **II** was constructed in the ionized form. Atomic point charges needed for the docking procedure were calculated through fitting of the electrostatic potential. Hydration energy (ΔE_{hyd}) was calculated using the SM5.4/A model⁶⁴ implemented in PC Spartan Pro. All calculations were performed in a Pentium II-MMX 350 MHz personal computer.

4.5. Docking calculations

Automated flexible molecular docking simulations were carried out using the DOCK⁶⁵ v4.0 program suite, which takes into account only the flexibility of the ligand. This seems a reasonable approximation here, since no major conformational change has been observed in the bound structures of small S1-directed inhibitors with a number of serine proteases of the chymotrypsin family.⁴ The docking procedure can be divided in four general stages: ligand/receptor preparation, site characterization, scoring grid calculation, and docking itself. The set of ligands to be docked consisted of the 21 amidine and guanidine derivatives,²³ shown in Figure 1, constructed

as described above. The final refined Lmm-TLE modeled structure was used as the receptor. In order to prepare the receptor for energetic grid calculation, all hydrogen atoms were added and then Kollman-type⁶⁶ atomic point charges were assigned. Preliminary site characterization was accomplished by running the autoMS script distributed within the DOCK package, which automatically generates a complementary sphere representation of cavities found in the receptor surface calculated by the *ms* program (program number 429, Quantum Chemistry Program Exchange, Indiana University). A better characterization of the S1 subsite in Lmm-TLE model was achieved by molecular graphics assisted editing of the spheres cluster generated by autoMS. This procedure was made in SwissPDB Viewer v.3.7b2 software and was guided by the analysis of the volume occupied by several S1-directed inhibitors bound to trypsin-like enzymes. In order to obtain a set of input parameters for the grid and dock programs able to reproduce the observed binding mode for the class of inhibitors considered in this work we have used as a guide the crystallographic complex of benzamidine and rat anionic trypsin (PDB code 1ANE)⁶⁷. Input parameters governing ligand flexibility, orientation, receptor matching, scoring, and minimization were varied until the crystallographic binding mode of benzamidine in rat trypsin could be reproduced with an RMS deviation of less than 1.0 Å. Docking calculations were performed in the SGI Origin 2000 server installed at FIOCRUZ/Scientific Computation Program (PROCC).

4.6. Analysis of docked complexes of amidine and guanidine competitive inhibitors of Lmm-TLE

Besides intermolecular energy terms (vdW and electrostatic) furnished by the dock program, the simulated complexes had their nonbonding interactions (hydrogen bonds and nonpolar contacts) mapped using the HBPLUS program.⁶² Schematic plots of these interactions were generated using the Ligplot software.⁶⁸ In addition, the docked complexes were inspected using molecular graphics in SwissPDB viewer v7.2b2 and the receptor solvent accessible area buried (AAB) upon binding was calculated for each inhibitor according to the following formula: $AAB = (A - B) - AB$, where A is receptor solvent accessible area, B is the ligand solvent accessible area and AB is the solvent accessible area of the complex. A probe radius of 1.4 Å, representing the vdW radius of water, was used in these calculations.

4.7. Binding affinity calculation background

The methodology employed by us stems from the additivity principle in calculating binding affinities,^{69,70} that is, different contributions to free energy of binding can be calculated separately and summed. The resulting 'master equation' usually accounts for the contributions due to solvent, conformational changes in the protein and ligand, the protein–ligand interactions and variation in the degrees of freedom of the diverse molecular motions involved.⁷³ Partial least square regression

(PLSR) algorithm⁷² can be used to weight out each of these terms in order to maximize the explanation of the observed property, in this case, pK_i . The latter approach has very straight relationship to quantitative structure activity relationship studies (QSAR).³⁵

It has been suggested that in the case of small S1-directed inhibitors of trypsin-like enzymes, the electrostatic portion of the intermolecular interaction energy is almost constant for a series of congeneric ligands, contributing very little to the variance in binding affinities.⁴³ Thus, we have opted to explore the ability of the scoring function from the docking program to approximate the major enthalpic component of the binding free energy in this case, that is the intermolecular vdW energy, while the other components related to the solvent contribution to the observed order in binding affinities are entered in the regression equation in the form of calculated ligand/complexes energetic and structural parameters like ΔE_{hyd} and AAB, respectively: $pK_i = k + aE_{\text{vdw}} + b\Delta E_{\text{hyd}} + cAAB$, where E_{vdw} is the vdW portion of the intermolecular scoring function of dock 4.0 program; ΔE_{hyd} is the hydration energy of the ligand calculated according to the SM5.4/A model.⁶⁴ AAB is the solvent accessible area of the protein that is buried upon inhibitor binding, k is a constant and a , b , and c are the regression coefficients. It is interesting to note that we have used ΔE_{hyd} in place of cLogP⁷² as a hydrophobicity parameter because ΔE_{hyd} has included ionic compounds in its parameterization and as a semi-empirical quantum chemical calculation, presents much less restriction about compound classes or ionization states to be submitted to calculations. On the other hand, cLogP has an intrinsic handicap with ionized compounds, since calculated values are derived from parameters fitted from experimental data with neutral or ionizable compounds in the neutral form from buffered solutions.⁷²

The PLS method models the experimental data (Y or dependent variable) in terms of calculated descriptors of the complex/inhibitors (X block or independent variables). This robust chemometric method is superior to the classical multiple linear regression (MLR) in being able to deal with a great number of (even highly correlated) X variables by reducing this large data set to a few latent variables (LV) maximum correlated to the Y variable.⁷¹ The importance of each original X variable to the LV is determined by its loading. The number of significant LVs to be used is usually determined by crossvalidation.⁷¹ This regression approach, although not based in the rigorous statistical thermodynamics grounds of more realistic simulations⁷⁴ is high computationally efficient and has been successfully applied in a number of cases.^{35,75}

4.8. Calculation of binding affinities

Experimental binding affinity data, expressed as the negative logarithm of the inhibitory constants (pK_i), for the series of amidine and guanidine competitive S1-directed inhibitors of Lmm-TLE were computed from the work of Magalhães et al.²⁰ (Table 1). The PLS method

implemented in the MSS/QSAR module of Sybyl v.6.8 from Tripos Inc. (St. Louis, MO, USA) was used to obtain the MLR-like regression coefficients of the model. The statistical significance of these was calculated by the bootstrap method.⁷⁶ Overall statistical relevance of the model was judged by the correlation coefficient (r^2), p value, standard deviation (s) and F -ratio. Predictive ability of the model was evaluated by the crossvalidation coefficient (Q^2) and standard error of predictions (SEP).

4.9. Calculation of the lipophilic potential over the surface around S1 subsite

For this analysis, besides the final Lmm-TLE model, 19 structures of trypsin-like enzymes were used: 6 trypsins (PDB codes 1BTY, 3PTB, 1CLO, 1ANE, 1EZZ, and 1DPO), 6 thrombins (PDB codes 1ETR, 1HRT, 2HNT, 1AWH, 1MHO, and 1AHT), 3 t-PAs (PDB codes 1RTF, 1A5H, and 1BDA), 2 plasmins (PDB codes 1BUI and 1BLM), CIs (PDB code 1ELV), and Factor Xa (PDB code 1HCG). Structures were selected in order to account for possible effects due to experimental conditions, inter-species variation, and post-translational processing. All manipulations and calculations described in this section were conducted in the Sybyl v.6.8 modeling package from Tripos, Inc. (St. Louis, MO, USA). First, the PDB structures were freed from any ligands and water molecules followed by addition of all hydrogen atoms. Next, the molecular surface around the S1 subsite (constituting the major cavity in the surface of these structures) was calculated by the MOLCAD module. Then, the LP generated from the atomic hydrophobicity constants derived by Ghose and Crippen⁷⁷ was mapped on the surface surrounding the S1 subsite, according to the method of Furet et al.⁴⁵ implemented in MOLCAD.

4.10. Analysis of the binding affinities of trypsin-like enzymes to benzamidine derivatives according to the relative lipophilicity of the S1 subsite cavity

At first, in order to have a simple measure of the S1 subsite lipophilicity, the LP was put in the same scale for all enzymes analyzed and the areas belonging to three arbitrary property ranges were measured: low (−0.19 to −0.10), medium (−0.10 to −0.01), and high (−0.01 to 0.09); each of these was expressed as percent of the total cavity surface area and the total subsite lipophilicity was considered as the sum over the medium and high LP percentages. Following, a classification of the 20 structures analyzed was proposed based on a complete linkage, hierarchical-clustering analysis⁷⁸ (HCA) of S1 subsite lipophilicities (MSS module, Sybyl molecular modeling package). Natural groups were identified by careful visual inspection of the dendrogram produced by the analysis. Finally, the pK_i values^{79,80} of several benzamidine-based inhibitors of trypsin, plasmin, thrombin, and Lmm-TLE were plotted to investigate the variation of binding affinities according to S1 subsite lipophilicities.

Acknowledgements

We thank Prof. R. B. De Alencastro and C. A. F. Oliveira for carefully reading this manuscript. We are in debt with Prof. O. A. C. Antunes and Prof. M. G. Albuquerque for valuable corrections and suggestions. We also thank the referee for generous suggestions.

References and notes

1. Rawlings, N. D.; O'Brien, E. A.; Barrett, A. J. *Nucleic Acids Res.* **2002**, *30*, 343.
2. Lesk, A. M.; Fordham, W. D. *J. Mol. Biol.* **1996**, *258*, 501.
3. Meloun, B.; Kluch, I.; Kostka, V.; Moravek, L.; Prusik, Z.; Vanacek, J.; Keil, B.; Sorm, F. *Biochim. Biophys. Acta* **1986**, *130*, 543.
4. Hedstrom, L. *Chem. Rev.* **2002**, *102*, 4501.
5. Schechter, A.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
6. Czapińska, H.; Otłowski, J. *Eur. J. Biochem.* **1999**, *260*, 571.
7. Colomb, E.; Guy, O.; Deprez, P.; Michel, R.; Figarella, C. *Biochim. Biophys. Acta* **1978**, *525*, 186.
8. Perona, J. J.; Craik, C. S. *Protein Sci.* **1995**, *4*, 337.
9. Stubbs, M. T.; Bode, W. *Trends Biochem. Sci.* **1995**, *20*, 23.
10. Markland, F. S. *Toxicol.* **1998**, *36*, 1749.
11. Matsui, T.; Fujimura, Y.; Titani, K. *Biochim. Biophys. Acta* **2000**, *1477*, 146.
12. Aguiar, A. S.; Alves, C. R.; Melgarejo, A.; Giovanni-De-Simone, S. *Toxicol.* **1996**, *34*, 555.
13. Silva, N. J.; Aird, S. D.; Seebart, C.; Kaiser, I. I. *Toxicol.* **1989**, *27*, 763.
14. Magalhães, A.; Fonseca, B. C. B.; Diniz, C. R.; Gilroy, G. P.; Richardson, M. *FEBS Lett.* **1993**, *329*, 116.
15. Itoh, N.; Tanaka, N.; Funakoshi, I.; Kawasaki, T.; Mihashi, S.; Yamashima, I. *Nucleic Acids Res.* **1988**, *16*, 10377.
16. Burkhardt, W.; Simth, G. F. H.; Su, J.-L.; Parikh, I.; Levine, H. *FEBS Lett.* **1992**, *297*, 297.
17. Zhang, Y.; Wisner, A.; Xiong, Y. L.; Bon, C. *J. Biol. Chem.* **1995**, *270*, 10246.
18. Parry, M. A.; Jacob, U.; Huber, R.; Wisner, A.; Bon, C.; Bode, W. *Structure* **1998**, *6*, 1195.
19. Castro, H. C.; Silva, D. M.; Craik, C.; Zingali, R. B. *Biochim. Biophys. Acta* **2001**, *1547*, 183.
20. Magalhães, A.; Monteiro, M. R.; Magalhães, H. P. B.; Marcos-Guia, M.; Rogana, E. *Toxicol.* **1997**, *35*, 1549.
21. Massova, I.; Pirkle, H.; Edwards, B. F. P.; Mobashery, S. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 3139.
22. Bartunik, H. D.; Summers, L. J.; Bartsch, H. H. *J. Mol. Biol.* **1989**, *210*, 813.
23. Katz, B. A.; Sprengeler, P. A.; Luong, C.; Verner, E.; Ehröd, K.; Kirtley, M.; Janc, I.; Spencer, J. R.; Breitenbacher, J. G.; Hui, H.; McGee, D.; Allen, D.; Martelli, A.; Mackman, R. L. *Chem. Biol.* **2001**, *8*, 1107.
24. Nishida, S.; Fujimura, Y.; Miura, S.; Ozaki, Y.; Usami, Y.; Suzuki, M.; Titani, K.; Yoshida, E.; Sugimoto, M.; Yoshioka, A.; Fukui, H. *Biochemistry* **1994**, *33*, 1843.
25. Hahn, B. S.; Yang, K. Y.; Park, E. M.; Chiang, I. M.; Kim, Y. S. *J. Biochem.* **1996**, *119*, 835.
26. Nikai, T.; Ohara, A.; Komori, Y.; Fox, J. W.; Sugihara, H. *Arch. Biochem. Biophys.* **1995**, *318*, 89.
27. Rawlings, N. D.; Barret, A. J. *Methods Enzymol.* **1994**, *244*, 19–61.
28. Sanderson, P. E. J. *Med. Res. Rev.* **1999**, *19*, 179.
29. Mares-Guia, M.; Shaw, E. *J. Biol. Chem.* **1965**, *240*, 1579.

30. Mares-Guia, M.; Figueiredo, A. F. S. *Biochemistry* **1970**, *9*, 3223.
31. Mares-Guia, M.; Shaw, E.; Cohen, W. *J. Biol. Chem.* **1967**, *242*, 5777.
32. Bode, W.; Schwager, P. *J. Mol. Biol.* **1975**, *98*, 693.
33. Mares-Guia, M.; Nelson, D. L.; Rogana, E. *J. Am. Chem. Soc.* **1977**, *99*, 2331.
34. Recantini, M.; Klein, T.; Yang, C.-Z.; McClarin, J.; Langridge, R.; Hansch, C. *Mol. Pharmacol.* **1986**, *29*, 436.
35. Ortiz, A. R.; Pisabarro, M. T.; Gago, F.; Wade, R. C. *J. Med. Chem.* **1995**, *38*, 2681.
36. Essex, J. W.; Severance, D. L.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **1997**, *101*, 9663.
37. Talhout, R.; Engberts, J. B. F. N. *Eur. J. Biochem.* **2001**, *268*, 1554.
38. Dullweber, F.; Stubbs, M. T.; Musil, D.; Stürzebecher, J.; Klebe, G. *J. Mol. Biol.* **2001**, *313*, 593.
39. Guimarães, C. R. W.; De Alencastro, R. B. *J. Phys. Chem. B* **2002**, *106*, 466.
40. Hammett, L. P. *Chem. Rev.* **1935**, *17*, 125.
41. Sharp, K. A.; Madam, B. *J. Phys. Chem. B* **1997**, *101*, 4343.
42. Rogana, E.; Penha-Silva, N.; Mares-Guia, M. *Brazilian J. Med. Biol. Res.* **1989**, *22*, 1177.
43. Checa, A.; Ortiz, A. R.; Pascual-Tereza, B.; Gago, F. *J. Med. Chem.* **1997**, *40*, 4136.
44. Tapparelli, C.; Metternich, R.; Ehrhardt, C.; Cook, N. S. *Trends Pharmacol. Sci.* **1993**, *14*, 366.
45. Furet, P.; Sele, A.; Cohen, N. C. *J. Mol. Graphics* **1988**, *6*, 182.
46. Di Cera, E.; Guinto, E. R.; Vindigni, A.; Dang, Q. D.; Ayala, Y. M.; Wuyi, M.; Tulinsky, A. *J. Biol. Chem.* **1995**, *270*, 22089.
47. Zhang, E.; Tulinsky, A. *Biophys. Chem.* **1997**, *63*, 185.
48. Dang, Q. D.; Di Cera, E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10653.
49. Wells, C. M.; Di Cera, E. *Biochemistry* **1992**, *31*, 11721.
50. Guinto, E. R.; Vindigni, A.; Ayala, Y. M.; Dang, Q. D.; Di Cera, E. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11185.
51. Lai, M.-T.; Di Cera, E.; Shafer, J. A. *J. Biol. Chem.* **1997**, *272*, 30275.
52. Pineda, A. O.; Savvides, S. N.; Waksman, G.; Di Cera, E. *J. Biol. Chem.* **2002**, *277*, 40177.
53. Huntington, J. A.; Esmon, C. T. *Structure* **2003**, *11*, 469.
54. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673.
55. Jones, D. T. *J. Mol. Biol.* **1999**, *292*, 195.
56. Peitsch, M. C. *Biochem. Soc. Trans.* **1996**, *24*, 274.
57. Perona, J. J.; Tsu, C. A.; Craik, C. S.; Fletterick, R. J. *J. Mol. Biol.* **1993**, *230*, 919.
58. Skordalakes, E.; Dodson, G. G.; Green, D. S.; Goodwin, C. A.; Scully, M. F.; Hudson, H. R.; Kakkar, V. V.; Deadman, J. J. *J. Mol. Biol.* **2001**, *311*, 549.
59. Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714.
60. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
61. van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992.
62. Hootn, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E. *Nature* **1996**, *381*, 272.
63. Dewar, M. J. S.; Zoebisch, E. G.; Horsley, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
64. Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 16385.
65. Kuntz, I. D. *Science* **1992**, *257*, 1078.
66. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
67. McDonald, I. K.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777.
68. Wallace, A. C. *J. Polym. Sci. Polym. Chem. Ed.* **1995**, *33*, 127.
69. Dill, K. A. *J. Biol. Chem.* **1997**, *272*, 701.
70. Ajay Murko, M. A. *J. Med. Chem.* **1995**, *38*, 4953.
71. Wold, S.; Sjöström, M.; Eriksson, L. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109.
72. Leo, A. *Chem. Rev.* **1993**, *93*, 1281.
73. Topliss, J. G.; Edwards, R. P. *J. Med. Chem.* **1979**, *21*, 1238.
74. Kollman, P. A. *Acc. Chem. Res.* **1996**, *29*, 461.
75. Head, R. D.; Smythe, M. L.; Oprea, T. I.; Walker, C. I.; Green, S. M.; Marshall, G. R. *J. Am. Chem. Soc.* **1996**, *118*, 3959.
76. Cramer, R. D.; Bunce, J. D.; Patterson, D. E.; Frank, I. E. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18.
77. Ghose, A. K.; Crippen, G. M. *J. Comp. Chem.* **1986**, *7*, 565.
78. Cormarck, R. M. *J. Royal Stat. Soc.* **1971**, *A134*, 321.
79. Andrews, J. M.; Roman, D. P., Jr.; Bing, D. H. *J. Med. Chem.* **1978**, *21*, 1202.
80. Coats, E. A. *J. Med. Chem.* **1973**, *16*, 1102.



ANEXO 3



large scale to remove remaining impurities of samples [6–9] and, more recently, it has been coupled to other methods to improve qualitative and quantitative proteomic recovery studies [10–12]. In addition to its technological interest, affinity chromatography is also an excellent instrument for understanding the mechanisms underlying enzyme and inhibitor interaction.

Benzamidine is a competitive inhibitor and its derivatives are known to inhibit trypsin-like proteases primarily through the interaction between the amidine moiety and an aspartate residue (D189, in trypsinogen numeration) in the enzyme's specificity pocket known as the S1 subsite [13,14]. Consequently, amidine derivatives, such as *p*-aminobenzamidine (PABA), have been prepared in different supports [15,16] and used as affinity ligands for the removal of proteases and the purification of trypsin and trypsin-like proteases such as thrombin [17,18], thrombin-like proteins [19], plasmin [20,21], urokinase [22], enterokinase [23,24], fibrinogen-clotting enzymes [25], acrosin [26,27], kallikrein [28], collagenase [29], clostripain [30] and parasite proteases [6–8].

The purpose of this study was to investigate qualitative and quantitatively the ability of two commercially available amidine derivatives resins to purify TLSPs from venoms of two important snakes widely distributed in South America. Snakes of the *Bothrops* genus are responsible for high frequency (>90%) of snake bite events in Brazil, while the *Lachesis* genus represents the most important vipers of the Amazon region [31]. We have shown that the two analogous commercial affinity columns used in our study present distinct binding specificities and this was sufficient to identify and purify in a high yield two related TLSPs. An investigation of the possible reasons supporting this amazing specificity was carried out at the atomic level with molecular modeling. In addition, it was demonstrated that the tandem alignment of the two columns, in a definite order, allows the purification of three other different proteins from the *Bothrops jararacussu* venom.

2. Experimental

2.1. Materials

B. jararacussu and *Lachesis muta rhombeata* venoms were obtained from Institute Vital Brazil serpentarium. Centricon P10 and high performance liquid chromatography (HPLC) filters were purchased from Amicon (CA, USA) and benzamidine Sepharose 6B (Am-PABA, Lot 254128) resin from Amersham Biosciences (New Jersey, USA). *p*-Aminobenzamidine-agarose (Si-PABA, Lot 96H9534), low range molecular weight markers, α -*N*-benzoyl-DL-arginine-*p*-nitroanilide (BAPNA) and 4-aminobenzamidine hydrochloride were from Sigma–Aldrich Química (Brasil). Standard molecular weight markers were from Biolabs Laboratories (New England, UK). Polyvinylidene difluoride

(PVDF) membrane was from Bio-Rad, sequence reagents from Shimadzu (Kyoto, Japan) and all other chemicals were of reagent grade.

2.2. Chromatographic assay

Chromatographic assays were performed on two commercial supports (Si-PABA and Am-PABA) grafted with the same amidine derivative but differing in the ligand coupling method reported by the manufacturer. Am-PABA contains *p*-aminobenzamidine covalently attached to Sepharose 6B by the epoxy coupling method and presents a spacer arm of 12 atoms [1,4-bis(2,3-epoxypropyl)butane] [32] while the Si-PABA presents the inhibitor coupled to agarose through the CNBr method and carries a spacer arm of eight atoms (glycylglycine) [33]. The same amount (500 mg) of each venom (*B. jararacussu* and *L. m. rhombeata*) was dissolved separately in water (6.0 mL), centrifuged (10,000 \times g, 60 min, 15 °C) and the supernatant buffer exchanged against 50 mM Tris–HCl buffer, pH 7.5, containing 0.5 M NaCl using Centricon-P10 filters. After filtration on 0.22 μ m filters, a half (about 250 mg) of each venom was applied separately to each column (3 mL packed resin) or to both columns mounted in tandem (Si-PABA followed by Am-PABA). All columns were previously equilibrated with 50 mM Tris–HCl, 0.5 M NaCl, pH 7.5 and after sample application, the columns were exhaustively washed in the same buffer. Those mounted in tandem were disconnected and eluted separately. Elution of bound protein from columns was routinely carried out with 1 mM HCl, pH 3.0, containing 0.5 M NaCl at a flow rate of 1 mL min⁻¹ and room temperature (25 °C). The protein content of the fractions was monitored by the absorbance change at 280 nm and the concentration estimated according to Lowry's method [34], using bovine serum albumin as a standard.

2.3. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using 12% polyacrylamide gels in Laemmli buffers [35] under reductive conditions. The gels were Coomassie blue G-250 or silver stained (Bio-Rad Laboratories, CA, USA). Rabbit muscle fructose-6-phosphate kinase (84.0 kDa), bovine serum albumin (66.2 kDa), chicken egg ovalbumin (45.0 kDa) and bovine erythrocyte carbonic anhydrase (29.0 kDa) were used as molecular weight markers.

2.4. Sequence analysis

NH₂-terminal amino acid sequence of peptide samples was determined by automated Edman degradation using a gas-phase protein microsequencer (Model PSQ-1; Shimadzu, Kyoto, Japan). N-terminal sequence analysis of the protein was performed on samples electrotransferred on PVDF membranes after SDS-PAGE and treated with 5 μ L of the mixture

of 0.6% tri-*n*-butylphosphine and 1.2 μL of 4-vinylpyridine in 80% isopropanol for simultaneous reduction and alkylation.

2.5. Enzymatic assays

Amidolytic activity assays were measured with BApNA as a substrate [36]. Briefly, upon addition of enzymes (100 μg), substrate digestion was followed by measuring the change in absorbance at 410 nm (BApNA). The substrate was digested in 50 mM Tris–HCl, pH 8.0, at 25 °C for 30 min. The activity was calculated using a molar absorption of $8800\text{M}^{-1}\text{cm}^{-1}$ [37].

2.6. Inhibition of *B. jararacussu* TLSP by benzamidine and *p*-aminobenzamidine

The initial rates of enzyme-catalyzed BApNA hydrolysis were estimated from reaction progress curves at 37 °C measured by the change in absorbance at 410 nm read in a microplate spectrophotometer (SPECTRAMax PLUS³⁸⁴, Molecular Devices, CA). Reaction mixtures (200 μL) consisting of 100 μL 2 \times buffer (100 mM Tris–HCl pH 8.0) + 40 μL substrate solution + 40 μL inhibitor solution were distributed as serial substrate dilutions in 96-well plates. The later were then pre-incubated at 37 °C prior to triggering reactions by the addition of 20 μL enzyme solution containing 1 μg protein. Substrate concentrations ranged from 93 to 2500 μM BApNA whereas inhibitor concentrations ranged from 10 to 1000 μM *p*-aminobenzamidine or benzamidine. At least three inhibitor concentrations and five substrate concentrations per inhibitor concentration (in addition to the control without inhibitor) were assayed. In order to estimate inhibition constants (K_i) it was employed the “ $K_{M,app}$ method” as previously described [38].

2.7. Modeling the “active” part of the immobilized ligands

Based on the docked complex between a thrombin-like enzyme from *Lachesis muta muta* (Lmm-TLE) and 4-carboethoxybenzamidine [39] we have modeled the “active” or functional portion of the ligand immobilized on the sorbents studied in this work. The “active” portion was modeled as the 4-aminobenzamidine core attached to the first four (Am-PABA) or five atoms (Si-PABA) donated by the spacer arm, in which the terminal group was truncated as a neutral methyl group. The size of the “active” portion of the affinity sorbent was selected based on: (i) the degree of similarity between the ligands/spacers immobilized on the resins; (ii) prior knowledge about the conformation of bound thrombin inhibitors and (iii) previous data obtained by our group [39]. A hundred steps of steepest descent followed by 1000 steps of conjugated gradient energy minimization were performed to optimize the geometry of the Lmm-TLE ligands, constructed as described above, already docked in a reasonable confor-

mation at the enzyme active site. For the steepest descent procedure, only enzyme atoms at 2.0 Å distance from any atom of the ligand were allowed to move along the minimization, while enzyme atoms at 6.0 Å radius from ligand atoms were restricted to contribute for energy evaluations during the minimization, but not allowed to move. For the conjugated gradient part of the minimization schedule, these distances were extended to 8.0 and 12.0 Å, respectively.

2.8. Docking

The structures minimized as described in later section were used as inputs with optimized geometries for the program FlexX. This program is an efficient flexible docking method that uses an incremental construction algorithm to place ligands into an active site [40]. Briefly, the ligand is constructed using an incremental approach in which a base fragment (the ligand core) is automatically selected and is placed into the active site (user-defined) followed by the attachment of other remaining ligand fragments, until reconstruction of the whole molecule is completed. In particular, the conformational flexibility of the ligand is covered by generating multiple conformations for each fragment and including all in the ligand building steps. Ligands were extracted from the minimized structures and the “active-site” for ligand docking was constituted by receptor atoms 6.5 Å away from any ligand atom. None restriction was done for positioning ligands and other program parameters were default. Along with the scoring functions implemented in FlexX, functional forms included in CS-Score module were used to create a consensus score and rank the top 15 more stable configurations of ligands docked at Lmm-TLE. From these, two configurations of each ligand were chosen for the subsequent detailed study and were minimized according to the previously described protocol. In addition to a favorable score, configurations that did retain the ubiquitous salt bridge between the D189 carboxylate and the amidinium group of S1-directed inhibitors were preferred over configurations that did not. All calculations were performed using Sybyl v6.8 molecular modeling package (Tripos Inc., Illinois, USA) running on the SGI Origin 2000 server installed at the Scientific Computation Program from Oswaldo Cruz Foundation (PROCC-FIOCRUZ).

3. Results and discussion

3.1. Purification of fibrinogenases

Both *B. jararacussu* and *L. m. rhombeata* venoms are used as immunogens to obtain bothropic and lachetic horse antivenom products widely used to protect patients during an envenomation accident [31]. Serine proteases are among the venom components causing the most deleterious effects on envenomed victims. These enzymes interfere with specific points of the coagulation cascade of vertebrates, impairing blood clotting and inducing hemorrhagic disorders along with

cardiovascular failure. Therefore, in order to investigate the properties of different benzamidine derivatives to purify or remove this interesting group of proteases, the venom was fractionated using identical chromatographic conditions on two commercial supports (Si-PABA and Am-PABA) grafted with the same amidine. A schematic representation of the multiple venom proteins isolation using the two benzamidine sorbents mounted in tandem is shown in Fig. 1. As seen in the Fig. 1a the Si-PABA sorbent retained a single broad band with 48–52 kDa from the *B. jararacussu* venom (BJ-48) while no band from the *L. m. rhombeata* venom was retained (Fig. 1b). On the other hand when the *B. jararacussu* venom is fractionated first by Am-PABA, a total of four proteins (48, 32, 18 and 6–8 kDa) are retained by this column (data not shown). As expected, when the Am-PABA is connected after the Si-PABA column, it does not retain the 48–52 kDa protein but retains the three other bands from the venom of *B. jararacussu* (Fig. 1e). In addition, this same sorbent is able to adsorb a single broad band with 45–47 kDa from the venom of *L. m. rhombeata* (LMR-47, Fig. 1f). Therefore, the alignment of Si-PABA followed by Am-PABA has the advantage to purify the BJ-48 and isolate the three other proteins on the subsequent column in a single step of fractionation.

The BJ-48 presented as its N-terminal the sequence **VVGGDXIPQVPFLAF**, whereas the LMR-47 the sequence **VIGGDEXNINEHRFLVALYDGLSGT**. Both proteins are serine proteinases and had been previously characterized by us using different approaches [36,41]. LMR-47 is also probably identical to the reported 41–47 kDa SDS-PAGE migrating species that was characterized as a thrombin-like enzyme/gyroxin analogue from the venom of *L. muta muta* [42]. The 18 kDa protein (**NEDEQTRVPKEKFLSSK-TYTKW**) was also identified as a segment from a serine protease, however the 32 kDa protein (**NNCPQDWLPM-NGLCYKIF DE**) presented high similarity with a recently identified galactose-binding protein from *B. jararacussu* venom [43]. Data listed in Table 1 demonstrate that the affinity approach used here to purify the LMR-47 protein was two times superior in recuperations and faster than that used previously by means of an isoelectrofocusing preparative method [36]. In addition, these results showed that despite both enzymes belong to the serine proteinase class and are associated with similar biological functions they present a selective binding to the two sorbents.

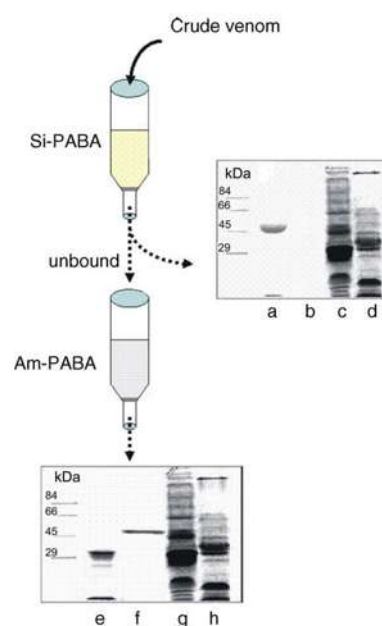


Fig. 1. Schematic representation of the multiple venom proteins isolation using the two benzamidine sorbents mounted in tandem. Venoms were applied on Si-PABA column and the non-adsorbed proteins were passed through the Am-PABA column. After exhaustively washing with the equilibrating buffer (50 mM Tris-HCl, 0.5 M NaCl, pH 7.5), the two columns were disconnected and the retained proteins eluted with 1 mM Tris-HCl buffer, pH 3.0, containing 0.5 M NaCl. The proteins were analyzed by SDS-polyacrylamide gel electrophoresis (12%). Lanes: (a) *B. jararacussu* protein retained by Si-PABA column; (b) Si-PABA column eluate; (c and g) total *B. jararacussu* venom proteins; (d and h) total *L. m. rhombeata* venom proteins; (e) *B. jararacussu* proteins retained by Am-PABA column; (f) *L. m. rhombeata* protein retained by Am-PABA column. Rabbit muscle fructose-6-phosphate kinase (84 kDa), bovine serum albumin (66 kDa), chicken egg ovalbumin (45 kDa) and bovine erythrocyte carbonic anhydrase (29 kDa) were used as molecular weight markers.

3.2. Studies of sorbent binding specificity

In order to evaluate the properties of both sorbents we have investigated the effect of the ionic strength, temperature and the amino acid competitor arginine on the elution of the *B. jararacussu* 48 kDa protein from Si-PABA. Elution of BJ-48 with buffer solutions bearing increasing salt concentration is

Table 1
Purification of fibrinogenases (thrombin-like serine proteases) from *Bothrops jararacussu* (BJ-48) and *Lachesis muta rhombeata* (LMR-47) venoms

Protease	Step	Total protein		Amidolytic activity		Purification (fold)	Yield (%)
		(mg)	(%)	Total (units)	Specific ^a (units/mg)		
BJ-48	Centricon-P10	248	100	1.454	0.101	1.00	100
	Si-PABA	6.5	2.6	0.111	0.297	2.92	76.2
LMR-47	Centricon-P10	246	100	1.570	0.060	1.00	100
	Am-PABA	5.3	2.0	0.100	0.180	3.00	62.5

^a One unity of amidolytic activity stands for 1 μ mol of α -N-benzoyl-DL-arginine-p-nitroanilide hydrolyzed per minute, at 25 °C in 50 mM Tris-HCl, pH 8.0 buffer.

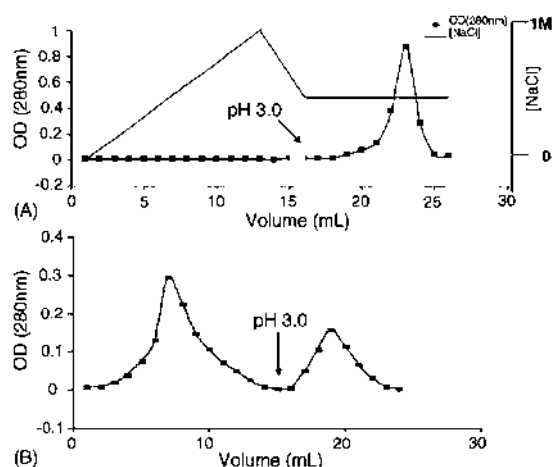


Fig. 2. Physicochemical properties of BJ-48 interaction with *p*-aminobenzamidine-agarose (Si-PABA). (A) Elution of BJ-48 adsorbed to Si-PABA resin (3 mL gel) by NaCl at various concentrations prepared in 50 mM Tris-HCl, pH 7.0, at a flow rate of 1 mL min⁻¹ and 25 °C followed by 1 mM HCl, pH 3.0, containing 0.5 M NaCl. (B) Elution of BJ-48 by arginine (0.5 M) prepared in 50 mM Tris-HCl, 0.5 M NaCl, pH 7.0, at a flow rate of 1 mL min⁻¹ and 25 °C followed by 1 mM HCl, pH 3.0, containing 0.5 M NaCl.

presented in Fig. 2A. BJ-48 is adsorbed at low ionic strength and cannot be desorbed by increasing the ionic strength of the eluent. This result is identical to that obtained by others [17] using bovine trypsin on dextran-coated silica supports. This observation indicates that the interaction between the proteins and the support functionalized with Si-PABA is highly specific and cannot be disrupted by high ionic strengths as would be expected if the interaction of the enzyme with the sorbent was solely due to the salt bridge between D189 and the positively charged amidine function [44].

Regarding the binding of snake proteins at different temperatures, we conducted one additional experiment at 6 °C. Neither qualitative nor quantitative difference was observed within the BJ-48 purification either if the purification experiments were developed in a cold room at 6 °C or at room temperature (25 °C, data not shown). Trypsin has been purified with *m*- or *p*-aminobenzamidine agarose at 6 °C because of the limitation of using agarose at high flow-rates and the tendency of trypsin to undergo autodegradation during purification. This is a fact not normally observed with the BJ-48 and LMR-47 proteins. As expected, BJ-48 desorption from the stationary phase took place by the competitor arginine (Fig. 2B). BJ-48 molecules adsorbed on Si-PABA were eluted with 0.5 M arginine solution, indicating that probably the same type of interactions occurred between the BJ-48 and the competing amino acid. Thus, the results described herein clearly show that the BJ-48 protein adsorb specifically to the Si-PABA since: (i) the two TLSPs (BJ-48 and LMR-47) present distinct affinity and behavior in the two sorbents and (ii) the elution of BJ-48 from Si-PABA occurs

by the competitor amino acid arginine and was temperature independent.

3.3. Computer-aided evaluation of the interaction between Lmm-TLE and the functional portion of sorbents at the atomic level

Evidence has accumulated in our lab that has allowed us to safely presume that the 45–47 kDa species isolated from *L. m. rhombata* venom using chromatography on Am-PABA resin is virtually identical to the orthologue from *L. m. muta* venom that had its primary structure determined [36,45]. In fact, the very existence of these subspecies seems not to be a consensus among herpetologists [46]. Therefore, we employed a model of Lmm-TLE previously obtained by our group [39] to study the interaction of the ligands immobilized on the resins with one of the proteins they bind specifically (LMR-47). Since both resins are claimed to be *p*-aminobenzamidine by their respective suppliers these were expected to present similar properties. Nevertheless, these sorbents presented striking distinct selectivity for the snake venoms proteins they are able to bind. One possible explanation for this observation could be based in possible differences in the affinity constants presented by the distinct enzymes for *p*-aminobenzamidine. However, inhibition data listed in Table 2 show that BJ-48 binds to this S1-directed inhibitor with comparable strength to Lmm-TLE/LMR-47. In addition, both enzymes favor, almost to the same extent, the presence of the amino group at position 4 of the benzamidine ring in order to increase binding affinity. This indicates that the S1 binding pocket of these enzymes is indeed very similar. These data encouraged us to assume that the distinct binding specificities shown by the commercial affinity sorbents based on *p*-aminobenzamidine have necessarily to be explained by the different spacer arms used to couple the *p*-aminobenzamidine moieties to the inert polymeric support (Fig. 3). Hence, we have isolated and constructed a model for the putative "active" part of the Si-PABA and Am-PABA resins and examined their most probable binding mode to Lmm-TLE.

Analysis of the top 15 configurations furnished by FlexX for each ligand docked on Lmm-TLE suggested that these could be gathered into two groups. For the ligand representing

Table 2
Comparison of inhibition constants (K_i) of S1-directed inhibitors of thrombin-like serine proteases (TLSPs) from *Bothrops jararacussu* (BJ-48) and *Lachesis muta muta* (Lmm-TLE) venoms

S1-directed inhibitor	K_i (μ M)	
	BJ-48 ^a	Lmm-TLE ^b
Benzamidine	110.8 \pm 0.3	53.7 \pm 0.1
<i>p</i> -Aminobenzamidine	17.0 \pm 0.3	13.3 \pm 0.4

^a Inhibition was observed over reaction progress curves for α -N-benzoyl-DL-arginine-*p*-nitroamide (BAPNA) hydrolysis at 37 °C in 50 mM Tris-HCl, pH 8.0. Values shown along with standard errors.

^b Constants at 25 °C using BapNA as substrate [45]. Values are shown as means \pm standard deviations.



the Am-PABA “active” part, one kind of configuration, represented by the third ranked conformer was characterized by: (i) a hydrogen bond between the *p*-aminobenzamidine NH (acceptor) and OH from S195 (donor); (ii) the projection of the terminal methoxyl group to the lipophilic region composed by W215, F214, Y228 and V227; and (iii) a polar interaction between the side chain hydroxyl and H57 (Fig. 4A). The flexibility furnished by the three freely rotatable bonds in the side chain from this ligand result in a number of variations of this archetypal binding mode. In one of these, an additional stabilizing hydrogen bond between the hydroxyl moiety and H57 could be attained (Fig. 4A, right panel). This later configuration (fifth in rank, Am-5) was selected for further energy minimization and detailed analysis. The second selected configuration (13th in rank, Am-13) differs from the first in that the directionality of the hydrogen bond between the NH and S195 was inverted as well as methoxyl points to the S1' subsite and makes a hydrogen bond with H57 while the hydroxyl points to the solvent (Fig. 4A, left panel). In the case of the ligand mimicking the Si-PABA sorbent, the glycyl-glycine side chain brought by the spacer arm makes a beta-sheet with the polypeptide segment from residues 192

through 194, as exemplified by the second ranked configuration (Si-2). The first carbonyl (counting from the benzamidine ring) makes a hydrogen bond with the backbones of R193 and D194, in addition to the OH from S195 (Fig. 4B, right panel). The second carbonyl makes a hydrogen bond with R60, which delimits the S1' subsite walls, while the terminal methyl points to the hydrophobic S1' floor, composed by L41 and C42. A variant of this configuration (seventh in rank, Si-7), in which the second carbonyl is inverted, makes a hydrogen bond with NH of R193 backbone (Fig. 4B, left panel).

We decided to inspect in detail at least two configurations of each resin surrogate because we felt that a premature decision about which binding mode would be more favorable could be misleading. Analysis of the final minimized configurations of the Am-PABA resin surrogate showed that both would represent equally acceptable binding modes for this ligand. Am-5 differs from Am-13 mainly in the fact that the methoxyl group in the former is oriented to S2/S3 subsites making additional van der Waals (vdW) contacts with W99 and R174, whereas this moiety in Am-13 is facing the S1' subsite, promoting two hydrogen bonds with the OH from

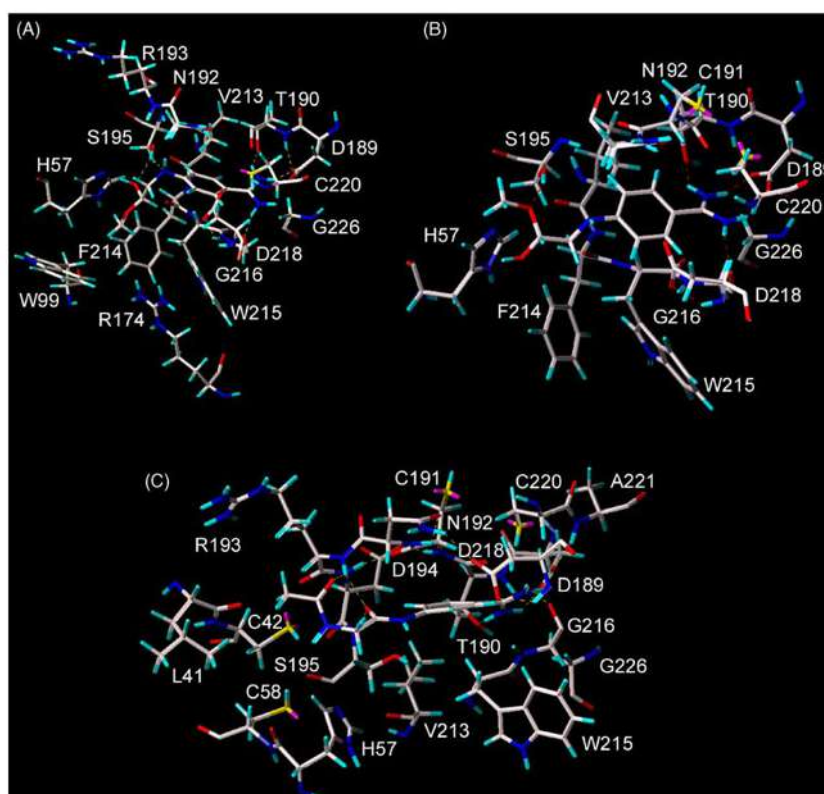


Fig. 5. Analysis of the minimized structures for the complexes between *Lachesis muta* thrombin-like enzyme (Lmm-TLE) and the resin surrogates on their most likely binding modes: (A) Am-5; (B) Am-13 and (C) Si-7. Enzyme residues and ligands colored according to CPK code (oxygen: red, carbon: gray, nitrogen: blue and sulfur: yellow).

S195, one through methoxyl oxygen as an acceptor and the other through the adjoining OH as a donor (Fig. 5A). This moiety in Am-13 is oriented to H57, but makes no hydrogen bond since the appropriate geometry cannot be achieved due to steric restrictions (Fig. 5B). Am-13 also makes vdW contacts with the R193 backbone atoms, while Am-5 does so with C191 backbone. Although the amidinium group is fixed in the face to face salt bridge with D189 carboxylate, the aromatic ring of one ligand is torsioned near 30° in relation to the other. The two studied configurations of the Si-PABA surrogate are almost identical, differing only in the ϕ dihedral of glycyl-glycine. In Si-7 the terminal carbonyl makes a hydrogen bond with NH from R193 backbone (Fig. 5C). Si-2 makes additional contacts with the guanidine group of R60 and the F214 backbone. Due to the better beta-sheet that can be accommodated by the Si-7 configuration we felt comfortable to propose that this configuration would represent the most probable binding mode for the Si-PABA surrogate.

A comparison between the binding mode of the Si-PABA surrogate ligand and the two possible binding modes for Am-PABA (Am-5 and Am-13) turns possible to speculate about why should LMR-47/Lmm-TLE bind to the later resin and not to the former. Si-PABA makes exclusive vdW interactions with L41 and C42 in S1', C58 in S3, A221 and D194 while Am-5 makes exclusive contacts with R174 guanidinium group, W99 side chain, F214 backbone and R193 backbone (Fig. 6). The Si-PABA surrogate ligand is able to make a total of three hydrogen bonds in addition to the bonds made by the amidinium moiety. Am-5 makes two hydrogen bonds and Am-13 makes only one. Since we have showed that LMR-47 is able to bind only to Am-PABA, none of these interactions can be invoked to explain the observed selectivity. Equivalently, the total contact area of these ligands cannot be summoned to explain the observed chromatographic behavior since Si-PABA has 380 \AA^2 compared to

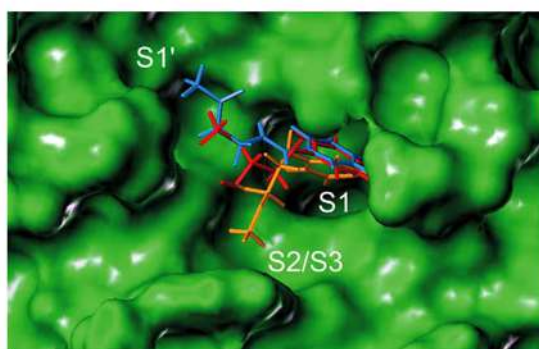


Fig. 6. Superposition of the putative binding modes of Am-PABA and Si-PABA surrogates over the *Lachesis muta muta* thrombin-like enzyme (Lmm-TLE) molecular surface. Ligands are shown as sticks (Am-5: orange, Am-13: red and Si-7: cyan). Enzyme surface (green) was calculated using a 1.4 \AA probe radius. Location of specificity pockets known as S1, S1', S2 and S3 are indicated.

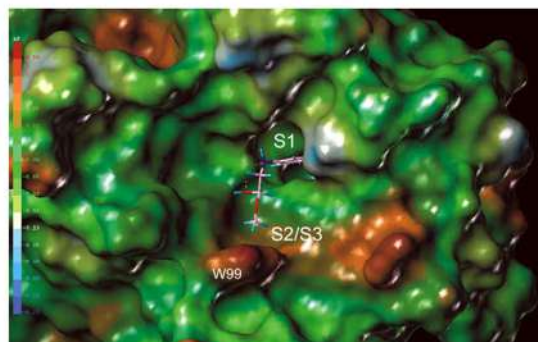


Fig. 7. Lipophilic potential of the *Lachesis muta muta* thrombin-like enzyme (Lmm-TLE) molecular surface in complex with the most likely binding-mode for the Am-PABA surrogate. High lipophilic areas are shaded in brown while high hydrophilic regions are shown in bluish tones (a lipophilicity scale is shown at left). Ligand is shown as CPK colored sticks with hydrogen bonds represented as dashed lines. The enzyme residue responsible for the high lipophilicity of the pocket receiving the methoxyl moiety of the ligand is detached.

$361\text{--}367 \text{ \AA}^2$ for the Am-PABA surrogate in the two possible conformations. The answer to this apparent contradiction must come from an unequal contribution of one of the interactions made by these ligands. The most peculiar contacts of the three analyzed complexes is the hydrophobic interaction of Am-PABA with the S3/S2 lipophilic patch, as can be observed in the binding mode represented by Am-5 configuration (Fig. 7). The importance of this interaction in defining the striking preference of LMR-47 for Am-PABA is in accordance to results recently obtained by us, which demonstrate the importance of the lipophilic potential to achieve high binding affinity to the S1 subsite of trypsin-like serine proteases [39].

3.4. Coupling of the two derivative columns to purify specific proteins

Our results indicated that the two columns possess distinct adsorptive capacity of venom proteins in spite of bovine trypsin and human plasmin being similarly retained by both columns (data not shown). Thus, in order to improve the purification process of these thrombin-like serine proteinases, the two columns were combined in tandem (using the same equilibrating buffer) as showed in the Fig. 1 (Si-PABA followed by Am-PABA). Using this procedure the BJ-48 kDa protein was obtained in high yield (2% of all venom proteins) in the first column (Si-PABA), while the second flow-through column was devoid of this enzyme and contained the 32 and 18 kDa proteins. This approach facilitated the fractionation of these additional proteins by the second column (Am-PABA) which could be easily purified to homogeneity using a further step of anion exchange high performance liquid chromatography [data not shown]. The experiments carried out with the *Lachesis* venom demonstrated the specificity of the sorbents.

4. Conclusion

We have demonstrated that similar commercial amidine derivatives supports display a strict selectivity and this could be used to purify serine proteases from snake venom with similar biological properties, in particular the BJ-48 from *B. jararacussu* and the LMR-47 from *L. m. rhombeata* venom. This specificity is probably due to the distinct chemical structures of the PABA spacer from the different sorbents, which contribute with dissimilar hydrophobic contacts to substrate binding clefts of the enzymes. The later translates into the particular chromatographic behavior observed, where BJ-48 elutes as single band from the Si-PABA sorbent while LMR-47 is exclusively adsorbed by the Am-PABA sorbent. Thus, individually the columns promote the purification of both proteins in a high yield and when aligned in tandem the isolation of other three proteins from *B. jararacussu* venom is possible.

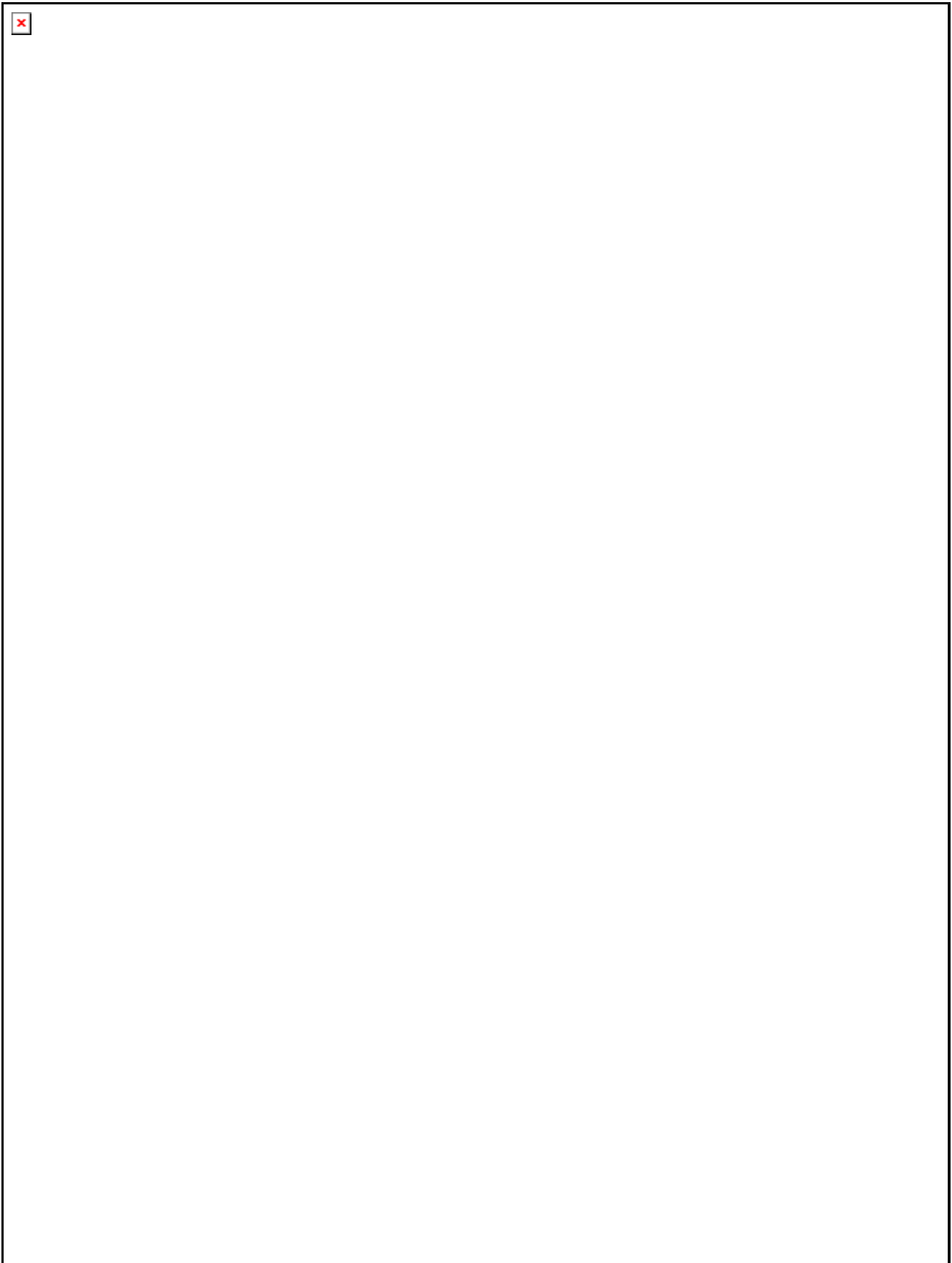
Acknowledgments

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/MCT-Brazil) and FIOCRUZ for financial support. FPS Jr is a CNPq-Post-graduate Fellowship recipient.

References

- [1] H.C. Castro, R.B. Zingali, M.G. Albuquerque, M. Pujol-Luz, C.R. Rodrigues, *Cell. Mol. Life Sci.* 61 (2004) 1.
- [2] S.Y. Cho, B.-S. Hahn, K.Y. Yang, Y.S. Kim, *Toxicon* 39 (2001) 499.
- [3] Y. Jin, Q.-M. Lu, J.-F. Wei, D.-S. Li, W.-Y. Wang, Y.-L. Xiong, *Toxicon* 39 (2001) 1203.
- [4] S.M.T. Serrano, C.A.M. Sampaio, R. Mentele, A.C.M. Camargo, E. Fink, *Thromb. Haemost.* 83 (2000) 438.
- [5] S. Nishida, Y. Fujimura, S. Miura, Y. Ozaki, Y. Usami, M. Suzuki, K. Titani, E. Yoshida, M. Sugimoto, A. Yoshioka, H. Fukui, *Biochemistry* 33 (1994) 1843.
- [6] S.S.M. Junior, C.O.A. Vianna, I. Ennes, S.A. Gomes, M.S. Freire, E.D. Silva, S.G. De-Simone, M.T.B. Moraes, *J. Chromatogr. B* 787 (2003) 303.
- [7] R.E. Silva-Lopez, S.G. De-Simone, *Exp. Parasitol.* 107 (2004) 173.
- [8] R.E. Silva-Lopez, J.A. Morgado-Diaz, C.R. Alves, S.G. De-Simone, *Parasitol. Res.* 93 (2004) 328.
- [9] R.E. Silva-Lopez, S.G. De-Simone, *Z. Naturforsch.* 59C (2004) 590.
- [10] W.C. Lee, K.H. Lee, *Anal. Biochem.* 324 (2004) 1.
- [11] J.V. Coumans, I. Humphery-Smith, C.G. dos Remedios, *Electrophoresis* 18 (1997) 1079.
- [12] R.K. Rasmussen, H. Ji, J.S. Eddes, R.L. Moritz, G.E. Reis, R.J. Simpson, D.S. Dorov, *Electrophoresis* 19 (1998) 809.
- [13] M. Mares-Guia, E. Shaw, *J. Biol. Chem.* 240 (1965) 1579.
- [14] H.F. Hixon, A.H. Nishikawa, *Arch. Biochem. Biophys.* 154 (1973) 501.
- [15] M. Ellouali, S. Khamlichi, J. Jozefonvicz, D. Muller, *J. Chromatogr.* 548 (1991) 255.
- [16] K. Nakamura, T. Suzuki, M. Hasegawa, Y. Kato, H. Sasaki, K. Inouye, *J. Chromatogr. A* 1009 (2003) 133.
- [17] K. Mizuno, K.T. Nakamura, K. Takada, S. Sakakibara, H. Matsuo, *Biochem. Biophys. Res. Commun.* 144 (1987) 807.
- [18] G. Schner, *Hoppe-Seyler's Z. Physiol. Chem.* 353 (1972) 810.
- [19] W.H. Holliman, L.J. Weiss, *J. Biol. Chem.* 251 (1976) 1663.
- [20] S.S. Huscin, *Arch. Biochem. Biophys.* 285 (1991) 373.
- [21] M.S. Runge, C. Bode, G.R. Matsueda, E. Haber, *Proc. Natl. Acad. Sci. U.S.A.* 84 (1987) 7659.
- [22] T. Someno, T. Saino, K. Karoh, H. Miyazaki, S. Ishii, *J. Biochem.* 97 (1985) 1493.
- [23] A. Kanamori, N. Seno, I. Matsumoto, *J. Chromatogr.* 363 (1986) 231.
- [24] D. Grant, A. Magge, J. Hermon-Taylor, *Eur. J. Biochem.* 88 (1978) 183.
- [25] M.L. Bender, M. Beuge-Canton, R. Blakely, L. Brubacher, J. Feder, C. Gunter, F. Kezdy, J. Killeffer Jr., T. Marshall, C. Miller, R. Roeske, J. Stoops, *J. Am. Chem. Soc.* 88 (1966) 5890.
- [26] C.R. Brown, *J. Reprod. Fertil.* 69 (1983) 289.
- [27] A.J. Adairan, T. Shoshani, M. Minuth, J.A. Award, J.B. Elce, R.A. Johnson, *Biol. Reprod.* 52 (1995) 490.
- [28] J.P. Girolami, C. Pecher, J.L. Bascands, G. Cabos, A. Adam, *J.M. Soc. Prep. Biochem.* 19 (1989) 75.
- [29] M. Kroger, H. Tschesche, *Gene* 196 (1997) 875.
- [30] N.E. Labrou, D.J. Rigden, *Eur. J. Biochem.* 271 (2004) 983.
- [31] B. Lomonte, *Toxicon* 23 (1985) 173.
- [32] Benzamidine Sepharose 4 Fast Flow Instruction, Amersham Biosciences (2000).
- [33] Bulletin A7155, *p*-Aminobenzamidine-agarose, Sigma Production Department.
- [34] O.H. Lowry, N.J. Rosebrough, L.A. Farr, R.J. Randall, *J. Biol. Chem.* 193 (1951) 265.
- [35] U.K. Laemmli, *Nature* 227 (1970) 680.
- [36] A.S. Aguiar, C.R. Alves, A. Melgarejo, S.G. De-Simone, *Toxicon* 34 (1996) 555.
- [37] B.F. Earlander, N. Kokowsky, W. Cohen, *Arch. Biochem. Biophys.* 95 (1961) 271.
- [38] T. Kakkar, H. Boxenbaum, M. Mayersohn, *Drug Metab. Dispos.* 27 (1999) 756.
- [39] F.P. Silva Jr., S.G. De-Simone, *Bioorg. Med. Chem.* 12 (2004) 2571.
- [40] B. Kramer, M. Rarey, T. Lengauer, *Proteins: Struct. Funct. Genet. Suppl.* 1 (1997) 221.
- [41] F.P. Silva Jr., H.L.M. Guedes, C. Oliveira, S.C. Bourguignon, I.L. Torriani, S.G. De-Simone, *FEBS J.*, submitted for publication.
- [42] A.M.V. Silveira, A. Magalhães, C.R. Diniz, E.B. De Oliveira, *Int. J. Biochem.* 21 (1989) 863.
- [43] D.D. de-Carvalho, S. Marangoni, J.C. Novello, *J. Prot. Chem.* 21 (2002) 43.
- [44] R. Talhout, J.B.F.M. Engberts, *Eur. J. Biochem.* 268 (2001) 1554.
- [45] A. Magalhães, B.C.B. Da Fonseca, C.R. Diniz, J. Gilroy, M. Richardson, *FEBS Lett.* 329 (1993) 116.
- [46] K.R. Zamudio, H.W. Greene, *Biol. J. Linn. Soc.* 62 (1997) 421.

ANEXO 4



ARTICLE IN PRESS

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

281

inhibited by TFPI or ATIII. In the final step of the coagulation cascade, the release of fibrinopeptides A and B by the action of thrombin over fibrinogen α and β chains, respectively, leads to the formation of fibrin monomers. An important regulatory pathway of the coagulation process results from the inactivation of factors Va and VIIIa by activated Protein C (aPC). The formation of the later is catalyzed by a complex between thrombin and thrombomodulin, a transmembrane protein from the endothelium.

Thrombin is a multifunctional protease, central to the coagulation process in mammals and other vertebrates. As such, thrombin must be under very tight control and consequently must interact with different substrates and molecular modulators. This diversity of functional interactions is also present in other serine proteases operating in the coagulation cascade, such as aPC and factors VIIa, IXa and Xa. Overall, efficiency of the coagulation system requires a rapid and localized response at the injury site that can be readily terminated. This is accomplished through a plethora of regulatory mechanisms among which sodium mediated allostery has emerged as the simplest and most intriguing [4]. The physiological role of sodium as a thrombin allosteric modulator was disclosed short after the characterization of the Na^+ dependent conformational change and activation of the enzyme was reported [5]. The fast form that is generated after Na^+ binding is considered procoagulant primarily because it cleaves fibrinogen with higher specificity than Protein C.

Oppositely, the Na^+ -free, slow form, is thought to be anticoagulant due to its higher specificity for Protein C in comparison to fibrinogen.

The activating effects of Na^+ over the hydrolytic activities of the coagulation proteases, FXa, thrombin and aPC, has been known for many years [6–8]. The coagulation factors FVIIa and FIXa are suspected to shelter a Na^+ binding site based upon the Dang and Di Cera's rule [9]. This rule was elaborated after the canonical thrombin Na^+ binding site (Fig. 1) was identified crystallographically through rubidium substitution [5,10]. It states that serine proteases of the chymotrypsin family (clan S1A) carrying a phenylalanine or a tyrosine residue at position 225 can properly provide a Na^+ binding site whereas family members presenting P225 cannot. Indeed, it was demonstrated that FIXa displays a 4-fold enhancement in specificity for synthetic amide substrates upon Na^+ addition to the reaction media [11].

The specificity of coagulation serine proteases is primarily dictated by the complementary interactions between the major pocket in the substrate binding cleft (the S1 subsite) and the side chain of the residue N-terminal to the scissile bond (P1 residue). Nonetheless, these enzymes display on their surfaces other sites involved in substrate recognition such as the exosite-I in thrombin and the Ca^{2+} binding site found in aPC and coagulation factors VIIa, IXa and Xa. These sites are distributed in analogous regions of the serine protease COOH-terminal domain. Exosite-I is delineated by loops

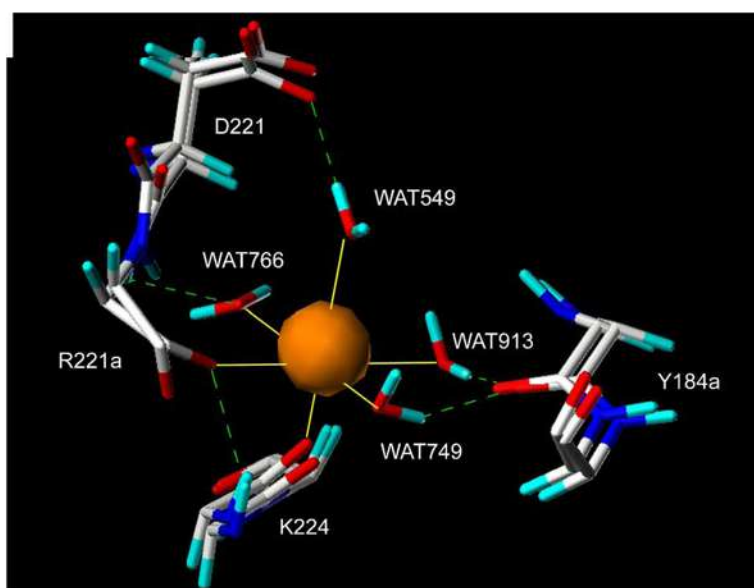


Fig. 1. Comparison of the Na^+ coordinating residues conformations in thrombin fast and slow forms. The Na^+ ion is plotted as an orange van der Waals sphere while protein residues are depicted as CPK-colored capped sticks. For the fast form, water molecules coordinating Na^+ are shown along with hydrogen bonding interactions (green dashed lines; 2 dashes per angstrom) and coordination bonds (yellow lines), which is composed of the carbonyl O atoms of residues 224 and 221a and four water molecules in thrombin or in the case of aPC and FXa by the carbonyl oxygen atoms of residues 185, 184a, 224 and 221a along with two water molecules. The site architecture is largely maintained upon sodium release by the thrombin fast form (pdb 1SG8) and transition to the slow form (pdb 1SGI). One important exception is the loop 1 backbone whose change in conformation culminating at residue R221a (depicted by deviations of 41.4° and 31.9° in ϕ and ψ torsion angles, respectively) moves the carbonyl oxygen atom in the slow form almost 1.0 Å away from its position relative to the fast form.

74–80 as well as part of the adjoining strands (residues 81–82 and 85–87) and is responsible for mediating essential interactions with the thrombin's natural pro-coagulant substrate, fibrinogen [12], as well as for providing a binding site for thrombomodulin [13]. The calcium binding site is formed by structures in the 70–80 loop and is typified by the one first described in trypsin [14]. This site is functionally relevant because it has been shown that a Ca^{2+} ion is required for efficient activation of Protein C by the thrombin–thrombomodulin complex [15,16] while it has also been shown to be a necessary co-factor or to significantly enhance the catalytic activities of factors VIIa and Xa, respectively [17,18].

The unique monovalent metal binding site in sodium activating coagulation enzymes is inserted in a cavity nearly opposite to the entrance of the S1 subsite and can be traced to a region located 15–16 Å away from the catalytic triad. In thrombin and FIXa, a network of water molecules has been described to link the Na^+ binding site to the S1 subsite [10]. Ala-scanning studies accompanied by the structural elucidation of carefully crystallized inhibitor-free R77aA thrombin mutants in the Na^+ bound and unbound states revealed that very subtle and localized conformational changes in loops 1 (184–189) and 2 (221–225) underlie sodium mediated activation [19]. Slow-fast transition results from the formation of the R187:D222 ionic pair allied to the optimal reorientation of D189 and to the significant changes in E192 orientation [20]. All of these changes are coupled to the rearrangement in the water network that connects the cation to S195 in the catalytic site. Interestingly enough, experimental data support an energetic linkage between exosite-I, the Na^+ binding site and the active site [13,21], although a definitive structural basis for this connection is still missing. Similarly, thermodynamic or functional linkage between the S1 subsite, and the Na^+ and Ca^{2+} binding sites has been shown for aPC, FIXa and FXa [22–24].

Recently, using computer-aided molecular surface analysis (CaMSA) our group evidenced the shape of the hydrophilic channel connecting the sodium binding site to the S1 subsite [25]. The importance of shape complementarity in biomolecular complex formation has been fostered by the knowledge accumulated through years studying molecular recognition processes [26]. Hence, in giving the opportunity to begin examining extensive Ala-scanning and other functional data at a more intuitive level, i.e., the molecular surface, we have envisaged that CaMSA would arise as a powerful methodology to bring novel insights on the mechanisms behind Na^+ recognition and allostery.

Here we report the application of CaMSA to the study of the sodium binding channel shapes of the human coagulation proteases (thrombin, factors VIIa, IXa, Xa and aPC). As Na^+ ions are rarely directly identified in crystallographic studies many of the structures reported in the past for coagulation proteases were not characterized as to their Na^+ binding state. Therefore, for the properly unbiased application of CaMSA methodology we had to carefully screen these structures for the presence of an unnoted Na^+ ion as well as other allosteric effectors, such as Ca^{2+} and exosite-I ligands. In addition to

the comparisons of the channel shape features among the enzymes we also describe, for the first time, an analysis of the Na^+ binding sites in factors VIIa and IXa. Except for thrombin, no other enzyme structure was available in both the fast and slow forms. Hence, the application of CaMSA to the characterization of the slow-fast transition had to be limited to thrombin.

2. Experimental

2.1. Structures and characterization of sodium and calcium binding sites

All molecular manipulations and calculations described in this section were performed in Sybyl v6.8 (Tripos Inc., St. Louis, MO). Structures for thrombin fast and slow forms, aPC and factors VIIa, IXa and Xa were retrieved from the PDB. The structures of aPC and factors VIIa, IXa and Xa were also searched for Ca^{2+} ions whenever these were not reported in the coordinate files. The binding sites for Ca^{2+} were characterized with respect to residues possibly involved in the coordination sphere through the aid of molecular graphics. Whenever it was not explicitly stated in the coordinate files, the presence of sodium ions that would have been incorrectly assigned to water oxygen atoms was investigated with the program WASP [27]. The program WASP was chosen for that purpose because it pertains to a class of methods based on valence calculations, which proved to be very accurate [27], and because it is readily available.

2.2. Modeling of the hydrogen bonding network in the S1 subsite and sodium binding water channel

For the calculations, only bound water molecules, belonging to the S1 subsite or to the sodium binding water channel, were retained, other ligands, except for Na^+ , were deleted from the coordinate files and hydrogen atoms were added. In order to efficiently account for hydrogen bonding interactions in the region around the S1 subsite and the sodium binding water channel we have employed the ANNEAL command with help of the “minimize subset” dialog of the Sybyl v6.8 graphical interface. The Tripos force-field was used with a non-bonded cutoff of 8.0 Å and Kollman all-atoms charges were assigned. A gradient inferior to 0.05 kcal/mol Å was used as the termination criterion. The minimization protocol, where only hydrogen atoms were allowed to move, consisted of two phases: initially, a maximum of 20 simplex steps followed by a hundred steps of steepest descent was employed; then, 1000 steps of conjugated gradients were performed.

2.3. Computer-aided molecular surface analysis (CaMSA)

As first described [25], CaMSA methodology essentially combines calculation of Connolly channel surfaces [28–30] with arbitrary indexes of cavity lipophilicity based on lipophilic molecular surface potentials [31]. Briefly, structures were minimized as stated in the above section and had their

ARTICLE IN PRESS

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

283

ligands (if any) and water molecules deleted, followed by addition of all hydrogen atoms. Addition of all polar and non-polar hydrogen atoms is a requirement for the correct assignment of atomic lipophilic potentials as they were originally derived [31]. Then, using MOLCAD program in Sybyl v6.8 (Tripos Inc., St. Louis), fast Connolly channel surfaces were calculated and the lipophilic potential mapped over the cavity surface. Lipophilic potential ranges were adjusted to -0.20 to $+0.10$ and areas under low (-0.20 to -0.10), medium (-0.10 to 0.00) and high (0.00 to $+0.10$) lipophilicity were mapped.

3. Results

3.1. Na^+ binding states and coordination partners in factors VIIa and IXa

The structures of coagulation serine proteases used in this study were selected to reflect different binding states for the active and allosteric sites (Na^+ , Ca^{2+} and inhibitors). The 17 structures (2 thrombins, 8 FXas, 1 aPC, 2 FIXas and 4 FVIIas) are listed in Table 1. Characterization of Na^+ binding states required careful inspection of solvent molecules in order to determine if a sodium atom was not confounded with isoelectronic water oxygen during crystallographic refinement. In the case of factors VIIa and IXa, where several coordinate files were screened for the presence of Na^+ , we selected a number of structures differing in crystallization conditions, space group and cocrystallized inhibitor. This was performed in order to have a proper representation of the molecular surface, discarding possible artifacts (Table 1).

To the best of our knowledge, the Na^+ binding sites in coagulation factors VIIa and IXa have not been described yet. Therefore, we embraced the characterization of Na^+ binding partners in these proteases. The conformations of residues involved in Na^+ coordination in thrombin (pdb 1SG8) and FXa (typified by pdb 1MQY) were compared with the conformations of residues at the structurally equivalent positions in FVIIa (embodied by pdb 1CVW) and FIXa (exemplified by pdb 1RFN) structures (Fig. 2). Contrary to the available inhibitor-free FVIIa structure (pdb 1KLL), the benzamide inhibited FVIIa structure (pdb 1CVW) showed a water molecule (WAT605) near to the Na^+ position in thrombin (1.63 \AA) and FXa (1.21 \AA). The WASP program probably did not recognize this water molecule as a candidate for a sodium ion because of its unfilled valence. At least one of the water ligands coordinating the monovalent cation in thrombin (WAT549) or FXa (WAT729) was missing in the FVIIa structure analyzed (water site 2, WAT-2, in Fig. 2). The conformations of FVIIa residues in loop 1 are more similar to those displayed by FXa than to the residues in the equivalent thrombin segment. This observation supports the presence of a Na^+ binding site in FVIIa, which is similar to FXa, where two of the water molecules coordinating Na^+ in thrombin are replaced by two carbonyls from residues in loop 1. Hence, in FVIIa, the metal coordination partners probably are the backbone carbonyls of Y184 (Y185 in FXa 1MQY) and S185 (D185a in FXa 1MQY) in loop 1, T221 (R221a in thrombin 1SG8 and R222 in FXa 1MQY) and K224 in loop 2, and the oxygen atoms of solvent molecules positioned at water site 1 (WAT-1) and WAT-2 (Fig. 2).

As can be perceived from Fig. 2, FIXa presented a poor superposition in loops 1 and 2 either with FXa or thrombin fast form. The highest RMSD occurs in positions 185 (185a in

Table 1
Analysis of lipophilic potentials for the S1 subsite containing cavity in sodium binding coagulation enzymes

Enzyme	PDB code	Ligands present?			Area of S1 subsite containing cavity				
		Na^+ (water)	Ca^{2+}	Inhibitor	Total ^a (\AA^2)	Lipophilic potential ^b (%)			Total lipophilic ^c
						Low	Medium	High	
Thrombin	1SG8	Y (N.A. ^d)	–	N	919.2	11.0	77.9	11.1	89.0
	1SGI	N (N.A.)	–	N	792.7	13.5	73.7	12.8	86.5
FXa	IKIG	N (N.A.)	N	Y, AS ^e	1080.7	7.9	84.0	8.1	92.1
	IHCG	Y (583)	N	Y ^f	463.9	2.7	91.0	6.4	97.4
	1MQ5	Y (733)	Y	Y, AS	489.8	9.1	70.9	20.0	90.9
	INFY	Y (24)	Y	Y, AS	434.4	8.9	85.4	5.8	91.1
	IKSN	Y (11)	Y	Y, AS	518.9	6.4	79.6	14.0	93.6
	IG2L	Y (500)	Y	Y, AS	582.4	7.6	75.1	17.3	92.4
	IIQG	Y (577)	Y	Y, AS	375.7	11.9	82.9	5.2	88.1
	IXKA	Y (680)	Y	Y, AS	409.6	11.3	73.8	15.0	88.8
	1AUT	N (N.A.)	N	Y, AS	374.4	17.9	81.6	0.4	82.0
FIXa	IPFX	N (N.A.)	N	Y, AS	421.5	6.6	88.9	4.5	93.4
	IRFN	N (N.A.)	Y	Y, AS	330.0	20.7	79.1	0.3	79.4
FVIIa	1CVW	N (N.A.)	Y	Y, AS	290.0	34.7	65.3	0.0	65.3
	IDAN	N (N.A.)	Y	Y, AS ^g	300.8	32.6	67.4	0.0	67.4
	IDVA	N (N.A.)	Y	Y ^h	281.6	32.9	67.1	0.0	67.1
	IKLJ	N (N.A.)	Y	N	336.6	24.1	75.9	0.0	75.9

^aAreas calculated from the total cavity detected by the MOLCAD module in Sybyl v.6.8, Tripos Inc. ^bAreas under the following lipophilicity ranges — low: -0.20 to -0.10 ; medium: -0.10 to 0.00 and high: 0.00 to 0.10 . Percentages are expressed in relation to total cavity area. ^cCalculated from the sum of areas within medium and high lipophilic potential. ^dNot applicable. ^eactive-site bound. ^fAlthough this structure has no added inhibitor the COOH-terminal of the FXa A-chain interacts in a substrate-like manner with the active site of a neighboring FXa molecule. ^gBound to soluble tissue factor. ^hPeptide exosite inhibitor.

ARTICLE IN PRESS

284

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

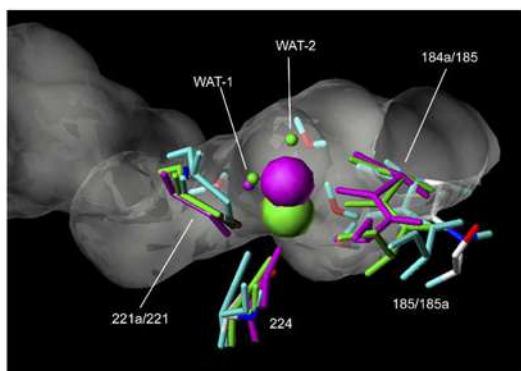


Fig. 2. Characterization of Na^+ binding sites in coagulation factors VIIa and IXa. The putative Na^+ binding partners in FVIIa (pdb 1CVW) and FIXa (pdb 1RFN) were compared with the known metal ligands in thrombin fast form and FXa (pdb 1MQY). Proteins are color-coded: thrombin — CPK, FXa — green, FVIIa — magenta and FIXa — cyan. Known (thrombin and FXa) and putative (FVIIa) sodium ions are shown as van der Waals spheres. Thrombin water molecules are shown with hydrogen atoms while waters pertaining to FXa or FVIIa are depicted only by their oxygen atoms. The thrombin channel surface is shown in gray and had its surface rendered transparent.

FXa) and 224. These particular conformations would be compatible either with flexibility in Na^+ binding site architecture, in similar fashion to aPC, or with an alternative sodium binding site in FIXa. This site would be a hybrid in which features of the Na^+ binding sites in thrombin and FIXa would be joined in a single framework. In particular, it would be possible that a water molecule bound to the carbonyl of H185 would occupy one of Na^+ coordination sites in FIXa, analogously to thrombin. The remaining ligands in the sodium coordination sphere would be represented by the carbonyl oxygen atoms of F184a, M221a and K224 and water molecules in WAT-1 and WAT-2.

3.2. CaMSA of the S1 subsite and the sodium binding water channel

As employed in the original CaMSA methodology [25], lipophilic potentials were mapped on the channel surfaces of the coagulation enzymes. It was shown that lipophilic potentials could be directly correlated to the binding affinities of trypsin-like proteases for S1 ligands. In this sense, we hypothesized that the conformational changes associated with the slow-fast transition and the resulting increase in catalytic rate could be related to an increase in surface lipophilic potential. Calculated areas for the different ranges of cavity lipophilic potentials for the sodium binding coagulation enzymes are listed in Table 1. There was a small increase in the lipophilic area of thrombin upon sodium binding. Similarly, there was a slight tendency of Ca^{2+} -free enzymes to present higher lipophilicity on their channel surfaces. These observation may reflect changes on the exposure of polar groups on the protein surface upon binding of Na^+ or Ca^{2+} .

Differences in solvent accessible surface (SAS) areas of the analyzed structures were also observed (Table 1) and were

related to the shape of the cavities depicted in Fig. 3. The cavities in FVIIa, FIXa and aPC are very similar and are essentially restricted to the S1 subsite and the adjoining hydrophilic channel linking the Na^+ binding site to the primary specificity pocket. Data in Table 1 also show a significantly higher SAS for most human FXa structures (pdb codes 1HCG, 1MQY, 1KSN and 1G2L) in relation to aPC and factors VIIa and IXa. This is a result of an extension of the cavity surface that reaches the regions corresponding to the S3 and S2 subsites.

Employing the same analysis to the catalytic domain of bovine FXa (pdb id 1KIG) a very complex surface is obtained. Differently from all human FXa structures, it involves the S1 subsite and the adjoining water channel as well as projections that reach almost all of the substrate binding cleft, from the non-primed (S2–S4) to the primed subsites (S1'–S3') near the Ca^{2+} binding site (not shown). The high sequence similarity (84.5% in 233 residues) between the orthologous bovine and human enzymes precludes the use of divergence in primary structure as an argument to explain these drastic differences in topographical features. The 3D structure of the bovine enzyme was reported without bound Ca^{2+} and Na^+ as well as any solvent molecules [32]. Perhaps, association of these observa-

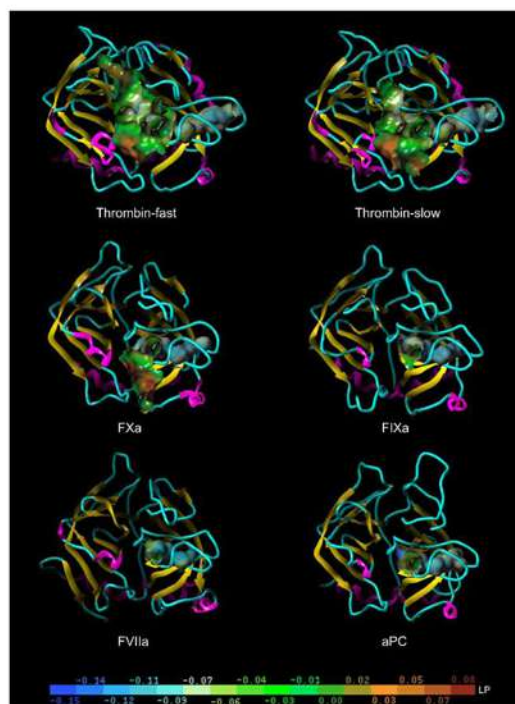


Fig. 3. General topography and lipophilicity of the major channels/cavities in human coagulation enzymes bearing a sodium binding site in the catalytic domain. On the bottom is shown the color-code for the lipophilic potential range: lipophilicity grows from bluish (hydrophilic) to brownish (lipophilic) shades. Protein secondary structure is color-coded and represented as ribbons.

ARTICLE IN PRESS

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

285

tions in terms of the presence or absence of such allosteric modulators could provide an explanation to the markedly distinct topographical features of the bovine FXa surface as a consequence of global conformational changes in the enzyme structure by effect of Ca^{2+} release.

A comparison of the structures surrounding the divalent metal binding sites in the FXa orthologues disclosed significant conformational changes possibly related to the release of the Ca^{2+} ion. First of all, it is important to note that G79 in the human enzyme is substituted for N79 in the bovine FXa. This gives a priori justification for the changes in the backbone conformation of the calcium binding loop, particularly in the segment 75–78. Nevertheless, there is a drastic reversal in the direction of the E80 carboxylate group, which points toward the solvent in the Ca^{2+} -free bovine enzyme. Whether the observed displacement in the loop engaged in Ca^{2+} binding (residues 70–80) would cause the abrogation of the metal binding site in bovine FXa is not certain but the reorientation in E80 side chain would surely be related to the absence of a bound Ca^{2+} ion. In fact, the human FXa structure reported without Ca^{2+} in the catalytic domain (pdb id 1HCG) the E80 side chain is in a conformation intermediary between the solvent exposed, as observed in the bovine enzyme, and the one involved in calcium coordination, as in all other Ca^{2+} -bound FXa catalytic domain structures.

It is clear from the data in Table 1 and inspection of Fig. 3 that both thrombin allosteric forms have a larger cavity SAS than the other coagulation enzymes (with exception of bovine FXa). Distinctly from aPC and human factors VIIa, IXa and Xa, the cavity surfaces for thrombin allosteric forms entail not only the S1 subsite and the water channel but also a great part of the substrate binding cleft, from S3 to S3'. But, more importantly, the free thrombin fast form presents a continuous surface connecting the sodium binding water channel and the S1 subsite to the region corresponding to the prime-side subsites (S1'–S4) and exosite-I, a feature not observed in the slow form (Fig. 4). Since the connection with the exosite-I is lost in the slow form, a marked shortening of the cavity SAS is observed. This fact is

accompanied by a pronounced narrowing of the entrance to the S1 subsite in the slow form.

Another feature of the surfaces compared in Fig. 3 is that, except for thrombin, the water channels in all other sodium binding coagulation proteases do not open to the enzyme surface. It is clear from the alignment of Na^+ -regulated coagulation enzymes in Fig. 5 that the unique three-residue insertion in the thrombin loop 184–189 (loop 1) leads to the aperture in the channel extremity opposite to the S1 subsite. Loop 1 and loop 2 support most of the architecture of the S1 subsite and the adjoining water channel. We have made a comparison of the backbone conformations for loop 1 in thrombin with the other coagulation proteases bearing a Na^+ binding site (Fig. 6). It clearly shows that this loop in the later enzymes is dislocated toward the channel lumen imposing a constriction that ends up by occluding the communication with the external part of the enzyme surface (e.g., FXa in Fig. 6A and C). In general, the maximal displacement occurs in position 222, e.g., 1.6 Å in the C α positions in aPC and thrombin compared to 1.14 Å RMSD in the whole loop (Fig. 6B).

3.3. Mapping of residues involved in topographical changes in thrombin surface upon Na^+ -mediated allosteric transition

Fig. 7 shows in detail some important residues responsible for the channel architecture and the underlying water network for thrombin slow and fast forms. As reported above, sodium release results in the narrowing of the S1 subsite entrance in thrombin. Inspection of the structures for the thrombin fast and slow forms revealed that a conformational change in E192 side chain is responsible for this effect (compare panels A and B in Fig. 7). The rearrangement in the water network associated with the movement in E192 side chain is also evident. Comparison of the panels in Fig. 7 also reveals that the rupture in the double hydrogen-bonding interaction between R187 and D222 upon transition of the fast to the slow form has virtually no effect over the channel shape near its aperture opposite to the S1 subsite.

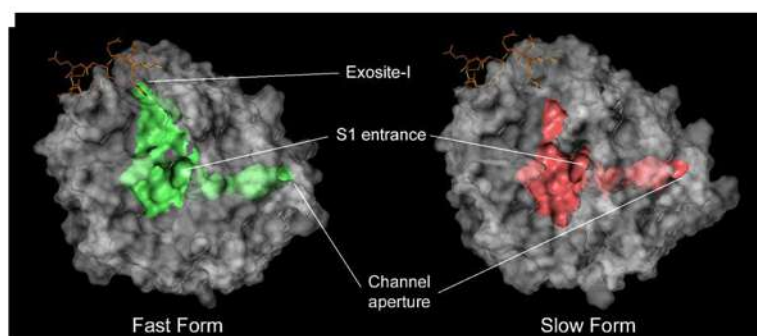


Fig. 4. Coverage of the cavity bearing the Na^+ binding channel in thrombin fast and slow forms. Solvent accessible surfaces (1.4 Å probe) for the whole molecules were made transparent to turn evident the range of the cavity bearing the water channel and specificity pockets. An exosite-I ligand, *N*-acetylthirudin[53–64]-Tyr63-sulfate (orange sticks), was transposed from the PDB structure 1HAH in order to locate this site on the thrombin surface.

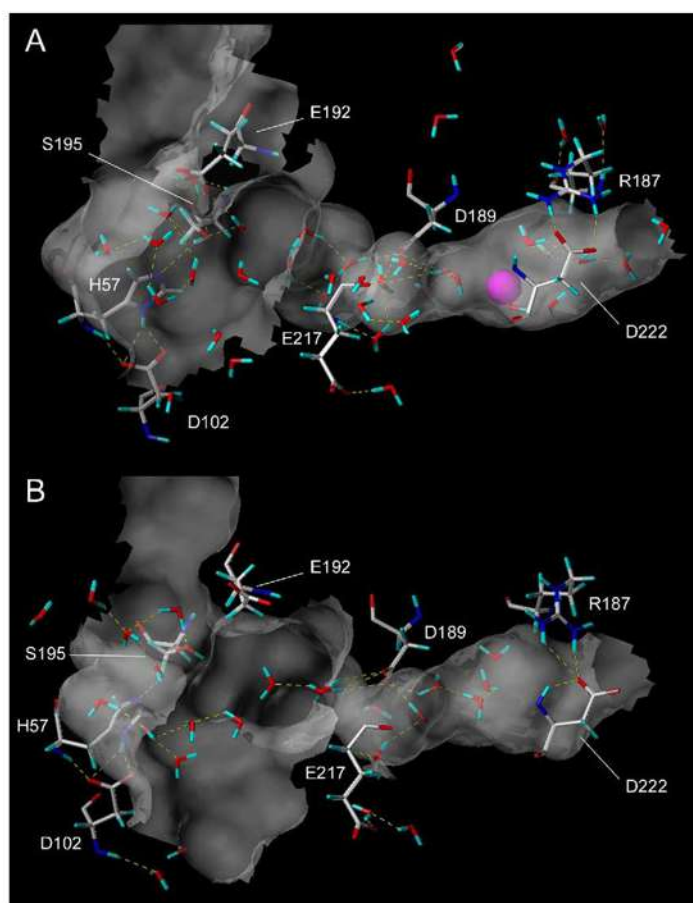


Fig. 7. Comparison of the channel shape in the fast and slow forms of thrombin. Some important residues responsible for the channel structure are identified in the figure along with bound water molecules pertaining to the hydrogen-bonded network surrounding the S1 subsite and sodium binding site. Note E192 orientation in the ligand-free slow and fast forms of thrombin resulting in the narrowing of S1 subsite entrance. The breakage of R187:D222 salt bridge in the slow form produces no observable difference in the subjacent region of the channel surface, as compared to the fast form. Also, observe the different pattern of hydrogen bonding interactions (dashed yellow lines) involving the water molecules associated with the slow–fast transition.

segment (not shown). Consequently, in the Na^+ -free aPC, one of the former Na^+ ligands in the structure reported by Schmidt et al. has its carbonyl oriented away from the Na^+ ion, preventing its role as a sodium coordination partner [22]. This suggests a high level of flexibility in loop 1 that seems to be unique to aPC.

Indeed, comparison of the loop 1 conformation in all coagulation proteases thought to bind Na^+ showed that its flexibility is considerably higher than in the other polypeptide segment (loop 2) contributing to the metal coordination (Fig. 2). Independently of the way that loop 1 contributes to Na^+ coordination, either through anchored water molecules (as in thrombin) or directly (as in FXa), it is the backbone conformation of loop 2 that defines the pivotal metal coordination site. What the Dang and Di Cera's rule does not cover is how the Na^+ site architecture may influence the Na^+ dissociation constants (K_d) and monovalent cation specificity

of human coagulation serine proteases. Moreover a different Na^+ site architecture may impact on the degree of activation for each enzyme in the presence or not of other modulators, such as substrate or inhibitor in S1 subsite, Ca^{2+} or an exosite-I ligand. In view of the variability in primary and tertiary structure of loop 1 in sodium activated coagulation proteases, it appears that this loop would furnish the fine tuning necessary to explain the different K_d for Na^+ reported in the literature for these enzymes. In fact, it has been shown that mutations in loop 1 are able to redesign the ionic specificity of thrombin [33,34].

4.2. CaMSA discloses novel features of the molecular recognition and allosteric responses in thrombin and other sodium-activated coagulation enzymes

Our hypothesis that new aspects of the molecular recognition of Na^+ in the water channel and changes

ARTICLE IN PRESS

288

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

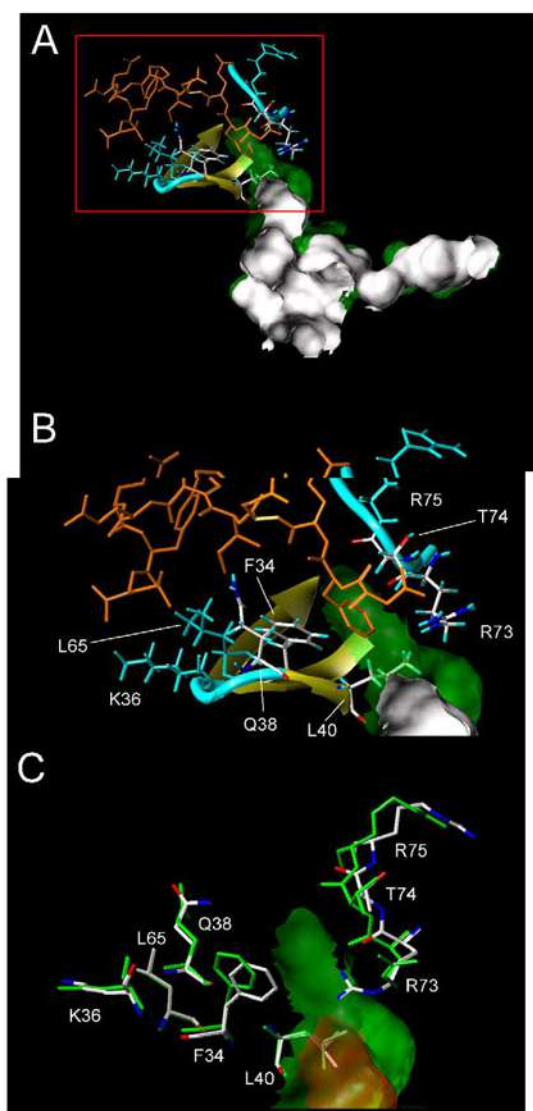


Fig. 8. Analysis of topographical changes responsible for the loss of communication between exosite-I and the active site in thrombin transition to the slow form. A-Mapping of residues in “hot region” next to exosite-I possibly involved in topographical changes upon slow–fast transition in thrombin. Cavity surface for the slow form is shown in solid white over the fast form cavity in transparent green. B-Detail of the red box in A. Residues in the hot region of thrombin fast form making contacts with the modeled exosite-I ligand are shown in sticks. Colored in cyan are general residues, simply making contacts with the exosite-I ligand and in CPK color-scheme are the residues directly involved in disrupting the exosite-I communication upon thrombin transition to the slow form. C-Conformational changes between residues in the fast (green sticks) and slow (CPK-colored sticks) thrombin forms. Hydrogen atoms were omitted for the sake of clarity.

associated with allosteric enhancement of catalytic rate could be revealed by CaMSA of coagulation serine proteases was supported by the results described in Section 3. Indeed, from

the cavities typified in Fig. 3, one major difference can be clearly noted which distinguishes thrombin from the other coagulation serine proteases that are able to bind a sodium ion. Whereas the sodium binding water channel in both slow and fast forms of thrombin opens up on the surface of the enzyme opposite to the S1 subsite, the same channel in FVIIa, FIXa, FXa and aPC closes just before reaching the Na^+ binding site. In the thrombin segment comprising residues 185–187 (K185, P186, D186a, E186b, G186c, K186d and R187) the backbone considerably deviates (maximum distance=3.5 Å) from the path followed by the homologous segment in other coagulation proteases. In this 7-residue segment, aPC lacks three of the thrombin insertions and conserves only a half of the remaining residues (D186a and R187) and all other sodium binding coagulation enzymes behave similarly (Fig. 5). A clear consequence is that the sodium binding site in thrombin is directly accessible to the solvent from both sides of the channel, while for the other enzymes it is not. Hence, the mechanism by which the sodium binding event in FVIIa, FIXa, FXa and aPC can be translated into a catalytic rate enhancement may be different from that operating in thrombin.

Inspection of the structures beneath the channels surfaces revealed that they collapse in the segment equivalent to the region delineated by the R187:D222 salt bridge, which has been postulated as a hallmark of the slow–fast transition in thrombin [20]. Based on the results of site-directed mutagenesis coupled to thermodynamic data this interaction has been claimed to be essential for maintaining the integrity of the channel near the Na^+ binding site. However, as can be perceived from comparison of the panels in Fig. 7, despite of the fast form presenting the most stabilizing salt-bridge, the topography on the sodium binding water channel for the slow as well as the fast forms of thrombin shows no significant alteration. Interestingly, the residues involved in this electrostatic interaction are weakly conserved in other sodium binding coagulation enzymes (Fig. 5). In aPC, D222 is substituted by leucine, which makes hydrophobic contacts with the aliphatic part of the R187 side chain. This interaction could partially substitute the stabilizing effect that the salt bridge has over the water channel in thrombin. R187 in aPC is dislocated 2.7 Å in relation to thrombin (from the $\text{C}\alpha$ positions) pointing out of the channel due to the steric blockage by the isobutyl moiety of the L222 side chain (Fig. 6B). In every FXa structure, the aliphatic part of the side chain of a lysine residue (K223) occupies position equivalent to the isobutyl moiety of L222 in aPC. In order to accommodate the bulkier K222 side chain, FXa seemed to have acquired a compensatory mutation in position 187, the R187Q (thrombin→FXa) substitution. The Q187 side chain being two atoms shorter than R187 frees space for K222 (Fig. 6C). A similar behavior is observed in FIXa and FVIIa structures, where the R187:D222 pair of thrombin is substituted for G187:K222 and G187:V222, respectively. The above analysis raises a question about the requirement of a strong stabilizing interaction between residues in positions 187 and 222 to stabilize the channel structure in coagulation enzymes other than thrombin.

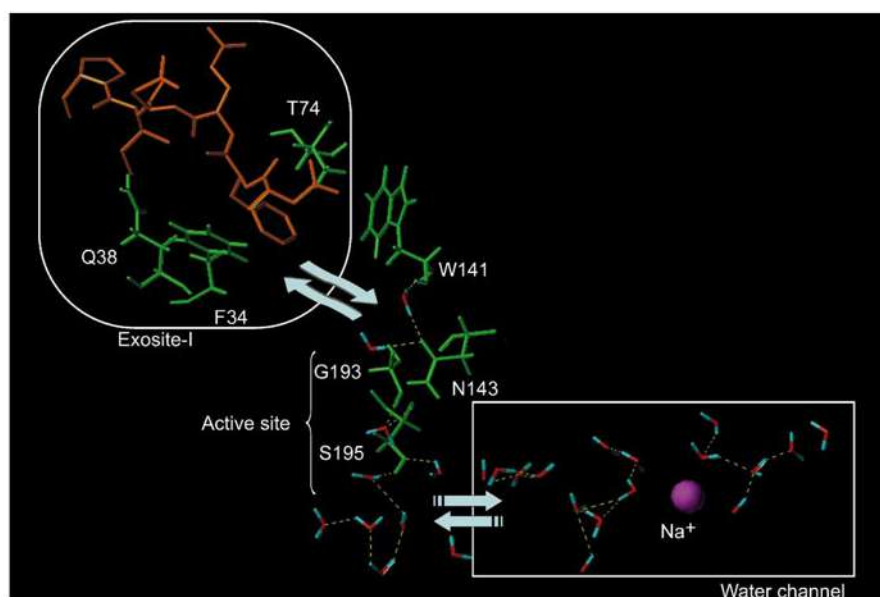


Fig. 9. Proposed structural mechanism for linking conformational changes in the exosite-I to the S1 subsite and adjoining water channel in thrombin. See text for thorough description. For clarity, some of the residues involved in binding the hirugen surrogate, *N*-acetylhirudin[53–64]-Tyr63-sulfate (orange sticks), in exosite-I were omitted (K36, L65, R73 and R75).

Residue D221 belongs to the periphery of the allosteric core in thrombin and fixes one of the water molecules involved in Na^+ coordination in the fast form [20]. It is also a key residue in the allosteric transduction, i.e., in transmitting changes in the sodium binding site to the S1 subsite. In factors VIIa, IXa and Xa, D221 is substituted for Ala whereas in aPC it is mutated to a Gly residue (Fig. 5). In these enzymes, the Asp carboxylate necessary to bind one of the water molecules that coordinate de sodium ion is missing. This indicates that either the water network found in thrombin channel must organize differently in the other coagulation enzymes in order to bind the Na^+ ion or a distinct set of binding partners must be involved in the Na^+ coordination. In support of the later hypothesis, in FXa and aPC, two out of four water molecules coordinating Na^+ in thrombin (waters bound to D221 and Y184a) are replaced by carbonyl groups from residues in loop 1. Noteworthy, the D221X substitutions (thrombin \rightarrow FVIIa, FIXa, FXa and aPC) are most probably compensatory mutations since they open up space for accommodating the alternative conformation of bulky side chains in position 187.

Other interactions, specific to each enzyme, are found to help in giving structural stabilization to the water channel in the absence of the salt bridge R187:D222 that occurs in thrombin. In aPC, strong hydrogen bonds are found between the guanidinium group of R187 and the G221 carbonyl (2.4 Å R187^{N-H} to ^{O=C}G221) and between the backbones of G186a and L222. In FXa, additional channel stabilization is provided by a hydrogen bond involving the Y225 and H185 side chains.

4.3. Role of position 192 in controlling the access to the S1 subsite entrance

It has been shown that the E192A mutation affects poorly Na^+ affinity and clearly has no impact over thrombin interaction with the synthetic substrate FPR [20]. Nevertheless, a significant rearrangement in E192 side chain has been detected in the original report of thrombin slow and fast forms structures [20]. At the time, it was proposed that E192 would have a role in connecting changes in the Na^+ binding site to the active site through the water network. Indeed, in both thrombin forms, E192 interacts with G142, N143, G193 and C191, but in the fast form E192 is orientated in such manner that it is able to communicate with S195 through a water-mediated hydrogen bonding network (Fig. 7). It is noteworthy that a similar E192 reorientation has been observed in several conditions denoting an inherent flexibility for the side chain of this residue: (i) when sodium binding is disrupted by the D221A/D222K mutations present in the thrombin variant ARK [35]; (ii) upon collapse of the S1 subsite pocket coupled to abrogation of sodium binding and allosteric transduction in the W215A/E217A thrombin mutant [36] and (iii) on PPACK binding to the fast form [20].

Still obscure is why Ala mutation of a residue suffering a major conformational change detected from the comparison of structures standing for thrombin slow and fast forms had no impact on Na^+ binding or allosteric transduction. A possible explanation for this apparent contradiction has been claimed in terms of a strict role for residue E192 in the recognition of larger substrates by thrombin such as Protein C [20]. They

ARTICLE IN PRESS

290

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

proposed that the rearrangement in E192 side chain upon sodium release would facilitate binding of the anticoagulant (slow) thrombin form to Protein C by minimizing the electrostatic repulsion with the acidic P3 and P3 residues. This hypothesis was recently refuted by results from site-directed thermodynamic analysis showing that E192 does not pertain to the Protein C binding epitope neither in the free-thrombin nor in the thrombomodulin complex [37].

At any rate, the cavity shapes obtained in our analysis clearly depict the influence of the residue in position 192 in the size of the S1 subsite opening in the different coagulation proteases sensitive to sodium binding (Fig. 3). From the alignment in Fig. 5, it was verified that while aPC retained the glutamate in 192, FIXa and FXa had it mutated to Q192. In addition to the loss of the E192 negative charge, Q192 side chain conformation is significantly altered by a modification in χ_1 that results in the relocation of its polar side chain away from the S1 subsite entrance, turning this pocket wider in FXa than in thrombin or aPC. A more drastic mutation in relation to thrombin occurred in FVIIa where E192 was mutated to K192.

4.4. A possible mechanism for the thermodynamic linkage between the exosite-I, S1 subsite and the Na⁺ binding site on thrombin

One major aim of our analysis was to extract more data from the available sodium-free and sodium-bound thrombin structures through CaMSA. As mentioned before, a remarkable result from our analysis was that the connection between exosite-I and S1 subsite is lost in thrombin slow form due to topographical changes over the region encompassing the prime-side subsites and exosite-I periphery (Fig. 4). The kinetic mechanism supporting the enhanced catalytic rate of the fast form involves greater diffusion rates of the substrate to the binding cleft in the enzyme and higher acylation rates (reviewed in Ref. [4]). Both features are compatible, respectively, with the greater SAS as provided by the additional surface extending to the exosite-I and with the slight increase in the channel lipophilicity in the sodium bound enzyme (Table 1). The later observation is in line with our previous conclusions about lipophilicity governing binding affinities in the S1 subsite in trypsin-like enzymes [25]. Moreover, changes in lipophilicity may be associated to tiny conformational changes in residues exposed to the surface resulting in lesser exposition of polar residues to the solvent. Such distinct properties of the fast and slow thrombin structures have not been described before.

By carefully inspecting the molecular surfaces and their underlying residues, it was possible to identify a “hot region” as being responsible for the extension of the channel surface to the exosite-I in the Na⁺ bound form of thrombin. F34, Q38 and L40 in the region corresponding to the two central β -strands comprising residues 29–36 and 38–47 from the N-terminal β -barrel, and residues R73 and T74 from a loop and a short β -strand that delineate part of the exosite-I constitute this hot region (Fig. 8B). As can be perceived from the superposition of

these residues in the ligand-free slow and fast forms depicted in Fig. 8C, upon sodium binding, a significant conformational change in R73 side chain occurs. This promotes the formation of two hydrogen bonding interactions with segment 151–153 in the C-terminal domain of the fast form that culminates in a displacement (RMSD) of 1.92 Å in the position of Q151 side chain.

The hot region responsible for disrupting the surface linkage between S1 subsite and exosite-I in thrombin slow form is located around 25 Å away from the sodium site. Therefore, one should expect to find an intricate set of interactions in thrombin responsible for the long range communication between exosite-I and the sodium binding site. W141, G193 and residues in the segment 151–152 appeared as interesting candidates for playing this role since they are in contact distance (4.0 Å) from residues in the hot region. Residues Q151, P152 and S153 are located at the C-terminal tip of the autolysis loop (141–152), while W141 is situated at the opposite tip. Although the central part of the autolysis loop is usually disordered, these residues act as anchors and present low to average temperature factors. On the other hand, G193 constitutes the oxyanion hole and therefore furnishes a direct communication with the catalytic apparatus, which in turn is linked to the sodium binding water channel through the S1 subsite. Residue W141, in turn, has been implicated in the past in the fluorescence increases that are detected on exosite-I binding by protease-activated receptor (PAR) peptides [38]. Recent experimental data give support to the role of W141 in connecting events in the active site to changes in exosite-I [39]. A series of amide H²/D studies showed that W141 loses SAS on PPACK binding, which suggests that W141, even though not contacting directly PPACK, may transmit alterations in the active site (and indirectly in the sodium binding site) to the exosite-I, mainly by contacting R73.

Our results combined with literature data permitted us to propose the mechanism delineated in Fig. 9. The Na⁺ binding event is thought to be transmitted to the active site S195 through a rearrangement in the water network in the hydrophilic channel adjacent to the S1 subsite [20]. N143 possibly senses such changes in the active site through the oxyanion hole (G193) and passes the information up to W141, in the exosite-I periphery, through water mediated hydrogen bonds. Conformational changes in W141 related to differential solvent exposure are then transmitted to residues in the hot region (F34, Q38, R73 and T74), which in turn, move their side chains toward the cavity and abrogate the continuity on the surface connecting the active site to the exosite-I. Finally, repositioning of residues in the hot region, including K36, L65, T74 and R75 are translated into different binding affinities displayed by the slow form for specific ligands in exosite-I, as hirugen and thrombomodulin.

Considering the above exposition, one further question arises: if residues in the hot region in addition to E192 are important for the topographical changes observed during slow fast transition, so why these residues do not present relevant thermodynamic signatures in thrombin allostery? The case of E192 has been discussed before, but the issue regarding F34,

ARTICLE IN PRESS

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

291

Q38, R73 and T74 remains obscure. Nevertheless, one could hypothesize that once these residues are located in the exosite-I, in order to detect a clear effect upon Ala mutation, activity assays should be performed with ligands also interacting with this site and not just with the active site like the FPR substrate used before [20]. Indeed, recent site-specific thermodynamic experiments have confirmed this hypothesis by showing that T74 among other thrombin residues (K36, L65 and R75) are responsible for hirudin binding with higher affinity for the fast form [40]. Other residues, such as L40, the 151–153 segment and W141 were not mutated previously [20] and therefore still wait for further experimental data in order to confirm their roles in thrombin allostery. At any rate, a multiple sequence alignment (not shown) reveals a high degree of conservation of residues in the hot region among several thrombin orthologues but not within other chymotrypsin-like serine proteases, giving additional support for their role in the function of thrombin. For instance, R73 is completely conserved from human to hagfish. On the other hand, F34 is strictly conserved on mammalian thrombins, being substituted by a tyrosine in orthologues from other vertebrates like chicken and fishes. Similarly T74 is conserved among mammalian species and zebrafish but is mutated to Ala in other vertebrates.

5. Concluding remarks

We have characterized the structure, topography and lipophilicity of the Na⁺ channel in the sole available structure of aPC and in all FVIIa, FIXa and FXa structures available in the PDB. Through application of CaMSA we have unveiled that the pore opening distal to the S1 subsite giving access to the Na⁺ binding site in thrombin is unique to this enzyme among other coagulation proteases activated by sodium ion. The shorter loop 1 and the lack of a R187:D222 salt bridge in aPC and factors VIIa, IXa and Xa could be invoked to rationalize the different channel shapes. Such remarkable distinction may have a bearing in the mechanism evolved to acquire affinity for monovalent cations and specificity for Na⁺ in coagulation enzymes.

A main goal of our analysis consisted in the comparison of the crystallographic structures representing the thrombin slow and fast forms in the search for obvious differences in the S1 pocket and on the shape and lipophilicity of the water channel that shelters the sodium binding site. From the analysis of molecular surfaces we disclosed previously unobserved conformational changes possibly associated with the slow–fast transition triggered by Na⁺ binding in thrombin. These conformational changes are mapped to residues R73, T74, R75 and F34 located in the exosite-I region.

Our general observations and the proposed structural mechanism for linking sites thermodynamically coupled in thrombin surface were supported by both published and unpublished experimental results we were unaware of by the time the analysis was performed [40]. Nevertheless, future work with molecular dynamics simulations of the slow–fast transition in thrombin as well as other sodium-activated coagulation proteases would provide a reasonable test for our conclusions

since crystallographic analyses have proved to be insufficient to map the tiny conformational changes involved in this process.

6. Note added in proof

During preparation of the manuscript a paper by Schmidt et al. on the Na⁺ site of FIXa was published [41]. Distinctly from us, these authors propose a Na⁺ binding site strictly equivalent to that on FXa, aPC and FVIIa. They also effectively show that the FIXa Na⁺ site is linked strongly to the S1 and FVIIa binding sites but only weakly to the Ca²⁺ site.

Acknowledgments

This work was supported by CNPq, FAPERJ and PAPES-FIOCRUZ. FPS Jr was a CNPq Dsc fellowship recipient. R.B.A. acknowledges a FAPERJ grant (E26-152–174/2002). We thank Dr. E. Di Cera (Washington University School of Medicine, St. Louis) for sharing unpublished results and fruitful discussions.

References

- [1] M. Hoffman, Remodeling the blood coagulation cascade, *J. Thromb. Thrombolysis* 16 (2003) 17–20.
- [2] E.W. Davie, K. Fujikawa, W. Kisiel, The coagulation cascade: initiation, maintenance, and regulation, *Biochemistry* 30 (1991) 10363–10370.
- [3] E.W. Davie, A brief historical review of the waterfall/cascade of blood coagulation, *J. Biol. Chem.* 278 (2003) 50819–50832.
- [4] E. Di Cera, Thrombin: a paradigm for enzymes allosterically activated by monovalent cations, *C.R. Biol.* 327 (2004) 1065–1076.
- [5] E. Di Cera, E.R. Guinto, A. Vindigni, Q.D. Dang, Y.M. Ayala, M. Wuyi, A. Tulinsky, The Na⁺ binding site of thrombin, *J. Biol. Chem.* 270 (1995) 22089–22092.
- [6] C. Orthner, D.P. Kosow, The effect of metal ions on the amidolytic activity of human factor Xa (activated Stuart–Prower factor), *Arch. Biochem. Biophys.* 185 (1978) 400–406.
- [7] C. Orthner, D.P. Kosow, Evidence that human alpha-thrombin is a monovalent cation-activated enzyme, *Arch. Biochem. Biophys.* 202 (1980) 63–75.
- [8] S.A. Steiner, F.J. Castellino, Kinetic studies of the role of monovalent cations in the amidolytic activity of activated bovine plasma protein C, *Biochemistry* 21 (1982) 4609–4614.
- [9] Q.D. Dang, E. Di Cera, Residue 225 determines the Na(+)-induced allosteric regulation of catalytic activity in serine proteases, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 10655–10656.
- [10] E. Zhang, A. Tulinsky, The molecular environment of the Na⁺ binding site of thrombin, *Biophys. Chem.* 63 (1997) 185–200.
- [11] Q.D. Dang, A. Vindigni, E. Di Cera, An allosteric switch controls the procoagulant and anticoagulant activities of thrombin, *Proc. Natl. Acad. Sci. U. S. A.* 92 (1995) 5977–5981.
- [12] T. Rose, E. Di Cera, Three-dimensional modeling of thrombin–fibrinogen interaction, *J. Biol. Chem.* 277 (2002) 18875–18880.
- [13] A.O. Pineda, A.M. Cantwell, L.A. Bush, T. Rose, E. Di Cera, The thrombin epitope recognizing thrombomodulin is a highly cooperative hot spot in exosite I, *J. Biol. Chem.* 277 (2002) 32015–32019.
- [14] W. Bode, P. Schwager, The single calcium-binding site of crystalline bovin beta-trypsin, *FEBS Lett.* 56 (1975) 139–143.
- [15] A.R. Rezaie, T. Mather, F. Sussman, C.T. Esmon, Mutation of Glu-80→Lys results in a protein C mutant that no longer requires Ca²⁺ for rapid activation by the thrombin–thrombomodulin complex, *J. Biol. Chem.* 269 (1994) 3151–3154.
- [16] W.G. Owen, C.T. Esmon, Functional properties of an endothelial cell cofactor for thrombin-catalyzed activation of protein C, *J. Biol. Chem.* 256 (1981) 5532–5535.

ARTICLE IN PRESS

292

F.P. Silva et al. / Biophysical Chemistry 119 (2005) 280–292

- [17] P. Wildgoose, D. Foster, J. Schiodt, F.C. Wiberg, J.J. Birktoft, L.C. Petersen, Identification of a calcium binding site in the protease domain of human blood coagulation factor VII: evidence for its role in factor VII-tissue factor interaction, *Biochemistry*, 32(1993)114–119.
- [18] A.R. Rezaie, C.T. Esmon, Asp-70→Lys mutant of factor X lacks high affinity Ca^{2+} binding site yet retains function, *J. Biol. Chem.* 269 (1994) 21495–21499.
- [19] A.O. Pineda, C.J. Carrell, L.A. Bush, S. Prasad, S. Caccia, Z.W. Chen, F.S. Mathews, E. Di Cera, Molecular dissection of Na^+ binding to thrombin, *J. Biol. Chem.* 279 (2004) 31842–31853.
- [20] A.O. Pineda, C.J. Carrell, L.A. Bush, S. Prasad, S. Caccia, Z.W. Chen, F.S. Mathews, E. Di Cera, Molecular dissection of Na^+ binding to thrombin, *J. Biol. Chem.* 279 (2004) 31842–31853.
- [21] M.T. Lai, E. Di Cera, J.A. Shafer, Kinetic pathway for the slow to fast transition of thrombin. Evidence of linked ligand binding at structurally distinct domains, *J. Biol. Chem.* 272 (1997) 30275–30282.
- [22] A.E. Schmidt, K. Padmanabhan, M.C. Underwood, W. Bode, T. Mather, S.P. Bajaj, Thermodynamic linkage between the S1 site, the Na^+ site, and the Ca^{2+} site in the protease domain of human activated protein C (APC). Sodium ion in the APC crystal structure is coordinated to four carbonyl groups from two separate loops, *J. Biol. Chem.* 277 (2002) 28987–28995.
- [23] M.C. Underwood, D. Zhong, A. Mathur, T. Heyduk, S.P. Bajaj, Thermodynamic linkage between the S1 site, the Na^+ site, and the Ca^{2+} site in the protease domain of human coagulation factor Xa. Studies on catalytic efficiency and inhibitor binding, *J. Biol. Chem.* 275 (2000) 36876–36884.
- [24] A.E. Schmidt, J.E. Stewart, S.P. Bajaj, Functional linkage between the Na^+ site, substrate site, and the Ca^{2+} site in the protease domain of human factor IXa (FIXa), *Blood* 102 (2003) 1074.
- [25] F.P. Silva Jr., S.G. De Simone, S1 subsite in snake venom thrombin-like enzymes: can S1 subsite lipophilicity be used to sort binding affinities of trypsin-like enzymes to small-molecule inhibitors? *Bioorg. Med. Chem.* 12 (2004) 2571–2587.
- [26] H.-J. Böhm, G. Schneider, R. Mannhold, et al., (Eds.), *Protein–Ligand Interactions: From Molecular Recognition to Drug Design*, Wiley–VCH GmbH and Co. KGaA, Weinheim, 2003.
- [27] M. Naylor, E. Di Cera, Valence screening of water in protein crystals reveals potential Na^+ binding sites, *J. Mol. Biol.* 256 (1996) 228–234.
- [28] M.L. Connolly, Solvent-accessible surfaces of proteins and nucleic acids, *Science* 221 (1983) 709–713.
- [29] J.S. Delaney, Finding and filling protein cavities using cellular logic operations, *J. Mol. Graph.* 10 (1992) 174–177.
- [30] T.E. Exner, M. Keil, G. Moeckel, J. Brickmann, Identification of substrate channels and protein cavities, *J. Mol. Biol.* 304 (1998) 340–343.
- [31] A. Ghose, G. Crippen, Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure–activity relationships I. Partition coefficients as a measure of hydrophobicity, *J. Comput. Chem.* 7 (1986) 565–577.
- [32] A. Wei, R.S. Alexander, J. Duke, H. Ross, S.A. Rosenfeld, C.H. Chang, Unexpected binding mode of tick anticoagulant peptide complexed to bovine factor Xa, *J. Mol. Biol.* 283 (1998) 147–154.
- [33] S. Prasad, K.J. Wright, D.B., L.A. Bush, A.M. Cantwell, E. Di Cera, Redesigning the monovalent cation specificity of an enzyme, *Proc. Natl. Acad. Sci. U. S. A.* 100 (2003) 13785–13790.
- [34] S. Prasad, A.M. Cantwell, L.A. Bush, P. Shih, H. Xu, E. Di Cera, Residue Asp-189 controls both substrate binding and the monovalent cation specificity of thrombin, *J. Biol. Chem.* 279 (2004) 10103–10108.
- [35] A.O. Pineda, E. Zhang, E.R. Guinto, et al., Crystal structure of the Thrombin mutant D221A/D222K: the Asp222:Arg187 ion-pair stabilizes the fast form, *Biophys. Chemist.* 112 (2004) 253–256.
- [36] A.O. Pineda, Z.-W. Chen, S. Caccia, A.M. Cantwell, S.N. Savvides, G. Waksman, F.S. Mathews, E. Di Cera, The anticoagulant thrombin mutant W215A/E217A has a collapsed primary specificity pocket, *J. Biol. Chem.* 279 (2004) 39824–39828.
- [37] H. Xu, L.A. Bush, A.O. Pineda, S. Caccia, E. Di Cera, Thrombomodulin changes the molecular surface of interaction and the rate of complex formation between thrombin and protein C, *J. Biol. Chem.* 280 (2005) 7956–7961.
- [38] M.M. Krem, E. Di Cera, Dissecting substrate recognition by thrombin using the inactive mutant S195A, *Biophys. Chemist.* 100 (2003) 315–323.
- [39] C.H. Croy, J.R. Koeppel, S. Bergqvist, E.A. Comives, Allosteric changes in solvent accessibility observed in thrombin upon active site occupation, *Biochemistry* 43 (2004) 5246–5255.
- [40] K.E. Mengwasser, L.A. Bush, P. Shih, A.M. Cantwell, E. Di Cera, Hirudin binding reveals key determinants of thrombin allostery, *J. Biol. Chem.* 280 (2005) 26997–27003.
- [41] A.E. Schmidt, J.E. Stewart, A. Mathur, S. Krishnaswamy, S.P. Bajaj, Na^+ site in blood coagulation factor IXa: effect on catalysis and factor VIIIa binding, *J. Mol. Biol.* 350 (2005) 78–91.

ANEXO 5

FEBS Journal

Página 1 de 1

Successful Submission Confirmation 29 Sep
05 19:01

Your manuscript has been successfully uploaded to FEBS Journal. You will receive future communications via e-mail.

Your manuscript number is: **FJ-05-0995**

Please make note of your manuscript number. You will receive an e-mail from FJ within 24 hours of the end of this process, confirming receipt of your submission.

[Print This Page](#)

[Return to the Author Centre](#)

[Log Out](#)

Manuscript Central™ v1.8 (patent pending). Copyright © ScholarOne, Inc., 2005. All Rights Reserved.
Manuscript Central is a trademark of ScholarOne, Inc. [Terms and Conditions of Use](#)
[ScholarOne Privacy Policy](#)

BJ-48, a highly glycosylated clotting serine protease from the *Bothrops jararacussu* venom: BIOCHEMICAL CHARACTERIZATION AND SOLUTION STRUCTURE ANALYSIS

Floriano P. Silva-Junior^{‡,¶}, Saulo C. Bourguignon[§], Anibal R. Melgarejo⁺, Cristiano L. P. Oliveira[&], Iris L. Torriani[&] and Salvatore Giovanni-De-Simone^{§,*}

[‡]Laboratório de Bioquímica de Proteínas e Peptídeos, Departamento de Bioquímica e Biologia Molecular, Instituto Oswaldo Cruz, FIOCRUZ, Av Brasil 4365, 21045-900, RJ, Brazil; [¶]Programa de Pós-graduação em Química Orgânica, Instituto de Química, Universidade Federal do Rio de Janeiro, RJ, Brazil; ⁺Instituto Vital Brazil, Niterói, RJ, Brazil; [&]Laboratório Nacional de Luz Síncrotron, Campinas, SP, Brazil; [§]Departamento de Biologia Celular e Molecular, Instituto de Biologia, Universidade Federal Fluminense, Niterói, RJ, Brazil.

Keywords: Thrombin-like, snake venom, serine protease, SAXS, N-linked glycosylation.

*To whom correspondence should be addressed: FAX: +55-21-5903495; E-mail: dsimone@ioc.fiocruz.br.

Abbreviations: N α -Tosyl-L-Arg-methyl ester (TAME); α -N-benzoyl-DL-Arg-methyl ester (BAME), α -N-benzoyl-DL-arginine-p-Nitroanilide (BApNA), L-trans-epoxysuccinyleucyl amido-(4-guanidino) butane (E-64); alpha 2-macroglobulin (α 2-M); *Dimorphandra mollins* trypsin inhibitor (DMTI); high performance liquid chromatography (HPLC); phenylmethanesulfonyl fluoride (PMSF); soybean trypsin inhibitor (SBTI), sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), N-tosyl-lysine chloromethyl ketone (TLCK), N-tosyl-L-phenylalanine

chloromethyl ketone (TPCK), small-angle x-ray scattering (SAXS), circular dichroism (CD), prothrombin time (PT).

Summary

BJ-48, a serine protease from the venom of *Bothrops jararacussu* was purified to homogeneity using affinity chromatography on p-aminobenzamidine-agarose followed by HPLC gel filtration (35-fold, 77% yield). BJ-48 presented an isoelectric point of 4.5 and an apparent molecular weight of 52 kDa by SDS-PAGE analysis and 48,036 Da by ESMS. The enzyme was shown to be highly glycosylated presenting 42% of N-linked carbohydrate (~12% sialic acid) with the composition Fuc(1):GalN(4):GlcN(5):Gal(1):Man(2). BJ-48 had optimal esterase activity at pH 7.5 and displayed maximum catalytic rate at 50°C. The kinetics of BJ-48 with chromogenic substrates indicates a higher specificity for arginine than lysine in the primary subsite and an apparent faster attack to ester than amide linkages. BJ-48 was shown to interfere with blood coagulation at prothrombin time (PT) assay conditions. The enzyme mostly resembles typical snake venom thrombin-like enzymes (SVTLEs) as it (i) displayed strict substrate specificity cleaving the B-beta chain of the bovine fibrinogen faster than the A-alpha while showing no effect on the gamma chain and (ii) presented 53% identity to the 22 N-terminal residues of reptilase and a close degree of similarity to other SVTLEs. Its hydrolytic activity was strongly inhibited by aprotinin and DTT while it was less affected by benzamidine, PMSF and TLCK. TPCK, 6-aminocaproic acid, E-64 and the macromolecular inhibitors SBTI, hirudin, infestin, DMTI and alpha2-macroglobulin were ineffective. SAXS and CD analysis of BJ-48 in solution buffered at pH 8.0 allowed the characterization of the glycoprotein as prolate particle with highly flexible domains.

Introduction

Bothrops jararacussu (common name jararacuçu) is the most corpulent and maybe the most feared snake belonging to the *Bothrops* genus. Although presenting comparable envenoming characteristics with other genus members, such as *B. jararaca*, *B. moojeni* and *B. alternatus*, the accident caused by jararacuçu is potentially more dangerous [1]. This is true primarily because of the large amounts of venom that can be inoculated in a single bite, but also by the lack of neutralizing effect of the common “anti-bothropic” horse serum, despite of the fact that *B. jararacussu* venom integrates the antigenic pool used in anti-venom production.

Although death is an unusual outcome of *Bothrops* envenoming (0.3-0.5% of cases), in all such cases it is related to acute renal and respiratory failure, hemorrhage, shock and/or sepsis [2]. Coagulation disorders are associated with many of these clinical pictures, among which consumption coagulopathy is the most prevalent. The later is characterized clinically by prolonged prothrombin or activated partial thromboplastin times (PT and APTT, respectively) due to depletion of circulating fibrinogen which results in an inability or difficulty of blood to coagulate. Venom components involved in this process are known to be proteolytic enzymes pertaining to the serine and metalloprotease classes [3].

These fibrinogen-degrading serine proteases are part of a larger group of snake venom serine proteases (SVSP) able to interfere with the prey haemostatic system. These enzymes exert their toxic effects by cleaving specific bonds in one or more of the following haemostatic system components: kininogen (kalikrein-like activity); coagulation factors V, VII, VIII or X; fibrin/fibrinogen [fibrin(ogeno)lytic activity] and plasminogen (plasminogen activators) [4]. Not uncommonly, fibrinogenolytic SVSPs are inaccurately called thrombin-like enzymes (as in [5]), whereas this term should be

reserved for enzymes actually displaying the ability to clot fibrinogen *in vitro*, alone or in combination with other actions over the haemostatic system.

At present, numerous snake venom thrombin-like enzymes (SVTLEs) have been described from members of the Crotalinae and Viperidae families [6]. SVTLEs are single-chain enzymes able to release fibrinopeptides A and/or B, from fibrinogen A α and B β chains, respectively. Moreover, they all present trypsin-like specificity, cleaving after Arg or Lys in susceptible protein substrates, with the great majority of them displaying esterolytic and amidolytic activities over small synthetic substrates bearing an Arg at P1 position (Schechter and Berger nomenclature, [7]). Accordingly, SVTLEs are inhibited by one or more of the classical trypsin-like serine protease inhibitors such as benzamidine and TLCK or even by more general serine protease inhibitors as PMSF and DFP. Another common, but still unexplored, characteristic of these enzymes is their thermostability. Although SVTLEs have been intuitively subjected to functional assays at 37°C, most members of the family have not been characterized as to their temperature optimum or thermoresistance but the majority of the few that have been so showed an amazing thermostability of their arginine esterase or amidase activities, e.g. [8, 9].

Interestingly, a rapid survey of the literature demonstrates the striking biochemical diversity of SVTLEs. In particular, the reported molecular weights (mostly often M_r estimations from SDS-PAGE analysis) of SVTLEs range from 25-28 kDa to values as high as 65-70 kDa. Meanwhile, the availability of dozens of SVTLEs primary structures reveals that the molecular weight of these enzymes should not be greater than 25-28 kDa considering the average length of their polypeptide chains. This discrepancy suggests a major role for post-translational modifications (PTMs) in the molecular diversity of SVTLEs, among which glycosilation appears as the most important, e.g. [10-13].

It has been shown that many SVTLEs, differently from thrombin, do not activate factor XIII [12, 14, 15]. Indeed, fibrinogen hydrolysis by most SVTLEs produces non-cross-linked fibrin gels that are easily removed from the injury site by the action of plasmin [16]. As a result, intravenous injection of SVTLEs often cause benign defibrination due to formation of soft fibrin clots that can be rapid removed from the circulation either by fibrinolysis or via the reticulo-endothelial system [17, 18]. This beneficial physiological effect have prompted the use of SVTLEs, such as Batroxobin [15] and Ancrod [19], in the prevention of thrombi formation and in the improvement of blood flow by reducing blood viscosity [20, 21].

Our group has previously studied the ability of benzamidine-based affinity sorbents to selectively purify serine proteases from venoms of different snake species, including BJ-48, a 48 kDa SVSP from *B. jararacussu* venom [22]. Here we report extensive molecular characterization and solution structure studies of BJ-48.

Results

Purification, molecular weight and isoelectric point

BJ-48 protease was initially purified from the venom of *B. jararacussu* on p-aminobenzamidine-agarose affinity column, apparently to homogeneity, as judged by the single band in SDS-PAGE analysis (not shown). However, ESMS analysis of such fraction demonstrated the presence of two components with m/z 48,036 and 3,862 (Fig. 1). This behavior was constant in all preparations obtained, despite of a preliminary step of ultra-filtration on Centricon P10 filters. HPLC gel filtration was thus envisaged to obtain BJ-48 in pure form. The chromatographic profile of BJ-48 sample at 280 nm does not show the contaminating peptide, possibly due to the lack of aromatic amino acid residues in its composition (Fig. 2). However, the chromatogram does reveal a predominant peak (peak II) at retention time compatible with a 90 kDa species and a

significant shoulder (peak I) corresponding to an even larger species at gel filtration buffer (**Fig. 2**).

These results indicate possible shape heterogeneity in BJ-48 structure present in affinity chromatography fractions. Interestingly, BJ-48 migrates as a diffuse 52-54 kDa species in reducing SDS-PAGE analysis (**Fig. 2, inset lane B**) while in the absence of 2-ME in SDS-PAGE sample buffer, BJ-48 migrates as a slightly smaller species with a shaper band at 50 kDa (**Fig. 2, inset lane C**). Accordingly, 2DE analysis of BJ-48 confirmed these results and indicated a pI 4.5 for BJ-48 (**not shown**). Overall, starting from 100 mg of lyophilized venom (36.1 mg solubilized protein), almost 1 mg was retained by the affinity column (2.64% of the total) and a final 35-fold purification factor was achieved after HPLC gel filtration (**Table I**).

Activity dependence on pH and temperature

The esterase activity (TAME hydrolysis) of BJ-48 showed the expected bell-shaped dependence on pH at 25 °C, with the maximum centered at 7.5. Activity was about 55% of maximal at pH 6.0 and about 15% at pH 9.0 while almost 81% of activity was observed at pH 8.0. For practical reasons, the optimal temperature for BJ-48 catalysis was evaluated in terms of amidolytic activity over BApNA. BJ-48 displayed progressively higher hydrolysis rates from 25°C up to 45-50°C. The enzyme lost almost 50% of its maximal activity at assay temperatures above 60°C, being completely inactivated only at 80°C.

Kinetic parameters for synthetic substrate hydrolysis

In **Table II** are presented the kinetic parameters obtained for BJ-48 catalyzed hydrolysis of small chromogenic substrates. Apparently, TAME was a better substrate for BJ-48 than BApNA. TAME was also preferred over BAME mainly due to a 3-fold

increase in K_{cat} . $N\alpha$ -Tos-Gly-Pro-ME, lacking Arg in P1, was poorly processed by BJ-48, even worse than the smaller BAME and TAME. Regarding tripeptide substrates, BJ-48 presented negligible activity over the p-nitroanilide bearing Lys at P1 while the Arg containing analog was hydrolyzed with efficiency comparable to the ester substrates.

Effects of synthetic and protein inhibitors

The effects of various inhibitors specific for different classes of proteases were investigated (Table III). The esterase activity of BJ-48 was neither affected by a metal chelator (EDTA) nor by a thiol-alkylating agent (E-64), while the other chemicals studied inhibited the activity from 28% to 36%. In particular, under the experimental conditions chosen, the enzymatic activity was decreased by reagents that (i) reduce disulfide bonds (i.e., dithiothreitol, 36%), or (ii) are specific for serine proteases (benzamidine, 28%; PMSF, 24%; TLCK, 28%). These observations indicate that BJ-48 is neither a metalloproteinase nor a cysteine proteinase, but suggest that it is a serine proteinase with disulfide bonds essential for its integrity. In support of this, aprotinin, another known serine protease inhibitor, was the only non-synthetic compound able to decrease the enzyme activity (by about to the half). Other inhibitors specific for trypsin-like enzymes (SBTI, DMTI and Infestin) and an exosite I-dependent thrombin inhibitor (hyrudin) were unable to interfere with BJ-48 activity at conditions assayed. Moreover, α 2-macroglobulin, a macromolecular inhibitor circulating in human blood, was inefficient in blocking the BJ-48 catalytic activity.

Fibrinogenolytic activity

The time-course of human fibrinogen (plasminogen-free) digestion by the purified BJ-48 proteinase was accompanied by SDS-PAGE analysis of the reaction mixture samples (Figure 3). BJ-48 rapidly hydrolyzed the B β chain of fibrinogen and

degraded the A α chain much slower, but it did not digest the Y chain. This result is supported by the elution profile in reverse-phase HPLC of fibrinopeptides released from the action of BJ-48 over fibrinogen (**Fig. 4C**). BJ-48 released peptides with the same retention time as of FPB and FPA standards (**Fig. 4A and 4B**). As shown in **Fig. 4C** the relative amounts of fibrinopeptides liberated by BJ-48 were proportional to the extent of degradation in fibrinogen chains (**Fig. 3**), with FPB being preferentially released. Apart from fibrinopeptides, no other peptides or single amino acids were detected. These observations indicate that BJ-48 actions over human fibrinogen mimic that of thrombin, including its outstanding specificity.

Procoagulant activity

The procoagulant activity of BJ-48 was evaluated by direct incubation in PT assays carried out with human blood. BJ-48 was able to reduce the normal PT (14-18 s) to the half at assay concentration of 35 $\mu\text{g/mL}$ and coagulated blood immediately at 80 $\mu\text{g/mL}$.

N-terminal sequence analysis

The N-terminal analysis of the BJ-48 polypeptide (~200 pmol) furnished an unambiguous sequence of 22 residues with two putative half-cystines (represented as X): VVGGDXIPQVPFLAFLYSEYFX. Cysteine identification is typically difficult while sequencing peptides using the Edman degradation reaction, because its phenylthiohydantoin derivative is unstable and for this reason cysteine is often reduced and alkylated with tri-n-butylphosphine and 4-vinylpyridine, respectively. A BLAST search against protein sequence databases identified statistically significant similarity (66-73% identity in 19-26 amino acids compared) with the N-terminal of three SVSPs. These were the hypothetical enzymes from *B. insularis* (tr-EMBL accession code

Q8QG86) and *B. jararaca* (tr-EMBL accession code Q5W958), and bothrombin (also known as reptilase, Swiss-Prot accession code P81661), a TLE isolated from *B. jararaca* venom. A multiple sequence alignment of this sequence with the N-terminal of several SVSPs (including many SVTLEs) evidenced the remarkable similarity of BJ-48 with this group of enzymes (**Fig. 5**). For instance, the BJ-48 N-terminal segment of 22 residues presented 79.7% and 68.5% similarity with the batroxobin and flavoxobin sequences, respectively.

BJ-48 glycosylation analysis

The glycosylation level of BJ-48 was assessed by the use of several glycosidases under denaturing conditions (**Fig. 6**). Under the action of a non-specific neuraminidase, BJ-48 had its M_r reduced by almost 8% (**Fig. 6, lane A**) indicating the presence of terminal sialic acid residues in the glycoprotein. Treatment of the desialylated protein with O-glycosidase had no effect in its electrophoretic mobility (**Fig. 6, lane B**) suggesting the absence of sialic acid-substituted Gal- β (1-3)-GalNAc- α - cores linked to Ser or Thr residues. Conversely, PNGase F, an enzyme capable of cleaving virtually all Asn-linked complex, hybrid, or high mannose oligosaccharides [with exception of α (1-3) core fucosylated], promoted a marked increase in electrophoretic mobility of denatured BJ-48 (**Fig. 6, lane C**), corresponding to a loss of 41.6% in the glycoprotein weight. The combined action of all above mentioned glycosidases in addition to β (1-4)-galactosidase and β -N-acetylglucosaminidase resulted in a species with the same M_r as the one produced by PNGase F action alone (**Fig. 6, lane D**). These results strongly suggest that all of the carbohydrate moieties in BJ-48 enzyme are linked to Asn residues.

The carbohydrate content of BJ-48 was also assessed by chemical methods. The sialic acid content was determined as $11.8 \pm 3 \%$, a figure somewhat higher than

estimated by the enzymatic method. The molar composition of neutral sugars present in BJ-48 was calculated as Fuc(1):GalN(4):GlcN(5):Gal(1):Man(2). Such composition is also consistent with a complex of the N-linked type.

Analysis of BJ-48 structure in solution

The CD spectrum of BJ-48 protein in solution buffered at pH 8 is shown in [Figure 7](#). It suggests a high content of random-coil structures in this glycoprotein. Such feature is completely unexpected for an active enzyme known to belong to the chymotrypsin fold. Further characterization of BJ-48 structure in solution was achieved by SAXS experiments. Scattering data for BJ-48 solutions at 4 mg/mL were collected at a synchrotron facility which enabled detection of a scattering vector range, $q = (4\pi/\lambda)\sin(\theta)$ (λ = wavelength and 2θ = scattering angle), equal to $0.0165 \text{ \AA}^{-1} < q < 0.2192 \text{ \AA}^{-1}$.

The plot of corrected intensity data as a function of the modulus of q is shown in [Figure 8A](#). From the GNOME desmeared fit of scattering data the indirect Fourier transform was performed and the resulting pair distance distribution function, $p(r)$, is shown in [Figure 8B](#). The maximum dimension (D_{max}) for the solvated BJ-48 particle was $\sim 130 \text{ \AA}$ as determined from the radius (r) value where $p(r)$ function returns to zero. From the pair distribution function we also obtained the radius of gyration (R_g) of the scattering particle which was calculated as $37 (\pm 1) \text{ \AA}$ for BJ-48 at pH 8.0. From this analysis of the $p(r)$ function we had an indication that the protein presents an elongated (prolate) shape. In addition, using ovalbumin as a secondary standard in the SAXS experiments it was possible to estimate the molecular weight of BJ-48 as $50 (\pm 3) \text{ kDa}$, which is in very good agreement with the expected 48 kDa value. This observation gave a strong indication that the measured sample was monodisperse and free of aggregates. The $I \cdot q^2 \times q^2$ plot (also known as Kratky plot) in [Figure 8C](#) indicated the presence of

flexible domains in BJ-48, corroborating the excess of random coils interpreted from CD spectrum.

Discussion

BJ-48 can be easily purified with only two steps

It was previously demonstrated that a commercial affinity sorbent based on p-aminobenzamidine, specifically produced through CNBr coupling of a glycil-glycine spacer arm to an agarose support, represent an efficient chromatographic medium to isolate the BJ-48 enzyme from *B. jararacussu* venom [22]. Subsequent characterization of the isolated serine protease by ESMS evidenced the presence of a contaminating 3,862 Da species that could not be detected *a priori* by gel electrophoresis (Fig. 1). Hence, an additional HPLC gel filtration step had to be performed in order to obtain BJ-48 in pure form (Fig 2). As indicated in Table I, this modification in the purification scheme of BJ-48 was responsible for a loss of 0.4% in the final protein content of the preparation while keeping the level of amidolytic activity recovered (76.2 %) and slightly improving the purification factor (from 29.6 to 34.9). Interestingly, the 4 kDa contaminant had not been separated by the preliminary ultrafiltration step using a filter with 10 kDa cutoff. These results are indicative of possible weakly-bound non-inhibitory complex between this species and BJ-48.

BJ-48 activity is highly resistant to pH changes and heating

BJ-48 esterase activity displayed the typical pH optimum of chymotrypsin family serine proteases which is primarily imposed by the presence of a His residue in the catalytic site [23]. Interestingly, BJ-48 was stable to long-term storage and repeating freeze-and-thaw cycles at solutions as acidic as 1 mM HCl (pH 3). This particular resistance to low pH values has been observed for other SVSPs, such as *Lachesis muta*

47 kDa TLE [22] and *Bothrops* protease A (BPA) [24]. Clearly, this is not the case of many of the orthologous mammalian enzymes. For instance, human α -thrombin eluted with low-pH buffers from beanzamidine-based affinity column needs to be immediately transferred to solutions buffered at neutral pH in order to maintain its full catalytic activity [25].

Other noticeably aspect of BJ-48 catalytic activity is its ability to cleave amide bonds of susceptible substrates with increasing rates at temperatures above 37°C and up to 50°C. Although temperature profiles for SVSPs are rarely found in the literature (see ref [26] for an exception), the more commonly reported thermostability assays indicate that the great majority of fibrinolytic SVSPs are resistant to varying degrees of heating above 37°C [8-9, 26-29]. For instance, the fibrinogen-clotting enzyme purified by Zaganelli and co-workers [8] was shown to be stable to treatment at 40°C, losing 70% of its activity after 15' heating at 70°C while, on the other side of the spectrum, BPA presented resistance to heating at 86°C for 10' [28]. One extreme case was that reported by Hung *et al.* where a β -fibrinogenase displayed heat-stability to about 95°C, with 65% of the activity remaining after 30' treatment [29]. The origin of such varying resistance to heating is currently unknown.

Protein thermostability has been positively associated with the number of disulfide bridges, solvent-exposed salt bridges and the presence of divalent metal binding sites in the protein structure [30,31], while the effect of the ratio of hydrophilic/hydrophobic residues exposed on the molecular surface is still controversial [30,32-33]. SVSPs are known to have six conserved disulfide bridges [4], the same number as the prototypical digestive trypsin and two more than thrombin, yet the later enzymes present optimum activity around 45 °C [34-35]. On the other hand, the interaction of carbohydrate moieties with trypsin, either by chemical attachment to the

enzyme structure [36] or through addition of the sugar to enzyme solution [37] produces a markedly enhancement in the protein thermostability. This fact suggests a role for the extensive glycosylation of BJ-48 (and many other SVSPs) in the observed high resistance to heating.

BJ-48 kinetic parameters reveals a striking specificity for Arg in P1

The catalytic properties of BJ-48 enzyme were further characterized by comparison of kinetic parameters evaluated for distinct ester and amide chromogenic substrates (Table II). BJ-48 was able to hydrolyze both p-nitroanilides and methyl esters of arginine, apparently with higher specificity for ester bonds, as judged by the comparison of K_{cat}/K_m parameters for BApNA and TAME. Caution is needed in this interpretation since amidase assays were performed with racemic BApNA, which is known to be less sensible to hydrolysis by trypsin-like enzymes than the enantiomeric pure L isomer due to the inhibitory action of the D isomer [38]. A slight preference for the tosyl over the benzoyl $N\alpha$ protecting group was also detected by comparing K_{cat} values for TAME and BAME hydrolysis. The requirement for a positively charged residue at P1 position of substrates, characteristic of trypsin-like enzymes, was suggested by the remarkable low activity of BJ-48 against $N\alpha$ -Tos-Gly-Pro-ME.

Unexpectedly, however, BJ-48 showed a strict specificity for Arg over Lys at P1, being unable to hydrolyze a tripeptidyl p-nitroanilide with Lys at P1. Distinctly, α -thrombin was able to hydrolyze the Lys containing substrate displaying an ordinary preference of less than 2 orders of magnitude for Arg. This observation is possibly reflective of the fact that $N\alpha$ -Tos-Gly-Pro-X-pNA is a substrate not as good for the snake venom enzyme as it is for the mammalian protease. The kinetic data for the Arg bearing substrate in Table II suggest this to be the result of BJ-48 presenting both lower binding affinity (2 orders of magnitude higher K_m) and catalytic efficiency (25-fold

higher K_{cat}) for this ligand. The former is probably related to a poor fit to the S2/S3 specificity pockets while the later can be the result of a more strained and energetic transition state.

Other BJ-48 properties revealed by inhibition studies

The identification of BJ-48 as a serine protease with trypsin-like specificity was further evidenced by the reduction in activity achieved by incubation with benzamidine and TLCK and other more general serine protease inhibitors, such as PMSF and aprotinin (Table III). Other relevant BJ-48 properties arise from data in Table III. First, disulfide bonds appear to be important for maintaining a folded, catalytic competent, structure in accordance with inhibitor studies with the SVTLEs Calobin II [39] and Jerdofibrase [40]. Second, BJ-48 resistance to inhibition by hirudin mostly probably results from the absence of an exosite I similar to the one found in thrombin. This result has a bearing on the mechanism of BJ-48 interaction with its biological substrate and other potential macromolecular ligands since thrombin's exosite I is also known to be responsible for fibrinogen recognition and for mediating interactions with co-factors, such as thrombomodulin [41]. Some SVTLEs have been shown to lack inhibition by hirudin and alternative "fibrinogen recognition exosites" have been proposed by molecular modeling [4].

Finally, BJ-48 was also refractory to inhibition by some known serine protease macromolecular inhibitors, including DMTI, Infestin and SBTI. For the homologous TSV-PA, absence of inhibition by SBTI was demonstrated to be fundamentally associated to steric hindrance at the S2' subsite as conferred by the bulk Phe side chain at position 193 of this SVSP [42]. More recently, an alternative SVSP resistance mechanism to SBTI has emerged from modeling studies on the structures of two almost identical enzymes from *A. actus* [43]. The proteins differ by the identity of a single

amino acid at position 35, either an Asn or an Asp. Only the variant with Asn is resistant to SBTI and based on molecular modeling it was suggested that the carbohydrate moiety linked to Asn³⁵ would be responsible for blocking SBTI binding.

BJ-48 is mostly probably a SVTLE

BJ-48 was shown to be a fibrinolytic enzyme. Its specificity was assessed by following the time-course of fibrinogen degradation by SDS-PAGE (Fig. 3) and through reverse-phase HPLC analysis of peptides liberated in the hydrolysis (Fig. 4). The results clearly demonstrate that BJ-48 acts specifically on fibrinogen to release preferentially FPB, but also FPA from chains B β and A α , respectively, leaving the γ chain intact. It was once believed that fibrinogen-degrading SVSP could be differentiated from fibrinolytic snake venom metalloproteases (SVMPs) through their specificity over fibrinogen [6]. According to this view, SVSPs were thought to degrade the B β chain with rare exceptions while SVMPs act solely over the A α chain. At present, the plethora of identified fibrinolytic SVSPs degrading A α , B β or both chains has irrefutably denied such association [4].

The *in vitro* procoagulant activity of BJ-48 was clearly demonstrated under PT assays. The enzyme was able to replace completely the thromboplastin reagent at BJ-48 assay concentration of 80 $\mu\text{g}/\text{mL}$ and could reduce coagulation time to the half at 35 $\mu\text{g}/\text{mL}$. The PT assay probes the extrinsic tenase (membrane bound tissue factor, activated factor VII and Ca^{2+} ions) and prothrombinase (activated factors X and V, along with Ca^{2+} ions and a phospholipid surface) complexes and the conversion of fibrinogen to fibrin clot, and is usually employed to measure the effect of an anticoagulant present on a patient blood circulation [44]. Nevertheless, under our experimental conditions, BJ-48 most probably affected PT by mimicking the fibrinogen clotting activity displayed by human thrombin, whose formation from prothrombin in

assayed blood is dependent upon the addition of Ca^{2+} ions and thromboplastin reagent (containing tissue factor and phospholipids).

N-terminal sequencing of BJ-48 revealed a clear homology to other SVSPs (**Fig. 5**). Moreover, BJ-48's first 22 amino acid residues showed the highest similarity to well characterized SVTLEs, such as reptilase [45], batroxobin [46] and flavoxobin [47]. These and the previously discussed results, altogether, strongly suggest that BJ-48 is actually a SVTLE. Three other SVTLEs have been isolated from *B. jararacussu* venom, all of them clearly structurally distinct from BJ-48 but presenting comparable specific clotting activities [9, 26, 48]. The fibrinogen-clotting enzyme reported by Zaganelli and co-workers [9] was shown to have 50.55 NIH U/mg. The coagulant specific activity of Jararacussin-I was determined to be 45.8 NIH U/mg by Bortoleto *et al.* [26]. The clotting factor isolated by Andrião-Escarso *et al.* [43] was reported to have 42.5 U/mg. Considering the usual *in vivo* defibrinogenating activity displayed by SVLEs, apparently, *B. jararacussu* venom glands evolved to present a redundant anticoagulant toxic effect probably due to its importance to achieving high lethal efficacy in envenoming.

BJ-48 is highly N-glycosylated

BJ-48 molecule was shown to be comprised of more than 40% of carbohydrates. Near 20% of these carbohydrate moieties are constituted by sialic acid residues (8% of the entire glycoprotein mass). Moreover, all of the carbohydrate moieties in BJ-48 are attached to Asn residues. N-linked oligosaccharides possess molecular weights ranging from 1884 Da, for a high-mannose carbohydrate ($\text{GlcNAc}_2\text{Man}_9$), to 2851 Da, for a sialylated triantennary complex saccharide ($\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_3\text{Gal}_3\text{NeuAc}_3$) [49]. Considering the BJ-48 carbohydrate composition and the high content of sialic acid residues, N-linked saccharides in this snake venom enzyme are presumably almost all of

the complex sialylated triantennary type. In this manner, the near 23 kDa of sugars in BJ-48 structure can be predicted to be distributed over 6-10 Asn residues exposed on its surface.

To the best of our knowledge, the most glycosylated SVSP reported to date is BPA, an enzyme known to be extremely resistant to heating [28]. BPA has 62% of its native molecular weight constituted by sugars, most of which were shown to be N-linked glycans. Other SVSPs, many of them belonging to the SVTLEs group, are known to carry a fair amount of glycans on their structures, most of them of the N-linked type [9-13; 15; 48; 50-51]. Carbohydrate contents in these enzymes vary between 5-30% of glycoprotein mass, most often corresponding to 20% of the molecular mass. Examples include the clotting enzyme from *B. jararacussu* venom purified by Zaganelli *et al.* [9] as well as SVTLEs from other snake species such as TL-BJ (*B. jararaca*) [11] and from other genera like the enzyme purified from *Agkistrodon saxatilis* venom by Koh *et al.* [10] and Gabonase (*Bitis gabonica*) [50].

N-linked glycans are known to affect protein structure in two ways: facilitating the protein folding process or stabilizing the mature protein [49,52]. More specifically, N-glycosylation has also been shown to influence the catalytic activities of many proteases from the serine as well as other catalytic types [53-55]. Giving such huge contribution to the molecular mass of BJ-48 (as in other SVSPs), N-linked carbohydrates would certainly produce an effect in its structures and, consequently, on function. Studies exploring this feature are currently being held in our group. In BPA, from the analysis of a partially N-deglycosylated form of the enzyme it was shown that glycosylation has no impact on the catalytic activity over small-chromogenic substrate, yet it presented a favorable contribution in the resistance to extreme pH values and a global stabilizing effect on the protein structure [28]. As mentioned before in this text,

maybe the most direct evidence for a biochemical function displayed by carbohydrate moieties in SVSPs came from the structural analysis of Zhu and co-workers on the crystal structures of two plasminogen activators from *A. actus* venom [43].

BJ-48 is a prolate, yet highly flexible particle in solution

At present, several reports of crystallization and data collection for SVTLEs have appeared in the literature [56-59], but to the best of our knowledge no structure has been deposited in the Protein Data Bank. Indeed, the PDB is plenty of 3D structures for several serine proteases including two SVSPs [43,60], whose homology to SVTLEs can be used to model these enzymes [61], yet some of their functional aspects are unique and may not be confidently modeled. Moreover, the high glycosylation of BJ-48 and many other SVTLEs prevent obtaining suitable crystals for high resolution structural studies. In this sense, the low-resolution SAXS technique emerged as an appropriate choice for avoiding the laborious and often disappointing protein crystallization steps. Protein solutions, essentially composed of light atoms (H, C, N and O) in water, display little contrast of electron density and therefore weak scattering power. For this reason the use of an X-ray source with very high flux, such as that provided by synchrotron radiation storage rings, is ideal for SAXS studies of protein solutions [62].

SAXS data collected for BJ-48 suggest a non-globular shape for this enzyme in solution. This result is in accordance to the hydrodynamic behavior of BJ-48 in gel filtration chromatography. Under gel filtration conditions, BJ-48 presents retention time compatible with a 90 kDa globular protein. In view of the ESMS measure of 48 kDa for BJ-48, the overestimated chromatographic value must be consequence a rupture in the relationship between hydrodynamic radius and molecular mass, as derived from our gel filtration calibration curve constructed with globular proteins. Both the prolate shape

presumed from the SAXS $p(r)$ function (**Fig. 8B**) and the molecular mass of 48 kDa obtained from extrapolation to I_0 are compatible with the previous conclusion.

Other property of the BJ-48 solution structure is its high flexibility, as indicated by CD spectrum (**Fig. 7**) and the Kratky plot derived from SAXS data (**Fig. 8C**). The latter gives information on the compactness of the protein structure [62]. For particles with compact shape and devoid of flexible domains these plots show a well-defined curve with a typical bell shape whereas for particles with random coil conformation the Kratky plot shows a characteristic plateau and a monotonic rise at higher angles (lower q). Structured particles with flexible domains will have a well defined maximum but also a descending curve that does not reach the horizontal axis.

The observation of a prolate shape for BJ-48 was completely unexpected. In SAXS studies for enzymes of the chymotrypsin family, $p(r)$ functions compatible with a globular shape are obtained, in agreement with structures determined by X-ray crystallography [63]. In view of the proteolytic activity exhibited by BJ-48 in the conditions used for SAXS measurements, the possibility of a denatured state for the protein can be ruled out. Given the extensive glycosylation on the BJ-48 surface, it can be hypothesized that the highly flexible carbohydrate chains would be having a dominating effect on the SAXS scattering pattern.

Conclusion

We have presented extensive biochemical characterization for BJ-48, a 48 kDa serine protease purified from the venom of *B. jararacussu*. The enzyme was shown to mimic thrombin actions over fibrinogen and preferentially hydrolyzes ester bonds. As a SVSP, it is strongly inhibited by aprotinin and DTT, yet it is resistant to hirudin, a thrombin inhibitor dependent on exosite I interaction. Other macromolecular inhibitors, such as SBTI, infestin and DMTI were ineffective. The molecular mass of BJ-48 is

more than 40% comprised of N-linked carbohydrate chains. The protein is highly flexible in aqueous solution buffered at pH 8 and presents a time-averaged prolate shape. Work in our group is currently being held to determine the functional relevance of the extensive glycosylation of BJ-48 as well as its contribution to the highly flexible and unexpected non-globular shape for the protein in solution.

Experimental Part

Materials and B. jararacussu venom

Sequence reagents, PTEE (polytetrafluoroethylene) membrane, polybrene and 6-aminocaproic acid were from Wako Pure Chemicals (Richmond, CA, U.S.A.). Shinpack-Diol 150 and CLC-ODS HPLC columns were obtained from Shimadzu (Kyoto, Japan). Benzamidine-agarose, fetuin, tri-n-butylphosphine, 4-vinylpyridine, HPLC marker proteins, chromogenic substrates, bovine and human fibrinogen, peptide substrates and inhibitors were from Sigma, Chemical Co., (St Louis, Mo, U.S.A.). Centricon P10 filters were obtained from Amicon (CA, U.S.A.) while the SDS-PAGE protein marker kit was from Biolabs Laboratories (New England, U.K.). *B. jararacussu* venom was extracted manually from several specimens kept at the Vital Brazil Institute. The venom was dried and stored frozen at -20°C until use.

Isolation of Bj-48 TLE by p-aminobenzamidine affinity chromatography

The initial isolation of BJ-48 TLE was achieved by affinity chromatography, essentially as described before [22]. Briefly, the dry venom (100 mg) was reconstituted in 3 mL equilibrating buffer and subjected to ultra-filtration in Centricon-10 filters. The retentate (> 10 kDa) was loaded into a 5 mL benzamidine-agarose column previously equilibrated with 50 mM Tris-HCl containing 0.5 M NaCl, pH 7.0 and, after extensive washing, the adsorbed proteins were eluted with 1 mM HCl, pH 3.0, containing 0.5 M NaCl.

Gel filtration high performance liquid chromatography

The concentrated post-benzamidine-agarose preparation was injected in a Shin pack Diol-150 (50 cm x 7.9 mm, I.D.) HPLC column previously equilibrated in 50 mM phosphate buffer pH 7.2. The protein was fractionated on an automatic HPLC system (Shimadzu, 6A model) at a flow rate of 1 mL/min, during 28 min at 25°C. For molecular mass characterization, the column was calibrated in the same buffer with the following markers: β -galactosidase (105 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa) and carbonic anhydrase (29 kDa).

Two-dimensional gel electrophoresis (2DE)

Proteins were separated in the first dimension using the Ettan IPGphor isoelectric focusing system (Bio-Rad). Then 120 μ g of protein was mixed with rehydration buffer containing 8 M urea, 2% 3-[(3-cholamidopropyl)dimethylamino]-1-propanesulfonic acid (CHAPS), 0.5% carrier ampholytes buffer, and 65 mM dithiothreitol (DTT). Rehydration and isoelectric focusing were performed according to the manufacturer's instructions using pH 4-10 IPG strips (13 cm). Following isoelectric focusing, proteins were reduced and alkylated by successive 15-min treatments with equilibration buffer containing 2% DTT followed by 2.5% iodoacetamide. Then, proteins were resolved in the second dimension on a 12.5% SDS-PAGE gel. Resolved proteins were then either transferred directly to PVDF membrane for N-terminal sequencing or stained with Coomassie blue.

Mass spectrometry (MS) analysis

Samples were dissolved in 50% (v/v) acetonitrile [containing 0.1% (v/v) TFA] and analyzed on a triple quadrupole mass spectrometer Micromass (Altrincham), model QUATTRO II, equipped with a standard electrospray probe, adjusted to ca. 5 μ l/min.

During all experiments the source temperature was maintained at 80°C and the needle voltage at 3.6 kV, applying a drying gas flow (nitrogen) of 200 L/h and a nebulizer gas flow of 20 L/h. The mass spectrometer was calibrated with intact horse heart myoglobin and its typical cone-voltage induced fragments.

N-terminal sequence determination

NH₂-terminal amino acid sequences of peptide samples were determined by automated Edman degradation using a gas-phase protein microsequencer (Model PSQ-1; Shimadzu, Kyoto, Japan) and the conditions described before [22]. Briefly, the toxin was immobilized on PVDF membrane and 5 µL of the mixture of 0.6% tri-n-butylphosphine and 1.2 µL of 4-vinylpyridine in 80% isopropanol was added for the simultaneous reduction and alkylation of the toxin successively followed by Edman degradation reaction. The optimal reaction temperature and time for the reaction were 60°C and 50 min respectively.

Sequence analysis

Similarity searches with the BJ-48 N-terminal sequence against UniProt knowledgebase release 5.3 (Swiss-Prot and tr-EMBL databases, releases 47.3 and 30.3, respectively) were performed using BLAST program implemented in ExPASy proteomics server (<http://br.expasy.org/>). Sequences of SVSPs were obtained through keyword search in Swiss-prot and tr-EMBL with the Sequence Retrieval System server (SRSWWW) at ExPASy, version 5.1.0. SVSPs and BJ-48 N-terminal sequences were multiple-aligned within ClustalX v1.81 and the alignment processed and formatted with BioEdit v5.0.6 tool.

Determination of optimal pH and temperature of activity

The optimal pH was determined incubating the enzyme (1 μg) for 30 min at room temperature with 0.25 mM TAME in different pH values. Enzyme activity was monitored in 100 mM of sodium citrate buffer (pH 5.0-6.0) and Tris-HCl (pH 7.0-9.0). For determining the optimal temperature, 2 μg BJ-48 in 20 μL water was added to 180 μL of a reaction mixture composed of 100 mM Tris-HCl, pH 7.5 + 0.1% DMF (Tris-buffer) + 1.25 mM BApNA and the change in absorbance at 410 nm monitored for 15 min at temperatures ranging from 25°C to 60°C. Specific activity was expressed in μmoles of formed product per minute per mg of protein.

Enzyme assays and determination of kinetic parameters

Chromogenic substrates TAME, BAME and BApNA were digested in 50mM Tris-HCl, pH 8.0, at 25°C for 30 min. After addition of enzyme (1 μg), digestion of substrate was followed by measurement of change in absorbance at 247 nm, 253 nm and 410 nm, respectively. The activities were calculated using the following molar extinction coefficients for TAME, BAME and BApNA, respectively: 409 $\text{M}^{-1}.\text{cm}^{-1}$, 1,150 $\text{M}^{-1}.\text{cm}^{-1}$ and 8,800 $\text{M}^{-1}.\text{cm}^{-1}$. For determination of Michaelis constant (K_m), maximal velocities (V_{max}) and catalytic constant (K_{cat}) at least fifteen substrate concentrations between 0.00125 mM and 2.5 mM were used.

Effects of various inhibitors on TAME hydrolysis

A 100 μl reaction volume of 50 mM Tris-HCl buffer, pH 8.0, containing 1 μg enzyme and the appropriate inhibitor at varying concentration were incubated for 30 min at room temperature (25 °C). Following incubation, 30 μl of 0.1 mM TAME (Tris-HCl buffer, pH 8.0) was added and the absorbance change recorded as described above, after 30 min. Control solutions, lacking inhibitors, were run simultaneously. Inhibition activity was calculated on the basis of at 4 experiments and is reported as per cent.

Fibrino(geno)lytic activity

The hydrolytic activity of the BJ-48 on fibrinogen was evaluated by SDS-PAGE and C18 HPLC. For electrophoretic analysis the time-course of fibrinogen hydrolysis was accompanied after incubating 1 μg of enzyme with 20 μg of human fibrinogen (30 to 180 min) at 37°C in 50 mM Tris-HCl pH 8.0 buffer. The incubation was stopped by mixing reaction mixtures with SDS-PAGE sample buffer and heating for 3' at 95 °C. For C18 HPLC analysis of peptides released upon fibrinogen hydrolysis by BJ-48, 1 μg of enzyme was incubated with 1 mg of human fibrinogen for 90 min at reaction conditions described above. The reaction was stopped incubating the reaction tubes in boiling water for 3 min. Denatured proteins were precipitated by centrifugation at 20.000g for 10'. Soluble peptides in the supernatant were analyzed in the HPLC.

Prothrombin time (PT)

Citrated blood (fresh collected human blood added of 3.2% sodium citrate) was centrifuged at 2000 RPM (table top clinical centrifuge) for 10-15 min and the plasma decanted. Prior to measuring the coagulation time, assay tubes containing 0-200 μL calcium thromboplastin were incubated (water-bath at 37°C for 3 min) with BJ-48 solutions ranging in concentration from 0 to 200 $\mu\text{g}/\text{mL}$. Finally, 100 μL citrated plasma were added and the coagulation time measured at regular intervals.

Determination of BJ-48 glycosylation type and content

The presence of carbohydrate was investigated using an enzymatic deglycosylation kit according to the manufacturer instructions (Sigma, Chemical Co., St Louis, Mo, U.S.A). Briefly, glycosidases were employed to virtually cleave all N- and most of the O-linked oligosaccharides from glycoproteins or glycopeptides. Fetuin was

used as an internal standard control and the reaction products were analyzed by SDS-PAGE as described above.

Determination of sialic acid content

The sialic acid content was determined according to the Dionex technical note 41 (Dionex Inc., Sunnyvale, CA). Briefly, two purified BJ-48 aliquots (125 μg) were dissolved in 1N acetic acid and heated (80°C) for 2h. D-glucuronic acid was added as an internal standard and the samples analyzed in a Dionex Biol C chromatographic system equipped with a PA-10 column. Chromatographic runs were performed at 1 mL.min⁻¹ flow applying a gradient of 7-30% solvent B (1M sodium acetate in 100 mM NaOH) in solvent A (100 mM NaOH).

Composition and proportion of monosaccharides

BJ-48 aliquots were dissolved in 200 μL of 2N acetic acid and heated (100°C) for 4h. Hydrolysed samples were subsequently lyophilised, dissolved in deionized water and injected in the Dionex Biol C system described above. The chromatography was developed using 200 mM NaOH and deionized water as mobile phases A and B, respectively. Elution was achieved by applying 9% A for 20 minutes. Fetuin was used as a control for the quantification of sugars and 2-deoxy-D-glucose was employed as an internal standard.

Circular dichroism (CD) analysis

CD spectra were recorded on a Jasco J720 spectro-polarimeter using a quartz cuvette with 0.2 cm path-length. Measurements were carried out at room temperature (25 °C) in 10 mM Tris-HCl buffer, pH 7.0, at a final protein concentration of 1 mM.

Small-angle x-ray scattering (SAXS) analysis

SAXS experiments were performed at the beamline of the National Synchrotron Laboratory, Campinas, Brazil [64]. The experimental setup included a temperature-controlled, 1.5-mm-diameter capillary tube sample holder [65] and a linear position-sensitive detector. Data acquisition was performed by taking five 600-s frames, using a sample detector distance of 446 mm and x-ray wavelength of 1.488 Å. All the data treatment of the scattering intensities was performed using the software package TRAT1D [64]. Usual correction for detector homogeneity, incident beam intensity, sample absorption, blank subtraction and intensities averaging were included in this routine. The output of this software provides the corrected experimental intensities and error values. Data analysis was performed using the computer programs GNOM [66]. The BJ-48 measurements were performed at a protein concentration of 4 mg.mL⁻¹ in a standard buffer (50 mM Tris-HCl, pH 8.0). In order to obtain the equivalent molecular weight of BJ-48 protein a standard sample of 5 mg.mL⁻¹ ovalbumin (42.2 kDa) was measured in the same experimental conditions.

General procedures

Protein contents were estimated either by the DC Bio-Rad protein assay kit using bovine serum albumin (BSA) as standard or by the absorbance at 280 nm (using the BJ-48 extinction coefficient $E^{(1\%, 1\text{cm})}$ of 13.4). The discontinuous SDS-PAGE was performed using Laemmli buffers [67] under reduction conditions. The gels were coomassie blue or silver stained (Bio-Rad kit). BSA (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (29 kDa), trypsinogen (24 kDa) and lactalbumin (14.4 kDa) were used as standards for characterization of molecular masses.

Acknowledgments

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support. Dr. Ligia Macedo for the generous gift of amounts of purified DMTI and Infestin inhibitors. Dr. Cristiane Martins for HPLC runs. Filipe Quirino and Claudia Conceição for chemical analysis of carbohydrates. Herbert Guedes for assistance in SAXS experiments. Carlos Correa and Eliane Nascimento for technical support. FPSJr was a CNPq Dsc. fellowship recipient.

References

1. Milani-Jr R, Jorge MT, Ferraz de Campos FP et al. (1997) Snake bites by jararacuçu (*Bothrops jararacussu*): clinicopathological studies of 29 proven cases in São Paulo state, Brazil, *Q J Med* **90**, 323-334.
2. Benvenuti LA, França FOS, Barbaro KC, Nunes JR & Cardoso JLC (2003) Pulmonary haemorrhage causing rapid death after *Bothrops jararacussu* snakebite: a case report, *Toxicon* **42**, 331-334.
3. Markland FS (1998) Snake venoms and the hemostatic system, *Toxicon* **36**, 1749-1800.
4. Serrano SMT & Maroun RC (2005) Snake venom serine proteinases: sequence homology vs. substrate specificity, a paradox to be solved, *Toxicon* **45**, 1115-1132.
5. Castro HC, Zingali RB, Albuquerque MG, Pujol-Luz M & Rodrigues CR (2004) Snake venom thrombin-like enzymes: from reptilase to now, *Cell. Mol. Life Sci.* **61**, 843-856.
6. Pirkle H (1998) Thrombin-like enzymes from snake venoms: An updated inventory, *Thromb Haemost* **79**, 675-683.

7. Schechter I & Berger A (1967) On size of active site in proteases. I. Papain, *Biochem Biophys Res Commun* **27**, 157-162.
8. Smolka MB, Marangoni S, Oliveira B & Novello JC (1998) Purification and partial characterization of a thrombin-like enzyme, Balterobin, from the venom of *Bothrops alternatus*, *Toxicon* **36**, 1059-1063.
9. Zaganelli GL, Zaganelli MGM, Magalhães A, Diniz CR & Lima ME (1996) Purification and characterization of a fibrinogen-clotting enzyme from the venom of jararacuçu (*Bothrops jararacussu*), *Toxicon* **34**, 807-819.
10. Koh Y-S, Chung K-H & Kim D-S (2001) Biochemical characterization of a thrombin-like enzyme and a fibrinolytic serine protease from snake (*Agkistrodon saxatilis*) venom, *Toxicon* **39**, 555-560.
11. Serrano SMT, Sampaio CAM, Mentele R, Camargo ACM & Fink E (2000) A novel fibrinogen-clotting enzyme, TL-BJ, from the venom of the snake *Bothrops jararaca*: purification and characterization, *Thromb Haemost* **83**, 438-444.
12. Huang Q-q, Teng M-k & Niu L-w (1999) Purification and characterization of two fibrinogen-clotting enzymes from five-pace snake (*Agkistrodon acutus*) venom, *Toxicon* **37**, 999-1013.
13. Aguiar AS, Alves CR, Melgarejo A & Giovanni-De-Simone S (1996) Purification and partial characterization of a thrombin-like/gyroxin enzyme from bushmaster (*Lachesis muta rhombeata*) venom, *Toxicon* **34**, 555-565.
14. Lu Q-M, Jin Y, Li D-S, Wang W-Y & Xiong Y-L (2000) Characterization of a thrombin-like enzyme from the venom of *Trimeresurus jerdonii*, *Toxicon* **38**, 1225-1236.

15. Holleman WH & Weiss LJ (1976) The thrombin-like enzyme from *Bothrops atrox* snake venom: properties of the enzyme purified by affinity chromatography on p-aminobenzamidine-substituted agarose, *J Biol Chem* **251**, 1663-1669.
16. Turpie AGG, Prentice CRM, McNicol GP, Douglas AS (1971) *In vitro* studies with Ancrod (Arvin), *Br J Haemat* **20**, 217-224.
17. Silberman S, Bernik MB, Potter EU & Kwaan HC (1973) Effects of Ancrod (Arvin) in mice: studies of plasma fibrinogen and fibrinolytic activity, *Br J Haemat* **24**, 101-103.
18. Mahir MS, Hynd JW, Flute PT & Dormondy JA (1987) Effects of defibrinogenation on the early patency rate of experimental small caliber arterial grafts, *Br J Surg* **74**, 508-510.
19. Nolan C, Hall LS & Barlow GH (1976) Ancrod, the coagulation enzyme from Malayan pit viper (*Agkistrodon rhodostoma*) venom. In *Methods in Enzymology* (Lorand L, ed), vol. 45, p. 205, Academic Press, New York.
20. Stocker KF (1988) Clinical trials with batroxobin. In *Hemostasis and Animal Venoms* (Pirkle H & Markland-Jr FS, eds), pp. 525-540, Marcel Dekker, New York.
21. Bell WR (1974) Defibrinogenation with Arvin in thrombotic disorders. In *Platelet and Thrombosis* (Sherry S & Sciabine A, eds), pp. 273-298, Urban-Schwarzenberg, Munich.
22. De-Simone SG, Correa-Neto C, Antunes OAC, De-Alencastro RB & Silva-Jr FP (2005) Biochemical and molecular modeling analysis of the ability of two p-aminobenzamidine-based sorbents to selectively purify serine proteases (fibrinogenases) from snake venoms, *J Chromatogr B* **822**, 1-9.

23. Hedstrom L (2002) Serine protease mechanism and specificity *Chem Rev* **12**, 4501-4523.
24. Reichl AP, Assakura MT & Mandelbaum FR (1983) Biophysical properties and amino acid composition of Bothrops protease A, a proteolytic enzyme isolated from the venom of the snake *Bothrops jararaca* (jararaca), *Toxicon* **21**, 421-427.
25. Nakamura K, Suzuki T, Hasegawa M, Kato Y, Sasaki H, Inouye K (2003) Characterization of p-aminobenzamidine-based sorbent and its use for high-performance affinity chromatography of trypsin-like proteases, *J. Chromatogr. A*, **1009**, 133-139.
26. Bortoleto RK, Murakami MT, Watanabe L, Soares AM & Arni RK (2002) Purification, characterization and crystallization of Jararacussin-I, a fibrinogen-clotting enzyme isolated from the venom of *Bothrops jararacussu*, *Toxicon* **40**, 1307-1312.
27. Nishida S, Fujimura Y, Miura S, Ozaki Y, Usami Y, Susuki M, Titani K, Yoshida E, Sugimoto A, Yoshioka A & Fukui H (1994) Purification and characterization of Bothrombin, a fibrinogen-clotting serine protease from the venom of *Bothrops jararaca*, *Biochemistry* **33**, 1843-1849.
28. Murayama N, Saguchi K, Mentele R, Assakura MT, Ohi H, Fujita Y, Camargo ACM, Higuchi S & Serrano SMT (2003) The unusual high molecular mass of Bothrops protease A, a trypsin-like serine peptidase from the venom of *Bothrops jararaca*, is due to its high carbohydrate content, *Biochim Biophys Acta* **1652**, 1-6.
29. Hung C-C, Huang K-F & Chiou S-H (1994) Characterization of one novel venom protease with β -fibrinogenase activity from the Taiwan habu (*Trimeresurus mucrosquamatus*): purification and cDNA sequence analysis, *Biochem Biophys Res Commun* **205**, 1707-1715.

30. Pack SP & Yoo YJ (2003) Protein thermostability: structure-based difference of residual properties between thermophilic and mesophilic proteins, *J Mol Catal B* **26**, 257-264.
31. Demidyuk IV, Zabolotskaya MV, Safina DR, Kostrov SV (2003) Molecular mechanisms of stabilization of proteolytic enzymes: A model of thermolysin-like microbial metalloproteases, *Russ J Bioorg Chem* **29**, 418-425.
32. Longo, MA & Combes, D (1999) Thermostability of modified enzymes: a detailed study, *J. Chem Technol Biotechnol* **74**, 25-32.
33. Mozhaev, VV; Meliknubarov, NS; Levitsky, VY; Siksnis VA, Martinek K (1992) High-stability to irreversible inactivation at elevated-temperatures of enzymes covalently modified by hydrophilic reagents - alpha-chymotrypsin, *Biotechnol. Bioeng.* **40**: 650-662.
34. Sipos T & Merkel JR (1970) An effect of calcium ions on the activity, heat stability and structure of trypsin, *Biochemistry* **9**, 2766-2775.
35. Leborgne S & Graber M (1994) Amidase activity and thermal-stability of human thrombin, *App. Biochem. Biotech.* **48**, 125-135.
36. Villalonga ML, Reyes G, Frago A, Cao R, Fernandez L, Villalonga R (2005) Chemical glycosidation of trypsin with O-carboxymethyl-poly-beta-cyclodextrin: catalytic and stability properties, *Biotechnol Appl Biochem* **41**, 217-223.
37. Fernandez L, Gomez L, Ramirez HL, Villalonga ML, Villalonga R (2005) Thermal stabilization of trypsin with glycol chitosan, *J. Mol Catal B* **34**, 14-17.
38. Junqueira RG, Mares-Guia M (1984) Competitive inhibition by the D-isomer in racemic mixtures used as substrate in kinetic studies: a simple method for data treatment. *Arch Biochem Biophys* **228**, 278-81.

39. Cho SY, Hahn B-S, Yang KY & Kim YS (2001) Purification and characterization of calobin II, a second type of thrombin-like enzyme from *Agkistrodon caliginosus* (Korean viper), *Toxicon* **39**, 499-506.
40. Jin Y, Lu Q-M, Wei J-F, Li D-S, Wang W-Y & Xiong Y-L (2001) Purification and characterization of jerdofibrase, a serine protease from the venom of *Trimeresurus jerdonii* snake, *Toxicon* **39**, 1203-1210.
41. Tsiang M, Jain AK, Dunn KE, Rojas ME, Leung LLK & Gibbs CS (1995) Functional mapping of the surface residues of human thrombin, *J Biol Chem* **270**, 16854-16863.
42. Braud S, Parry MAA, Maroun R & Bon, C (2000) The contribution of residues 192 and 193 to the specificity of snake venom serine proteinases, *J Biol Chem* **275**, 1823-1828.
43. Zhu Z, Liang Z, Zhang T, Zhu Z, Xu W & Teng M (2005) Crystal structure and amidolytic activities of two glycosylated snake venom serine proteases, *J Biol Chem* **280**, 10524-10529.
44. Kini RM & Banerjee Y (2004) Dissection approach: a simple strategy for the identification of the step action of anticoagulant agents in the blood coagulation cascade, *J Thromb Haemost* **3**, 170-191.
45. Nishida S, Fujimura Y, Miura S, Ozaki Y, Usami Y, Suzuki M, Titani K, Yoshida E, Sugimoto M, Yoshioka A, Fukui H (1994) Purification and characterization of bothrombin, a fibrinogen-clotting serine protease from the venom of *Bothrops jararaca*, *Biochemistry* **33**, 1843-1849.
46. Itoh N, Tanaka N, Mihashi S, Yamashina I (1987) Molecular cloning and sequence analysis of cDNA for batroxobin, a thrombin-like snake venom enzyme, *J Biol Chem* **262**, 3132-3135.

47. Shieh T-C, Kawabata S-I, Kihara H, Ohno M, Iwanaga S (1988) Amino acid sequence of a coagulant enzyme, flavoxobin, from *Trimeresurus flavoviridis* venom, *J Biochem* **103**, 596-605.
48. Andrião-Escarso SH, Sampaio SV, Cunha OAB, Marangoni S, Oliveira B & Giglio JR (1997) Isolation and characterization of a new clotting factor from *Bothrops jararacussu* (jararacuçu) venom, *Toxicon* **35**, 1043-1052.
49. Imperiali B & O'Connor S (1999) Effect of N-linked glycosylation on glycopeptide and glycoprotein structure, *Curr. Opin. Chem. Biol.* **3**, 643-649.
50. Pirkle H, Theodor I, Miyada D & Simmons G (1986) Thrombin-like enzyme from the venom of *Bitis gabonica*: purification, properties and coagulant actions, *J. Biol. Chem.* **261**, 8830-8835.
51. Aragon-Ortiz F & Gubensek F (1993) A thrombin-like enzyme from bushmaster (*Lachesis muta stenophrys*) venom, *Toxicon* **31**, 1435-1443.
52. Wujek P, Kida E, Walus M, Wisniewski KE & Golabek AA (2004) N-glycosylation is crucial for folding, trafficking and stability of human tripeptidyl-peptidase I, *J. Biol. Chem.* **279**, 12827-12839.
53. Aertgeerts K, Ye S, Shi L, Prasad SG, Witmer D, Chi E, Sang B, Wijnands RA, Webb DR & Swanson RV (2004) N-linked glycosylation of dipeptidyl peptidase IV (CD26): Effects on enzyme activity, homodimer formation and adenosine deaminase binding, *Prot. Sci.* **13**, 145-154.
54. Barinka C, Sacha P, Sklenar J, Man P, Bezouska K, Slusher BS & Konvalinka J (2004) Identification of the N-glycosylation sites on glutamate carboxypeptidase II necessary for proteolytic activity, *Prot. Sci.* **13**, 1627-1635.

55. Yoshimasu MA, Tanaka T, Ahn J-K & Yada RY (2004) Effect of N-linked glycosylation on the aspartic proteinase porcine pepsin expressed from *Pichia pastoris*, *Glycobiology* **14**, 417-429.
56. Vieira DF, Watanabe L, Sant'ana CD, Marcussi S, Sampaio SV, Soares AM, Arni RK (2004) Purification and characterization of jararassin-I, a thrombin-like enzyme from *Bothrops jararaca* snake venom, *Acta Biochim Biophys Sin* **36**, 798-802.
57. Bortoleto RK, Murakami MT, Watanabe L, Soares AM, Arni RK (2002) Purification, characterization and crystallization of Jararacussin-I, a fibrinogen-clotting enzyme isolated from the venom of *Bothrops jararacussu*, *Toxicon* **40**, 1307-1312.
58. Watanabe L, Vieira DF, Bortoleto RK, Arni RK (2002) Crystallization of bothrombin, a fibrinogen-converting serine protease isolated from the venom of *Bothrops jararaca*, *Acta Crystal. Sec. D* **58**, 1036-1038.
59. Liu SJ, Huang QQ, Zhu XY, Teng MK, Niu LW (1999) Purification, characterization, crystallization and preliminary X-ray diffraction of acuthrombin-B, a thrombin-like enzyme from *Agkistrodon acutus* venom, *Acta Crystal. Sec. D* **55**, 1193-1197.
60. Parry MAA, Jacob U, Huber R, Wisner A, Bon C, Bode W The crystal structure of the novel snake venom plasminogen activator TSV-PA: a prototype structure for snake venom serine proteinases, *Structure* **6**, 1195-1206.
61. Silva-Jr FP & De Simone SG (2004) S1 subsite in snake venom thrombin-like enzymes: can S1 subsite lipophilicity be used to sort binding affinities of trypsin-like enzymes to small-molecule inhibitors, *Bioorg Med Chem* **12**, 2571-2487.

62. Svergun DI & Koch MJH (2003) Small-angle scattering studies of biological macromolecules in solution, *Rep Prog Phys* **66**, 1735-1782.
63. Caracciolo G, Amiconi G, Bencivenni L, Boumis G, Caminiti R, Finocchiaro E, Maras B, Paolinelli C, Congiu Castellano A (2001) Conformational study of proteins by SAXS and EDXD: the case of trypsin and trypsinogen, *Eur Biophys J Biophys Lett* **30**, 163-170.
64. Oliveira, CLP (2003) "TRAT1D - Computer Program for SAXS Data Treatment", LNLS technical Manual MT 01/2003.
65. Cavalcanti LP, Plivelic TS, Kellermann G, Neuenschwander R, Oliveira CLP & Torriani IL (2004) Two new sealed sample cells for small angle x-ray scattering from macromolecules in solution and complex fluids using synchrotron radiation, *Rev. Sci. Inst.* **75**, 4541-4546.
66. Svergun DI, Petoukhov MV & Koch MHJ (2001) Determination of domain structure of proteins from X-ray solution scattering, *Biophys J* **80**, 2946-2953.
67. Laemmli UK (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4, *Nature* **227**, 680 1970.
68. Campos ITN, Amino R, Sampaio CAM, Auerswald EA, Friedrich T, Lemaire HG, Schenkman S, Tanaka AS (2002) Infestin, a thrombin inhibitor presents in *Triatoma infestans* midgut, a Chagas' disease vector: gene cloning, expression and characterization of the inhibitor, *Insect Biochem. Mol. Biol.* **32**, 991-997.
69. Mello GC, Oliva MLV, Sumikawa JT, Machado OLT, Marangoni S, Novello JC, Macedo MLR (2001) Purification and characterization of a new trypsin inhibitor from *Dimorphandra mollis* seeds, *J. Prot. Chem.* **20**, 625-632.

TABLES

Table I: Purification of BJ-48 enzyme from *B. jararacussu* venom.

Purification Steps	Total Protein		Amidolytic activity		Yield (%)	Purification factor (n-fold)
	(mg)	(%)	Total (units)	Specific* (units/mg)		
Crude venom	36.1	100.0	0.21	0.043	100.0	1
Affinity column	0.93	2.6	0.16	0.172	76.2	29.6
HPLC gel filtration	0.80	2.2	0.16	0.202	76.2	34.9

*One unity of amidolytic activity stands for 1 μmol of BApNA hydrolyzed per minute, at 25 °C in Tris-HCl 50 mM, pH 8.0 buffer.

Table II- Values of steady state parameters for the BJ-48 and thrombin catalyzed hydrolysis of ester, amides and tripeptidyl anilides in 50 mM Tris-HCl, pH 8.0 at 25 °C.

Enzyme	Substrate	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m ($\text{s}^{-1} \cdot \text{M}^{-1}$)
BJ-48	BApNA	8.3×10^2	3.7×10^{-2}	4.5×10^1
	BAME	4.0×10^2	1.5	3.8×10^3
	TAME	6.0×10^2	4.5	7.6×10^3
	N α -Tos-Gly-Pro-ME	3.0×10^3	1.6×10^{-3}	0.53
	N α -Tos-Gly-Pro-Lys-pNA	Not detectable		
	N α -Tos-Gly-Pro-Arg-pNA	3.6×10^2	2.0	5.5×10^3
α -Thrombin	N α -Tos-Gly-Pro-Lys-pNA	4.0×10^1	2.2×10^1	5.4×10^5
	N α -Tos-Gly-Pro-Arg-pNA	4.2	4.8×10^1	1.1×10^7

Values for K_m were calculated from the linear regression data derived from the Lineweaver-Burk plot for each substrate. Data are presented as means (n=3) and standard errors are inferior to 5%. K_{cat} values were calculated assuming that 48 mg of protein represents 1 μmol of enzyme.

Table III: Effects of inhibitors on the BJ-48 ester hydrolysis activity.

6.1.1.a.1.1 Inhibitor	Concentration	Residual Activity (%)
Control	—	100
Aprotinin	0.3 μ M	53
DTT	1 mM	64
TLCK	0.1 mM	72
Benzamidine	1 mM	72
PMSF	1 mM	76
SBTI	0.2 μ M	90
Chymostatin	100 μ M	97
E-64	10 mM	98
6-aminocaproic acid	1.0 mM	100
TPCK	0.1 mM	100
EDTA	10 mM	100
Infestin ^a	0.5 nM	100
DMTI ^b	3.5 μ g/mL	100
Hyrudin	0.2 μ M	100
α 2-Macroglobulin	0.2 μ M	100

^aInfestin is a kazal-type thrombin inhibitor from midguts of *Triatoma infestans* [68].

^bDMTI is a Kunitz trypsin family inhibitor isolated from *D. mollis seeds* [69].

FIGURES

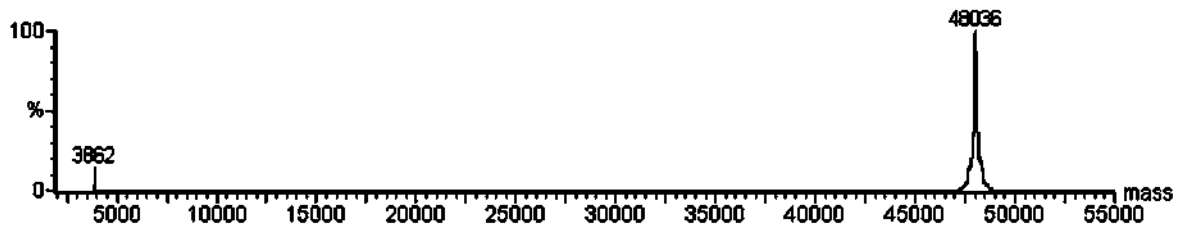


Fig. 1

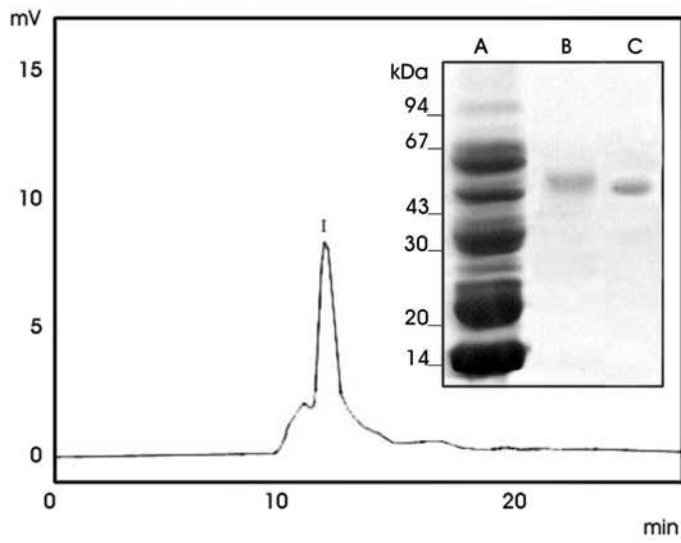


Fig. 2

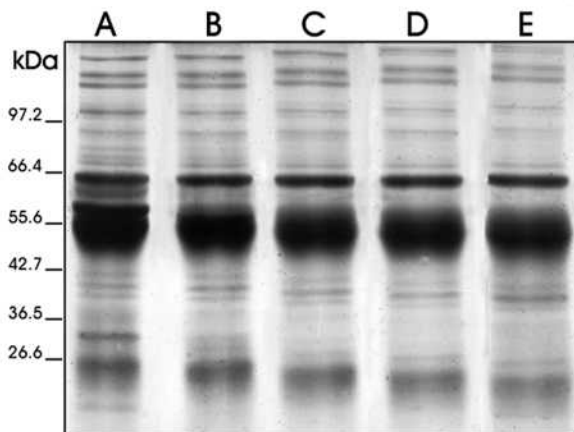
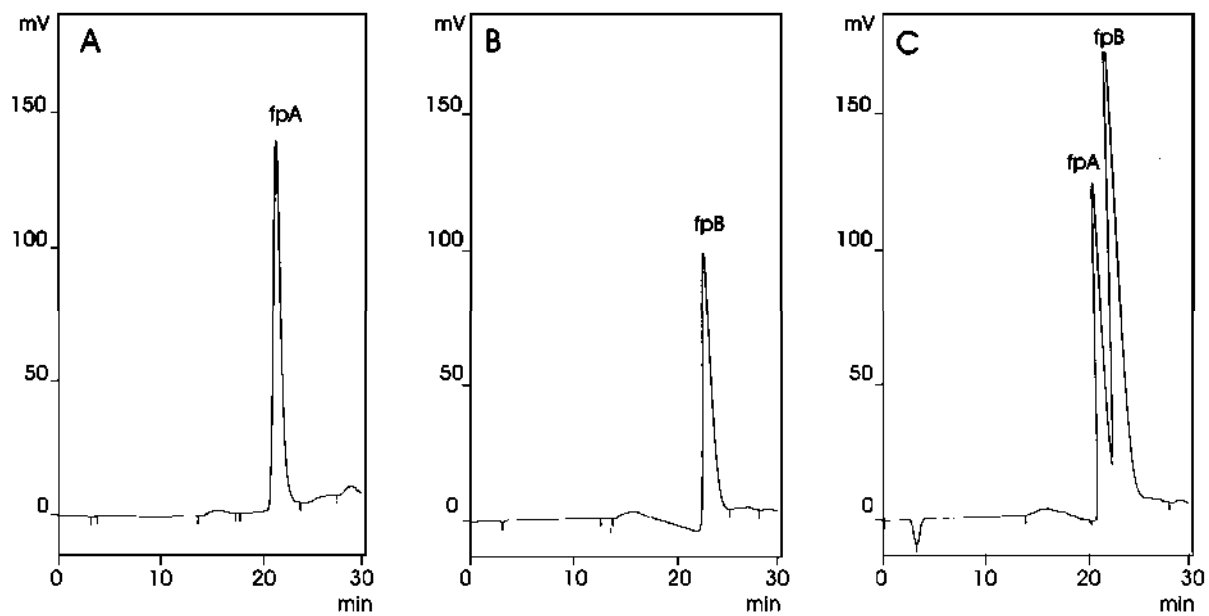


Fig. 3

**Fig. 4**

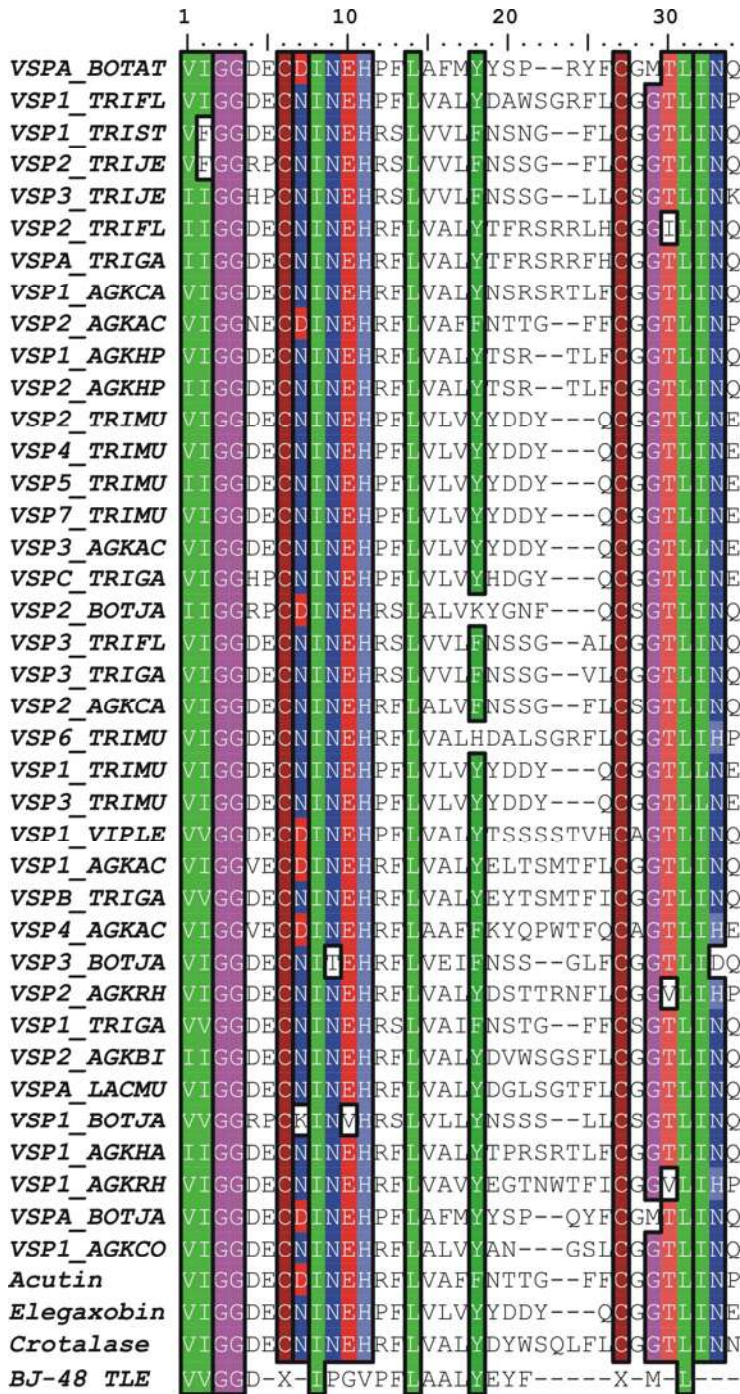


Fig. 5

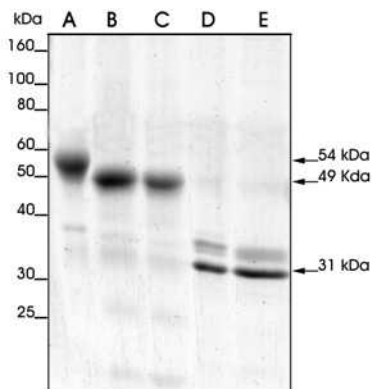
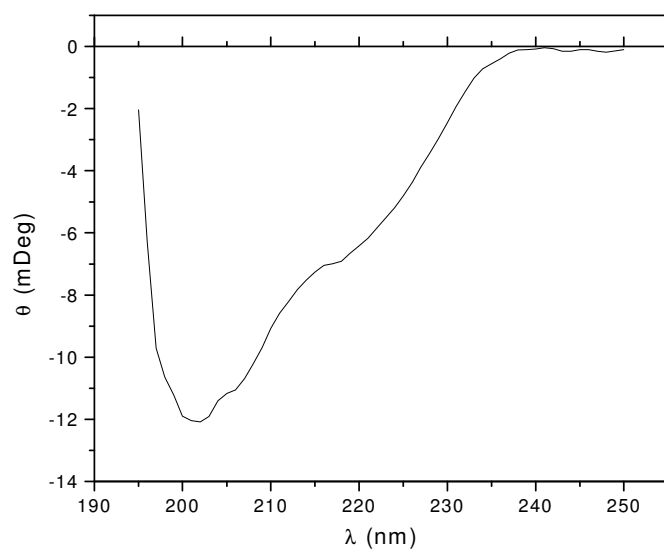


Fig. 6**Fig. 7**

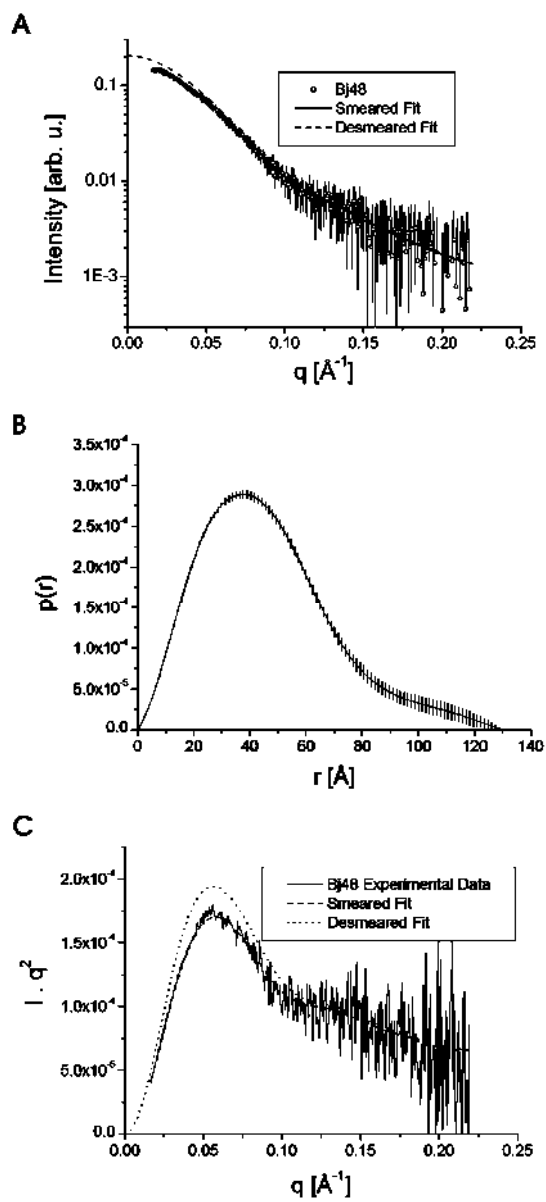


Fig. 8

LEGENDS TO FIGURES

Fig. 1- Transformed ESMS spectrum of HPLC-purified BJ-48 enzyme. For details, see text.

Fig. 2- HPLC gel filtration profile of benzamidine-affinity fraction of *Bothrops jararacussu* venom. HPLC was performed in a Shinpack Diol-150 column and the chromatogram recorded at 280 nm. The major peak (II) was collected, concentrated and examined by 12% SDS-PAGE (inset). Molecular weight standard proteins are indicated at left. Lane A- crude *B. jararacussu* venom (100 µg); lanes B and C, gel filtration peak II (5 µg) treated and non-treated with 2-ME, respectively.

Fig. 3- SDS-PAGE (10%) analysis of human fibrinogen degradation by BJ-48. Lane A, intact fibrinogen; lanes B-E, fibrinogen incubated with BJ-48 at 37 °C for 30, 60, 120 and 180 minutes, respectively.

Fig. 4- HPLC analysis of fibrinopeptides release upon fibrinogen hydrolysis by BJ-48. Fibrinopeptide release was analyzed by C18 reverse chromatography at 220 nm. (A) Human fibrinopeptide A (fpA) standard. (B) Human fibrinopeptide B (fpB) standard. (C) profile of reaction products obtained after human fibrinogen incubation with BJ-48 at 37 °C for 1h.

Fig. 5- Amino acid sequence alignment of BJ-48 and other SVSPs. Proteases are identified by their accession codes in Swiss-Prot or common names. The alignment is boxed and colored according to 70% conservation threshold and the BLOSUM62 matrix.

Fig. 6- Analysis of BJ-48 glycosylation type and content. The products of several glycosidases reactions with denatured BJ-48 were subjected to reducing 12% SDS-PAGE analysis. Lane A: control glycoprotein with no enzyme added; lanes B-E show the result overnight incubation (37°C) of glycoprotein with: neuraminidase (B), neuraminidase + O-glycosidase (C), PNGase F (D), PNGase F + O-glycosidase + α -2(3,6,8,9)-neuraminidase + β (1-4)-galactosidase + β -N-acetylglucosaminidase (E). Molecular weight standards are indicated at left. The M_r of the major BJ-48 deglycosylation products are indicated by arrows at right.

Fig. 7- CD spectrum of BJ-48. Spectrum was recorded at 25 °C for a 1 mM protein solution buffered at pH 7.5.

Fig. 8- SAXS analysis of BJ-48 structure at pH 8.0. (A) Experimental SAXS curve; (B) distance distribution function $p(r)$ and (C) *Kratky* plot. The curves obtained from experimental scattering data were calculated using the program GNOM [66].

ANEXO 6**Floriano Paes Silva Junior**

De: bj@biophysics.org
Enviado em: terça-feira, 11 de outubro de 2005 12:07
Para: Herbert Leonel de Matos Guedes; Floriano Paes Silva Júnior; Carlos Corrêa Netto; Cristiane Cardoso Martins de Salles; Cristiano Luis Pinto de Oliveira; Iris Torriani; Salvatore Giovanni De Simone
Assunto: BIOPHYSJ Acknowledgment of Manuscript BIOPHYSJ/2005/075663

MS ID#: BIOPHYSJ/2005/075663

MS TITLE: Structural characterization of BJ-48, a thrombin-like enzyme from Bothrops jararacussu venom: A STUDY BY SMALL ANGLE X-RAY SCATTERING (SAXS)

Dear Dr. De Simone:

Thank you for sending the above-mentioned manuscript, which has been entered into our online submission system (<http://submit.biophysics.org>).

The paper has now been assigned to an Associate Editor, who will then assign it to the appropriate Editor. Your paper will be reviewed entirely online. After review, you will be able to view the decision online by visiting our site.

Sincerely,

Dr. Robert Callender
Editor-in-Chief

Structural characterization of BJ-48, a thrombin-like enzyme from *Bothrops jararacussu* venom: A STUDY BY SMALL ANGLE X-RAY SCATTERING (SAXS)

Herbert L. M. Guedes^{1,2,‡}, Floriano P. Silva-Jr^{1,3,‡}, Carlos Correa Netto^{1,5}, Cristiane M. C. de Salles¹, Cristiano L. P. de Oliveira^{6,7}, Íris Torriani^{6,7}, Salvatore Giovanni De Simone^{1,4,*}.

¹ Laboratório de Bioquímica de Proteínas e Peptídeos, Departamento de Bioquímica e Biologia Molecular, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Av. Brasil, 4365, 21045-900 Rio de Janeiro, RJ, Brazil.

² Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.

³ Departamento de Química Orgânica, Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.

⁴ Departamento de Biologia Celular e Molecular, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Niterói, RJ, Brazil.

⁵ Instituto Vital Brazil, Niterói, RJ, Brazil.

⁶ Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Síncrotron, Campinas, Brazil.

⁷ Instituto de Física, Universidade Estadual de Campinas, Campinas, Brazil.

* Corresponding author. Fax: +55 21 25903495

E-mail address: dsimone@ioc.fiocruz.br (S. G. De-Simone)

‡These authors contributed equally to this work.

ABSTRACT Thrombin-like enzymes (TLEs) are important components of snake venoms due to their involvement in coagulopathies occurring on envenoming. Structural characterization of this group of serine proteases is of utmost importance for better understanding their unique properties. However, the high carbohydrate content of some members of this group prevents successful crystallization for structural determination. Circumventing this difficulty, the structure of BJ-48, a highly glycosylated TLE from *Bothrops jararacussu* venom, was studied in solution using small angle X-ray scattering (SAXS). BJ-48 presented a prolate, yet rigid shape at pH 8.0 while it assumed a flexible U-shape upon acidification to the natural venom pH at 6.0. BJ-48 tertiary structure at pH 6.0 was maintained on heating up to 52 °C and was completely lost at 70 °C. BJ-48 showed no modification on its shape after complex formation with the inhibitor benzamidine. The possible existence of two pH-induced folding states for BJ-48 was discussed.

Keywords: snake venom; serine protease; SVTLE; solution structure; pH-induced conformational changes; folding states; protein shape.

INTRODUCTION

The *Bothrops jararacussu* snake is found in tropical forests, swamps and river banks of south and southeast states of Brazil (1,2). The venom from *B. jararacussu* has lethal potency much higher than venoms of most other *Bothrops* species (3,4). The venom of these snakes has proteolytic, coagulant and hemorrhagic effects (5). Local envenoming is characterized by pain, swelling, bleeding, occasionally blisters, abscess formation and necrosis (6). Coagulation disorders occur in almost all fatal cases, and death has been related to acute renal failure, acute respiratory failure, hemorrhage, shock and sepsis (6).

Venom serine proteases affect prey hemostasis by issuing specific cleavages in factors involved in coagulation, fibrinolysis and in the kallikrein-kinin systems (7,8). Some of these venom proteases are often referred to as thrombin-like enzymes (SVTLEs) due to their ability to cleave fibrinogen, releasing fibrinopeptide A, fibrinopeptide B or both. Nevertheless, in most cases, these fibrin monomers are unable to be cross-linked by factor XIII, resulting in aberrant clots that are independent of plasmin activity for dissolution. The *in vivo* significance of this activity is dependent upon the plasmatic concentration of the enzyme but in common envenomations it usually causes an anticoagulant effect. Some of these SVTLEs also participate in diverse biological processes such as activation of plasminogen, protein C, factors V and X, platelet aggregation and releasing hypotensive Lys-bradykinin or bradykinin from kininogen (8,9).

SVTLEs show a high level of similarity within them and to a lesser degree with mammalian serine proteases such as trypsin and thrombin (10). SVTLEs and trypsin-like enzymes have similar primary structures and are believed to share a common ancestor (11). SVTLEs have been extensively studied, yet there is a lack of three dimensional structural information concerning this group of proteases. Homology models generated for several SVTLEs have fostered the belief that these proteins would display a globular structure similar to other members of the chymotrypsin family, e.g. (10). However, SVTLEs may show significant differences in the molecular shape mainly due to length and composition variations in surface loops (12) and to distinct carbohydrate content (8).

Recently, we have characterized a venom thrombin-like enzyme (BJ-48) from *B. jararacussu* (13) which is capable of cleaving either the α and β chain of fibrinogen but show preference for the β chain. The N-terminal is most similar to batroxobin, a TLE isolated from *Bothrops atrox* venom. Moreover, BJ-48 shows distinct physical-chemical properties from trypsin and thrombin. For instance, it has greater stability to low pH and heating than its mammalian counterparts (14,15). During the purification procedure, BJ-48 can be eluted from the benzamidine-agarose affinity column using acid pH and a rapid increase of pH to prevent loss of enzymatic activity is not necessary as described for thrombin (15). The optimum temperature for BJ-48 amidase activity is 50-55°C and it presents activity until 70°C.

Structural information on BJ-48 is essential to further characterize this protein and better comprehend the relationship between fold and enzymatic function. However, BJ-48 is extensively glycosylated, which poses difficulties to its structural characterization by crystallography. In this context, small angle X-ray scattering (SAXS) may help us to solve the molecular envelope of BJ-48 in solution. Previously, we have deduced a prolate shape from BJ-48 SAXS curves at pH 8.0 (13). This conformation is strikingly different from the globular crystallographic structures of thrombin and trypsin.

In the present work we report the effect of pH, temperature and benzamidine on the molecular envelope of BJ-48 as determined by small-angle X-ray scattering (SAXS)

studies. At first, SAXS data were collected at the natural pH of *B. jararacussu* venom (pH 6.0) and compared to the data obtained at pH 8.0, as described before (13). In addition, we have evaluated the effect of heating (from 10 °C to 75 °C) on the structure of the enzyme at pH 6.0 and the effect of benzamidine (a reversible active-site inhibitor) on BJ-48 shape at pH 8.0. *Ab initio* models of BJ-48 were constructed from SAXS data at pH 6.0 and 8.0.

MATERIALS AND METHODS

***B. jararacussu* venom and pH determination**

B. jararacussu venom was extracted manually from ten specimens kept at the Vital Brazil Institute and the pH value determined using a Corning pH/Ion analyzer 350.

Benzamidine-agarose affinity column

BJ-48 was purified as described before (16). Briefly, the venom (100 mg) was equilibrated in 50 mM Tris-HCl containing 0.5 M NaCl, pH 7.0 and subjected to affinity chromatography on a benzamidine-agarose column. The adsorbed proteins were eluted with 1 mM HCl (pH 3.0), containing 0.5 M NaCl. The eluant was collected on ice and immediately concentrated using P10 centrifugal filters.

Gel filtration high performance liquid chromatography

The concentrated post-benzamidine-agarose preparation was injected in HPLC (Shin pack Diol-150 column-Shimadzu) previously equilibrated in 50 mM phosphate buffer pH 7.2. The protein was fractionated on an automatic HPLC system (Shimadzu, 6A model) at a flow rate of 1 mL/min, during 28 min at 25°C. The fraction containing BJ-48 was dialyzed against 50 mM phosphate buffer pH 6.0 or 50 mM Tris-HCl pH 8.0 and then concentrated using P10 centrifugal filters until protein concentration of 4 mg/mL. For molecular mass characterization the column was calibrated in the same buffer with the following markers: β -galactosidase (105 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa) and carbonic anhydrase (29 kDa).

Protein determination

Protein content of samples was determined by the Bio-Rad protein assay or from the absorbance at 280 nm. Bovine serum albumin was used as standard.

X-ray scattering experiments, equipment set up and data analysis

Small angle X-ray scattering experiments were performed at the SAXS beamline of the National Synchrotron Laboratory, Campinas, Brazil (17). The experimental setup included a temperature-controlled, 1.5 mm diameter capillary tube sample holder and a linear position-sensitive detector. Data acquisition was performed by taking five 600 s frames, using a sample detector distance of 446 mm and 849 mm and X-ray wavelength of 1.488 Å which enabled detection of a q range ($q=(4\pi/\lambda)\sin(\theta)$, λ = wavelength and 2θ = scattering angle) equal to $0.0165 \text{ \AA}^{-1} < q < 0.2192 \text{ \AA}^{-1}$. All the data treatment of the scattering intensities was performed using the software package TRAT1D (18). Usual correction for detector homogeneity, incident beam intensity, sample absorption, blank subtraction and intensities averaging were included in this routine. The output of this software provides the corrected experimental intensities and error values. Data analysis and model calculations were performed using the computer programs GNOM (19) and GASBOR (20).

Effect of pH, temperature and benzamidine on BJ-48 structure by SAXS

The BJ-48 measurements were performed at a protein concentration of 4 mg/ml in two different pH values: 50 mM Tris-HCl buffer pH 8.0 and 50 mM phosphate buffer pH 6.0 at 37°C. Measurements of denatured BJ-48 (devoid of enzymatic activity) at pH 6.0 and 7.5 were also performed for comparison with SAXS curves for native BJ-48 samples at pH 6.0 and 8.0. In order to determine the effect of temperature on BJ-48 structure at pH 6.0, the sample was heated from 10 °C to 75 °C (10 °C, 25 °C, 32 °C, 37 °C, 42 °C, 47 °C, 52 °C and 75 °C). After the heating, the sample was immediately cooled to 25°C and a new SAXS data set was collected. The effect of benzamidine on BJ-48 structure at pH 8.0 was performed by previously incubating the protein (4 mg/mL) with 2 mM of this inhibitor. In this case, the blank was made with the inhibitor at the same concentration.

Determination of molecular mass by SAXS

The equivalent molecular weight of BJ-48 protein was determined using a standard sample of 5 mg/ml of ovoalbumin (42.2 kDa) measured in the same experimental conditions. Molecular masses of BJ-48 in different conditions studied were obtained from the quotient between their intensity extrapolated to $q = 0$ and the $I_{(0)}$ value corresponding to ovoalbumin.

RESULTS

Effect of pH

The pH determined for the crude *B. jararacussu* venom was 6.0. Consequently, SAXS data by BJ-48 was measured at pH 6.0 using synchrotron radiation. The superposition of experimental (corrected and normalized) SAXS curves for BJ-48 in pH 8.0 (13) and 6.0 are shown in Figure 1A. The values of the radius of gyration (R_g) and maximum dimension (D_{max}) were determined from the pair distribution function ($p(r)$) using the program GNOM (Fig. 1B). For pH 8.0, it was obtained $R_g = 37 \pm 1 \text{ \AA}$ and $D_{max} = 130 \text{ \AA}$, and for pH 6.0 the values derived were $R_g = 27.9 \pm 0.5 \text{ \AA}$ and $D_{max} = 82 \text{ \AA}$. The R_g and the D_{max} suggest a prolate elongated shape for pH 8.0. At pH 6.0, the pair distribution function indicates a non-globular protein as the curve does not have a single maximum. The Kratky plot ($I \cdot q^2 \times q$) indicates that BJ-48 presents a compact form at pH 8.0 and possesses higher flexibility at pH 6.0 (Fig. 1C). SAXS experiments for denatured BJ-48 (without enzymatic activity) demonstrated a clearly different scattering pattern (results not shown). Noteworthy, the Kratky plot shows an ascending curve that undoubtedly characterizes a random coil or non-structured protein (Fig. 2). It is worth saying that in both pH values BJ-48 is soluble and aggregation is not observed.

Molecular mass of BJ-48

The molecular masses of BJ-48 in pH 8.0 and 6.0 were $50 \pm 3 \text{ kDa}$ and $55 \pm 3 \text{ kDa}$, respectively, these values are in good agreement with those determined using SDS-PAGE and MALDI-TOF (13). In both cases, BJ-48 was monomeric as other members of serine proteases family. However, the molecular mass of the BJ-48 determined by gel filtration chromatography at pH 7.2 is 90 kDa (13), suggesting a prolate shape for the protein. The later observation is in accordance with our SAXS measurements at pH 8.0.

Effect of temperature on BJ-48 structure

The structural changes of BJ-48 enzyme induced by heating were accompanied by SAXS experiments. Our experiments conducted at pH 6.0 showed that in temperatures

lower than 52 °C BJ-48 preserves its shape and R_g . However, at 75 °C it was observed an increase in R_g and flexibility. The lowering of the temperature back to 25 °C also induced a significant increase in R_g and flexibility. Therefore, this observation suggests that above 75 °C an irreversible denaturation of the protein is induced as the reduction of the temperature did not promote the refolding of the protein (**Fig. 3**).

Effect of benzamidine on BJ-48 structure

Using SAXS experiments, we evaluated the potential of benzamidine binding to alter the structure of BJ-48. The scattering pattern and the resulting $p(r)$ function for BJ-48 at pH 8.0 in the presence of 2 mM benzamidine are essentially unaltered upon comparison with the curves for free BJ-48 in solution buffered at the same pH (**Fig. 4A** and **4B**). This result indicates that formation of a complex with benzamidine produces no significant change in the shape of BJ-48 enzyme.

Molecular envelope of BJ- 48

The low resolution of BJ-48 molecular shape in pH 8.0 and 6.0 were determined using GASBOR program. We have calculated five independent models for each pH and the result is shown in Figure 5. In agreement with the $p(r)$ function, the BJ-48 model at pH 8.0 has a prolate shape (**Fig. 5A**). For pH 6.0, the generated *ab initio* model indicates that the protein has a U-shape in this condition (**Fig. 5B**). These results suggest that two different conformational states for BJ-48 are induced by pH variation. The prolate form (**Fig. 5A**) is most stable in alkaline pH while the U-shape form (**Fig. 5B**) predominates in acid pH and is associated with a lower activity.

DISCUSSION

BJ-48 can be purified from *B. jararacussu* venom with high yield in only two chromatographic steps (13,16). Nevertheless, its high level of glycosylation (N-linked carbohydrates represent 42% of the total molecular mass) has until now rendered crystallization trials unsuccessful. Therefore, we have applied SAXS experiments to help shed some light on the structural features of this SVTLE.

The present study indicates possible significant structural differences for this serine protease in relation to other members of the chymotrypsin family. All known crystallographic structures of the members of this family are globular (e.g. thrombin (21)). As demonstrated in this study, BJ-48 has a compact, yet non-globular shape at pH 8.0 and a flexible U-shape in pH 6.0 (see **Fig. 1** and **Fig. 2**). Indeed, such increase in flexibility observed at pH 6.0 is possibly associated with some loss of tertiary structure as it is accompanied with reduction of proteolytic activity at this pH (13).

BJ-48 is fully active at elevated temperatures (> 37 °C) with the optimum activity displayed between 50-55°C. A complete loss of enzymatic activity at pH 7.5 occurs only at temperatures above 70°C (13). These biochemical properties were completely supported by the SAXS data obtained in the present work showing that the tertiary structure of this enzyme is retained at temperatures below 52 °C and is only completely lost at 75 °C (**Fig. 3**).

Benzamidine is a reversible inhibitor of trypsin and thrombin. We have previously determined that BJ-48 is inhibited by benzamidine with a K_i of 111 μ M (16). In agreement with crystallographic studies with other trypsin-like enzymes (22), the SAXS data presented in Figure 4 shows that benzamidine binding does not induce modification in the tertiary structure of BJ-48.

BJ- 48 presents two pH-induced conformational states

The structures of globular proteins can be classified as native, molten globule, premolten globule and unfolded (random coil) conformations (folding states), according to their secondary structure content and overall flexibility (23). The molten globule state shows the typical structure of native proteins; however the premolten globule and unfolded conformations do not have rigid tertiary structure. There are several examples in the literature of natively unfolded and intermediate folded proteins. For instance, the 50S and 30S ribosomal proteins are unfolded in solution (24), but transform into a rigid well-folded conformation in the functional ribosome (23). Using SAXS radiation, the different folding intermediates can be differentiated by the observed increase in hydrodynamic radius and accompanying decrease in the degree of compactness of the protein. Moreover, the Kratky plot of a globular protein has a characteristic maximum for the native and molten globule states that cannot be observed for the premolten globule and unfolded conformations (23). For example, alkaline denaturation of aspergillopepsin II, a non-pepsin-type acid protease occurs at pH 6.4 as observed by SAXS and pass through folding intermediate state (25).

The BJ-48 solution structure in pH 6.0 may suggest that the protein is in an intermediate folded state due to the increase in the protein flexibility when compared with pH 8.0, yet it is not accompanied by an increase in R_g as expected. Unfortunately the BJ-48 shows a complex CD spectrum at pH 7.5 (13), therefore it is not possible to use the secondary structure content to characterize this state as molten globule, premolten globule or other state. Nevertheless, it can be assured that the folding state for BJ-48 at pH 6.0 is completely different from the unfolded state observed in BJ-48 devoid of enzymatic activity (denatured). Kratky plots of denatured BJ-48 in solutions buffered at pH 6.0 or pH 7.5 lacked a characteristic maximum, showing much higher flexibility than the catalytic competent BJ-48 at pH 6.0. Moreover, the thermo stability of BJ-48 at pH 6.0 observed by SAXS is similar to the thermo stability of BJ-48 at pH 7.5 in enzymatic studies. This fact indicates that BJ-48 folding intermediate maintained the thermo stability displayed by the native protein.

In SAXS experiments, bovine trypsin shows the same R_g at acidic (≤ 5.2) and neutral-to-basic pH (≥ 7.0), but presents a peak in R_g at pH 6.2 (26). In accordance with our results, these data could possibly indicate the presence of an intermediate folding state for trypsin in this pH. Nevertheless, differences do exist between trypsin and BJ-48 intermediates. The Kratky plot for trypsin in pH 6.0 shows a more compact structure when compared with BJ-48 at the same pH. Moreover, trypsin has its R_g increased while BJ-48 presents a decrease in R_g when moving from basic to acidic pH. Giving the similarity between the two proteases in the amino acid sequence level, these results are unexpected. Whether such discrepancies in structural modifications induced by acid denaturation for trypsin and BJ-48 are related to the high carbohydrate content of the later enzyme is a question that can be addressed by future SAXS experiments with deglycosylated BJ-48.

The detection of two different states induced by pH variation (basic and acid pH) using fluorescent spectroscopy was also observed for acutolysin D, a metallo- protease from *Agkistrodon acutus* venom (27). This finding may suggest a general mechanism of regulation of proteolytic activity of venom enzymes. Possibly, in the slightly acid pH of venom (pH 6.0 as in *B. jararacussu* species) the proteases are partially folded with low activity, which could serve to avoid protein degradation in the snake gland. After snake bite, proteases are inoculated in the prey's blood circulation and exposed to higher pH (i.e., pH 7.4). This would trigger the transition from a less active folding intermediate

(or nearly unfolded state) to the fully catalytically competent native conformation would occur by a pH-induced conformational change.

CONCLUSION

The present study demonstrated particular and interesting structural modifications occurring with BJ-48 thrombin-like enzyme on pH change. We must state that some of these particularities revealed by SAXS can be the result of the extensive N-glycosylation on the enzyme molecular surface but experiments with BJ-48 at the venom pH have clearly given structural information about a novel folding state for the protein in solution. From the analysis, we conclude that a U-shape characterizes the less active enzyme in pH 6 while keeping the thermo stability of the fully competent enzyme with a prolate shape in pH 7.5-8.0. We propose that such behavior for BJ-48 enzyme in solution could be the result of a regulatory mechanism that has evolved in snake venom glands to control undesired proteolysis. Besides, we could not detect any conformational change in solution shape of BJ-48 upon binding to benzamidine and showed that the high thermostability of this enzyme is given by the resistance of its tertiary structure to high temperatures.

This work was supported by Fundação Oswaldo Cruz, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro for financial support. We are also grateful to the staff of the SAXS beam line at LNLS. H.L.M. Guedes and F.P. Silva-Jr were a CNPq Dsc. fellowship recipient.

REFERENCES

1. de Lema, T., M. E. Fabián-Beurmann, and M. L. Araújo. 1980. Lista de répteis encontrados na região da grande Porto Alegre, Estado do Rio Grande do Sul, Brasil. *Iheringia Ser. Zool.* 55: 27–36.
2. Campbell, J. A., and W. W. Lamar. 1989. *The Venomous Reptiles of Latin America*. Cornell Univ. Press, Ithaca.
3. Schottler, W. H. 1951. Toxicity of the principal snake venoms of Brazil. *Am. J. Trop. Méd. Hyg.* 31: 489-99.
4. [Sanchez, E. F., T. V. Freitas, D. L. Ferreira-Alves, D. T. Velarde, M. R. Diniz, M. N. Cordeiro, G. Agostini-Cotta, and C. R. Diniz.](#) 1992. Biological activities of venoms from South American snakes. *Toxicon* 30: 95-103.
5. Bucarechi, F., [S. H. Herrera, S. Hyslop, E. C. Baracat, and R. J. Vieira.](#) 2001. Snakebites by *Bothrops spp* in children in Campinas, São Paulo, Brazil. *Rev. Inst. Med. Trop. São Paulo* 43: 329-33.
6. Milani Junior, R., [M. T. Jorge, F. P. de Campos, F. P. Martins, A. Bousso, J. L. Cardoso, L. A. Ribeiro, H. W. Fan, F. O. Franca, I. S. Sano-Martins, D. Cardoso, C. I. de Fernandez, J. C. Fernandes, V. L. Aldred, M. P. Sandoval, G. Puerto, R. D. Theakston, and D. A. Warrell.](#) 1997. Snake bites by the jararacucu (*Bothrops jararacussu*): clinicopathological studies of 29 proven cases in Sao Paulo State, Brazil. *Q. J. Med.* 90:323-334.
7. [Lu, Q., J. M. Clemetson, and K. J. Clemetson.](#) 2005. Snake venoms and hemostasis. *J. Thromb. Haemost.* 3: 1791-1799.
8. [Serrano, S.M. and R. C. Maroun.](#) 2005 Snake venom serine proteinases: sequence homology vs. substrate specificity, a paradox to be solved. *Toxicon* 45: 1115-1132.

9. [Matsui, T., Y. Fujimura, and K. Titani.](#) 2000. Snake venom proteases affecting hemostasis and thrombosis. *Biochim. Biophys. Acta.* 1377:136-156.
10. [Silva-Jr, F.P., and S. G. De-Simone.](#) 2004. S1 subsite in snake venom thrombin-like enzymes: can S1 subsite lipophilicity be used to sort binding affinities of trypsin-like enzymes to small-molecule inhibitors? *Bioorg. Med. Chem.* 12:2571-2587. Erratum in: *Bioorg. Med. Chem.* 12: 3755.
11. [Lesk, A.M., and W. D. Fordham,](#) 1996. Conservation and variability in the structures of serine proteinases of the chymotrypsin family. *J. Mol. Biol.* 258: 501-537.
12. Castro, H.C., [D. M. Silva, C. Craik, and R. B. Zingali.](#) 2001. Structural features of a snake venom thrombin-like enzyme: thrombin and trypsin on a single catalytic platform? *Biochim. Biophys. Acta.* 1547: 183-95.
13. Silva-Jr, F.P., S. C. Bourguignon, A. R. Melgarejo, C. L. P. Oliveira, I. L. Torriani, S. G. De-Simone. 2005. BJ-48, a highly glycosylated clotting serine protease from the *Bothrops jararacussu* venom: Biochemical characterization and solution structure analysis. *FEBS Journal.* Submitted.
14. Outzen, H., I. G. Berglund, A. O. Smalas, and N. P. Willassen. 1996. Temperature and pH sensitivity of trypsins from atlantic salmon (*Salmo salar*) in comparison with bovine and porcine trypsin. *Comp. Biochem. Physiol.* 115:33-45.
15. Nakamura, K., T. Suzuki, M. Hasegawa, Y. Kato, H. Sasaki, and K. Inouye. 2003. Characterization of p-aminobenzamidine-based sorbent and its use for high-performance affinity chromatography of trypsin-like proteases. *J. Chromatogr. A* 1009:133-139.
16. [De-Simone, S.G., C. Correa-Netto, O. A. C. Antunes, R. B. De-Alencastro, and F. P. Silva-Jr.](#) 2005. Biochemical and molecular modeling analysis of the ability of two p-aminobenzamidine-based sorbents to selectively purify serine proteases (fibrinogenases) from snake venoms. *J. Chromatogr. B.* 822:1-9.
17. Kellermann, G., F. Vicentin, E. Tamura, M. Rocha, H. Tolentino, A. Barbosa, A. Craievitch, and I. L. Torriani. 1997. The Small-Angle X-ray Scattering Beamline of the Brazilian Synchrotron Light Laboratory. *J. Appl. Crystallogr.* 30:880–883.
18. Oliveira, C. L. P., D. R. dos Santos, G. Kellermann, T. Plivelic, and I. L. Torriani. 1997. Technical Note, LNLS, Campinas, Brazil.
19. Semenyuk, V., and D. I. Svergun. 1991. GNOM—a program package for small-angle scattering data processing. *J. Appl. Crystallogr.* 24:537–540.
20. Svergun, D.I., M. V. Petoukhov, and M. H. Koch. 2001. Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* 80:2946-2953.
21. Bode, V., V. Mayr, U. Baumann, R. Huber, S. R. Stone, and J. Hofsteenge. 1989. The refined 1.9 Å crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* 8: 3467-3475.
22. Bode, W., and P. Schwager. 1975. The refined crystal structure of bovine beta-trypsin at 1.8 Å resolution. II. Crystallographic refinement, calcium binding site, benzamidine binding site and active site at pH 7.0. *J. Mol. Biol.* 98:693-717.
23. [Uversky, V.N. 2002.](#) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* 11: 739-756.
24. [Yusupov, M. M., G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. Cate, and H. F. Noller.](#) 2001. Crystal structure of the ribosome at 5.5Å resolution. *Science* 292:883-96.
25. [Kojima, M., M. Tanokura, K. Maeda, Y. Kimura, H. Amemiya, and K. Takahashi.](#) 2000. pH-dependent unfolding of aspergillopepsin II studied by small-angle X-ray scattering. *Biochemistry* 39:1364-1372.

26. [Caracciolo, G., G. Amiconi, L. Bencivenni, G. Boumis, R. Caminiti, E. Finocchiaro, B. Maras, C. Paolinelli, and A. Congiu Castellano.](#) 2000. Conformational study of proteins by SAXS and EDXD: the case of trypsin and trypsinogen. *Eur. Biophys. J.* 30:163-170.
27. [Xu X.L., X. H. Liu, B. Wu, Y. Liu, W. Q. Liu, Y. S. Xie, and Liu Q. L.](#)http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=15211502&query_hl=18 2004. Metal-ion- and pH-induced conformational changes of acutolysin D from *Agkistrodon acutus* venom probed by fluorescent spectroscopy. *Biopolymers* 74:336-344.

Legends to Figures

FIGURE 1 SAXS analysis of BJ-48 structure at pH 8.0 and 6.0. (A) Small-angle scattering pattern. (B) Pair distribution function $[p(r)]$. (C) Kratky plot. The $p(r)$ function and fitted curves were calculated from experimental scattering data using the program GNOM (19). For details, see text.

FIGURE 2 Effect of pH in BJ-48 compactness. Kratky plots of native BJ-48 at pH 6.0 and pH 8.0. The curves for denatured BJ-48 (without enzymatic activity) at pH 6.0* and pH 7.5* are plotted for reference.

FIGURE 3: Thermo stability of BJ-48 structure. The values for radius of gyration of BJ-48 at pH 6.0 are plotted for varying temperatures at which SAXS data were collected. For details, see text.

FIGURE 4: Effect of benzamidine binding in BJ-48 structure at pH 8.0. (A) Small-angle scattering pattern in the presence and absence of benzamidine. (B) Pair distribution function for BJ-48 in the presence and absence of benzamidine.

FIGURE 5: Molecular envelopes of BJ-48 at pH 8.0 and 6.0. (A) BJ-48 at pH 8.0 and (B) BJ-48 at pH 6.0. *Ab initio* models represent an average of five dummy residue models (DRM) calculations using GASBOR (20). The light gray spheres in (B) represent the space accessible for the DRM as judged by the degree of conservation of a given dummy residue among the five models.

Curriculum Vitae

Floriano Paes Silva-Jr (Silva-Jr, F.P.)

Nascido em 10 de Fevereiro de 1978, no Rio de Janeiro, RJ, Brasil.

Formação acadêmica

1- Doutor em Química Orgânica, Instituto de Química, UFRJ, Rio de Janeiro, RJ, Brasil.

Título da Tese: "ESTUDOS BIOQUÍMICO-COMPUTACIONAIS EM ASPARTIL PROTEASES CATEPSINA D-SÍMILE DE SCHISTOSOMA MANSONI E EM SERINIL PROTEASES ATIVAS SOBRE A COAGULAÇÃO SANGUÍNEA", sob a orientação dos Prof. Dr. Ricardo Bicca de Alencastro e Salvatore Giovanni De Simone, desenvolvida nas seguintes dependências: Departamento de Bioquímica e Biologia Molecular, IOC/FIOCRUZ; Departamento de Química Orgânica, IQ/UFRJ e Centro de Biologia Molecular Estrutural do Laboratório Nacional de Luz Síncrotron CeBiME/LNLS.

Período: Junho de 2002 a Dezembro de 2005.

Auxílio Financeiro: Bolsa de Doutorado do CNPq.

2- Farmacêutico Industrial, Faculdade de Farmácia, UFF, Niterói, RJ, Brasil.

Período: Agosto de 1997 a Fevereiro de 2002.

Outras Atividades: Bolsista de Iniciação Científica do programa PIBIC/CNPq no Laboratório de Bioquímica de Proteínas e Peptídeos, Depto. de Bioquímica e Biologia Molecular, IOC/FIOCRUZ, sob a orientação do Prof. Dr. Salvatore Giovanni de Simone; Bolsista do Programa de Bolsas de Verão LNLS/MCT no Centro de Biologia Molecular Estrutural do Laboratório Nacional de Luz Síncrotron CeBiME/LNLS e Bolsista de Iniciação Científica do programa PIBIC/CNPq, no Depto. de Biologia Celular e Molecular, IB/CEG/UFF, sob a orientação do Prof. Dr. Saulo Cabral Bourguignon.

Áreas de interesse

Bioquímica

Química de Macromoléculas – Proteínas

Enzimologia

Química Orgânica

Físico-química orgânica

Estrutura, conformação e estereoquímica

Farmácia

Publicações Científicas (2002-2005)

Artigos completos em periódicos indexados

SILVA-JR, F.P.; ALEXANDRE, G.; C.M.C.; OLIVEIRA, C.L.P.; TORRIANE, I.L.; DE SIMONE, S.G. Solution structure characterization of LMR-47, a thrombin-like enzyme from *Lachesis muta rhombeata* venom (em preparação).

SILVA-JR, F.P.; GARVEY, L.; DI CERA, E.; DE SIMONE, S.G. Further functional characterization of BJ-48 and LMR-47, two snake venom thrombin-like enzymes (em preparação).

SILVA-JR, F.P.; GUEDES, H.L.M.; QUIRINO, F.; DE SIMONE, S.G. Role of the extensive glycosylation in the structure and function of snake venom serine proteases (em preparação).

SILVA-JR, F.P.; ALCON, M.E.T.A.; PIERCE, R.; DE SIMONE, S.G. SmCD2, a novel member of the *Schistosoma mansoni* aspartyl protease (SmAP) gene family (em preparação).

SILVA-JR, F.P.; AGUIAR, A.S.; MELGAREJO, A.R.; BOURGUIGNON, S.C.; OLIVEIRA, C.P.; TORRIANE, I.L.; DE SIMONE, S.G. BJ-48 TLE, a highly glycosylated thrombin-like enzyme from the *Bothrops jararacussu* venom, BIOCHEMICAL CHARACTERIZATION AND SOLUTION STRUCTURE ANALYSIS. *FEBS J.* (submetido).

GUEDES, H.L.M.; **SILVA-JR, F.P.**; CORREA NETTO, C.; SALLES, C.M.C.; OLIVEIRA, C.L.P.; TORRIANE, I.L.; DE SIMONE, S.G. Structural characterization of BJ-48, a thrombin-like enzyme from *Bothrops jararacussu* venom: A STUDY BY SMALL ANGLE X-RAY SCATTERING (SAXS). *Biophys. J.* (submetido).

DE SIMONE, S. G.; CORREA NETTO, C.; **SILVA-JR, F.P.** A Simple and fast affinity chromatography procedure to purify beta-galactoside binding lectins. *J. Chromatogr. B* (aceito em 21/11/2005).

SILVA-JR, F.P.; ELENNA, J.; FERREIRA, M.L.; MASCARENHAS, Y.P; SOUZA, M.V.N.V.; Thatyana, R.A.; WARDELL, J.L.; WARDELL, S.M.S.V. Experimental and Theoretical Structure Characterization of Two Isoniazid Derivatives. *J. Mol. Struct.* (aceito em 16/11/2005).

SILVA-JR, F.P.; ANTUNES, O.A.C.; ALENCASTRO, R.B.; DE SIMONE, S.G. The Na⁺ binding channel of human coagulation proteases: Novel insights on the structure and allosteric modulation revealed by molecular surface analysis. *Biophys. Chem.*, v. 119, p. 280-292 (2005).

DE SIMONE, S.G.; CORREA NETTO, C.; ANTUNES, O.A.C.; ALENCASTRO, R.B.; **SILVA-JR, F.P.** Biochemical and molecular modeling analysis of the ability of two p-aminobenzamidine-based sorbents to selectively purify serine proteases (fibrinogenases) from snake venoms. *J. Chromatogr. B*, v. 822, p. 1-9 (2005).

SILVA-JR, F.P.; DE SIMONE, S.G. S1 subsite in snake venom thrombin-like enzymes: can S1 subsite lipophilicity be used to sort binding affinities of trypsin-like enzymes to small-molecule inhibitors?. *Bioorg. Med. Chem.*, v. 12, p. 2571-2587 (2004).

SILVA-JR, F.P.; RIBEIRO, F.; KATZ, N.; DE SIMONE, S.G. Exploring the subsite specificity of *Schistosoma mansoni* aspartyl hemoglobinase through comparative molecular modeling. *Febs Lett.*, v. 514, p. 141-148 (2002).

SILVA-JR, F.P.; VEYL, F.Z.; CLOS, J.; DE SIMONE, S.G. Molecular modeling approaches for determining gene function: application to a putative poly-A binding protein from *Leishmania amazonensis* (LaPABP). *Mem. Inst. Oswaldo Cruz*, v. 97, p. 335-341 (2002).

Capítulos de livros

SILVA-JR, F.P. Capítulo 6: Bioinformática e Modelagem Molecular Aplicada à Predição e Análise da Estrutura de Proteínas. In: DE SIMONE, S.G. (Org.). *A Arte da Purificação e Caracterização da Estrutura de Proteínas e Peptídeos*. São Paulo, 2005 (em avaliação).

SILVA-JR, F.P.; DE SIMONE, S.G. Capítulo 2: Fundamentos de Estrutura de Proteínas. In: DE SIMONE, S.G. (Org.). *A Arte da Purificação e Caracterização da Estrutura de Proteínas e Peptídeos*. São Paulo, 2005, (em avaliação).

DE SIMONE, S.G.; **SILVA-JR, F.P.** Proteases e Metabolismo de Aminoácidos na Quimioterapia. In: BORÉM, A.; SANTOS, F.R.; ALMEIDA, M.R. (Org.). *Biotechnologia de A a Z*. Viçosa-MG, 2003, v. -, p. 77-93.

Resumos em Anais e apresentações em reuniões científicas

Apresentou 29 trabalhos em Congressos e reuniões científicas, sendo 22 em forma de painel e 7 em forma oral. Publicou 3 trabalhos de caráter científico em anais científicos.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)