

Laboratório Nacional de Computação Científica  
Programa de Pós Graduação em Modelagem Computacional

**ANÁLISE DO IMPACTO DO POLIMORFISMO  
GENÉTICO DO SUBTIPO C DO HIV-1 NA  
INTERAÇÃO DA PROTEASE VIRAL COM O  
INIBIDOR NELFINAVIR POR MODELAGEM E  
DINÂMICA MOLECULAR**

Por

**Rosemberg de Oliveira Soares**

PETRÓPOLIS, RJ - BRASIL

JANEIRO DE 2009

# **Livros Grátis**

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

Milhares de livros grátis para download.

**ANÁLISE DO IMPACTO DO POLIMORFISMO GENÉTICO DO  
SUBTIPO C DO HIV-1 NA INTERAÇÃO DA PROTEASE VIRAL  
COM O INIBIDOR NELFINAVIR POR MODELAGEM E  
DINÂMICA MOLECULAR**

**Rosemberg de Oliveira Soares**

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO LABORATÓRIO  
NACIONAL DE COMPUTAÇÃO CIENTÍFICA COMO PARTE DOS REQUISI-  
TOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM  
MODELAGEM COMPUTACIONAL COM ENFÂSE EM BIOINFORMÁTICA E  
BIOLOGIA COMPUTACIONAL

Aprovada por:

---

Prof. Marcelo Soares, Ph.D  
(Presidente)

---

Prof. Helio José Correa Barbosa, Ph.D

---

Prof. Carlos Frederico Fontes, Ph.D

---

Prof. Pedro Pascutti, D.Sc

---

Prof. Ricardo Bicca de Alencastro, Ph.D

PETRÓPOLIS, RJ - BRASIL  
JANEIRO DE 2009

Soares, Rosemberg de Oliveira

S676a           Análise do impacto do polimorfismo genético do subtipo C do HIV-1 na interação da protease viral com o inibidor nelfinavir por modelagem e dinâmica molecular / Rosemberg de Oliveira Soares. Petrópolis, RJ. : Laboratório Nacional de Computação Científica, 2009.

xxviii, 177 p. : il.; 29 cm

Orientadore(s): Marcelo Soares e Laurent Dardenne

Dissertação (Mestrado) – Laboratório Nacional de Computação Científica, 2009.

1. Enzimas Proteolíticas. 2. HIV-1. 3. Dinâmica Molecular. 4. Nelfinavir. I. Soares, Marcelo. II. LNCC/MCT. III. Título.

CDD 572.76



“Não se pode ensinar alguma coisa a alguém,  
pode-se apenas auxiliá-la a descobrir por si  
mesmo.” (*Galileu Galilei*)

“Por esta paz que nós nos transmitimos,  
por este amor que repartimos...  
Por isso e muito mais eu digo:  
Obrigado.”

À minha amiga e namorada Amanda, pelo  
carinho, força e por estar sempre ao meu  
lado em todos os momentos que se fizeram  
necessários.

# Agradecimentos

Agradecer àqueles que foram importantes na minha jornada é uma grande tarefa e desde já peço desculpa aos que não se encontrarem neste texto. Agradeço aos meus queridos orientadores, Laurent Dardenne e Marcelo Soares, que durante esses longos anos de convivência, confiaram, acreditaram, incentivaram e valorizaram o meu trabalho permitindo assim o meu desenvolvimento científico.

Agradeço ao Prof. Pedro G. Pascutti, por me receber em seu laboratório, com cordialidade, alegria e por sua ampla dedicação a todos os seus estudantes.

Ao amigo e Co-orientador Paulo Ricardo Batista, por me ensinar o que é fazer ciência ao longo desses anos e por suas sempre ótimas contribuições no aprimoramento deste trabalho.

As amigas Silvane e Simone, por participarem ativamente na edição deste trabalho, contribuindo para a obtenção de uma dissertação com resultados e escritas claros e de fácil entendimento.

Aos Professores das disciplinas que cursei, que com muita cordialidade me forneceram informações altamente relevantes para o meu desenvolvimento acadêmico e intelectual.

Resumo da Dissertação apresentada ao LNCC/MCT como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (Mestrado)

**ANÁLISE DO IMPACTO DO POLIMORFISMO GENÉTICO DO  
SUBTIPO C DO HIV-1 NA INTERAÇÃO DA PROTEASE VIRAL  
COM O INIBIDOR NELFINAVIR POR MODELAGEM E  
DINÂMICA MOLECULAR**

Rosemberg de Oliveira Soares

Janeiro , 2009

**Orientador:** Marcelo Soares, Ph.D

**Co-orientador:** Laurent Dardenne, D.Sc

O HIV pode ser dividido em HIV-1 e HIV-2. Aquele, por sua vez, pode ser dividido nos grupos: M, N e O. O grupo M, que representa 90% das infecções, foi dividido em vários subtipos (A, B, C, D, F, G, H, J e K). Sabe-se hoje que o subtipo mais circulante no mundo (a maior parte na África) é o C, entretanto o mais estudado é o B (prevalente nos EUA e Europa)

Diversas etapas do ciclo replicativo do HIV-1 têm sido identificadas como alvos para intervenção farmacológica. Um dos principais alvos é a enzima aspartilprotease (PR); é ela que processa as poliproteínas virais Gag e Gag-Pol e sua inibição resulta na formação de partículas virais não infecciosas, sendo atualmente 10 inibidores utilizados em clínica. No entanto, o aparecimento de resistência a esses inibidores leva à falha terapêutica, tendo sido identificados e estudados vários resíduos que se apresentam mutados em isolados resistentes. Uma dessas mutações de resistência é a D30N, que consiste numa mutação primária de resistência exclusiva ao nelfinavir descrita em pacientes soropositivos infectados pelo subtipo B. Entretanto, observações clínicas e laboratoriais mostraram que vírus do subtipo C com a mutação D30N (CD30N) têm baixíssima ocorrência clínica e adaptabilidade

reduzida *in vitro*.

Para tentar entender as diferenças causadas pela mutação D30N nos subtipos B e C, foi estudada a interação da PR destes vírus com o peptídeo KARVLAEAM (análogo ao substrato natural de clivagem entre a proteína do capsídeo (CA) e a proteína p2 do HIV-1) e com o inibidor nelfinavir. Também foi estudada a PR CD30N com as mutações compensatórias N83T e N88D, encontradas *in vitro* e *in vivo* respectivamente, que se manifestam quando o subtipo C sofre a mutação D30N. Este trabalho teve como objetivo estudar os mecanismos moleculares e atômicos dos efeitos da mutação D30N na PR dos subtipos B e C.

Os resultados mostram que o inibidor e o esqueleto peptídico dos modelos BD30N e CD30N/N83T sofreram as maiores variações, em relação à estrutura inicial. Embora os mutantes CD30N e CD30N/N88D não tenham sofrido variação semelhante, eles apresentaram, assim como os outros dois mutantes, uma redução na intensidade das ligações de hidrogênio que ocorrem entre a PR e o inibidor que estão localizadas próximas à região catalítica e aos flaps. Além disso, todos os mutantes apresentaram redução em seus contatos hidrofóbicos ocorridos na interação receptor/ligante. Alguns dados obtidos indicam que a alça de uma das cadeias é altamente imóvel no modelo CD30N sugerindo que a mutação D30N prejudica o contato do “flap” com o substrato no subtipo C. Além disso, a análise da estrutura das PRs, interagindo com o substrato, indicou que o mutante CD30N tem uma de suas regiões de  $\alpha$ -hélice desestruturada, o que pode estar diretamente associado a não clivagem do substrato.

O nosso trabalho provê importantes “insights” sobre o efeito da mutação D30N na estrutura da PR do subtipo C, bem como na sua interação com o substrato e com o inibidor. Tais dados corroboram e explicam, ao menos em parte, a menor ocorrência da mutação estudada naquele variante genético do HIV-1.

Abstract of Dissertation presented to LNCC/MCT as a partial fulfillment of the requirements for the degree of Master of Sciences (Mestrado)

**ANALYSE THE IMPACT OF GENETIC POLYMORPHISM OF  
SUBTYPE C OF HIV-1 PROTEASE INHIBITORS IN THE  
INTERACTION VIRAL WITH THE INHIBITOR NELFINAVIR BY  
MODELING AND MOLECULAR DYNAMICS**

Rosemberg de Oliveira Soares

January, 2009

**Advisor:** Marcelo Soares, Ph.D

**Co-advisor:** Laurent Dardenne, D.Sc

The human immunodeficiency virus (HIV) can be divided into HIV-1 and HIV-2. The former can be divided into groups: M, N and O. Group M, which represents 90% of infections, is divided into several subtypes (A, B, C, D, F, G, H, J and K). It is known today that the most prevalent subtype in the world (and in Africa) is the subtype C, although the most studied is B (prevalent in the U.S. and Western Europe).

Several stages the HIV-1 replicating cycle have been identified as a target for pharmacologic intervention. One of the main targets is the enzyme aspartyl protease (PR), which processes the viral polyprotein Gag and Gag-Pol. Its inhibition results in the formation of non-infectious virus particles. Currently 10 PR inhibitors are used in clinic. However, the emergence of resistance to these inhibitors leads to a therapeutic failure. Several mutated amino acid residues that are present in resistant isolates have been identified. One of such resistance mutations is the D30N, which confers primary resistance exclusively to nelfinavir, has been described in patients infected with subtype B. However, clinical and laboratory studies showed that virus of subtype C with the mutation D30N (CD30N) has low

incidence in clinical and reduced adaptability *in vitro*.

To try to understand these differences caused by mutation D30N in subtypes B and C, we studied the interaction of these PRs with the peptide KARVLAEAM (analogous to the natural substrate of cleavage between the protein the capsid (CA) and p2 of HIV-1) and with the inhibitor nelfinavir. We have also studied the PR CD30N with the compensatory mutations N83T or N88D, found *in vitro* and *in vivo*, respectively, which occur when the subtype C acquires the mutation D30N. This work aimed to study the molecular and atomic mechanisms of mutation D30N in the PR of subtypes B and C.

The results showed that the inhibitor and backbone of models BD30N and CD30N/N83T possessed the greatest variation, with respect to the initial structure. Although the mutants CD30N and CD30N/N88T have not suffered similar variations, they showed, as well as the other two mutants, a reduction in the intensity of the h-bonds that occur between PR and inhibitor which are located near the catalytic and the flaps regions. Also, all mutants had reduced hydrophobic contacts between the receptor and the ligand. Some data indicated that the flap of one of the chains is highly immobile in a model CD30N suggesting the mutation D30N impairs the contact of flap with the substrate in subtype C. Also, the analysis of the PR structure interacting with the substrate, indicated that the CD30N mutant has one of its  $\alpha$ -helix regions unstructured, which can be directly associated with substrate cleavage.

Our work provides important insights in to the effect of D30N mutation in the PR structure of the subtype C, and on its interaction with the substrate and the inhibitor. These data confirm and explain, at least in part, the smaller incidence of the studied mutation in that genetic subtype of HIV-1.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Um Breve Histórico . . . . .	1
1.1.1	As primeiras Informações . . . . .	1
1.1.2	Teorias da Primeira década . . . . .	3
1.1.3	Panorama Atual: As Contradições e Desafios Persistem . . . . .	5
1.2	O HIV . . . . .	7
1.2.1	A Origem do Vírus . . . . .	7
1.2.2	A Estrutura do HIV . . . . .	8
1.2.3	O ciclo viral . . . . .	10
1.2.4	Variação genética . . . . .	18
1.3	Alvos Moleculares para a Terapia Anti-retroviral . . . . .	18
1.3.1	Trascriptase reversa (RT) . . . . .	20
1.3.2	Integrase . . . . .	22
1.3.3	Fusão e Entrada . . . . .	22
1.3.4	Protease do HIV (PR) . . . . .	25
1.4	Inibidores da Protease do HIV . . . . .	32
1.4.1	Interações Receptor-Ligante . . . . .	38
1.4.2	Resistência aos Inibidores de Protease . . . . .	41
1.5	Vacinas . . . . .	41
1.6	Subtipos do HIV-1 . . . . .	43
1.7	Aspectos Gerais da Mutação D30N . . . . .	45



1.8	Dinâmica Molecular & Protease do HIV-1 . . . . .	48
<b>2</b>	<b>Objetivos</b>	<b>51</b>
2.1	Objetivos Gerais . . . . .	51
2.2	Objetivos Específicos . . . . .	51
<b>3</b>	<b>Métodos</b>	<b>52</b>
3.1	Modelagem Comparativa (MC) . . . . .	53
3.2	Otimização das Geometrias Moleculares . . . . .	56
3.3	Equilibração do Sistema . . . . .	57
3.4	Ensemble . . . . .	57
3.5	Dinâmica Molecular Clássica (DM) . . . . .	58
3.6	Função Energia Potencial ou Campo de Força . . . . .	60
3.7	Correções das Interações de Longo Alcance . . . . .	68
3.8	Condições de Fronteira . . . . .	71
3.9	Parâmetros das Simulações . . . . .	73
3.9.1	Análise de Estruturas e Geração de Figuras . . . . .	76
3.10	Topologia dos inibidores . . . . .	77
3.11	Desvio da Raiz Média Quadrática (RMSD) . . . . .	77
3.12	Flutuação da Raiz Média Quadrática . . . . .	78
3.13	Raio de Giro e Momento de Inércia . . . . .	78
3.14	Ligações de Hidrogênio . . . . .	78
3.15	Área de Interação Intermolecular . . . . .	79
<b>4</b>	<b>Resultados do Estudo de Casos – Subtipos B e C Complexados ao Nelfinavir</b>	<b>81</b>
4.1	Interação entre a PR e o nelfinavir . . . . .	81
4.1.1	Propriedades Globais . . . . .	81
4.1.2	Flutuações dos Resíduos . . . . .	84
4.1.3	Representação 3D do RMSD para os átomos do backbone . . . . .	88
4.1.4	Variações nas alças da Protease . . . . .	90

4.1.5	Variações estruturais na Compactação da PR Quando Complexada ao Inibidor . . . . .	92
4.1.6	Comportamento do Nelfinavir durante a DM . . . . .	94
4.1.7	Trajétórias das Dinâmicas . . . . .	97
4.1.8	Ligações de hidrogênio entre a PR e o Nelfinavir . . . . .	99
4.1.9	Superfície de Contato Intermolecular . . . . .	104
4.1.10	Padrões das estruturas secundárias . . . . .	110
<b>5</b>	<b>Resultados do Estudo de Casos – Subtipos B e C Complexados ao Substrato</b>	<b>112</b>
5.1	Interação entre a PR e o Substrato . . . . .	112
5.1.1	Propriedades Globais . . . . .	112
5.1.2	Flutuações dos Resíduos . . . . .	113
5.1.3	Representação 3D do RMSD para os átomos do esqueleto . .	118
5.1.4	Variações nas alças da Protease . . . . .	120
5.1.5	Variações Estruturais na Compactação da PR Complexada ao Substrato . . . . .	122
5.1.6	Comportamento do Substrato durante a DM . . . . .	124
5.1.7	Trajétórias das Dinâmicas . . . . .	128
5.1.8	Ligações de hidrogênio entre a PR e o substrato . . . . .	130
5.1.9	Contatos Moleculares entre a Proteína e o Ligante . . . . .	135
5.1.10	Padrões das estruturas secundárias . . . . .	140
<b>6</b>	<b>Discussão</b>	<b>142</b>
<b>7</b>	<b>Conclusões</b>	<b>149</b>
<b>8</b>	<b>Perspectivas</b>	<b>151</b>

<b>Referências Bibliográficas</b>	152
<b>Apêndice</b>	
<b>A Parametrização do Ligante</b>	164

# Lista de Figuras

## Figura

- 1.1 A – Microfotografia de uma partícula viral do HIV após purificação e concentração (barra 50 nm). B – Esquema da estrutura da partícula viral do HIV e localização de algumas das principais proteínas virais. Extraído de Pereira (2000).  
Obs: Nesta figura a sigla ARN foi usada em substituição a RNA. . . . . 10
- 1.2 Interação do HIV-2 com um linfócito T. A cabeça da seta indica partículas virais. Barra 0,5  $\mu\text{m}$ . Extraído de Pereira (2000) . . . . . 11
- 1.3 Representação esquemática do processo de fusão viral. Adaptado de [http://hivmedicine.aidsportugal.com/05\\_HAART\\_3.php](http://hivmedicine.aidsportugal.com/05_HAART_3.php). . . . . 11
- 1.4 Representação do ciclo de replicação do HIV-1: 1. Ataque à célula hospedeira, 2. Processo de fusão, 3. Transcrição Reversa, 4. Integração, 5. Tradução, 6. Brotamento e 7. Maturação. Extraído de Alterman (2001) . . . . . 16
- 1.5 Representação esquemática dos alvos moleculares de bloqueio da replicação do HIV. Extraído de Batista (2006) . . . . . 19

1.6	Ilustração em 3D da protease do HIV complexada com o inibidor nelfinavir (código 1OHR no PDB). As cadeias estão na representação de estrutura secundária. A cor cinza identifica as $\alpha$ -hélices, a marrom as folhas- $\beta$ e branco os loops. Estão em destaque na proteína os dois ácidos aspárticos que classificam a protease. O nelfinavir segue acrescido de seus raios de Van der Waals em transparente onde o vermelho representa átomos de oxigênio, cinza carbono, azul nitrogênio e amarelo enxofre. . . . .	26
1.7	Em A: O desenho da cadeia principal da estrutura da protease do HIV-1; os flaps (F) e sítio ativo (A) são mostrados. Em B: Detalhe do sítio ativo da HIV-1 protease com o “fireman’s grip”; As pontes de hidrogênio são mostradas em linha verdes. (adaptado de Ingr et al. (2003)) . . . . .	27
1.8	O esquema acima dá o tamanho relativo e a identificação de cada sub-sítio e seus respectivos resíduos associados. O sítio catalítico da aspartil protease de retrovírus clivaria a seqüência peptídica do substrato procurando por uma subseqüência predominante de resíduos hidrofóbicos. Adaptado de TÖSZÉR et al. (2000). . . . .	30
1.9	Provável mecanismo de ação das proteases de HIV envolvendo a clivagem da ligação peptídica. Extraído de Batista (2006). . . . .	31
1.10	Representação de oito inibidores comerciais de HIV-1 PR atualmente disponíveis e aprovados pelo FDA. . . . .	34
1.11	Esquema geral de ligações hidrogênio entre HIV-1 PR e um substrato representativo. As flechas indicam a ligação a ser clivada. Os resíduos na porção amino-terminal da ligação a ser clivada são chamados P1-Pn, e aqueles na porção carboxi-terminal, de P1'-Pn', de acordo com Schechter & Berger(1967). Os símbolos S4-S4' denotam os bolsões correspondentes a cada sub-sítio da enzima. Adaptado de Wlodawer e Erickson (1993). . . . .	35

1.12	Decréscimo na mortalidade dos casos de AIDS associado à terapia com inibidores de HIV tipo 1 para indivíduos norte-americanos. Adaptado de Palella et al. (1998). . . . .	37
1.13	Representação esquemática do processo de reação química com e sem enzima. . . . .	40
1.14	Estimativa do número de pessoas infectadas com HIV em 2007. Extraído de Un aids (2007). . . . .	44
3.1	Alinhamento comparativo das proteases utilizadas como referência. Os símbolos qualificam o alinhamento quanto a similaridade: traço representa que os aminoácidos das estruturas são idênticos; dois pontos representa que há uma substituição por um aminoácido conservativo; um ponto quando as substituições não são conservativas. . . . .	54
3.2	Diagrama de Ramachandran de um dos modelos criados, no qual estão mostradas as regiões fisicamente permitidas para os aminoácidos de acordo com os ângulos $\phi$ e $\psi$ . Em vermelho estão as regiões mais favorecidas, em amarelo escuro as regiões favorecidas, em amarelo claro as regiões pouco favorecidas e em branco as regiões desfavorecidas. . . . .	55
3.3	Ligação peptídica de dois aminoácidos consecutivos. Os ângulos torcionais $\phi$ e $\psi$ permitem que cada plano formado por esta ligação gire em relação aos seus vizinhos. . . . .	56
3.4	Representação da energia potencial de ligação para um par n qualquer. O comprimento da ligação oscila com o tempo segundo um potencial harmônico em torno de uma posição de equilíbrio (imagem adaptada de Silva (2003)). . . . .	62
3.5	Representação da energia potencial angular para um terço de átomos ligados n qualquer cujo ângulo entre eles varia harmonicamente em torno do ângulo de equilíbrio (imagem adaptada de Silva (2003)). . . . .	63

3.6	Representação da energia potencial diedral imprópria. Um carbono $\alpha$ quiral está representado aqui. Como na maioria dos campos de força o átomo de hidrogênio (representado para efeitos didáticos) fica integrado no $C\alpha$ . Este potencial cuida de manter a estabilidade tetraédrica deste grupo de átomos (imagem adaptada de Silva (2003)). . . . .	64
3.7	Representação da energia potencial torcional própria. Nesta representação, o anel benzênico (no plano diédrico 1_2_3) pode assumir seis ângulos de torção (mínimos) em relação ao plano do diedro 2_3_4 (imagem adaptada de Silva (2003)). . . . .	65
3.8	Representação atômica usando-se os raios de Van der Waals (Rvdw), os quais são definidos para cada átomo pela metade da distância em que seu potencial de Lennard-Jones vale zero (o gráfico não está na mesma escala que o desenho) (imagem adaptada de Silva (2003)).	66
3.9	Representação esquemática da energia potencial de Coulomb entre dois átomos de cargas opostas. No gráfico da função de Coulomb, o eixo x corresponde à distância (em $\text{\AA}$ ) e o eixo y corresponde à energia (KJ/mol) (Adaptado de Silva (2003)) . . . . .	67
3.10	Representação do potencial de Coulomb com e sem a aplicação do Campo de Reação. Em A, figura representativa das considerações para cada tratamento eletrostático. Em B, representação dos efeitos da aplicação do campo de reação na função de Coulomb para partículas de mesma carga com e sem o campo de reação, utilizando um raio de corte de $16 \text{\AA}$ . Imagem adaptada de Goliatt (2007). . . . .	70
3.11	Enzima protease do HIV em solvente explícito. As cadeias da protease estão na representação de estrutura secundária. A cor rosa identifica as $\alpha$ -hélices, amarelo folhas- $\beta$ e verde ou branco os loops. As moléculas de água estão na representação de varetas, coloridas pelo tipo de átomos: branco (H) e vermelho (O). . . . .	71

3.12	Condições periódicas de contorno retangular. Uma proteína e seu solvente com suas oito imagens replicadas. Uma esfera de raio $R_c$ é traçada a partir do átomo $i$ , o qual interage com a imagem do átomo $j$ ( $j'$ ). . . . .	72
3.13	Alinhamento das seqüências das PRs construídas. As colunas em destaque representam os polimorfismos encontrados entre os subtipos B e C (T12S, I15V, L19I, M36I, R41K, H69K, L89M e I93L); em vermelho estão as mutações D30N, N83T e N88D; e em rosa a diferença entre a seqüência 1OHR e a consenso. . . . .	74
3.14	Ligações de hidrogênio. Dados $r$ (distância) e $\alpha$ (ângulo) são três as convenções para determinação dessas ligações. Em (a) $r$ é tomado entre D-A inferior a $3,6 \text{ \AA}$ , com $\alpha$ variando entre $0 - 60^\circ$ e em (b) $r$ é tomado entre H-A inferior a $2,7 \text{ \AA}$ , com $\alpha$ variando entre $120 - 180^\circ$ . A combinação de critérios entre (a) e (b) também pode ser utilizada.	79
3.15	Superfície Molecular e Superfície Intermolecular. Representação bidimensional. Uma ponta de prova esférica corre por sobre a superfície de van der Waals (SvdW), gerando superfícies de contato (SC) e superfícies de reentrância (SR), que somadas compõem a superfície excluída ao solvente ou superfície molecular (SM). O centro de prova desenha a superfície acessível ao solvente (SAS). A superfície Intermolecular (SI) seria a intersecção das SAS de moléculas diferentes dentro do espaço definido pela ponta de prova. . . . .	80
4.1	Desvio Médio Quadrático (RMSD) da PR, em relação à estrutura inicial (pós - minimização de energia), no tempo (10 ns de simulação): consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). . . . .	83
4.2	Em A, RMSF do ConsB (verde) e ConsC (rosa). Em B, diferenças entre o RMSF do subtipo C e B (ConsC-ConsB). . . . .	85



4.3	Em A, RMSF do ConsB (verde) e BD30N (vermelho). Em B, diferenças entre o RMSF do mutante BD30N e ConsB (BD30N-ConsB).	86
4.4	Em A, RMSF de quatro sistemas: ConsB (verde), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o RMSF do subtipo B e os demais. CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D-consB (amarelo) . . . . .	87
4.5	Representação 3D das flutuações do RMSD do esqueleto peptídico da PR. O RMSF foi representado por um tubo com a espessura variada, onde quanto mais espesso o tubo, maior o desvio. Foram coloridos por estrutura secundária: em vermelho, $\alpha$ -hélice; em azul, folhas- $\beta$ e em cinza, loops e estruturas randômicas. As esferas representam as mutações: em vermelho, D30N; em azul, N83T e em amarelo, N88D. . . . .	89
4.6	Distância entre as pontas das alças e o centro catalítico calculada para cada sistema. Em A, distância referente a alça da cadeia A (ILE 50). Em B distância referente a cadeia B (ILE 149). Em verde ConsB, em vermelho BD30N, em rosa ConsC, em preto CD30N, em azul CD30N/N83T e em amarelo CD30N/N88D. A figura foi invertida visando facilitar a visualização do gráfico. . . . .	91
4.7	Representação do Raio de Giro de todos os sistemas durante a DM. ConsB (verde), BD30N (vermelho), ConsC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o subtipo B e os demais. ConsB-BD30N (vermelho), ConsC-ConsB (rosa), CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D (amarelo) . . . . .	92
4.8	Estrutura 3D do nelfinavir com seus grupos de interação P2, P1, P1' e P2'. . . . .	94

4.9	Sobreposições das estruturas do inibidor durante a DM. Representação das ligações químicas do nelfinavir, coloridas pelo tipo de átomo: branco (H), verde (C), azul (N), vermelho (O) e amarelo (S). . . . .	95
4.10	Flutuações do RMSF calculadas para os átomos do inibidor em cada sistema: consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D(amarelo). Os grupos P2' (resíduos 1 a 8), P1' (resíduos 9 a 22), P1 (resíduos 23 a 36), P2 (resíduos 37 a 52) do inibidor estão indicados, tanto no gráfico quanto na representação. . . . .	96
4.11	“Snapshots” da estrutura das PRs, representando somente o esqueleto peptídico, demonstrando a trajetória das PRs durante os 8 <i>ns</i> de simulação. . . . .	98
4.12	Esquema em 2D do sítio ativo da estrutura cristalográfica 1OHR. Estão representados os átomos do inibidor, exceto os hidrogênios. Em verde etão representadas as ligações de hidrogênio, com seus respectivos comprimentos. Cada aminoácido que está presente na superfície intemolecular de interação da PR com o nelfinavir está representado em vermelho.	
	* – Os resíduos seguidos por <b>(A)</b> representam os aminoácidos de 1 a 99 e os seguidos por <b>(B)</b> representam os de 100 a 198. . . . .	103
4.13	Área da Superfície de Contato Intermolecular entre a PR e o nelfinavir. Em barra estão representados os valores médios e em linha o desvio-padrão. Somente os contatos com mais de 5 Å <sup>2</sup> de área foram representados. . . . .	105
4.14	Padrões de estrutura secundária por resíduo, ao longo da dinâmica. Em branco, estão representados os resíduos em conformação aleatória; em verde, resíduos em curvas; em amarelo, resíduos em voltas; em azul, resíduos em α-hélice; em vermelho, resíduos em folha-β . .	111

5.1	Desvio Médio Quadrático (RMSD) da PR complexada ao substrato, em relação à estrutura inicial (pós – minimização de energia), no tempo (10 ns de simulação): consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). . . . .	113
5.2	Em A, RMSF do ConsB (verde) e ConsC (rosa). Em B, diferenças entre o RMSF do subtipo C e B (ConsC-ConsB). . . . .	115
5.3	Em A, RMSF do ConsB (verde) e BD30N (vermelho). Em B, diferenças entre o RMSF do mutante BD30N e ConsB (BD30N-ConsB). . . . .	116
5.4	Em A, RMSF de quatro sistemas: ConsB (verde), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o RMSF do subtipo B e os demais. CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D-consB (amarelo) . . . . .	117
5.5	Representação 3D das flutuações do RMSD do esqueleto peptídico da PR. O RMSF do backbone foi representado por um tubo com a espessura variada, onde quanto mais espesso o tubo, maior o desvio. Foram coloridos por estrutura secundária: em vermelho, $\alpha$ -hélice; em azul, folhas- $\beta$ e em cinza, loops e estruturas randômicas. As esferas representam as mutações: em vermelho, D30N; em azul, N83T e em amarelo, N88D. . . . .	119
5.6	Distância entre as pontas das alças e o centro catalítico calculada para cada sistema. Em A, distância referente a alça da cadeia A (ILE 50). Em B distância referente a cadeia B (ILE 149). Em verde ConsB, em vermelho BD30N, em rosa ConsC, em preto CD30N, em azul CD30N/N83T e em amarelo CD30N/N88D. A figura foi invertida visando facilitar a visualização do gráfico. . . . .	121

5.7	Representação do Raio de Giro de todos os sistemas durante a DM. ConsB (verde), BD30N (vermelho), ConsC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o subtipo B e os demais. ConsB-BD30N (vermelho), ConsC-ConsB (rosa), CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D (amarelo) . . . . .	122
5.8	Estrutura 3D do substrato CA/p2 com seus grupos de interação P5, P4, P3, P2, P1, P1', P2', P3' e P4'. . . . .	124
5.9	Sobreposições das estruturas do substrato durante a DM. Representações coloridas pelo tipo de átomo: branco (H), verde (C), azul (N), vermelho (O) e amarelo (S). . . . .	126
5.10	Flutuações do RMSF calculadas para os átomos do substrato em cada sistema: consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D(amarelo). Os grupos (P5, P4, P3, P2, P1, P1', P2', P3' e P4') do substrato estão indicados, tanto no gráfico quanto na representação. . . . .	127
5.11	“Snapshots” da estrutura das PRs, representando somente o esqueleto peptídico, demonstrando a trajetória das PRs durante os 10 <i>ns</i> de simulação. . . . .	129
5.12	Esquema em 2D do sítio ativo da estrutura cristalográfica 1F7A. Estão representados os átomos do inibidor, exceto os hidrogênios. Em verde estão representadas as ligações de hidrogênio, com seus respectivos comprimentos. Cada aminoácido que está presente na superfície intemolecular de interação da PR com o nelfinavir está representado em vermelho.	

\* – Os resíduos seguidos por **(A)** representam os aminoácidos de 1 a 99 e os seguidos por **(B)** representam os de 100 a 198. . . . . 134

5.13	Área da Superfície de Contato Intermolecular entre a PR e o substrato. Em barra estão representados os valores médio e em linha o desvio-padrão. Somente os contatos com mais de $5 \text{ \AA}^2$ de área foram representados. . . . .	136
5.14	Padrões de estrutura secundária por resíduo, ao longo da dinâmica. Em branco estão representados os resíduos em conformação aleatória; em verde, resíduos em curvas; em amarelo, resíduos em voltas; em azul, resíduos em $\alpha$ -hélice; em vermelho, resíduos em folha- $\beta$ . . . . .	141

# Lista de Tabelas

## Tabela

1.1	Genes e Proteínas do HIV . . . . .	17
1.2	As nove seqüências peptídicas reconhecidas pela PR . . . . .	28
1.3	Inibidores de protease (PIs) . . . . .	33
1.4	Características Fenotípicas dos Clones B e C relativos à inibição pelo nelfinavir (Adaptado de Gonzalez et al. (2004)) . . . . .	47
1.5	Ensaio de competição <i>in vitro</i> dos clones mutantes do HIV- 1. As duplas foram colocadas num mesmo meio de cultura e após um tempo foi avaliado qual estava em maior quan- tidade (Adaptado de Gonzalez et al. (2004)) . . . . .	48
3.1	Seqüências Proteicas utilizadas como molde na Construção dos Modelos . . . . .	54
3.2	Ficha técnica das condições gerais das simulações realizada por MD das proteases B e C . . . . .	76
4.1	Valores Médios do RMSD (nm) do esqueleto da protease complexadas ao nelfinavir . . . . .	84
4.2	Valores médios dos momentos de inércia durante a DM . . . . .	93
4.3	Número médio de Ligações de Hidrogênio entre a PR e o nelfinavir. . . . .	99
4.4	Ligações de hidrogênio (%) entre a PR e o nelfinavir. . . . .	102
4.5	Área de Contato ( $\text{Å}^2$ ) entre a PR e o nelfinavir . . . . .	108
4.6	Área de Contato ( $\text{Å}^2$ ) Hidrofóbica entre a PR e o nelfinavir	109

5.1	Valores Médios do RMSD (nm) do esqueleto da protease complexadas ao substrato . . . . .	113
5.2	Valores médios dos momentos de inércia dos modelos complexado ao substrato durante a DM . . . . .	123
5.3	Número médio de Ligações de Hidrogênio entre a PR e o substrato. . . . .	130
5.4	Ligações de hidrogênio (%) entre a cadeia A da PR e o substrato. . . . .	132
5.5	Ligações de hidrogênio (%) entre a cadeia B da PR e o substrato. . . . .	133
5.6	Área de Contato ( $\text{\AA}^2$ ) entre a PR e o substrato . . . . .	138
5.7	Área de Contato ( $\text{\AA}^2$ ) Hidrofóbica entre a PR e o substrato	139

# Lista de Siglas e Abreviaturas

- AIDS: Síndrome da Imunodeficiência Adquirida
- APV: Amprenavir
- ATV: Atazanavir
- CRF: Formas Recombinantes Circulantes
- CTE: Elemento de Transporte Constitutivo
- DNA: Ácido Desoxirribonucléico
- FDA: “Food and Drug Administration”
- FPV: Fosamprenavir
- HAART: Terapia Anti-Retroviral de Alta Eficácia
- HIV: Vírus da Imunodeficiência Humana
- HTVL-I: Vírus de Leucemia de Células T Humana Tipo I
- IDV: Indinavir
- KS: Sarcoma de Kaposi
- LAV: Vírus associados à Linfadenopatia
- LPV: Lopinavir
- MD: Dinâmica Molecular
- MC: Modelagem Comparativa
- NFV: Nelfinavir
- NRTI: Inibidores Nucleosídeos da Transcriptase Reversa
- NNRTI: Inibidores Não-Nucleosídeos da Transcriptase Reversa
- PDB: Protein Dat Bank
- PCP: Pneumonia por *Pneumocystis carinii*



- PIC: Complexo de Pré-Integração
- PME: Particle Mesh Ewald
- PR: Protease do HIV
- RC-CR: Raio de Corte com Campo de Reação
- RC-T: Raio de Corte Truncado
- RMN: Ressonância Magnética Nuclear
- RMSD: Desvio Quadrático Médio
- RMSF: Flutuação Média Quadrática
- RNA: Ácido Ribonucléico
- RSV: Vírus do Sarcoma de Rous
- RTV: Ritonavir
- RT: Transcriptase Reversa
- SQV: Saquinavir
- SIV: Vírus da Imunodeficiência de Símios
- SI: Indutor de Sincício

# Capítulo 1

## Introdução

### 1.1 Um Breve Histórico

#### 1.1.1 As primeiras Informações

A Síndrome da Imunodeficiência Adquirida (AIDS) é a designação dada à doença provocada pela infecção causada pelo vírus da Imunodeficiência Humana (HIV) em humanos e foi reconhecida em 1981. Os primeiros casos foram reconhecidos nos Estados Unidos e na Europa devido à identificação de um elevado número de pacientes com uma forma mais agressiva de Sarcoma de Kaposi (KS) – tipo benigno de câncer que tende a ocorrer em pessoas idosas (que hoje se sabe é uma das manifestações clínicas da AIDS) (Hymes et al., 1981), e da identificação clínica, em jovens homossexuais do sexo masculino, de uma rara infecção pulmonar, a PCP (pneumonia por *Pneumocystitis carini* – fungo cuja denominação foi posteriormente mudada para *Pneumocystitis jiroveci*) (Friedman-Kien et al., 1981). Todos esses pacientes apresentavam uma reduzida resposta de linfócitos T à estimulação por antígenos ou mitógenos, ou seja, uma severa imunodeficiência celular (Ginzburg et al., 1985). O termo AIDS foi então empregado para definir clinicamente as diversas manifestações desta doença. Embora a palavra síndrome seja por definição “a reunião de sinais e sintomas que ocorrem em conjunto e que caracterizam uma doença ou uma perturbação”, a AIDS é considerada uma doença em si e caracterizada pelo comprometimento do sistema imunológico do organismo.

Em busca de esclarecimentos os cientistas iniciaram uma tarefa difícil, desco-

brir qual era o agente causador da nova doença. A causa dessa síndrome ainda era desconhecida, mas a suspeita de um novo retrovírus era uma idéia atrativa, uma vez que alguns retrovírus como o FLV (“Feline Leukemia Virus”) sabidamente induziam imunodeficiência, assim como leucemias/linfomas em seus hospedeiros. Além disso, este vírus possuía similaridades entre os retrovírus humanos previamente descobertos, os HTLV-I e HTLV-II (Marx, 1983). Em maio de 1983 o agente etiológico da AIDS foi considerado isolado pelos pesquisadores Luc Montagnier (Barresinoussi et al., 1983), na França e Robert Gallo (Gallo et al., 1983), nos EUA, recebendo os nomes de LAV (lymphadenopathy-associated vírus) e HTVL-III (Human T-Lymphotropic Vírus Type III) respectivamente nos dois países. Em maio de 1986, o Comitê Internacional de Taxonomia de Vírus decidiu que os dois nomes (LAV e HTVL-III) deveriam ser deixados de lado e recomendou o termo HIV (“Human Immunodeficiency Virus” ou Vírus da Imunodeficiência Humana) (Coffin e et al, 1986).

Um segundo vírus associado a AIDS, denominado Vírus da Imunodeficiência Humana Tipo 2, foi isolado de indivíduos do oeste da África (Clavel et al., 1986) e parece ser mais relacionado com o SIV<sub>sm</sub>, (isolado de “sooty mangabeys” do centro-oeste da África) do que com o HIV-1 (Huet et al., 1990). Os genomas do HIV-1 e do HIV-2 diferem de 50 a 65% um do outro em relação à seqüência nucleotídica. Enquanto o HIV-1 é encontrado nas diferentes regiões do mundo, como o tipo predominante, o HIV-2 é restrito a áreas endêmicas, tais como África Ocidental e Índia, e foi também detectado em outras regiões da África, Europa e América do Sul (Cock et al., 1993).

A transmissão do HIV pode ocorrer através de relações sexuais, de sangue e produtos sangüíneos contaminados, de órgãos e sêmen doados e perinatalmente. A transmissão sexual é uma das principais formas de disseminação da doença, representando 75-85% das infecções pelo HIV em adultos (Baba et al., 1996). Atualmente a epidemia de AIDS se expande por todas as partes do mundo e sua incidência entre mulheres na idade reprodutiva aumenta gradualmente desde o iní-

cio da epidemia. Estudos recentes apontam que no fim de 2007 havia cerca de 30 a 36 milhões de pessoas infectadas pelo HIV, e destes, cerca de 22 milhões vivem na África sub-saariana (Unaid, 2007). Estimativas globais apontam para 4,3 a 6,4 milhões de novos casos de infecção pelo HIV-1 nos próximos 5 anos (Stover et al., 2006).

### 1.1.2 Teorias da Primeira década

Nos primeiros anos da década de 1980 a AIDS era praticamente desconhecida. O que todos sabem é que a partir de casos de uma pneumonia atípica e de um câncer raro foi identificado uma nova e misteriosa síndrome. A epidemia que já vinha afetando a população dos países africanos havia até então passado despercebido dos cientistas ocidentais.

Em um contexto de infecções, desnutrição, guerras civis e miséria não foi possível compreender que uma nova ameaça a humanidade estava a caminho. Os epidemiologistas acabaram, por quase uma década, seguindo teorias equivocadas o que permitiu o avanço da epidemia. Tais teorias foram denominadas de fábulas dos 4 Hs (Bastos, 2006).

- **O Primeiro H** – Os anos 1960 foram para os países, ditos ocidentais, um momento marcado por grandes transformações sociais e protestos, com o surgimento de movimentos sociais de grande envergadura, entre eles a cena gay norte americana. Emergindo de um contexto de repressão e ilegalidade, a cena gay se moveu no sentido de exaltação da liberdade sexual, orgulho por ser homossexual e engajamento político. Este movimento veio a influenciar vários outros de caráter semelhante. Foram exatamente esses grupos que foram profundamente atingidos pela disseminação do HIV, por volta do final da década de 1970 (tendo seu reflexo na década seguinte).

A partir da disseminação do HIV nesses grupos, uma combinação de falta de conhecimento e preconceito ajudou a construir a fábula do primeiro H. A idéia inicial era a de que a nascente síndrome não seria decorrente

de uma patologia infecciosa, mas devido ao uso de substâncias consumidas por tais grupos: principalmente o nitrito de amila (mais conhecido como popper). Entretanto, o isolamento do vírus, em 1983, tornou a hipótese de caráter não infeccioso indefensável. No mesmo período, um outro grupo de pesquisadores afirmava que a síndrome seria uma infecção exclusiva de homens **homossexuais**, mas a suposta essencialidade da síndrome gay foi desmontada quando, em 1983, foi verificado que a doença poderia ser transmitida heterossexualmente, do homem para mulher (Harris et al., 1983). Surgiu, então, uma terceira idéia equivocada – a suposta exclusividade da relação anal como modo de transmissão do HIV.

- ***O Segundo H*** – Uma segunda teoria preconceituosa foi inferir que havia uma preferência do vírus HIV por **haitianos** quando, na verdade, a disseminação do vírus na população haitiana estava estritamente relacionada com a exploração sexual vivida por aquele país e de séculos de injustiças e violações dos seus direitos (Bastos, 2006).
- ***O Terceiro H*** – Com a inovação dos processos biológicos, envolvidos na produção dos “hemoderivados” (produtos biológicos derivados do sangue) muitas pessoas passaram a fornecer sangue em troca de compensação financeira (denominados de doadores profissionais). Esses doadores eram em sua maioria oriundos das camadas mais pobres e marginalizadas de cada sociedade. Como se trata de uma população vivendo em condições precárias, não é de se estranhar que a taxa de infecção pelo HIV seja mais elevada do que no resto da população. Além disso, na produção de hemoderivados são utilizadas diversas amostras de vários doadores para a produção de um único lote de cada hemoderivado específico, o que aumentam as chances de uma contaminação.

Por essas razões e pelo fato dos **hemofílicos** receberem determinado hemoderivado de forma repetida ao longo da vida, um indivíduo desse grupo

possui grandes chances de ser infectado. Assim, no final da década de 1970 e início da década de 1980 milhares desses pacientes que haviam recebido transfusões sanguíneas foram infectados.

No Brasil, sob a liderança de Herbert de Souza, o Betinho, implementou-se em meados da década de 1980 uma reforma da legislação relativa a doações e a operações de banco de sangue.

- ***O Quarto H*** – Mesmo após vários equívocos um último iria surgir. Esse foi referente aos usuários de drogas injetáveis. Um dos fatos mais curiosos nesse caso é que apenas os usuários de heroína eram considerados como parte desse grupo. Entretanto, a cocaína injetada estava bastante presente na cena de drogas nos Estados Unidos e Europa ocidental na década de 1980 (quando foi formulada essa idéia), ou seja, uma boa parte dos usuários de drogas injetáveis, cujas taxas de infecção são elevadas, foi simplesmente ignorada.

Sabe-se hoje que esse grupo está fortemente associado à transmissão do HIV devido ao compartilhamento de agulhas e seringas. Paradoxalmente, cabe observar que um dos elementos chave desse compartilhamento ocorreu devido à implementações de legislações que tinham como finalidade evitar o uso de drogas injetáveis, embora a prática do compartilhamento desses instrumentos tenha ocorrido também em países que jamais adotaram tais legislações como, por exemplo, o Brasil.

### **1.1.3 Panorama Atual: As Contradições e Desafios Persistem**

Face às contradições e hesitações dos governos centrais (por ex., EUA), coube às comunidades mais afetadas diretamente (como a comunidade gay norte-americana) lutar contra a AIDS. A resposta impressa por essas comunidades foi, em geral, bem sucedida devido à sua capacidade de mobilização. Um dos momentos mais tenso dessa luta foi a luta dessas comunidades pelo acesso a novos medicamentos.

Em 1996 com o anúncio da combinação de diferentes medicamentos com propriedades anti-retrovirais no tratamento da infecção pelo HIV, popularmente conhecido como coquetel, pôs-se por terra uma série de mitos sobre a AIDS. Um deles era o de que não era possível reconstituir as funções do sistema imunológico, uma vez que essas funções estivessem comprometidas pelo contínuo ataque movido pelo HIV. O sistema imunológico que emergiu da era pós-coquetel ou pós-HAART (Terapia Antiretroviral de Alta Eficiência) se mostrou bastante dinâmico, ou seja, livre por algum tempo dos ataques empreendidos pelo HIV, o sistema imunológico se mostrou plenamente capaz de recompor sua estrutura e função, com uma redução ou mesmo eliminação das doenças secundárias (denominadas também de oportunistas). Livre dessas doenças as pessoas vivendo com AIDS puderam retomar suas vidas cotidianas sem os sobressaltos dos sucessivos episódios de doença.

Um dos desafios da era pós-HAART é transformar a terapia em uma medida de saúde pública, ou seja, um recurso vital que chegue exatamente aos que dela necessitam. O Brasil foi o primeiro país em desenvolvimento a instituir a HAART em larga escala e a custo zero para o paciente. Da perspectiva do Brasil, de uns poucos países em desenvolvimento ou da maioria dos países desenvolvidos, a maioria das pessoas vivendo com AIDS, devidamente identificadas como tais, recebe a medicação anti-retroviral, que lhes garante uma vida plena e substancialmente mais longa, ainda que não isenta de riscos e incômodos.

O advento da HAART possibilitou que as pessoas infectadas pudessem retomar suas vidas plenamente e nisso reside uma nova dimensão da prevenção – a prevenção passa a ser dirigida não só a evitar que as pessoas se infectem, mas também a evitar que pessoas que já estão vivendo com HIV não se reinfectem com o vírus (o que se traduz em riscos substanciais), não se infectem com outros agentes infecciosos e não infectem seus parceiros. Novos desafios surgem e passam a ter um caráter múltiplo e complexo.

Além disso, alguns dos pacientes dependentes de drogas estão freqüentemente infectados por outros vírus e bactérias e essas co-infecções tornam bastante difícil

o trabalho clínico. Paralelamente a isso, temos o aumento dos comportamentos de risco e o recrudescimento das infecções sexualmente transmissíveis. Em outras palavras, a AIDS passou a ser vista, por alguns grupos, como uma infecção trivial similar a um resfriado.

Uma nova geração de estratégias preventivas deve lidar com os desafios renovados da era pós-HAART. É importante mencionar os riscos de desenvolvimento da resistência viral e conseqüentemente a transmissão de cepas resistentes a novos indivíduos, o que dificulta ou até mesmo impossibilita a adoção de quaisquer opções terapêuticas.

Os benefícios advindos da HAART não devem ser motivos para o descaso e a indiferença. Se as conquistas advindas da HAART se traduzirem em complacência poderemos estar diante de uma epidemia com força renovada com a disseminação de vírus resistentes.

## **1.2 O HIV**

### **1.2.1 A Origem do Vírus**

Em todas as células vivas, a informação genética está presente no ácido desoxirribonucléico (DNA) (Hershey e Chase, 1952). Com sua estrutura química, essa substância carrega o código genético com as instruções para a produção de todas as proteínas que dão forma e função a cada organismo. A molécula que age diretamente na tradução de proteínas é o ácido ribonucléico (RNA), tal molécula copia a informação dos genes do DNA e a transmite para as proteínas atuando assim como um mensageiro. Esse padrão de fluxo da informação genética do DNA para o RNA e do RNA para a proteína ficou mundialmente conhecida como o Dogma Central da Biologia Molecular (Crick, 1970). Entretanto, alguns estudos com oncovírus (primeiro vírus causador de tumor descoberto, também chamado de vírus do Sarcoma de Rous – RSV) mostraram violações a esse Dogma, indicando que o RNA viral pode ser retro transcrito em DNA na célula do hospedeiro (Temin e Mizutani, 1970). Desse modo, foi descrita a transcrição reversa de RNA em DNA



viral, catalisada por uma enzima viral chamada de transcriptase reversa. O DNA formado a partir do RNA do vírus é integrado ao DNA da célula hospedeira para permitir a produção de novas partículas virais. Esse é o mesmo processo utilizado pelo vírus HIV.

Os vírus com capacidade de realizar a transcrição reversa biológica foram agrupados em uma família chamada *Retroviridae*. Tal família é compreendida principalmente por vírus que infectam eucariontes e ao gênero dos Lentivirus, que inclui vírus que possuem genomas complexos. Alguns lentivírus foram identificados infectando somente primatas não-humanos e são conhecidos coletivamente como vírus da imunodeficiência de símios (SIV).

Após o seqüenciamento dos genes dos vírus simianos da imunodeficiência em chimpanzés e de mais de 30 espécies de macacos e, a seguir a, compilação das “árvores genealógicas”, a teoria mais aceita é que em algum lugar da África Central o precursor simiano do vírus HIV tenha surgido em chimpanzés que tenham comido carne ou lutado com macacos infectados. Acredita-se que dois vírus de macacos estão envolvidos no processo, já que o vírus dos guenons (*Cercopithecus nictitans*) estão mais próximos da parte do genoma que contém o código para o envelope protéico enquanto que o vírus do mangabey (*Cercocebus torquatus*) apresentou mais similaridade em segmento diferente. As distâncias evolutivas entre o HIV do tipo 1, 2 e diferentes tipos de SIVs propõem que o HIV teria se originado a partir do SIV. Entretanto o HIV-1 possui uma seqüência homóloga altamente similar ao SIV<sub>cpz</sub> encontrado em chimpanzés do centro-oeste da África, enquanto o SIV<sub>sm</sub> encontrado em macacos sooty mangabey possuem maior similaridade com o HIV-2 (Lemey et al., 2003).

### 1.2.2 A Estrutura do HIV

O HIV é um vírus esférico, com um diâmetro de 95 – 120 nm. Seu exterior apresenta uma bicamada lipídica que é derivada da membrana celular do hospedeiro acrescentada com proteínas Env do vírus. Além disso, proteínas adicionais

incluindo as do complexo de histocompatibilidade principal derivadas da célula do hospedeiro podem ser encontradas na superfície viral. Tais proteínas podem servir como alvo para a neutralização de vários anticorpos. Nessa bicamada estão inseridas as glicoproteínas gp41 e gp120 (Fig. 1.1), que permitem a fusão viral através da interação com o receptor CD4 presente nas células T e com receptores celulares de quimiocina (CCR5). O capsídeo do vírus é formado principalmente por proteínas de matriz (p17) e por proteínas capsídicas (p24) (Fig. 1.1), estas organizadas em formato cônico. Já o nucleocapsídeo (camada mais interna do vírus) é composto por um genoma de duas fitas simples de RNA circundado por proteínas estruturais (Fig. 1.1). Nele também estão localizadas as enzimas virais transcriptase reversa, integrase e protease, essenciais para a replicação do HIV (Crandall, 1999a).

Quanto ao genoma, o HIV está estruturado segundo as características básicas de todos os retrovírus estudados. São compreendidos, principalmente, pelos genes *gag* (grupo antígeno seqüências específico) – que codificam as proteínas estruturais do capsídeo e nucleocapsídeo do vírus; *pol* (polimerase) – codificando a transcriptase reversa, integrase, protease e ribonuclease (para a replicação viral) e *env* (envelope) – que codifica as glicoproteínas do envoltório viral (Crandall, 1999a).

O HIV também apresenta alguns genes próprios. Esses genes adicionais (*tat*, *ver*, *vif*, *vpu*, *nef*, *vpr*) codificam proteínas acessórias e regulatórias do vírus. As proteínas acessórias são necessárias em diferentes passos da maturação da partícula viral, afetando a infectividade, a replicação e a montagem do vírus. As proteínas regulatórias são responsáveis por regular diferentes estágios da replicação viral (Vaishnav e Wong-Staal, 1991).

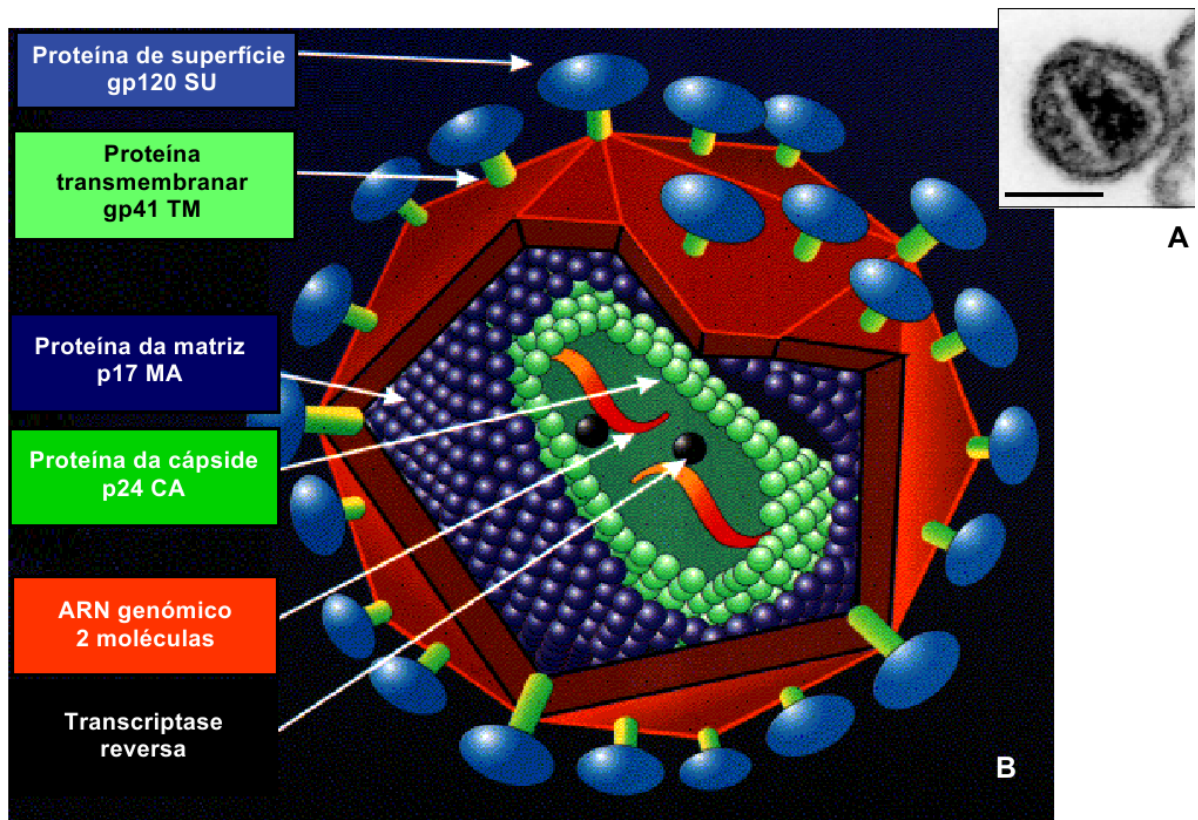


Figura 1.1: A – Microfotografia de uma partícula viral do HIV após purificação e concentração (barra 50 nm). B – Esquema da estrutura da partícula viral do HIV e localização de algumas das principais proteínas virais. Extraído de Pereira (2000).

Obs: Nesta figura a sigla ARN foi usada em substituição a RNA.

### 1.2.3 O ciclo viral

O ciclo do HIV é composto por sete etapas: entrada, síntese do DNA, integração, transcrição, processamento e transporte do RNA, síntese de proteínas e maturação.

- **Entrada** – Dentro do organismo do hospedeiro, o HIV irá interagir preferencialmente com as células imunológicas que possuam o marcador de superfície CD4 – linfócitos T e macrófagos – (Fig. 1.2). Atualmente, um modelo propõe que a ligação da gp120 à CD4 e depois a CCR5 permite que a molécula viral gp41 introduza sua cabeça hidrofóbica do terminal

amina na membrana celular iniciando a fusão do envelope viral com a célula (Fig. 1.3) (Wild et al., 1992).

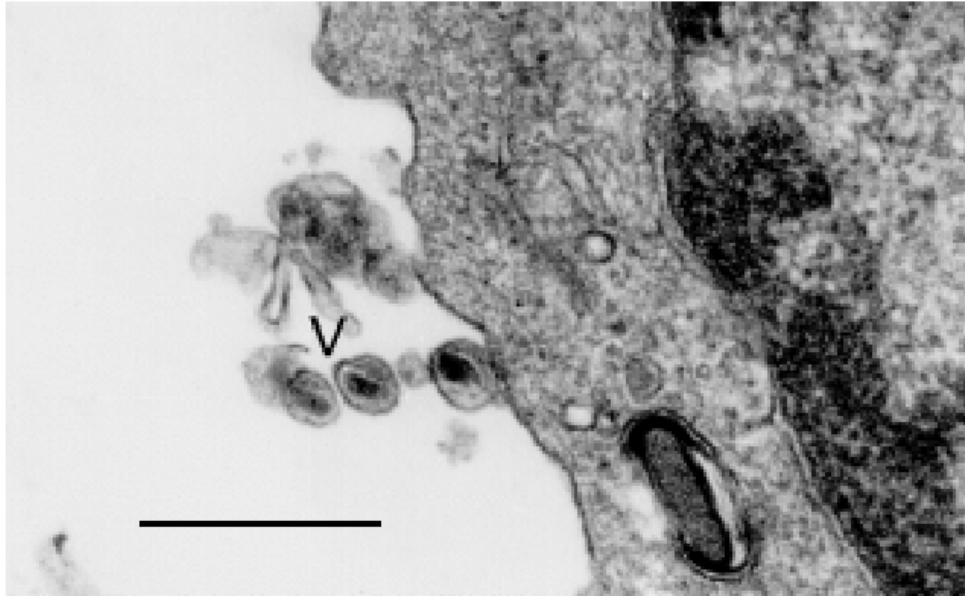


Figura 1.2: Interação do HIV-2 com um linfócito T. A cabeça da seta indica partículas virais. Barra 0,5  $\mu\text{m}$ . Extraído de Pereira (2000)

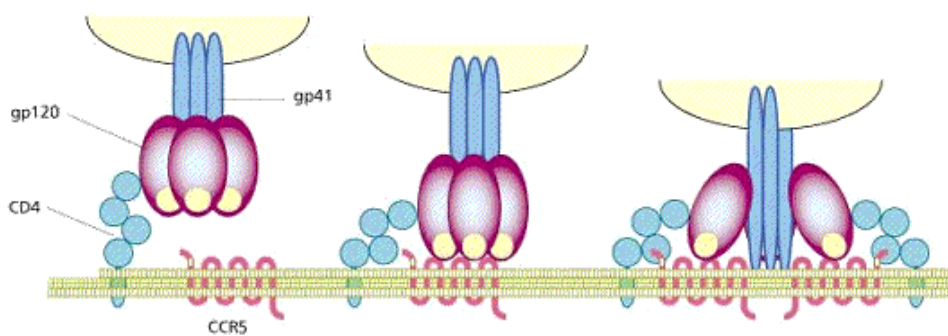


Figura 1.3: Representação esquemática do processo de fusão viral. Adaptado de [http://hivmedicine.aidsportugal.com/05\\_HAART\\_3.php](http://hivmedicine.aidsportugal.com/05_HAART_3.php).

A proteína Env do HIV tem algumas propriedades estruturais e funcionais que a distinguem das de outros retrovírus. Ela pode ser dividida nas regiões

variável (V) e constante (C). A variação na região V está relacionada com seleção de diversidade, onde algumas dessas seleções podem ser causadas por uma resposta ao sistema imune e outras podem refletir uma alteração no tipo de co-receptor utilizado no processo de fusão. O principal receptor do HIV são as moléculas de superfície CD4. Entretanto alguns tipos de HIV também requerem outras proteínas de superfície ou co-receptores. Todos os HIVs (e SIVs) usam o CD4 como receptor, já o tipo de co-receptores varia de acordo com o vírus. No caso do HIV-1, dois tipos principais podem ser identificados, os vírus que usam o co-receptor conhecido como CCR5 (vírus M-trópico) e os que usam o co-receptor conhecido como CXCR4 (vírus T-trópico) (Coakley et al., 2005). É importante perceber que ambos os vírus se replicam eficientemente em células T CD4<sup>+</sup> *in vivo*, e certamente tais células são o alvo principal para a infecção.

A maioria dos casos isolados no início da doença são de pacientes infectados com vírus M-trópico. Com o tempo, o vírus evolui para o tipo T-trópico (Indutor de Sincício – SI) (Crandall, 1999a). Essa mudança está associada a alterações localizadas na região V, principalmente numa região denominada de V3. Tais mudanças estão relacionadas com a habilidade do vírus SI infectar um maior número de células. O aparecimento dos SI, em geral, está correlacionado com o avanço da doença e conseqüentemente à morte do paciente, mas não se sabe ainda se o vírus SI é a causa do declínio do paciente ou se é uma conseqüência.

- **Síntese do DNA** – Após a fusão da membrana ocorre, no citosol da célula hospedeira, a transcrição reversa do RNA viral em DNA, catalisada pela enzima RT (Crandall, 1999a). A proteína acessória Vif é importante durante essa fase por facilitar os estágios iniciais da transcrição reversa ou os eventos de desnudamento viral e desmontagem do capsídeo (Gonçalves et al, 1996).

Duas características do processo de síntese do DNA retroviral são muito

importantes para a evolução biológica do vírus. Primeiro, o processo de transcrição reversa do HIV permite uma alta frequência de recombinação gênica, ou seja, uma troca genética pode ocorrer quando uma mesma célula é infectada por dois vírus distintos. Segundo, a alta taxa de erro do processo de transcrição reversa, se deve basicamente a dois fatores: os retrovírus utilizam três enzimas para sua replicação (transcriptase reversa, DNA polimerase e RNA polimerase), e a falta de um sistema de reparação das bases nucleotídicas colocadas incorretamente pela RT.

- **Integração** – Uma vez sintetizado, o DNA viral é transportado para o núcleo como parte de um complexo de pré-integração (PIC), que inclui as proteínas integrase, proteína de matriz (p17), RT e Vpr (Miller et al., 1997). A proteína do capsídeo (p24) não aparece como parte do PIC, embora ela contribua para a estrutura do PIC em outros retrovírus (Bowerman et al., 1989). A localização nuclear do PIC é dirigida pela proteína acessória Vpr, que também interfere no controle normal do ciclo celular, detendo o crescimento das células infectadas (Re et al., 1995).

Após ser transportado ativamente para o núcleo, o DNA viral é covalentemente integrado ao genoma do hospedeiro através da atividade catalítica da enzima integrase, que também penetra no núcleo, dando origem ao provírus. O provírus pode permanecer inativo por meses ou anos, com pouca ou nenhuma atividade protéica viral ou pode produzir novos vírions, levando a infecção ao estado produtivo (latência).

- **Transcrição** – Esta fase do ciclo de vida do vírus é iniciada com a síntese de transcritos de RNA mensageiro (mRNA) processados e não-processados, que são transportados para fora do núcleo para a tradução. Inicialmente, são sintetizados pequenos fragmentos de RNA que sofrem processamento e codificam as proteínas regulatórias Tat, Rev e Nef. A proteína Tat é essencial nesse processo, pois funciona como um ativador transcricional

que se liga ao transcrito nascente de RNA (Crandall, 1999a).

- **Processamento e Transporte do RNA** – Com relação ao RNA nascente, ele primeiro recebe uma cauda poli (A) (aproximadamente 200 nucleotídeos), em seguida sofre uma outra modificação importante para seu processamento, um processo denominado de *splicing*. Entretanto, somente uma parte das moléculas do RNA viral pode sofrer o *splicing* antes de ser transportado para o citoplasma.

Normalmente, o mRNA celular que não é processado é retido no núcleo onde futuramente pode ser processado ou degradado. Contudo, os transcritos de mRNA do HIV que contêm íntrons funcionais são necessários no citoplasma para a síntese e empacotamento da Gag e da Gag-Pol. Para solucionar esse problema, o mRNA viral possui uma seqüência conhecida como elemento de transporte constitutivo responsivo a Rev (RRE) que permite que mRNAs que não sofreram o processo de *splicing* possam migrar para o citoplasma. Tal processo é mediado pela proteína acessória Rev. Dessa forma, a Rev funciona como um interruptor entre a síntese inicial de mRNA completamente processado (codificando Tat, Rev e Nef) e a síntese posterior de mRNAs não-processados (codificando as proteínas Gag e Gag-Pol). É por isso que no início da transcrição, quando há ausência da proteína Rev, ou seja, todos os mRNA sofrem o processo de *splicing*, somente algumas proteínas são sintetizadas (Tat, Rev e Nef) (Crandall, 1999a).

- **Síntese de Proteínas** – As proteínas do vírion são produzidas utilizando a maquinaria celular da célula hospedeira, e a seguir juntam-se ao RNA genômico viral na membrana da célula e preparam-se para constituir a partícula.

No retículo endoplasmático rugoso (RER) são sintetizadas moléculas de Env e CD4 e a ligação prematura das duas no local pode inibir a trans-

locação de Env para a membrana celular ou a formação de um complexo TM-SU plenamente funcional (Crandall, 1999a). Assim, o receptor CD4 é designado a ser removido do RER pela proteína viral acessória Vpu, que se liga a moléculas de CD4 e sinaliza sua degradação (Margottin et al., 1998). Já a poliproteína Gag é sintetizada nos ribossomos a partir de mRNA não-processado. Um deslocamento traducional da poliproteína Gag resulta na geração de pequenas quantidades de proteínas precursoras Gag-Pol, que se associam com a Gag na membrana celular. Ainda no interior da célula hospedeira, aproximadamente 1500-2000 cópias da poliproteína Gag se auto-associam e se ligam a duas fitas de RNA que não sofreram processamento antes de ligar-se na membrana celular através do domínio de matriz da Gag, culminando na montagem do capsídeo imaturo (Bonhomme et al., 2003). Após montar-se, o capsídeo sai da célula por brotamento, adquirindo sua membrana, e assim liberando um vírion imaturo com capsídeo em forma de esfera.

- **Maturação** – Os componentes da partícula juntam-se na membrana da célula e esta começa a fabricar mini-bolhas que saem da célula levando consigo todas as proteínas virais e o RNA necessário para formar novas partículas. Ao se libertarem da célula, os vírus ainda estão imaturos e, nesta fase, são incapazes de infectar outras células, tendo que passar por um processo de amadurecimento para se tornarem infecciosos.

As células CD4 frequentemente não sobrevivem à invasão pelo HIV. Elas desintegram-se devido ao elevado número de vírus germinados ou porque o sistema imunológico do corpo reconhece as proteínas de envelope viral na membrana da célula e destrói as células danificadas. Como as células CD4 são uma parte essencial do sistema imunológico, a sua destruição pode provocar imunodeficiência profunda.

Depois do novo vírus deixar a célula (brotamento), uma outra enzima viral, a protease, cliva as poliproteínas Gag e Gag-Pol. As proteínas individuais



libertadas são remontadas para formar um vírus estruturado e maduro. Este vírus pode agora infectar outras células. A seguir são ilustradas a Figura 1.4 com as características gerais do ciclo de replicação do HIV-1; e a Tabela 1.1 que sumariza os principais genes do HIV, seus produtos e suas respectivas funções

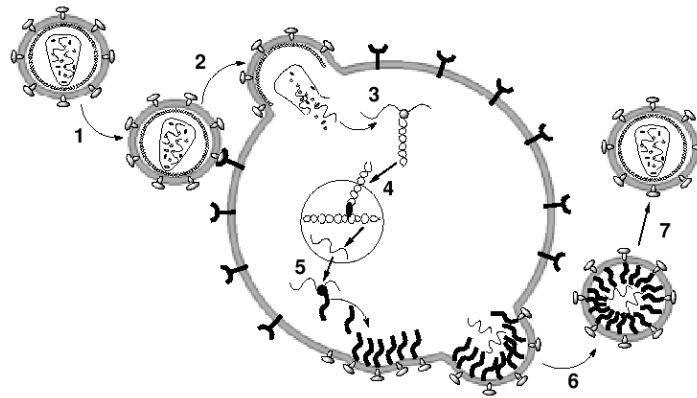


Figura 1.4: Representação do ciclo de replicação do HIV-1: 1. Ataque à célula hospedeira, 2. Processo de fusão, 3. Transcrição Reversa, 4. Integração, 5. Tradução, 6. Brotamento e 7. Maturação. Extraído de Alterman (2001)

Tabela 1.1: Genes e Proteínas do HIV

Gene	Produtos	Funções
<i>env</i>	SU (gp120) TM (gp41)	Ligação ao receptor Fusão e associação da membrana viral à celular
<i>pol</i>	RT (p66, p51)	Síntese de DNA
<i>vif</i>	p23 Vif	Acentua a infectividade do vírus, facilita os estágios iniciais da transcrição reversa e o evento do desnudamento viral
<i>vpr</i>	p15 Vpr	Atua no processo de condução do PIC ao núcleo. Também interfere no ciclo celular, detendo o crescimento das células infectadas
<i>pol</i>	IN (p32)	Integra o DNA viral ao genoma do hospedeiro
<i>tat</i>	p14 Tat	Acelera o processo inicial da transcrição
<i>rev</i>	p19 Rev	Transporte de RNAs não processados
<i>nef</i>	p27 Nef	É importante na ativação das células T e no aumento da replicação viral
<i>vpu</i>	p16 Vpu	regulação negativa de CD4
<i>gag</i>	MA (P17) CA (p24) p2 NC (p7) P1 P6	Matriz Capsídeo ? Nucleocapsídeo ? ?
<i>pol</i>	PR (p10)	Clivagem da Gag e Gag-pol

#### **1.2.4 Variação genética**

A variabilidade genética do HIV pode ser demonstrada em todo o genoma viral e, conseqüentemente, em todas as proteínas virais. Variantes de HIV são gerados randomicamente durante a replicação do vírus e então selecionados naturalmente após competirem com outras variantes para replicarem e escaparem da resposta imunológica do hospedeiro. Como a transcriptase reversa (RT) não apresenta um mecanismo de correção, isso faz com que ela seja a principal responsável pela variabilidade do HIV (Reid et al., 2005), permitindo substituições, deleções e inserções nucleotídicas nas duas cópias de RNA genômico viral. Entretanto não só a transcriptase reversa (RT), mas também a alta taxa de replicação viral tem grande influência no processo de variabilidade genética viral.

Os dois principais mecanismos responsáveis por produzir variantes virais são as mutações (incluindo mutações pontuais, deleções, inserções e duplicações) e as recombinações. Mesmo antes do uso de fármacos no tratamento anti-HIV, foi previsto que essas modificações poderiam dar origem a vírus resistentes (Levy et al., 1985).

A diversidade do HIV-1 é em torno de 6% num mesmo indivíduo podendo chegar a 50% entre indivíduos de diferentes regiões geográficas (Sabino, 1996). Com essa alta taxa de variação é de se esperar que o vírus apresente características diferentes.

### **1.3 Alvos Moleculares para a Terapia Anti-retroviral**

A identificação do HIV como agente causador da AIDS, o aumento da incidência de várias doenças em pacientes imunodeficientes e o impacto sócio-econômico dessa infecção, são fatores importantes no desenvolvimento de pesquisas por novos agentes anti-virais e novas modalidades de quimioterapia anti-viral. Nas últimas décadas, vários grupos têm buscado ativamente agentes químicos capazes de interferir em diferentes estágios da replicação viral (Fig. 1.5). Vários destes agentes já em uso clínico demonstram significativos resultados na inibição do HIV-1.

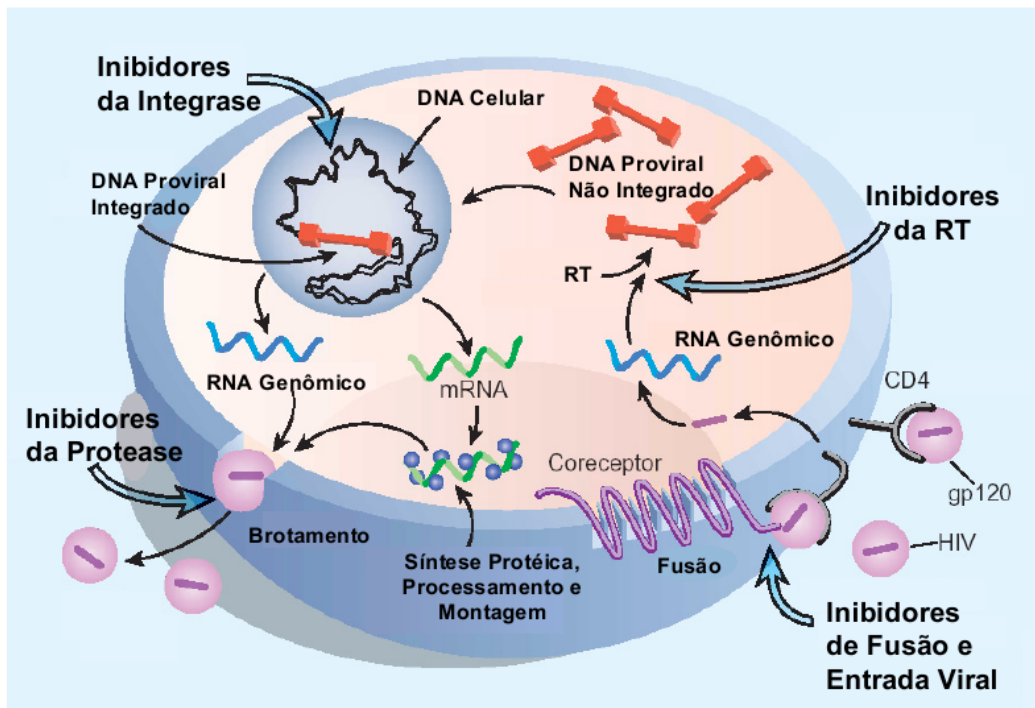


Figura 1.5: Representação esquemática dos alvos moleculares de bloqueio da replicação do HIV. Extraído de Batista (2006)

O desenvolvimento da HAART para o tratamento de indivíduos infectados pelo HIV encontra vários desafios tais como: a integração do vírus na maquinaria da célula-alvo, a alta taxa de mutação da transcriptase reversa do HIV-1, e a cronicidade da replicação viral. Estes representam obstáculos para o desenvolvimento de terapias anti-retrovirais específicas e efetivas. Estas terapias tornaram-se possíveis a partir de um maior conhecimento do vírus e da patogênese da doença, favorecendo a produção de uma grande variedade de agentes que possam ter um papel efetivo no tratamento de indivíduos infectados.

O agente anti-viral ideal deve interromper seletivamente alguma etapa específica e essencial do ciclo replicativo viral, sem afetar de forma significativa o metabolismo da célula hospedeira. A dependência da maquinaria da célula hospedeira para a replicação viral dificulta a identificação de inibidores que eliminem o agente patogênico sem causar efeitos colaterais. Os vírus, por serem geneticamente

simples, não possuem uma grande variedade de mecanismos para o desenvolvimento de resistência a agentes terapêuticos. Por outro lado, podem sofrer mutações muito rapidamente e em consequência, desenvolver resistência a um determinado agente.

Uma série de trabalhos e pesquisas tem sido realizada ao longo dos últimos 30 anos nos níveis biológico, bioquímico e estrutural do HIV. Diversas etapas do ciclo replicativo têm sido assim identificadas (como visto na seção anterior) como alvos potenciais para a terapia anti-retroviral. Entretanto, até 2005, somente 4 classes de inibidores tinham sido liberados para o uso clínico, estando disponíveis atualmente 27 drogas para tratamento da infecção pelo HIV-1. Nestas, existem 13 inibidores nucleosídeos/nucleotídeos da transcriptase reversa (NRTI), 3 inibidores não-nucleosídeos da transcriptase reversa (NNRTI), 10 inibidores da protease (PI) e 1 inibidor de fusão (Cohen, 2005).

### 1.3.1 Transcriptase reversa (RT)

O primeiro alvo molecular reconhecido e utilizado para o tratamento da AIDS foi a enzima RT.

- ***Inibidores Análogos Nucleosídeos da Transcriptase Reversa (NRTIs)*** – Os análogos nucleosídeos da RT são conhecidos também como inibidores nucleosídeos da RT. Tais inibidores atuam como substratos competitivos em relação aos substratos naturais como os desoxirribonucleotídeos (dNTPs). Esses nucleosídeos são formados por uma substituição do grupamento 3'-hidroxi por um NH<sub>3</sub>, hidrogênio ou outro grupo que quando incorporados ao DNA impedem a ligação 3'-5' fosfodiéster, ou seja, a incorporação de análogos nucleosídeos aborta a síntese de DNA proviral, dado que as pontes de fosfodiésteres não podem ser formadas.

Enquanto o AZT e D4T são análogos da timidina, o DDC e 3TC são da citosina. A combinação de AZT e D4T não faz sentido visto que ambos os fármacos competem pelas mesmas bases; isto se aplica também para o DDC e o 3TC. Já o DDI é um análogo da inosina, que é transformado

em didesoxiadenosina; e o abacavir é um da guanina. Há um elevado grau de resistência cruzada entre os análogos nucleosídeos. Eles são componentes importantes de quase todas as combinações terapêuticas. São potentes inibidores da replicação viral e são rapidamente absorvidos quando tomados por via oral. Contudo, têm um grande espectro de efeitos colaterais, entre eles mielotoxicidade, acidose láctica, polineuropatia e pancreatite. As queixas incluem fadiga, cefaleias e uma variedade de problemas gastrointestinais tais como desconforto abdominal, náusea, vômitos e diarreia (Galli, Ridolfo et al., 2002).

- ***Inibidores Não-Nucleosídeos da Transcriptase Reversa (NNRTIs)*** – Tal como para os análogos nucleosídeos, a enzima-alvo dos NNRTI's é a RT. Ligam-se diretamente e de forma não competitiva à enzima. Esses inibidores não-competitivos em relação à captação do substrato não afetam significativamente outros sistemas enzimáticos. Os compostos formados se ligam com pequenas diferenças à região hidrofóbica da subunidade p66, região da palma, e podem indiretamente afetar resíduos no sítio ativo da polimerase ou diminuir a mobilidade do subdomínio “polegar” da enzima (Kontorinis e Dieterich, 2003). Dados na literatura têm descrito que a grande importância desses agentes é sua relativa baixa toxicidade e alta especificidade para inibir a transcriptase reversa do HIV-1 (Loya et al., 2002), embora não apresentem boa atividade contra o HIV-2. Entretanto, sua eficácia pode ser comprometida pelo rápido desenvolvimento de resistência viral *in vitro*. A rápida emergência de vírus resistentes é resultado de alterações pontuais na RT, por substituição de resíduos, ocorrendo diminuição no efeito antiviral. Os quatro NNRTI's existentes são: nevirapina, delavirdina, efavirenz e intellence. A delavirdina, entretanto, já tem seu uso descontinuado devido a sua alta toxicidade.

### 1.3.2 Integrase

A integrase é uma das três enzimas chaves codificadas pelo gene *pol*. Ela está envolvida na integração do DNA proviral no genoma do hospedeiro. Embora provavelmente não exista integrase nas células humanas, a aprovação para o desenvolvimento de novos fármacos dentro desta classe está difícil e progressivamente mais lento que dos demais (Debyser et al., 2002). O inibidor de integrase S-1360 foi desenvolvido pela Shionogi/Glaxo apresenta dados preliminares promissores (Yoshinaga et al., 2002). Em estudos *in vitro*, o S-1360 é eficaz contra um leque de estirpes virais, incluindo mutantes resistentes a todos os NNRTIs e PIs. Parece haver sinergismo com zidovudina, lamivudina, nevirapina ou nelfinavir. Em estudos com animais (ratos e cães) o fármaco demonstrou pouca toxicidade. A molécula é muito pequena o que torna possível a formulação oral. O fármaco foi bem tolerado em voluntários saudáveis (Fujiwara, 2002).

### 1.3.3 Fusão e Entrada

Três etapas são cruciais para a entrada do HIV-1 na célula-alvo:

- A ligação ao CD4 (alvo dos inibidores de ligação);
- A ligação aos co-receptores (alvo dos antagonistas dos co-receptores)
- E a fusão do vírus nestas células (alvo dos inibidores da fusão).

Cada passo da entrada do HIV pode, teoricamente, ser inibido. Todas as três classes de fármacos, designados inibidores da ligação, antagonistas dos co-receptores e inibidores da fusão são atualmente denominadas como inibidores de entrada. Esses inibidores atuam no bloqueio das mudanças conformacionais da glicoproteína de envelope (env). Os estudos com inibidores de entrada apresentam grande adesão, devido à sua ação compreender os momentos mais precoces do ciclo replicativo viral.

- ***Inibidores de ligação***

O BMS-806 é um inibidor da ligação inicial que, sendo independente de

co-receptores, se liga específica e irreversivelmente à gp120, prevenindo assim a ligação do HIV ao receptor CD4. O entusiasmo por esta molécula foi um pouco abalado pela demonstração que diferentes estirpes de HIV possuem sensibilidades diferentes ao fármaco, indicando potencial para rápida resistência à sua atuação.

O Pro-542 é uma proteína solúvel, que também previne a ligação do HIV ao receptor CD4, por se ligar à proteína gp120. Os ensaios de fase I mostraram boa tolerância e diminuição da carga viral após administração de uma única dose (Jacobson et al., 2004).

- ***Antagonistas dos co-receptores***

O SCH-C é um antagonista do CCR5 com biodisponibilidade oral e potente atividade *in vitro* contra várias estirpes de HIV (Strizki et al., 2001). Em voluntários saudáveis os efeitos adversos, tais como arritmias, ocorreram, sobretudo, com as doses mais elevadas. Este problema parece estar ausente nas doses mais baixas. Num estudo piloto com 12 doentes com HIV, em que se utilizou SCH-C durante 10 dias, 10 doentes tiveram reduções na carga viral superiores a 0,5 log e dois superior a 1,0 log (Reynes et al., 2002). Este efeito persistiu mesmo alguns dias após terminar a terapêutica. No entanto, também já foram descritos mutantes de escape para o SCH-C, que possuem resistência cruzada com outros antagonistas do CCR5 (Riley et al., 2002).

O SCH-D foi liberado para a fase de teste em 2005, é referido como sendo mais potente e melhor tolerado do que o SCH-C, o que aumentaria a probabilidade de desenvolvimento clínico (Chen et al., 2002). No entanto, já há relatos de resistência, provavelmente por mutações no gene *env*.



O Pro-140 é um antagonista do CCR5 que atua como um anticorpo monoclonal. A idéia por detrás desta abordagem é a de que anticorpos geneticamente modificados podem ligar-se eles próprios aos co-receptores CCR5, impossibilitando o HIV de os utilizar. Em estudos com animais, a administração de dose única atingiu reduções significativas na carga viral que foram dose-dependentes (Franti et al., 2002). Uma preocupação com este tipo de fármaco é a de que o uso de um inibidor do CCR5 poderia propiciar a emergência de vírus HIV que comecem a usar alternativamente o co-receptor CXCR4 que, em alguns estudos, tem sido associado a uma progressão mais avançada da doença.

- ***Inibidores de Fusão***

O T-20 (Enfuvirtide, Fuzeon) é o protótipo dos inibidores de fusão. É um peptídeo relativamente grande constituído por 36 aminoácidos e, por isso necessita de ser administrado por injeção sub-cutânea. Liga-se a uma estrutura intermediária da proteína gp41, que surge durante a fusão do vírus com a célula alvo. Os primeiros estudos demonstraram que a simples adição de T-20 como “monoterapia” não seria muito benéfica (Lalezari et al., 2002). Há diversos efeitos colaterais e com o aumento da duração da terapêutica podem ocorrer graves lesões na região onde se aplicam as injeções, podendo resultar na interrupção do tratamento.

Doentes com uma carga viral bem controlada ou que ainda têm opções com a HAART “clássica” não necessitam do T-20. No entanto, o fármaco parece ser muito bom para terapêutica de resgate. Embora não exista ainda nenhum dado de estudos clínicos finalizados, análises estatísticas dos estudos em desenvolvimento mostraram ganhos de sobrevivência de 1,5 anos comparando com a HAART otimizada (Hornberger e et al, 2003). A combinação do T-20 a outros inibidores da entrada, que sejam sinérgicos entre si, e com os outros tipos de HAART, irá com certeza ser mais eficaz

na inibição do ciclo de vida viral que a HAART tradicional atualmente disponível, entretanto o preço continua a ser um aspecto importante dessa combinação, pois o tratamento custa quase o dobro com a adição do T-20.

O T-1249 foi o segundo inibidor de fusão desenvolvido. Entretanto os Laboratórios Gilead e Roche anunciaram, de forma independente, a descontinuação dos estudos de desenvolvimento do inibidor de fusão T-1249. Curiosamente os resultados obtidos com o T-1249 em estudos clínicos tinham sido muito bons, devido a isso analistas acreditam que a principal razão para essa atitude dos laboratórios Gilead e Roche/Trimeris seria a perspectiva de baixas vendas desses produtos (Vitória, 2004).

#### 1.3.4 Protease do HIV (PR)

A aspartil protease do HIV é indispensável para a maturação do vírus da AIDS, fazendo desta proteína um dos alvos para terapia anti-retroviral. Por se tratar do principal objeto de estudo desta dissertação, será abordada mais detalhadamente.

A PR é um homodímero de 99 aminoácidos, contendo, por cadeia, uma  $\alpha$ -hélice (normalmente do resíduo 87 ao 93) e nove folhas- $\beta$ , como pode ser visto na Figura 1.6.

As primeiras proteínas retrovirais que tiveram suas estruturas resolvidas por difração de raio-X foram a protease do RSV (Miller et al., 1989) e a protease do HIV-1 (Navia et al., 1989).

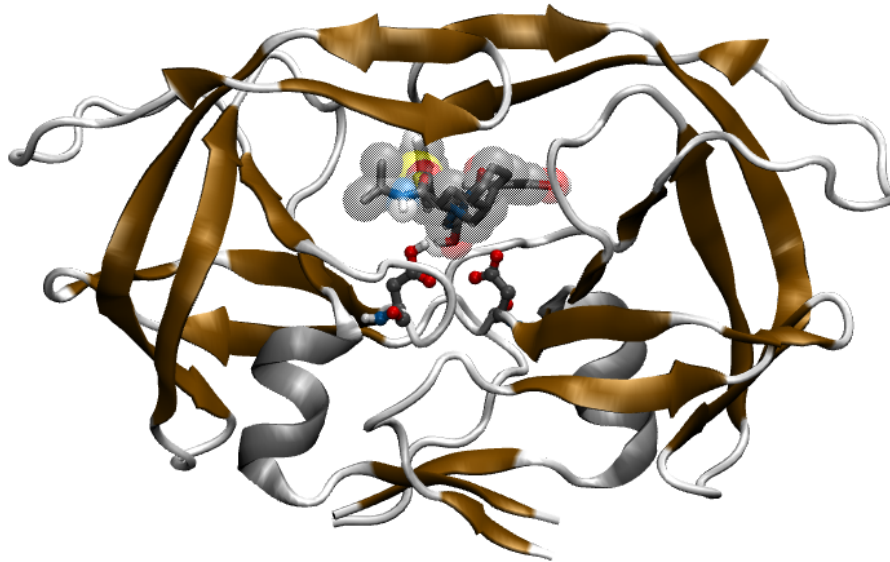


Figura 1.6: Ilustração em 3D da protease do HIV complexada com o inibidor nelfinavir (código 1OHR no PDB). As cadeias estão na representação de estrutura secundária. A cor cinza identifica as  $\alpha$ -hélices, a marrom as folhas- $\beta$  e branco os loops. Estão em destaque na proteína os dois ácidos aspárticos que classificam a protease. O nelfinavir segue acrescido de seus raios de Van der Waals em transparente onde o vermelho representa átomos de oxigênio, cinza carbono, azul nitrogênio e amarelo enxofre.

Do ponto de vista mecânico, as duas grandes alças/flaps (ida: do resíduo 43 ao 50; volta: do resíduo 51 ao 58), localizadas na parte superior da protease (vide Fig. 1.7) abrem-se, permitindo o encaixe da poliproteína em seu sítio ativo.

O sítio ativo é composto por uma tríade catalítica (Asp25-Thr26-Gly27) localizada numa alça interna entre as fitas- $\beta$  (Fig. 1.7.), cuja estrutura é mantida por um conjunto de ligações de hidrogênio característico das aspartil proteases, chamado “fireman’s grip” (Ingr et al., 2003). Os dois grupos carboxílicos dos resíduos Asp25 simetricamente relacionados são aproximadamente coplanares e formam um contato íntimo envolvendo seus átomos de oxigênio que estão localizados mais próximos (OD1).

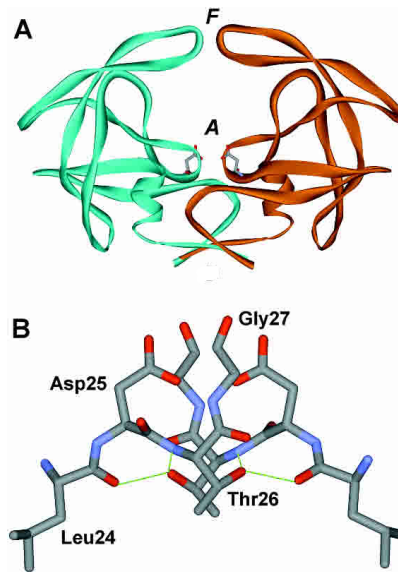


Figura 1.7: Em A: O desenho da cadeia principal da estrutura da protease do HIV-1; os flaps (F) e sítio ativo (A) são mostrados. Em B: Detalhe do sítio ativo da HIV-1 protease com o “fireman’s grip”; As pontes de hidrogênio são mostradas em linha verdes. (adaptado de Ingr et al. (2003))

A ligação de seqüências específicas do substrato peptídico aos subsítios da aspartil protease permite que ocorra a quebra da ligação peptídica de poliproteínas tornando-as funcionais. A protease do HIV-1 é responsável pelo processamento pós-traducional das poliproteínas Gag e Gag-pol para formar as proteínas estruturais e enzimas virais, ou seja, a protease desempenha um papel de “tesoura molecular” ao clivar diversas poliproteínas em proteínas menores e funcionais, de forma a gerar proteínas maduras. O sítio ativo da PR reconhece e cliva nove diferentes seqüências que podem ser visualizadas na Tabela 1.2 (Pearl e Taylor, 1987). O processamento incompleto das poliproteínas resulta na formação de partículas virais imaturas (não infecciosas).

Tabela 1.2: As nove seqüências peptídicas reconhecidas pela PR

Proteínas	Seqüência
MA/CA	SQNY/PIVQ
CA/p2	ARVL/AEAM
p2/NC	ATIM/MQRG
NC/p1	RQAN/FLGK.
p1/p6	PGNF/LQSR
TF/PR	SFNF/PQIT
PR/RT	TLNF/PISP
RT (interno)	AETE/YVDG
RT/IN	RKVL/FLDG

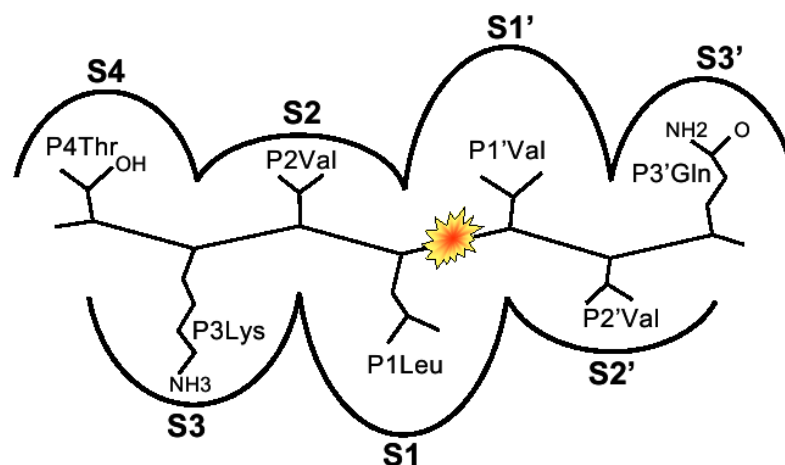
Dadas as conseqüências potencialmente catastróficas de uma ativação prematura, esta é retardada até que a partícula viral imatura esteja completamente formada e desligada da célula hospedeira. Assim, a protease é expressa como um produto de fusão Gag-Pol e é liberada dessa poliproteína precursora por um mecanismo autocatalítico, através da dimerização dos domínios da protease fusionada, quando a concentração dessas proteínas é aumentada, dado o encapsulamento pela partícula viral (Debouck et al., 1987).

Acredita-se também que as cisteínas 67 e 95, altamente conservadas entre as proteases de HIV, possuem um papel regulatório no processo de maturação do vírus. “Estudos mostraram (Davis et al., 1996) que a modificação dessas cisteínas, por vários agentes, pode levar a enzimas menos ativas”. Aparentemente essas cisteínas podem ser oxidadas no precursor Gag-Pol, ainda no citoplasma celular, prevenindo sua ativação prematura no citoplasma. Enzimas celulares, tais como tioltransferases, poderiam então reverter essas modificações quando o precursor se acumula na membrana celular, restaurando a atividade da enzima pouco antes do brotamento do vírion imaturo.

- ***Mecanismo de Catálise***

A PR do HIV, como visto anteriormente, se encarrega de quebrar longas seqüências peptídicas que emergem dos ribossomos da célula infectada pelo vírus, dando origem assim à proteínas que comporão um novo vírion a ser liberado (a matriz, o capsídeo e o nucleocapsídeo, incluindo as enzimas protease, transcriptase reversa e integrase).

De forma geral, a PR possui um sítio ativo composto de 7 sub-sítios (Fig. 1.8), divididos em 4 à direita e 3 à esquerda do ponto de clivagem do substrato, ficando usualmente o sub-sítio S1/S1' formado pelos resíduos ASP25/ASP25'.



S4 – Asp29, Asp30, Met46, Ile47, Gly48, Gln58, Leu76  
 S3 – Arg8', Leu23', Asp29, Gly48, Gly49, Ile50, Thr80', Pro81', Val82'  
 S2 – Ala28, Asp29, Asp30, Val32, Ile47, Gly49, Ile50, Leu76, Ile84  
 S1 – Arg8', Leu23', Asp25', Asp25', Gly27, Gly49, Ile50, Thr80', Pro81', Val82', Ile84'  
 S1' – Arg8, Leu23, Asp25, Asp25', Gly27', Gly49', Ile50', Thr80, Pro81, Val82, Ile84  
 S2' – Ala28', Asp29', Asp30', Val32', Ile47', Gly49', Ile50', Leu76', Ile84'  
 S3' – Arg8, Leu23, Asp29', Gly48', Gly49', Ile50', Thr80, Pro81, Val82

Figura 1.8: O esquema acima dá o tamanho relativo e a identificação de cada sub-sítio e seus respectivos resíduos associados. O sítio catalítico da aspartil protease de retrovírus clivaria a seqüência peptídica do substrato procurando por uma subsequência predominante de resíduos hidrofóbicos. Adaptado de TÖSZÉR et al. (2000).

Estudos cinéticos e estruturais indicaram que ambos os resíduos Asp25 e Asp124 (também conhecido como Asp25') no sítio ativo participam da catálise, em estados opostos de protonação (Hyland e et al, 1991). O mecanismo de ação atualmente aceito para as proteases de HIV está mostrado na Fig. 1.9, onde uma molécula de água interage, por ligações hidrogênio, com os resíduos catalíticos Asp25 e Asp25'.

Primeiramente o substrato se liga à enzima e forma o complexo enzima-substrato (Fig. 1.9A). Em seguida, a molécula de água ataca o carbono éster do substrato e o aspartato protonado transfere um próton para o substrato e o outro aspartato aceita um próton da água (Fig 1.9A e Fig 1.9B). A partir dessa etapa, o mecanismo não é muito bem compreendido e várias vias têm sido propostas. A mais aceita é que o aspartato protonado

pela água doe um próton para o intermediário e o outro aceita um próton do complexo intermediário (Fig. 1.9C). A ligação C-N do substrato é então quebrada (Fig. 1.9D).

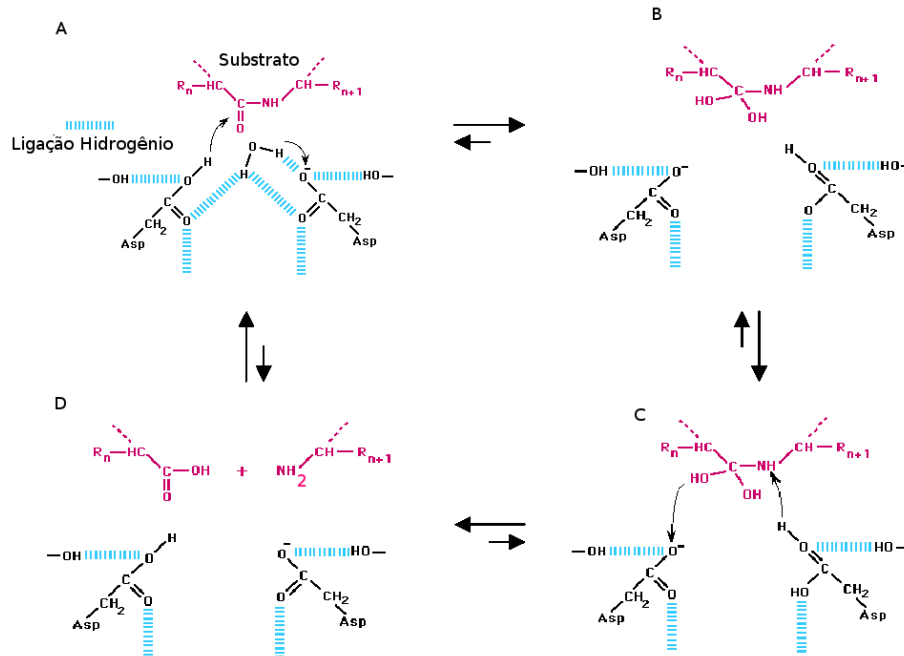


Figura 1.9: Provável mecanismo de ação das proteases de HIV envolvendo a clivagem da ligação peptídica. Extraído de Batista (2006).

As ligações entre as Ile50 e Ile50' e os inibidores também são intermediadas por uma molécula de água. Acredita-se que essa molécula de água não é necessária para a conservação da conformação ativa; contudo, tem papel importante no controle da abertura e fechamento das alças e no aumento da afinidade entre enzima e substrato. A presença de uma molécula de água estruturada nesta região tem sido observada em várias estruturas obtidas por difração de raios-X da protease do HIV-1 com inibidores (Wlodawer e Vondrasek, 1998) e também em estudos de RMN de complexos enzima-inibidor (Grzesiek et al.).

Esse mecanismo de interação característico tornou-se referência para o desenho de novos fármacos. Muitas classes de inibidores têm sido desenhadas



de forma a substituir as interações específicas dessas moléculas de água estruturadas com ambas as alças da protease do HIV-1, sendo dessa forma específicos às proteases retrovirais.

Estudos com as alças da protease por ressonância magnética (Freedberg et al., 2002) e por dinâmica molecular (Batista et al., 2006) indicam que elas são altamente móveis e podem adotar um grande número de conformações já na escala de ns.

#### 1.4 Inibidores da Protease do HIV

A palavra fármaco vem do grego *phármakon* e significa substância química empregada como medicamento. De forma geral, um fármaco é uma molécula que atua no organismo com o fim de produzir ou bloquear uma resposta biológica. Tal micromolécula interage no organismo com um determinado receptor. Esse receptor pode ser uma enzima ou proteína, responsável por algum efeito biológico que se deseja ativar ou bloquear. Os fármacos agem em uma região específica do receptor, conhecida como sítio ativo.

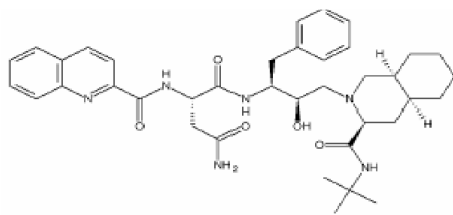
O reconhecimento molecular fármaco-receptor depende da complementaridade de características estruturais e físico-químicas das moléculas interagentes que influenciam diretamente a melhor energia de interação ligante-receptor. As características estruturais determinam os arranjos espaciais moleculares; já as características físico-químicas definem o grau de afinidade e especificidade do ligante pelo biorreceptor.

Em 1995, os inibidores de protease foram desenvolvidos e passaram a inibir drasticamente a replicação do HIV no organismo dos pacientes infectados. O primeiro inibidor aprovado pelo FDA para uso em humanos foi o saquinavir. Atualmente, existem outros 9 inibidores de protease aprovados como visto na tabela 1.3 (Hoffman e Mulcahy, 2006).

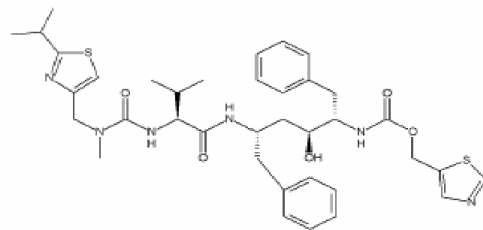
Tabela 1.3: **Inibidores de protease (PIs)**

<b>Nome Genérico</b>	<b>Nome Comercial</b>	<b>Sigla</b>
saquinavir	Invirase	SQV
ritonavir	Novir	RTV
indinavir	Crixivan	IDV
nelfinavir	Viracept	NFV
amprenavir	Agenerase	APV
lopinavir	Kaletra	LPV
atazanavir	Reyataz	ATV
fosamprenavir	Lexiva	FPV
duronavir	Prezista	DRV
tipranavir	Aptivus	–

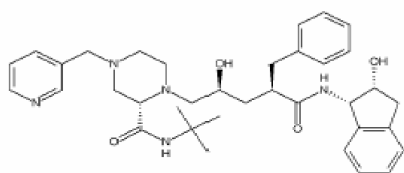
Os 9 primeiros inibidores são peptidomiméticos, isto é, simulam a estrutura e a conformação de alguns resíduos de aminoácidos do substrato com modificações no local onde seria a ligação peptídica. Tal ligação é substituída com o propósito de tornar os fármacos não hidrolisáveis pela enzima, inibindo competitivamente a ligação do substrato natural. Encontram-se destacadas na Figura 1.10 as estruturas químicas de oito inibidores de protease desenvolvidos e utilizados no tratamento anti-HIV: saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, lopinavir, atazanavir e tipranavir. O último inibidor é um fármaco não-peptidomimético e tem como intuito possibilitar uma melhor biodisponibilidade e farmacocinética. Esse composto simula a presença da molécula de água necessária para a catálise, estabilização e orientação do substrato no sítio ativo da protease. Na Fig. 1.11 pode ser visto um esquema geral das ligações hidrogênio que se formam entre os átomos da cadeia principal da enzima e do substrato/inibidor.



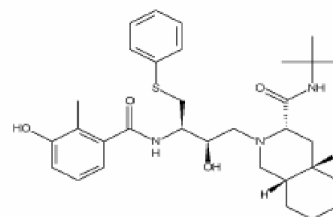
Saquinavir



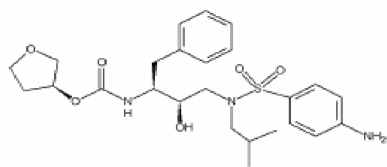
Ritonavir



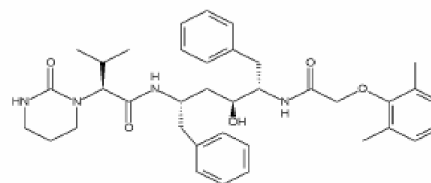
Indinavir



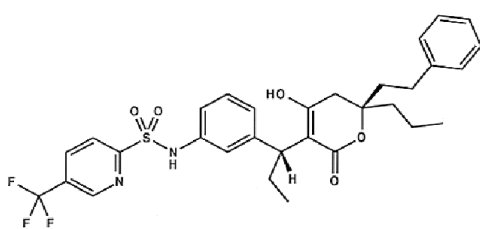
Nelfinavir



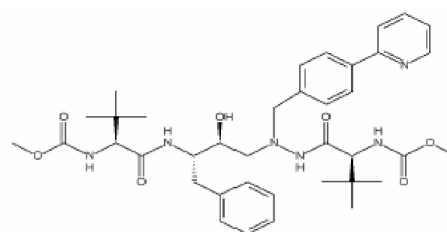
Amprenavir



Lopinavir



Tipranavir



Atazanavir

Figura 1.10: Representação de oito inibidores comerciais de HIV-1 PR atualmente disponíveis e aprovados pelo FDA.

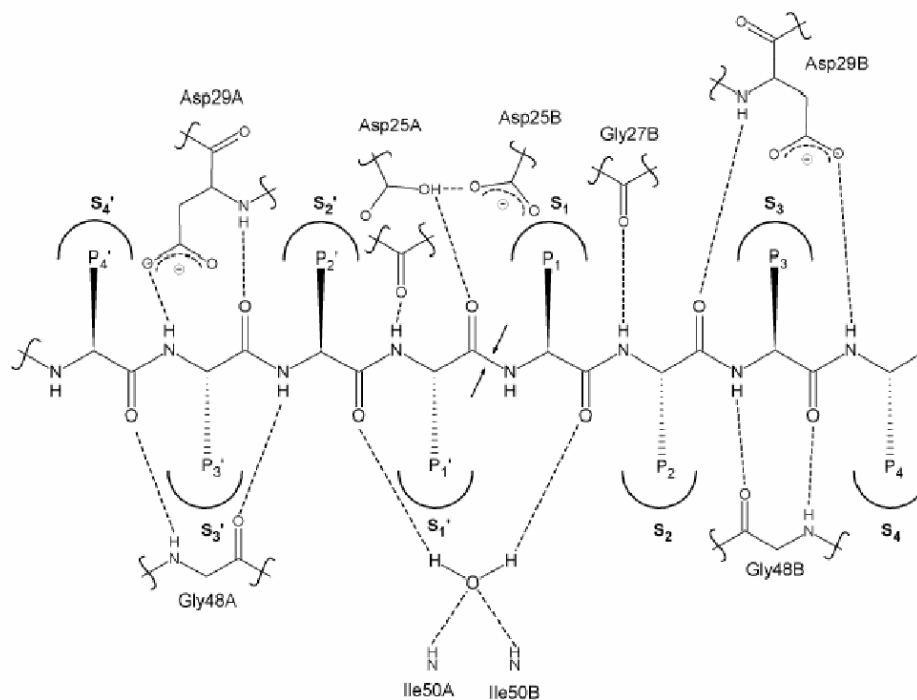


Figura 1.11: Esquema geral de ligações hidrogênio entre HIV-1 PR e um substrato representativo. As flechas indicam a ligação a ser clivada. Os resíduos na porção amino-terminal da ligação a ser clivada são chamados P1-Pn, e aqueles na porção carboxi-terminal, de P1'-Pn', de acordo com Schechter & Berger(1967). Os símbolos S4-S4' denotam os bolsões correspondentes a cada sub-sítio da enzima. Adaptado de Wlodawer e Erickson (1993).

Muito da pesquisa atual em drogas aplicadas ao combate do HIV concentra-se na inibição da atividade bioquímica de sua protease. O objetivo é projetar fármacos que venham a ligar-se à protease, covalentemente ou não, ajustando-se ao sítio ativo de forma a satisfazer ao máximo a complementaridade estrutural e eletrostática entre ambos.

A HAART representou um grande avanço no tratamento de indivíduos infectados pelo HIV-1, reduzindo o número de partículas de HIV-1 na corrente sanguínea (quando avaliado pela carga viral) e elevando a contagem de células T. Atualmente a combinação de dois anti-RT e um anti-PR, ficou conhecida popularmente como coquetel “anti-AIDS”. Esta nova abordagem no tratamento dos portadores de HIV pode ser justificada pelo sucesso que tal terapia obteve, e a Fig. 1.12 mostra um resumo disto. Porém, este sucesso só pode ser mantido à custa do contínuo surgimento de novos inibidores, tanto de protease quanto de RT, uma vez que a própria seleção natural se encarrega de fazer surgir vírus cada vez mais resistentes aos fármacos anti-HIV presentes no mercado.

É importante citar que este tratamento, embora reduza consideravelmente a infecção viral, não leva à cura da doença. Os pacientes submetidos ao uso do HAART, mesmo apresentando baixos níveis de vírus circulante, podem ainda transmitir o HIV-1 por meio de relações sexuais desprotegidas e outras vias.

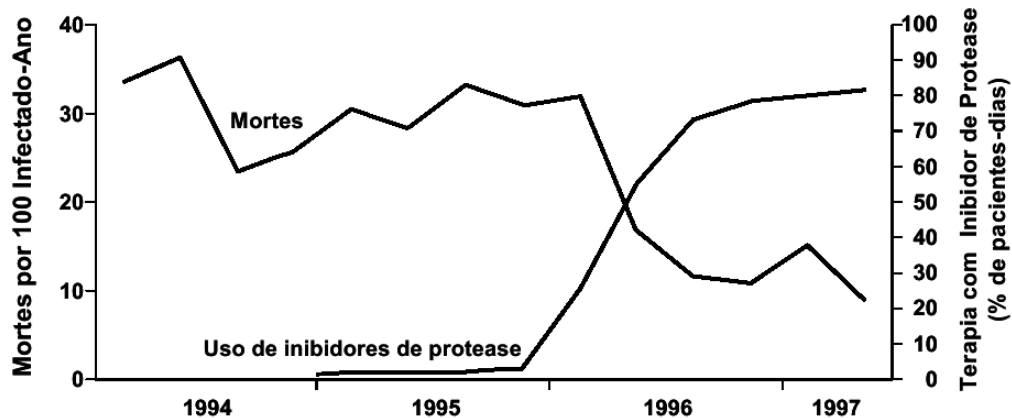


Figura 1.12: Decréscimo na mortalidade dos casos de AIDS associado à terapia com inibidores de HIV tipo 1 para indivíduos norte-americanos. Adaptado de Palella et al. (1998).

A modelagem computacional tem sido usada para identificar possíveis compostos que se ligam especificamente na região de encaixe de substrato da protease. O desenho racional de fármacos tem nos inibidores de PR um bem sucedido exemplo e ilustra significativamente a aplicação do estudo que correlaciona a função de uma proteína com sua estrutura (Roberts et al., 1990). Uma das principais características bioquímicas dos fármacos inibidores é seu caráter hidrofóbico, que aliado à baixa especificidade da PR implica em uma posologia muito alta, contínua e sistemática para que o medicamento tenha a eficiência desejada. Decorrendo deste tratamento efeitos colaterais tais como diarreia, náusea, desconforto gastrointestinal e erupções de sangue na pele (Deeks et al., 1997). Portanto, nem sempre um potente inibidor, tanto teoricamente quanto *in vitro*, pode vir a ser um possível candidato a medicamento.

### 1.4.1 Interações Receptor-Ligante

A termodinâmica descreve as relações entre as várias formas de energia e de como a energia afeta a matéria ao nível macroscópico. Além disso, é capaz de determinar se um processo físico pode ocorrer e quais processos ocorrem espontaneamente. Com isso, ela é essencial na compreensão de como macromoléculas biológicas adquirem a sua conformação nativa, de como as redes metabólicas funcionam, como as enzimas atuam como catalisadores e como moléculas conseguem atravessar membranas biológicas.

A variação de entalpia ( $\Delta H$ ) é uma propriedade físico-química importante para descrever sistemas biológicos:

$$\Delta H = \Delta U + P\Delta V \quad (1.1)$$

Onde:

$\Delta U$  → Variação de energia interna

$P$  → Pressão constante

$\Delta V$  → Variação de Volume

Em geral, nos processos bioquímicos a variação de entalpia se reduz a:

$$\Delta H = q \quad (1.2)$$

Onde  $q$  é calor gerado ou calor absorvido no processo. Processos que liberam energia ( $\Delta H < 0$ ) tendem a ser mais espontâneos. Mas esse não é o único critério para determinar a espontaneidade de reações. Existem processos que, apesar de serem endotérmicos ( $\Delta H > 0$ ), são espontâneos. Outro critério que determina a espontaneidade das reações chama-se de entropia ( $S$ ). A entropia está associada ao estado de desordem do sistema. Quanto maior a desordem, maior a entropia. Assim, se a variação de entropia ( $S_f - S_i$ ) é maior que zero implica em um aumento da desordem do sistema.

Geralmente nos processos termodinâmicos naturais, a energia se “degrada” (liberar calor  $\rightarrow \Delta H < 0$ ) de uma forma organizada para uma desordenada (aumentar a desordem  $\rightarrow \Delta S > 0$ ). Através da combinação de entropia e entalpia obtém-se uma outra grandeza termodinâmica que determina se um processo é espontâneo ou não. Esta grandeza é chamada de Energia Livre de Gibbs ( $G$ ). O valor da energia livre de Gibbs é dado pela equação:

$$\Delta G = \Delta H - T\Delta S \quad (1.3)$$

Onde:

$\Delta G \rightarrow$  Variação da energia livre de Gibbs

$\Delta H \rightarrow$  Variação de entalpia

$\Delta S \rightarrow$  Variação de entropia

$T \rightarrow$  Temperatura absoluta (em Kelvin)

Se  $\Delta G < 0$ , o processo é espontâneo (exergônico), para  $\Delta G = 0$  temos processos em equilíbrio termodinâmico, e se  $\Delta G > 0$  o processo é não-espontâneo (endergônico).

A função das enzimas é acelerar a taxa com que as reações químicas exergônicas ocorrem (Fig. 1.13), porém uma enzima sozinha não pode promover reações químicas que tenham  $\Delta G$  positivo.

Porém a maioria das ações metabólicas, em sistemas vivos, são constituídas por seqüências de reações espontâneas e não-espontâneas. Este fenômeno constitui a base termodinâmica para a ação das vias metabólicas e só é possível devido à aditividade das variações de energia livre que permite que uma reação endergônica possa ser viabilizada por uma reação exergônica.





Se  $\Delta G1 + \Delta G2 < 0$



Reações acopladas através do intermediário D

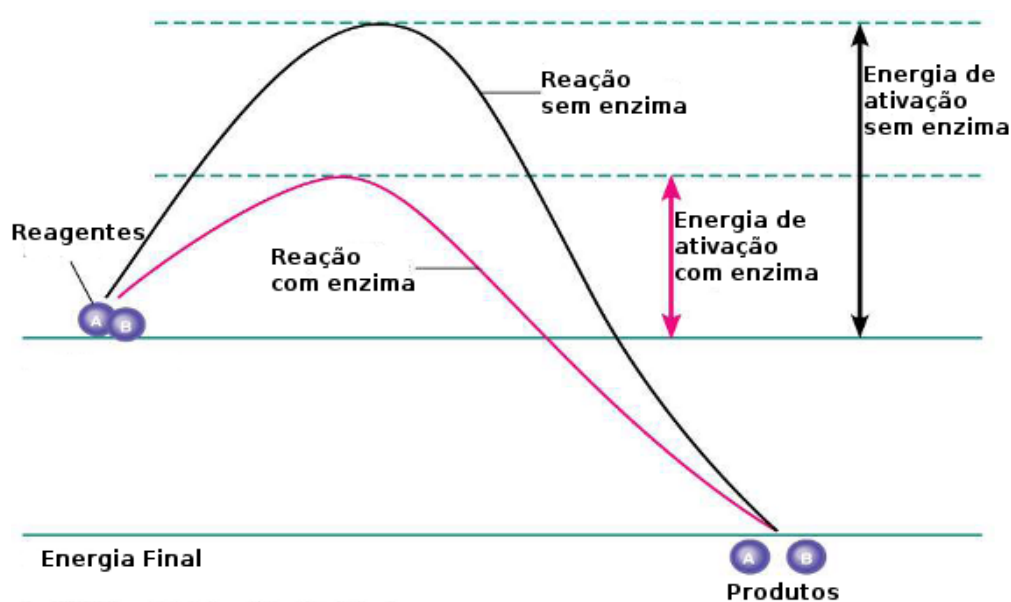


Figura 1.13: Representação esquemática do processo de reação química com e sem enzima.

As interações de van der Waals, eletrostáticas e ligações de hidrogênio são importantes para a estabilização dos complexos biomoleculares. As interações de van der Waals são importantes na atração de moléculas apolares, devido principalmente às forças de dispersão, enquanto as forças eletrostáticas resultam da interação entre cargas e dependem da constante dielétrica do meio e da distância entre suas cargas. O solvente também possui um papel importante nas interações

receptor-ligante; moléculas de água conservadas em sítios de ligação de proteínas podem modificar características estruturais destes sítios, além de poderem mediar ligações de hidrogênio entre o ligante e a molécula receptora.

#### **1.4.2 Resistência aos Inibidores de Protease**

Um obstáculo ao uso eficaz numa terapia de longo prazo com inibidores de PR é o surgimento e desenvolvimento de resistência e conseqüente perda da atividade do fármaco. A alta taxa de replicação do vírus ( $10^{10}$  partículas virais/dia), aliada às altas taxas de substituições incorretas em posições de nucleotídeos durante a transcrição reversa ( $10^{-4}$ /nt) devido à falta de uma atividade de correção da enzima, e a variabilidade genética do vírus, têm levado à identificação de variantes de HIV que exibem susceptibilidade variada aos inibidores.

Algumas mutações selecionadas durante a terapia com anti-retrovirais conferem resistência fenotípica mensurável, e por isso são chamadas mutações “primárias”; enquanto outras compensam a diminuição replicativa do vírus e podem ser associadas à resistência a fármacos, ou causam resistência somente quando em combinação com outras mutações sendo “secundárias”. No entanto, quais mutações são primárias e quais são secundárias não é estritamente definido e algumas mutações podem ser consideradas primárias para um inibidor e secundária para outros. Além disso algumas mutações podem exercer um papel dinâmico (quando se apresentam isoladas) e compensatório (quando estão acopladas a outra) na restauração da atividade catalítica da enzima mutada.

### **1.5 Vacinas**

Novas ferramentas moleculares em virologia e em imunologia, novos sistemas de expressão genética, descobertas recentes da infecção e patogênese do HIV e estudos promissores de hipotéticas vacinas em animais proporcionaram argumentos suficientes para o desenvolvimento de uma vacina para a AIDS, segura e eficaz. No entanto, constatou-se que tal eficácia não foi, ainda, possível e que a complexa

biologia do HIV torna a tarefa ainda mais complicada.

Apesar de existirem alguns determinantes antigênicos acessíveis, não se verifica, no entanto, uma resposta imunológica eficiente. Tal fato é devido à já citada extraordinária capacidade do vírus para ludibriar o sistema imunológico por meio das inúmeras mutações que sofre.

Existem muitos parâmetros inerentes à resposta imunitária induzida por uma vacina que determinam a sua maior ou menor capacidade para proteger o hospedeiro da infecção e da doença. Entre estes encontram-se a especificidade, as propriedades funcionais e a magnitude. No entanto, o fator mais relevante é o tempo. A duração da resposta imunitária relacionada à infecção inicial e disseminação do vírus é particularmente importante no caso da infecção pelo HIV. De fato, quanto maior for o grau de replicação viral no hospedeiro menores serão as probabilidades de o sistema imunológico reagir contra as adversidades promovidas.

Em se tratando do HIV-1, tem sido difícil de produzir vacinas capazes de provocar respostas de anticorpos que neutralizem a grande variedade de estirpes de vírus encontrados nos pacientes. Contudo, as respostas dos linfócitos T citotóxicos podem ser capazes de fazer frente a uma grande variedade de estirpes de vírus e parecem ajudar a controlar o HIV-1 em pessoas infectadas e, por isso, as atenções hoje centram-se nas vacinas capazes de produzir respostas pelos linfócitos citotóxicos. De maneira geral as vacinas anti-HIV estimulam os linfócitos B a produzir anticorpos que se liguem ao vírus, impedindo sua penetração nas células T CD4+, e estimulam o linfócito T CD8 a produzir substâncias (citocinas) que destroem células já infectadas pelo HIV.

A eficácia da vacina diminui ou aumenta de acordo com o nível em que se encontra a infecção, ou seja, a capacidade para produzir anticorpos específicos após uma infecção avançada se torna pior. É importante que a vacinação seja dada tão cedo quanto possível no decurso de uma infecção. Por outro lado, em doentes que se apresentem com uma situação de imunodeficiência avançada, é preferível adiar a vacinação até que seja iniciada a terapêutica anti-retroviral para maximizar o

efeito da vacina.

O objetivo a longo prazo é desenvolver uma vacina que seja 100% eficaz e que proteja todas as pessoas da infecção. Entretanto, mesmo uma vacina parcialmente eficaz poderia reduzir significativamente a quantidade de pessoas que contraem a infecção pelo HIV, evitando a contaminação de várias pessoas, o que já causaria impacto positivo no controle da epidemia da AIDS.

## 1.6 Subtipos do HIV-1

Desde a identificação do HIV-1 e HIV-2 como agentes etiológicos da AIDS, ambos os vírus têm sido intensivamente caracterizados por meio do seqüenciamento de seus genomas. Dada sua alta taxa de mutação, cepas de HIV circulando globalmente exibem um extraordinário grau de diversidade, que pode influenciar aspectos da sua biologia, tais como infectividade, transmissibilidade e imunogenicidade. Seqüências derivadas de cepas de HIV-1 e 2 têm sido historicamente classificadas em grupos e subtipos com base em sua relação filogenética. Contudo, enquanto o HIV-1 é o principal responsável pela pandemia da AIDS no mundo, o HIV-2 está associado a uma infecção endêmica no oeste da África (Lemey et al., 2003). Daí o objeto de estudo dessa dissertação ser uma enzima do HIV-1.

A partir das seqüências dos genes *env* e *gag*, o HIV-1 foi dividido, em 1993, em dois grupos: M (Major) e O (Outlier). Posteriormente, um novo “grupo” de HIV-1 foi identificado com características filogenéticas intermediárias entre os subtipos principais (grupo M) e o grupo O, e foi denominado de grupo N (New) (Crandall, 1999b). Os subtipos de HIV-1 formam um conjunto de seqüências dentro do grupo M que são classificadas de acordo com a análise filogenética, eles se diferenciam em 30% no gene *env* e 15% no gene *gag*.

Dentro do grupo M (aproximadamente 90% dos casos de HIV/AIDS no mundo), há pelo menos 9 subtipos distintos (A, B, C, D, F, G, H, J e K) e mais de 30 formas recombinantes circulantes (CRF) já foram identificadas. Entre alguns subtipos também pode ser identificada uma estrutura filogenética que permite a

separação em sub-subtipos como, por exemplo, o subtipo F que pode ser subdivido em F1 e F2 (Triques et al., 1999).

O grupo O tem chamado atenção devido à sua facilidade de escapar dos testes diagnósticos. Embora acredite-se que ele possua vários subtipos, assim como o grupo M, a distinção genética de tais subtipos tem sido muito mais difícil.

Dos 33 milhões estimados de pessoas infectadas com o HIV no mundo, mais de 22 milhões estão na África, principalmente na África Sub-Saariana. No ano de 2007 foi estimado que 75% das novas infecções que ocorreram em todo mundo aconteceram na África Sub-Saariana e 15% na Ásia (Fig. 1.14), ou seja, a África e a Ásia juntas foram responsáveis por 90% das novas infecções do mundo (Unaid, 2007).

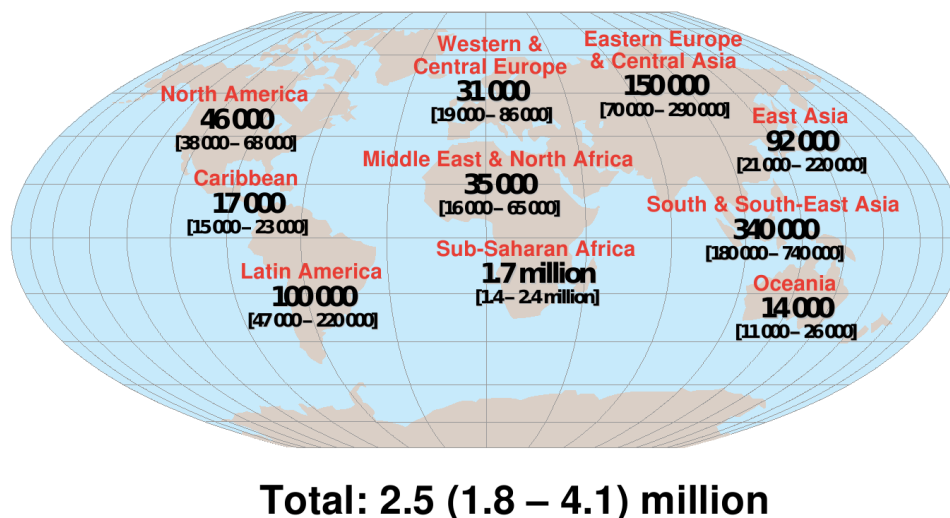


Figura 1.14: Estimativa do número de pessoas infectadas com HIV em 2007. Extraído de Unaid (2007).

Os subtipos do HIV-1 prevalentes na África e na Ásia não são os mesmos que prevalecem na América do Norte e na Europa Ocidental. Nessas regiões desenvolvidas, o subtipo B é responsável pela vasta maioria das infecções por HIV, enquanto

que na África e Ásia os subtipos A e C são os mais encontrados. Os subtipos não-B, em 2004, eram responsáveis por 90% das infecções do mundo; o subtipo A por 12%, o C 50%, o D 3%, o G 6%, e os demais subtipos juntos totalizaram 1% e as formas recombinantes juntas eram responsáveis por 18% (Helemar et al., 2006). No Brasil, o subtipo B é o principal (50% das infecções), mas o F (18%), C (30%) e o D são também encontrados (Soares et al., 2003). Todos os inibidores da PR disponíveis até o momento foram desenvolvidos para o subtipo B, ou seja, eles encontram-se altamente direcionados e moldados ao sítio ativo desse subtipo. As proteases dos subtipos A e C tiveram suas constantes de inibição ( $k_i$ ) para os inibidores ritonavir, saquinavir, nelfinavir e indinavir comparadas ao subtipo B (Velazquez-Campoy et al., 2001). O valor de  $k_i$  foi menor para todos os inibidores no subtipo B, demonstrando uma melhor inibição desse subtipo comparado à do A e C. O  $k_i$  dos inibidores saquinavir e nelfinavir foi 2 a 2,5 vezes maior nos subtipos A e C comparados ao do B.

O quadro se complica ainda mais quando se considera que o padrão de mutações, gerado por pressão dos inibidores, pode ser diferente entre os vários subtipos. Por exemplo, a posição 36 do subtipo B contém uma metionina, sendo que a mutação M36I aparece como consequência de tratamento anti-retroviral. Já nos subtipos não-B, a mesma posição contém uma isoleucina na quase totalidade das seqüências isoladas. Um outro exemplo que pode ser citado é que variantes do subtipo F, comuns no Brasil, não desenvolvem as mutações I84V e L90M, comumente encontradas em subtipos B (Caride et al., 2001). Esse padrão de resistências secundárias que são consideradas polimórficas em proteases não-B poderia levar a uma falha terapêutica prematura quando comparado com o subtipo B (Perno et al., 2001).

## 1.7 Aspectos Gerais da Mutação D30N

A mutação D30N consiste na substituição de um único aminoácido, o asparato, no códon de posição 30 do gene da protease pelo aminoácido asparagina. Esta

substituição ocorre próxima ao sítio ativo da enzima e é uma mutação primária de resistência exclusiva ao nelfinavir. Já a mutação L90M é a substituição da lisina no códon de posição 90 do gene da protease por uma metionina. Esta mutação ocorre distante do sítio ativo e não é exclusiva ao nelfinavir.

Embora as mutações D30N e L90M sejam de resistência ao nelfinavir, a D30N ocorre mais comumente no subtipo B, enquanto L90M ocorre mais no subtipo C (Gonzalez et al., 2004). Além disso a mutação D30N, quando ocorre no subtipo C, geralmente está associada à mutação N88D (substituição da asparagina no códon 88 por um ácido aspártico).

Experimentos de mutagênese sítio-dirigida provaram que a obtenção *in vitro* do vírus CD30N não foi possível, exceto depois de acumulada a mutação espontânea N83T (substituição da asparagina na posição 83 por treonina) (Gonzalez et al., 2004). Esta funciona como uma mutação compensatória, pois é selecionada para restaurar a viabilidade do clone CD30N. As características fenotípicas desse mutantes podem ser vistas nas tabelas 1.4 e 1.5.

A tabela 1.4 nos fornece uma quantificação do número de vezes que o mutante é mais resistente ao nelfinavir em relação ao controle. Dentre os clones o que possui maior resistência é o BD30N (42 vezes mais resistente do que o subtipo B) (Gonzalez et al., 2004). No subtipo C, o mutante CD30N/N88D tem menor resistência do que o clone CD30N/N83T, entretanto o mutante CD30N/N83T somente é observado *in vitro*. Por outro lado, nota-se que a capacidade replicativa é diminuída pela mutação D30N. Observa-se ainda que a capacidade replicativa perdida com ela é parcialmente recuperada pela mutação N88D no subtipo C. Contudo, os vírus do subtipo B são menos sensíveis aos efeitos da D30N que os do subtipo C, visto que a sua capacidade replicativa não é tão prejudicada.

As vantagens conferidas à PR devido às mutações dependem tanto do impacto da mutação sobre a capacidade replicativa como no ganho da resistência ao inibidor. Analisando a tabela 1.4, nota-se que a relação custo-benefício, no subtipo B, é melhor que no subtipo C quando estes sofrem a mutação D30N.

Tabela 1.4: **Características Fenotípicas dos Clones B e C relativos à inibição pelo nelfinavir** (Adaptado de Gonzalez et al. (2004))

Clone	IC <sub>50</sub> ( $\mu$ M) $\pm$ DP (resistência)	Cap. Rep. Relativa
B tipo selvagem	41.75 $\pm$ 9	100%
BD30N	1743 $\pm$ 87.6 (41,7)	86,5%
BL90M	178 $\pm$ 6.8 (4,28)	91,9%
CD30N/N83T	1447.5 $\pm$ 50 (34,6)	62,0%
CD30N/N88D	636.5 $\pm$ 21 (15,2)	72,4%
CL90M	241 $\pm$ 82 (5,7)	78,9%

Características fenotípicas observadas *in vitro* para os clones mutantes do HIV-1. IC<sub>50</sub> significa a quantidade de fármaco necessária para inibir 50% do número máximo de clones que podem ser inibidos. Quanto mais “resistente” for o clone, mais fármaco será necessário para inibi-lo e maior será o IC<sub>50</sub>. Ao lado do IC<sub>50</sub> temos o desvio padrão e entre parênteses podem ser visualizados o número de vezes que o IC<sub>50</sub> do nelfinavir foi maior em relação ao controle. Na última coluna pode-se observar a capacidade replicativa dos clones mutantes em relação ao seu respectivo tipo selvagem.

Na tabela 1.5, observa-se que o clone CD30N/N83T é sempre menos adaptado em relação aos outros e que a mutação D30N também torna o vírus do subtipo B menos adaptado. Entretanto, é bom lembrar que isso ocorre na ausência do inibidor.

Observando as duas tabelas percebe-se ainda que os vírus CD30N/N83T e CD30N/N88D são menos adaptados e com menor capacidade de replicação, do que o vírus selvagem e do que o mutante CL90M. Assim, as diferentes prevalências da D30N e L90M em pacientes que não respondem bem ao tratamento com nelfinavir nos subtipos B e C são explicadas, ou seja, determinantes genotípicos específicos do subtipo C dirigem o genótipo de resistência ao nelfinavir dos vírus do subtipo C preferencialmente para a aquisição da L90M porque ela tem um menor “custo” para a atividade viral e não requer nenhuma mutação compensatória.



Tabela 1.5: **Ensaio de competição *in vitro* dos clones mutantes do HIV-1.** As duplas foram colocadas num mesmo meio de cultura e após um tempo foi avaliado qual estava em maior quantidade (Adaptado de Gonzalez et al. (2004))

<b>Ensaio de competição</b>	
<b>BD30N</b> em competição com	Resultado da competição
CD30N/N83T	BD30N é mais adaptado
BL90M	BD30N é menos adaptado
B	BD30N é menos adaptado
<b>CD30N/N83T</b> em competição com	Resultado da competição
BD30N	CD30N/N83T é menos adaptado
CL90M	CD30N/N83T é menos adaptado
C	CD30N/N83T é menos adaptado

Estes dados sugerem que a mutação D30N possa ser deletéria para o subtipo C levando a uma baixa taxa replicativa e tornando o vírus incapaz de estabelecer uma infecção produtiva. Isso pode ser um fenótipo específico do subtipo C e não estar generalizado para todos os subtipos não-B. Há indícios de que duas vias diferentes são usadas por subtipo B e não-B na aquisição de mutações de resistência a fármacos e para aumentar a capacidade de replicação. Este conhecimento pode influenciar estratégias terapêuticas futuras e na interpretação de genótipos de resistência onde estas variantes prevalecem.

## 1.8 Dinâmica Molecular & Protease do HIV-1

Muitos trabalhos em DM têm estudado a protease do HIV-1 com a finalidade de entender o comportamento dessa enzima a nível atômico. Com objetivo de familiarizar o leitor com os estudos já realizados serão descritos, em linhas gerais, algumas das linhas seguidas por alguns pesquisadores.

Em um desses manuscritos foi elucidado como uma enzima simétrica se ajusta para reconhecer seus substratos assimétricos (Jeyabalan et al., 2000). Observou-se

que quatro regiões no esqueleto peptídico da proteína de cada monômero rearranjam-se de maneiras distintas na interação com o substrato. Além disso, pôde se perceber que as ligações de hidrogênio entre o substrato e a proteína que são intermediadas por moléculas de água possuem um papel de grande importância na interações entre receptor-ligante contribuindo principalmente para a adaptação do substrato assimétrico. As análises observadas forneceram um exemplo detalhado do reconhecimento proteína-proteína e ofereceu desse modo um paradigma, onde a enzima sofre mudanças conformacionais para reconhecer o substrato.

De posse do conhecimento acima muitos pesquisadores iniciaram um processo para tentar compreender o mecanismo de resistência sofrido pela PR. Vários desses estudos sugerem que a resistência aos inibidores de protease esteja associada ao movimento de abertura de suas alças, o que permite uma interação menos efetiva com o inibidor, mas relativamente boa com o substrato. Dessa forma vários são os trabalhos que tentam compreender esse mecanismo de abertura por meio de DM. Em 2006 Tóth et al. analisaram a dinâmica das alças com detalhes atômicos em proteases livres de ligante, já Batista et al. (2006) estudaram diferentes subtipos de protease interagindo com o inibidor ritonavir e puderam perceber que os subtipos não B sofrem maiores variações em sua estrutura em relação à estrutura inicial (principalmente nas regiões das alças) além de apresentarem redução, no tempo de permanência, de algumas ligações de hidrogênio.

Outros cientistas buscaram entender a influência que as mutações de resistência causam no comportamento da protease. Nessa linha Erickson e Burt (1996) estudaram os resíduos do sítio ativo e perceberam que eles frequentemente conferem mutação de resistência. Contudo alguns resíduos fora do sítio ativo também sofrem mutação que conferem resistência. Duas dessas mutações (D30N, N88S), a primeira dentro e a segunda fora do sítio ativo, foram estudadas por simulações de dinâmica molecular por Wartha et al. (2005). Este trabalho reporta a idéia de que mutações que ocorrem no sítio ativo são responsáveis por alteração direta entre o sítio e o inibidor enquanto mutações localizadas fora do sítio ativo contribuiriam

indiretamente para alterar o equilíbrio conformacional da protease.

Recentemente um trabalho veio sugerir um novo mecanismo de resistência (Murzycki et al., 2008). A hipótese é que os contatos hidrofóbicos da protease podem alterar a flexibilidade conformacional da proteína. Dessa forma mutações em um desses resíduos afetariam as interações hidrofóbicas existentes em cada monômero da proteína contribuindo diretamente no processo de resistência.

De modo geral os trabalhos de DM que estudam a protease do HIV-1 vêm tentando compreender o comportamento provocado pelas mutações de resistência na PR do subtipo B ou tentam entender as diferenças entre as PRs de diferentes subtipos do HIV-1. Este trabalho tem como diferencial tentar compreender o comportamento provocado pela mutação D30N nas PRs dos subtipos B e C. Até o momento não se tem notícias de nenhuma publicação envolvendo estudos por DM da mutação D30N na PR do subtipo C, assim como das mutações compensatórias N88D e N83T (observadas respectivamente *in vivo* e *in vitro* no subtipo C).

# Capítulo 2

## Objetivos

### 2.1 Objetivos Gerais

Utilizar técnicas computacionais de Modelagem Comparativa e Dinâmica Molecular, visando compreender os mecanismos pelos quais a mutação D30N altera as características funcionais da interação entre a PR do subtipo B e C e o ligante.

### 2.2 Objetivos Específicos

Conforme já mencionado, o subtipo C do HIV-1 que contém a mutação D30N tem baixíssima ocorrência clínica. Tal comportamento não ocorre para o subtipo B. Para tentar entender as diferenças causadas pela mutação D30N nos subtipos B e C foram estabelecidos os seguintes objetivos específicos:

- Construir os modelos estruturais teóricos das proteases BD30N, C, CD30N, CD30N-N83T, CD30N-N88D por modelagem comparativa.
- Executar etapas de dinâmica molecular em solvente explícito nos modelos.
- Analisar as diferenças e semelhanças entre as PR's do subtipo B e C estabelecendo padrões de interação protease/inibidor e protease/substrato para cada subtipo estudado.
- Obter informações que poderão auxiliar na obtenção de um inibidor que seja mais eficiente para a protease do subtipo C.

# Capítulo 3

## Métodos

O aprimoramento das técnicas de biologia molecular permitiu que fossem realizados projetos de sequenciamento em grande escala. O grande desenvolvimento das técnicas computacionais ocorreu devido ao avanço dos recursos computacionais em termos de hardware e software. Atualmente vários programas são comercializados através de grandes companhias e até mesmo de laboratórios acadêmicos. Além disso existem softwares fundamentados na filosofia de software livre, isto é, de código aberto e sem custos para o usuário. Um destes foi utilizado neste trabalho, o GROMACS (Van Gunsteren e Berendsen, 1987), para realizar as simulações de DM.

Métodos experimentais como a difração de raios-X em cristais ou Ressonância Magnética Nuclear (RMN), possuem um papel indispensável na correta compreensão dos mecanismos bioquímicos das proteínas, na elaboração de modelos experimentais, na construção de relações filogenéticas, no desenho de fármacos, no desenvolvimento de vacinas, dentre vários outros. Entretanto, muitas proteínas não possuem sua estrutura tridimensional determinada. Por isso as ferramentas computacionais têm grande utilidade na predição de estruturas tridimensionais de proteínas que ainda não foram resolvidas por métodos experimentais. Tais técnicas consistem em comparar a seqüência primária da proteína de interesse à proteínas homólogas já resolvidas experimentalmente. Essa metodologia é conhecida como Modelagem Comparativa (MC) (Marti-Renom et al., 2000).

As estruturas geradas por MC podem ser utilizadas como ponto de partida para o estudo dinâmico de proteínas utilizando-se a técnica de dinâmica molecular (DM), a qual nos permite acompanhar a trajetória de cada um dos átomos de uma proteína ao longo do tempo, permitindo a observação das mudanças que ocorrem nessa estrutura.

Ambas as técnicas (MC e DM) foram utilizadas nesse trabalho para que fossem obtidas as estruturas tridimensionais das proteases do subtipo B tipo selvagem (consB), BD30N, C tipo selvagem (consC), CD30N, CD30N/N83T e CD30N/N88D do HIV-1 complexadas ao inibidor nelfinavir e ao substrato KARVLAEEAM, e para que, em seguida, o comportamento destas fosse analisado ao longo do tempo.

### 3.1 Modelagem Comparativa (MC)

Detalhes tridimensionais de biomoléculas de estrutura desconhecida podem ser preditos pela técnica de MC. Esta técnica baseia-se na observação de que a estrutura de proteínas homólogas são bem conservadas evolutivamente (Holm e Sander, 1996). O passo inicial na modelagem comparativa é escolher a proteína de referência, ou seja, aquela que tem similaridade seqüencial ou estrutural suficiente com a seqüência de interesse para que a sua estrutura tridimensional seja utilizada como molde para a construção de um modelo tridimensional teórico.

O banco de dados utilizado para a escolha das estruturas de referência foi o PDB – Protein Data Bank (Berman, 2008), onde podem ser encontradas estruturas protéicas e não protéicas que foram resolvidas por métodos experimentais (i.e., raios-X, RMN) e são descritas por um conjunto de coordenadas atômicas. Nesse trabalho foram levados em conta os seguintes aspectos para a escolha das referências: **(i)** se o ligante/substrato estava complexoado à enzima; **(ii)** a qualidade da estrutura experimental.

O código de cada uma das proteínas no PDB, o ligante/substrato ao qual elas estão complexadas, sua resolução e a sua seqüência se encontram na Tabela 3.1. Na Fig. 3.1 pode ser visto o alinhamento dessas seqüências.

Tabela 3.1: Seqüências Proteicas utilizadas como molde na Construção dos Modelos

<b>Código PDB: 1OHR</b> <b>Protease do subtipoB do HIV-1</b> <b>Inibidor: Nelfinavir</b> <b>Resolução: 2.1 Å</b>	PQITLWQRPLVTIKIGGQLKEALLDTGAD DTVLEEMSLPGRWKPKMIGGIGGFIKVRQ YDQILIEICGHKAIGTVLVGPTPVNIIGR NLLTQIGCTLNF
<b>Código PDB: 1F7A</b> <b>Protease do subtipoB do HIV-1</b> <b>Substrato: KARVLAEAMS</b> <b>Resolução: 2.0 Å</b>	PQITLWKRPLVTIRIGGQLKEALLNTGAD DTVLEEMNLPKWKPKMIGGIGGFIKVRQ YDQIPVEICGHKAIGTVLVGPTPVNIIGR NLLTQIGCTLNF

1OHR	1 PQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGI	50
	:     :     :     :     :	
1F7A	1 PQITLWKRPLVTIRIGGQLKEALLNTGADDTVLEEMNLPKWKPKMIGGI	50
1OHR	51 GGFIVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRLLTQIGCTLNF	99
	:     :     :     :     :	
1F7A	51 GGFIVRQYDQIPVEICGHKAIGTVLVGPTPVNIIGRLLTQIGCTLNF	99

Figura 3.1: Alinhamento comparativo das proteases utilizadas como referência. Os símbolos qualificam o alinhamento quanto a similaridade: traço representa que os aminoácidos das estruturas são idênticos; dois pontos representa que há uma substituição por um aminoácido conservativo; um ponto quando as substituições não são conservativas.

Como as PRs do subtipo B e C são muito semelhantes, diferindo basicamente em apenas 8 aminoácidos, foi utilizado o programa de visualização e manipulação de estruturas protéicas e modelos, o Swisspdbviewer (Guex e Peitsch, 1997). Com ele, foram criadas as estruturas protéicas das seqüências escolhidas. Os arquivos dos modelos gerados foram então verificados com o programa PROCHECK (Laskowski et al., 1993), que permite a análise do comportamento global e individual dos aminoácidos através do diagrama de Ramachandran (Fig. 3.2) – que analisa a torção dos ângulos  $\phi$  e  $\psi$  da cadeia principal de uma determinada proteína (Fig. 3.3) – onde o percentual e os aminoácidos distribuídos dentro das regiões do gráfico identificam a qualidade estereoquímica da estrutura tridimensional da proteína. Um modelo é considerado bom quando mais de 90% dos seus

aminoácidos estão localizados nas regiões mais favorecidas. Neste trabalho todos os modelos criados apresentaram cerca de 94% dos seus resíduos nessas regiões. Na Fig. 3.2 pode-se observar um dos diagramas obtidos, todos os demais apresentaram comportamento semelhante.

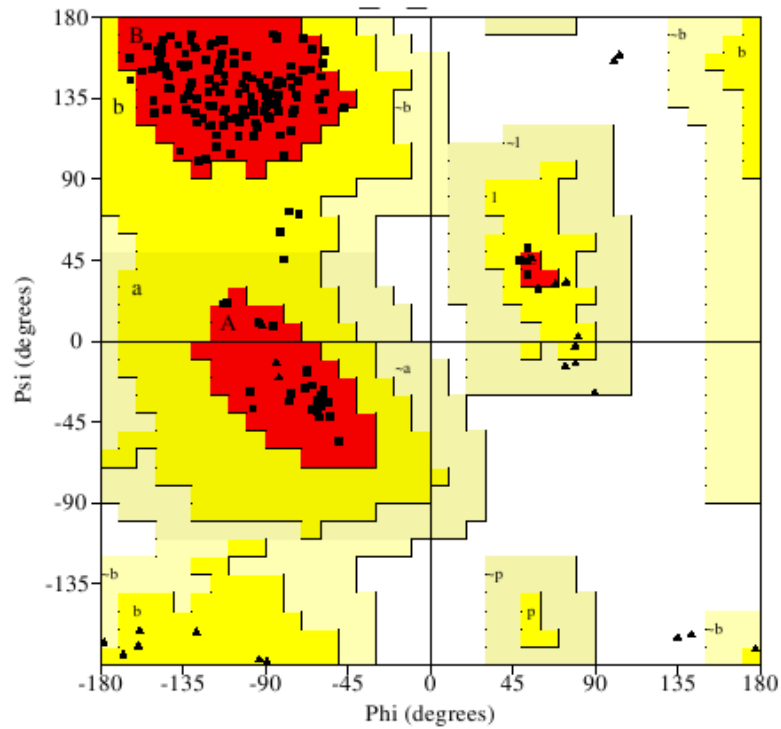


Figura 3.2: Diagrama de Ramachandran de um dos modelos criados, no qual estão mostradas as regiões fisicamente permitidas para os aminoácidos de acordo com os ângulos  $\phi$  e  $\psi$ . Em vermelho estão as regiões mais favorecidas, em amarelo escuro as regiões favorecidas, em amarelo claro as regiões pouco favorecidas e em branco as regiões desfavorecidas.



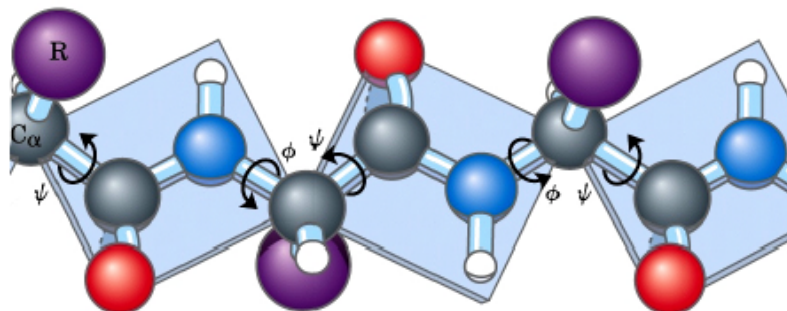


Figura 3.3: Ligação peptídica de dois aminoácidos consecutivos. Os ângulos torcionais  $\phi$  e  $\psi$  permitem que cada plano formado por esta ligação gire em relação aos seus vizinhos.

### 3.2 Otimização das Geometrias Moleculares

O processo de otimização consiste em encontrar um conjunto de coordenadas do sistema que minimize a energia potencial do sistema, de tal modo que este conjunto esteja associado a um mínimo local, ou global, da hipersuperfície de energia potencial (pontos no espaço de configurações onde todas as forças sobre os átomos do sistema estão balanceadas). Isso é feito através de ajustes nas posições atômicas, relaxando as distorções nas ligações químicas, nos ângulos entre ligações, nas interações eletrostáticas e nos contatos de van der Waals. Essa eliminação de deformações é importante para que, durante a dinâmica molecular, as energias potenciais localmente armazenadas não se convertam em energia cinética, aumentando muito a temperatura local podendo promover mudanças conformacionais

indesejáveis na macromolécula.

Dentre os métodos clássicos de minimização, encontramos: o algoritmo de *steepest-descent* (i.e., método de máximo declive) (Deift e Zhou, 1993) e o método de gradientes conjugados (Atkinson, 1988). O primeiro converge rapidamente quando se está longe do mínimo de energia, mas possui convergência lenta quando se está próximo do mesmo. Já o segundo método é indicado para um refinamento final, quando se está próximo do mínimo.

### 3.3 Equilibração do Sistema

Após o processo de minimização duas dinâmicas de equilibração são necessárias. A primeira consiste na equilibração do solvente. Nessa etapa aplicam-se restrições às posições dos átomos pesados da macromolécula em estudo (proteína + ligante). Isto permite que os átomos do solvente fiquem livres para a acomodação em camadas de solvatação e interajam corretamente com a proteína. Sem essa etapa é possível que se introduzam artefatos de simulação, onde a proteína pode sofrer mudanças conformacionais não “naturais” devido a uma má estruturação das camadas de solvatação.

A segunda etapa é a equilibração do sistema como um todo (nenhuma restrição é aplicada para quaisquer tipos de átomo). Nessa etapa o sistema se acomoda às condições termodinâmicas impostas à simulação, como temperatura, volume, pressão e a possíveis modificações introduzidas (e.g., mutações).

### 3.4 Ensemble

O estado termodinâmico de um sistema é definido por um pequeno número de grandezas como o número de moléculas  $N$ , o volume  $V$  e a temperatura  $T$ . O agrupamento de configurações que tem os mesmos valores destas grandezas formam o ensemble<sup>1</sup> correspondente. No limite termodinâmico ( $N \rightarrow \infty$ ) todos os ensembles são equivalentes, porém as simulações são feitas com sistemas finitos e

---

<sup>1</sup> Esta palavra foi introduzida na mecânica estatística por Gibbs (1902) e significa agrupamento, conjunto, etc.

por isto, para estudar certo tipo de propriedade, um ensemble pode ser vantajoso em relação a outro. Como exemplos de ensembles pode-se citar o canônico, em que os estados de equilíbrio são caracterizados por serem mínimos na função energia livre de Helmholtz (NVT constantes), e o ensemble isobárico-isotérmico no qual os estados de equilíbrio são caracterizados por serem mínimos na função energia livre de Gibbs (NPT constantes).

Os sistemas biológicos costumam ser simulados no ensemble NPT, procurando-se manter tanto a pressão quanto a temperatura com valores constantes, próximos aos encontrados nos organismos em estudo.

### 3.5 Dinâmica Molecular Clássica (DM)

Dinâmica molecular clássica (DM) é uma técnica com a qual se estudam os movimentos em um sistema de partículas por simulação. Ela pode ser empregada tanto em sistemas de átomos ou moléculas, como em sistemas macromoleculares. Seus elementos essenciais são o conhecimento do potencial de interação entre as partículas e das equações de movimento clássicas que governam a dinâmica dessas partículas. O potencial de interação deve incluir vários termos que descrevem as interações entre átomos e moléculas.

Como resultado da dinâmica molecular obtêm-se as energias e trajetórias para todas as partículas (ou átomos) e para o sistema como um todo, a partir das quais várias propriedades podem ser calculadas. O tempo deixa de ser contínuo, sendo discretizado nos sistemas moleculares em passos menores do que o período das vibrações dos átomos de hidrogênio, o movimento molecular mais rápido.

Na DM é calculada a força  $F_i$  que atua sobre cada partícula  $i$  a partir da derivada da função da energia potencial  $E(r_i)$ , que descreve a interação entre as partículas, em relação a sua posição  $r_i$ , sendo  $i = 1, 2, 3, \dots, N$ , até o número total de átomos.

$$F_i = - \frac{\partial E(\{r_i\})}{\partial r_i} \quad (3.1)$$

A aceleração a que cada átomo está submetida é obtida dividindo-se a força  $F_i$  pela massa  $m_i$ :

$$a_i = \frac{F_i}{m_i} \quad (3.2)$$

Substituindo a aceleração no algoritmo de Verlet (Verlet, 1967) se determina a posição das partículas em incrementos de tempo  $dt$ :

$$r_i(t + \delta t) = 2r_i(t) - r_i(t - \delta t) + a_i(t)\delta t^2 \quad (3.3)$$

Embora a predição das novas posições no instante  $t + \delta t$  sejam computadas somente a partir das posições nos instantes  $t$ ,  $t - \delta t$  e das forças  $F_i$  sobre cada partícula no instante  $t$ , as velocidades são necessárias para o cálculo da energia cinética que, somada à energia potencial, dá a energia total do sistema. Elas são obtidas a partir do algoritmo de Verlet para a propagação das velocidades:

$$v_i(t) = \frac{r_i(t + \delta t) - r_i(t - \delta t)}{2\delta t} \quad (3.4)$$

No algoritmo de Leap-Frog (Hockney, 1970) a velocidade é incluída explicitamente na determinação das novas posições e possui uma maior precisão numérica sendo as velocidades calculadas em  $(t \pm \frac{\Delta t}{2})$  e as posições em  $(t \pm \Delta t)$ . Esse algoritmo introduz pequenas modificações nas equações anteriores, de forma a reduzir as diferenças entre grandes termos, quando as velocidades são calculadas.

$$r_i(t + \delta t) = r_i(t) + v_i(t + \frac{\delta t}{2})\delta t \quad e \quad v_i(t + \frac{\delta t}{2}) = v_i(t - \frac{\delta t}{2}) + a_i(t)\delta t \quad (3.5)$$

As simulações por dinâmica molecular utilizam as equações da mecânica clássica, ou seja, as equações de Newton, para cada partícula e em cada incremento no tempo. O processo mais dispendioso computacionalmente é, em geral, a avaliação das forças para obtenção das acelerações. O tempo gasto no cálculo dessas

forças depende da complexidade da função de energia potencial de interação entre as partículas, como será descrito a seguir. Em sistemas moleculares, as funções potenciais mais realistas são compostas de diversos termos. Na DM são tomadas as derivadas espaciais de cada termo e somadas para se obter a força resultante e a aceleração, sobre cada átomo e a cada novo conjunto de coordenadas. As acelerações são inseridas no algoritmo de Verlet, ou em um similar, para a predição das novas posições e em seguida o processo se repete.

### **3.6 Função Energia Potencial ou Campo de Força**

Um sistema molecular seria descrito mais exatamente pela resolução quântica do problema para cada geometria. Porém, o elevado número de átomos das macromoléculas biológicas e as atuais limitações computacionais requerem a sua representação física através de uma função potencial relativamente simples, descrevendo de forma clássica a interação entre as unidades atômicas. Em dinâmica molecular e cálculos de mecânica molecular, as interações intra e intermoleculares são calculadas de uma forma simplificada utilizando uma parametrização empírica. Esses parâmetros incluem massas dos átomos, cargas, ângulos, diedros, potenciais entre átomos não ligados (van der Waals e interações eletrostáticas), etc. Frequentemente se refere ao conjunto desses parâmetros de ajuste, juntamente com as respectivas funções, como “campo de forças molecular”. Na simulação da dinâmica molecular, as forças que atuam sobre cada átomo são obtidas calculando-se a primeira derivada do potencial em relação às posições desses átomos.

A partir dessas forças, resolvem-se as equações do movimento para descrever como as posições atômicas variam com o tempo. A cada passo da dinâmica as forças são então reavaliadas. Diversas funções potenciais têm sido propostas na literatura. Como fatores comuns entre elas aparecem termos harmônicos para descrever ligações covalentes entre pares de átomos e ângulos entre ligações químicas vizinhas, um termo torcional para descrever rotações em torno de ligações e termos que descrevem interações entre átomos não ligados.

Os campos de forças mais comumente empregados em simulações de DM são o GROMOS (Van Gunsteren e Berendsen, 1987), AMBER (Weiner et al., 1984), CHARMM (Brooks et al., 1983) e MMFF94 (Halgren, 1996). Desses, um que é adequado para a simulação de proteínas em solução aquosa é o chamado GROMOS96. Uma forma funcional deste campo de forças é mostrada na Equação 3.6 que se encontra no final desta seção, contudo primeiro serão mostrados e explicados, cada termo pertencente a esta equação.

O primeiro termo da EQ. 3.6 abrange a interação entre dois átomos ligados, por meio de um potencial harmônico, reproduzindo classicamente o movimento vibracional entre pares de átomos ligados covalentemente.

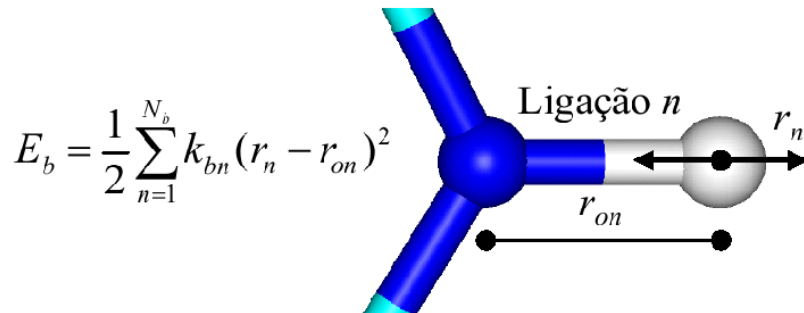


Figura 3.4: Representação da energia potencial de ligação para um par  $n$  qualquer. O comprimento da ligação oscila com o tempo segundo um potencial harmônico em torno de uma posição de equilíbrio (imagem adaptada de Silva (2003)).

Na equação acima,  $N_b$  é o número total de pares de átomos ligados e para cada par são definidos os valores específicos em que  $k_{bn}$  é a constante de força específica,  $r_n$  é a distância entre o par e  $r_{on}$  é o parâmetro que define o comprimento de equilíbrio da ligação.

O segundo termo da EQ. 3.6 simula a variação angular entre três átomos ligados em seqüência, por um potencial harmônico.

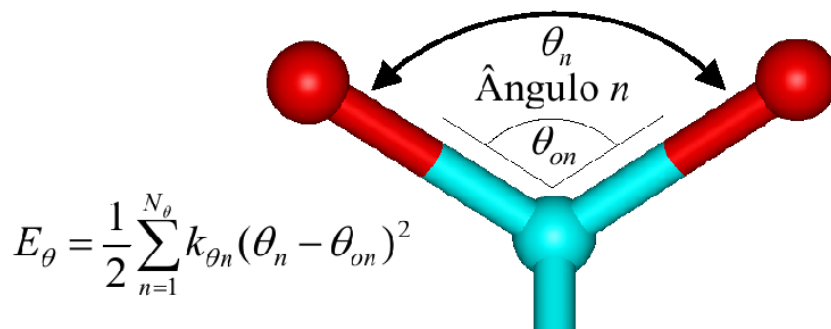


Figura 3.5: Representação da energia potencial angular para um terço de átomos ligados n qualquer cujo ângulo entre eles varia harmonicamente em torno do ângulo de equilíbrio (imagem adaptada de Silva (2003)).

$N_{\Theta}$  é o número total de trios de átomos,  $k_{\Theta_n}$  a constante de força angular,  $\Theta_{on}$  o ângulo de equilíbrio e  $\Theta_n$  a variação angular.

O terceiro termo da EQ. 3.6 descreve a variação torcional imprópria ou diedral imprópria, envolvendo quatro átomos não ligados sequencialmente, sendo um central ligado aos outros três, sobre os quais se pode imaginar planos que compõem um ângulo diedral, que variaria harmonicamente. Dão conta de manter a planaridade dos átomos de hidrogênio em relação ao anel benzênico e a da ligação peptídica, por exemplo, assim como a quiralidade para certos átomos de carbono.



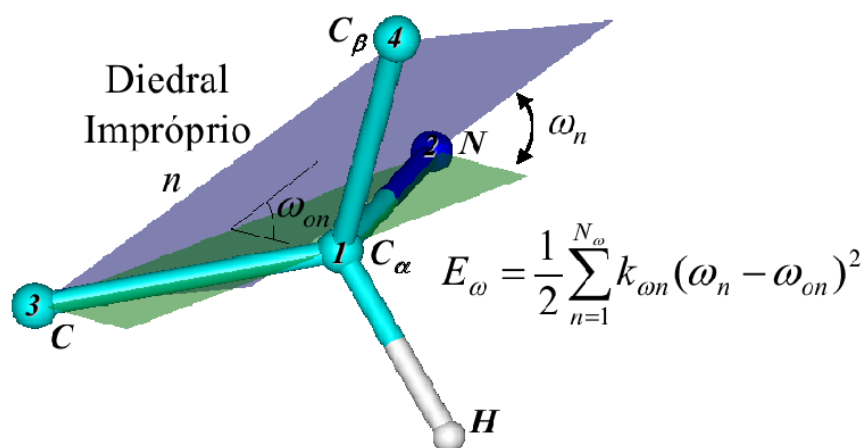


Figura 3.6: Representação da energia potencial diedral imprópria. Um carbono  $\alpha$  quiral está representado aqui. Como na maioria dos campos de força o átomo de hidrogênio (representado para efeitos didáticos) fica integrado no  $C\alpha$ . Este potencial cuida de manter a estabilidade tetraédrica deste grupo de átomos (imagem adaptada de Silva (2003)).

$N_{\omega}$  é o número total de quaternos de átomos envolvidos em diedrais impróprios,  $k_{\omega n}$  é a constante de força angular diedral impróprio,  $\omega_{on}$  o ângulo diedral de equilíbrio e  $\omega_n$  sua variação angular. Na variação do diedro impróprio, um carbono  $\alpha$  quiral está representado com o átomo de hidrogênio integrado a ele. Este potencial procura manter a estabilidade tetraédrica deste grupo de átomos.

O quarto termo ligado da EQ. 3.6 descreve potencial do tipo periódico, conhecido também como potencial torcional ou diedral próprio, portanto não harmônico como os outros três potenciais para átomos ligados. É formado por quatro átomos em seqüência, cujos três primeiros formariam um plano em ângulo diedral com o plano formando pelos três últimos, de tal forma que este ângulo possa assumir posições preferenciais em função de sua multiplicidade ( $m_n$ , número de mínimos para a torção de uma ligação química específica), que pode ser 2, um a cada  $180^\circ$ ; 3, um a cada  $120^\circ$ ; e 6, um a cada  $60^\circ$ ). Está associado às configurações que este conjunto de átomos pode assumir.

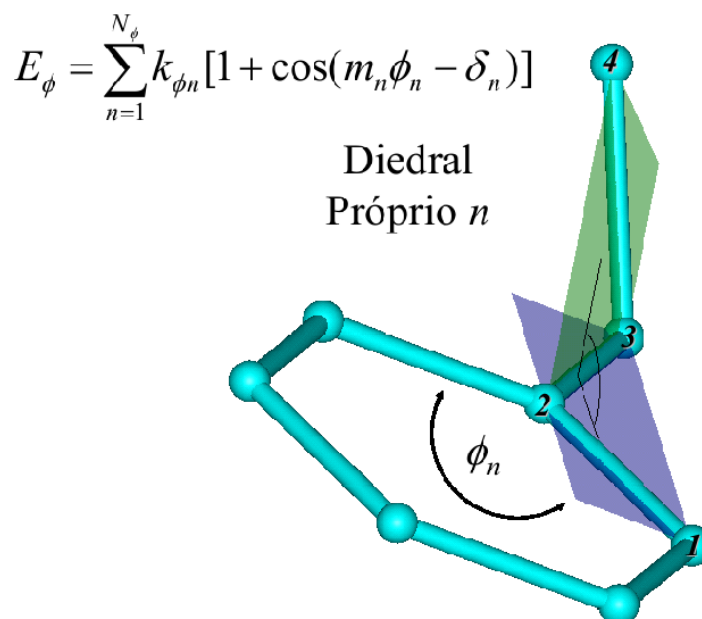


Figura 3.7: Representação da energia potencial torcional própria. Nesta representação, o anel benzênico (no plano diédrico 1\_2\_3) pode assumir seis ângulos de torção (mínimos) em relação ao plano do diedro 2\_3\_4 (imagem adaptada de Silva (2003)).

Na equação acima,  $k_{\phi_n}$  é a constante que define a barreira de torção,  $m_n$  a multiplicidade de mínimos,  $\phi_n$  a variação diedro angular e  $\delta_n$  sua diferença de fase que pode ser  $0^\circ$  ou  $180^\circ$ , deslocando-se assim o ponto de máxima para este potencial.

O quinto termo da EQ. 3.6 representa a energia potencial de Lennard-Jones. Este termo descreve interações entre átomos não ligados quimicamente. É um potencial de curto alcance (unidades de  $\text{Å}$ ) com um termo repulsivo, em virtude do princípio de exclusão de Pauli, o qual diz que dois elétrons não podem compartilhar os mesmos 4 números quânticos, e um atrativo (associado às interações de van der Waals), derivado principalmente pelas forças de dispersão, conhecidas também como forças de London, que surgem a partir de pequenos dipolos oscilantes criados pela flutuação da nuvem eletrônica.

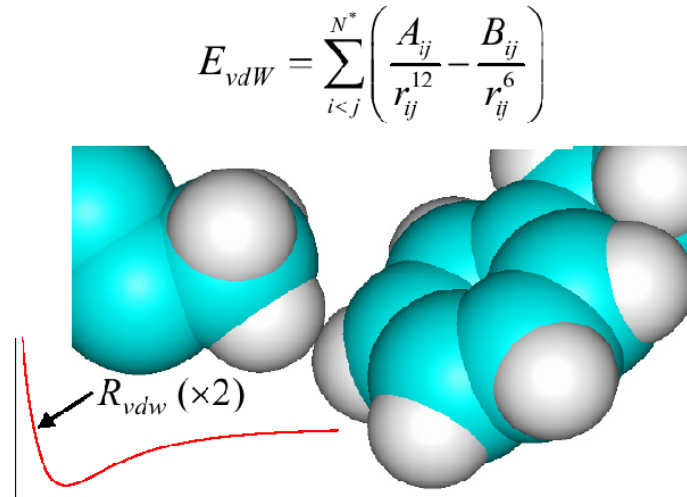


Figura 3.8: Representação atômica usando-se os raios de Van der Waals ( $R_{vdw}$ ), os quais são definidos para cada átomo pela metade da distância em que seu potencial de Lennard-Jones vale zero (o gráfico não está na mesma escala que o desenho) (imagem adaptada de Silva (2003)).

Para um dado átomo  $i$  é computada sua energia de interação com os átomos  $j$ , de um total de  $N^*$ , pois são excluídos seus primeiros e segundos vizinhos ligados, que são tratados pelos potenciais anteriores, ao passo que para os terceiros vizinhos parâmetros menores no potencial de Lennard-Jones são considerados. Os parâmetros  $A_{ij}$  e  $B_{ij}$  dependem dos tipos de pares de átomos (e se são terceiros vizinhos ou não) e  $r_{ij}$  é a distância entre eles.

O sexto termo da EQ. 3.6 descreve a energia potencial eletrostática. Na definição dos parâmetros de carga para os átomos segundo um campo de forças, valores de cargas parciais lhes são atribuídos de forma a se tentar reproduzir a diferença de eletronegatividade que eventualmente existe entre dois átomos ligados covalentemente. As pontes de hidrogênio são de origem essencialmente eletrostática (no campo de forças elas são descritas pelo termo eletrostático e LJ) e tem grande importância na estabilização de estruturas secundárias das proteínas e na especificidade entre ligante e proteína.

$$E_{el} = \sum_{i < j}^{N^*} \frac{q_i q_j}{4\pi \xi \xi_0 r_{ij}}$$

Na equação acima,  $q_i$  e  $q_j$  são as cargas parciais para dois átomos quaisquer do sistema, cuja distância entre eles é  $r_{ij}$ ,  $\xi_0$  é a permissividade dielétrica do vácuo livre,  $\xi$  a constante dielétrica relativa do meio e, para um dado átomo  $i$ ,  $N^*$  é o número restante de átomos  $j$ , subtraído seus primeiros e segundos vizinhos ligados, sendo sua interação como os terceiros vizinhos reduzida por uma constante. Contudo, o mais comum em DM é aplicar-se um raio de corte (RC), isto é, dado um átomo  $i$  a partir do qual se delimita uma esfera de raio RC, tomam-se para cálculo de interação eletrostática e de van der Waals somente os átomos não ligados contidos nesta esfera.

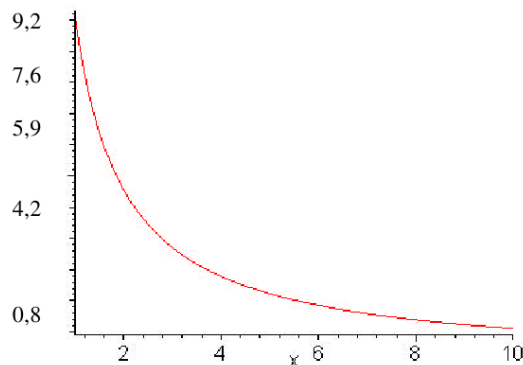
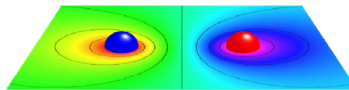


Figura 3.9: Representação esquemática da energia potencial de Coulomb entre dois átomos de cargas opostas. No gráfico da função de Coulomb, o eixo  $x$  corresponde à distância (em  $\text{\AA}$ ) e o eixo  $y$  corresponde à energia (KJ/mol) (Adaptado de Silva (2003))

Desta forma, os termos combinados para energia potencial compõem a equação do Campo de Forças.

$$E = E_b + E_{\Theta} + E_{\omega} + E_{\phi} + E_{vdW} + E_{el} \quad (3.6)$$

Energia Potencial de Ligação:  $E_b = \frac{1}{2} \sum_{n=1}^{N_b} k_{bn}(r_n - r_{on})^2$

Energia Potencial de Ligação Angular:  $E_{\Theta} = \frac{1}{2} \sum_{n=1}^{N_{\Theta}} k_{\Theta n}(\Theta_n - \Theta_{on})^2$

Energia Potencial Torcional Imprópria:  $E_{\omega} = \frac{1}{2} \sum_{n=1}^{N_{\omega}} k_{\omega n}(\omega_n - \omega_{on})^2$

Energia Potencial Diedral Agular:  $E_{\phi} = \sum_{n=1}^{N_{\phi}} k_{\phi n}(1 + \cos(m_n \phi_n - \delta_n))$

Energia Potencial de Lennard-Jones:  $E_{vdW} = \sum_{i<j}^N \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right)$

Energia Potencial Eletrostática:  $E_{el} = \sum_{i<j}^{N^*} \frac{q_i q_j}{4\pi \epsilon \epsilon_0 r_{ij}}$

Os quatro primeiros termos tratam das interações entre átomos ligados e respondem por interações essencialmente de origem quântica (como, por exemplo, nas ligações químicas covalentes). Os termos de Lennard-Jones e de Coulomb, ainda que simplificados, tratam das interações de longa distância, entre átomos distantes quatro vizinhos ou mais e, como mencionado antes, com um tratamento especial para o terceiro. Todos estes potenciais atuam de forma independente uns dos outros e sobre todos os átomos do sistema.

### 3.7 Correções das Interações de Longo Alcance

A estabilidade conformacional de uma proteína resulta de um arranjo de interações locais e não-locais. As contribuições eletrostáticas são as principais contribuições de longo alcance.

A contribuição para as interações de van der Waals em distâncias superiores a 0,8 nm é bem próxima de zero. Entretanto, a contribuição para o termo

de Coulomb é importante a distâncias mais longas. Como as interações de longo alcance, a princípio, são de alcance infinito, quando se faz o cálculo da contribuição eletrostática (Coulomb) que cada átomo sofre, deve-se fazer o somatório das interações entre todos os pares de átomos. Como os sistemas em DM são de dezenas ou centenas de milhares de átomos, isso acarretaria em um alto custo computacional, inviabilizando esta técnica. São empregadas técnicas para limitar o número de partículas envolvidas nestes cálculos. O método clássico para essa limitação é o Raio de Corte Truncado (RC-T), em que há interrupção nas interações a partir de uma determinada distância. Este método pode atribuir um grande ruído no sistema (série de imprecisões e artefatos) por eliminar um número grande de interações importantes, quando o valor do Raio de Corte for insuficiente. Foram desenvolvidos métodos que reduzem drasticamente este ruído, dentre eles o Raio de Corte com Campo de Reação (RC-CR) e o Particle Mesh Ewald (PME). Como descrito por Batista et al. (2006) o RC-CR demonstrou ser a melhor técnica para simular os efeitos da PR do HIV, portanto nesse trabalho tal técnica foi escolhida para nossas simulações.

O tratamento das interações eletrostáticas utilizando o RC-CR faz com que o potencial de Coulomb tenha o valor zero no Raio de Corte escolhido, ao assumir uma constante dielétrica para o ambiente além do Raio de Corte. O potencial para a interação de Coulomb passa a ser:

$$V_{crf} = f \frac{q_i q_j}{r_{ij}} \left[ 1 + \frac{\xi_{rf} - 1}{2\xi_{rf} + 1} \frac{r_{ij}^3}{r_c^3} \right] - f \frac{q_i q_j}{r_c} \frac{3\xi_{rf}}{2\xi_{rf} + 1} \quad (3.7)$$

onde  $q_i$  e  $q_j$  são as cargas dos átomos envolvidos,  $\xi_{rf}$  é a constante dielétrica para o Campo de Reação,  $r_c$  é o Raio de Corte e  $r_{ij}$  é a distância entre os átomos envolvidos.

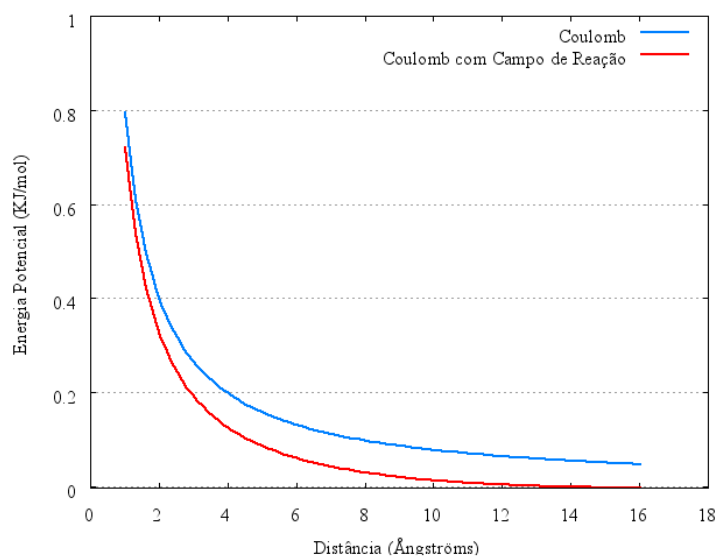
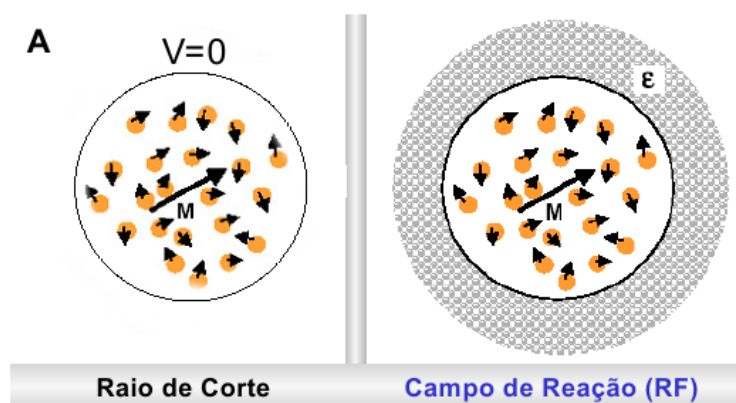


Figura 3.10: Representação do potencial de Coulomb com e sem a aplicação do Campo de Reação. Em A, figura representativa das considerações para cada tratamento eletrostático. Em B, representação dos efeitos da aplicação do campo de reação na função de Coulomb para partículas de mesma carga com e sem o campo de reação, utilizando um raio de corte de 16 Å. Imagem adaptada de Goliatt (2007).

### 3.8 Condições de Fronteira

Um sistema pode ser simulado sob diversas condições de fronteira. O ideal seria que o meio circundante fosse infinito, mas isto é impossível de ser modelado, ou então que fosse grande o bastante para ser definido fisicamente como “infinito”, de maneira que se pudessem desprezar os efeitos de borda, algo de altíssimo custo computacional e, portanto, inviável. Simulações podem ser feitas em meio efetivo (em que os parâmetros do campo de forças são elaborados para simular o efeito do solvente implicitamente), sem solvente explícito não há o problema das condições de fronteira. Porém, a simulação molecular, em busca de uma melhor representação de sistemas biológicos (reações que ocorrem em meio aquoso), exige, para uma gama bem grande de estudos, a presença do solvente tratado explicitamente (Fig. 3.11).

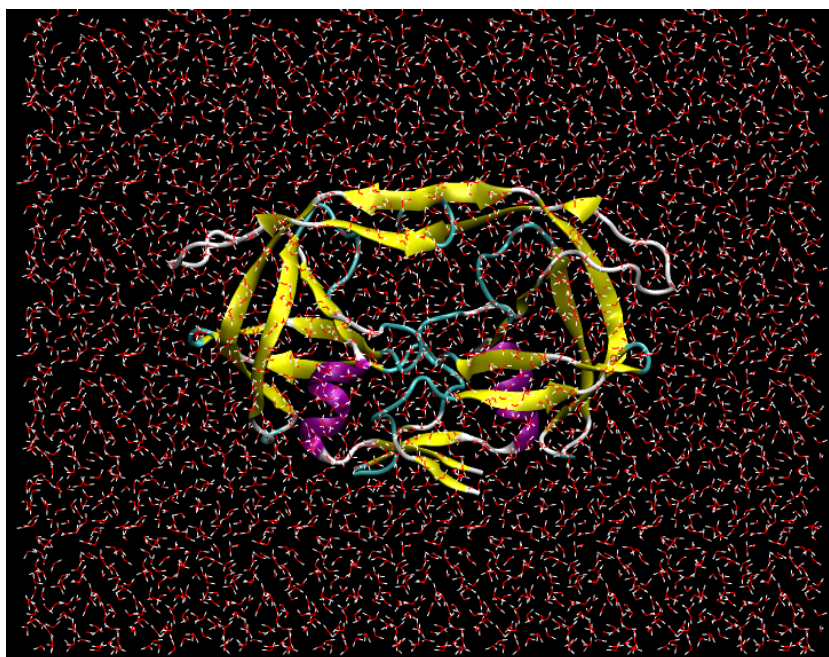


Figura 3.11: Enzima protease do HIV em solvente explícito. As cadeias da protease estão na representação de estrutura secundária. A cor rosa identifica as  $\alpha$ -hélices, amarelo folhas- $\beta$  e verde ou branco os loops. As moléculas de água estão na representação de varetas, coloridas pelo tipo de átomos: branco (H) e vermelho (O).

Assim, uma forma de se abordar este problema é pela inclusão das Condições Periódicas de Contorno (PBC), cuja geometria pode ser cúbica, octaédrica, entre



outras. Esse método impede que um átomo que se encontra próximo a borda realize um número menor de interações que um átomo que se encontra dentro da caixa. Em uma dinâmica com moléculas explícitas de água, aplicam-se as PBC a uma caixa contendo a molécula de interesse e o número apropriado de moléculas do solvente, que é circundada por imagens periódicas de si mesma. Dessa forma, quando uma molécula se move na caixa central, suas imagens periódicas em cada uma das caixas vizinhas se movem exatamente da mesma forma. Então, se uma molécula deixa a caixa central, uma de suas imagens irá entrar pela face oposta como esquematizado abaixo.

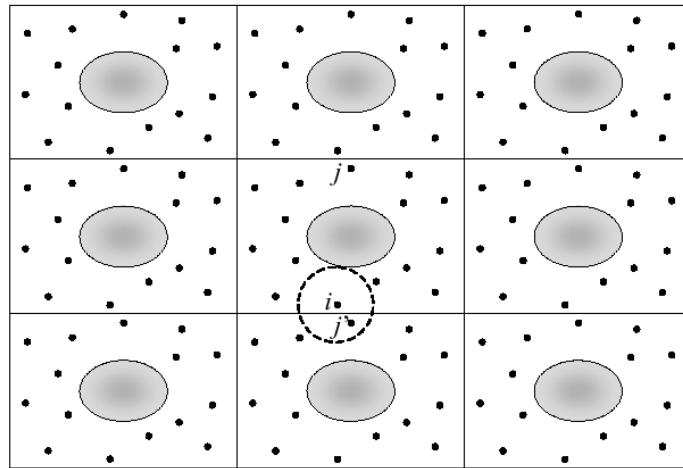


Figura 3.12: Condições periódicas de contorno retangular. Uma proteína e seu solvente com suas oito imagens replicadas. Uma esfera de raio  $R_c$  é traçada a partir do átomo  $i$ , o qual interage com a imagem do átomo  $j$  ( $j'$ ).

Além disso, o uso das PBC implica a necessidade de se usar raio de corte ( $R_c$ ), de tal forma que um dado átomo do sistema não possa ver outro átomo de sua imagem.

$$d > \frac{R_c}{2}$$

Onde  $d$  é a distância mínima entre a proteína e a parede da caixa, a ser preenchida, normalmente, por moléculas de solvente. Na Fig. 3.12 está uma representação bidimensional da PBC, neste caso um sistema de geometria retangular.

### 3.9 Parâmetros das Simulações

Levando em consideração os resultados obtidos por Gonzalez et al. (2004), foram propostas 12 estruturas da protease do HIV-1:

- subtipoB+inibidor, subtipoB+substrato;
- subtipoC+inibidor, subtipoC+substrato;
- mutante BD30N+inibidor, mutante BD30N+substrato;
- mutante CD30N+inibidor, mutante CD30N+substrato;
- mutante CD30N/N83T+inibidor, mutante CD30N/N83T+substrato;
- mutante CD30N/N88D+inibidor, mutante CD30N/N88D+substrato;

Para o subtipoB+inibidor não foi criado nenhum modelo. Neste caso a estrutura cristalográfica 1OHR (usada como molde) foi utilizada. As sequências utilizadas para a construção dos modelos consB+substrato, consC+substrato e consC+inibidor foram as sequências consenso propostas por Soares et al. (2003). Para os respectivos mutantes foram utilizadas essas mesmas sequências acrescidas das mutações em estudo. Na Fig 3.13 pode-se visualizar o alinhamento das sequências de cada uma das estruturas.

		10	20	30	40				
1OHR	PQITLWQRPLV	TIK	GGQ	LKEAL	LDTGAD	TVLEEM	SLPGR	WKPKMIGGI	50
B_D30N	PQITLWQRPLV	TIK	GGQ	LKEAL	LDTGAD	N	TVLEEM	NLPGR	WKPKMIGGI
consC	PQITLWQRPLV	S	IKVGGQ	LKEAL	LDTGAD	TVLEE	I	NLPGR	WKPKMIGGI
C_D30N	PQITLWQRPLV	S	IKVGGQ	LKEAL	LDTGAD	N	TVLEE	I	NLPGR
CD30N_N83T	PQITLWQRPLV	S	IKVGGQ	LKEAL	LDTGAD	N	TVLEE	I	NLPGR
CD30N_N88D	PQITLWQRPLV	S	IKVGGQ	LKEAL	LDTGAD	N	TVLEE	I	NLPGR

		60	70	80	90					
1OHR	GGFIKVRQYDQ	ILIEICGH	KAIGTV	LVGPT	PVNI	I	GRN	LLTQ	IGCTLNF	99
B_D30N	GGFIKVRQYDQ	ILIEICGH	KAIGTV	LVGPT	PVNI	I	GRN	LLTQ	IGCTLNF	99
consC	GGFIKVRQYDQ	ILIEICG	K	KAIGTV	LVGPT	PVNI	I	GRN	MLTQ	IGCTLNF
C_D30N	GGFIKVRQYDQ	ILIEICG	K	KAIGTV	LVGPT	PVNI	I	GRN	MLTQ	IGCTLNF
CD30N_N83T	GGFIKVRQYDQ	ILIEICG	K	KAIGTV	LVGPT	P	V	T	I	GRN
CD30N_N88D	GGFIKVRQYDQ	ILIEICG	K	KAIGTV	LVGPT	PVNI	I	GR	D	MLTQ

Figura 3.13: Alinhamento das seqüências das PRs construídas. As colunas em destaque representam os polimorfismos encontrados entre os subtipos B e C (T12S, I15V, L19I, M36I, R41K, H69K, L89M e I93L); em vermelho estão as mutações D30N, N83T e N88D; e em rosa a diferença entre a seqüência 1OHR e a consenso.

Essas estruturas foram construídas através do programa Swiss-PdbViewer e posteriormente foram submetidas a cálculos de minimização de energia no programa Gromacs 3.3.1.

O campo de forças utilizado foi o GROMOS96 (Van Gunsteren et al., 1996) e os parâmetros de carga escolhidos para os resíduos foram padrões (representações de carga mais prováveis quando o resíduo se encontra em pH fisiológico). Para todos os modelos, os resíduos Arg e Lys ficaram protonados, carga +1, enquanto as His ficaram neutras. Os resíduos Glu e Asp ficaram em estados desprotonados, carga -1, exceto o aminoácido catalítico, Asp124, que foi protonado com a finalidade de mimetizar o comportamento real da enzima. Todas as simulações foram realizadas em caixas cúbicas contendo a descrição explícita de moléculas de água e considerando uma camada de solvatação de no mínimo 15 Å em cada dimensão da macromolécula.

Como essa é uma proteína carregada, foram adicionados contra-íons ao sistema para torná-lo neutro. Na verdade, estes sistemas foram neutralizados pela substituição de moléculas de água por íons de cloro (Cl<sup>-</sup>) nas posições mais favo-

ráveis eletrostaticamente. Em geral cada sistema foi composto por cerca de 40.000 átomos.

Antes da dinâmica molecular, foram feitos 5000 passos de minimização pelo método *steepest-descent* com restrição de posição para os átomos pesados do esqueleto da proteína, 5000 passos de minimização ainda pelo *steepest-descent* com todos os átomos flexíveis e 5000 passos de minimização pelo método de gradientes conjugados. Todos os cálculos envolvendo o GROMACS foram feitos em precisão dupla. Isso resultou que o processo de otimização parasse antes dos 5000 passos de interações (entre 1500 e 3500 passos no primeiro método e entre 1 e 10 passos no segundo) porque o limite de precisão numérica do processador havia sido alcançado.

Após o processo de minimização foram feitas duas etapas de equilíbrio, uma DM de 500 ps com restrição de posição para os átomos pesados e outra DM de 2 ns sem restrição de posição. Após a etapa de equilíbrio, foram feitas simulações de 8 ns de DM para posteriores análises. As minimizações das energias potenciais por mecânica molecular e as simulações de DM foram feitas com o pacote de programas GROMACS, versão 3.3.1, utilizaram o modelo de água SPC flexível (Van Gunsteren e Berendsen, 1987) e PBC.

As simulações de MD foram de 8 ns, com dt igual a 2 fs, à temperatura de 300 K e pressão igual de 1 atm, modelo de água SPC, registrando-se coordenadas espaciais atômicas e energias a cada 2ps. Foram usados raios de corte de 14 Å para Coulomb, 12 Å para Van der Waals e a constante dielétrica  $\xi_{rf} = 54$ , valores escolhidos em função da temperatura, do sistema estudado e do modelo de água utilizado (Batista et al., 2006; Smith e Van Gunsteren, 1994). Foram também aplicados os algoritmos de constrição: SETTLE (Miyamoto e Kollman, 1992), para o solvente, e LINCS (apenas para comprimentos de ligações, Hess et al. (1997)) para a protease e ligante. Estes algoritmos são os mais rápidos até agora implementados para simulações de MD. Só com as constrições de distância de ligações (aplicada a todos os tipos de ligações) é que é possível elevar o passo de integração no tempo de

0,5 para 2,0 fs. Para o tratamento das interações eletrostáticas de longa distância foi usado o campo de reação.

As condições das simulações de dinâmica molecular das proteases dos subtipos B e C estão sumariadas na Tabela 3.2.

Tabela 3.2: **Ficha técnica das condições gerais das simulações realizada por MD das proteases B e C**

<b>Tempo de simulação</b>	8ns (4000000 passos de 2fs)
<b>Programa</b>	Gromacs 3.3.1
<b>Campo de forças</b>	Gromos96 (43a1)
<b>Interação eletrostática</b>	RC_CR
<b>Raio de Corte (Coulomb)</b>	14Å
<b>Interação de van der Waals</b>	Raio de Corte
<b>Raio de corte (Van der Waals)</b>	12Å
<b>Restrição das ligações da Proteína</b>	LINCS
<b>Restrição das moléculas de água</b>	SETTLE
<b>Modelo de água</b>	SPC
<b>Outros</b>	Temperatura constante de 300 k (Berendsen) Pressão Constante de 1 atm (Berendsen) Condições Periódicas de Contorno (PBC)

### 3.9.1 Análise de Estruturas e Geração de Figuras

Para analisar as estruturas geradas após os cálculos de otimização e dinâmica molecular, foi utilizado o programa VMD (programa de visualização molecular, animação e análise de grandes sistemas biomoleculares). Os gráficos de variação de energia, da variação do RMSD, das distâncias entre átomos e da formação de ligações de hidrogênio foram analisados com o programa XMGRACE e posteriormente gerados no programa Microsoft Office Excel. Já os gráficos que descrevem o comportamento das enzimas ao longo da dinâmica foram gerados no programa MOLMOL. As análises de superfície de contato intermolecular foram feitas com o programa SURFMDS (Souza et al., versão atualizada de 2007), um programa desenvolvido pelo Laboratório de Modelagem e Dinâmica Molecular (LMDM) do Instituto de Biofísica Carlos Chagas Filho da UFRJ.

### 3.10 Topologia dos inibidores

Tendo em vista que no campo de forças GROMOS96 (Van Gunsteren et al., 1996), implementado no programa GROMACS 3.3.1 (Van der Spoel et al., 2001), não há parâmetros para compostos diferentes de aminoácidos, proteínas e ácidos nucléicos, houve a necessidade de parametrizar o ligante a fim de obtermos arquivos com o conjunto de informações necessárias para a execução do programa. Tal parametrização me foi cedida pelo Laboratório de Modelagem e Dinâmica Molecular (LMDM) do Instituto de Biofísica Carlos Chagas Filho da UFRJ. O grupo do LMDM tomou como base a topologia para o campo de forças GROMOS87 (Van Gunsteren e Berendsen, 1987) gerada automaticamente pelo servidor PRODRRG (Aalten et al., 1996), sendo utilizado deste campo de forças alguns parâmetros de ângulos, ligações e outros parâmetros previamente determinados para outras estruturas similares. Para o cálculo dos parâmetros não disponíveis (i.e., cargas parciais atômicas), foram feitos cálculos *ab initio* utilizando o programa GAUSSIAN94 (Frisch et al., 1995), aplicando a base B3LYP/6-31G\*\* com a opção CHELPG, assumindo carga total nula para os fármacos.

### 3.11 Desvio da Raiz Média Quadrática (RMSD)

O desvio padrão ou Desvio da Raiz Média Quadrática (RMSD do inglês *Root Mean Square Deviation*) (Spiegel, 1994) é o resultado do cálculo realizado para a comparação de dois conjuntos de dados. Sua aplicação neste trabalho é o cálculo para a comparação das coordenadas dos átomos (excluindo o solvente) ao longo do tempo das simulações em relação à estrutura inicial como referência. Nesse caso temos um gráfico do desvio global da estrutura em relação à conformação inicial versus tempo.

$$\text{Desvio Padrão} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - X)^2}$$

onde  $n$  é o número total de amostras,  $X_i$  é a coordenada tridimensional de cada átomo da amostra  $i$  e  $X$  a coordenada de referência. O resultado destes

cálculos dá uma idéia do quanto a proteína variou ao longo do tempo em relação à estrutura de referência.

### 3.12 Flutuação da Raiz Média Quadrática

Outra análise com base no Desvio Padrão é a Flutuação da Raiz Média Quadrática (*RMSF*). Este método difere em relação ao primeiro pelo fato de agora tomarmos com referência a estrutura média na Dinâmica Molecular e calcular a flutuação de cada resíduo de aminoácido (média dos desvios de cada átomo do resíduo em relação a esta estrutura média). Nesse caso temos um gráfico do desvio ou flutuação conformacional de cada resíduo versus sua posição na cadeia polipeptídica que pode permitir observar as variações locais, ou seja, quais os resíduos tiveram maior variação.

### 3.13 Raio de Giro e Momento de Inércia

Uma análise sobre o formato adquirido pela proteína ao longo da simulação por Dinâmica Molecular é possível monitorando o comportamento da proteína nos eixos x, y e z, e calculando os principais momentos de inércia. O momento de inércia pode ser calculado multiplicando-se o raio de giro pela massa da proteína. O raio de giro  $R_g$  é calculado como:

$$R_g = \frac{\sum_i \|r_i\|^2 m_i}{\sum_i m_i}$$

onde  $m_i$  é a massa do átomo  $i$  e  $r_i$  é a distância do átomo  $i$  em relação ao centro de massa da molécula. O momento de inércia é dado por:

$$I = \sum_i m_i R_g^2$$

### 3.14 Ligações de Hidrogênio

As ligações de hidrogênio são importantes interações não-covalentes existentes nos sistemas biológicos, contribuindo para a especificidade das interações

receptor-ligante e para a manutenção das conformações bioativas de macromoléculas e das estruturas secundárias  $\alpha$ -hélices e as folhas- $\beta$  de proteínas. Nas biomoléculas essas interações são formadas entre átomos mais eletronegativos que o carbono, como o oxigênio e o nitrogênio, que atuam como aceitadores e doadores de hidrogênio (Fig. 3.14). A energia envolvida em ligações de hidrogênio é da ordem de 1 a 6 *kcal/mol*, razoavelmente superior à energia de ruído térmico, da ordem de 0,6 *kcal/mol*, para temperatura ambiente, porém inferiores às ligações iônicas e, bem menos ainda, às ligações covalentes.

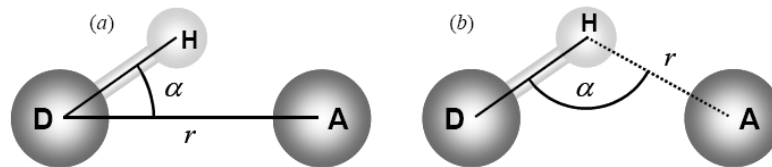


Figura 3.14: Ligações de hidrogênio. Dados  $r$  (distância) e  $\alpha$  (ângulo) são três as convenções para determinação dessas ligações. Em (a)  $r$  é tomado entre D-A inferior a 3,6 Å, com  $\alpha$  variando entre  $0 - 60^\circ$  e em (b)  $r$  é tomado entre H-A inferior a 2,7 Å, com  $\alpha$  variando entre  $120 - 180^\circ$ . A combinação de critérios entre (a) e (b) também pode ser utilizada.

### 3.15 Área de Interação Intermolecular

Uma outra análise feita foi a área da superfície do sítio ativo da proteína acessível ao ligante, calculando as áreas médias de contato de cada resíduo aminoácido com o respectivo ligante. Nessa análise fica caracterizada a importância de cada resíduo aminoácido para a interação protease-ligante, dando um importante subsídio para o desenvolvimento de um fármaco. Esta análise foi feita utilizando o programa SURFMDS (citado na seção 3.9.1). A Superfície Molecular do sítio ativo foi desenhada a partir de uma ponta de prova esférica correndo sobre a proteína com seus átomos representados por suas esferas de van der Waals (Connolly,



1983a). Dessa maneira foram calculadas a área de superfície acessível ao solvente (SAS) e a área de superfície excluída ao solvente (SES) (Connolly, 1983b) (Fig. 3.15).

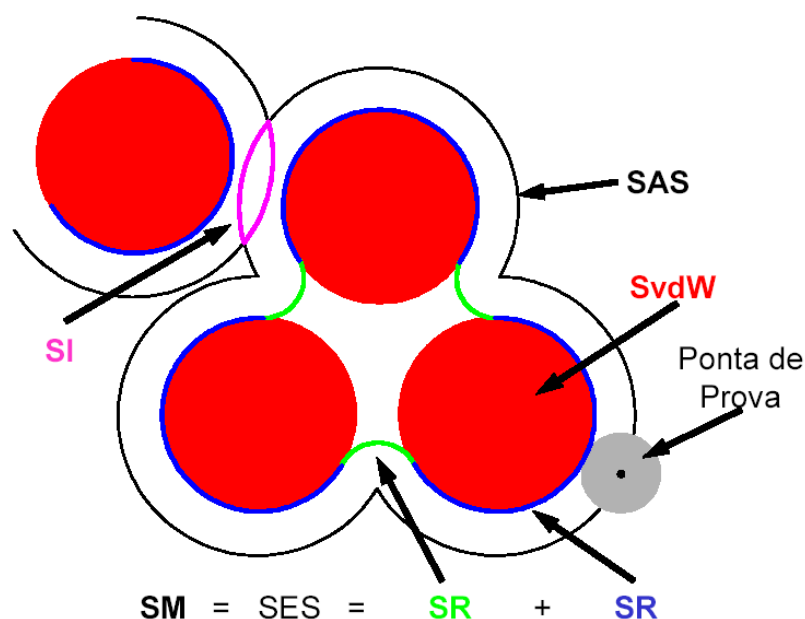


Figura 3.15: Superfície Molecular e Superfície Intermolecular. Representação bidimensional. Uma ponta de prova esférica corre por sobre a superfície de van der Waals (SvdW), gerando superfícies de contato (SC) e superfícies de reentrância (SR), que somadas compõem a superfície excluída ao solvente ou superfície molecular (SM). O centro de prova desenha a superfície acessível ao solvente (SAS). A superfície Intermolecular (SI) seria a intersecção das SAS de moléculas diferentes dentro do espaço definido pela ponta de prova.

A partir das superfícies moleculares da proteína e do ligante é possível determinar a superfície intermolecular como sendo a superfície de intersecção entre as SAS do ligante e a SAS da proteína, ou seja, somam-se as áreas dos átomos da proteína e do ligante próximos o suficiente para que não haja espaço para uma molécula de água. Neste trabalho foi empregado um raio de esfera de prova igual a 1,4Å.

# Capítulo 4

## Resultados do Estudo de Casos – Subtipos B e C Complexados ao Nelfinavir

### 4.1 Interação entre a PR e o nelfinavir

#### 4.1.1 Propriedades Globais

Tomando a protease do subtipo B (código de acesso no PDB: 1OHR) complexada com o ligante nelfinavir como molde para a construção dos modelos teóricos das estruturas tridimensionais das PR's mutantes, foram feitas simulações para reproduzir a dinâmica de interação entre a PR e o inibidor. Os seis sistemas complexados com o nelfinavir - consB, BD30N, consC, CD30N, CD30N/N83T e CD30N/N88D - foram examinados por simulações de DM durante 10 ns. Entretanto os dois primeiros nanosegundos foram para equilibração do sistema sendo apenas os 8 ns finais considerados para a etapa de análise. Certos parâmetros foram analisados a fim de se estabelecer relações entre os subtipos B e C que levem a entender as diferenças causadas pela mutação D30N em cada subtipo. Todas as 6 simulações foram realizadas com a protease em solvente explícito e contra-íons, utilizando condições periódicas de contorno, perfazendo cerca de 40.000 átomos cada sistema.

Para avaliar a estabilidade das simulações precisamos atentar para o com-

portamento dinâmico dos sistemas. Como pode ser visto na figura 4.1, o desvio quadrático médio (RMSD), ao longo das dinâmicas, dos átomos do esqueleto peptídico, com relação às estruturas iniciais (minimizadas) de quase todos os modelos possuem uma variação menor que 0,3 nm ao fim dos 8 ns. Somente os mutantes BD30N e CD30N/N83T apresentaram desvio maior ou igual que 0,3 nm. Esta análise permite acompanhar dinamicamente as variações da estrutura em relação à estrutura antes da DM.

As PR's consB, consC, CD30N e CD30N/N88D se mostraram mais estáveis, principalmente após 5 ns. Por outro lado, o mutante BD30N apresentou os maiores desvios durante toda a DM, chegando a mais de 0,35 nm em alguns momentos. Além disso, este sistema, após 9 ns apresentou um novo crescimento no desvio.

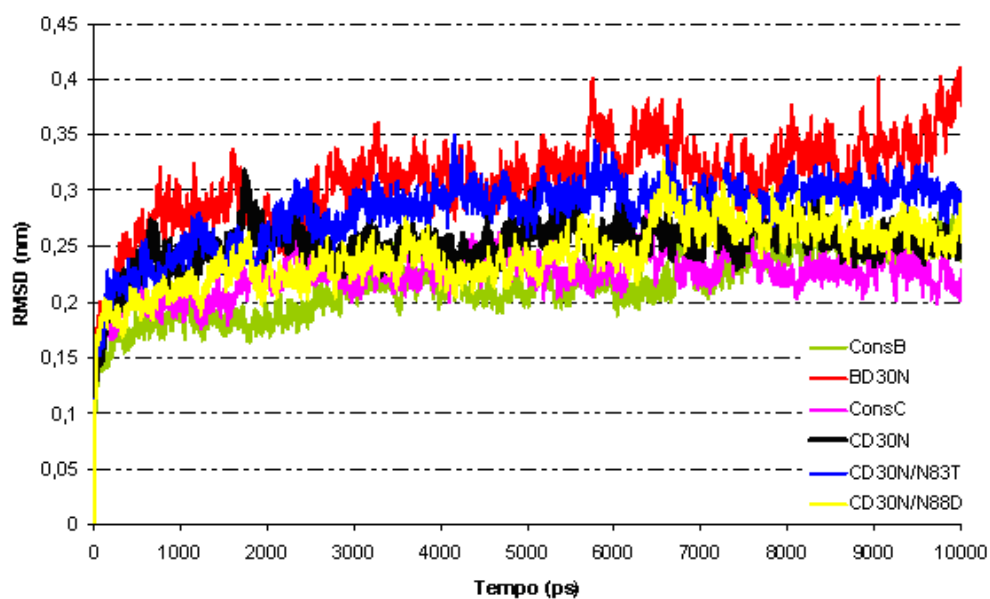


Figura 4.1: Desvio Médio Quadrático (RMSD) da PR, em relação à estrutura inicial (pós – minimização de energia), no tempo (10 ns de simulação): consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo).

Tabela 4.1: Valores Médios do RMSD (nm) do esqueleto da protease complexadas ao nelfinavir

<b>B</b>	<b>BD30N</b>	<b>C</b>	<b>CD30N</b>	<b>CD30N/N83T</b>	<b>CD30N/N88D</b>
0.22 ± 0.02	0.37 ± 0.04	0.22 ± 0.01	0.25 ± 0.01	0.29 ± 0.01	0.25 ± 0.02

A média dos RMSDs, calculada para os últimos 8 *ns* de simulação, varia de 2,2 Å a 3,7 Å. Os modelos consB e consC apresentaram o menor desvio médio – 2,2 Å – e os modelos BD30N e CD30N/N83T obtiveram os maiores desvios médios – 3,7 Å e 2,9 Å respectivamente – (Tabela 4.1).

#### 4.1.2 Flutuações dos Resíduos

Para melhor entender a origem dos desvios observados nas análises do item anterior, têm-se os gráficos da flutuação do RMSD por resíduo (RMSF) o que permite uma análise da contribuição de cada resíduo nos últimos 8 *ns* de simulação. Analisando tais contribuições (Figs. 4.2, 4.3, 4.4), temos que as regiões onde há maior variação para a PR BD30N são as próximas às isoleucinas 149 (também chamada de 50') e 50 (Fig. 4.3), entretanto as PRs mutantes do subtipo C não tiveram sua maior variação nesse resíduo, mas sim na glicina 16 (Fig. 4.4). Além disso, os desvios sofridos nas isoleucinas 50 e 50' de todas as PRs mutantes do subtipo C (0,25 nm) foram menores do que os sofridos pelo BD30N (acima de 0,35 nm). Essa região tem uma importância ímpar na ligação do substrato/inibidor, pois estes resíduos estabilizam o substrato/inibidor no sítio ativo, através de ligações hidrogênio intermediadas por moléculas de água estruturadas (que estão presentes nas estruturas determinadas por cristalografia); uma outra região da PR, próxima a GLY115, também apresentou grande desvio (Figs. 4.2, 4.3, 4.4). É importante notar que a região dos ASP catalíticos (25 e 124) possui uma variação pequena em todos os sistemas, essa região é essencial para a interação do substrato/inibidor.

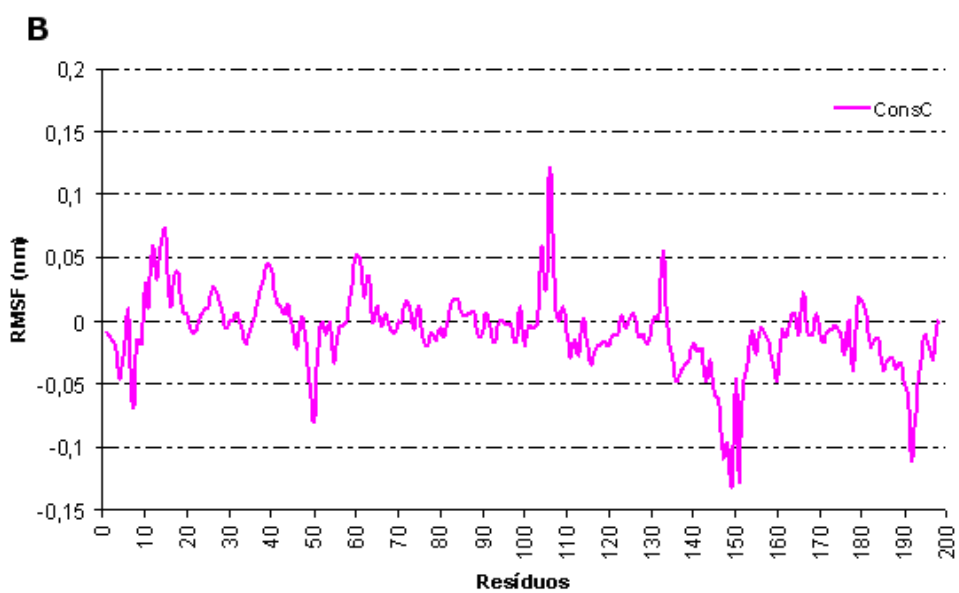
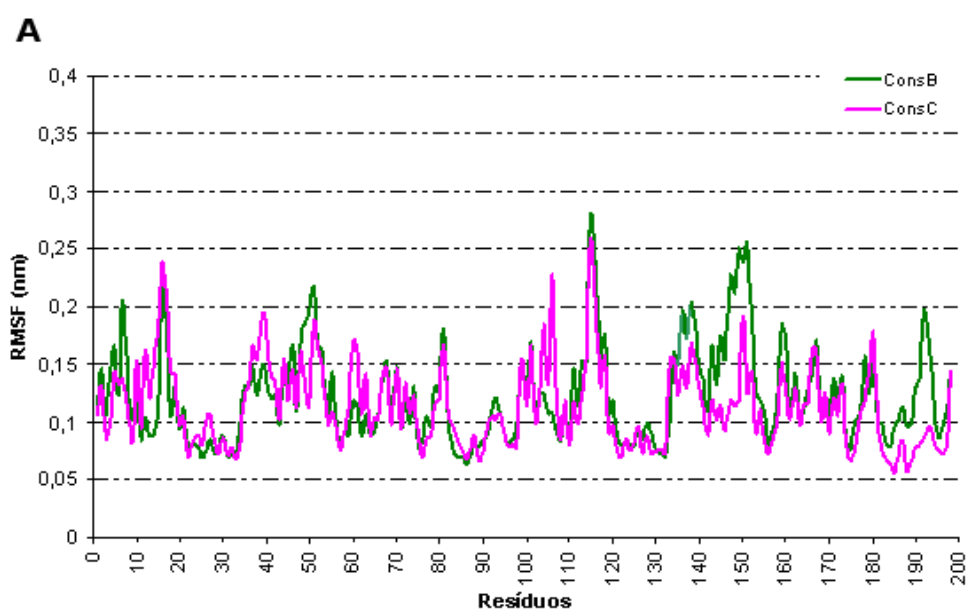


Figura 4.2: Em A, RMSF do ConsB (verde) e ConsC (rosa). Em B, diferenças entre o RMSF do subtipo C e B (ConsC-ConsB).

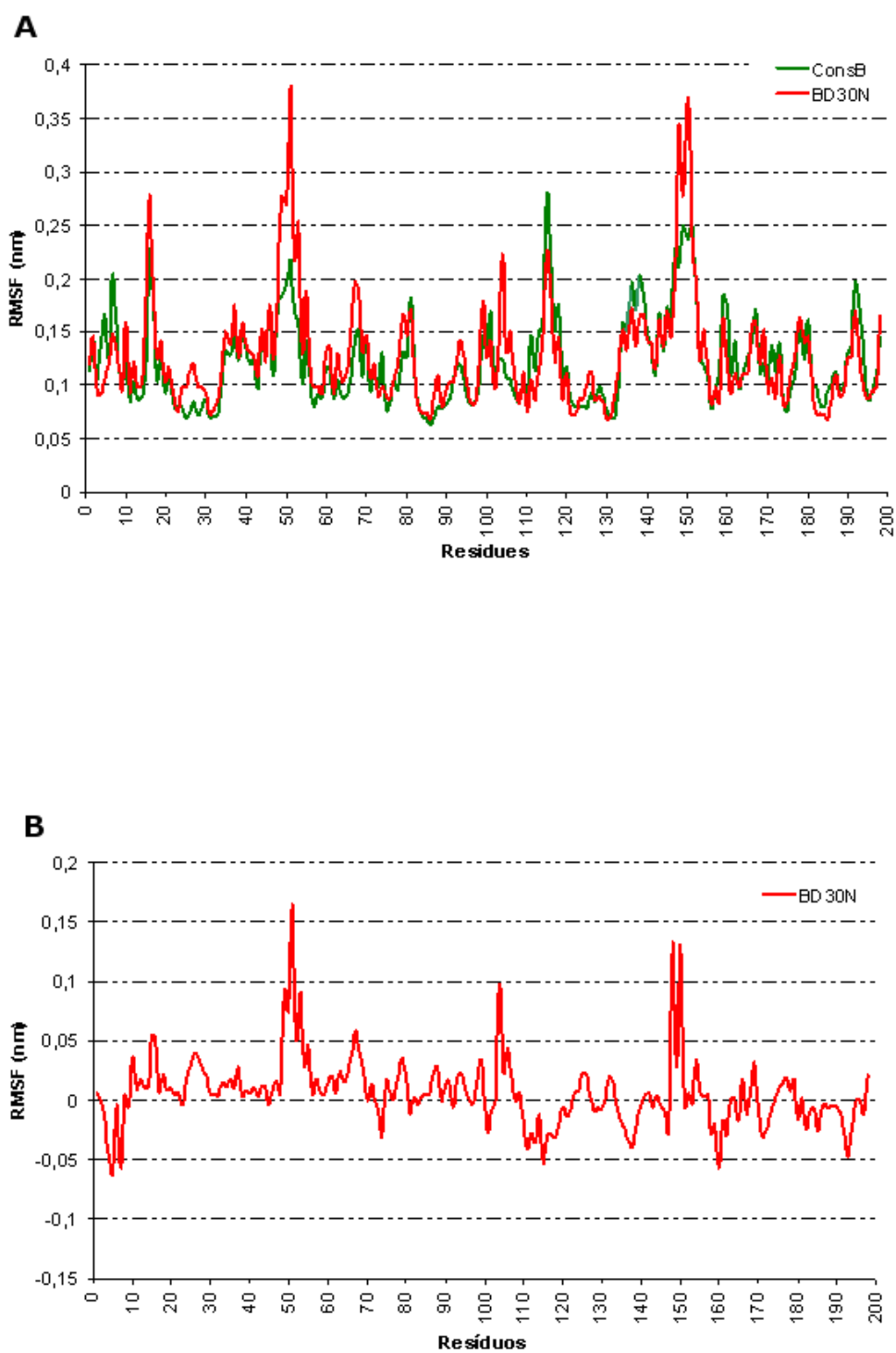


Figura 4.3: Em A, RMSF do ConsB (verde) e BD30N (vermelho). Em B, diferenças entre o RMSF do mutante BD30N e ConsB (BD30N-ConsB).

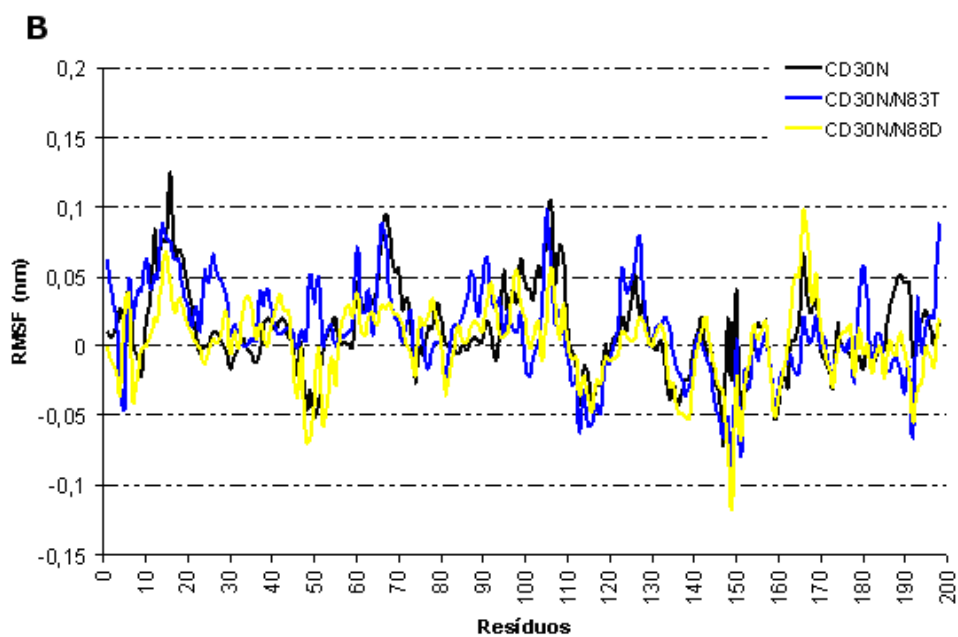
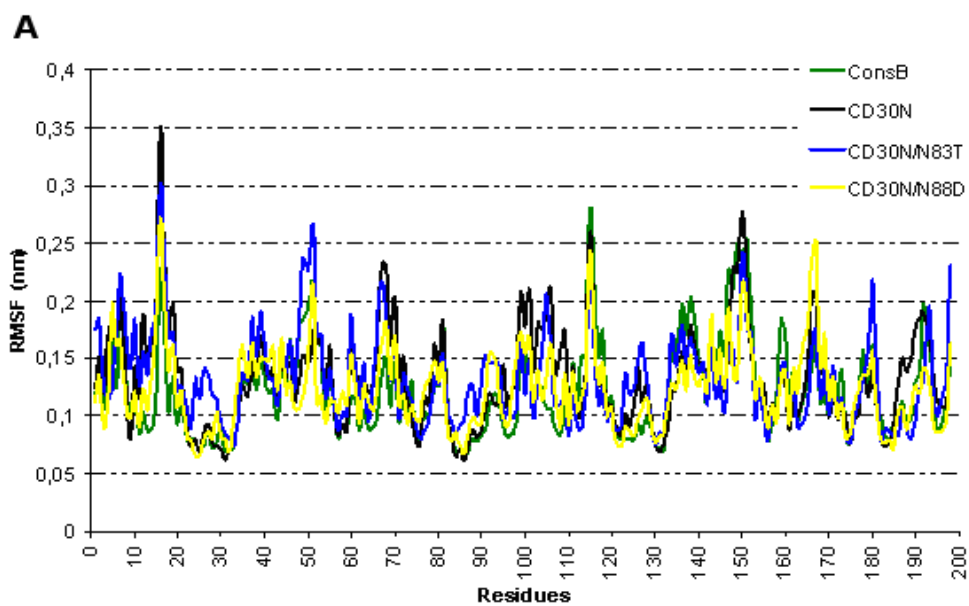


Figura 4.4: Em A, RMSF de quatro sistemas: ConsB (verde), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o RMSF do subtipo B e os demais. CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D-consB (amarelo)



### 4.1.3 Representação 3D do RMSD para os átomos do backbone

É possível ter uma idéia da flexibilidade do esqueleto peptídico das proteínas através de uma representação gráfica 3D do valor do RMSD associado ao mesmo. Apesar do inibidor estar no sítio ativo, nessa análise só tratamos da representação dos desvios da PR. Os desvios respectivos ao inibidor durante a DM serão tratados em uma seção a seguir.

Com a representação 3D do RMSD do esqueleto peptídico na figura 4.5, pode-se constatar visualmente que no mutante BD30N, os desvios na estrutura foram maiores, principalmente na região dos flaps. Nas PR's mutantes do subtipo C, o desvio foi ligeiramente menor na região das alças quando comparado ao observado no mutante BD30N.

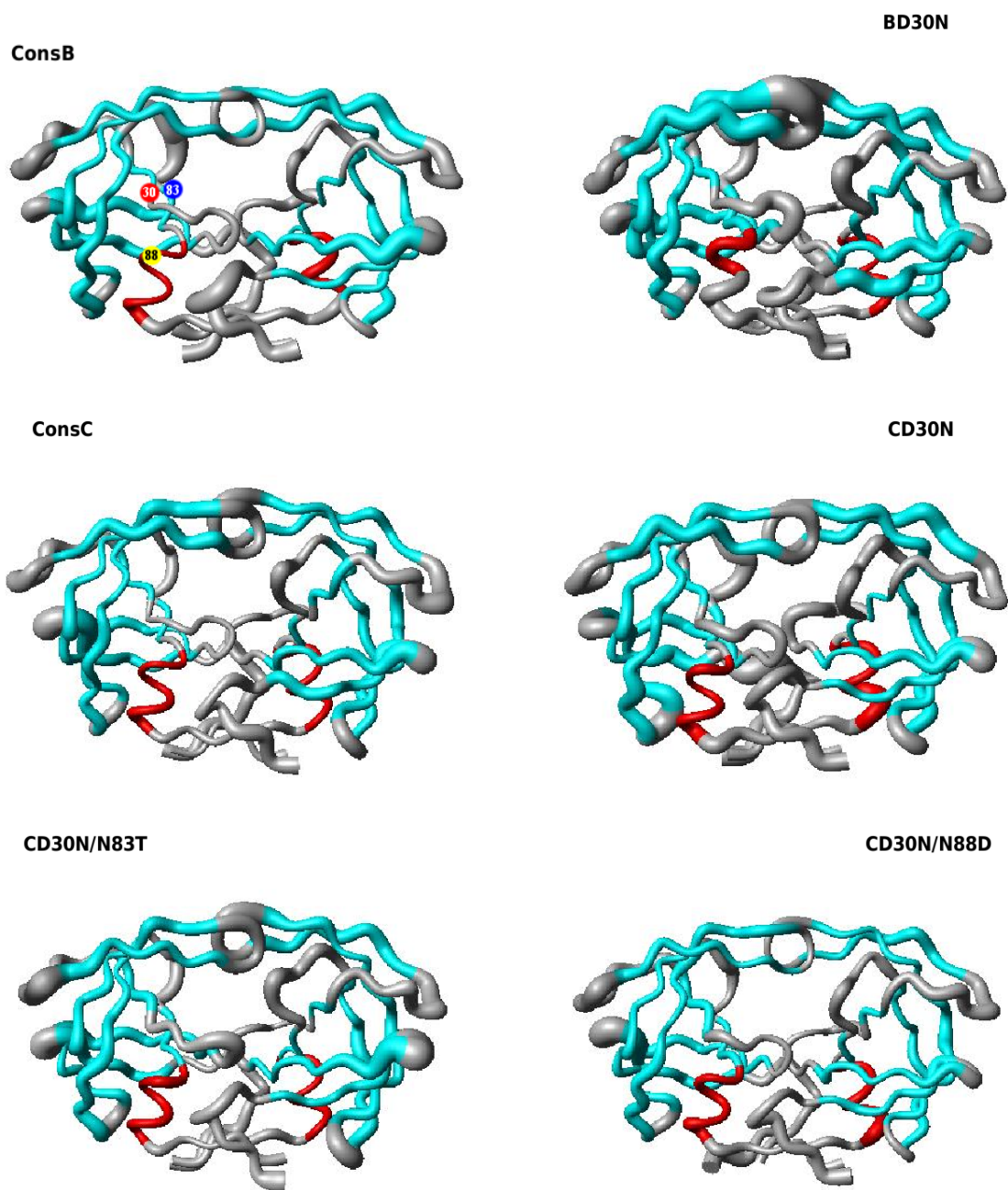


Figura 4.5: Representação 3D das flutuações do RMSD do esqueleto peptídico da PR. O RMSF foi representado por um tubo com a espessura variada, onde quanto mais espesso o tubo, maior o desvio. Foram coloridos por estrutura secundária: em vermelho,  $\alpha$ -hélice; em azul, folhas- $\beta$  e em cinza, loops e estruturas randômicas. As esferas representam as mutações: em vermelho, D30N; em azul, N83T e em amarelo, N88D.

#### 4.1.4 Variações nas alças da Protease

Uma questão importante quando se estuda a variação das alças é saber se houve abertura das mesmas. Para isso foi calculada a distância entre as pontas das alças (resíduos ILE50 e ILE149) e o centro catalítico.

Quando analisamos a Fig. 4.6 temos a variação temporal da distância entre as isoleucinas (ponta das alças) e o centro catalítico da PR. Esta medida representa como se comportaram essas alças durante a DM. Nota-se claramente que nos modelos consC e CD30N/N88D as alças não se afastaram dos Aspárticos catalíticos. No caso do CD30N enquanto a cadeia A manteve a distância invariável a cadeia B apresentou uma redução na abertura.

Já no consB ocorreu a separação (aumento na distância) das alças a partir de 5 *ns*, devido ao aumento na distância da alça da cadeia B. Entretanto somente as distâncias superiores a 1.6 *nm* são consideradas como tendo as alças na conformação semi aberta (Spinelli et al., 1991). Daí temos que somente os mutantes BD30N e CD30N/N83T obtiveram uma abertura significativa em suas alças. Além disso, comparando a abertura das alças desses mutantes pode-se perceber que a abertura no BD30N é maior que no CD30N/N83T.

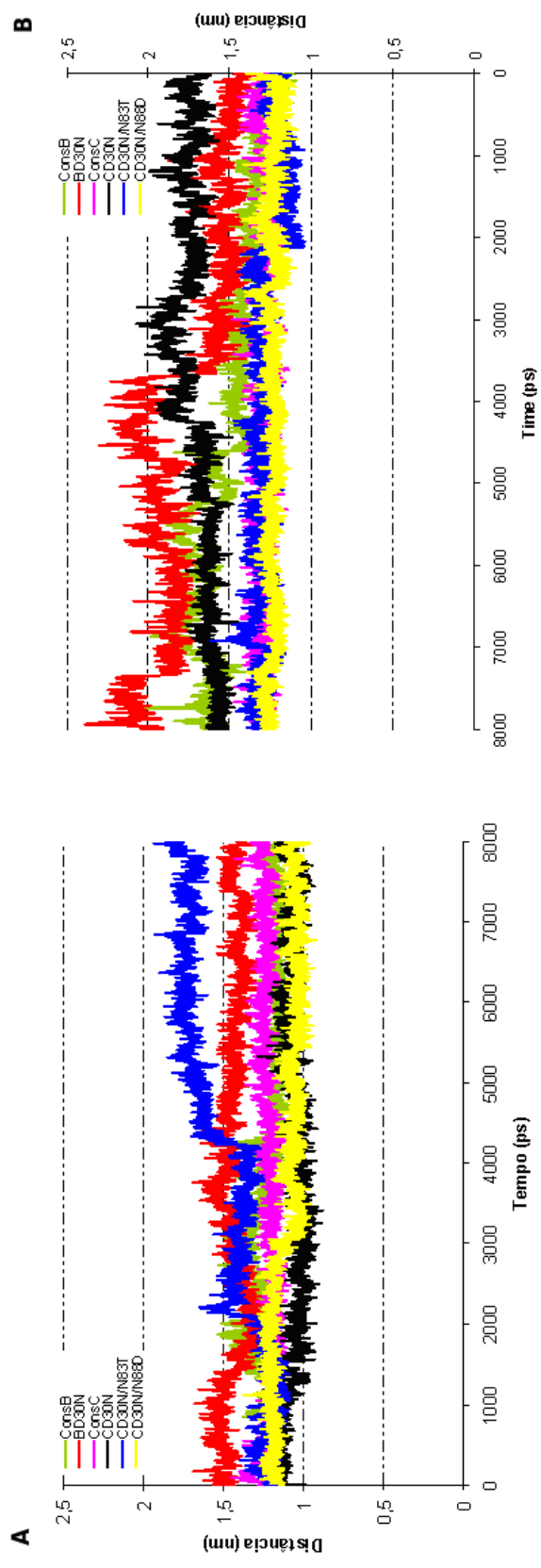


Figura 4.6: Distância entre as pontas das alças e o centro catalítico calculada para cada sistema. Em A, distância referente a alça da cadeia A (ILE 50). Em B distância referente a cadeia B (ILE 149). Em verde ConsB, em vermelho BD30N, em rosa ConsC, em preto CD30N, em azul CD30N/N83T e em amarelo CD30N/N88D. A figura foi invertida visando facilitar a visualização do gráfico.

#### 4.1.5 Variações estruturais na Compactação da PR Quando Complexada ao Inibidor

Para acompanhar a dinâmica das alterações na forma da proteína, monitorando os três eixos cartesianos (X, Y e Z), foram calculados os raios de giro para cada sistema. Através dessa análise é possível saber se houve compactação, expansão ou manutenção da forma da proteína ao longo da dinâmica.

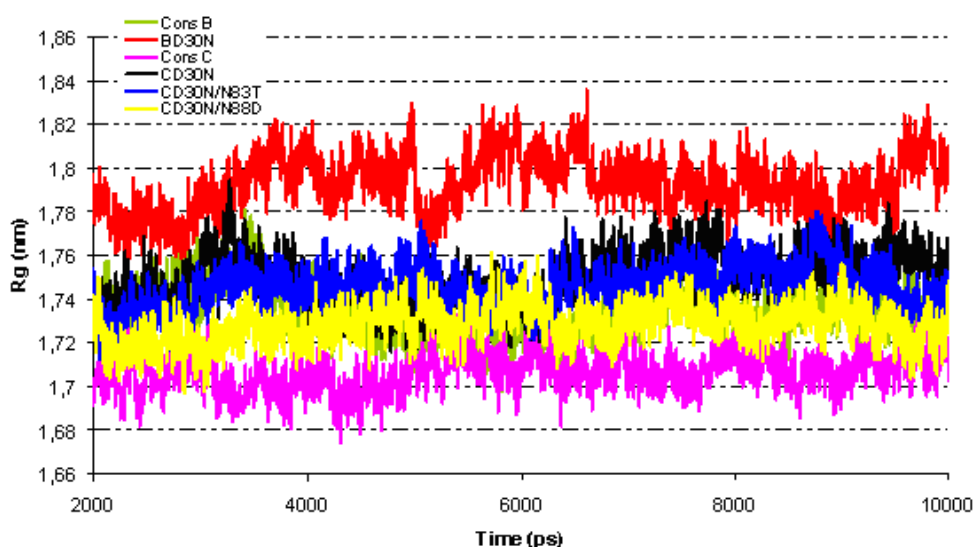


Figura 4.7: Representação do Raio de Giro de todos os sistemas durante a DM. ConsB (verde), BD30N (vermelho), ConsC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o subtipo B e os demais. ConsB-BD30N (vermelho), ConsC-ConsB (rosa), CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D (amarelo)

Com a Fig. 4.7 fica evidente a diferença entre o BD30N e os demais sistemas. Através dessa análise a mudança na forma dessa proteína fica evidente. O sistema BD30N ficou com valores de raio de giro em torno de 1,8 *nm* ao final da simulação enquanto os outros sistemas ficaram com valores abaixo de 1,76 *ns*. No caso dos

demais mutantes houve uma abertura das alças de forma mais branda em comparação com o BD30N, onde houve uma maior abertura levando a uma mudança excessiva na forma. Essa mudança pode ser acompanhada, tanto com os gráficos de raios de giro quanto com as análises das variações espaciais (RMSD, RMSD 3D e RMSF) das posições dos átomos durante a simulação.

Os momentos de inércia  $I$  para os eixos  $x$ ,  $y$  e  $z$  com valores próximos indicam uma forma esférica, enquanto dois valores similares e um componente muito menor indicam um elipsóide. A compactação total de uma proteína pode ser definida por um valor  $\alpha$ , calculado da seguinte forma (Tieleman et al., 2000):

$$\alpha = \frac{2I_1 - I_2 - I_3}{I_1 + I_2 + I_3} \quad (4.1)$$

Quanto mais próximo do zero for esse valor, mais esférica é a proteína. Os momentos de inércia  $I_1$ ,  $I_2$  e  $I_3$  se referem aos eixos  $x$ ,  $y$  e  $z$ ; sendo  $I_1$  o maior componente. Na tabela 4.2 podemos acompanhar o comportamento médio dos momentos de inércia durante a DM.

Tabela 4.2: **Valores médios dos momentos de inércia durante a DM**

Modelos				
	$I_1$	$I_2$	$I_3$	$\alpha$
<b>ConsB</b>	57405.7 $\pm$ 904.9	48557.8 $\pm$ 1148.2	24278.9 $\pm$ 483.4	0.32228
<b>BD30N</b>	61874.1 $\pm$ 1016.8	52549.2 $\pm$ 1176.2	24731.3 $\pm$ 436.5	0.33392
<b>ConsC</b>	54520.5 $\pm$ 765.7	45312.8 $\pm$ 721.1	25494.5 $\pm$ 379.9	0.30506
<b>CD30N</b>	58286.9 $\pm$ 1043.2	48782.8 $\pm$ 847.3	24778.2 $\pm$ 418.7	0.32623
<b>CD30N/N83T</b>	58314.8 $\pm$ 910.1	47666.5 $\pm$ 856.7	25010.1 $\pm$ 445.4	0.33554
<b>CD30N/N88D</b>	56834.5 $\pm$ 852.4	46984.3 $\pm$ 799.5	24729.4 $\pm$ 390.3	0.32637

O modelo BD30N apresenta o maior valor médio de  $I_1$  e  $I_2$  enquanto o consC possui o menor. Todos os outros sistemas apresentaram momentos de inércia similares. Quando se calcula a compactação total da PR, complexada ao nelfinavir, obtendo-se um valor  $\alpha$ , nota-se que o modelo C apresenta o menor valor, ou seja, possui uma forma um pouco mais esférica que as demais PRs.

#### 4.1.6 Comportamento do Nelfinavir durante a DM

Para melhor entender as interações do inibidor com a PR durante a DM, foram feitas análises que dizem respeito à estabilidade do inibidor no sítio ativo da proteína. Para ocorrer uma melhor interação da PR com o inibidor, ou seja, para que a afinidade do inibidor para com a PR seja alta, é necessário que o inibidor fique o mais estável possível no sítio ativo da enzima. A figura 4.8 mostra cada grupo (P2, P1, P1', P2') do inibidor.

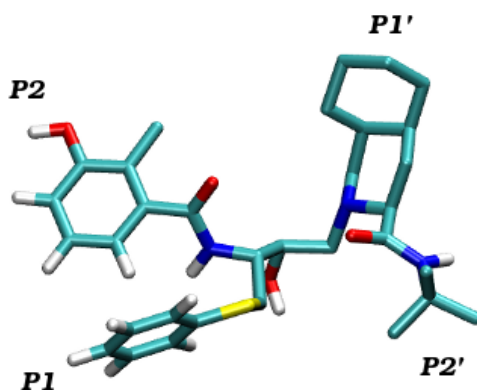


Figura 4.8: Estrutura 3D do nelfinavir com seus grupos de interação P2, P1, P1' e P2'.

Já a figura 4.9 dá uma idéia da estabilidade de cada um desses grupos, nessa figura foram sobrepostas estruturas do inibidor a cada 100 ps. Pode-se perceber que o inibidor está mais instável nos grupos P1 e P2 das PRs CD30N/N83T e BD30N,

Com o propósito de analisar os desvios de cada átomo do inibidor durante a dinâmica, foram calculados os RMSF para esses átomos, que foram então separados por grupo no inibidor (Fig. 4.10).

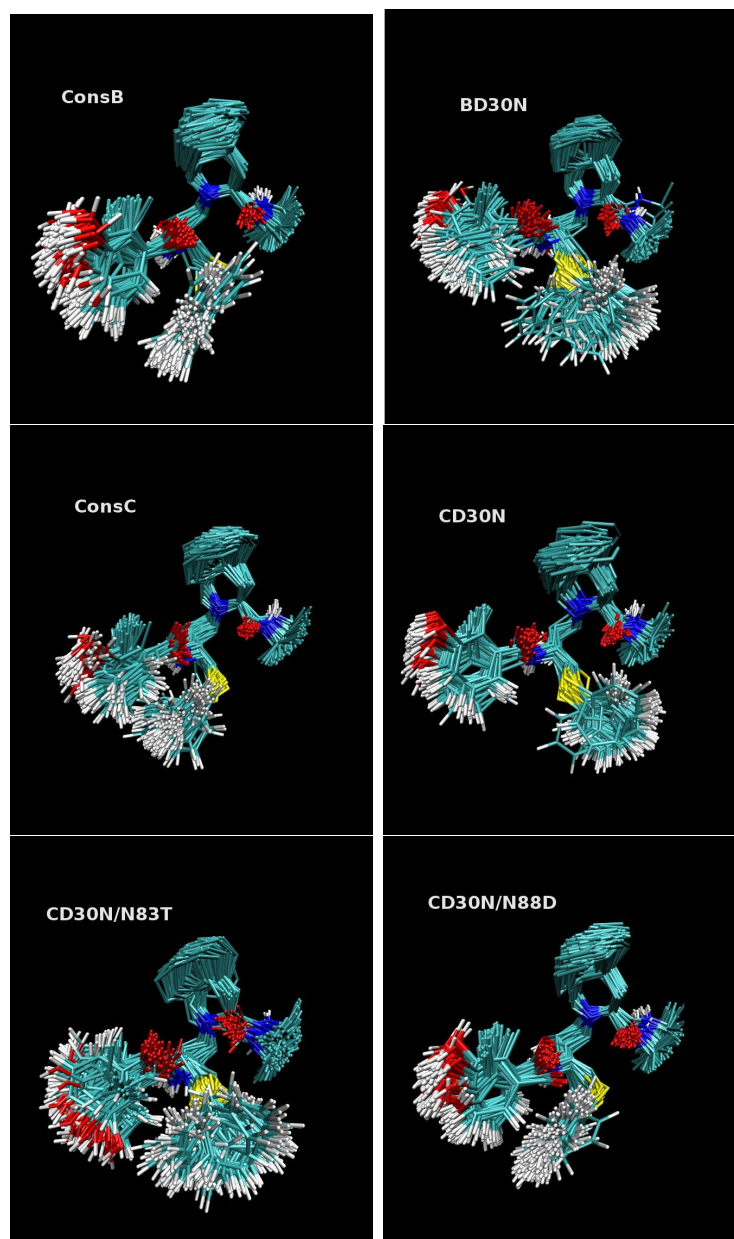


Figura 4.9: Sobreposições das estruturas do inibidor durante a DM. Representação das ligações químicas do nelfinavir, coloridas pelo tipo de átomo: branco (H), verde (C), azul (N), vermelho (O) e amarelo (S).



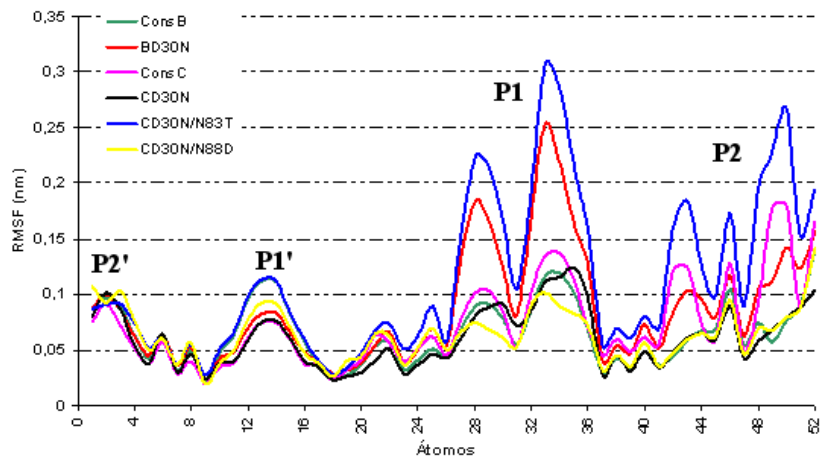


Figura 4.10: Flutuações do RMSF calculadas para os átomos do inibidor em cada sistema: consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D(amarelo). Os grupos P2' (resíduos 1 a 8), P1' (resíduos 9 a 22), P1 (resíduos 23 a 36), P2 (resíduos 37 a 52) do inibidor estão indicados, tanto no gráfico quanto na representação.

Focando-se no comportamento do nelfinavir durante as simulações pode-se perceber que o RMSF, por átomo, do inibidor que mostrou maior flutuação foi o do nelfinavir do mutante CD30N/N83T (Fig. 4.10). Todas as PR's apresentaram o nelfinavir com um maior desvio na região P1 exceto a PR consC que apresentou o inibidor com maior desvio na região P2. Pode-se perceber também que o átomo 21 (O21) que é o responsável pela interação com os aspárticos catalíticos da protease, manteve-se bem estável em todas as estruturas.

#### 4.1.7 Trajetórias das Dinâmicas

Para a melhor demonstração das trajetórias da dinâmica da PR foram tiradas “fotos instantâneas” dos quadros, a cada 2 *ns* e estas foram expostas em seqüência como em um filme de película. Isso permite acompanhar a evolução do sistema de uma forma simples e satisfatória. Além disso, os modelos foram alinhados de modo a formarem uma matriz. Com isso, suas colunas permitiram a comparação estrutural de cada modelo a cada 2 *ns*.

Na Figura 4.11, pode-se notar que a partir de 4 *ns*, ocorre um leve afastamento entre as alças no mutante BD30N e em 6 *ns* para o modelo CD30N/N83T. Embora já nos primeiros nanossegundos tenha ocorrido uma ligeira abertura dos “flaps ” no CD30N, essa conformação com as alças abertas flutuou com poucas variações de 0 a 8 *ns*.

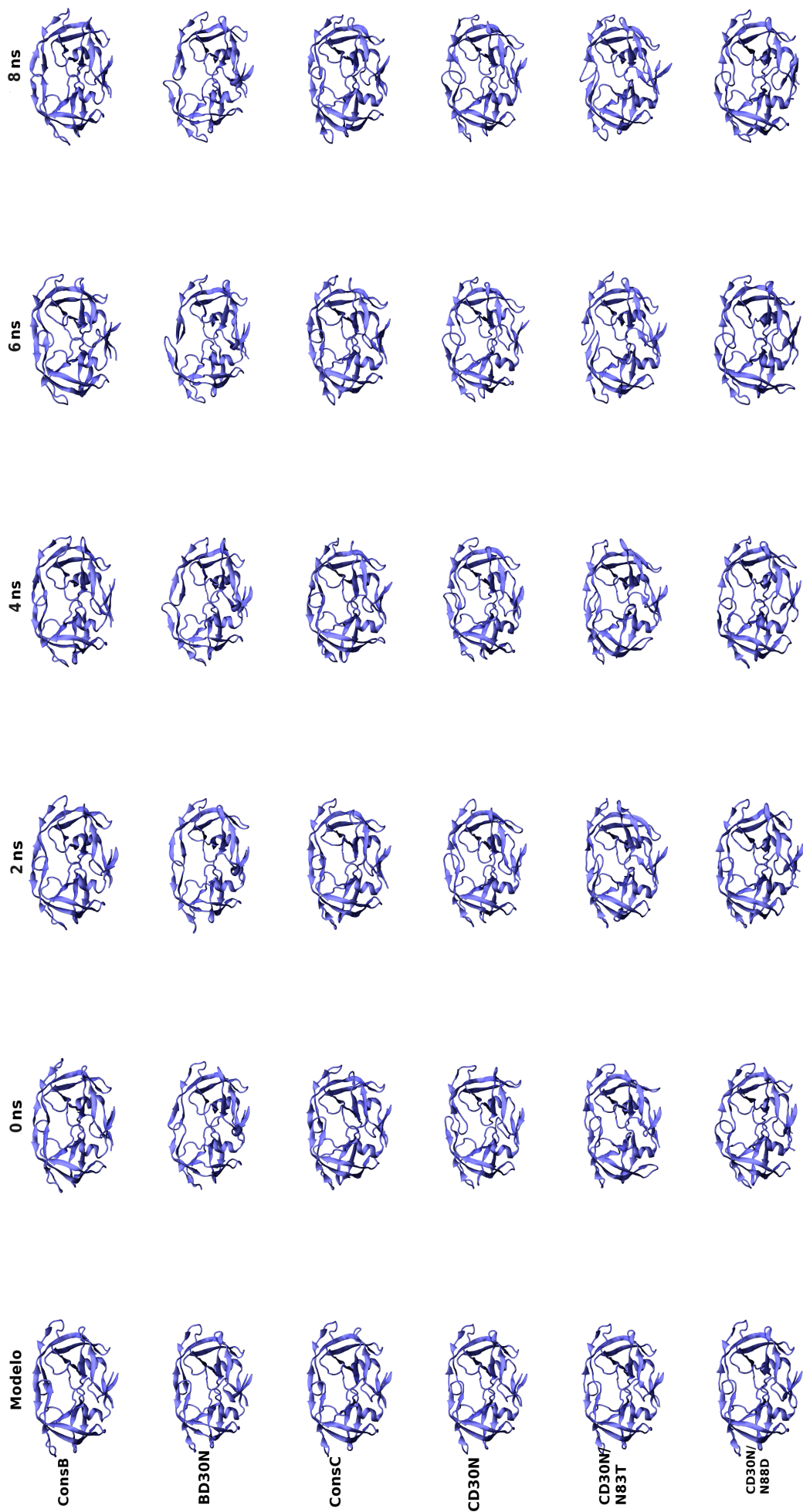


Figura 4.11: “Snapshots” da estrutura das PRs, representando somente o esqueleto peptídico, demonstrando a trajetória das PRs durante os 8 ns de simulação.

#### 4.1.8 Ligações de hidrogênio entre a PR e o Nelfinavir

As interações não-covalentes são essenciais para a manutenção da estrutura da proteína, para os processos de reconhecimento e para as interações proteína/ligante e proteína/substrato. As ligações de hidrogênio são um tipo especial de interação entre átomos não-ligados e possuem um papel muito importante na afinidade do inibidor pela proteína. Nesta dissertação, foram investigadas as ligações de hidrogênio entre a PR e o nelfinavir e entre a PR e o substrato. Essas ligações de hidrogênio foram classificadas em (D) diretas e intermediadas pela água (WM).

Na Tabela 4.3 se encontra o número médio de ligações de hidrogênio entre a PR e o nelfinavir, tanto diretas quanto intermediadas por moléculas de água.

Além disso, foi identificada e discriminada cada ligação durante os 8 *ns* da dinâmica (Tabela 4.4).

Tabela 4.3: **Número médio de Ligações de Hidrogênio entre a PR e o nelfinavir.**

Modelos						
	ConsB	BD30N	ConsC	CD30N	CD30N/N83T	CD30N/N88D
<b>D</b>	3.42	2.39	3.11	3.02	3.70	3.49
<b>WM</b>	0.44	0.84	1.74	1.72	2.24	0.59
<b>Total</b>	3.86	3.23	4.85	4.74	5.94	4.08

Número médio de ligações de Hidrogênio entre a PR e o nelfinavir durante a simulação. As ligações foram separadas em: diretas (D), intermediadas por água (WM) e total.

A Tabela 4.3 mostra uma diminuição do número de ligações de hidrogênio diretas comparados o controle B e o mutante BD30N. Na Tabela 4.3 também se verifica que o total de ligações de hidrogênio ( $D + WM$ ) do mutante BD30N é menor que todos os outros sistemas. Já quando comparamos o consC e os seus respectivos mutantes observa-se que o número de ligações diretas não aumenta significativamente em nenhum dos modelos. Além disso, quando foram analisadas as ligações intermediadas por água somente o CD30N/N88D apresentou uma redução significativa. Note que embora o modelo CD30N/N83T tenha apresentado um

aumento no número de ligações total esse aumento não é significativo, visto que, ele foi devido a aumentos não relevantes nas ligações diretas e indiretas. Por outro lado, os mutantes do subtipo C possuem maior número de ligações que o mutante BD30N.

Observando a Tabela 4.4, podemos destacar a formação de ligações de hidrogênio com os ácidos aspárticos (ASP25OD2 e ASPH124OD2), aminoácidos fundamentais para a ligação do nelfinavir. Pode ser visto que a porcentagem de formação de pontes de hidrogênio entre o inibidor (NEL200O21) e esse aminoácido é alta para todos os modelos. Porém, os consB e C apresentaram porcentagens de formação de pontes de hidrogênio, com os ácidos aspárticos, muito próximas de 100%.

Outro resultado importante visualizado na Tabela 4.4 é que a mutação D30N reduz em quase todos os mutantes a ligação direta entre o aminoácido da posição 30 (ASN/ASP30N) e o O46 do nelfinavir (NEL202O46). Ocorrendo o mesmo com a ligação entre o N37 do inibidor e o ASP25 da PR (NEL202N37 – ASP25OD2).

Após mapearmos todas as ligações de hidrogênio que ocorreram durante DM por mais de 25% do tempo (2 *ns*) comparamos essas com as encontradas no cristal. Pode-se perceber na figura 4.12 que a interação existente entre o ASP30 e o O46 já se encontrava presente na estrutura cristalográfica. Entretanto o resíduo ASP30 sofre uma alteração no átomo que interage com o átomo do inibidor (repare que no cristal essa interação ocorre através do átomo OD2 enquanto na DM a maior permanência se deu com o átomo N).

A ligação de hidrogênio entre o ASP25 e o átomo O21 observado durante quase toda DM também está presente na estrutura cristalográfica (Fig 4.12). Além disso, o mesmo átomo ASP25 faz um ligação com o N37 o qual destacamos a sua menor permanência de tempo nos mutantes quando comparados aos consensos. Repare que a ligação entre a Gly27 e o N37, encontrada no cristal, não permanece por mais de 2 *ns*, facilitando a ligação do N37 ao ASP25.

O outro áspartico catalítico (ASP124) também interage com o átomo O21 do nelfinavir, contudo à medida que a ligação ASP25–O21 é mais duradoura a

permanência ASP124–O21 é um pouco menos frequente.

Tabela 4.4: Ligações de hidrogênio (%) entre a PR e o nelfinavir.

Substícios		Átomos													
		Inibidor - Receptor		ConsB		BD30N		ConsC		CD30N		CD30N/N83T		CD30N/N88D	
		D	WM	D	WM	D	WM	D	WM	D	WM	D	WM	D	WM
P2'	NEL199O8 -ILE149N	*	*	*	*	*	*	*	*	*	*	0.9	52.8	0.5	2.8
P1'	NEL200O21 -ALA28N	*	*	*	*	0.1	*	*	*	*	*	27.6	*	*	*
P1'	NEL200O21 -ASP124OD2	3.8	*	*	*	29.9	*	*	*	*	*	*	*	4.4	*
P1'	NEL200O21 -ASP25OD1	7.8	*	62.9	*	40.6	*	62.4	0.5	76.1	0.2	44.8	*	*	*
P1'	NEL200O21 -ASP25OD2	100	*	73.8	*	99.9	*	57.5	0.4	62.1	0.3	71.5	*	*	*
P2	NEL202O46 -ASP29N	31.9	0.7	25.9	6.6	9.4	5.9	29.4	7.7	27.1	1.3	14.3	6.5	6.3	6.3
P2	NEL202O46 -ASN/ASP30N	26	4.9	*	*	12.7	56.6	1.7	30.7	2.8	0.1	11.9	12.4	11.9	6.3
P2	NEL202O46 -GLY48N	16.6	11.7	0.1	1.1	2.2	2.9	28.3	13	13.7	14.3	20.3	12.4	20.3	12.4
P2	NEL202O40 -GLY49N	1.1	5.1	*	*	0.1	*	9	40	0.7	*	0.2	*	0.2	*
P2	NEL202O40 -ILE50N	*	*	21.5	38.2	1.9	87.2	0.2	13.7	0.5	0.5	4.9	*	4.9	*
P2	NEL202O40 -ILE149N	*	*	*	*	*	*	*	*	8.1	41.3	*	*	*	*
P2	NEL202O40 -GLY150N	*	*	*	*	*	*	*	*	26.6	52.3	*	*	*	*
P2	NEL202N37 -ASP25OD1	*	*	*	*	*	*	34.7	17.8	*	*	30	*	30	*
P2	NEL202N37 -ASP25OD2	81.8	*	*	*	69.7	*	27.3	21	*	*	59	*	59	*
P2	NEL202O46 -ASP29OD1	12.2	2	17.9	0.8	6.4	7.3	31.6	7.9	39.3	2.4	10.9	5.1	10.9	5.1
P2	NEL202O46 -ASP29OD2	18.3	3.6	23.5	0.9	3.3	3.5	19.3	8.7	45.4	4.3	23.4	7.8	23.4	7.8

Percentagem da Permanência das ligações de hidrogênio entre a PR e o nelfinavir durante os 8 *ns* de simulação de DM. Somente ligações de hidrogênio que permaneceram mais de 25% do tempo testado foram representadas. Quando a ligação de hidrogênio de um dos modelos apresentava valor maior que 25% também foi mostrado a ligação dos demais modelos, mesmo se estes tivessem menos de 25%.

\* – Os asteriscos denotam ausência de ligações detectadas.

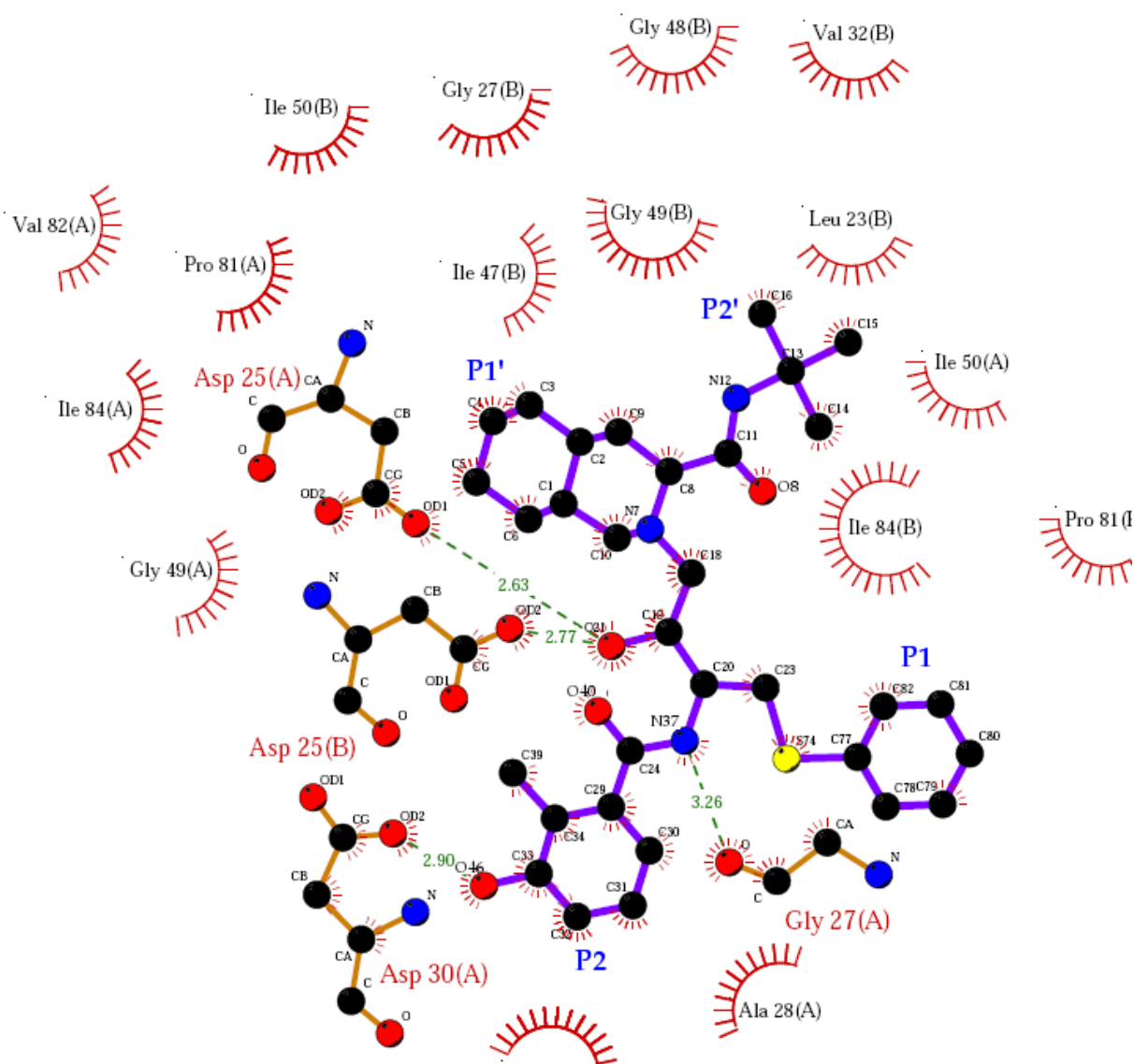


Figura 4.12: Esquema em 2D do sítio ativo da estrutura cristalográfica 1OHR. Estão representados os átomos do inibidor, exceto os hidrogênios. Em verde estão representadas as ligações de hidrogênio, com seus respectivos comprimentos. Cada aminoácido que está presente na superfície intemolecular de interação da PR com o nelfinavir está representado em vermelho.

\* – Os resíduos seguidos por (A) representam os aminoácidos de 1 a 99 e os seguidos por (B) representam os de 100 a 198.



#### 4.1.9 Superfície de Contato Intermolecular

Para melhor entender a ligação do nelfinavir às PR's dos diferentes subtipos, foram calculadas as áreas de interação entre a PR e o inibidor, sendo discriminada por resíduo de aminoácido presente na interface PR/inibidor. Com isso, pôde-se verificar padrões de interação e identificar os resíduos envolvidos nesta interface, além dos que estão formando ligações de hidrogênio com o inibidor (discutidos na seção 4.1.8).

Pode-se observar três regiões bem definidas em torno dos resíduos 25, 50 e 82, na cadeia A, e três regiões bem definidas em torno dos resíduos 124, 149, 181, na cadeia B (Fig. 4.13), definindo um padrão de interação destas proteases com o nelfinavir. Essas regiões estão centradas nos resíduos Asp25/124 (centro catalítico), Ile50/149 (alças) e Val82/181 (laterais do sítio ativo). Nota-se que o padrão de interação é semelhante entre os subtipos, sendo praticamente os mesmos resíduos participantes da interação PR/inibidor, com algumas pequenas variações.

A mutação D30N diminuiu significativamente a área de contato entre o nelfinavir e o mutante BD30N na região ao redor do sítio catalítico da cadeia A (resíduos 25, 28, 30, 47, 80, 81, 82 e 84). Já o mutante CD30N/N83T teve seus contatos significativamente diminuídos somente próximo ao aspártico 124 (resíduos 124, 126, 127, 128). Por outro lado, o modelo CD30N/N88D demonstrou comportamento semelhante ao consB. Já o CD30N apresentou uma redução significativa somente nos resíduos 122 e 181, e o consC no 30 e 122.

A assimetria observada entre as cadeias A e B era esperada, visto que o nelfinavir é assimétrico quanto a seus grupos laterais.

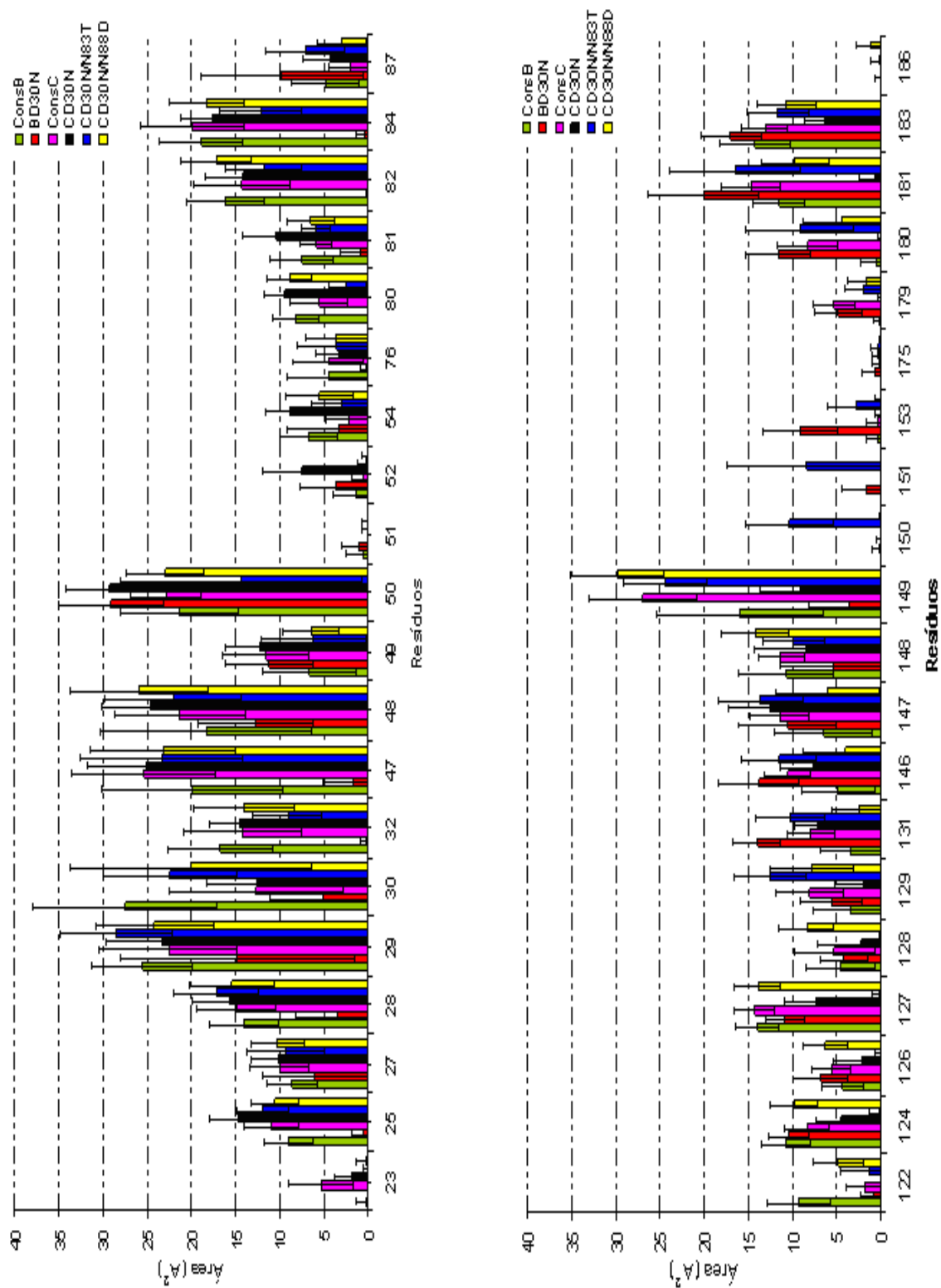


Figura 4.13: Área da Superfície de Contato Intermolecular entre a PR e o nelfinavir. Em barra estão representados os valores médios e em linha o desvio-padrão. Somente os contatos com mais de 5 Å<sup>2</sup> de área foram representados.

Para melhor comparação, a área total de cada um dos subsítios das diversas PR's foi calculada. Os subsítios das PR's foram definidos como as regiões da proteína que interagiram com as regiões do inibidor P2', P1', P1 e P2. Assim, temos os subsítios S2', S1', S1, S2. O resultado obtido encontra-se nas Tabelas 4.5 e 4.6. Pode-se observar que ocorrem diferenças significativas na área de contato dos subsítios dessa enzima para os diferentes modelos. De maneira geral, ocorre uma redução nos contatos dos modelos mutantes em relação ao seu tipo selvagem (Tab. 4.5). O modelo BD30N foi o que apresentou a menor área de contato, seguidos pelos CD30N e CD30N/N83T. O CD30N/N88D foi o único mutante que apresentou maior área que os consB e consC, o que não era esperado, visto ser este um mutante resistente.

A partir da área de contato hidrofóbica entre PR e inibidor, foi possível observar que nos mutantes do subtipo C (CD30N, CD30N/N83T e CD30N/N88D) boa parte da redução na área de contato total deu-se devido á perda de contatos hidrofóbicos (Tab. 4.6). Nota-se que no mutante CD30N/N83T a redução nos contatos hidrofóbicos foi maior que a redução total indicando um ligeiro aumento nos contatos hidrofílicos desse mutante. Além disso, pode-se perceber que o modelo CD30N/N88D não sofreu redução nos contatos hidrofóbicos, o que justifica a não-redução na área de contato total.

Por outro lado, embora o mutante BD30N tenha apresentado a maior redução nos contatos hidrofóbicos (Tab. 4.6) em relação ao seu tipo selvagem, tal decréscimo representou apenas 50% da redução da área de contato total. Dessa forma, a mutação D30N reduz significativamente a interação dos contatos hidrofóbicos e hidrofílicos no subtipo B, mas no subtipo C reduz somente os hidrofóbicos.

Podemos notar que para quase todos os modelos os maiores valores de área encontram-se nos subsítios S1', S1 e S2. Esses resultados ressaltam a importância desses subsítios no planejamento de novos fármacos. Vale ressaltar que as regiões do inibidor P1 e P2 foram as que sofreram os maiores RMSFs (Fig 4.10).

As regiões S2' e S2 apresentaram redução de contato em todos os mutantes

quando comparados com seus respectivos consensos. Contudo o modelo CD30N foi o que mais teve sua área reduzida no subsítio S2', sendo seguido pelo CD30N/N83T, BD30N e CD30N/N88D. Por outro lado, notou-se que o modelo BD30N foi o que apresentou uma redução mais drástica na região S2.

Tabela 4.5: Área de Contato ( $\text{\AA}^2$ ) entre a PR e o nelfinavir

Área de Contato Intermolecular entre a PR e o nelfinavir. Entre parênteses podem ser visualizados os aumentos (+) ou decréscimos (-) de cada um dos subsítio dos mutantes em relação ao seu respectivo tipo selvagem.

Subsítio	ConsB	BD30N	ConsC	CD30N	CD30N/N83T	CD30N/N88D
<b>S2'</b>	108.5	91.1 (-17.4)	121.3	69.9 (-51.4)	101.4 (-19.9)	116.4 (-4.9)
<b>S1'</b>	190.1	120.2 (-69.9)	169.9	167.5 (-2.4)	143.2 (-26.7)	190.6 (+20.7)
<b>S1</b>	174.1	170.9 (-3.2)	168.9	131.6 (-37.3)	167.3 (-1.6)	176.3 (+7.4)
<b>S2</b>	145.7	78.5 (-67.2)	169.7	146.3 (-23.4)	168.5 (-1.2)	162.1 (-7.6)
<b>Área Total</b>	618.4	460.7 (-157.7)	629.8	515.3 (-114.5)	580.4 (-49.4)	645.4 (+15.6)

**S2** - Ile50, Ile52, Ile81, Ile82, Ala127, Ala128, Asp129, Asp131, Ile146, Gly147, Gly148, Ile149, Gly150, Gly151, Ile183

**S1** - Leu23, Asp25, Gly27, Ala28, Val32, Ile47, Gly48, Ile50, Gly52, Ile54, Thr80, Pro81, Val82, Ile84, Leu122, Asp124, Gly126, Ala127, Gly147, Gly148, Ile149, Val181, Ile183

**S1'** - Asp25, Gly27, Ala28, Gly48, Gly49, Ile50, Arg107, Leu122, Asp124, Ala127, Val131, Ile146, Gly147, Ile149, Gly151, Ile153, Pro178, Thr179, Pro180, Val181, Ile183

**S2'** - Asp25, Gly27, Ala28, Asp29, Asp30, Val32, Ile47, Gly48, Gly49, Ile50, Ile84, Arg87, Asp124, Ile149, Gly150, Val181

Tabela 4.6: Área de Contato ( $\text{Å}^2$ ) Hidrofóbica entre a PR e o nelfinavir

Área de Contato Hidrofóbica entre a PR e o nelfinavir. Entre parênteses podem ser visualizados os aumentos (+) ou decréscimos (-) de cada um dos subsítio dos mutantes em relação ao seu respectivo tipo selvagem.

Subsítio	ConsB	BD30N	ConsC	CD30N	CD30N/N83T	CD30N/N88D
<b>S2'</b>	75.7	58.7 (-17)	77.5	37.7 (-39.8)	46.8 (-30.7)	68.3 (-9.2)
<b>S1'</b>	123.8	58.1 (-65.7)	103.5	108.2 (+4.7)	83.0 (-20.5)	126.3 (+22.8)
<b>S1</b>	115.5	144.9 (+29.4)	110.6	97.2 (-13.4)	126.4 (+15.8)	120.3 (+9.7)
<b>S2</b>	52.4	23.2 (-29.2)	86.2	52.6 (-33.6)	71.9 (-14.3)	66.3 (-19.9)
<b>Total</b>	367.4	284.9 (-82.5)	377.8	295.7 (-82.1)	328.1 (-49.7)	381.2 (+3.4)

Resíduos hidrofóbicos – Ala, Leu, Val, Ile, Pro, Phe, Met e Trp

#### 4.1.10 Padrões das estruturas secundárias

Para identificar os padrões de estruturas secundárias assumidos pela protease ao longo da DM, usamos o programa DSSP (“Database of Secondary Structure in Proteins” – Kabsch e Sander (1983)) implementado no GROMACS. Com esse programa podemos identificar, através de diferentes cores, quais as conformações (por resíduo) assumidas pela proteína no tempo.

Na Figura 4.14, mostramos o diagrama que define os padrões de estrutura secundária assumidos pela protease ao longo da simulação. Nota-se que a proteína permanece bem estável durante toda a dinâmica, mantendo suas regiões de  $\alpha$ -hélice, folhas- $\beta$  e loops muito bem definidas. Entretanto, algumas dessas regiões mostraram comportamentos distintos entre as diferentes PRs. Uma dessas diferenças foi observada no modelo BD30N. Tal modelo teve sua conformação mais desestabilizada que as demais PRs na região de  $\alpha$ -hélice. Esse dado evidencia que a desestabilização na conformação helicoidal pode ter participação no processo de interação entre protease e ligante, ou seja, quanto maior a variação nessa região menor será a afinidade entre a PR e o ligante (neste caso, o nelfinavir). É importante lembrar que embora essa região não interaja diretamente com o nelfinavir é justamente dentro dela que ocorre o polimorfismo L89M (encontrado entre os subtipos B e C) e isso pode estar associado à diferença de comportamento observada. Nota-se que a mutação D30N desestabiliza a região de  $\alpha$ -hélice no subtipo B, mas não no C

Outra observação importante é a variação da região em torno do resíduo 5. Tal região permaneceu boa parte do tempo nos modelos BD30N e CD30N/N83T na conformação de loop enquanto nos demais não.

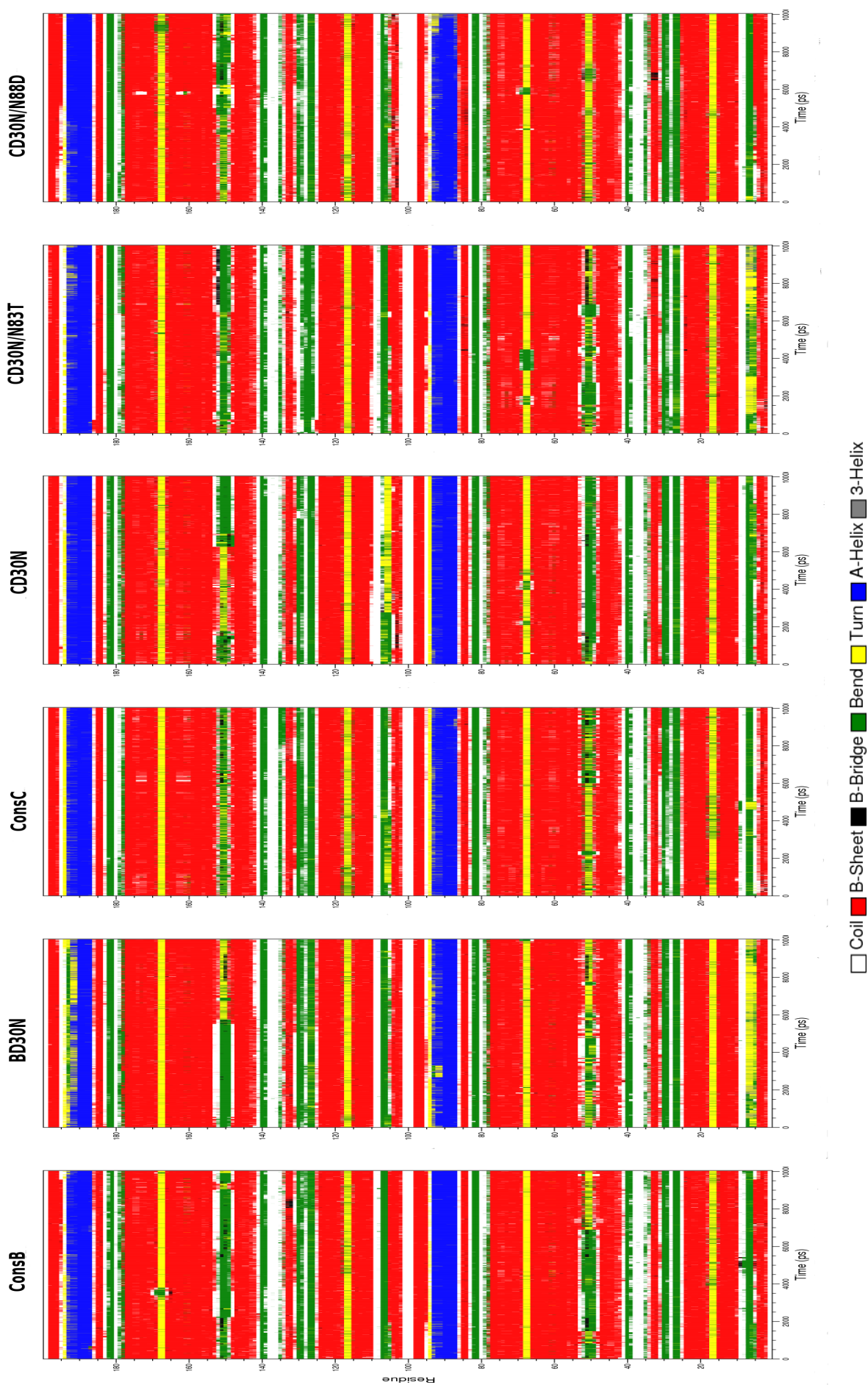


Figura 4.14: Padrões de estrutura secundária por resíduo, ao longo da dinâmica. Em branco, estão representados os resíduos em conformação aleatória; em verde, resíduos em curvas; em amarelo, resíduos em voltas; em azul, resíduos em hélices; em vermelho, resíduos em folhas- $\beta$



# Capítulo 5

## Resultados do Estudo de Casos – Subtipos B e C Complexados ao Substrato

### 5.1 Interação entre a PR e o Substrato

#### 5.1.1 Propriedades Globais

As seis PR's complexadas ao substrato foram estudadas através de simulações de dinâmica molecular. Alguns parâmetros foram analisados a fim de se estabelecer relações entre subtipos B e C que levem a entender as diferenças, na interação com o substrato, causadas pelos polimorfismos naturais em cada subtipo.

Para avaliar a estabilidade das simulações precisamos atentar para o comportamento dinâmico dos sistemas. Como pode ser visto na Fig. 5.1, o desvio quadrático médio (RMSD), ao longo das dinâmicas, dos átomos do esqueleto peptídico, com relação às estruturas iniciais (minimizadas) do consB, consC, CD30N, CDD30N/N88D, tende a um mesmo valor (0,3 nm) ao fim dos 8 ns. Entretanto, os mutantes BD30N e CD30N/N83T após 6 ns apresentaram um aumento no desvio, com relação aos demais, com o BD30N chegando em um dado momento a 0,45 nm.

Assim como no capítulo anterior, foram analisados os últimos 8 ns de cada simulação. A média dos RMSDs variou de 2,5 Å a 3,1 Å. O modelo consC tem o menor desvio médio (2,5 Å) e o modelo CD30N/N83T tem o maior desvio médio

(3,1 Å).

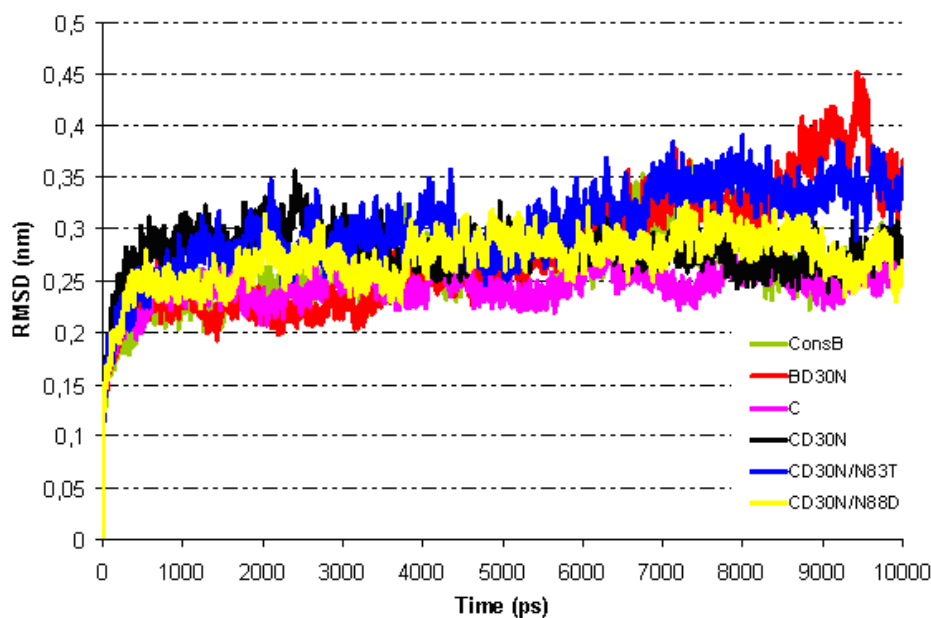


Figura 5.1: Desvio Médio Quadrático (RMSD) da PR complexada ao substrato, em relação à estrutura inicial (pós – minimização de energia), no tempo (10 ns de simulação): consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo).

Tabela 5.1: Valores Médios do RMSD (nm) do esqueleto da protease complexadas ao substrato

<b>B</b>	<b>BD30N</b>	<b>C</b>	<b>CD30N</b>	<b>CD30N/N83T</b>	<b>CD30N/N88D</b>
0.27 ± 0.02	0.29 ± 0.05	0.25 ± 0.01	0.28 ± 0.02	0.31 ± 0.03	0.27 ± 0.02

### 5.1.2 Flutuações dos Resíduos

Pelas Figs. 5.2, 5.3 e 5.4, tem-se uma visão das regiões que mais flutuaram ao longo da dinâmica molecular. Analisando tais contribuições temos que as regiões

onde há maior variação para a PR BD30N é próxima a glicina 16 e às isoleucinas 50 e 50' (Fig. 5.3 A). Entretanto quando analisamos as diferenças entre as variações dos resíduos do consB e do mutante BD30N (Fig. 5.3 B) nota-se que as principais variações ocorreram próximo as isoleucinas 50 e 50'. Porém tal variação foi menor que seu próprio desvio quando comparado à PR complexada ao nelfinavir – acima de 0,35 nm – (ver Fig. 4.3).

Quando comparamos os resultados do consB e os demais subtipos, percebemos uma grande diferença nos resíduos 50 e 50'. Contudo as diferenças apresentam-se menores no consC – aproximadamente 0.1 nm – (Fig. 5.2 B) e maiores no mutante CD30N/N83T – aproximadamente 0.15 nm – (Fig. 5.4 B). Além disso, um dado interessante ocorreu na comparação entre o consB e consC. Nesse caso pôde-se perceber que as principais diferenças entre consC e consB complexado ao substrato (Fig. 5.2 B) apresentaram comportamentos opostos aos observados quando essas mesmas proteases estão complexadas ao nelfinavir (Fig. 4.2 B).

Outras regiões que tiveram uma diferença significativa em relação ao consB foram regiões próximas ao resíduo 178 das proteases BD30N – aproximadamente 0.1 nm – (Figs. 5.3 B) e CD30N/N83T – aproximadamente 0.15 nm – (Figs. 5.4 B). Essas regiões estão localizadas próximas a uma das  $\alpha$ -hélice da PR. O resíduo 16 da protease CD30N/N88D também apresentou uma alta flutuação – aproximadamente 0.35 nm – (Figs. 5.4 A). Entretanto, a diferença em relação ao consB (Figs. 5.4 B) foi menor do que as citadas anteriormente – menos de 0.1 nm.

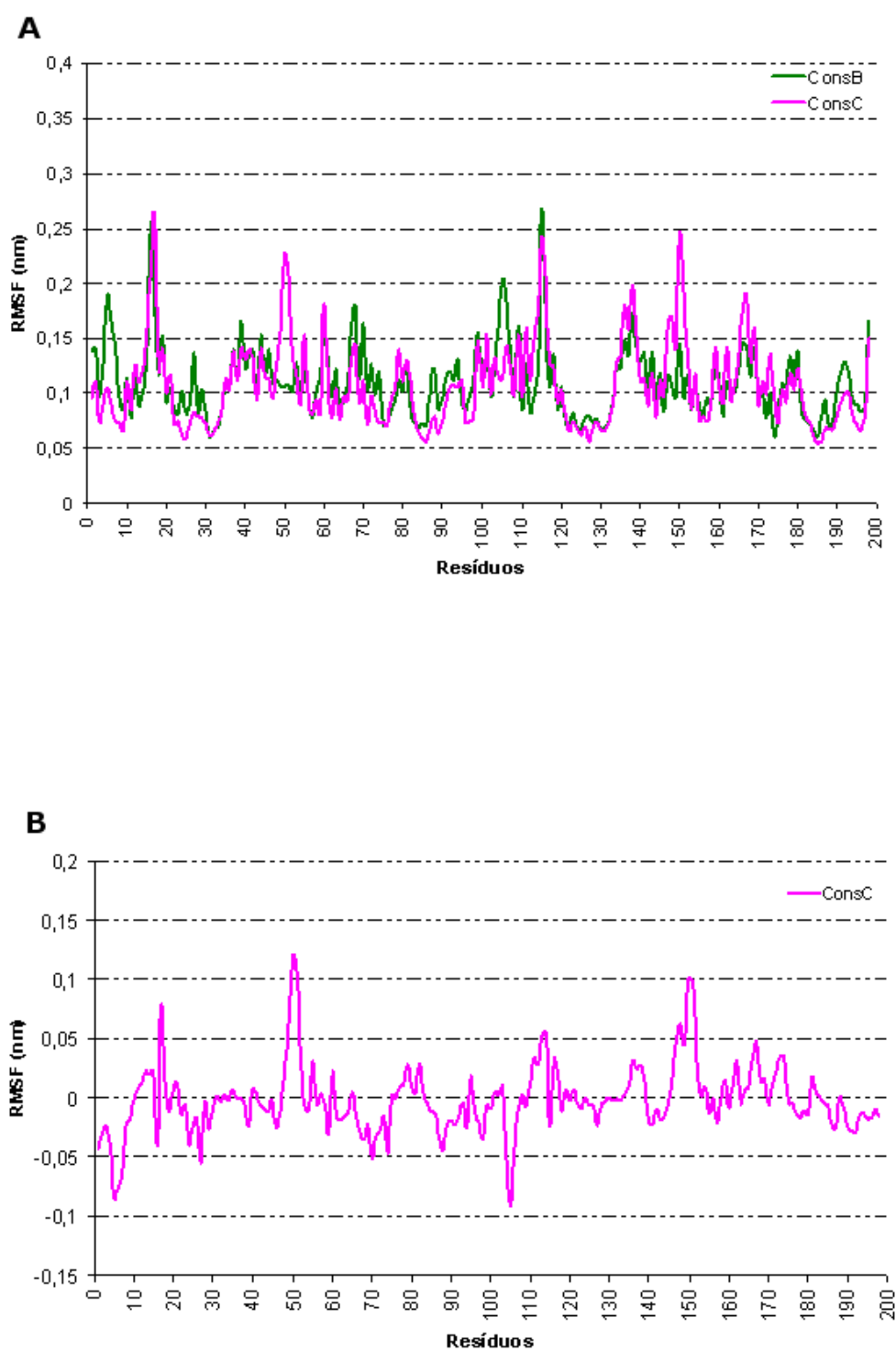


Figura 5.2: Em A, RMSF do ConsB (verde) e ConsC (rosa). Em B, diferenças entre o RMSF do subtipo C e B (ConsC-ConsB).

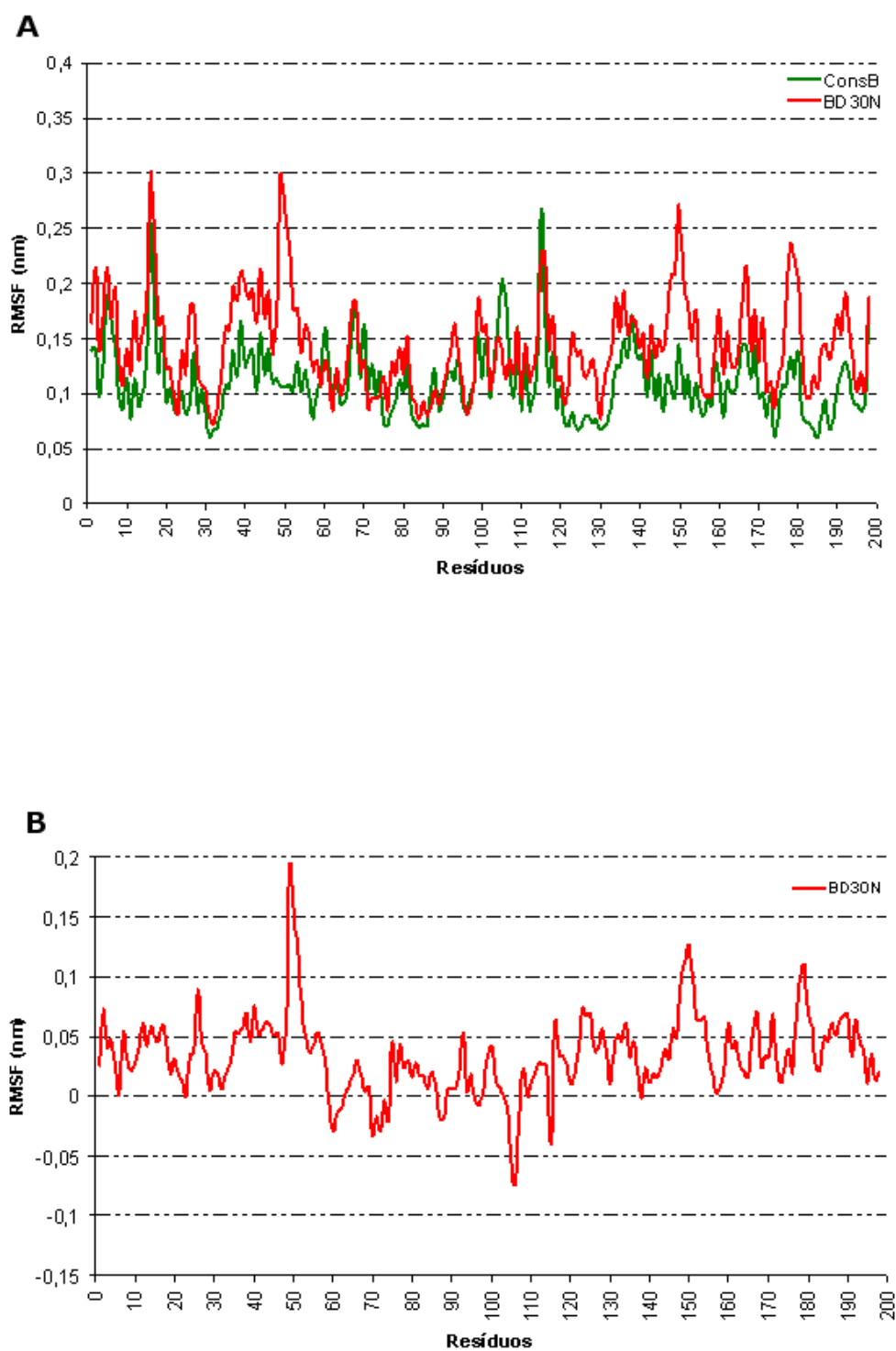


Figura 5.3: Em A, RMSF do ConsB (verde) e BD30N (vermelho). Em B, diferenças entre o RMSF do mutante BD30N e ConsB (BD30N-ConsB).

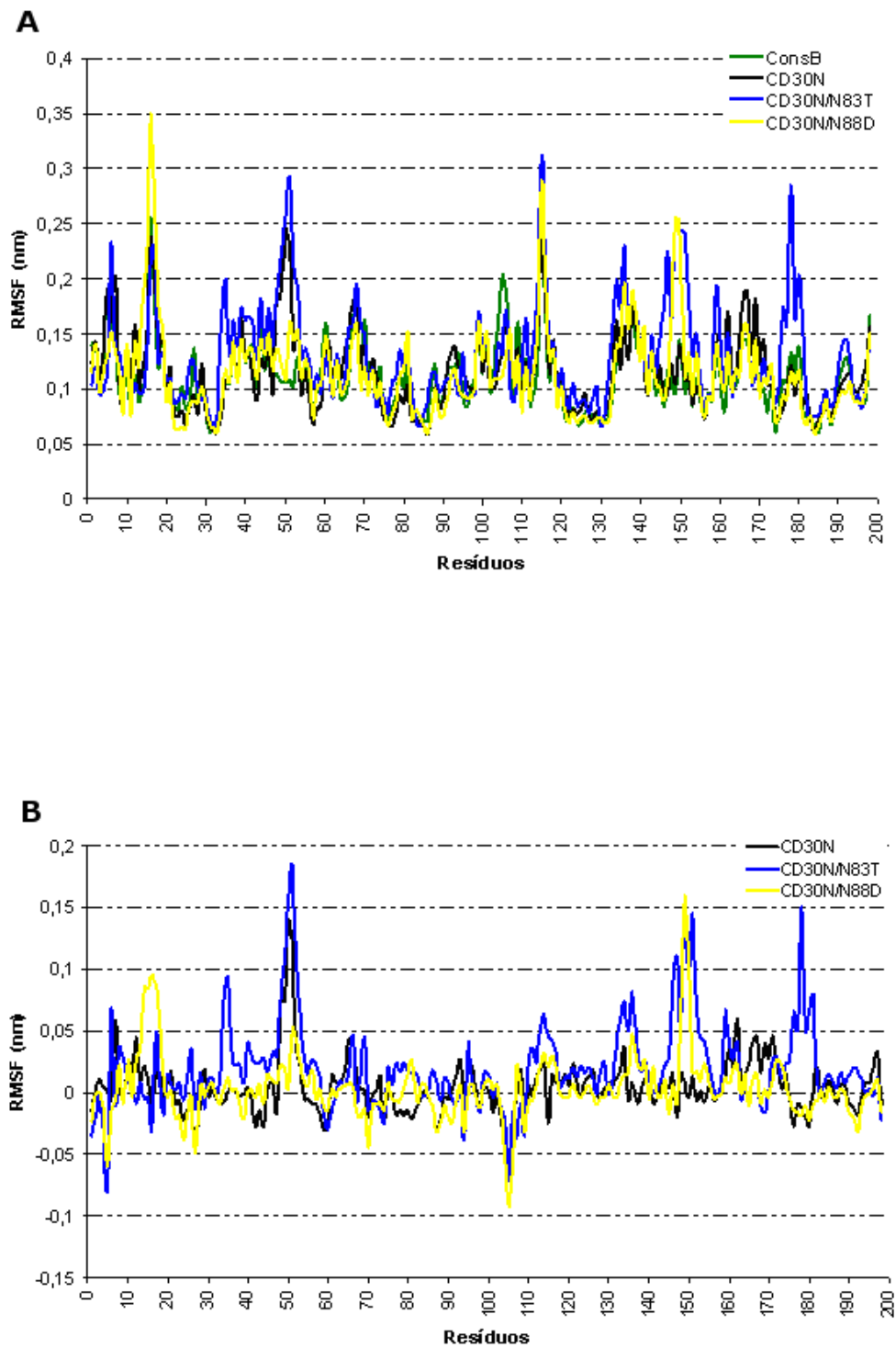


Figura 5.4: Em A, RMSF de quatro sistemas: ConsB (verde), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o RMSF do subtipo B e os demais. CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D-consB (amarelo)

### 5.1.3 Representação 3D do RMSD para os átomos do esqueleto

Assim como na seção 4.1.3, com o auxílio do programa MolMol foram criadas imagens espaciais do RMSD dos átomos do esqueleto peptídico da PR, para cada sistema. Na Figura 5.5, pode-se constatar que as diferenças observadas nas alças das PR's são bem menores do que aquelas observadas na Figura 4.5 (PR's complexadas com o nelfinavir). Particularmente as alças da BD30N complexada com o substrato permanecem mais estáveis do que quando esta proteína mutante é complexada com o inibidor nelfinavir.

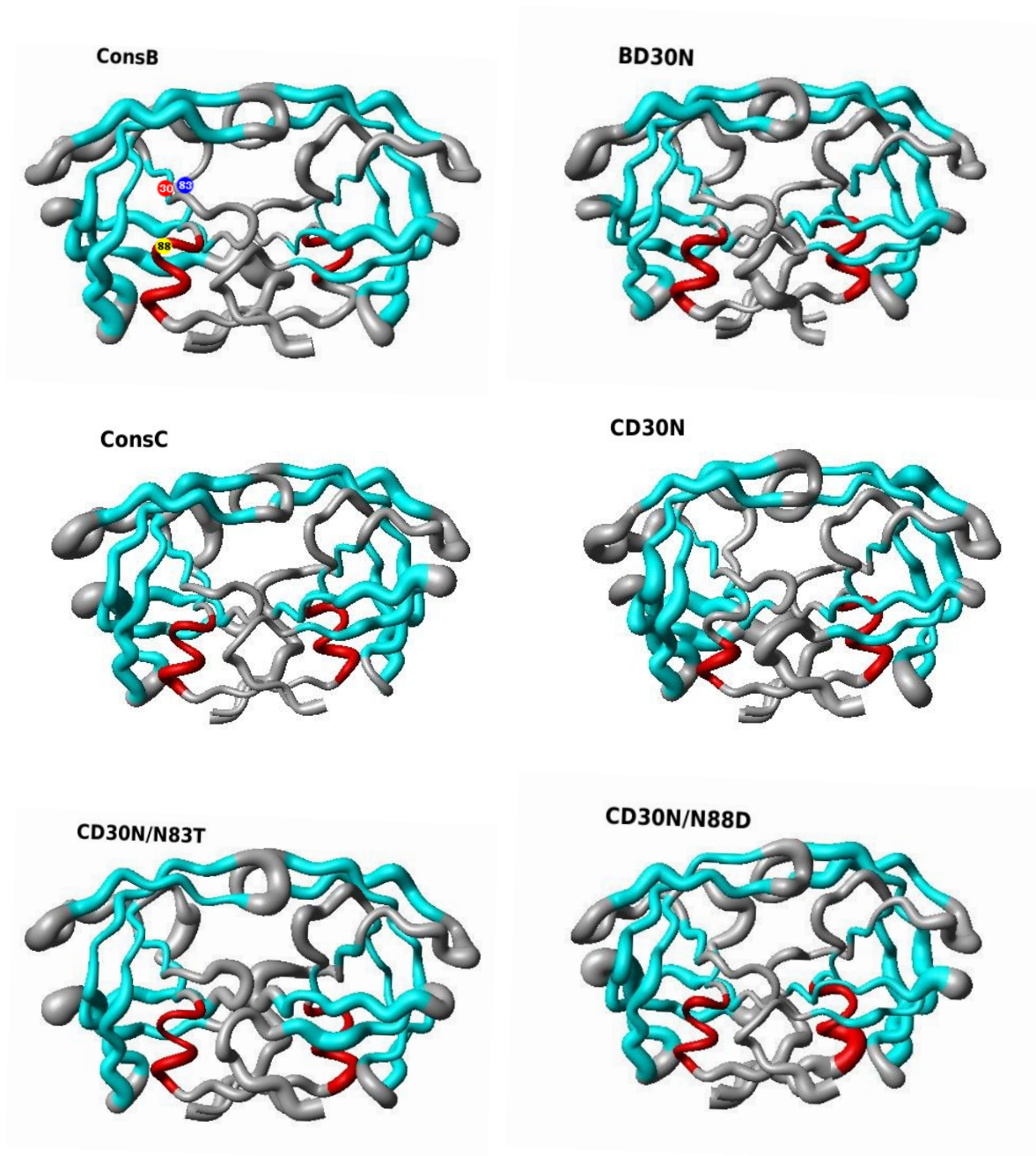


Figura 5.5: Representação 3D das flutuações do RMSD do esqueleto peptídico da PR. O RMSF do backbone foi representado por um tubo com a espessura variada, onde quanto mais espesso o tubo, maior o desvio. Foram coloridos por estrutura secundária: em vermelho,  $\alpha$ -hélice; em azul, folhas- $\beta$  e em cinza, loops e estruturas randômicas. As esferas representam as mutações: em vermelho, D30N; em azul, N83T e em amarelo, N88D.



#### 5.1.4 Variações nas alças da Protease

Para analisar a variação das alças, e saber se houve abertura das mesmas, foi calculado a distância entre suas pontas (resíduos ILE50 e ILE149) e o centro catalítico (Fig. 5.6).

Esta análise representa como se comportaram as alças durante a DM. Nota-se claramente que a alça da cadeia A não sofreu nenhuma variação significativa para todos os modelos. Já na cadeia B ocorreu um rápido afastamento da alça no mutante CD30N/N88D em 3 *ns*, contudo a partir daí permanecendo estável até o fim da DM. O mesmo ocorreu para o mutante BD30N em 6 *ns*. Além disso, note que embora o CD30N/N83T não tenha apresentado as maiores distâncias ele foi o único que apresentou uma tendência ao processo de abertura de suas alças. Todos os outros mantiveram suas pontas bem estáveis enquanto ele lentamente vai distânciando suas alças do centro catalítico.

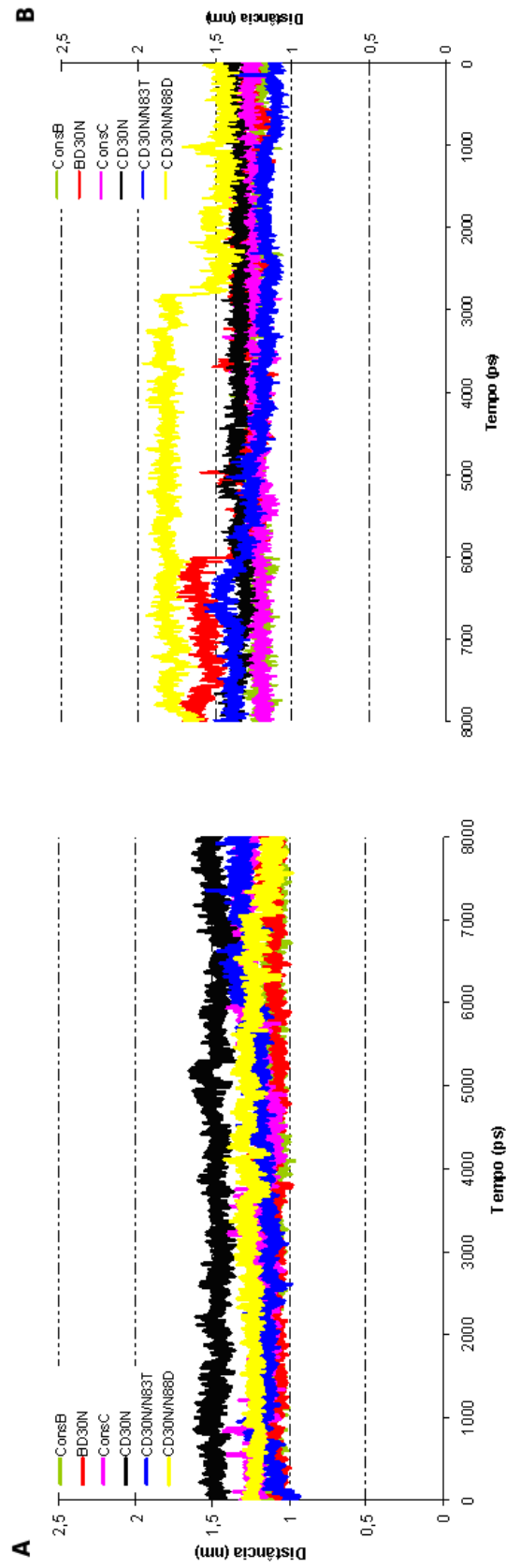


Figura 5.6: Distância entre as pontas das alças e o centro catalítico calculada para cada sistema. Em A, distância referente a alça da cadeia A (ILE 50). Em B distância referente a cadeia B (ILE 149). Em verde ConsB, em vermelho BD30N, em rosa ConsC, em preto CD30N, em azul CD30N/N83T e em amarelo CD30N/N88D. A figura foi invertida visando facilitar a visualização do gráfico.

### 5.1.5 Variações Estruturais na Compactação da PR Complexada ao Substrato

Foram calculados os raios de giro para cada sistema para acompanhar as alterações sofridas pela proteína.

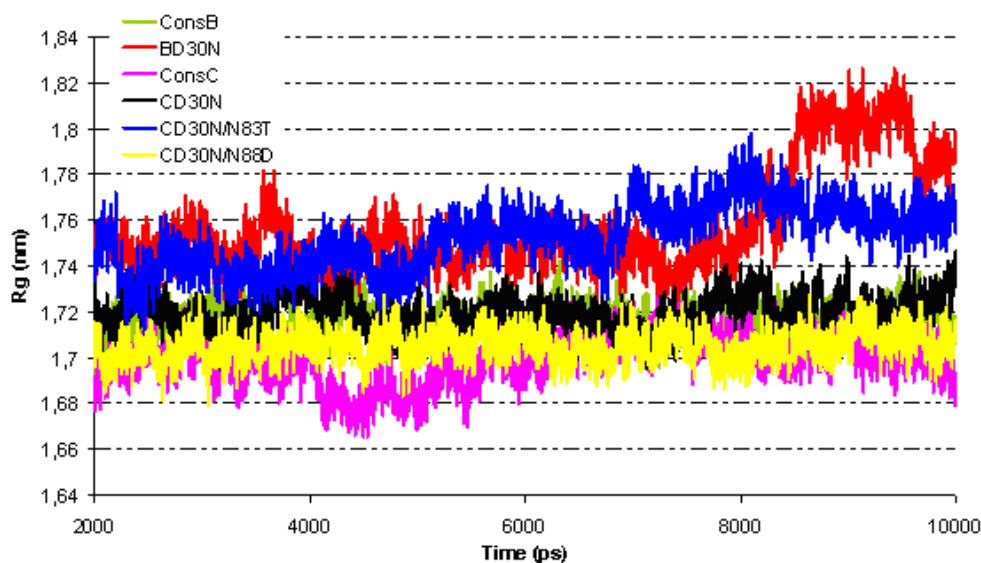


Figura 5.7: Representação do Raio de Giro de todos os sistemas durante a DM. ConsB (verde), BD30N (vermelho), ConsC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D (amarelo). Em B, diferenças entre o subtipo B e os demais. ConsB-BD30N (vermelho), ConsC-ConsB (rosa), CD30N-ConsB (preto), CD30N/N83T-ConsB (azul), CD30N/N88D (amarelo)

No gráfico da Fig. 5.7 os valores para os modelos BD30N e CD30N/N83T diferem um pouco dos valores dos demais modelos. Todos os modelos se mantiveram bem estáveis, entretanto após 8 ns os mutantes BD30N e CD30N/N83T sofreram uma alteração maior.

Essa mudança pode ser acompanhada, tanto com os gráficos de raios de giro quanto com as análises das variações espaciais (RMSD, RMSD 3D e RMSF) das

posições dos átomos durante a simulação.

A tabela 5.2 contém o comportamento médio dos momentos de inércia durante a MD e os valores de  $\alpha$ .

Tabela 5.2: **Valores médios dos momentos de inércia dos modelos complexado ao substrato durante a DM**

<b>Modelos</b>				
	$I_1$	$I_2$	$I_3$	$\alpha$
<b>ConsB</b>	55885.0 $\pm$ 622.3	46239.8 $\pm$ 719.9	24921.6 $\pm$ 411.9	0.31963
<b>BD30N</b>	59416.1 $\pm$ 1848.3	49921.0 $\pm$ 1618.5	24437.8 $\pm$ 339.6	0.33244
<b>ConsC</b>	53926.8 $\pm$ 882.4	44859.1 $\pm$ 862.1	25499.7 $\pm$ 353.8	0.30168
<b>CD30N</b>	56190.2 $\pm$ 683.9	44467.1 $\pm$ 713.0	26624.2 $\pm$ 409.8	0.32439
<b>CD30N/N83T</b>	58428.7 $\pm$ 1203.9	49504.8 $\pm$ 1053.4	24081.0 $\pm$ 446.1	0.32778
<b>CD30N/N88D</b>	55016.6 $\pm$ 608.5	44606.3 $\pm$ 684.7	25492.4 $\pm$ 428.1	0.31918

De acordo com a tabela as PR's mais esféricas são as consC, CD30N/N88D e consB respectivamente, e o valor de  $\alpha$  nos leva a inferir que a mutação D30N, tanto isoladamente ou com as mutações compensatórias N83T e N88D, diminui a esfericidade das proteases. De um modo geral, essas variações estruturais globais ocorrem durante toda a DM e devem-se em parte à flexibilidade da protease e sua tendência a ajustar-se à forma do substrato ou outro ligante. O maior "empacotamento" de uma protease pode sugerir uma maior estabilidade desta com o substrato.

### 5.1.6 Comportamento do Substrato durante a DM

Para entender as interações do substrato com a PR foram analisadas a estabilidade do substrato no sítio ativo da protease. Enquanto a figura 5.8 mostra cada grupo do substrato de clivagem (P4 ,P3, P2, P1, P1', P2',P3', P4') a Fig. 5.9 da uma idéia da estabilidade de cada um desses grupos durante a DM.

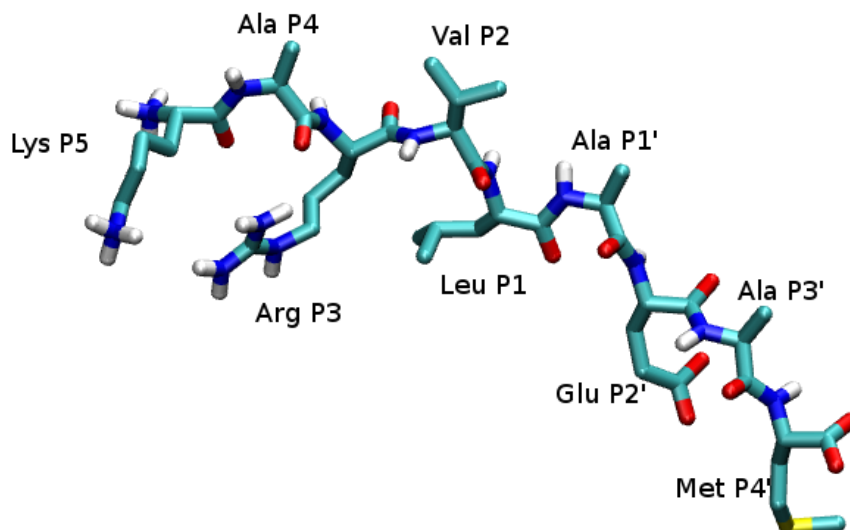


Figura 5.8: Estrutura 3D do substrato CA/p2 com seus grupos de interação P5, P4, P3, P2, P1, P1', P2', P3' e P4'.

O RMSF, por resíduo, (Fig. 5.10) do substrato do mutante BD30N mostrou que essa mutação aumenta a flexibilidade do esqueleto peptídico do ligante entre os resíduos 202 (P2) e 204 (P1'); e diminui a mobilidade nos outros. Tal fato foi observado pelo maior RMSF do substrato do BD30N em relação ao do consB. No subtipo C, o resultado varia: o substrato do CD30N tem a mobilidade diminuída em quase todos os resíduos em relação ao do consC; já o substrato do mutante CD30N/N88D tem as menores flutuações nos resíduos 203 (P1) e 206 (P3'); por último o substrato do mutante CD30N/N83T tem sua variação aumentada em todos os resíduos exceto no 203 e 204 (P1 e P1').

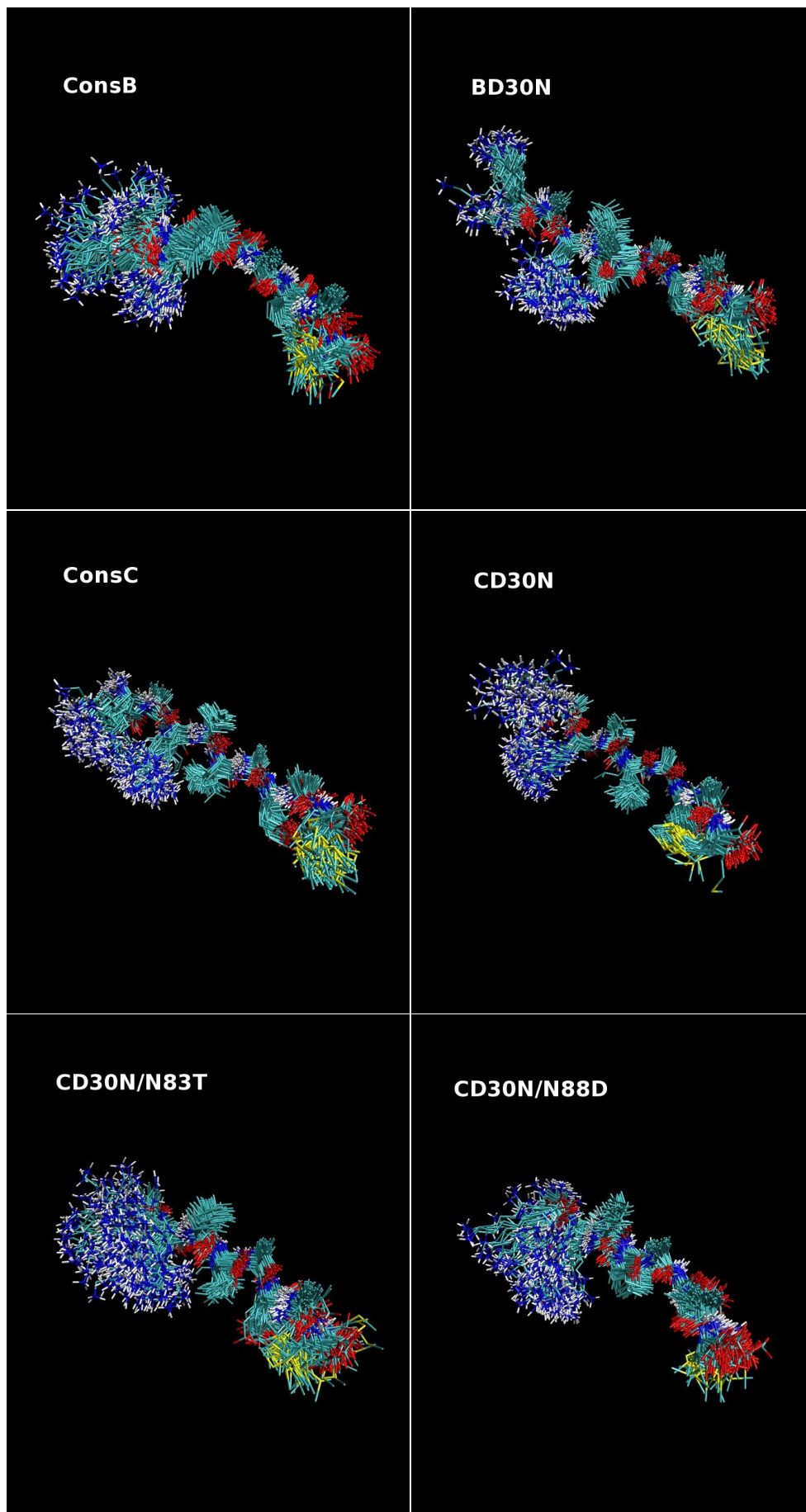


Figura 5.9: Sobreposições das estruturas do substrato durante a DM. Representações coloridas pelo tipo de átomo: branco (H), verde (C), azul (N), vermelho (O) e amarelo (S).

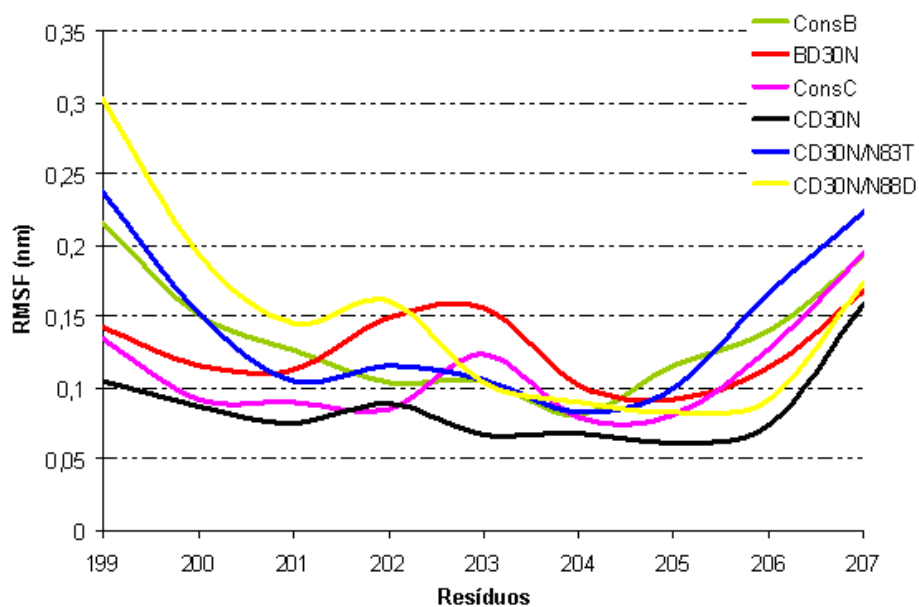


Figura 5.10: Flutuações do RMSF calculadas para os átomos do substrato em cada sistema: consB (verde), BD30N (vermelho), consC (rosa), CD30N (preto), CD30N/N83T (azul) e CD30N/N88D(amarelo). Os grupos (P5, P4, P3, P2, P1, P1', P2', P3' e P4') do substrato estão indicados, tanto no gráfico quanto na representação.

Verifica-se que no mutante BD30N as regiões mais flexíveis, em relação ao controle, são poucas e se restringem ao centro, e nos mutantes CD30N/N83T e CD30N/N88D elas são quase todas e estão espalhadas pelas diversas partes do substrato.



### 5.1.7 Trajetórias das Dinâmicas

As trajetórias de todos os sistemas complexados ao substrato estão demonstradas na figura 5.11. Verifica-se que logo nos processos de otimização e equilíbrio ocorreu uma transformação estrutural do modelo CD30N/N88D que se manteve ao longo da dinâmica. A partir de 4 *ns* tem início um leve afastamento dos “flaps” da estrutura CD30N/N83T. Já o mutante BD30N permaneceu estável durante quase toda a dinâmica somente iniciando o processo de abertura de suas alças nos últimos nanosegundos, indicando que a mutação D30N induz a uma menor afinidade também ao substrato, entretanto possivelmente melhor que as dos demais mutantes, já que estes apresentam um afastamento das alças já aos 4 *ns* de simulação. Como era esperado, os modelos B e C permaneceram estáveis durante toda a DM.

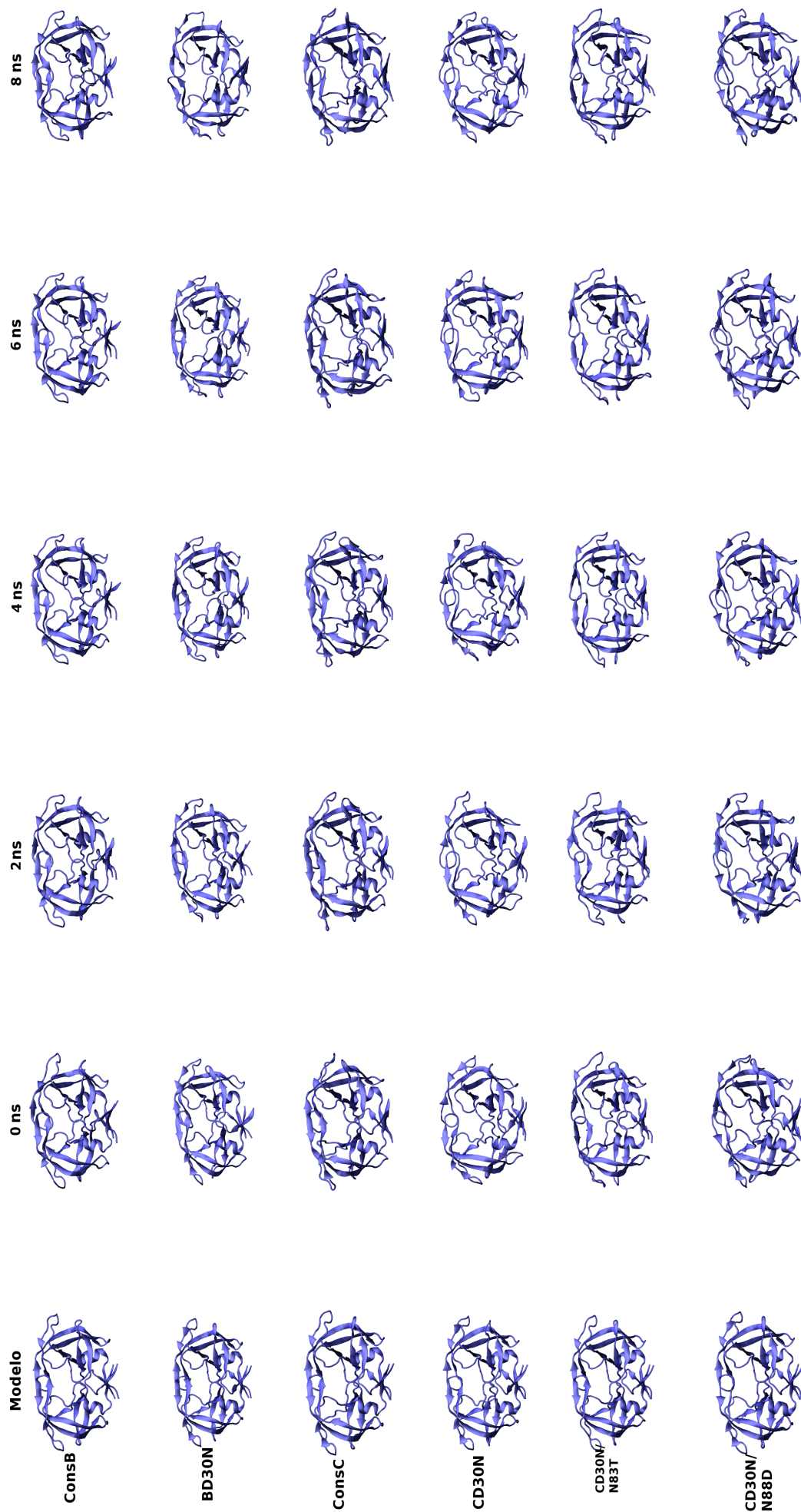


Figura 5.11: “Snapshots” da estrutura das PRs, representando somente o esqueleto peptídico, demonstrando a trajetória das PRs durante os 10 ns de simulação.

### 5.1.8 Ligações de hidrogênio entre a PR e o substrato

Verificou-se as ligações de hidrogênio entre as diversas PRs e o substrato. As análises foram feitas considerando as ligações de hidrogênio diretas, entre proteína e substrato (D), e as intermediadas por moléculas de água (WM).

Tabela 5.3: **Número médio de Ligações de Hidrogênio entre a PR e o substrato.**

Modelos						
	ConsB	BD30N	ConsC	CD30N	CD30N/N83T	CD30N/N88D
<b>D</b>	11.48	13.63	5.04	9.29	10.05	9.98
<b>WM</b>	2.20	2.21	3.10	2.81	2.01	2.79
<b>Total</b>	13.68	15.84	8.14	12.10	12.06	12.77

Número médio de ligações de Hidrogênio entre a PR e o substrato durante a simulação. As ligações foram separadas em: diretas (D), intermediadas por água (WM) e total.

A Tabela 5.3 mostra uma ligeira diminuição no número médio das ligações de hidrogênio indiretas dos mutantes do subtipo C quando comparados com o consC. Já quando comparamos os mutantes do subtipo C com o mutante BD30N percebe-se que todos aqueles possuem menor média total que este. Contudo dois resultados se mostraram mais evidentes. O primeiro foi a diferença encontrada entre os subtipos B e C. Nesse caso pode-se perceber que há uma redução de cerca de 6 ligações de hidrogênio diretas no modelo C. O segundo foi que todos os mutantes aumentaram significativamente o número médio de Pontes de Hidrogênio diretas com relação ao observado entre o substrato e os consensos. Fato não observado na interação com o nelfinavir.

As Tabelas 5.4 e 5.5 mostram os resultados obtidos para cada um dos sistema. Nelas estão discriminados os resíduos do substrato e da PR que estão interagindo.

Na Tabela 5.4, nota-se uma ligação, em P4, muito bem conservada, nas proteases consenso, com o resíduo 48 do flap. Esta ligação apresentou uma redução considerável (na sua forma direta) no BD30N e CD30N/N88D com relação às PRs selvagens e esteve totalmente ausente nos mutantes CD30N e CD30N/N83T.

Outra redução importante foi observada entre o resíduo 48 e os subsítios P3 e P2. Além disso, a ligação entre a GLY52 e P3 também foi reduzida em todos os sistemas quando comparados ao consB. Na cadeia B (Tabela 5.5), um comportamento similar ocorreu. Nota-se que a ligação do resíduo 149 (50') da alça com P1' é significativamente reduzida nos mutantes BD30N, CD30N, CD30N/N83T e CD30N/N88D, em relação ao seu respectivo consenso, e a ligação entre ARG186 e P2' foi basicamente observada no consB.

Com intuito de analisarmos se as ligações de hidrogênio citadas acima são mesmo de grande relevância para a interação protease/substrato verificamos se as mesmas também se encontravam na estrutura cristalográfica 1F7A (Fig. 5.12). Pode-se perceber nessa figura que a interação existente em P4 entre o Gly48 (GLY48N) e a Ala 200 (ALA200O) já se encontrava presente na estrutura cristalográfica com uma distância de 3.21 Å. O mesmo ocorreu com a ligação existente em P3 entre os resíduos Gly48 (GLY48O) e Arg201 (ARG201NH1), nesse caso a distância encontrada foi de 2.97 Å. Além disso, notou-se que hora o oxigênio da Gly48 faz uma ligação com o nitrogênio terminal da Arg201, hora com o nitrogênio terminal da Val 202 (ligação não encontrada na estrutura cristalográfica).

Um outro fato que chamou atenção quando comparamos ligações presentes no cristal com as que ocorreram durante a DM foi o que ocorreu em P2', entre o nitrogênio do resíduo 129 (ASP129N ou ASN129) e os oxigênios do resíduo 205 (GLU205OE1 ou GLU205OE1), nesse caso as ligações estiveram presentes de forma direta, cerca de 50% do tempo no consB, enquanto no consC tal ligação ocorreu, no mesmo período tempo, porém de forma indireta (Tab 5.5).

Tabela 5.4: Ligações de hidrogênio (%) entre a cadeia A da PR e o substrato.

Átomos		ConsB		BD30N		ConsC		CD30N		CD30N/N83T		CD30N/N88D	
Substílios	Substrato - Receptor	D	WM	D	WM	D	WM	D	WM	D	WM	D	WM
P4	ALA2000 -GLY48N	82.8	6	6.5	43	92.4	*	*	*	*	*	44.3	2.9
P4	ALA200N -ASP29OD1	0.1	0.2	57	0.6	*	*	*	*	0.1	0.1	*	*
P4	ALA200N -ASP29OD2	*	0.2	71.2	1	*	*	*	*	*	*	*	*
P4	ALA200N -MET46O	*	*	*	*	61.8	11.9	*	*	*	*	*	*
P3	ARG201O -ASP29N	5	26.6	6.7	2.4	*	1.4	80.8	1.1	*	*	12.5	0.1
P3	ARG201N -ASP29OD1	2.3	1.1	44.7	7.1	*	*	40.1	5.2	*	*	1.5	0.8
P3	ARG201N -ASP29OD2	2.6	0.4	1.6	11.1	*	*	31.9	3.5	*	*	1.8	0.1
P3	ARG201NH1 -GLY48O	32.2	0.3	4.3	0.3	6.2	1.6	*	*	3.1	1.2	4.6	0.2
P3	ARG201NH1 -GLY52O	42.3	0.5	*	*	*	1.1	*	*	0.6	0.2	0.1	*
P2	VAL202O -GLY49N	*	*	0.1	1.2	1.3	0.3	1.9	57.7	*	*	*	*
P2	VAL202O -ILE50N	*	*	63.4	6.9	*	*	*	68.3	79.8	1.2	3.9	0.2
P2	VAL202N -GLY48O	96.1	*	76.3	*	70.6	*	*	*	45.3	7.2	66.8	1.2
P1	LEU203O -ILE50N	*	*	*	*	*	*	*	*	*	*	82.3	11.9
P1	LEU203N -ASP25OD1	94.6	*	46.5	*	*	*	*	*	54.2	10.5	*	*
P1	LEU203N -ASP25OD2	93.1	*	53.6	*	*	*	*	*	35.9	6.2	*	*
P1'	ALA204N -ASP25OD1	*	*	78.3	*	*	*	*	*	20.2	4.9	0.3	5.7
P1'	ALA204N -ASP25OD2	*	*	64.7	0.2	*	*	*	*	54.1	9.8	*	18.1

Percentagem da Permanência das ligações de hidrogênio entre a PR e o substrato durante os 8 *ns* de simulação de DM. Somente ligações de hidrogênio que permaneceram mais de 25% do tempo testado foram representadas. Quando a ligação de hidrogênio de um dos modelos apresentava valor maior que 25% também foi mostrado a ligação dos demais modelos, mesmo se estes tivessem menos de 25%.

\* – Os asteriscos denotam ausência de ligações detectadas.

Tabela 5.5: Ligações de hidrogênio (%) entre a cadeia B da PR e o substrato.

Substítios	Átomos													
	Substrato - Receptor		ConsB		BD30N		ConsC		CD30N		CD30N/N83T		CD30N/N88D	
	D	WM	D	WM	D	WM	D	WM	D	WM	D	WM	D	WM
P1		*	*	*	*	*	*	*	92.1	1.3	92.3	*	*	*
P1'		*	68.4	*	*	15.9	17.6	5.3	95.9	*	2.4	*	*	*
P2'		27.6	*	85.4	*	73	2.6	95.9	*	*	93.2	*	85.2	2.4
P2'		18.9	*	72.2	*	22	1	95.1	*	*	73.5	*	74.4	1.2
P2'		*	*	64	1.3	*	*	*	*	*	*	*	*	*
P2'		*	*	39.6	2.9	*	*	*	*	*	*	*	*	*
P2'		41.6	*	33.4	0.1	0.5	42.8	64.4	*	*	96.7	*	81.5	*
P2'		34.3	*	60.4	0.1	0.1	7.8	32.2	*	*	*	*	15.9	*
P2'		*	*	42.1	0.7	*	*	65.4	0.9	0.9	91.3	0.4	83	0.3
P2'		40.4	5.8	77.7	0.6	*	*	32.5	0.6	0.6	54.7	*	15.9	*
P2'		61.3	0.1	*	*	*	*	*	*	*	*	*	*	*
P2'		68.5	1.2	0.3	1.1	*	*	*	*	*	6.7	20.8	*	*
P2'		93.7	*	6.2	0.2	*	*	*	*	*	35	20.3	*	*
P3'		*	*	42	0.3	*	46.7	74.8	2	2	*	3	97	0
P3'		33	*	49	*	57.1	*	98.1	*	*	41.1	0.2	55.4	*
P4'		*	*	40.2	1.6	*	*	*	*	*	*	*	*	*

Percentagem da Permanência das ligações de hidrogênio entre a PR e o substrato durante os 8 *ns* de simulação de DM. Somente ligações de hidrogênio que permaneceram mais de 25% do tempo testado foram representadas. Quando a ligação de hidrogênio de um dos modelos apresentava valor maior que 25% também foi mostrado a ligação dos demais modelos, mesmo se estes tivessem menos de 25%.

\* – Os asteriscos denotam ausência de ligações detectadas

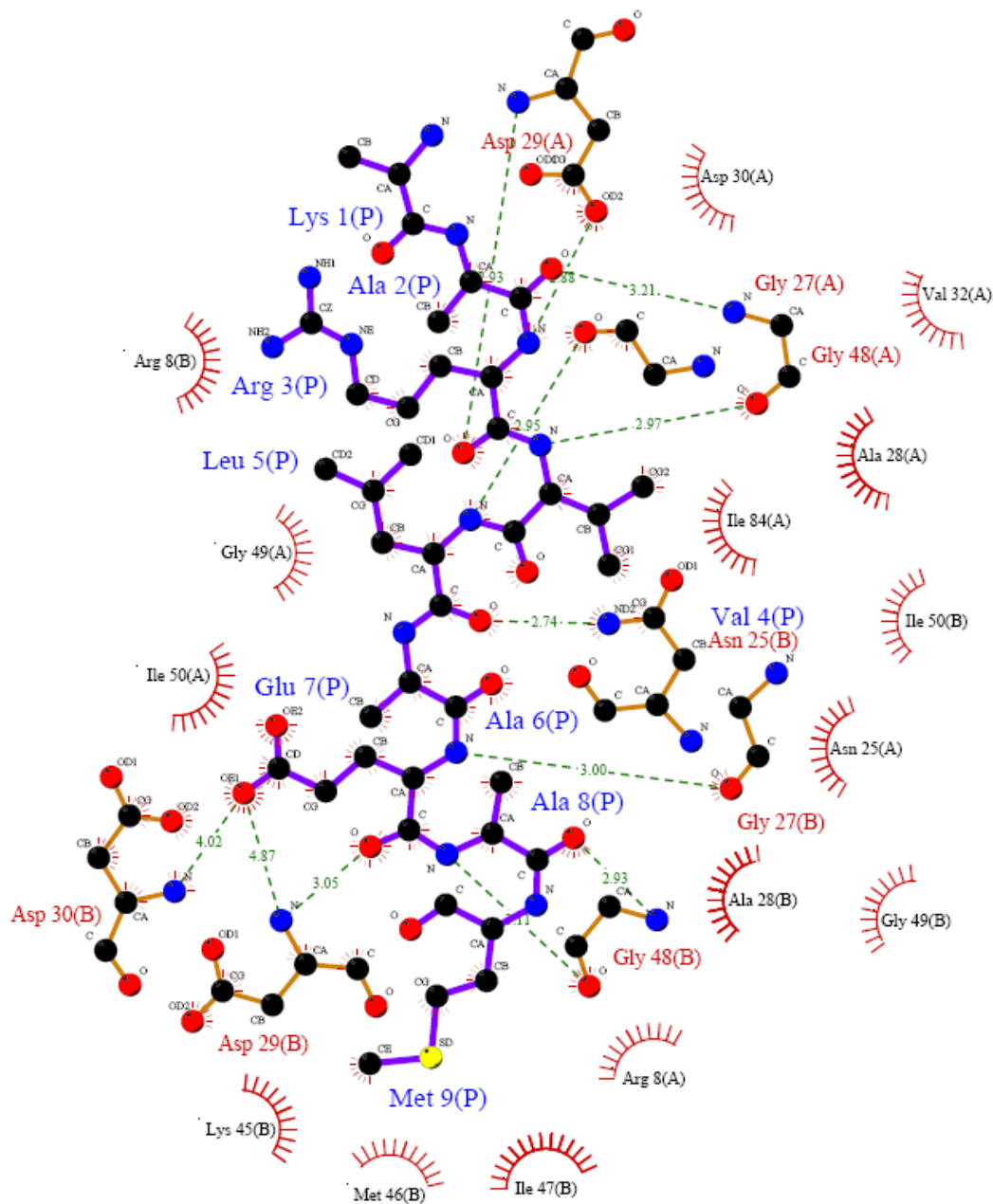


Figura 5.12: Esquema em 2D do sítio ativo da estrutura cristalográfica 1F7A. Estão representados os átomos do inibidor, exceto os hidrogênios. Em verde estão representadas as ligações de hidrogênio, com seus respectivos comprimentos. Cada aminoácido que está presente na superfície intemolecular de interação da PR com o nelfinavir está representado em vermelho.

\* – Os resíduos seguidos por (A) representam os aminoácidos de 1 a 99 e os seguidos por (B) representam os de 100 a 198.

### 5.1.9 Contatos Moleculares entre a Proteína e o Ligante

Para melhor entender a ligação do substrato às PR's dos diferentes subtipos, foram calculadas as áreas de interação entre a PR e o substrato, sendo cada uma discriminada por resíduo de aminoácido presente na interface PR/substrato.

Pode-se observar quatro regiões bem definidas em torno dos resíduos 8, 25, 50 e 82, na cadeia A e quatro regiões em torno dos resíduos 107, 124, 149, 181, na cadeia B (Fig. 5.13), definindo um padrão de interação destas proteases com o substrato. Essas regiões estão centradas nos resíduos Arg8/107 (entradas do sítio ativo), Asp25/124 (centro catalítico), Ile50/149 (flaps) e Val82/181 (laterais do sítio ativo). Nota-se que o padrão de interação é semelhante entre os subtipos, sendo praticamente os mesmos resíduos participantes da interação PR/substrato, com algumas pequenas exceções.

A mutação D30N diminui significativamente a área de contato entre o substrato e o mutante CD30N/N83T no resíduo 30. Além disso, os mutantes do subtipo C têm menor área média de interação com o substrato nos resíduos 23, 46, 53, 178 e 179. Já no mutante BD30N essa diminuição se dá apenas em torno da região 181, onde todos os resíduos próximos a essa região (180, 181, 183 e 186) mantiveram menos contato com o substrato em relação ao consB.

A assimetria observada entre as cadeias A e B era esperada, visto que o substrato é assimétrico quanto a seus grupos laterais.



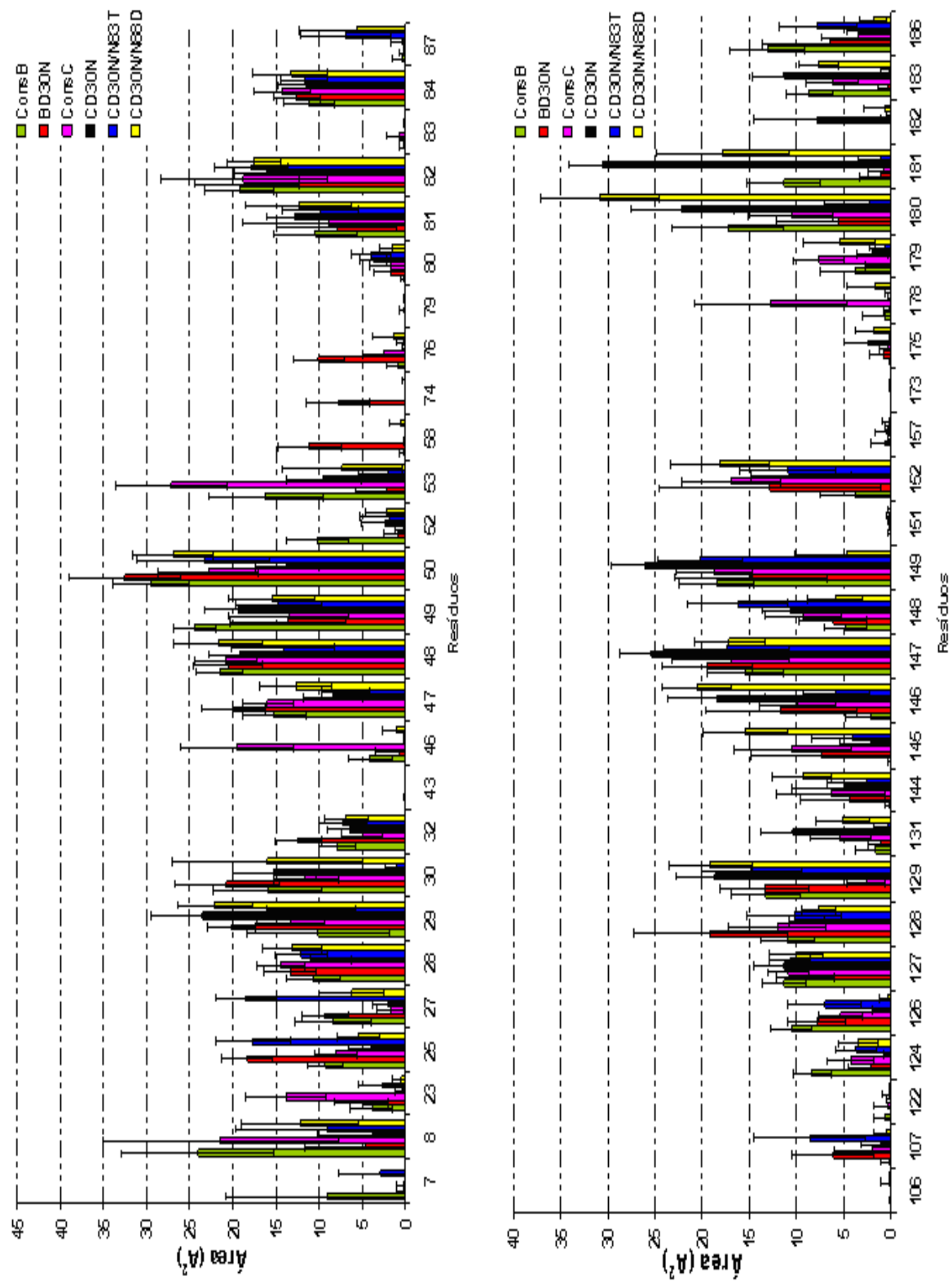


Figura 5.13: Área da Superfície de Contato Intermolecular entre a PR e o substrato. Em barra estão representados os valores médio e em linha o desvio-padrão. Somente os contatos com mais de 5 Å<sup>2</sup> de área foram representados.

No resultado obtido para o cálculo de área total e hidrofóbica (Tabela 5.6), os diferentes modelos obedeceram ao mesmo critério de redução observados quando as PRs interagiram com o nelfinavir. Entretanto, as PRs apresentaram quase o dobro de contato com o substrato. Novamente, o mutante CD30N/N88D apresentou uma área de contato maior que o consC. Podemos ver que o modelo CD30N/N83T foi o que apresentou menor área, o que implica numa pior interação com o substrato, o que poderia explicar o porque do seu baixo poder replicativo *in vitro*. É importante lembrar que esse também foi o modelo que obteve a menor área hidrofóbica (Tabela 5.7).

Além disso, podemos observar que em geral os maiores valores de área foram encontrados nos subsítios S2, S1, S1' e S2', ressaltando a importância desses subsítios. Contudo se observarmos o RMSF, por resíduo, do substrato (Fig 5.10) iremos perceber que as regiões do substrato, complexado à PR BD30N, onde ocorreram as maiores variações em relação ao do consB foram as regiões P2, P1 e P1'. Tais regiões interagem com os subsítios S2, S1 e S1' da PR e pode-se notar que foram nessas regiões onde ocorrem as principais reduções na área de contato desse mutante. Já o RMSF do substrato do CD30N/N83T mostrou que as principais variações em relação ao do consC ocorrem nas extremidades (P5, P4, P3, P2, P2', P3', P4') e nota-se que em quase todas essas regiões ocorrem reduções significativas na interação intermolecular (S5, S4, S2, S2', S3' e S4') desse mutante. Por outro lado embora o substrato do sistema CD30N/N88D tenha apresentado as maiores diferenças no RMSF, em relação ao consC, nos resíduos P5, P4, P3 e P2 sua redução mais significativa na interação intermolecular ocorreu no subsítio S4. Na análise do sistema CD30N notou-se que as interações com as regiões P do substrato (S5, S4, S3, S2, S1) tiveram um aumento em sua área de contato enquanto as interações com as regiões P' (S1', S2', S3' e S4') obtiveram um decréscimo.

Tabela 5.6: Área de Contato ( $\text{Å}^2$ ) entre a PR e o substrato

Área de Contato Intermolecular entre a PR e o substrato. Entre parênteses podem ser visualizados os aumentos (+) ou decréscimos (-) de cada um dos subsítio dos mutantes em relação ao seu respectivo tipo selvagem.

Subsítio	ConsB	BD30N	ConsC	CD30N	CD30N/N83T	CD30N/N88D
<b>S5</b>	54.5	114.9 (+60.4)	52.8	66.4 (+13.8)	32.1 (-20.7)	81.3 (+28.5)
<b>S4</b>	87.9	119.4 (+31.5)	91.8	113.6 (+21.8)	54.7 (-37.1)	73.8 (-18)
<b>S3</b>	112.1	100.4 (-11.7)	77.9	128.6 (+50.7)	89.6 (+11.7)	120.1 (+42.2)
<b>S2</b>	173.9	141.3 (-32.6)	146.6	157.2 (+10.6)	134.9 (-11.7)	142.6 (-4)
<b>S1</b>	170.8	111.4 (-59.4)	128.0	157.6 (+29.6)	141.9 (+13.9)	154.7 (+26.7)
<b>S1'</b>	166.9	128.6 (-38.3)	131.6	126.3 (-5.3)	139.4 (+7.8)	129.7 (-1.9)
<b>S2'</b>	149.5	141.3 (-8.2)	139.7	102.8 (-36.9)	138.3 (-1.4)	132.2 (-7.5)
<b>S3'</b>	123.7	105.9 (-17.8)	110.9	87.5 (-23.4)	67.3 (-43.6)	121.6 (+10.7)
<b>S4'</b>	49.7	53.3 (+3.6)	74.2	56.8 (-17.4)	40.9 (-33.3)	73.9 (-0.3)
<b>Total</b>	1089	1016.7 (-72.3)	953.5	996.8 (+43.3)	839.1 (-114.4)	1029.9 (+76.4)

**S5** – Gly27, Asp29, Asp30, Lys45, Lys 45, Met46, Ile47, Gly49, Gly51, Phe53, Gln58, Thr74, Leu76, Arg87, Arg107, Pro180, Val181  
**S4** – Gly27, Ala28, Asp29, Asp30, Val32, Lys45, Met46, Ile47, Gly48, Gly49, Ile50, Phe53, Leu76, Arg87, Ile149, Pro180, Val181  
**S3** – Asp25, Gly27, Ala28, Asp29, Asp30, Gly48, Gly49, Ile50, Gly52, Phe53, Arg107, Pro178, Pro180, Val181, Asn182  
**S2** – Asp25, Gly27, Ala28, Asp29, Asp30, Val32, Ile47, Gly48, Gly49, Ile50, Ile54, Val82, Ile84, Ile149, Pro180, Val181  
**S1** – Asp25, Gly27, Ala28, Asp29, Asp32, Gly48, Gly49, Ile50, Ile54, Pro81, Ile84, Arg107, Asp124, Gly126, Ala127, Val131, Gly147, Gly148, Ile149, Thr179, Pro180, Val181, Ile183  
**S1'** – Leu23, Asp25, Ala28, Gly49, Ile50, Ile54, Thr80, Pro81, Val82, Ile84, Asp124, Gly126, Ala127, Ile146, Gly147, Gly148, Ile149  
**S2'** – Lys7, Arg8, Leu23, Asp25, Ile50, Pro81, Val82, Gly126, Ala127, Asp128, Asp129, Val131, Ile146, Gly147, Gly148, Ile149, Arg186  
**S3'** – Lys7, Arg8, Leu23, Ile50, Pro81, Val82, Gly126, Asp128, Asn129, Met145, Ile146, Gly147, Gly148, Ile149, Phe152, Arg186  
**S4'** – Lys7, Arg8, Pro81, Val82, Asp128, Asp129, Val131, Lys144, Met145, Ile146, Gly147, Phe152

Tabela 5.7: Área de Contato ( $\text{Å}^2$ ) Hidrofóbica entre a PR e o substrato

Área de Contato Hidrofóbica entre a PR e o substrato. Entre parênteses podem ser visualizados os aumentos (+) ou decréscimos (-) de cada um dos subsítio dos mutantes em relação ao seu respectivo tipo selvagem.

Subsítio	ConsB	BD30N	ConsC	CD30N	CD30N/N83T	CD30N/N88D
<b>S5</b>	27.8	42.8 (+15)	48.2	25.9 (-22.3)	3.7 (-44.5)	33.4 (-14.8)
<b>S4</b>	57.1	65.4 (+8.3)	64.8	55.2 (-9.6)	8.2 (-56.6)	36.9 (-27.9)
<b>S3</b>	43.8	28.7 (-15, 1)	32.7	58.9 (+26.2)	26 (-6.7)	58.3 (+25.6)
<b>S2</b>	98.6	96.0 (-2.6)	85.4	117.1 (+31.7)	92.8 (+7.4)	82.7 (-2.7)
<b>S1</b>	89.9	58.5 (-31.4)	68.9	121.8 (+53.6)	89.6 (+20.7)	109.5 (+40.6)
<b>S1'</b>	110.2	90.2 (-20)	94.2	89.6 (-4.6)	98.7 (+4.5)	99.3 (+5.1)
<b>S2'</b>	46.6	56.4 (+9.8)	67.9	57.0 (-10.9)	54.9 (-13)	72.0 (+4.1)
<b>S3'</b>	53.9	63.2 (+9.3)	72.5	50.6 (-21.9)	36.0 (-36.5)	77.7 (+5.2)
<b>S4'</b>	18.1	26.3 (+8.2)	45.4	22.9 (-22.5)	15.8 (-29.6)	44.9 (-0.5)
<b>Total</b>	546	527.7 (-18.3)	580	599 (+19)	425.7 (-154.3)	614.7 (+34.7)

**Resíduos hidrofóbicos – Ala, Leu, Val, Ile, Pro, Phe, Met e Trp**

### 5.1.10 Padrões das estruturas secundárias

Na Fig. 5.14 podemos identificar, através de diferentes cores, quais as conformações (por resíduo) assumidas pela PR no tempo. Assim como na interação com o inibidor, pode-se notar que os modelos mantiveram-se bem estáveis durante a DM.

Como já descrito anteriormente, a instabilidade na região de  $\alpha$ -hélice pode ter participação no processo de interação entre protease e ligante. Pode-se observar que quando as PRs interagiram com o substrato, o modelo CD30N foi o que teve as maiores variações na região de  $\alpha$ -hélice (Fig. 5.14). Nota-se que a mutação D30N desestabiliza a região de  $\alpha$ -hélice somente quando o subtipo C apresenta-se sem as mutações compensatórias N83T e N88D. Isso implica que tais mutações recuperam a estabilidade dessa região. É importante lembrar que o mutante CD30N não é encontrado *in vivo* nem *in vitro*.

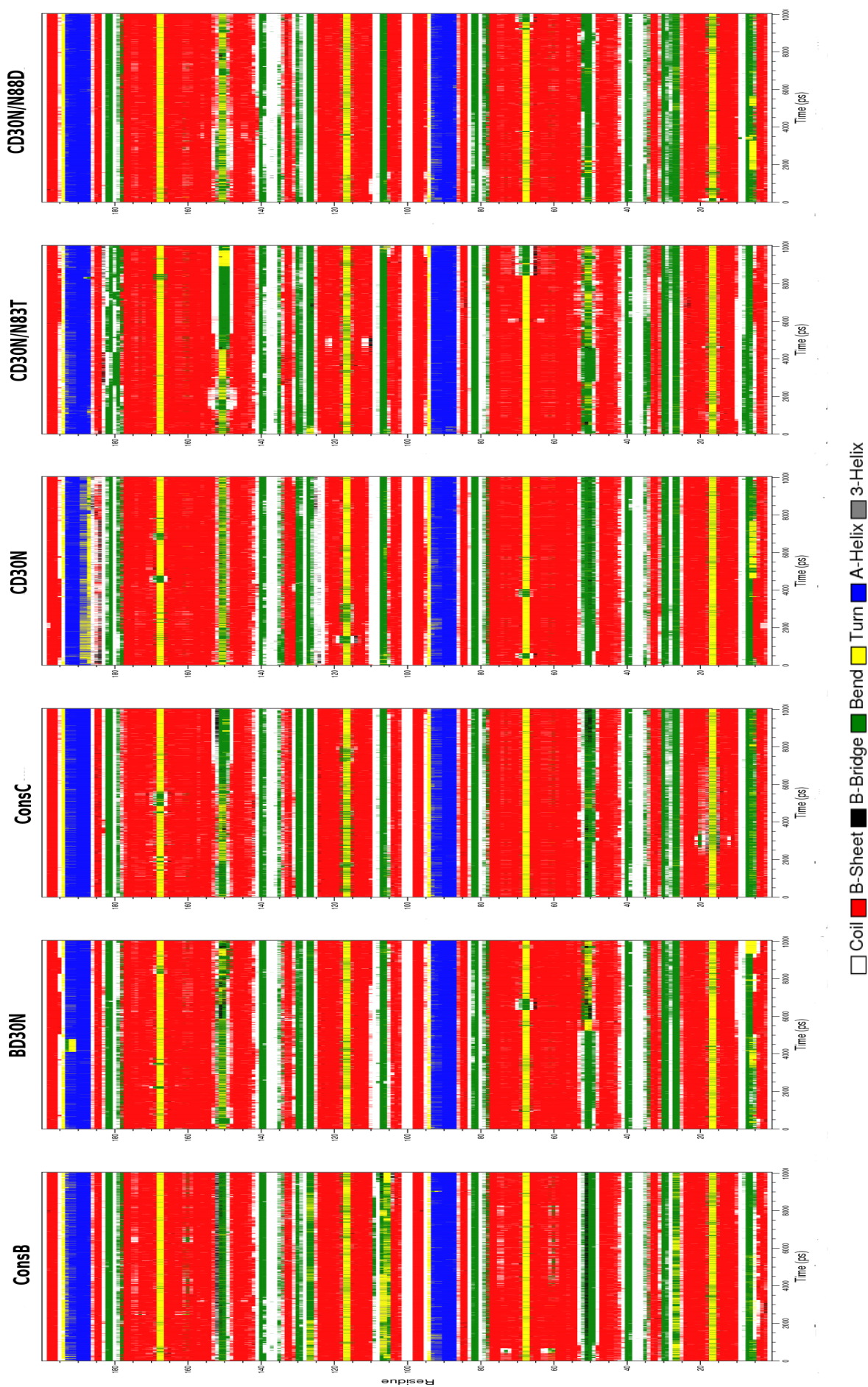


Figura 5.14: Padrões de estrutura secundária por resíduo, ao longo da dinâmica. Em branco estão representados os resíduos em conformação aleatória; em verde, resíduos em curvas; em amarelo, resíduos em voltas; em azul, resíduos em hélices; em vermelho, resíduos em folhas- $\beta$ .

# Capítulo 6

## Discussão

A Modelagem e a Dinâmica Molecular constituem uma ferramenta muito poderosa para o estudo de sistemas biomoleculares em escala atômica, tanto do ponto de vista estrutural quanto do ponto de vista funcional, permitindo a análise dos átomos de forma dinâmica.

Este trabalho permitiu conhecer um pouco mais acerca da interação protease-inibidor e protease-substrato em HIV-1 dos subtipos B e C. Em estudos por dinâmica molecular em solvente explícito, monitorou-se o desvio das posições atômicas do esqueleto peptídico com relação às posições iniciais. Verificou-se uma alta variação do RMSD logo nas primeiras centenas de picosegundos, o que pode ser atribuído à mudanças de um ambiente cristalino (experimental) para um ambiente líquido (simulação). A estrutura cristalográfica é obtida em estado sólido e em baixas temperaturas, enquanto na DM não há empacotamento de estruturas e a temperatura é ambiente, no nosso caso (300 K). Com isto, justificamos a grande variação inicial do RMSD (Figs. 4.1 e 5.1) observada para todas as nossas simulações até os primeiros 2000 ps. Após esse tempo, as estruturas tendem a estabilizar e passam a flutuar em torno de conformações médias. Por isso optou-se por analisar os últimos 8000 ps de cada sistema.

Variações estruturais após essa fase inicial podem estar associadas a acomodações que os modelos sofrem devido a interação com o ligante. Devido às diferenças entre as seqüências das PRs dos subtipos B e C, e às mutações sofridas

por elas, os arranjos conformacionais podem ocorrer de forma bem distinta.

As altas variações do RMSD para o modelo BD30N e CD30N/N83T (estas PRs são as que possuem os maiores IC<sub>50</sub> de acordo com os ensaios fenotípicos (Tabela 1.4) obtidos por Gonzalez et al. (2004)), observadas na figura 4.1, estão associadas à resistência ao nelfinavir, tal afirmação pode ser reforçada pelas análises que focaram a abertura das alças. Nesses resultados foram constatados que esses modelos foram os que mais sofreram abertura em seus “flaps” contribuindo com a idéia de que as altas variações observadas foram devido ao movimento das suas alças. Entretanto, encontrar variações próximas para os mutantes CD30N e o CD30N/N88D (Tabela 4.1) não era esperado. Visto que o CD30N não é encontrado *in vitro* esperavamos uma variação elevada indicando uma possível desnaturação dessa protease. Isso pode evidenciar que a inviabilidade do modelo CD30N está antes da ligação com o ligante, deste modo, a simulação do complexo PR/ligante para este modelo pode não resultar em resultados robustos. Nas variações observadas quando as PRs estão complexadas ao substrato, pode-se perceber que o modelo CD30N/N83T foi o que apresentou maior RMSD, o que pode ser associado a um ineficiente processo de acomodação corroborando com as características fenotípicas dos clones B e C (Tabela 1.4) e com os ensaios de competição (Tabela 1.5), obtidos por Gonzalez et al. (2004). Naquele trabalho, foi observado que o mutante CD30N/N83T era o de menor capacidade replicativa, e também o que perdia as competições em cultura de células contra outros mutantes ou o tipo selvagem do subtipo C.

Enquanto o RMSD dá informações sobre o desvio global de uma estrutura em função do tempo e com relação a estrutura inicial, o RMSF dá informações sobre a flutuação local por átomo ou por grupo de átomos em função da sua posição inicial, permitindo distinguir regiões flexíveis e mais estáveis.

Na presença do inibidor, verificou-se que as altas variações observadas no RMSD do mutante BD30N são devido, principalmente, à variação sofrida pelos resíduos 50 e 50' (alças) (Fig 4.3 A). Este fato está associado ao processo de



abertura dos flaps. O modelo CD30N/N83T, embora tenha sofrido altas variações, com relação a estrutura inicial, nas regiões das alças, obteve variações menores quando comparadas com o BD30N. Por outro lado, o CD30N/N83T apresentou um RMSD mais elevado que os demais modelos no sítio ativo (Fig 4.4 B). Tanto a variação na região das alças como a na região catalítica podem estar associadas ao processo de resistência. Enquanto uma alta variação nos flaps pode permitir que o ligante fique instável no sítio ativo, uma pequena variação na região catalítica já é suficiente para prejudicar o processo de acomodação.

Quando foram analisados o RMSF dos resíduos dos subtipos complexados ao substrato, notou-se uma redução na variação dos flaps do BD30N e uma menor flutuação na região catalítica do CD30N/N83T comparado às variações sofridas na presença do nelfinavir. Tais dados corroboram a idéia de que PRs resistentes à inibidores possuem um mecanismo discriminatório do inibidor em relação ao substrato natural, ou seja, se ligam menos eficiente ao primeiro, mas conseguem reconhecer relativamente bem o segundo. Além disso, como era esperado, nenhuma PR sofreu abertura das alças.

Ainda com o objetivo de analisar as variações observadas nas PRs, foram calculados os raios de giro de cada modelo. Através dessa análise foi possível saber se houve compactação ou expansão na forma da proteína. Observou-se que o BD30N quando ligado ao nelfinavir foi o que mais sofreu variação em sua forma (Fig. 4.7). Isso corrobora a idéia de que a alta variação observada nas alças está diretamente associada ao processo de abertura das mesmas. Quando observadas essas variações da PR interagindo com o substrato, todos os modelos permaneceram sem grandes modificações na sua forma. Tal fato era esperado visto que não houve abertura das alças das PRs quando complexadas ao substrato. Somente uma pequena variação foi observada no mutante BD30N ao final da DM.

Para se ter uma idéia visual da abertura dos flaps foram plotadas figuras das trajetórias das simulações a cada 2 ns de todas as PRs (Fig. 4.11 e 5.11). Quando as PRs se encontram interagindo com o nelfinavir (Fig. 4.11), pode-se

visualizar que as alças do BD30N começam a se abrir em 4 *ns*. A partir de 6 *ns* já é possível ver suas alças abertas, e também um leve afastamento das alças do CD30N/N83T. Embora o CD30N tenha afastado levemente suas pontas logo nos primeiros nanosegundos, no resto da DM esse modelo manteve-se estável, não dando continuidade ao processo de abertura de suas pontas. Quando observadas essas trajetórias para as PRs interagindo com o substrato (Fig. 5.11), nota-se que os modelos permaneceram sem grandes modificações e nenhum deles apresentou abertura das alças.

Ao analisar as ligações de hidrogênio, observamos que os subtipos B e C têm um número médio de ligações próximo quando estão complexado ao inibidor (Tabela. 4.3). Entretanto, apresentam números bem distintos quando na presença do substrato. Na interação com o substrato o subtipo C mostrou um menor número de ligações que o B (aproximadamente 6 ligações). Estes dados podem ajudar a explicar a menor adaptabilidade natural (“Fitness”) do subtipo C comparada aos demais subtipos do grupo **M** relatada por Ariën et al. (2005).

Contudo alguns mutantes apresentaram um número médio de ligações de hidrogênio maior do que os seus respectivos consensos (Tabela 5.3). Isso nos remete à idéia de que o simples aumento ou diminuição na média das pontes de hidrogênio pode ou não significar um melhor processo de interação entre receptor e ligante. Por isso, foram mapeadas as ligações mais persistentes durante a DM e pôde-se perceber que a mutação D30N provoca nos mutantes, quando comparados aos tipos selvagens, uma diminuição significativa na ligação entre O46 do nelfinavir e o nitrogênio do ASN30 ou ASP30 (ASN/ASP30N–NEL202O46) e entre o O21 do inibidor e o oxigênio delta do ASP25 (NEL200O21–ASP25OD2) (Tabela 4.4). Um outro ponto notável é a diminuição na prevalência da ligação de hidrogênio entre o N37 do inibidor e o oxigênio delta do ASP25 (NEL202N37–ASP25OD2) para todos os mutantes. Comparamos essas ligações com as encontradas na estrutura cristalográfica e verificamos que apenas a última não estava presente. Com isso tentamos validar que as principais ligações de hidrogênio, que estão localizadas no

cristal, tem o tempo de permanência reduzido nos mutantes.

Com relação à interação com o substrato, todos os resíduos das PRs mutantes que tiveram diminuição significativa na ligação de hidrogênio estavam localizados nas alças (GLY48, ILE149, MET46, GLY52). Ao compararmos com as ligações encontradas no cristal, nota-se que duas já estão presentes (GLY48N–ALA200O e GLY48O–ARG201NH1) uma, embora não esteja presente (Gly48O–Val202N), está bem indicada visto que ora o oxigênio da Gly48 faz uma ligação com o nitrogênio terminal da Arg201, ora com o nitrogênio terminal da Val 202. Além disso, o comportamento apresentado pelo nitrogênio do ASP129 e os oxigênios do GLU205 apresentaram comportamentos bem distintos entre os consensos B e C. No primeiro a ligação permaneceu na sua forma direta, enquanto no segundo na forma indireta. Todas estas evidências podem estar fortemente relacionadas com a maior instabilidade dos complexos com PRs mutantes.

Com relação a área de superfície de contato intermolecular, nós podemos notar que nossos modelos teóricos são dinamicamente consistentes com o modelo cristalográfico, uma vez que o mesmo perfil da região de contato foi encontrado para todos os sistemas. Um fato bem evidente nas análises dos resultados e consenso na literatura é a interação assimétrica das cadeias da protease. Sabendo-se que nenhum substrato natural é simétrico, espera-se que análises envolvendo sítio ativo sejam distintas entre as cadeias. Ao fazermos esta análise, constatamos que a área de interação dos modelos mutantes é bastante diminuída em seus subsítios quando as PRs estão interagindo com o nelfinavir (Tabelas 4.5), exceto para o CD30N/N88D, e esta diminuição, nos mutantes do subtipo C, deve-se basicamente à perda de contatos hidrofóbicos (Tabelas 4.6). Além disso, cerca de 50% da redução nos contatos do mutante BD30N também são devido à perda de contatos hidrofóbicos. O fato do modelo CD30N/N88D apresentar maior área de contato que o consC nos leva a pensar que a mutação compensatória N88D restitui a área de contato hidrofóbica perdida pelo CD30N.

Seguindo a análise da superfície de interação intermolecular, notou-se que o

mutante BD30N foi o que apresentou a maior redução na sua superfície de contato (Tabelas 4.5). Além disso, um dos subsítios onde houve redução da área de contato ocorreu numa região onde o inibidor manteve-se muito instável (P2) (Fig. 4.10). Tal fato proporciona uma maior liberdade para uma determinada região do nelfinavir cuja instabilidade é grande, logo prejudica o processo de acomodação do ligante.

Quando as proteases estão interagindo com o substrato nota-se que somente há redução significativa no modelo CD30N/N83T (Tab. 5.6), nesse sistema observou-se que essa redução também se deu por redução nos contatos hidrofóbicos. Além disso, o resíduo 30 desse modelo apresentou uma área de contato intermolecular reduzida quando comparado com os outros, indicando que a mutação D30N e N83T juntas no subtipo C alteram significativamente a interação entre o resíduo 30 e o substrato.

Não o bastante é importante ressaltar os resultados encontrados para os mutantes CD30N/N83T e CD30N/N88D. Enquanto o primeiro teve sua área de contato com o ligante (inibidor ou substrato) diminuída em relação ao consC, o segundo teve sua área aumentada. Tal fato está de acordo com os resultados experimentais obtidos por Gonzalez et al. (2004). Nesse estudo foi observado que o mutante CD30N/N83T é mais resistente que o CD30N/N88D. Todavia aquele possui menor adaptabilidade que este.

Geralmente, estruturas secundárias periódicas, como  $\alpha$ -hélices ou folhas- $\beta$ , apresentam maior estabilidade pelo fato das potenciais ligações de hidrogênio do esqueleto proteico apresentarem-se todas satisfeitas. Da mesma forma, segmentos conectando essas estruturas secundárias, como voltas e loops, são geralmente mais flexíveis por apresentarem um menor número de ligações de hidrogênio satisfeitas no esqueleto proteico. De um modo geral esse comportamento foi verificado em nossas simulações e pode ser observado nas Figuras 4.14 e 5.14. Entretanto, uma das regiões de  $\alpha$ -hélice teve sua conformação desestabilizada no modelo BD30N, quando ligado ao nelfinavir, e no CD30N quando a PR está complexada ao substrato. Embora essa região não tenha participação direta na interação protease

ligante a variação, na conformação helicoidal, pode influenciar indiretamente no processo de interação entre PR e ligante, provocando uma interação pouco eficiente. No primeiro caso isso colaboraria com a resistência, enquanto no segundo isso pode representar uma das causas da inviabilidade (observada *in vivo* e *in vitro*) desse mutante. Contudo uma análise mais acurada deve ser feita afim de melhor qualificar a relação existente entre atividade e estabilidade da  $\alpha$ -hélice.

Deve-se ponderar que o substrato utilizado é somente um dos substratos da protease e que eles têm taxa de afinidade e catálise diferentes. Os resultados obtidos não podem ser generalizados para o comportamento de protease porque a afinidade com outro substrato pode ser distinta. Estudos adicionais devem ser conduzidos para melhor avaliar outros sítios de clivagem.

Por outro lado ressaltamos que o substrato natural analisado neste estudo (KARVLAEEAM), que corresponde ao sítio de clivagem CA/p2 da proteína Gag, é totalmente conservado nos subtipos B e C (na verdade ele é quase universalmente invariável no grupo M). Portanto, nossas simulações mimetizam a interação real entre a PR do subtipo C e seu substrato natural.

Do conjunto de resultados discutidos neste capítulo, espera-se ter justificado o objetivo principal que norteou esta dissertação, ou seja, explicar as diferenças que a mutação de resistência D30N provoca nos subtipos B e C.

# Capítulo 7

## Conclusões

Este trabalho traz uma importante contribuição para se compreender o efeito da mutação D30N nos subtipos B e C do vírus HIV-1. Através de simulações mimetizando o ambiente natural de atuação da PR, ficaram claras algumas diferenças entre as proteases dos subtipos B e C. As estruturas dos mutantes BD30N e CD30N/N83T mostraram-se mais instáveis durante toda a dinâmica, ou seja, a acomodação dos inibidores contra esses modelos se mostrou menor.

Foi possível também identificar as regiões de interação das PR's com o inibidor nelfinavir e com o substrato, e através do cálculo da área da superfície de contato intermolecular, pôde-se perceber uma grande diferença entre as áreas de contato de todos os modelos estudados, contribuindo com a idéia de que a resistência esteja associada também a uma redução dos contatos hidrofóbicos. Por outro lado, uma menor afinidade entre receptor e ligante não pode ser considerada analisando-se simplesmente o número médio de ligações de hidrogênio.

Em nosso trabalho, a menor interação do mutante BD30N com o inibidor e do CD30N/N83T pelo substrato foi confirmada e caracterizada ao nível molecular. De uma forma geral, os mutantes CD30N e CD30N/N88D se comportaram de maneira análoga ao consC. A principal diferença encontrada para o CD30N foi uma desestabilização na sua região de  $\alpha$ -hélice quando esta interage com o substrato. Tal fato pode estar associado à ineficiência no processo de clivagem, entretanto por não termos dados fenotípicos desse mutante (ele não pôde ser obtido experimen-

talmente (Gonzalez et al., 2004)), as análises obtidas tomam um caráter de previsibilidade. Já o mutante CD30N/N88D, embora com comportamento semelhante ao consC, apresentou-se com maior RMSD, maior raio de giro e principalmente com suas ligações de hidrogênio entre GLY48N–ALA200O e ALA200N–MET46O menos intensas.

Concluimos que a mutação D30N no subtipo C, quando não acompanhada de outras mutações primárias e/ou secundárias, pode ter um comportamento defectivo no seu processo de clivagem. Levantamos a hipótese que a mutação D30N no subtipo C, interfere no reconhecimento do substrato natural e também tem o potencial de aumentar a resistência ao inibidor nelfinavir.

# Capítulo 8

## Perspectivas

Como perspectivas deste trabalho destacamos:

- Realização de estudos de Cálculo de Energia Livre;
- Proposição de alterações na estrutura do nelfinavir que procurem aumentar a especificidade de protótipos de inibidor para o subtipo C, visando posterior síntese dos candidatos mais promissores;
- Analisar a interação das proteases dos subtipos B e C com outros substratos.



## Referências Bibliográficas

- D. M. V. Aalten, R. Bywater, J. Findlay, M. Hendlich, R. Hooft, e G. Vriend. ProdrG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. **Journal of Computer Aided Molecular Design**, 10(3):255–262, 1996.
- M. Alterman. **Design and Synthesis of HIV-1 Protease Inhibitors**. Tese de Doutorado, ACTA Universitatis Upsaliensis Uppsala / Suécia, 2001.
- K.K. Ariën, A. Abraha, M.E. Quinones-Mateu, L. Kestens, G. Vanham, e E.J.Arts. The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, group O and HIV-2 isolates. **J. Virol**, 79(14):8979–8990, 2005.
- A. K. Atkinson. **An introduction to numerical analysis**, Capítulo 8. 2 edição, 1988.
- T. W. Baba, A. M. Trichel, e et al. Infection and aids in adult macaques after nontraumatic oral exposure to cell-free SIV. **Science**, 272:1486–1489, 1996.
- F. Barresinoussi, J. C. Chermann, e et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune-deficiency syndrome (aids). **Science**, 220(4599):868–871, 1983.
- F. I. Bastos. **AIDS na Terceira Década**. Editora Fiocruz, 1 edição, 2006.
- P. R. Batista. Estudos computacionais da protease do HIV-1: Abertura das alças e diferenças entre subtipos. Dissertação de Mestrado, Instituto de Biofísica Carlos Chagas Filho - Universidade Federal do Rio de Janeiro / Brasil, 2006.

- P. R. Batista, A. Wilter, E. H Durham, e P. G. Pascutti. Molecular dynamics simulations applied to the study of subtype of HIV-1 protease common to Brazil, Africa and Asia. **Cell Biochem Biophys**, 44(3):395–404, 2006.
- H. M. Berman. The protein data bank: a historical perspective. **Acta Cryst**, A64:88–95, 2008.
- M. Bonhomme, S. Wong, e et al. The ph dependence of hiv-1 capsid assembly and its interaction with cyclophilin a. **Biophys Chem**, 105(1):67–77, 2003.
- B. Bowerman, P. O. Brown, e et al. A nucleoprotein complex mediates the integration of retroviral dna. **Genes Dev**, 3(4):469–478, 1989.
- B. R. Brooks, R. E. Bruccoreli, B. D. Olafson, D. J. States, e et al. Charmm: a program for macromolecular energy minimization and dynamics calculation. **Journal of Computational Chemistry**, 4(2):187–217, 1983.
- E. Caride, K. Hertogs, e et al. Genotypic and phenotypic evidence of different drugresistance mutation patterns between b and non-b subtype isolates of human immunodeficiency virus type 1 found in brazilian patients failing haart. **Virus Genes**, 23(2):193–202, 2001.
- Z. Chen, B. Hu, e et al. Hiv-1 mutants less susceptible to sch-d, a novel small-molecule antagonist of ccr5. In **9th CROI**, Seattle, USA, 2002.
- F. Clavel, D. Guetard, e et al. Isolation of a new human retrovirus from west african patients with aids. **Science**, 233(4761):343–346, 1986.
- E. Coakley, C. J. Petropoulos, e J. M. Whitcomb. Assessing chemokine co-receptor usage in hiv. **Current opinion in infectious diseases**, 18:9–15, 2005.
- K. M. De Cock, G. Adjoriolo, E. Ekpini, T. Sibailly, J. Kouadio, M. Maran, e et al. Epidemiology and transmission of hiv-2. why there is no hiv-2 pandemic. **JAMA**, 270(17):2083–2086, 1993.

- J. M. Coffin e et al. New name for aids virus - what to call the aids virus. **Nature**, 321(6065):10–10, 1986.
- J. Cohen. Aids treatment. a step toward cheaper anti-hiv therapy. **Science**, 307(5710):653, 2005.
- M. L. Connolly. Analytical molecular-surface calculation. **Journal of Applied Crystallography**, 16:548–558, 1983a.
- M. L. Connolly. Solvent-accessible surfaces of proteins and nucleic-acids. **Science**, 221(4612):709–713, 1983b.
- K. A. Crandall. **The Evolution of HIV**, Capítulo 1. 1 edição, 1999a.
- K. A. Crandall. **The Evolution of HIV**, Capítulo 2. 1 edição, 1999b.
- F. Crick. Central dogma of molecular biology. **Nature**, 227:561–563, 1970.
- D. A. Davis, K. Dorsey, e et al. Regulation of hiv-1 protease activity through cysteine modification. **Biochemistry**, 35(7):2482–2488, 1996.
- C. Debouck, J. G. Gorniak, e et al. Human immunodeficiency virus protease expressed in escherichia coli exhibits autoprocessing and specific maturation of the gag precursor. **Science**, 84(24):8903–8906, 1987.
- Z. Debyser, Peter Cherepanov, Bénédicte Van Maele, Erik De Clercq, e Myriam Witvrouw. In search of authentic inhibitors of hiv-1 integration. **Antivir Chem Chemother**, 13(1):1–15, 2002.
- S. G. Deeks, M. Smith, e et al. HIV-1 protease inhibitors - a review for clinicians. **Jama-Journal of the American Medical Association**, 277(2):145–153, 1997.
- X. Deift e Zhou. A steepest descent method for oscillatory riemann-hilbert problems. asymptotics for the MKdV equation. **Ann. of Math**, 137(2):295–368, 1993.

- J. W. Erickson e S. K. Burt. Structural mechanisms of HIV drug resistance. **Annu. Rev. Pharmacol. Toxicol**, 36:545–571, 1996.
- M. Franti, T. O’neill, e et al. Pro 542 (cd4-igg2) has a profound impact on HIV-1 replication in the hu-pbl-scid mouse model. In **9th CROI**, eattle, USA, 2002.
- D. I. Freedberg, R. Ishima, J. Jacob, e et al. Rapid structural fluctuations of the free HIV protease flaps in solution: relationship to crystal structures and comparison with predictions of dynamics calculations. **Protein Science**, 11(2):211–232, 2002.
- A. Friedman-Kien, L. Laubenstein, e et al. Kaposi’s sarcoma and pneumocystis pneumonia among homosexual men-new york and california. **MMWR**, 30: 305–308, 1981.
- M. J. Frisch, G. W. Trucks, e et al. Gaussian94, revision b.1. Pittsburgh: Gaussian, 1995.
- T. Fujiwara. Phase 1 multiple oral dose safety and pharmacokinetic study of s-1360, an HIV integrase inhibitor with healthy volunteers. In **XIV International AIDS Conference**, Barcelona, Spain, 2002.
- R. C. Gallo, P. S. Sarin, E. P. Gelmann, e et al. Isolation of human t cell leukemia virus in acquired immune deficiency syndrome (aids). **Science**, 220:865–867, 1983.
- J. W. Gibbs. **Elementary Principles in Statistical Mechanics**, Capítulo XIII. 3 edição, 1902.
- H. M. Ginzburg, S. H. Weiss, M. G. MacDonald, R. L. Hubbard, e et al. Htlv-iii exposure among drug users. **Cancer research**, 45:4605–4608, 1985.
- P. V. Z. C. Goliatt. Técnicas de bioinformática e modelagem computacional aplicadas ao estudo do genoma de trypanossoma cruzi e de enzimas consideradas de interesse no tratamento da doença de chagas: Estudo particular das cruzipaínas

- 1 e 2. Dissertação de Mestardo, Grupo de Modelagem Molecular de Sistemas Biológicos - Laboratório Nacional de Computação Científica / Brasil, 2007.
- J. Gonçalves e et al. Role a vif in human immunodeficiency virus type 1 reverse transcription. **J Virol**, 70(12):8701–8709, 1996.
- J. Gonzalez, R. M. Brindeiro, R. S. Aguiar, H. S. Pereira, C. M. Abreu, e et al. Impact of nelfinavir resistance mutations on in vitro phenotype, fitness, and replication capacity of human immunodeficiency virus type 1 with subtype b and c proteases. **J Virol**, 48(9):3552–3555, 2004.
- S. Grzesiek, A. Bax, L. K. Nicholson, e et al. Nmr evidence for the displacement of a conserved interior water molecule in HIV protease by a nonpeptide cyclic urea-based inhibitor.
- N. Guex e M. C. Peitsch. Swiss-model and the swiss-pdb viewer: An environment for comparative protein modeling. **Electrophoresis**, 18(15):2714–2723, 1997. <http://expasy.org/spdbv/>.
- T. A. Halgren. Merck molecular force field. i. basis, form, scope, parametrization and performance of mmff94. **Journal of Computational Chemistry**, 17 (5&6):490–519, 1996.
- C. Harris, C. B. Small, e et al. Epidemiologic notes and reports immunodeficiency among female sexual partners of males with acquired immune deficiency syndrome (aids) - new york. **MMWR**, 31(52):697–698, 1983.
- J. Helemar, E. Gouws, PD. Ghys, e S. Osmanov. Global and regional distribution of HIV-1 genetic subtypes and recombinantes in 2004. **AIDS**, 20(16):13–23, 2006.
- A. D. Hershey e M. Chase. Independent functions of viral protein and nucleic acid in growth of bacteriophage. **J. Gen. Physiol**, 36:39–56, 1952.

- B. Hess, H. Bekker, e et al. Lincs: A linear constraint solver for molecular simulations. **Journal of Computational Chemistry**, 18(12):1463–1472, 1997.
- R. W. Hockney. The potential calculation and some applications. **Methods in Computational Physics**, 9:136–211, 1970.
- C. Hoffman e F. Mulcahy. Inibidores das proteases (IP's), 2006. [http://hivmedicine.aidsportugal.com/05\\_HAART\\_2\\_4.php](http://hivmedicine.aidsportugal.com/05_HAART_2_4.php) (acessado em 15/03/2008).
- L. Holm e C. Sander. Mapping the protein universe. **Science**, 273(2):595–602, 1996.
- J. Hornberger e et al. Clinical prognosis of enfuvirtide in combination with an optimized background regimen among categories of baseline cd4+ cell count and HIV antiretroviral resistance. In **Abstract 478, 2nd IAS 2003**, Paris, 2003.
- T. Huet, R. Cheynier, e et al. Genetic organization of a chimpanzee lentivirus related to HIV-1. **Nature**, 345(6273):356–359, 1990.
- L. J. Hyland e et al. Human immunodeficiency virus-1 protease. 1.initial velocity studies and kinetic characterization of reaction intermediates by 18o isotope exchange. biochemistry. **Comal Physics**, 30:8441–8453, 1991.
- K. B. Hymes, J. B. Greene, e et al. Kaposi sarcoma in homosexual men - a report of 8 cases. **Lancet**, 2(8247):598–600, 1981.
- M. Ingr, T. Uhlikova, e et al. Kinetics of the dimerization of retroviral proteases: the "fireman's grip" and dimerization. **Science**, 12(10):2173–2182, 2003.
- J. M. Jacobson, R. J Israel, I. Lowy, e et al. Single-dose safety, pharmacology, and antiviral activity of the human immunodeficiency virus (HIV) type 1 entry inhibitor pro 542 in HIV-infected adults. **Antimicrobial Agents and Chemotherapy**, 48(2):423–429, 2004.

- M. P. Jeyabalan, E. Nalivaika, e C. A. Schiffer. How does a symmetric dimer recognize an asymmetric substrate? a substrate complex of HIV-1 protease. **Annual Reviews of Biophysics and Biomolecular Structure**, 301:1207–1220, 2000.
- W. Kabsch e C. Sander. Dictionary of secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. **Biopolymers**, 22:2577–2637, 1983.
- N. Kontorinis e D. T. Dieterich. Toxicity of non-nucleoside analogue reverse transcriptase inhibitors. **Seminars in liver disease**, 23:173–182, 2003.
- J. Lalezari, DeJesus E, Northfelt D, e et al. A week 48 assessment of a randomized, controlled, open-label phase ii trial (t20-206) evaluating 3 doses of t-20 in pi-experienced, nrti-naïve patients infected with hiv-1. In **Abstract 418, 9th CROI**, Seattle, USA, 2002.
- R. A. Laskowski, M. W. Macarthur, D. S. Moss, e J. M. Thornton. Procheck: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, 26:283–291, 1993.
- P. Lemey, O. G. Pybus, B. Wang, N. K. Saksena, M. Salemi, e A. Vandamme. Tracing the origin and history of the hiv-2 epidemic. **PNAS**, 100(11):6588–6592, 2003.
- J. A. Levy, L. s. Kaminsky, e et al. Infection by the retrovirus associated with the acquired immunodeficiency syndrome. clinical biological, and molecular features. **Ann Intern Med**, 103(5):694–699, 1985.
- S. Loya, A. Rudi, e Y. Kashman A. Hizi. Mode of inhibition of hiv-1 reverse transcriptase by polyacetylenetriol, a novel inhibitor of rna and dna directed dna polymerases. **The Biochemical journal**, 362:685–692, 2002.
- F. Margottin, S. P. Bour, e et al. A novel human wd protein, h-beta trcp, that interacts with hiv-1 vpu connects cd4 to the er degradation pathway through an f-box motif. **Mol Cell**, 1(4):565–574, 1998.

- M. A. Marti-Renom, A. C Stuart, e et al. Comparative protein structure modeling of genes and genomes. **Annu. Rev. Biophys Biomol Struct.**, 29:291–325, 2000.
- J. L. Marx. Human t-cell leukemia virus linked to aids. **Science**, 220:806–809, 1983.
- M. Miller, M. Jaskolski, e et al. Crystal structure of a retroviral protease proves relationship to aspartic protease family. **Nature**, 337(6207):576–579, 1989.
- M. D. Miller, C. M. Farnet, e et al. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. **J. Virol**, 71(7):5382–5390, 1997.
- S. Miyamoto e P. A. Kollman. Settle - an analytical version of the shake and rattle algorithm for rigid water models. **Journal of Computational Chemistry**, 13(6):952–962, 1992.
- J. E. F. Murzycki, W. R. P. Scott, e C. A. Schiffer. Hydrophobic sliding: Possible mechanism for drug resistance in human immuno deficiency virus type 1 protease. **Cell Press**, 15:225–233, 2008.
- M. A. Navia, P. M. Fitzgerald, e et al. Three-dimensional structure of aspartyl protease from human immunodeficiency virus hiv-1. **Nature**, 337(6208):615–620, 1989.
- F. J. Palella, K. M. Delaney, A. C. Moorman, M. O. Loveless, e et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. hiv outpatient study investigators. **New England Journal of Medicine**, 338:853–860, 1998.
- L. H. Pearl e W. R. Taylor. A structural model for the retroviral proteases. **Nature**, 329:351–354, 1987.



- J. M. A. Pereira. Vírus da imunodeficiência humana - introdução histórica, 2000.  
<http://www.ordembilogos/Biologias/N1.html> (acessado em 30/06/2006).
- C. F. Perno, A. Cozzi-Lepri, e et al. Secondary mutations in the protease region of human immunodeficiency virus and virologic failure in drug-naive patients treated with protease inhibitor-based therapy. **J. Infect Dis.**, 184(8):983–991, 2001.
- F. Re, D. Braaten, E. K. Franke, J. Luban, e et al. Human immunodeficiency virus type 1 vpr arrests the cell cycle in G2 by inhibiting the activation of p34cdc2-cyclin b. **J. Virol**, 69(11):6859–6864, 1995.
- P. Reid, H. Maclnnes, M. E. Cong, e et al. Natural resistance of human immunodeficiency virus type 2 to zidovudine. **Virology**, 336:251–264, 2005.
- J. R. Reynes, R. Rouzier, e et al. Sch c: safety and antiviral effects of a ccr5 receptor antagonist in hiv-1- infected subjects. In **9th CROI**, Seattle, USA, 2002.
- J. Riley, L. Wojcik, e et al. Genotypic and phenotypic analysis of in vitro generated hiv-1 escape isolates to the ccr5 antagonist sch-c. In **9th CROI**, Seattle, USA, 2002.
- N. A. Roberts, J. A. Martin, e et al. Rational design of peptide-based hiv proteinase inhibitors. **Science**, 248(4953):358–361, 1990.
- E. C. Sabino. Subtipos de HIV no brasil, 1996.  
<http://www.hemolnline.com.br/artigos/subipos.php>(acessado em 02/05/2006).
- A. W. S. Silva. **Estudo Por Modelagem E Dinâmica Molecular Da Protease De Variantes Do Vírus Da Imunodeficiência Humana Tipo 1 Resistentes A Drogas Antivirais**. Tese de Doutorado, Instituto de Biofísica Carlos Chagas Filho - Universidade Federal do Rio de Janeiro / Brasil, 2003.

- P. E. Smith e W. F. Van Gunsteren. consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300k. **Journal of Chemistry Physics**, 100(4):3169–3174, 1994.
- M. A. Soares, T. de Oliveira, R. M. Brindeiro, A. Tanuri, e et al. A specific subtype c of human immunodeficiency virus type 1 circulates in brazil. **AIDS**, 17:11–21, 2003.
- G. L. S. C. Souza, D. E. B. Gomes, e A. W. S. Silva. Surfmds. Desenvolvido no LMDM – Laboratório de Modelagem e Dinâmica Molecular do Instituto de Biofísica Carlos Chagas Filho, UFRJ, versão atualizada de 2007.
- M. R. Spiegel. **Estatística**, páginas 74–75, 106–110. 3 edição, 1994.
- S. Spinelli, Q. Z. Liu, R. J. Poljak, e et al. The three-dimensional structure of the aspartyl protease from the hiv-1 isolate bru. **Biochemie**, 73:1391–1396, 1991.
- J. Stover, S. Bertozzi, J. Gutierrez, N. Walker, K. A. Stanecki, R. Greener, e et al. The global impact of scaling up hiv/aids prevention programs in low- and middle-income countries. **Science**, 311(5766):1474–1476, 2006.
- J. M. Strizki, S. Xu, e et al. Sch-c (sch 351125), an orally bioavailable, small molecule antagonist of the chemokine receptor ccr5, is a potent inhibitor of hiv-1 infection in vitro and in vivo. **PNAS**, 98(22):12718–23, 2001.
- H. N. Temin e S. Mizutani. Rna-dependent dna polymerase in virions of rous sarcoma virus. **Nature**, 226:1211–1213, 1970.
- D. P. Tieleman, D. Spoel, e et al. Molecular dynamics of dodecylphosphocholine micelles at three different aggregate sizes: Micellar structure and chain relaxation. **J. Phys.**, 104(27):6380–6388, 2000.
- K. Triques, A. Bourgeois, S. Saragosti, N. Vidal, E. Mpoudi-Ngole, N. Nzilambi, e et al. High diversity of hiv-1 subtype f strains in central africa. **Virology**, 259: 99–109, 1999.

- Unaided. Aids epidemic update. Relatório Técnico **7**, 2007.
- Y. N. Vaishnav e F. Wong-Staal. The biochemistry of aids. **Annu Rev Biochem**, 60:577–630, 1991.
- D. Van der Spoel, E. Lindahl, B. Hess, e et al. Biomolecular simulation: The gromos96 manual and user guide. **Journal of Molecular Modeling**, 7, 8: 306–317, 2001.
- W. F. Van Gunsteren e H. J. C. Berendsen. **Groningen Molecular Simulation (GROMOS) Library Manual**. Groningen: BIOMOS b.v., 1987.
- W. F. Van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, e et al. **Biomolecular Simulation: The GROMOS96 Manual and User Guide**. Zürich, Groningen: vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v., 1996.
- A. Velazquez-Campoy, M. J. Todd, e et al. Catalytic efficiency and vitality of hiv-1 proteases from african viral subtypes. **PNAS**, 98(11):6062–6067, 2001.
- L. Verlet. Computer experiments on classical fluids. i. thermodynamical properties of lennard-jones molecules. **Physical Review**, 159:98–103, 1967.
- M. A. A. Vitória. Gilead e roche anunciam interrupção de estudos com novas drogas anti-retrovirais de resgate, 2004. <http://www.hiv.org.br> (acessado em 04/04/2006).
- F. Wartha, A. H. C. Horn, H. Meiselbach, e H. Sticht. Molecular dynamics simulations of hiv-1 protease suggest different mechanisms contributing to drug resistance. **J. Chem. Theory Comput**, 1:315–324, 2005.
- S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, e et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. **Journal of the American Chemical Society**, 106:765–784, 1984.

- C. Wild, T. Oas, C. MacDanal, D. Bolognesi, e T. Matthews. A synthetic peptide inhibitor of human immunodeficiency virus replication: Correlation between solution structure and viral inhibition. **Proc. Natl. Acad. Sci**, 89:10537– 10541, 1992.
- A. Wlodawer e J. W. Erickson. Structure-based inhibitors of hiv-1 protease. **Annual Reviews of Biophysics and Biomolecular Structure**, 62:543–585, 1993.
- A. Wlodawer e J. Vondrasek. Inhibitors of hiv-1 protease: a major success of structure-assisted drug design. **Annual Reviews of Biophysics and Biomolecular Structure**, 27(1):249–284, 1998.
- T. Yoshinaga, A. Sato, e et al. S-1360: in vitro activity of a new HIV-1 integrase inhibitor in clinical development. In **9th CROI**, Seattle, USA, 2002.

# Apêndice A

## Parametrização do Ligante

[ moleculetype ]

; Name nrexcl

NEL 3

[ atoms ]

; nr	type	resnr	resid	atom	cgnr	charge	mass
1	CH3	1	NEL	C1	14	0.000	15.0350
2	CH3	1	NEL	C2	15	0.000	15.0350
3	CH3	1	NEL	C3	16	0.000	15.0350
4	C	1	NEL	C4	13	0.000	12.0110
5	N	1	NEL	N5	12	-0.280	14.0067
6	H	1	NEL	H6	41	0.280	1.0080
7	C	1	NEL	C7	11	0.380	12.0110
8	O	1	NEL	O8	17	-0.380	15.9994
9	CH1	2	NEL	C9	8	0.000	13.0190
10	CH2	2	NEL	C10	9	0.000	14.0270
11	CH1	2	NEL	C11	2	0.000	13.0190
12	CH2	2	NEL	C12	3	0.000	14.0270
13	CH2	2	NEL	C13	4	0.000	14.0270
14	CH2	2	NEL	C14	5	0.000	14.0270
15	CH2	2	NEL	C15	6	0.000	14.0270

16	CH1	2	NEL	C16	1	0.000	13.0190
17	CH2	2	NEL	C17	10	0.000	14.0270
18	N	2	NEL	N18	7	0.000	14.0067
19	CH2	2	NEL	C19	18	0.000	14.0270
20	CH1	2	NEL	C20	19	0.150	13.0190
21	OA	2	NEL	O21	21	-0.548	15.9994
22	H	2	NEL	H22	42	0.398	1.0080
23	CH1	3	NEL	C23	20	0.000	13.0190
24	CH2	3	NEL	C24	23	0.150	14.0270
25	S	3	NEL	S25	34	-0.300	32.0600
26	C	3	NEL	C26	35	0.150	12.0110
27	C	3	NEL	C27	36	-0.100	12.0110
28	C	3	NEL	C28	37	-0.100	12.0110
29	C	3	NEL	C29	38	-0.100	12.0110
30	C	3	NEL	C30	39	-0.100	12.0110
31	C	3	NEL	C31	40	-0.100	12.0110
32	HC	3	NEL	H32	48	0.100	1.0080
33	HC	3	NEL	H33	49	0.100	1.0080
34	HC	3	NEL	H34	50	0.100	1.0080
35	HC	3	NEL	H35	51	0.100	1.0080
36	HC	3	NEL	H36	52	0.100	1.0080
37	N	4	NEL	N37	22	-0.280	14.0067
38	H	4	NEL	H38	43	0.280	1.0080
39	C	4	NEL	C39	24	0.380	12.0110
40	O	4	NEL	O40	25	-0.380	15.9994
41	C	4	NEL	C41	26	0.000	12.0110
42	C	4	NEL	C42	27	-0.100	12.0110
43	C	4	NEL	C43	28	-0.100	12.0110
44	C	4	NEL	C44	29	-0.100	12.0110

45	C	4	NEL	C45	30	0.150	12.0110
46	OA	4	NEL	O46	32	-0.548	15.9994
47	C	4	NEL	C47	31	0.000	12.0110
48	CH3	4	NEL	C48	33	0.000	15.0350
49	HC	4	NEL	H49	45	0.100	1.0080
50	HC	4	NEL	H50	46	0.100	1.0080
51	HC	4	NEL	H51	47	0.100	1.0080
52	H	4	NEL	H52	44	0.398	1.0080

[ bonds ]

	ai	aj	funct
16	11	2	gb_26 ; CH1 - CH1
16	15	2	gb_26 ; CH1 - CH2
16	17	2	gb_26 ; CH1 - CH2
11	12	2	gb_26 ; CH1 - CH2
11	10	2	gb_26 ; CH1 - CH2
12	13	2	gb_26 ; CH2 - CH2
13	14	2	gb_26 ; CH2 - CH2
14	15	2	gb_26 ; CH2 - CH2
18	9	2	gb_20 ; N - CH1
18	17	2	gb_20 ; N - CH2
18	19	2	gb_20 ; N - CH2
9	10	2	gb_26 ; CH1 - CH2
9	7	2	gb_26 ; CH1 - C
7	5	2	gb_9 ; C - N
7	8	2	gb_4 ; C - O
5	4	2	gb_9 ; N - C
5	6	2	gb_2 ; N - H

4	1	2	gb_26 ; C - CH3
4	2	2	gb_26 ; C - CH3
4	3	2	gb_26 ; C - CH3
19	20	2	gb_26 ; CH2 - CH1
20	23	2	gb_26 ; CH1 - CH1
20	21	2	gb_17 ; CH1 - OA
23	37	2	gb_20 ; CH1 - N
23	24	2	gb_26 ; CH1 - CH2
21	22	2	gb_1 ; OA - H
37	39	2	gb_9 ; N - C
37	38	2	gb_2 ; N - H
24	25	2	gb_30 ; CH2 - S
39	40	2	gb_4 ; C - O
39	41	2	gb_26 ; C - C
41	42	2	gb_15 ; C - C
41	47	2	gb_15 ; C - C
42	43	2	gb_15 ; C - C
42	49	2	gb_3 ; C - HC
43	44	2	gb_15 ; C - C
43	50	2	gb_3 ; C - HC
44	45	2	gb_15 ; C - C
44	51	2	gb_3 ; C - HC
45	47	2	gb_15 ; C - C
45	46	2	gb_12 ; C - OA
47	48	2	gb_26 ; C - CH3
46	52	2	gb_1 ; OA - H
25	26	2	gb_29 ; S - C
26	27	2	gb_15 ; C - C
26	31	2	gb_15 ; C - C



27	28	2	gb_15 ; C	- C
27	32	2	gb_3 ; C	- HC
28	29	2	gb_15 ; C	- C
28	33	2	gb_3 ; C	- HC
29	30	2	gb_15 ; C	- C
29	34	2	gb_3 ; C	- HC
30	31	2	gb_15 ; C	- C
30	35	2	gb_3 ; C	- HC
31	36	2	gb_3 ; C	- HC

[ pairs ]

;ai	aj	funct
-----	----	-------

16	13	1
16	9	1
16	19	1
11	14	1
11	18	1
11	7	1
12	15	1
12	9	1
12	17	1
13	10	1
14	17	1
15	18	1
15	10	1
18	5	1
18	8	1
18	23	1

18	21	1
9	4	1
9	20	1
9	6	1
10	17	1
10	5	1
10	8	1
10	19	1
17	7	1
17	20	1
7	1	1
7	2	1
7	3	1
7	19	1
4	8	1
1	6	1
2	6	1
3	6	1
8	6	1
19	37	1
19	24	1
19	22	1
20	39	1
20	25	1
20	38	1
23	40	1
23	41	1
23	26	1
23	22	1

21	37	1
21	24	1
37	42	1
37	47	1
37	25	1
24	39	1
24	27	1
24	31	1
24	38	1
39	43	1
39	45	1
39	48	1
39	49	1
40	42	1
40	47	1
40	38	1
41	44	1
41	46	1
41	38	1
41	50	1
42	45	1
42	48	1
42	51	1
43	47	1
43	46	1
44	48	1
44	52	1
44	49	1
45	50	1

47	52	1
47	49	1
47	51	1
46	48	1
46	51	1
25	28	1
25	30	1
25	32	1
25	36	1
26	29	1
26	33	1
26	35	1
27	30	1
27	34	1
27	36	1
28	31	1
28	35	1
29	32	1
29	36	1
30	33	1
31	32	1
31	34	1
49	50	1
50	51	1
32	33	1
33	34	1
34	35	1
35	36	1

[ angles ]

	ai	aj	ak	funct		
	11	16	15	2	ga_14; CH1 - CH1 - CH2	110.0000 111.0000
	11	16	17	2	ga_14; CH1 - CH1 - CH2	110.0000 111.0000
	15	16	17	2	ga_12; CH2 - CH1 - CH2	110.0000 109.5000
	16	11	12	2	ga_14; CH1 - CH1 - CH2	110.0000 111.0000
	16	11	10	2	ga_14; CH1 - CH1 - CH2	110.0000 111.0000
	12	11	10	2	ga_12; CH2 - CH1 - CH2	110.0000 109.5000
	11	12	13	2	ga_14; CH1 - CH2 - CH2	110.0000 111.0000
	12	13	14	2	ga_14; CH2 - CH2 - CH2	110.0000 111.0000
	13	14	15	2	ga_14; CH2 - CH2 - CH2	110.0000 111.0000
	16	15	14	2	ga_14; CH1 - CH2 - CH2	110.0000 111.0000
	9	18	17	2	ga_20; CH1 - N - CH2	120.0000 116.0000
	9	18	19	2	ga_20; CH1 - N - CH2	120.0000 116.0000
	17	18	19	2	ga_20; CH2 - N - CH2	120.0000 116.0000
	18	9	10	2	ga_12; N - CH1 - CH2	110.0000 109.5000
	18	9	7	2	ga_12; N - CH1 - C	110.0000 109.5000
	10	9	7	2	ga_12; CH2 - CH1 - C	110.0000 109.5000
	11	10	9	2	ga_12; CH1 - CH2 - CH1	110.0000 109.5000
	16	17	18	2	ga_12; CH1 - CH2 - N	110.0000 109.5000
	9	7	5	2	ga_18; CH1 - C - N	120.0000 115.0000
	9	7	8	2	ga_29; CH1 - C - O	120.0000 121.0000
	5	7	8	2	ga_32; N - C - O	120.0000 124.0000
	7	5	4	2	ga_30; C - N - C	120.0000 122.0000
	7	5	6	2	ga_31; C - N - H	70.0000 123.0000
	4	5	6	2	ga_17; C - N - H	90.0000 115.0000
	5	4	1	2	ga_12; N - C - CH3	110.0000 109.5000
	5	4	2	2	ga_12; N - C - CH3	110.0000 109.5000

5	4	3	2	ga_12; N - C - CH3	110.0000	109.5000
1	4	2	2	ga_12; CH3 - C - CH3	110.0000	109.5000
1	4	3	2	ga_12; CH3 - C - CH3	110.0000	109.5000
2	4	3	2	ga_12; CH3 - C - CH3	110.0000	109.5000
18	19	20	2	ga_12; N - CH2 - CH1	110.0000	109.5000
19	20	23	2	ga_14; CH2 - CH1 - CH1	110.0000	111.0000
19	20	21	2	ga_12; CH2 - CH1 - OA	110.0000	109.5000
23	20	21	2	ga_12; CH1 - CH1 - OA	110.0000	109.5000
20	23	37	2	ga_12; CH1 - CH1 - N	110.0000	109.5000
20	23	24	2	ga_14; CH1 - CH1 - CH2	110.0000	111.0000
37	23	24	2	ga_12; N - CH1 - CH2	110.0000	109.5000
20	21	22	2	ga_11; CH1 - OA - H	95.0000	109.5000
23	37	39	2	ga_30; CH1 - N - C	120.0000	122.0000
23	37	38	2	ga_17; CH1 - N - H	90.0000	115.0000
39	37	38	2	ga_31; C - N - H	70.0000	123.0000
23	24	25	2	ga_15; CH1 - CH2 - S	110.0000	113.0000
37	39	40	2	ga_32; N - C - O	120.0000	124.0000
37	39	41	2	ga_18; N - C - C	120.0000	115.0000
40	39	41	2	ga_29; O - C - C	120.0000	121.0000
39	41	42	2	ga_26; C - C - C	100.0000	120.0000
39	41	47	2	ga_26; C - C - C	100.0000	120.0000
42	41	47	2	ga_26; C - C - C	100.0000	120.0000
41	42	43	2	ga_26; C - C - C	100.0000	120.0000
41	42	49	2	ga_24; C - C - HC	90.0000	120.0000
43	42	49	2	ga_24; C - C - HC	90.0000	120.0000
42	43	44	2	ga_26; C - C - C	100.0000	120.0000
42	43	50	2	ga_24; C - C - HC	90.0000	120.0000
44	43	50	2	ga_24; C - C - HC	90.0000	120.0000
43	44	45	2	ga_26; C - C - C	100.0000	120.0000

43	44	51	2	ga_24; C	- C	- HC	90.0000	120.0000
45	44	51	2	ga_24; C	- C	- HC	90.0000	120.0000
44	45	47	2	ga_26; C	- C	- C	100.0000	120.0000
44	45	46	2	ga_26; C	- C	- OA	100.0000	120.0000
47	45	46	2	ga_26; C	- C	- OA	100.0000	120.0000
41	47	45	2	ga_26; C	- C	- C	100.0000	120.0000
41	47	48	2	ga_26; C	- C	- CH3	100.0000	120.0000
45	47	48	2	ga_26; C	- C	- CH3	100.0000	120.0000
45	46	52	2	ga_11; C	- OA	- H	95.0000	109.5000
24	25	26	2	ga_3 ; CH2	- S	- C	110.0000	100.0000
25	26	27	2	ga_26; S	- C	- C	100.0000	120.0000
25	26	31	2	ga_26; S	- C	- C	100.0000	120.0000
27	26	31	2	ga_26; C	- C	- C	100.0000	120.0000
26	27	28	2	ga_26; C	- C	- C	100.0000	120.0000
26	27	32	2	ga_24; C	- C	- HC	90.0000	120.0000
28	27	32	2	ga_24; C	- C	- HC	90.0000	120.0000
27	28	29	2	ga_26; C	- C	- C	100.0000	120.0000
27	28	33	2	ga_24; C	- C	- HC	90.0000	120.0000
29	28	33	2	ga_24; C	- C	- HC	90.0000	120.0000
28	29	30	2	ga_26; C	- C	- C	100.0000	120.0000
28	29	34	2	ga_24; C	- C	- HC	90.0000	120.0000
30	29	34	2	ga_24; C	- C	- HC	90.0000	120.0000
29	30	31	2	ga_26; C	- C	- C	100.0000	120.0000
29	30	35	2	ga_24; C	- C	- HC	90.0000	120.0000
31	30	35	2	ga_24; C	- C	- HC	90.0000	120.0000
26	31	30	2	ga_26; C	- C	- C	100.0000	120.0000
26	31	36	2	ga_24; C	- C	- HC	90.0000	120.0000
30	31	36	2	ga_24; C	- C	- HC	90.0000	120.0000

[ dihedrals ]

; ajuste manualmente

; ai	aj	ak	al	funct			
16	11	15	17	2	gi_2 ;	CH1 - X - Y - CH2	80.0000 35.26400
11	10	12	16	2	gi_2 ;	CH1 - X - Y - CH1	80.0000 35.26400
18	17	9	19	2	gi_2 ;	N - X - Y - CH2	80.0000 35.26400
9	7	10	18	2	gi_2 ;	CH1 - X - Y - N	80.0000 35.26400
7	5	9	8	2	gi_1 ;	C - X - Y - O	40.0000 0.00000
5	4	6	7	2	gi_1 ;	N - X - Y - C	40.0000 0.00000
4	1	3	2	2	gi_2 ;	C - X - Y - CH3	80.0000 35.26400
20	21	19	23	2	gi_2 ;	CH1 - X - Y - CH1	80.0000 35.26400
23	20	24	37	2	gi_2 ;	CH1 - X - Y - N	80.0000 35.26400
37	23	38	39	2	gi_1 ;	N - X - Y - C	40.0000 0.00000
39	37	41	40	2	gi_1 ;	C - X - Y - O	40.0000 0.00000
41	39	42	47	2	gi_1 ;	C - X - Y - C	40.0000 0.00000
42	41	43	49	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
43	42	44	50	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
44	43	45	51	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
45	44	46	47	2	gi_1 ;	C - X - Y - C	40.0000 0.00000
47	48	41	45	2	gi_1 ;	C - X - Y - C	40.0000 0.00000
26	25	31	27	2	gi_1 ;	C - X - Y - C	40.0000 0.00000
27	26	28	32	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
28	27	29	33	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
29	28	30	34	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
30	29	31	35	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
31	26	30	36	2	gi_1 ;	C - X - Y - HC	40.0000 0.00000
31	26	27	28	2	gi_1 ;	C - X - Y - C	40.0000 0.00000
27	26	31	30	2	gi_1 ;	C - X - Y - C	40.0000 0.00000



26	27	28	29	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
27	28	29	30	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
28	29	30	31	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
29	30	31	26	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
42	41	47	45	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
47	41	42	43	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
41	47	45	44	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
41	42	43	44	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
42	43	44	45	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
43	44	45	47	2	gi_1 ; C	- X - Y - C	40.0000	0.00000	
10	11	12	13	1	gd_17 ; X	- CH1 - CH2 - Y	1.4000	0.0000	3
12	11	10	9	1	gd_17 ; X	- CH1 - CH2 - Y	1.4000	0.0000	3
16	15	14	13	1	gd_17 ; X	- CH2 - CH2 - Y	1.4000	0.0000	3
16	17	18	19	1	gd_19 ; X	- CH2 - N - Y	0.2400	180.0000	6
17	16	15	14	1	gd_17 ; X	- CH1 - CH2 - Y	1.4000	0.0000	3
11	16	17	18	1	gd_17 ; X	- CH1 - CH2 - Y	1.4000	0.0000	3
11	12	13	14	1	gd_17 ; X	- CH2 - CH2 - Y	1.4000	0.0000	3
11	10	9	7	1	gd_17 ; X	- CH2 - CH1 - Y	1.4000	0.0000	3
12	11	16	17	1	gd_17 ; X	- CH1 - CH1 - Y	1.4000	0.0000	3
12	13	14	15	1	gd_17 ; X	- CH2 - CH2 - Y	1.4000	0.0000	3
18	9	7	5	1	gd_20 ; X	- CH1 - C - Y	0.2400	0.0000	6
18	19	20	23	1	gd_17 ; X	- CH2 - CH1 - Y	1.4000	0.0000	3
9	18	19	20	1	gd_19 ; X	- N - CH2 - Y	0.2400	180.0000	6
9	7	5	4	1	gd_4 ; X	- C - N - Y	8.0000	180.0000	2
7	9	18	19	1	gd_19 ; X	- CH1 - N - Y	0.2400	180.0000	6
7	5	4	1	1	gd_4 ; X	- N - C - Y	8.0000	180.0000	2
19	20	23	37	1	gd_17 ; X	- CH1 - CH1 - Y	1.4000	0.0000	3
19	20	21	22	1	gd_12 ; X	- CH1 - OA - Y	0.3000	0.0000	3
20	23	37	39	1	gd_19 ; X	- CH1 - N - Y	0.2400	180.0000	6

```
20  23  24  25  1  gd_17 ; X  - CH1 - CH2 - Y 1.4000 0.0000  3
23  37  39  41  1  gd_4  ; X  - N   - C   - Y 8.0000 180.0000  2
23  24  25  26  1  gd_13 ; X  - CH2 - S   - Y 0.7000 0.0000  3
37  39  41  42  1  gd_1  ; X  - C   - C   - Y 1.4000 180.0000  2
24  25  26  31  1  gd_13 ; X  - S   - C   - Y 0.7000 0.0000  3
47  45  46  52  1  gd_2  ; X  - C   - OA  - Y 1.7000 180.0000  2
```

```
; Include Position restraint file
```

```
#ifdef POSRES
```

```
#include "nelokposre.itp"
```

```
#endif
```

# 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](#)