

**Modelos de QSAR-4D de uma Série de  
Análogos do Raloxifeno: um Modulador  
Seletivo do Receptor de Estrogênio**

*Ana Carolina Rennó Sodero*

Rio de Janeiro – Março de 2007

# **Livros Grátis**

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

Milhares de livros grátis para download.

# **Modelos de QSAR-4D de uma Série de Análogos do Raloxifeno: um Modulador Seletivo do Receptor de Estrogênio**

*Ana Carolina Rennó Sodero*

Dissertação de Mestrado apresentada  
ao Programa de Pós-Graduação em  
Química Orgânica do Instituto de  
Química da Universidade Federal do  
Rio de Janeiro como parte dos  
requisitos necessários à obtenção do  
título de Mestre em Ciências

Orientadoras:

Dra. Magaly Girão Albuquerque (Instituto de Química / UFRJ)

Dra. Nelilma Correia Romeiro (Faculdade de Farmácia / UFRJ)

Rio de Janeiro – Março de 2007

# **Modelos de QSAR-4D de uma Série de Análogos do Raloxifeno: um Modulador Seletivo do Receptor de Estrogênio**

**Ana Carolina Rennó Sodero**

Dissertação de Mestrado submetida ao corpo docente do Departamento de Química Orgânica do Instituto de Química da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Mestre em Ciências.

Aprovada por:

---

Prof. Dra. Magaly Girão Albuquerque (Orientadora/Presidente)  
(Instituto de Química / UFRJ)

---

Dra. Nelilma Correia Romeiro (Orientadora)  
(Faculdade de Farmácia / UFRJ)

---

Prof. Dr. Carlos Mauricio Rabello de Sant'Anna  
(Instituto de Ciências Exatas / UFRRJ)

---

Prof. Dra. Helena Carla Castro  
(Instituto de Biologia / UFF)

---

Prof. Dr. Marcio Contrucci Saraiva de Mattos  
(Instituto de Química / UFRJ)

---

Prof. Dr. Carlos Rangel Rodrigues (Suplente)  
(Faculdade de Farmácia / UFRJ)

---

Prof. Dr. Joaquim Fernando Mendes da Silva (Suplente)  
(Instituto de Química / UFRJ)

Rio de Janeiro – Março de 2007

SODERO, Ana Carolina Rennó

Modelos de QSAR-4D de uma Série de Análogos do Raloxifeno:  
um Modulador Seletivo do Receptor de Estrogênio / Ana Carolina  
Rennó Soderó – 2007

xviii, 112 p.

Dissertação: Mestrado em Ciências (Química Orgânica) –  
Universidade Federal do Rio de Janeiro, Instituto de Química, Rio de  
Janeiro, 2007.

Orientadoras: Dra. Magaly Girão Albuquerque, Dra. Nelilma  
Correia Romeiro.

1. QSAR,
2. Raloxifeno,
3. Modulador Seletivo do Receptor de Estrogênio,
4. Câncer de Mama.

I. Universidade Federal do Rio de Janeiro - IQ

Dissertação realizada no Departamento de Química Orgânica do Instituto de Química da Universidade Federal do Rio de Janeiro sob a orientação da Dra. Magaly Girão Albuquerque e da Dra. Nelilma Correia Romeiro com apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

À minha querida avó Zilda  
pelos anos de dedicação, incentivo  
e pelo exemplo de vida.  
*(in memoriam)*

## AGRADECIMENTOS

Aos meus pais João Carlos e Maria Aida, meus maiores incentivadores. Obrigada pelo amor incondicional e por serem fonte de apoio nos momentos difíceis.

Às minhas irmãs Daniela e Fabrízia e àqueles que se fizeram irmãos, Marcelo e Gustavo. Obrigada por estarem sempre presentes em minha vida.

Aos lindinhos da dinda Paulo Victor e João Marcelo.

Aos meus familiares, em especial aos meus queridos primos Gugu, Guigui e Renata, pela ajuda permanente.

Ao Tutty, meu cachorrinho querido, meu sonho de infância.

Aos amigos Alessandra, minha amiga-divã, e André, pelos sinceros conselhos e por sempre estarem dispostos a me ouvir e apoiar.

À orientadora, amiga e incentivadora Nelilma, pela presença e auxílio constante, obrigada por seu carinho e ajuda.

À orientadora Magaly, pela oportunidade, por ter me acolhido da melhor forma possível no Laboratório de Modelagem Molecular e pela paciência sem limites.

Ao Professor Ricardo Bicca, afinal de contas “nossos professores são muito mais queridos que os professores dos outros...”

À Monique, pela amizade sincera, pela forma doce de tratar a todos e pelas inúmeras contribuições profissionais e também pessoais.

Ao Bruno e Paulo, os viciados em café. Obrigada pelas vezes em que o café saiu cheirosinho da cafeteira, mesmo que tenha sido preciso rolar no chão e esperar. Vocês foram motivos das minhas melhores risadas, além do grande aprendizado sobre o mundo da Modelagem Molecular.

Aos amigos Samuel, Rafael e Lucas. Nossas discussões foram sempre muito valiosas.

Aos demais amigos da 609: Jackeline, Thiago, Luiz Felipe, André, Jocley, Fabiano e Nadja. Obrigada pela convivência sempre agradável.

Aos amigos do IQ Ana Paula, Silvia, Alessandro, Renato, Márcia, Ana Paula Bernardo, Juan, Alessandra e Carol.

Ao Fleming, por me mostrar tantas coisas diferentes, pelas inúmeras discussões científicas e pelos momentos mais felizes. Obrigada por ter tornado minha vida muito mais divertida.

À Banca Examinadora, obrigada por aceitar o convite, pelas críticas e pelas contribuições positivas ao trabalho.

Aos órgãos de fomento CNPq, FAPERJ, CAPES, FUJB e ao IQ por tornarem possível a realização do mestrado.



"Sinto-me nascido a cada momento para a eterna novidade do mundo"

Fernando Pessoa

## RESUMO

A metodologia de correlação quantitativa estrutura-atividade em quatro dimensões (QSAR-4D) foi aplicada a 54 análogos do raloxifeno, um antagonista do receptor de estrogênio  $\alpha$ , sintetizados e avaliados farmacologicamente como inibidores da proliferação de células MCF-7 estimulada pelo hormônio  $17\beta$ -estradiol (Grese *et al.*, 1997a). Foram utilizados 41 compostos para a construção dos modelos (i.e. conjunto de treinamento) e 13 compostos para a validação externa (i.e. conjunto de teste). As conformações obtidas por simulação de Dinâmica Molecular (DM) foram sobrepostas, de acordo com os três alinhamentos testados, em uma caixa virtual tridimensional composta por células de 2 e 1Å. Os descritores (*Grid Cell Occupancy Descriptors, GCODs*) se referem à frequência de ocupação de cada célula por diferentes tipos de átomos. Os modelos foram gerados pela técnica combinada de algoritmos genéticos (*Genetic Algorithm, GA*) e mínimos quadrados parciais (*Partial Least Square, PLS*) e avaliados por validação cruzada. Os modelos de QSAR-4D obtidos foram coerentes com o mecanismo de ação proposto para o raloxifeno, permitindo uma previsão quantitativa da atividade biológica da série em estudo e auxiliando no planejamento de novos análogos estruturais do raloxifeno.

**Palavras-chave:** QSAR, Raloxifeno, Modulador Seletivo do Receptor de Estrogênio, Câncer de Mama.

## ABSTRACT

Four-dimensional quantitative structure-activity relationship (4D-QSAR) analysis was applied on a series of 54 raloxifene analogs (Grese *et al.*, 1997a), an estrogen receptor  $\alpha$  antagonist, which was synthesized and evaluated as estrogen-stimulated proliferation of MCF-7 cells. The training set was composed by 41 compounds and the test set was composed by 13. The conformations of each analogue, sampled from a molecular dynamic simulation, were placed in a grid cell lattice for each of the three trial alignments, considering grid cell sizes of 2 and 1Å. The grid cell occupancy descriptors (*GCODs*) correspond to the frequency of occupation of each grid cell by different atom types. The models were generated by a scheme that combines genetic algorithms (GA) with partial least squares (PLS) regression and were evaluated by cross-validation. The obtained 4D-QSAR models are in agreement with the proposed action mechanism for raloxifene. This study allowed a quantitative prediction of compounds potency and supported the design of new raloxifene analogs.

**Keywords:** QSAR, Raloxifene, Selective Estrogen Receptor Modulator, Breast Cancer.

# SUMÁRIO

<b>RESUMO</b> .....	<b>VIII</b>
<b>ABSTRACT</b> .....	<b>IX</b>
<b>SUMÁRIO</b> .....	<b>X</b>
<b>LISTA DE SIGLAS E ABREVIATURAS</b> .....	<b>XIII</b>
<b>ÍNDICE DE FIGURAS</b> .....	<b>XIV</b>
<b>ÍNDICE DE TABELAS</b> .....	<b>XVIII</b>
<b>1 INTRODUÇÃO</b> .....	<b>1</b>
1.1 CÂNCER DE MAMA E ESTROGÊNIOS .....	1
1.2 MODULADORES SELETIVOS DE RECEPTORES DE ESTROGÊNIO .....	3
1.3 RECEPTORES DE ESTROGÊNIO.....	8
1.4 MECANISMO DE AÇÃO DE LIGANTES DE RECEPTORES DE ESTROGÊNIO	10
1.5 RELAÇÃO ENTRE ESTRUTURA E ATIVIDADE DOS LIGANTES DE RECEPTORES DE ESTROGÊNIO.....	13
1.5.1 Estrogênios Naturais.....	13
1.5.2 Estrogênios Semi-Sintéticos.....	17
1.5.3 Estrogênios Não-Esteroidais.....	18
1.5.4 Moduladores Seletivos dos Receptores de Estrogênio .....	20
1.5.4.1 Tamoxifeno e 4-hidroxitamoxifeno .....	20
1.5.4.2 Raloxifeno .....	24
1.6 METODOLOGIAS DE CORRELAÇÃO QUANTITATIVA ESTRUTURA- ATIVIDADE .....	30
1.6.1 Metodologia de QSAR-4D .....	33

<b>2</b>	<b>OBJETIVOS .....</b>	<b>35</b>
<b>3</b>	<b>METODOLOGIA.....</b>	<b>36</b>
3.1	DADOS BIOLÓGICOS: SÉRIE DE DERIVADOS DO RALOXIFENO .....	36
3.2	CONSTRUÇÃO E OTIMIZAÇÃO DAS ESTRUTURAS TRIDIMENSIONAIS E OBTENÇÃO DAS CARGAS ATÔMICAS PARCIAIS .....	38
3.3	AMOSTRAGEM CONFORMACIONAL POR SIMULAÇÃO DE DINÂMICA MOLECULAR .....	39
3.4	DEFINIÇÃO DOS ELEMENTOS DE INTERAÇÃO FARMACOFÓRICA .....	40
3.5	DEFINIÇÃO DO TAMANHO DAS CÉLULAS DA CAIXA VIRTUAL TRIDIMENSIONAL.....	41
3.6	DEFINIÇÃO DOS ALINHAMENTOS.....	41
3.7	DEFINIÇÃO DOS TIPOS DE OCUPAÇÃO DAS CÉLULAS .....	43
3.8	OBTENÇÃO DAS VARIÁVEIS INDEPENDENTES .....	44
3.9	REDUÇÃO DOS DADOS .....	44
3.10	OBTENÇÃO DOS MODELOS DE QSAR-4D.....	45
3.11	VALIDAÇÃO INTERNA DOS MODELOS .....	47
3.12	SELEÇÃO DOS MELHORES MODELOS .....	48
3.13	VALIDAÇÃO EXTERNA DOS MODELOS E OBSERVAÇÃO DOS <i>OUTLIERS</i> .....	48
3.14	SELEÇÃO DA CONFORMAÇÃO “BIOATIVA” .....	49
<b>4</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>50</b>
4.1	ANÁLISE DA SIMULAÇÃO POR DINÂMICA MOLECULAR .....	50
4.2	SELEÇÃO DO MELHOR ALINHAMENTO.....	50
4.3	ANÁLISE DOS MELHORES MODELOS DE QSAR-4D DO ALINHAMENTO 1.54	
4.3.1	Análise do Modelo Representativo do Alinhamento 1.....	56

4.3.1.1	Análise dos Valores Residuais e Identificação dos <i>Outliers</i> do Modelo Representativo do Alinhamento 1 .....	63
4.3.1.2	Validação Externa do Modelo Representativo do Alinhamento 1 .....	67
4.4	ANÁLISE DOS MELHORES MODELOS DE QSAR-4D DO ALINHAMENTO 2 .....	70
4.4.1	Análise do Modelo Representativo do Alinhamento 2.....	73
4.4.1.1	Análise dos Valores Residuais e Identificação dos <i>Outliers</i> do Modelo Representativo do Alinhamento 2 .....	78
4.4.1.2	Validação Externa do Modelo Representativo do Alinhamento 2 .....	83
4.5	NOVOS COMPOSTOS BASEADOS NA ANÁLISE DE QSAR-4D .....	86
<b>5</b>	<b>CONCLUSÕES E PERSPECTIVAS .....</b>	<b>92</b>
<b>6</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>93</b>
<b>ANEXO I</b>	<b>.....</b>	<b>103</b>
<b>ANEXO II</b>	<b>.....</b>	<b>104</b>
<b>ANEXO III</b>	<b>.....</b>	<b>105</b>
<b>ANEXO III (CONT.)</b>	<b>.....</b>	<b>106</b>
<b>ANEXO IV</b>	<b>.....</b>	<b>107</b>
<b>ANEXO V</b>	<b>.....</b>	<b>107</b>
<b>ANEXO VI</b>	<b>.....</b>	<b>109</b>
<b>ANEXO VII</b>	<b>.....</b>	<b>109</b>
<b>ANEXO VII (CONT.)</b>	<b>.....</b>	<b>111</b>
<b>ANEXO VIII</b>	<b>.....</b>	<b>112</b>

## LISTA DE SIGLAS E ABREVIATURAS

AF	<i>Activation Function</i> (Função de Ativação)
Ala	Alanina
Arg	Arginina
Asp	Ácido aspártico
DM	Dinâmica Molecular
DNA	<i>Deoxyribonucleic Acid</i> (Ácido Desoxirribonucleico)
DP <sub>res</sub>	Desvio-Padrão dos Valores Residuais
ER	<i>Estrogen Receptor</i> (Receptor de Estrogênio)
ERE	<i>Estrogen Response Elements</i> (Elementos Responsivos ao Estrogênio)
FDA	<i>Food and Drug Administration</i> (Agência Controladora de Drogas e Alimentos)
GA	<i>Genetic Algorithms</i> (Algoritmos Genéticos)
GFA	<i>Genetic Function Approximation</i> (Aproximação da Função Genética)
GCODs	<i>Grid Cell Occupancy Descriptors</i> (Descritores de Ocupação das Células)
Glu	Ácido glutâmico
Gly	Glicina
His	Histidina
Ile	Isoleucina
IPEs	<i>Interaction Pharmacophore Elements</i> (Elementos de Interação Farmacofórica)
IR	Independente do Receptor
LBD	<i>Ligand-Binding Domain</i> (Domínio de Ligação do Ligante)
Leu	Leucina
LOO	<i>Leave-One-Out Cross-Validation</i> (Método Validação Cruzada por Exclusão Individual Exaustiva)
LOF	<i>Lack of Fit</i> (Fator de Desajuste)
LSE	<i>Least Square Error</i> (Erro de Mínimos Quadrados)
MAP kinase	<i>Mitogen-Activated Protein Kinase</i> (Proteína Quinase Ativada por Mitógenos)
Met	Metionina
mRNA	<i>Messenger Ribonucleic Acid</i> (Ácido Ribonucleico Mensageiro)
PDB	<i>Protein Data Bank</i>
Phe	Fenilalanina
PLS	<i>Partial Least Square</i> (Mínimos Quadrados Parciais)
QSAR	<i>Quantitative Structure-Activity Relationship</i> (Correlação Quantitativa Estrutura-Atividade)
RBA	<i>Relative Binding Affinity</i> (Afinidade Relativa de Ligação)
RMS	<i>Root Mean Square</i> (Raiz dos Mínimos Quadrados)
SERM	<i>Selective Estrogen Receptor Modulator</i> (Modulador Seletivo do Receptor de Estrogênio)
Thr	Treonina
Tyr	Tirosina
Trp	Triptofano
Val	Valina

## ÍNDICE DE FIGURAS

<b>Figura 1</b>	Estruturas químicas do 17 $\beta$ -estradiol e do colesterol, precursor dos hormônios esteroidais.....	02
<b>Figura 2</b>	Estrutura química do etamoxitriptol.....	04
<b>Figura 3</b>	Estruturas químicas de alguns <i>SERMs</i> (A) de primeira e (B) de segunda geração.....	07
<b>Figura 4</b>	Comparação entre as seqüências de aminoácidos dos receptores de estrogênio $\alpha$ e $\beta$ (Adaptado de Klinge, 2000). Entre parênteses estão indicados o percentual de homologia entre as seqüências dos domínios A/B, C, D, E e F.....	09
<b>Figura 5</b>	Representação esquemática do mecanismo de ação de ligantes (agonistas e antagonistas) do receptor de estrogênio (Adaptado de McDonnel, 2004).....	12
<b>Figura 6</b>	Estrutura comum dos hormônios esteroidais, representada pelo composto hipotético ciclopentaperhidrofenantreno. A numeração é IUPAC.....	14
<b>Figura 7</b>	Estrutura química dos estrogênios endógenos 17 $\beta$ -estradiol, estrona e estriol.....	14
<b>Figura 8</b>	Representação da estrutura tridimensional do <i>LBD</i> do <i>ER-<math>\alpha</math></i> complexado ao 17 $\beta$ -estradiol obtida por cristalografia de raios-X (código <i>PDB</i> : 1ERE): (A) principais estruturas secundárias ao redor do ligante; (B) interações por ligação hidrogênio do ligante com o <i>LBD</i> .....	16
<b>Figura 9</b>	Tratamento de derivados da estrona com acetilato de potássio em amônia líquida para produzir compostos mais potentes quando administrados por via oral (Loose-Mitchell & Stancel, 2003).....	17
<b>Figura 10</b>	Estruturas químicas do núcleo estilbeno e do dietilestilbestrol, um composto sintético não-esteroidal com potente atividade agonista dos receptores de estrogênio..	18
<b>Figura 11</b>	Representação da estrutura tridimensional da região do <i>LBD</i> do <i>ER-<math>\alpha</math></i> complexada ao dietilestilbestrol obtida por cristalografia de raios-X (código <i>PDB</i> : 3ERD): (A) principais estruturas secundárias ao redor do ligante; (B) interações por ligação hidrogênio do ligante com os resíduos Glu353, Arg394 e His524 e uma molécula de água e interações de van der Waals do ligante com os resíduos Met343, Leu525, His524, Gly521 e Met421.....	19
<b>Figura 12</b>	Estruturas químicas dos análogos Ia e Ib do tamoxifeno (Kraft <i>et al.</i> , 1999).....	21
<b>Figura 13</b>	Estrutura química do 4-hidroxitamoxifeno.....	21
<b>Figura 14</b>	Representação da estrutura tridimensional do complexo formado entre o 4-hidroxitamoxifeno e a região do <i>LBD</i> do <i>ER-<math>\alpha</math></i> obtida por cristalografia de raios-x (código <i>PDB</i> : 3ERT): (A) principais estruturas secundárias ao redor do ligante; (B) interações por ligação hidrogênio entre o ligante e o <i>LBD</i> .....	23
<b>Figura 15</b>	Representação das interações de van der Waals entre o 4-hidroxitamoxifeno e alguns resíduos do <i>ER-<math>\alpha</math></i> observadas por cristalografia de raios-X (código <i>PDB</i> : 3ERT).....	23



<b>Figura 16</b>	Estruturas químicas do (A) tamoxifeno e do (B) raloxifeno. As setas indicam as características marcantes no raloxifeno quando comparado ao tamoxifeno (Adaptado de Grese <i>et al.</i> , 1997b).....	25
<b>Figura 17</b>	Estruturas químicas dos análogos VI e VII do raloxifeno.....	27
<b>Figura 18</b>	Representação da estrutura tridimensional da região do <i>LBD</i> do <i>ER-<math>\alpha</math></i> complexado ao raloxifeno obtida por cristalografia de raios-X (código <i>PDB</i> : 1ERR): (A) principais estruturas secundárias ao redor do ligante; (B) interações por ligação hidrogênio do ligante com o <i>LBD</i> .....	29
<b>Figura 19</b>	Comparação entre as diferentes orientações adotadas pela cadeia lateral do resíduo His524 do <i>ER-<math>\alpha</math></i> quando em complexo com: (A) o 17 $\beta$ -estradiol (código <i>PDB</i> : 1ERE) e (B) o raloxifeno (código <i>PDB</i> : 1ERR). Esta diferença permite uma orientação adequada deste resíduo para manter uma interação por ligação hidrogênio com o ligante.....	29
<b>Figura 20</b>	Distribuição dos valores de atividade ( $pIC_{50exp}$ ) dos 54 compostos derivados do raloxifeno selecionados para o estudo de QSAR-4D (Grese <i>et al.</i> , 1997a).....	38
<b>Figura 21</b>	Gráfico dos valores de energia potencial (kcal/mol) <i>versus</i> tempo de simulação (ps) por simulação de dinâmica molecular em temperatura constante (310K) para o composto 1.....	50
<b>Figura 22</b>	Valores de $q^2_{ajustado}$ <i>versus</i> número de termos dos modelos dos alinhamentos 1 a 3 em células de 2Å.....	53
<b>Figura 23</b>	Valores de $q^2_{ajustado}$ <i>versus</i> número de termos dos modelos dos alinhamentos 1 a 3 em células de 1Å.....	53
<b>Figura 24</b>	Representação da conformação “bioativa” do composto <b>1</b> (cinza), obtida pelo modelo 1B9 de QSAR-4D, sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do <i>ER-<math>\alpha</math></i> (código <i>PDB</i> 1ERR), pelo alinhamento 1. A ocupação dos <i>GCODs</i> em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os <i>GCODs</i> descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).....	60
<b>Figura 25</b>	Representação da conformação “bioativa” do composto <b>54</b> (cinza), obtida pelo modelo 1B9 de QSAR-4D, sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do <i>ER-<math>\alpha</math></i> (código <i>PDB</i> 1ERR), pelo alinhamento 1. A ocupação dos <i>GCODs</i> em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os <i>GCODs</i> descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).....	60
<b>Figura 26</b>	Valores de atividade experimentais ( $pIC_{50exp}$ ) <i>versus</i> valores de atividade calculados ( $pIC_{50calc}$ ) dos 41 análogos do raloxifeno utilizados para a construção do modelo 1B9 de QSAR-4D.....	65
<b>Figura 27</b>	Valores dos resíduos das potências dos 41 análogos do raloxifeno utilizados para a construção do modelo 1B9 de QSAR-4D.....	65

<b>Figura 28</b>	Representação da conformação “bioativa” do composto <b>9</b> (cinza), identificado como <i>outlier</i> , obtida pelo modelo 1B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do <i>ER-α</i> (código <i>PDB 1ERR</i> ), pelo alinhamento 1. A ocupação dos <i>GCODs</i> em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os <i>GCODs</i> descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).	67
<b>Figura 29</b>	Valores de atividade experimentais ( $pIC_{50exp}$ ) versus valores de atividade calculados ( $pIC_{50calc}$ ) dos 13 compostos utilizados como conjunto de teste para a validação externa do modelo 1B9 de QSAR-4D.....	68
<b>Figura 30</b>	Valores dos resíduos dos 13 compostos do conjunto de teste utilizados para a validação externa do modelo 1B9 de QSAR-4D.....	69
<b>Figura 31</b>	Representação da conformação “bioativa” do composto <b>51</b> (cinza), identificado como <i>outlier</i> , obtida pelo modelo 1B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do <i>ER-α</i> (código <i>PDB 1ERR</i> ), pelo alinhamento 1. A ocupação dos <i>GCODs</i> em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os <i>GCODs</i> descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).	70
<b>Figura 32</b>	Representação da conformação “bioativa” do composto <b>1</b> (cinza) obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do <i>ER-α</i> (código <i>PDB 1ERR</i> ), pelo alinhamento 2. A ocupação dos <i>GCODs</i> em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os <i>GCODs</i> descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,11,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).....	76
<b>Figura 33</b>	Representação da conformação “bioativa” do composto <b>54</b> (cinza) obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do <i>ER-α</i> (código <i>PDB 1ERR</i> ), pelo alinhamento 2. A ocupação dos <i>GCODs</i> em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os <i>GCODs</i> descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,11,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).....	76
<b>Figura 34</b>	Valores de atividade biológica experimentais ( $pIC_{50exp}$ ) versus calculados ( $pIC_{50calc}$ ) dos 41 derivados do raloxifeno utilizados para a construção do modelo 2B9 de QSAR-4D.....	80
<b>Figura 35</b>	Valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) dos 41 derivados do raloxifeno utilizados para a construção do modelo 2B9 de QSAR-4D.....	80

- Figura 36** Representação da conformação “bioativa” do composto **2** (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER- $\alpha$*  (código *PDB 1ERR*), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).. 82
- Figura 37** Representação da conformação “bioativa” do composto **49** (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER- $\alpha$*  (código *PDB 1ERR*), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).. 82
- Figura 38** Representação da conformação “bioativa” do composto **52** (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER- $\alpha$*  (código *PDB 1ERR*), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).. 83
- Figura 39** Valores de atividade experimentais ( $pIC_{50exp}$ ) versus valores de atividade calculados ( $pIC_{50calc}$ ) dos 13 compostos utilizados como conjunto de teste para a validação externa do modelo 2B9 de QSAR-4D..... 84
- Figura 40** Valores residuais dos 13 compostos do conjunto de teste utilizados para a validação externa do modelo 2B9 de QSAR-4D..... 85
- Figura 41** Representação da conformação “bioativa” do composto **51** (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER- $\alpha$*  (código *PDB 1ERR*), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).. 86
- Figura 42** Representação esquemática das modificações estruturais propostas para o raloxifeno, baseadas no modelo 1B9 de QSAR-4D. As letras R, R' e R'' correspondem aos diferentes substituintes presentes nos análogos propostos..... 88
- Figura 43** Representação da conformação inicial do análogo **C2** (cinza) sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER- $\alpha$*  (código *PDB 1ERR*), pelo alinhamento 1. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa). 90

## ÍNDICE DE TABELAS

<b>Tabela 1</b>	Inibição da proliferação de células MCF-7 ( $IC_{50}$ ) do raloxifeno e de alguns derivados (Grese <i>et al.</i> , 1997a).....	26
<b>Tabela 2</b>	Estrutura química e valores de $pIC_{50}$ ( $-\log IC_{50}$ ) da série de 54 derivados do raloxifeno (Grese <i>et al.</i> , 1997a). Os composto sublinhados correspondem àqueles do conjunto de teste.....	37
<b>Tabela 3</b>	Classificação dos Elementos de Interação Farmacofórica (Hopfinger <i>et al.</i> , 1997).....	40
<b>Tabela 4</b>	Alinhamentos testados na construção dos modelos de QSAR-4D. A numeração do raloxifeno foi obtida automaticamente no programa Hyperchem e não é IUPAC.....	42
<b>Tabela 5</b>	Critérios para o pré-tratamento do banco de dados.....	45
<b>Tabela 6</b>	Valores de $R^2$ , $R^2_{ajustado}$ , $q^2$ e $q^2_{ajustado}$ dos modelos com 7 a 9 termos obtidos pelos alinhamentos 1 a 3 em células de 2 e 1Å. A numeração do raloxifeno foi obtida automaticamente no programa Hyperchem e não é IUPAC.....	52
<b>Tabela 7</b>	Melhores modelos do alinhamento 1 em células de 1Å.....	55
<b>Tabela 8</b>	Matriz de correlação cruzada entre os <i>GCODs</i> do modelo 1B9 obtido pelo método de QSAR-4D em célula de 1Å a partir do alinhamento 1.....	57
<b>Tabela 9</b>	Valores de atividade experimental ( $pIC_{50exp}$ ) e calculada ( $pIC_{50calc}$ ) e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) do modelo 1B9 de QSAR-4D para o conjunto de treinamento.....	64
<b>Tabela 10</b>	Valores de atividade experimental e calculada e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ), do modelo 1B9 de QSAR-4D para o conjunto de teste.....	68
<b>Tabela 11</b>	Melhores modelos do alinhamento 2 em células de 1Å.....	71
<b>Tabela 12</b>	Matriz de correlação dos resíduos das atividades dos modelos 2B7, 2B8 e 2B9 de QSAR-4D, obtidos com célula de 1Å.....	71
<b>Tabela 13</b>	Matriz de correlação cruzada entre os <i>GCODs</i> do modelo 2B9 obtido pelo método de QSAR-4D em célula de 1Å.....	74
<b>Tabela 14</b>	Valores de atividade experimental ( $pIC_{50exp}$ ) e calculada ( $pIC_{50calc}$ ) e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) do modelo 2B9 de QSAR-4D para o conjunto de treinamento.....	79
<b>Tabela 15</b>	Valores de atividade experimental e calculada e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) do modelo 2B9 de QSAR-4D para o conjunto de teste.....	84
<b>Tabela 16</b>	Novos compostos sugeridos para síntese e avaliação como <i>SERMs</i> .....	89
<b>Tabela 17</b>	Distância interatômicas entre os heteroátomos das hidroxilas fenólicas e do grupo amino da cadeia lateral da conformação do raloxifeno observada por cristalografia de raios-X (código do <i>PDB</i> : 1ERR) e dos compostos das séries A, B e C.....	91

# 1 INTRODUÇÃO

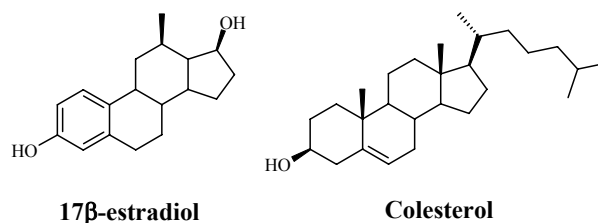
## 1.1 CÂNCER DE MAMA E ESTROGÊNIOS

O câncer é definido como o crescimento celular descontrolado. As células cancerosas podem invadir outros tecidos (metástase) do corpo, o que caracteriza o câncer invasivo. De acordo com a Organização Mundial de Saúde (*World Health Organization, WHO*), a cada ano, mais de 11 milhões de pessoas no mundo são diagnosticadas com câncer e as estimativas são de 16 milhões de novos casos por ano até 2020. O câncer causa 7 milhões de óbitos a cada ano no mundo, correspondendo a 12,5% do total de óbitos ([www.who.int/cancer/en/](http://www.who.int/cancer/en/)). Os tipos mais comuns de câncer entre mulheres do mundo inteiro são: o câncer de mama e o câncer cervical ([www.who.int/cancer/en/](http://www.who.int/cancer/en/)).

No Brasil, as estimativas para o ano de 2006 indicaram 472.050 novos casos de câncer. Os tipos mais incidentes, com exceção do câncer de pele não melanoma, são os cânceres de próstata e pulmão no sexo masculino e mama e colo do útero no sexo feminino, acompanhando o mesmo perfil da magnitude observada no mundo. Em 2006 foram estimados 237.480 casos novos para o sexo feminino, sendo os tumores de mama o segundo em incidência na população brasileira (estima-se cerca de 49 mil novos casos) ([www.inca.gov.br/estimativa/2006/](http://www.inca.gov.br/estimativa/2006/)).

Os estrogênios (e.g. 17 $\beta$ -estradiol, Figura 1) são hormônios sexuais femininos sintetizados a partir do colesterol (Figura 1) e pertencem à classe dos hormônios esteroidais (Lewis & Jordan, 2005). Em mulheres na idade fértil, os ovários são os principais responsáveis pela produção de estrogênios, enquanto que durante a gravidez, a placenta se torna a principal fonte de estrogênios. Na pós-menopausa, entretanto, a maior produção ocorre em regiões extraglandulares, como por exemplo, no tecido adiposo (Brueggemeier *et al.*, 1995).

Embora atuem predominantemente como hormônios sexuais, os estrogênios também estão envolvidos na regulação de diversos processos celulares não associados à função reprodutiva, em regiões como o sistema nervoso central e o sistema cardiovascular (McDonnell, 2000).



**Figura 1.** Estruturas químicas do 17β-estradiol e do colesterol, precursor dos hormônios esteroidais.

Os efeitos fisiológicos dos estrogênios são mediados por receptores nucleares específicos, denominados de receptores de estrogênio (*Estrogen Receptor, ER*). A interação desses hormônios com o *ER* ocorre numa região do receptor denominada de domínio de ligação do ligante (*Ligand-Binding Domain, LBD*), iniciando uma série de eventos moleculares que culminam na ativação ou repressão dos genes alvos (Brzozowski *et al.*, 1997).

Ensaio com culturas de células do tipo MCF-7, uma linhagem de células humanas de câncer de mama positiva para receptor de estrogênio, têm sido extensivamente utilizados na caracterização e diferenciação de agonistas (estimuladores do crescimento) e antagonistas (inibidores do crescimento) de receptores de estrogênio (Kraft *et al.*, 1999).

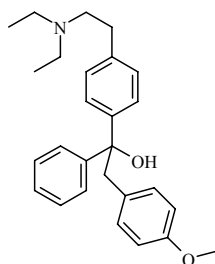
Terapeuticamente, os estrogênios têm como indicações mais comuns a anticoncepção em mulheres na idade fértil e a terapia de reposição hormonal em mulheres na pós-menopausa, exercendo neste último caso, a profilaxia da perda óssea e a atenuação dos sintomas vasomotores (Loose-Mitchell & Stancel, 2003). Entretanto, como um dos alvos dos

estrogênios é o tecido mamário, onde estimulam a proliferação celular, eles apresentam como efeito colateral o estímulo do crescimento do carcinoma mamário hormônio-dependente (Brueggemeier *et al.*, 1995). Além desse efeito, podem também provocar outros efeitos, tais como tromboembolismo, alteração do metabolismo de lipídios e carboidratos, hipertensão e doença da vesícula biliar (Loose-Mitchell & Stancel, 2003).

Assim, a necessidade de compostos que exercessem os efeitos benéficos dos estrogênios nos ossos e no coração, mas que atuassem como antagonistas nas mamas e no útero, evitando os efeitos indesejáveis dos estrogênios nestes tecidos, levou ao desenvolvimento de uma nova classe de fármacos, denominados moduladores seletivos de receptores de estrogênio (*Selective Estrogen Receptor Modulator, SERM*) (McDonnell, 2000; Lewis & Jordan, 2005). Este termo se refere à capacidade do fármaco em exercer efeitos diferenciados nos alvos regulados pelo estrogênio, *i.e.*, exercer ação agonista em alguns tecidos (*e.g.* ossos, fígado e sistema cardiovascular) e antagonista em outros (*e.g.* cérebro e mama), podendo ainda apresentar ação agonista/antagonista mista (*e.g.* útero).

## 1.2 MODULADORES SELETIVOS DE RECEPTORES DE ESTROGÊNIO

Em 1956, Lerner e colaboradores relataram as propriedades farmacológicas do primeiro antagonista puro de *ER*, um composto sintético não-esteroidal denominado etamoxitri-fetol (Figura 2) (Jordan, 2001). Contudo, não houve grande avanço no desenvolvimento clínico de um fármaco desprovido de atividade agonista porque os estrogênios são importantes na manutenção da densidade óssea e na diminuição do colesterol circulante. Um antagonista puro poderia prevenir o câncer de mama, mas levaria a uma pré-disposição da mulher à osteoporose e às doenças coronarianas.



**Etamoxitriphenol**

**Figura 2.** Estrutura química do etamoxitriphenol.

O desenvolvimento da primeira geração de *SERMs*, que inclui o tamoxifeno (Figura 3), forneceu uma oportunidade única para o estudo clínico e farmacológico de um agonista parcial de *ER* com atividade antagonista no tecido mamário (Jordan, 2001). Em 1973, no Reino Unido, e em 1977, nos Estados Unidos, o tamoxifeno foi aprovado para o tratamento do câncer de mama. Este fármaco, apesar de não provocar redução significativa de fraturas nos ossos, reduz a incidência de câncer de mama, apesar de poder aumentar a incidência de câncer no útero (Lewis & Jordan, 2005; Jordan, 2001; Park & Jordan, 2002). Atualmente, o tamoxifeno é o fármaco de primeira escolha para o tratamento de câncer de mama em estado avançado em mulheres na pós-menopausa (Park & Jordan, 2002).

No início dos anos 80, o raloxifeno (Figura 3), um *SERM* de segunda geração, foi desenvolvido e, em 1997, foi aprovado nos Estados Unidos como fármaco para o tratamento e prevenção da osteoporose em mulheres na pós-menopausa, sendo seguro para os tecidos das mamas e do útero (Grese *et al.*, 1997a). Este fármaco, além de manter a densidade óssea, reduz as fraturas no esqueleto e diminui o colesterol circulante (Jordan, 2001).

Em 1999, o Instituto Nacional do Câncer nos Estados Unidos iniciou um estudo clínico denominado *Study of Tamoxifen and Raloxifene, STAR*, com o objetivo de comparar a eficácia do tamoxifeno e do raloxifeno na redução da incidência de câncer de mama em mulheres na pós-menopausa com alto risco de desenvolverem esta doença. O estudo envolveu



19000 mulheres e foi concluído no segundo semestre de 2006 (Jordan, 1999; Pappas & Jordan, 2002). Além disso, o raloxifeno também está sendo avaliado como fármaco para prevenção de doenças coronarianas (Jordan, 2001).

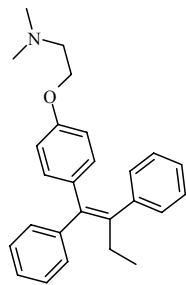
Os resultados iniciais do estudo *STAR* demonstraram que tanto o tamoxifeno quanto o raloxifeno reduzem o risco do surgimento de câncer de mama invasivo e apresentam menor risco de ocasionar trombose e catarata (Vogel *et al.*, 2006). O risco do surgimento de outros tipos de câncer, fraturas e outras doenças cardiovasculares é semelhante para ambos os fármacos. O estudo confirmou o benefício do raloxifeno na redução do câncer de mama invasivo e indicou este fármaco como sendo tão ativo quanto o tamoxifeno para tal finalidade. Contudo, diferente do tamoxifeno, o raloxifeno não reduz o risco do câncer de mama não-invasivo.

A compreensão das relações estrutura-atividade levou a modificações químicas na estrutura do tamoxifeno, tanto por alterações das cadeias laterais, originando novos análogos do tamoxifeno, tais como toremifeno, droloxifeno e idoxifeno (Figura 3), quanto pela alteração da estrutura do anel trifeniletileno não-esteroidal do tamoxifeno, gerando estruturas contendo “anéis rígidos” como o anel benzotiofeno, exemplificado pelos derivados raloxifeno e arzoxifeno, e o anel benzopirano e indol, exemplificado pelo derivado acolbifeno e pipendoxifeno respectivamente (Figura 3). Todas essas estruturas não-esteroidais são classificadas como *SERMs*, pois exibem atividade agonista/antagonista mista dependente do tecido (Johnston, 2005).

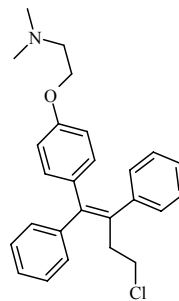
Apesar de todas as ações benéficas dos *SERMs* citados, esses fármacos ainda não são os moduladores seletivos ideais, pois não se pode prever quando o fármaco atuará como agonista ou antagonista no sistema nervoso central ou se os efeitos benéficos desses agentes nos lipídeos séricos levam a uma queda da mortalidade provocada por doenças cardiovasculares. Além disso, os *SERMs* disponíveis atualmente não podem ser utilizados no

tratamento de distúrbios vasomotores ou outra condição associada à depleção de estrogênio, podendo até mesmo exacerbar tais eventos (McDonnell, 2000).

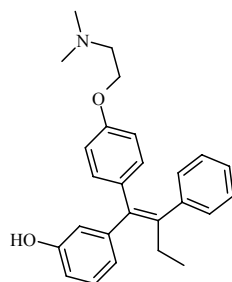
**(A) SERMs de 1ª Geração**



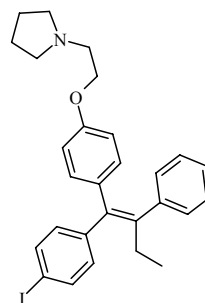
**Tamoxifeno**



**Toremifeno**

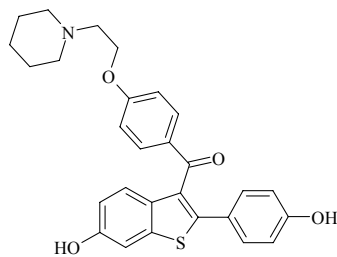


**Droloxifeno**

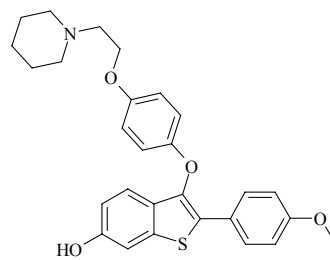


**Idoxifeno**

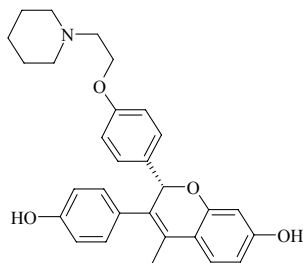
**(B) SERMs de 2ª Geração**



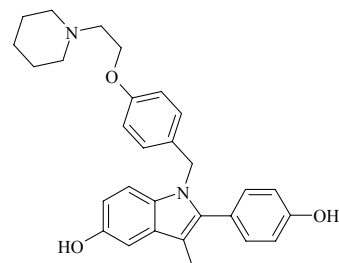
**Raloxifeno**



**Arzoxifeno**



**Acolbifeno**



**Pipendoxifeno**

**Figura 3.** Estruturas químicas de alguns *SERMs* (A) de primeira e (B) de segunda geração.

### 1.3 RECEPTORES DE ESTROGÊNIO

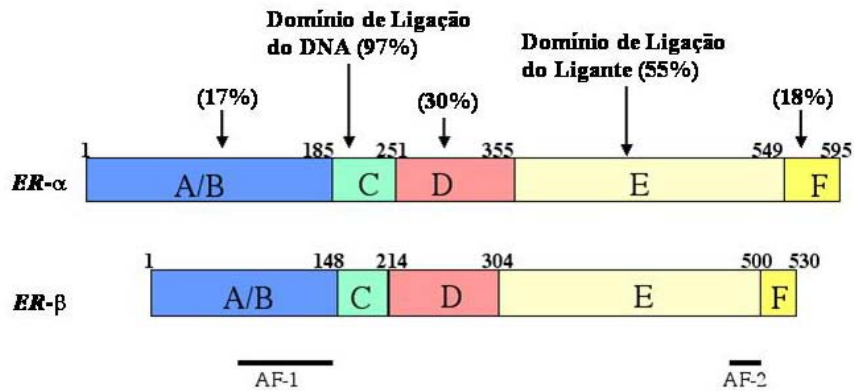
A descoberta do receptor de estrogênio (*Estrogen Receptor, ER*) se deu no início dos anos 60, e o subsequente conhecimento adquirido nos estudos farmacológicos foi aplicado na identificação de tumores de mama que responderiam à terapia endócrina (Jordan, 2001). Posteriormente, foi descoberto que os receptores de estrogênio existem em duas formas distintas, denominadas  $\alpha$  e  $\beta$  (*ER- $\alpha$*  e *ER- $\beta$* ) (Johnston, 2005).

Ambos os receptores são membros da superfamília dos receptores nucleares de fatores de transcrição induzidos pelo ligante. A distribuição dos *ER  $\alpha$*  e *ER  $\beta$*  nos tecidos é bastante diferenciada, implicando em funções fisiológicas distintas (Johnston, 2005). Alguns tecidos, como os ovários, expressam aproximadamente a mesma quantidade de *ER- $\alpha$*  e *ER- $\beta$* . Outros tecidos expressam mais o *ER- $\alpha$* , incluindo o útero, a hipófise e os testículos, enquanto a próstata e a bexiga apresentam maior expressão do *ER- $\beta$* . Os rins e o fígado expressam apenas o *ER- $\alpha$* , enquanto que o pulmão e o cérebro expressam apenas o *ER- $\beta$*  (Kuiper *et al.*, 1997).

Os receptores de estrogênio possuem uma arquitetura modular composta por quatro domínios: A/B, C, D e E/F. A região N-terminal do domínio A/B contém a função de ativação 1 de transcrição independente do ligante, denominada *AF-1 (Activation Function-1)*; o domínio C é a região de ligação do DNA; o domínio D corresponde a uma região de dobradiça (*hinge*); e o domínio E/F contém a função de ativação 2 de transcrição dependente do ligante, denominada como *AF-2 (Activation Function-2)*, onde o domínio E é a seqüência específica para o domínio de ligação do ligante (*Ligand Binding Domain, LBD*) (Nilsson & Koehler, 2005).

Uma comparação da seqüência de aminoácidos dos receptores de estrogênio  $\alpha$  e  $\beta$  (Klinge, 2000) indica que ambos apresentam o domínio de ligação do DNA (domínio C) praticamente idêntico (97% de homologia), reconhecendo as mesmas regiões regulatórias do

DNA. Contudo, no domínio de ligação do ligante (domínio E), a sequência diverge consideravelmente (55% de homologia) (Figura 4).



**Figura 4.** Comparação entre as seqüências de aminoácidos dos receptores de estrogênio  $\alpha$  e  $\beta$  (Adaptado de Klinge, 2000). Entre parênteses estão indicados o percentual de homologia entre as seqüências dos domínios A/B, C, D, E e F.

Apesar de todos os ligantes do *ER* se ligarem exclusivamente na região C-terminal do *LBD* (domínio E), ligantes com estruturas moleculares relativamente distintas em peso molecular, conformação molecular e propriedades físico-químicas são reconhecidos pelo *LBD* (Shiau *et al.*, 1998). Kuiper e colaboradores (1997) analisaram as afinidades de diversos tipos de ligantes pelos receptores de estrogênio  $\alpha$  e  $\beta$  e constataram que, embora o 17 $\beta$ -estradiol apresente afinidade equivalente por ambos os receptores, existem diferenças significativas entre agonistas e antagonistas de *ER*. Por exemplo, enquanto o tamoxifeno apresenta afinidade relativa semelhante para o *ER- $\alpha$*  e o *ER- $\beta$* , seu metabólito ativo, o 4-hidroxitamoxifeno apresenta afinidade consideravelmente maior para o *ER- $\beta$*  do que para o *ER- $\alpha$* . Por outro lado, o dietilestilbestrol, um estrogênio sintético não-esteroidal, possui maior afinidade para o *ER- $\alpha$*  do que para o *ER- $\beta$* .

Um pré-requisito para a modulação da expressão dos genes alvo pelos receptores de estrogênio é a interação desses receptores com outras proteínas, denominadas de co-fatores, que reconhecem elementos específicos de resposta ao estrogênio (*Estrogen Response Elements, ERE*) nas regiões promotoras dos genes alvo. A ativação da transcrição pelo *ER* é mediada por pelo menos duas funções de ativação (*AFs, Activation Functions*) distintas, *AF-1* na região N-terminal do domínio A/B e *AF-2* no *LBD*, como citado anteriormente. A atividade de *AF-1* é regulada pela via da proteína quinase ativada por mitógenos (*Mitogen-Activated Protein Kinase, MAP kinase*), envolvendo inúmeras proteínas que constituem uma cascata de transcrição de sinais ativada por estímulos diversos, como fatores de crescimento, enquanto que a atividade do *AF-2* é regulada pela ligação do ligante (Shiau *et al.*, 1998). O balanço relativo entre a quantidade de proteínas co-ativadoras e co-repressoras, num determinado tipo de célula, também influencia uma dada resposta do *ER* a um ligante em particular (Johnston, 2005).

#### 1.4 MECANISMO DE AÇÃO DE LIGANTES DE RECEPTORES DE ESTROGÊNIO

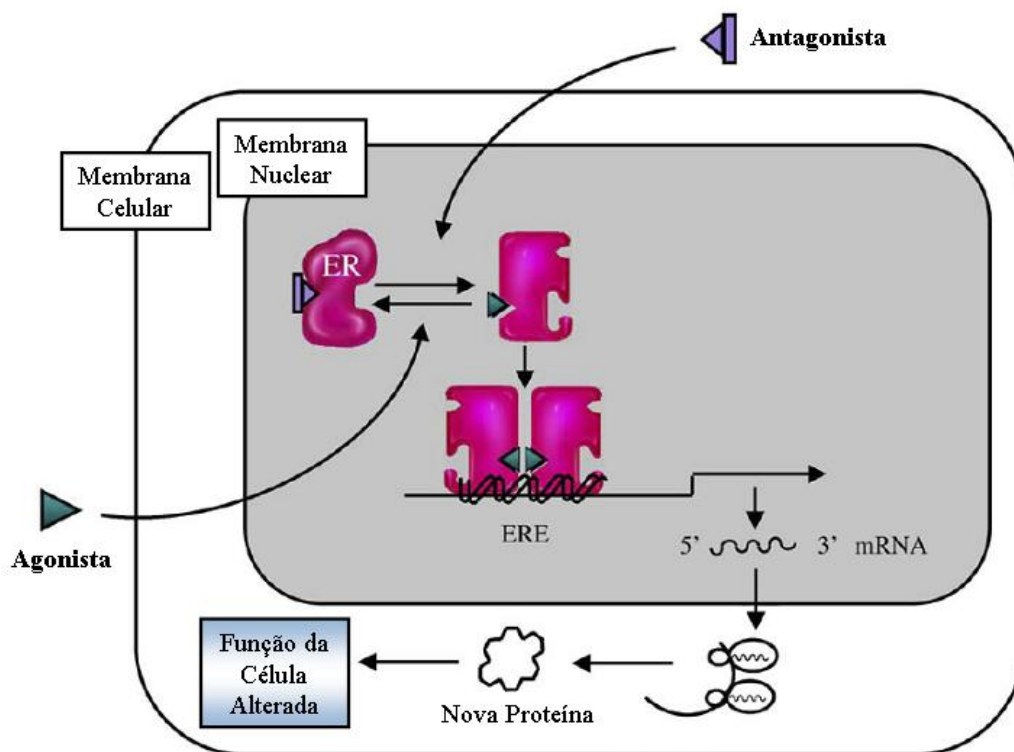
Os hormônios esteróides agem nas células alvo para regular a expressão gênica e a biossíntese de proteínas, via a formação de complexos esteróide-receptor (Brueggemeier *et al.*, 1995). No mecanismo de ação proposto, a atividade do complexo é controlada pelo encaixe de um estrogênio endógeno ao receptor de estrogênio no sítio de ligação do ligante (*LBD*), e que um antagonista bloqueia a ação do estrogênio endógeno devido a sua ligação no mesmo sítio de ligação (Jordan, 2001).

Os receptores de estrogênio (*ER- $\alpha$*  e *ER- $\beta$* ), localizados no núcleo das células alvo na forma de monômeros, são fatores de transcrição ativados pelo ligante. O estrogênio endógeno ou um agonista, após penetrar na célula por difusão passiva pela membrana plasmática, liga-

se ao *ER* promovendo uma alteração conformacional que resulta na dimerização do *ER*, aumentando a afinidade e taxa de ligação ao ácido desoxirribonucleico (*deoxyribonucleic acid, DNA*). O complexo agonista-*ER* interage com seqüências específicas do *DNA* responsivas ao estrogênio (*Estrogen Response Elements, ERE*), e com diversos fatores de transcrição nuclear. Essa ligação do complexo ao *DNA* inicia a transcrição da seqüência de *DNA* para a produção do ácido ribonucleico mensageiro (*mRNA, messenger ribonucleic acid*). A elevação do nível de *mRNA* leva a um aumento da síntese de proteínas no retículo endoplasmático, que resulta na regulação do crescimento, diferenciação e função celular (Brueggemeier *et al.*, 1995) (Figura 5).

O complexo ligante-*ER* também atua em conjunto com diversas proteínas que influenciam o acesso à maquinaria genética nuclear. Estes fatores, que facilitam a ativação gênica com o complexo ligante-*ER*, são denominados co-ativadores. Ao contrário, outras proteínas semelhantes, conhecidas como co-repressoras, atuam ao prevenir a transcrição através da ligação do *ER* (Bryant, 2001).

A ligação de um antagonista ao *ER* também promove a dimerização do receptor e ligação ao *DNA*. Entretanto, a ligação do antagonista induz a uma conformação diferente daquela observada quando o receptor se liga a um agonista. A conformação do *ER* induzida pelo antagonista facilita a ligação de proteínas co-repressoras, ao invés de proteínas co-ativadoras (Loose-Mitchell & Stancel, 2003) (Figura 5).



**Figura 5.** Representação esquemática do mecanismo de ação de ligantes (agonistas e antagonistas) do receptor de estrogênio (Adaptado de McDonnel, 2004).

Em 1997, Brzozowski e colaboradores relataram a estrutura tridimensional da região do *LBD* do receptor de estrogênio em complexo com o estrogênio endógeno (o hormônio  $17\beta$ -estradiol) obtida por cristalografia de raios-X e depositadas no *Protein Data Bank (PDB)* (Berman *et al.*, 2000), um banco de dados de proteínas com acesso livre *on line* ([www.rcsb.org](http://www.rcsb.org)). A estrutura 3D do domínio de ligação do ligante (*LBD*) do *ER* (hélices H3-H12) é semelhante àquelas observadas para outros receptores nucleares. O *LBD* é formado por um sanduíche  $\alpha$ -hélice de três camadas antiparalelas contendo uma camada central do núcleo de três hélices (H5/6, H9 e H10) colocada entre duas camadas adicionais de hélices (H1-4 e H7, H8, H11). Este arranjo das hélices torna possível a formação de uma cavidade para a ligação do ligante. Elementos estruturais secundários, compostos por uma estrutura de



duas folhas  $\beta$  antiparalelas unidas por uma alça (S1 e S2) e uma  $\alpha$ -hélice (H12), estão localizados nesta subunidade ao lado das três camadas principais (Brzozowski *et al.*, 1997).

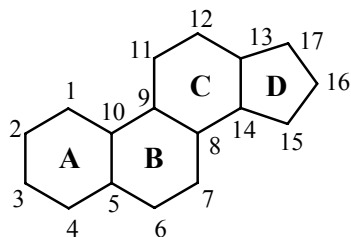
A hélice H12, posicionada sobre o *LBD* quando o receptor está ligado ao estrogênio, é crítica para o recrutamento de co-ativadores para o sítio do *AF-2* e subsequente início da atividade da RNA polimerase. O reposicionamento da hélice H12, após a ligação do ligante, foi proposto como uma característica que diferencia as ações agonista e antagonista do ligante sobre o *ER- $\alpha$*  (Brzozowski *et al.*, 1997). O raloxifeno, ao se ligar no sítio hidrofóbico do *LBD*, impede a reorientação da H12 devido à presença de uma cadeia lateral ausente no estradiol, impedindo, conseqüentemente, a ligação do receptor aos co-ativadores (Lewis & Jordan, 2005).

O *ER* é único entre os receptores esteróides na habilidade em reconhecer uma grande variedade de compostos não-esteroidais. Esta promiscuidade pode ser atribuída ao tamanho da cavidade do *LBD* que apresenta um volume disponível quase que duas vezes maior do que o volume molecular do  $17\beta$ -estradiol (Brzozowski *et al.*, 1997).

## 1.5 RELAÇÃO ENTRE ESTRUTURA E ATIVIDADE DOS LIGANTES DE RECEPTORES DE ESTROGÊNIO

### 1.5.1 Estrogênios Naturais

Os esteróides são estruturalmente relacionados ao hidrocarboneto ciclopentaperhidrofenantreno, que contém um anel do cinco membros (D) e um conjunto de três anéis (A, B e C) relacionado ao fenantreno (Figura 6). Os hormônios esteroidais mais importantes possuem uma ligação dupla nas posições 4 e 5 ou 5 e 6 (Brueggemeier *et al.*, 1995) entre outras características.



**Figura 6.** Estrutura comum dos hormônios esteroidais, representada pelo composto hipotético ciclopentaperhidrofenantreno. A numeração é IUPAC.

Os estrogênios endógenos ou naturais (*e.g.* 17 $\beta$ -estradiol, estrona e estriol, Figura 7) são hormônios esteroidais contendo 18 átomos de carbono, uma hidroxila na posição 3 do anel A, que é aromático, e uma hidroxila (em orientação  $\beta$ ) ou uma carbonila de cetona na posição 17 do anel D (Figura 7) (Loose-Mitchell & Stancel, 2003).



**Figura 7.** Estrutura química dos estrogênios endógenos 17 $\beta$ -estradiol, estrona e estriol.

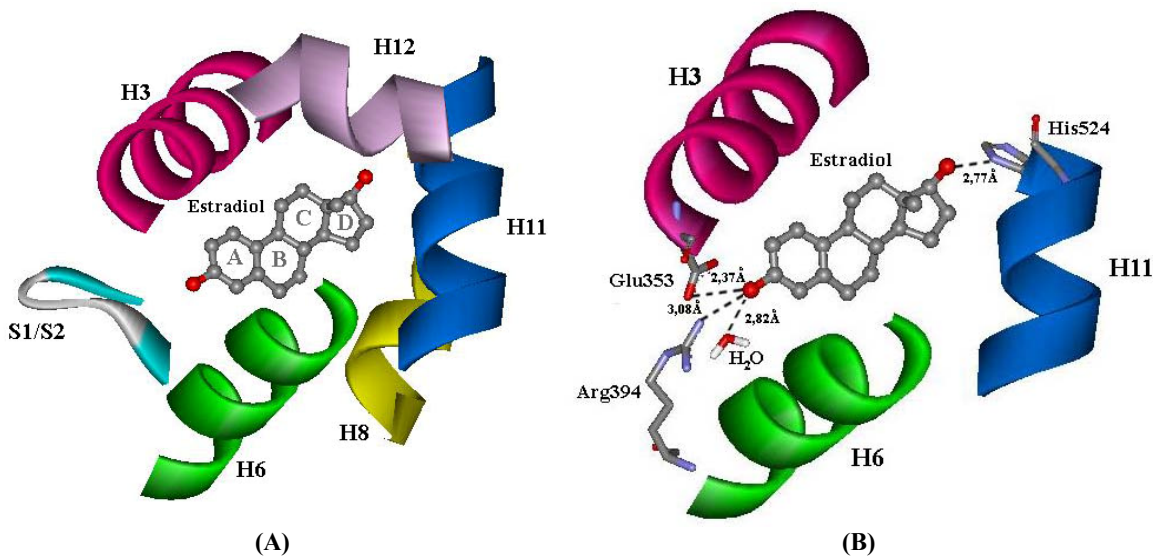
O anel fenólico A é o componente estrutural principal dos estrogênios, sendo responsável pela alta afinidade de ligação aos receptores de estrogênio. Em geral, substituintes alquila no anel fenólico A dificultam a interação estrogênio-receptor, enquanto que substituintes no anel C ou D podem ser tolerados (Loose-Mitchell & Stancel, 2003).

Dentre os estrogênios endógenos, o estradiol é o mais potente, seguido pela estrona e pelo estriol, quando administrados via subcutânea. Quando a administração é feita por via oral, o estriol é o mais potente, seguido pelo estradiol e pela estrona. Isto ocorre porque o 17 $\beta$ -estradiol e a estrona são metabolizados, principalmente no fígado, sendo interconvertidos enzimaticamente pela estradiol desidrogenase. Além disso, tanto o 17 $\beta$ -estradiol quanto a

estrona são convertidos em estriol pelo complexo enzimático do citocromo P450 (Brueggemeier *et al.*, 1995).

Segundo Brzozowski e colaboradores (Brzozowski *et al.*, 1997), a cavidade do sítio de ligação do 17 $\beta$ -estradiol está separada do ambiente externo e corresponde a uma grande parte da região hidrofóbica do *LBD* do *ER*. Esta região hidrofóbica é composta por resíduos das  $\alpha$ -hélices H3 (Met342 a Leu354), H6 (Trp383 a Arg394), H8 e a alça adjacente (Val418 a Leu428), H11 (Met517 a Met528), H12 (Leu539 a His547) e a curva S1/S2 (Leu402 a Leu410) (Figura 8).

Um aspecto crucial na interação estradiol-*ER* é o envolvimento completo do hormônio pela região hidrofóbica no domínio de ligação do ligante, devido à localização estratégica da hélice H12 do *ER* (Johnston, 2005). No complexo estradiol-*ER*, a H12 se acomoda sobre o sítio de ligação e é acondicionada entre as  $\alpha$ -hélices H3, H5/6 e H11. Embora a hélice H12 não faça contato direto com o estradiol, ela funciona como uma “tampa” do sítio de ligação, expondo sua superfície hidrofóbica interna para o hormônio ligado (Figura 8A). Esta disposição específica da hélice H12, observada em todas as estruturas conhecidas de ligantes ligados ao *LBD*, parece ser um pré-requisito para a ativação da transcrição gênica, gerando uma região *AF-2* capaz de interagir com as proteínas co-ativadores (Brzozowski *et al.*, 1997).



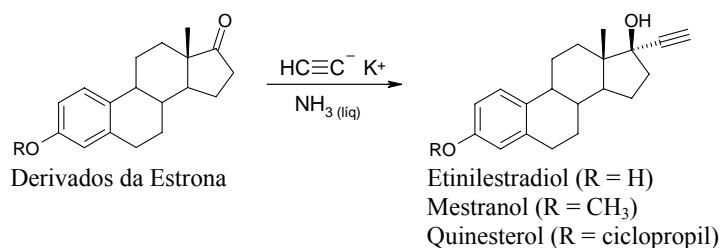
**Figura 8.** Representação da estrutura tridimensional do *LBD* do *ER- $\alpha$*  complexado ao 17 $\beta$ -estradiol obtida por cristalografia de raios-X (código *PDB*: 1ERE): (A) principais estruturas secundárias ao redor do ligante; (B) interações por ligação hidrogênio do ligante com o *LBD*.

O reconhecimento do hormônio pelo *LBD* é obtido por um conjunto de interações específicas por ligação hidrogênio e pela complementaridade do sítio hidrofóbico do *LBD* ao caráter apolar do estradiol (Figuras 8A e 8B). O estradiol se liga na cavidade formada entre as hélices H11, H3 e H6, numa conformação de baixa energia. A hidroxila fenólica do anel A localiza-se entre as hélices H3 e H6 e realiza interações por ligação hidrogênio direta com os resíduos Glu353 e Arg394 e uma molécula de água. A hidroxila 17 $\beta$  do anel D interage por ligação hidrogênio com o resíduo His524 da hélice H11.

O anel D realiza contatos inespecíficos com os resíduos Ile424, Gly521 e Leu525. A combinação de interações específicas polares e inespecíficas contribui para a habilidade do *ER* em reconhecer seletivamente os ligantes e ligar o estradiol com afinidade subnanomolar em relação a uma grande variedade de esteróides endógenos.

### 1.5.2 Estrogênios Semi-Sintéticos

Os estrogênios naturais são pouco potentes quando administrados por via oral, pois são facilmente degradados por microorganismos encontrados no trato gastrintestinal, além de sofrerem rápida metabolização hepática. Um método para aumentar a potência do 17 $\beta$ -estradiol é a proteção da hidroxila da posição 17 $\beta$  (Brueggemeier *et al.*, 1995). Por exemplo, o tratamento da estrona (ou derivados) com acetiletos de potássio em amônia líquida resulta em compostos com substituinte  $\alpha$ -etnil na posição C<sub>17</sub> (Figura 9), que apresentam maior biodisponibilidade oral devido à menor metabolização hepática por mecanismo de primeira passagem hepático (Brueggemeier *et al.*, 1995; Loose-Mitchell & Stancel, 2003).

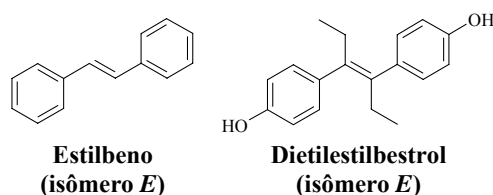


**Figura 9.** Tratamento de derivados da estrona com acetileto de potássio em amônia líquida para produzir compostos mais potentes quando administrados por via oral (Loose-Mitchell & Stancel, 2003).

A inserção de grupos hidroxila nas posições 6, 7 ou 11 reduz a potência dos compostos esteroidais endógenos. A introdução de insaturação no anel B e o aumento do número de átomos de carbono do anel D também reduzem a potência dos compostos. No caso do estradiol, a remoção da hidroxila da posição 3 ou 17 ou a epimerização da hidroxila da posição 17 $\beta$  para a configuração  $\alpha$  resulta em derivados menos potentes. (Brueggemeier *et al.*, 1995).

### 1.5.3 Estrogênios Não-Esteroidais

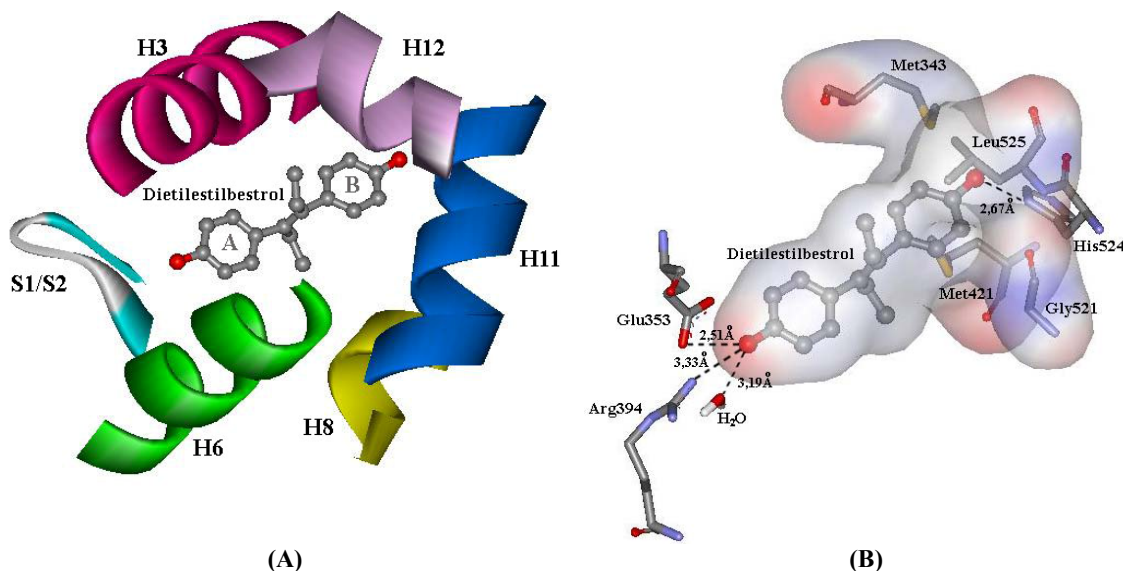
O núcleo esteróide não é essencial para a atividade agonista sobre os receptores de estrogênio. Muitos derivados do estilbeno, como por exemplo o dietilestilbestrol (Figura 10), um composto sintético não-esteroidal, são potentes agonistas do receptor de estrogênio, sendo utilizados regularmente na terapêutica.



**Figura 10.** Estruturas químicas do núcleo estilbeno e do dietilestilbestrol, um composto sintético não-esteroidal com potente atividade agonista dos receptores de estrogênio.

O isômero *E* do dietilestilbestrol (Figura 5) tem atividade agonista dez vezes maior do que o isômero *Z*, sendo também mais semelhante estruturalmente ao estradiol endógeno do que o isômero *Z* (Brueggemeier *et al.*, 1995). O *E*-dietilestilbestrol é tão potente quanto o estradiol na maioria dos ensaios farmacológicos, porém é mais potente por via oral e tem tempo de meia-vida mais longo no organismo. Esse fármaco, em particular, não é mais amplamente utilizado na clínica médica, mas apresenta grande importância histórica, devido ao seu uso como estrogênio oral ativo e de baixo custo, o que representou um marco no desenvolvimento de fármacos eficazes para tratamentos endócrinos (Loose-Mitchell & Stancel, 2003).

Shiau e colaboradores (Shiau *et al.*, 1998) analisando-se a estrutura tridimensional do dietilestilbestrol complexado ao *LBD* do *ER- $\alpha$* , depositada no banco de proteínas *PDB* (Berman *et al.*, 2000), perceberam que, assim como estradiol, o dietilestilbestrol interage dentro do *LBD* numa cavidade predominantemente hidrofóbica composta dos resíduos das hélices H3, H6, H7, H8, H11 e H12 e a curva S1/S2 (Figura 11).



**Figura 11.** Representação da estrutura tridimensional da região do *LBD* do *ER- $\alpha$*  complexada ao dietilestilbestrol obtida por cristalografia de raios-X (código *PDB*: 3ERD): **(A)** principais estruturas secundárias ao redor do ligante; **(B)** interações por ligação hidrogênio do ligante com os resíduos Glu353, Arg394 e His524 e uma molécula de água e interações de van der Waals do ligante com os resíduos Met343, Leu525, His524, Gly521 e Met421.

O dietilestilbestrol se liga ao *LBD* do *ER- $\alpha$*  (Figura 11) de modo semelhante aquele observado para o  $17\beta$ -estradiol (Figura 8A). Um dos anéis fenólicos do dietilestilbestrol, denominado anel A em analogia ao anel A do estradiol, localiza-se próximo às  $\alpha$ -hélices H3 e H6, ficando acomodado pelas cadeias laterais dos resíduos Phe404, Ala350, Leu387 e Leu391, enquanto que a hidroxila fenólica forma três interações por ligação hidrogênio com os resíduos Glu353 e Arg394 e uma molécula de água. O outro anel fenólico do dietilestilbestrol, o anel B, está ligado próximo às hélices H7, H8 e H11 adjacente ao local dos anéis C e D do estradiol. Este anel forma uma interação por ligação hidrogênio com o resíduo de His524, além de contatos de van der Waals, não somente com os resíduos Gly521 e Leu525, tal como o anel D de estradiol, mas também com os resíduos Met343, Leu346 e Met421 (Figura 11).

O dietilestilbestrol apresenta interações dentro do *LBD* que o estradiol não faz, por exemplo, os dois grupos etila do dietilestilbestrol se encaixam em cavidades não ocupadas

pelo estradiol, adjacentes à face  $\alpha$  do anel B e à face  $\beta$  do anel C do estradiol, resultando em contatos apolares adicionais com as cadeias laterais de Ala350, Leu384, Phe404 e Leu428, contribuindo para a alta afinidade do dietilestilbestrol pelo receptor.

O *ER* é capaz de utilizar os mesmos resíduos para formar todas as interações por ligação hidrogênio e de van der Waals observada com os diversos agonistas com os quais se liga. Esta capacidade de adaptação é, provavelmente, o resultado do grande volume molecular do sítio de ligação e de sua aparente plasticidade estrutural (Shiau *et al.*, 1998).

#### **1.5.4 Moduladores Seletivos dos Receptores de Estrogênio**

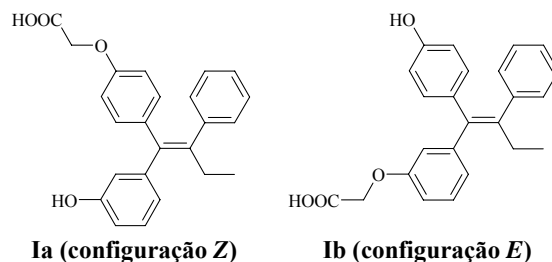
##### 1.5.4.1 Tamoxifeno e 4-hidroxitamoxifeno

O tamoxifeno, um *SERM* de primeira geração, é um antagonista não-esteroidal do estrogênio que é efetivo no tratamento e prevenção do câncer de mama. Embora preocupações tenham surgido em virtude do aumento de seus efeitos adversos em tecidos saudáveis, os riscos parecem ser pequenos em relação ao benefício substancial na redução dos eventos relatados do câncer de mama (Johnston, 2005).

Em analogia ao dietilestilbestrol, a disposição relativa dos anéis aromáticos em torno da ligação dupla na estrutura do tamoxifeno parece ser determinante para o seu perfil farmacológico. Kraft e colaboradores (1999) realizaram estudos de afinidade de ligação relativa pela competição com o hormônio  $17\beta$ -estradiol (*Relative Binding Affinity*, *RBA*) e de inibição da proliferação de células MCF-7 estimuladas pelo mesmo hormônio ( $IC_{50}$ ) com os análogos Ia e Ib (Figura 12) do tamoxifeno. Os autores observaram que apesar do composto Ia ter apresentado valores modestos de potência ( $RBA = 0,4$ ;  $IC_{50} = 1040$  nM), o composto Ib foi inativo ( $RBA < 0,05$ ;  $IC_{50} > 10000$  nM). Sugere-se que esta diferença seja consequência de uma preferência do *ER* por triariletilenos com configuração *Z* (caso do composto Ia) em

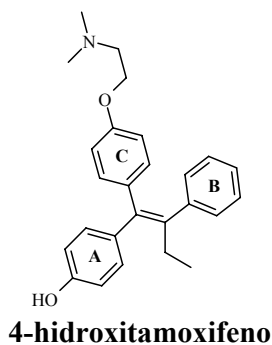


analogia ao isômero *Z* do dietilestilbestrol, em detrimento da configuração *E* (caso do composto Ib).



**Figura 12.** Estruturas químicas dos análogos Ia e Ib do tamoxifeno (Kraft *et al.*, 1999).

O metabolismo do tamoxifeno é complexo e envolve reações de *N*-desmetilação, hidroxilação aromática,  $\alpha$ -hidroxilação da cadeia lateral e *N*-oxidação, com a participação das enzimas do sistema citocromo P450. Embora o tamoxifeno seja o fármaco mais utilizado contra o câncer, sua dose ótima é desconhecida. Suas ações e efeitos colaterais podem ser, em parte, devido aos seus metabólitos ativos (Monteiro *et al.*, 2003), principalmente o 4-hidroxitamoxifeno (Figura 13), o metabólito mais potente em humanos, que apresenta afinidade 25-50 vezes maior pelo receptor de estrogênio do que o tamoxifeno (Loose-Mitchell & Stancel, 2003).



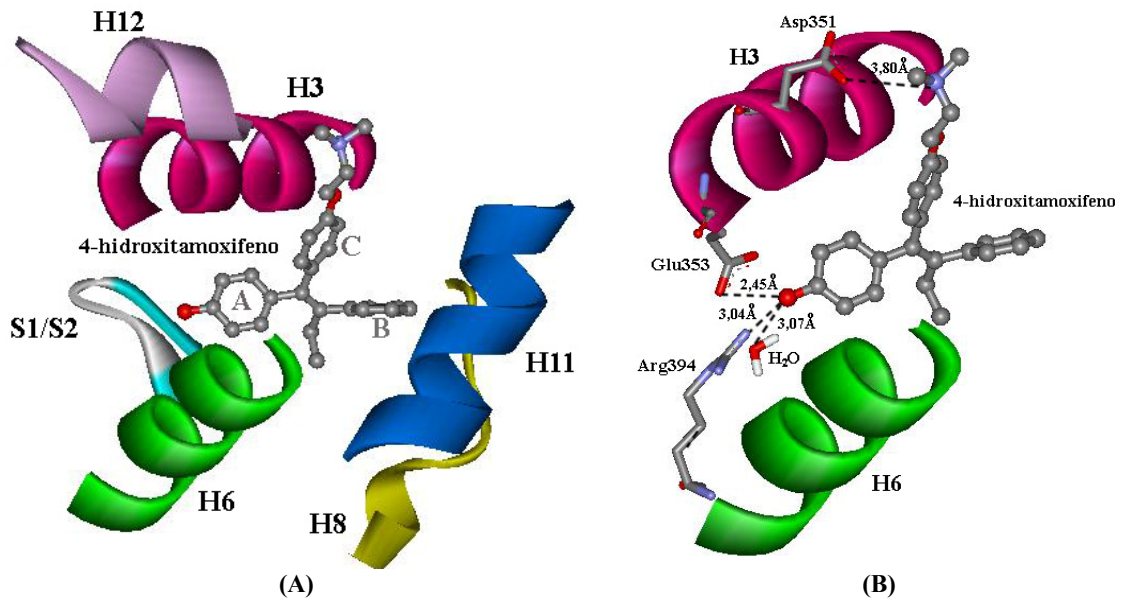
**Figura 13.** Estrutura química do 4-hidroxitamoxifeno.

Estudos de cristalografia de raios-X do complexo do 4-hidroxitamoxifeno com o ER- $\alpha$  (Shiau *et al.*, 1998) demonstram que o anel A do 4-hidroxitamoxifeno está ligado, aproximadamente, no mesmo local que o anel A do dietilestilbestrol se liga, próximo às hélices H3 e H6, com a hidroxila fenólica mantendo as mesmas três interações por ligação hidrogênio com as cadeias laterais dos resíduos Glu353 e Arg394 e com uma molécula de água (Figura 14A).

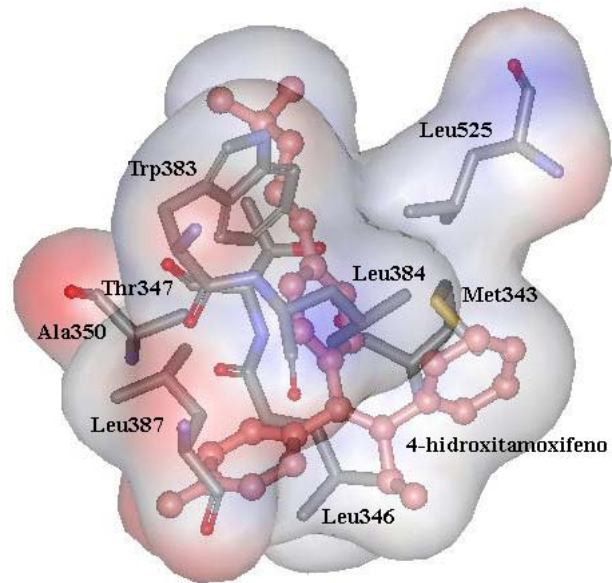
O anel C do 4-hidroxitamoxifeno faz interações de van der Waals com as cadeias laterais dos resíduos Met343, Leu346, Thr347, Ala350, Trp383, Leu384, Leu387 e Leu525. O grupo dimetilaminoetóxi faz interações de van der Waals com os resíduos Thr347 e Trp383 (Figura 15) e interação iônica com o grupamento carboxilato do resíduo de Asp351 (Figura 14B).

A ligação do 4-hidroxitamoxifeno ao ER- $\alpha$  promove uma alteração na conformação do *LBD* em relação àquela observada para o estradiol (Shiau *et al.*, 1998). Esta alteração conformacional envolve, principalmente, a reorientação da hélice 12 que é impedida pelo 4-hidroxitamoxifeno de se posicionar sobre o sítio de ligação. Dessa forma, o arranjo espacial final dos resíduos do sítio de ligação estabiliza a conformação do *LBD* complexado ao 4-hidroxitamoxifeno, permitindo que a hélice H12 alcance a região da função de ativação 2 (*AF-2*), prevenindo sua interação com o co-ativador.

Embora o 4-hidroxitamoxifeno não interaja diretamente com nenhum resíduo da hélice H12, sugere-se que seus efeitos sobre o posicionamento da hélice H12 contribuem para o antagonismo deste composto sobre o ER- $\alpha$  (Shiau *et al.*, 1998).



**Figura 14.** Representação da estrutura tridimensional do complexo formado entre o 4-hidroxitamoxifeno e a região do *LBD* do *ER-α* obtida por cristalografia de raios-x (código *PDB*: 3ERT): (A) principais estruturas secundárias ao redor do ligante; (B) interações por ligação hidrogênio entre o ligante e o *LBD*.

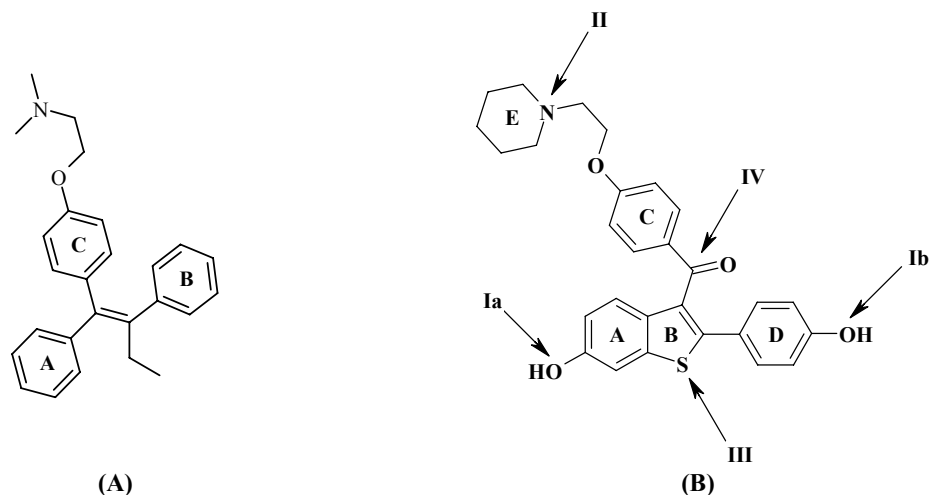


**Figura 15.** Representação das interações de van der Waals entre o 4-hidroxitamoxifeno e alguns resíduos do *ER-α* observadas por cristalografia de raios-X (código *PDB*: 3ERT).

#### 1.5.4.2 Raloxifeno

O raloxifeno é um antagonista do receptor de estrogênio que age especificamente contra os efeitos mitogênicos do estradiol em tecidos reprodutivos, enquanto mantém efeitos estrogênicos benéficos em outros tecidos (Brzozowski *et al.*, 1997). O raloxifeno tem ação antagonista no *ER* em tecidos mamários e uterinos, enquanto apresenta efeitos agonistas nos ossos e sistema cardiovascular. Devido a este perfil de especificidade tecidual, o raloxifeno é considerado um *SERM*. O raloxifeno foi aprovado pelo *FDA* (*Food and Drug Administration*) para o tratamento e prevenção da osteoporose em mulheres na pós-menopausa (Grese *et al.*, 1997a) e atualmente tem sido avaliado para o tratamento e prevenção do câncer de mama (Jordan, 1999; Pappas & Jordan, 2002).

Grese e colaboradores (1997b) relataram as características moleculares do raloxifeno, em relação ao tamoxifeno, responsáveis pelo aumento da seletividade tecidual e potência (Figura 16). Alterações nesses elementos estruturais afetam o perfil da atividade biológica descrito para o raloxifeno. Quatro características são marcantes no raloxifeno quando comparado ao tamoxifeno: (I) a presença das duas hidroxilas fenólicas ausentes no tamoxifeno, (II) a presença de um grupo amina de caráter básico também presente na extremidade da cadeia lateral do tamoxifeno, (III) a presença do grupo 2-fenil-benzotiofeno que incorpora a estrutura do estilbeno (isômero *E*), presente no tamoxifeno, porém com maior restrição conformacional, e (IV) a inserção de um grupo carbonila no início da cadeia lateral, conferindo maior liberdade conformacional à cadeia lateral.



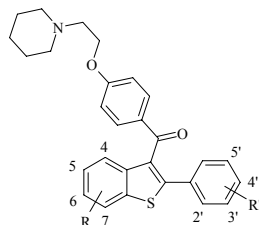
**Figura 16.** Estruturas químicas do (A) tamoxifeno e do (B) raloxifeno. As setas indicam as características marcantes no raloxifeno quando comparado ao tamoxifeno (Adaptado de Grese *et al.*, 1997b).

Em estudos de correlação estrutura-atividade do raloxifeno e de seus derivados, Grese e colaboradores (Grese *et al.*, 1997a) evidenciaram os efeitos dos substituintes do anel benzotiofeno e da fenila por ensaios de afinidade de ligação relativa pela competição com o hormônio  $17\beta$ -estradiol (*Relative Binding Affinity, RBA*) e por ensaios de inibição da proliferação de células MCF-7 estimuladas pelo mesmo hormônio ( $IC_{50}$ ).

Foi demonstrada a importância das duas hidroxilas fenólicas do raloxifeno para a ligação ao receptor. Os substituintes 6-OH e 4'-OH, na estrutura do raloxifeno, parecem mimetizar os substituintes 3-OH e  $17\beta$ -OH, respectivamente, na estrutura do estradiol. A relevância do grupo hidroxila na posição 6 é demonstrada pela depleção desse grupo (composto II, Tabela 1) ou pela sua substituição pelo grupo metoxila (composto III, Tabela 1). O grupo hidroxila da posição 4' parece ter uma importância menor, visto que sua eliminação diminui, de forma menos intensa, a potência do derivado obtido (composto IV, Tabela 1). No derivado 5-OH (composto V, Tabela 1), é interessante notar que, embora este composto se ligue com relativa afinidade ao ER ( $RBA= 0,1$ ), esta interação não produz efeitos antagonistas relevantes na proliferação de células MCF-7 ( $IC_{50}= 100$  nM). Isto indica que o grupo 6-OH é o componente molecular chave na ligação ao ER e que esta interação é muito importante na

regulação dos efeitos antiproliferativos desses compostos. A importância deste grupo também é evidenciada por estudos com o tamoxifeno e o 4-hidroxitamoxifeno, que é o metabólito mais potente (Grese *et al.*, 1997b).

**Tabela 1.** Inibição da proliferação de células MCF-7 (IC<sub>50</sub>) do raloxifeno e de alguns derivados (Grese *et al.*, 1997a).

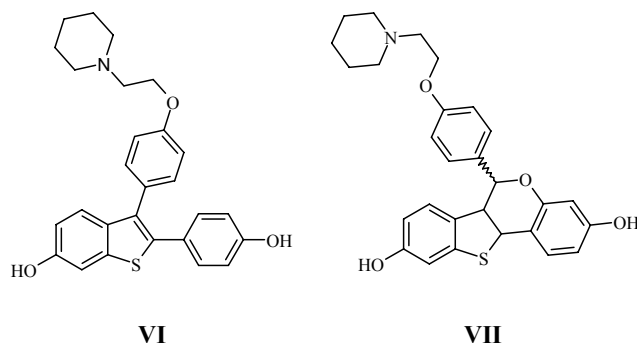


Composto	R	R'	IC <sub>50</sub> (nM)
Raloxifeno	6-OH	4'-OH	0,2
II	6-H	4'-OH	35
III	6-OMe	4'-OH	250
IV	6-OH	4'-H	2,5
V	5-OH	4'-OH	100

Estudos com substituintes nos anéis aromáticos do grupo trifeniletileno, como aqueles realizados para o tamoxifeno (Robertson *et al.*, 1982; Grese *et al.*, 1997b), demonstraram a importância do grupo amino básico terminal. Pesquisas com o raloxifeno e seus derivados também indicaram que o grupo dialquilamino é crítico para a atividade antagonista sobre o *ER*, visto que a troca do átomo de nitrogênio por um átomo de carbono ou por um átomo de nitrogênio não-básico resulta na perda total da atividade antagonista dos compostos (Grese *et al.*, 1997b).

Outra distinção entre raloxifeno e tamoxifeno é a presença do grupo carbonila entre o estilbeno e a cadeia lateral básica do raloxifeno. Grese e colaboradores (Grese *et al.*, 1997b) propuseram que a orientação da cadeia lateral no tamoxifeno e no composto VI (Figura 17), que se encontra no mesmo plano do estilbeno em ambos os casos, contribui para a estimulação de células do útero, enquanto que no raloxifeno e no composto VII (Figura 17),

onde a cadeia lateral está fora do plano do estilbeno, produzem pouca ou nenhuma estimulação. Os efeitos biológicos diferenciados indicam que a orientação da cadeia lateral básica apresenta um importante papel no perfil farmacológico do raloxifeno e do composto VII. Entretanto, esta diferença também pode ser devida às diferentes distâncias interatômicas observadas entre os grupos hidroxila e amina.



**Figura 17.** Estruturas químicas dos análogos VI e VII do raloxifeno.

Em 1997, a estrutura tridimensional do raloxifeno complexado ao *LBD* do *ER- $\alpha$*  obtida por cristalografia de raios-X, disponível sob o código 1ERR no banco de dados *PDB* (Berman *et al.*, 2000), evidenciou a possível conformação ativa do fármaco e suas interações com a proteína (Brzozowski *et al.*, 1997). O raloxifeno se liga ao mesmo sítio de ligação do estradiol dentro do *LBD*, onde a hidroxila fenólica na posição 6 do benzotiofeno mimetiza a hidroxila fenólica da posição 3 do anel A do estradiol, visto que ambas estão no sítio polar posicionados entre as hélices H3 e H6 (Figura 18A).

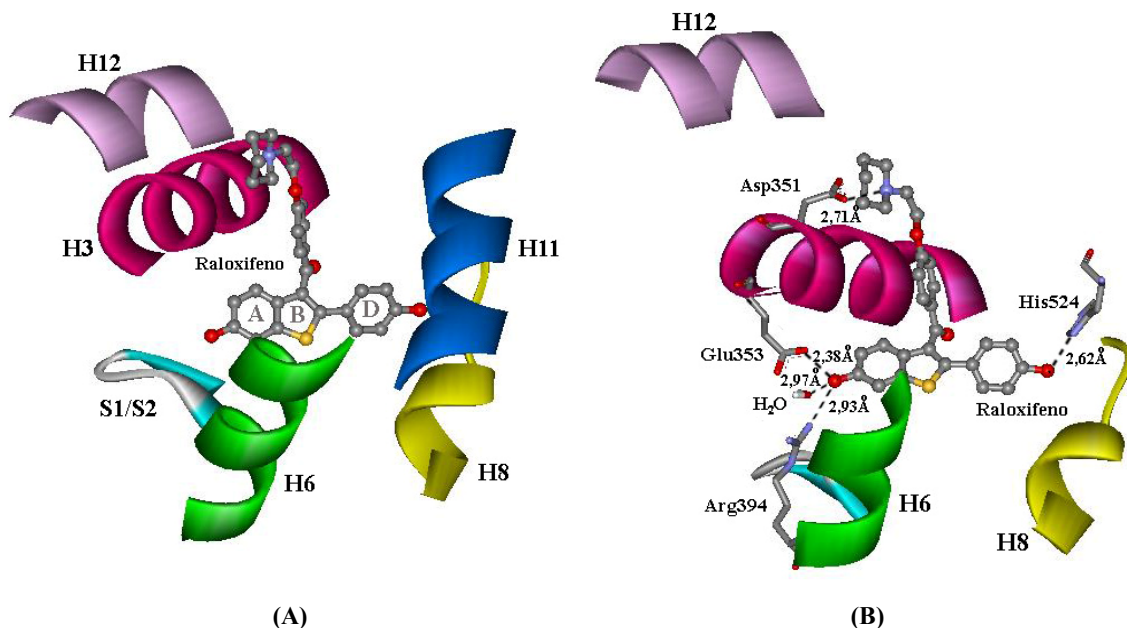
O anel D do raloxifeno está envolvido principalmente por contatos apolares semelhantes àqueles observados para o anel D do estradiol. A cadeia lateral do raloxifeno apresenta intensos contatos hidrofóbicos com as hélices H3 e H5/6, H11 e a alça entre H11 e H12 (Figura 18B). O raloxifeno também está ancorado à proteína por interações por ligação hidrogênio direta entre: o nitrogênio protonado do anel piperidina (anel E) e o resíduo

Asp351; a hidroxila fenólica do anel A e os resíduos Glu353 e Arg394 e uma molécula de água (Figura 18B). Além disso, a cadeia lateral do resíduo His524 sofre uma rotação para compensar a orientação diferente do grupo 4'-hidroxila do anel D do raloxifeno (Figura 19B), em comparação ao grupo 17 $\beta$ -hidroxila do anel D do estradiol (Figura 19A), o que permite uma orientação adequada para a manutenção da interação por ligação hidrogênio com o anel imidazólico do resíduo His524.

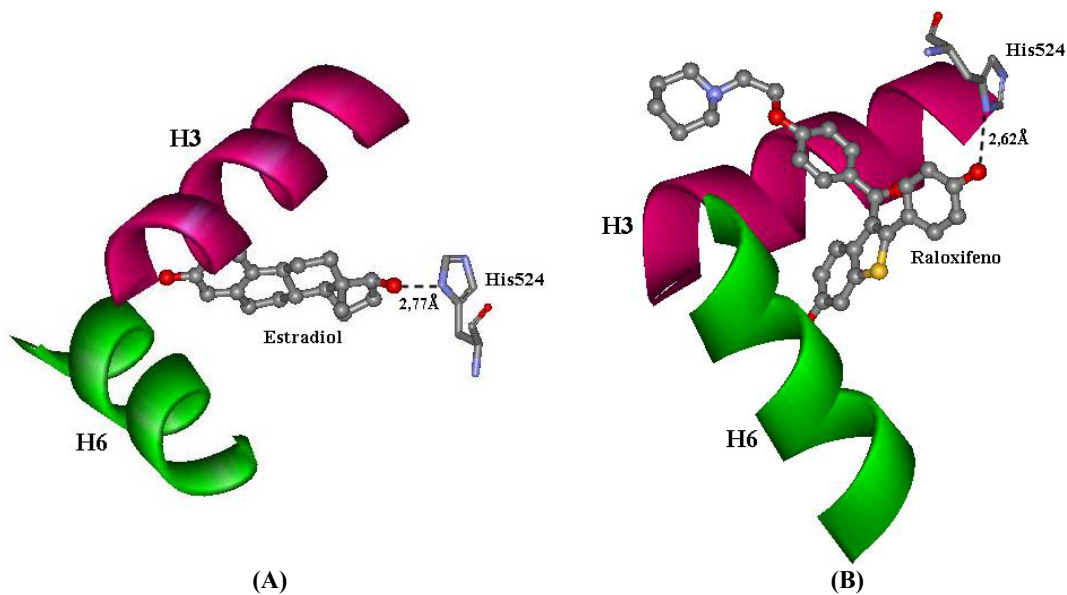
Contudo, a cadeia lateral do raloxifeno (ausente na estrutura do estradiol) é muito longa para ficar contida na cavidade de ligação e, conseqüentemente, ela desloca a hélice H12 e se projeta em direção ao exterior do sítio, entre as hélices H3 e H11 (Figuras 18A e 18B). Esse deslocamento da hélice H12 é uma característica geral dos ligantes do *ER* com ação antagonista, independente dos compostos apresentarem estrutura esteroideal ou não, mas desde que possuam uma cadeia lateral relativamente extensa.

Como o posicionamento da hélice H12 sobre o *LBD* é impedido pela cadeia lateral do raloxifeno, a hélice H12 se posiciona num sulco formado pela H5 e o grupo C-terminal da H3. A complementaridade deste sulco hidrofóbico com a superfície interna da H12 sugere que esta posição do complexo com o raloxifeno representa a real conformação, ao invés de um artefato produzido pelas condições de cristalização. Baseado nessas informações, conclui-se que a atividade antagonista do raloxifeno está baseada na sua habilidade em prevenir uma conformação de *AF-2* competente (Brzozowski *et al.*, 1997).





**Figura 18.** Representação da estrutura tridimensional da região do *LBD* do *ER- $\alpha$*  complexado ao raloxifeno obtida por cristalografia de raios-X (código *PDB*: 1ERR): (A) principais estruturas secundárias ao redor do ligante; (B) interações por ligação hidrogênio do ligante com o *LBD*.



**Figura 19.** Comparação entre as diferentes orientações adotadas pela cadeia lateral do resíduo His524 do *ER- $\alpha$*  quando em complexo com: (A) o 17 $\beta$ -estradiol (código *PDB*: 1ERE) e (B) o raloxifeno (código *PDB*: 1ERR). Esta diferença permite uma orientação adequada deste resíduo para manter uma interação por ligação hidrogênio com o ligante.

## 1.6 METODOLOGIAS DE CORRELAÇÃO QUANTITATIVA ESTRUTURA-ATIVIDADE

Os primeiros a sugerir que o efeito de uma mudança na constituição química de uma molécula (estrutura) acarretaria uma alteração em sua ação fisiológica (atividade) foram Crum-Brown e Fraser em 1868 (Livingstone, 1995; Kubinyi, 1993). Porém, a metodologia conhecida como Correlação Quantitativa Estrutura-Atividade (*QSAR*, *Quantitative Structure-Activity Relationship*), somente foi estabelecida na década de 60, principalmente através dos trabalhos de Hansch e Fujita (Hansch & Fujita, 1964; Hansch, 1969) que em 1964 demonstraram que a atividade biológica poderia ser linearmente correlacionada com diferentes parâmetros físico-químicos, sendo expressa pela Equação (1).

$$\log 1/C = a\pi + b\delta + cEs + d \quad \text{Equação (1)}$$

Na Equação (1), **C** corresponde à concentração molar do composto necessária para produzir uma resposta biológica definida,  **$\pi$**  corresponde à contribuição hidrofóbica,  **$\delta$**  corresponde ao efeito eletrônico, **Es** corresponde ao efeito estérico, **d** corresponde à constante e **a**, **b** e **c** aos respectivos coeficientes determinados por análise de regressão.

Mais tarde, Hansch propôs a extensão do modelo linear para o parabólico, ao notar que a resposta biológica de compostos com caráter hidrofóbico estagnava ou sofria decréscimo após atingir um valor máximo. Isto porque moléculas com alta hidrofobicidade podem ficar retidas nas barreiras lipofílicas e, por isso, têm menor probabilidade de atingir os sítios biológicos de ação (Hansch, 1969). Um modelo parabólico proposto por Hansch com base numa série de dados pressupõe uma relação de segunda ordem da hidrofobicidade,

expressa como o logaritmo do coeficiente de partição entre octanol e água ( $\log P$ ), com a atividade biológica (Debnath, 2001), conforme descrita pela Equação 2.

$$\log 1/C = a(\log P)^2 + b \log P + c \quad \text{Equação (2)}$$

Na década de 70, Kubinyi desenvolveu o modelo bilinear que descreve a dependência não linear da atividade biológica sobre o caráter hidrofóbico, onde o termo  $\beta$  corresponde ao movimento do composto em sistemas multicompartimentados, tal qual o sistema biológico, conforme descrito pela Equação 3. Em alguns casos, o modelo bilinear demonstrou melhores resultados quando comparado ao modelo parabólico (Kubinyi, 1977; Tavares, 2004; Debnath, 2001).

$$\log 1/C = a \log P - b \log(\beta P + 1) + c \quad \text{Equação (3)}$$

Estas metodologias e outras correlacionadas são denominadas de QSAR clássico ou QSAR em uma (1D) ou duas (2D) dimensões, visto que no QSAR-1D a atividade biológica é correlacionada com propriedades físico-químicas (experimentais ou calculadas) como, por exemplo, valores de  $\log P$ ,  $pK_a$  etc, enquanto que no QSAR-2D a atividade biológica é correlacionada com descritores topológicos como por exemplo índice de conectividade molecular (Vedani *et al.*, 2006).

Na década de 80, surgiram diversas abordagens sugerindo uma nova metodologia referida como QSAR tridimensional (3D) (Debnath, 2001). Os descritores do QSAR 3D estão relacionados ao cálculo de propriedades correspondentes à estrutura 3D dos ligantes (método independente do receptor, IR) ou do complexo ligante-receptor (método dependente do

receptor, DP). Tais descritores podem ser classificados em globais (*e.g.* moleculares) e locais (*e.g.* atômicos) (Albuquerque *et al.*, 2006).

Em 1980, Hopfinger incorporou a análise conformacional das estruturas e a obtenção de descritores globais 3D da forma molecular nos estudos de QSAR, criando o método denominado de Análise da Forma Molecular (*Molecular Shape Analysis, MSA*) (Hopfinger, 1980).

Em 1988, Cramer e colaboradores descreveram a metodologia de Análise Comparativa do Campo Molecular (*Comparative Molecular Field Analysis, CoMFA*), o método de QSAR-3D mais comumente utilizado. Esse método demonstra que a atividade biológica dos compostos pode ser correlacionada com as energias estérica e eletrostática (descritores locais) (Cramer *et al.*, 1988). Comparado à análise de Hansch (QSAR-2D), o CoMFA é mais adequado para descrever as interações ligante-receptor, pois considera as propriedades dos ligantes na sua (suposta) conformação bioativa (Kubinyi, 1997).

A metodologia de QSAR-4D (Hopfinger *et al.*, 1997; Albuquerque *et al.*, 1998; Romeiro *et al.*, 2005; Cunha *et al.*, 2005) é uma abordagem de QSAR-3D, que utiliza uma amostragem conformacional obtida por meio de simulação por Dinâmica Molecular (DM), reduzindo a dificuldade em identificar a conformação bioativa.

Com a intenção de aperfeiçoar o poder preditivo das equações de QSAR, dimensões adicionais foram incorporadas aos métodos, como por exemplo, no caso do trabalho de Vedani & Dobler (2002) que relataram a metodologia de QSAR 5D, onde a teoria do encaixe induzido é considerada, ou seja, a adaptação do sítio de ligação do receptor a cada ligante individual. Recentemente, foi desenvolvido o método de QSAR 6D, também desenvolvido pelo grupo de Vadani (Vedani *et al.*, 2005), que considera ainda vários modelos de solvatação simultaneamente.

### 1.6.1 Metodologia de QSAR-4D

A metodologia de QSAR-4D é uma evolução no que diz respeito aos métodos de QSAR-3D, onde a quarta dimensão corresponde à amostragem conformacional obtida por meio de simulação por Dinâmica Molecular (DM) (Albuquerque *et al.*, 2006).

A grande vantagem da análise de QSAR-4D está na sua habilidade em incorporar flexibilidade conformacional ao ligante, em explorar múltiplos alinhamentos e na avaliação exaustiva de grupos farmacofóricos como parte do processo de construção e de otimização do modelo de QSAR (Hong & Hopfinger, 2003).

Da mesma forma que os métodos de QSAR-3D, o QSAR-4D também pode ser aplicado a problemas independente e dependente do receptor. Quando a estrutura tridimensional do receptor não é incluída no estudo de QSAR, este é denominado independente do receptor (IR). Quando a estrutura do complexo ligante-receptor é utilizada para o desenvolvimento do modelo quantitativo, o método é denominado dependente do receptor (DR).

No método de QSAR-4D, os descritores locais são definidos como a frequência de ocupação dos diferentes tipos de átomos nas células unitárias durante a DM, de acordo com cada alinhamento. Estes átomos podem ser definidos de acordo com as interações que podem ocorrer no sítio ativo e estão relacionados aos grupos farmacofóricos (Hopfinger *et al.*, 1997; Albuquerque *et al.*, 2006).

As conformações obtidas de cada composto são alinhadas em uma caixa tridimensional e os descritores são computados e correlacionados com a atividade biológica pelo uso da técnica denominada Aproximação da Função Genética (*Genetic Function Approximation, GFA*) (Rogers & Hopfinger, 1994), que combina Algoritmos Genéticos

(*Genetic Algorithms, GA*) e o método de análise de regressão dos Mínimos Quadrados Parciais (*Partial Least Square, PLS*) (Hopfinger *et al.*, 1997; Albuquerque *et al.*, 2006).

Os modelos podem ser avaliados por procedimentos de validação interna e externa. A validação interna mais comumente empregada é denominada *leave-one-out (LOO)*, na qual um composto é removido do conjunto de treinamento e sua atividade é predita. Na validação externa, um conjunto de teste, ou seja, um determinado número de compostos que não foi incluído no processo de geração dos modelos, é utilizado para predições “reais” (Hopfinger *et al.*, 1997; Albuquerque *et al.*, 1998; Romeiro *et al.*, 2005; Albuquerque *et al.*, 2006).

## 2 OBJETIVOS

Os estrogênios são largamente utilizados como anticoncepcionais em mulheres na idade fértil e na terapia de reposição hormonal em mulheres na pós-menopausa. Porém, um dos alvos dos estrogênios é o tecido mamário, onde estimula a proliferação das células e promovem o crescimento do carcinoma mamário hormônio-dependente (Brueggemeier *et al.*, 1995). Com isso, muitos grupos de pesquisa têm sua atenção voltada para a prevenção e cura do câncer de mama. O raloxifeno, um fármaco utilizado para a prevenção da osteoporose em mulheres na pós-menopausa, tem se mostrado eficaz em relação a esse tipo de câncer (Vogel *et al.*, 2006).

Este trabalho objetiva a construção de modelos de QSAR-4D (IR) para derivados do raloxifeno, sintetizados e avaliados como inibidores da proliferação de células MCF-7 estimulada pelo hormônio  $17\beta$ -estradiol (Grese *et al.*, 1997a). Espera-se que esta análise permita uma maior compreensão das interações moleculares do raloxifeno e seus análogos com o *ER- $\alpha$* , auxiliando no planejamento de novos compostos que apresentem melhor perfil farmacológico.

### 3 METODOLOGIA

#### 3.1 DADOS BIOLÓGICOS: SÉRIE DE DERIVADOS DO RALOXIFENO

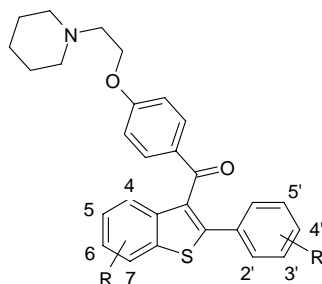
Grese e colaboradores (1997a) descreveram uma série de derivados do raloxifeno, contendo diferentes substituintes no núcleo 2-arilbenzotiofeno deste composto. Desta série, 54 compostos, incluindo o raloxifeno (composto **1**), foram utilizados para a análise de QSAR-4D, sendo que 41 compostos (1–7, 9-13, 15, 16, 19, 21–25, 27, 28, 29, 30, 32, 34–39, 42–45, 47, 49, 50 e 52–54), selecionados aleatoriamente, foram usados como conjunto de treinamento, enquanto que 13 compostos (8, 14, 17, 18, 20, 26, 31, 33, 40, 41, 46, 48 e 51) foram usados como conjunto de teste na etapa de validação externa.

A Tabela 2 descreve a estrutura química dos 54 compostos e suas respectivas potências (variáveis dependentes), as quais são definidas pelos valores de  $pIC_{50}$ . Os valores de  $pIC_{50}$  correspondem a  $-\log IC_{50}$  (M). Os valores de  $IC_{50}$  indicam a dose mínima capaz de inibir 50% da proliferação de células MCF-7 estimuladas pelo hormônio  $17\beta$ -estradiol (em concentração de  $10^{-11}$ M).

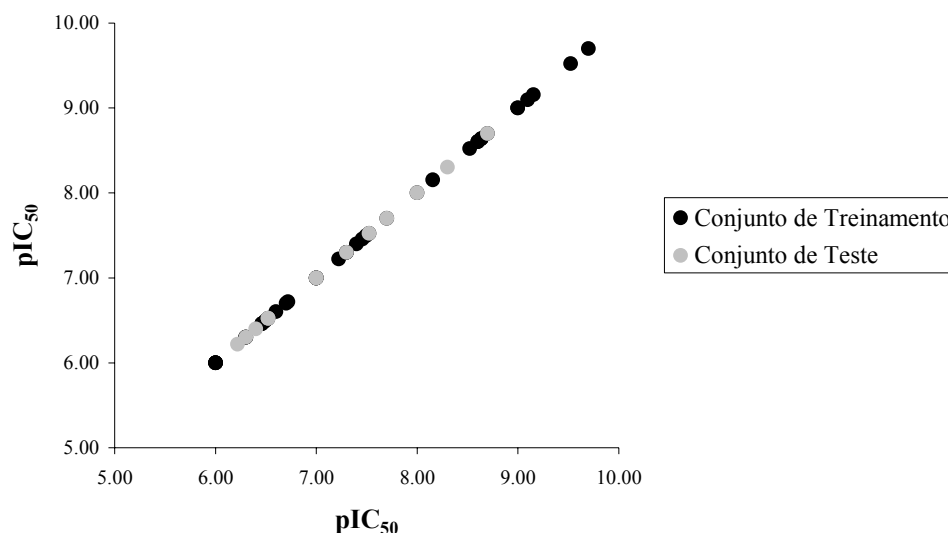
A Figura 20 apresenta a distribuição homogênea dos compostos em relação à atividade antagonista ao  $ER-\alpha$ , onde se verifica que os valores de  $pIC_{50}$  variam de 6,00 a 9,70, ou seja, compreende uma faixa de aproximadamente três unidades logarítmicas.



**Tabela 2.** Estrutura química e valores de  $pIC_{50}$  ( $-\log IC_{50}$ ) da série de 54 derivados do raloxifeno (Grese *et al.*, 1997a). Os composto sublinhados correspondem àqueles do conjunto de teste.



#	R	R'	$pIC_{50}$	#	R	R'	$pIC_{50}$
1	6-OH	4'-OH	9,70	28	6-COMe	4'-OH	7,22
2	6-OH	3'-F, 4'-OH	9,52	29	4,7-di(Me), 6-OH	4'-OH	7,00
3	6-OH	2'-Me	9,15	30	H	4'-OMe	7,00
4	6-OH	4'-C≡CH	9,10	<u>31</u>	5-OH	4'-OH	7,00
5	6-OH	3'-Me, 4'-OH	9,00	32	6-OH	4'-Ph	7,00
6	6-OH	4'-Cl	9,00	<u>33</u>	6-OH	4'-CH <sub>2</sub> SEt	7,00
7	6-OH	2'-Me, 4'-OH	8,70	34	6-OH	3',5'-di(Me), 4'-OH	7,00
<u>8</u>	6-OH	2'-OMe, 4'-OH	8,70	35	4-OH	4'-OH	6,72
9	6-OH	3'-Cl, 4'-OH	8,64	36	6-OH	4'-CONH <sub>2</sub>	6,70
10	6-OH	4'-F	8,64	37	6-OMe	4'-OH	6,60
11	6-OH	H	8,60	38	H	H	6,52
12	6-OH	3'-F	8,60	39	6-OMe	4'-OMe	6,52
13	5-F, 6-OH	4'-OH	8,52	<u>40</u>	6-Me	4'-OH	6,52
<u>14</u>	6-OH	4'-Et	8,30	<u>41</u>	7-OH	4'-OH	6,52
15	6-OH	4'-CH=CH <sub>2</sub>	8,15	42	6-OH	4'-CO <sub>2</sub> H	6,49
16	6-OH	2'-OH	8,00	43	5,6,7-tri(OMe)	4'-OMe	6,49
<u>17</u>	6-OH	4'- <i>n</i> -Bu	8,00	44	4,6-di(OH)	4'-OH	6,46
<u>18</u>	6-OH	4'-CONMe <sub>2</sub>	7,70	45	5,6-di(OH)	4'-OH	6,40
19	6-C≡CH	4'-OH	7,70	<u>46</u>	6-OH	4'-NO <sub>2</sub>	6,30
<u>20</u>	6-OH	4'- <i>i</i> -Pr	7,52	47	4,5-benzo, 6-OH	4'-OH	6,30
21	6-CO <sub>2</sub> Me	4'-OH	7,52	<u>48</u>	6-OMe	3',4'-OCH <sub>2</sub> O-	6,30
22	6-OH	4'-COMe	7,49	49	5,7-di(Me), 6-OH	4'-OH	6,30
23	H	4'-OH	7,46	50	6-OMe	4'-CH <sub>2</sub> OH	6,22
24	6-OH	4'-CONHMe	7,40	<u>51</u>	6-OH	4'-OMe	6,00
25	6-OH	4'-Me	7,30	52	6-CONH <sub>2</sub>	4'-OH	6,00
<u>26</u>	6-OH	4'-CO <sub>2</sub> Me	7,30	53	6-Cl	4'-OH	6,00
27	6-OH	4'-CO <sub>2</sub> Et	7,30	54	6-OH	4'-CF <sub>3</sub>	6,00



**Figura 20.** Distribuição dos valores de atividade ( $pIC_{50exp}$ ) dos 54 compostos derivados do raloxifeno selecionados para o estudo de QSAR-4D (Grese *et al.*, 1997a).

### 3.2 CONSTRUÇÃO E OTIMIZAÇÃO DAS ESTRUTURAS TRIDIMENSIONAIS E OBTENÇÃO DAS CARGAS ATÔMICAS PARCIAIS

A análise de QSAR-4D apresentada neste trabalho consiste em doze etapas, todas descritas por Hopfinger e colaboradores (Hopfinger *et al.*, 1997).

A estrutura 3D de cada derivado foi construída no programa *Hyperchem*<sup>®</sup> 7.0 (2002) por modificações apropriadas na estrutura 3D do raloxifeno complexado com o *ER-α* obtida por cristalografia de raios-X e disponível no *PDB* sob o código 3ERR (Brzozowski *et al.*, 1997). O átomo de nitrogênio do grupo piperidina foi protonado e os átomos de hidrogênio correspondentes foram adicionados. Todas as etapas de cálculo de otimização geométrica foram realizadas no vácuo, sem qualquer restrição geométrica, pelo método de mecânica molecular, utilizando o campo de força MM+, disponível no programa *Hyperchem*<sup>®</sup> 7.0 (2002), até atingir norma de gradiente inferior a 0,1 kcal/mol·Å.

Em seguida, as estruturas foram otimizadas pelo método de mecânica quântica semi-empírico usando o Hamiltoniano AM1 (*Austin Model 1*), disponível no programa

*Hyperchem*<sup>®</sup> 7.0 (2002), até atingir norma de gradiente inferior a 0,1 kcal/mol·Å. Esta etapa foi utilizada para computar as cargas atômicas parciais, necessárias para o cálculo do termo de energia não-ligante eletrostático na etapa da amostragem conformacional por Dinâmica Molecular (DM).

O formato de saída das estruturas obtidas após o cálculo AM1 foi modificado para o formato do programa *Molsim* 3.0 (1994) para ser utilizado como formato de entrada no programa de DM.

### 3.3 AMOSTRAGEM CONFORMACIONAL POR SIMULAÇÃO DE DINÂMICA MOLECULAR

Na análise de QSAR-4D, a quarta dimensão corresponde à amostragem conformacional, obtida por simulação de DM utilizando o campo de força MM2 do programa *Molsim* 3.0 (1994) incorporado no programa *4D-QSAR* (2000). Essa amostragem conformacional corresponde a um conjunto estatisticamente representativo de conformações distintas de baixa energia.

No método de DM, uma trajetória (*i.e.*, conformações como uma função do tempo) de um sistema molecular é gerada pela integração simultânea das equações de movimento de Newton para todos os átomos do sistema. A simulação por DM requer o cálculo do gradiente da energia potencial, que portanto, deve ser uma função diferencial das coordenadas atômicas (van Gunsteren & Berendsen, 1990; van Gunsteren & Mark, 1992).

As variáveis mais importantes na simulação por DM são o tempo e a temperatura em que são realizadas. O tempo de simulação é o produto do número de períodos de simulação pelo tempo de duração de cada período, onde este deve ser o menor possível para que não haja grandes variações de energia potencial durante a simulação (Hopfinger *et al.*, 1997). A simulação dos derivados foi realizada com tempo de simulação de 100 ps, o que é

considerando suficiente para a pesquisa do espaço conformacional de moléculas com até 100 átomos (van Gunsteren & Berendsen, 1990) e intervalos de 0,001 ps para uma amostragem de 10000 períodos. Os arquivos de trajetória foram salvos a cada 50 períodos de simulação, totalizando 2000 conformações para cada molécula.

A temperatura inicial e de referência utilizadas foram 310 K, sendo a mesma utilizada para os testes *in vitro* (Thompson *et al.*, 1988). A temperatura foi mantida constante pelo acoplamento do sistema a um “banho térmico”, de acordo com o modelo proposto por Berendsen (1984), usando tempo de relaxação de 0,1 ps.

### 3.4 DEFINIÇÃO DOS ELEMENTOS DE INTERAÇÃO FARMACOFÓRICA

A metodologia de QSAR-4D permite a definição de sete tipos de elementos de interação farmacofórica (*Interaction Pharmacophore Elements, IPEs*), que correspondem aos tipos de átomos de cada molécula, de acordo com o tipo de interação que podem realizar com os átomos dos aminoácidos do sítio receptor (Tabela 3) (Hopfinger *et al.*, 1997).

**Tabela 3.** Classificação dos Elementos de Interação Farmacofórica (Hopfinger *et al.*, 1997).

Definição	Código	Número
Qualquer Tipo de Átomo	qa	0
Átomo Apolar	ap	1
Átomo Polar Positivo	p+	2
Átomo Polar Negativo	p-	3
Átomo Aceptor de Ligação Hidrogênio	alh	4
Átomo Doador de Ligação Hidrogênio	dlh	5
Átomo em Sistema Aromático	ar	6

### 3.5 DEFINIÇÃO DO TAMANHO DAS CÉLULAS DA CAIXA VIRTUAL TRIDIMENSIONAL

Outra etapa da metodologia de QSAR-4D é a definição de uma caixa virtual tridimensional, composta de células unitárias cúbicas, em torno de cada conjunto de conformações de cada composto. Em geral, essas células apresentam 1 ou 2 Å de aresta, não sendo permitidos números fracionados. Quanto menor o tamanho da célula, maior o número de variáveis independentes gerado e mais complexa a etapa de pré-tratamento dos dados. Porém, o modelo se torna mais acurado em relação aos tipos de átomos selecionados.

No presente trabalho, os alinhamentos foram inicialmente testados em uma caixa com células de 2 Å de aresta e, em seguida, os modelos foram refinados utilizando-se uma caixa com células de 1 Å de aresta.

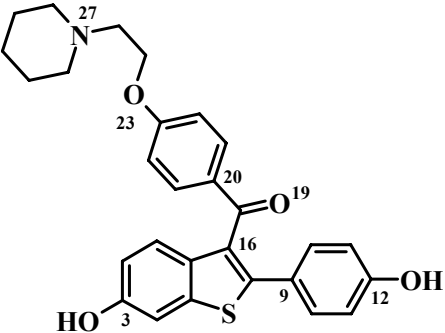
### 3.6 DEFINIÇÃO DOS ALINHAMENTOS

O conjunto de conformações gerado na simulação por DM foi utilizado como entrada nesta etapa da metodologia do QSAR-4D. O alinhamento pressupõe que todos os derivados possuem o mesmo modo de ligação (*i.e.*, orientação relativa) na interação com o receptor e consiste na escolha de três átomos selecionados na mesma ordem. Em geral, a seleção dos átomos é baseada em dados obtidos por estudos de relação estrutura-atividade ou, quando estes não estão disponíveis, por átomos ou grupos funcionais conservados em todas as estruturas químicas. O primeiro átomo selecionado é posicionado na origem das coordenadas cartesianas e a molécula é girada até que o segundo átomo se posicione no eixo x e o terceiro, no eixo y (Hopfinger *et al.*, 1997), de modo que os três átomos fiquem no plano formado pelos eixos x e y.

Dessa forma, propôs-se três alinhamentos usando-se três átomos do raloxifeno comuns a todos os derivados (Tabela 4). Os átomos dos alinhamentos 1 e 2 foram selecionados em função da interação com o receptor, ou seja, pela sua importância farmacofórica. No alinhamento 3 foram selecionados três átomos próximos no espaço, que não correspondem a grupos farmacofóricos, de forma a permitir que grupos farmacofóricos sejam selecionados, visto que os átomos do alinhamento apresentam pouca variação de ocupação, e por este motivo, em geral, não são selecionados como células importantes no modelo.

O Alinhamento 1 (Tabela 4) leva em consideração os átomos de carbono ligados às hidroxilas fenólicas e o átomo de nitrogênio da cadeia lateral. O Alinhamento 2 (Tabela 4), proposto por Cunha e colaboradores (2004), corresponde ao átomo de carbono ligado à hidroxila fenólica do benzotiofeno, o átomo de carbono do grupo fenila diretamente ligado ao anel benzotiofeno e o carbono do anel aromático diretamente ligado ao grupo piperidiniletoxila. O Alinhamento 3 (Tabela 4) envolve os dois átomos de carbono ligados ao átomo de carbono do grupo carbonila e ao átomo de oxigênio da carbonila.

**Tabela 4.** Alinhamentos testados na construção dos modelos de QSAR-4D. A numeração do raloxifeno foi obtida automaticamente no programa Hyperchem e não é IUPAC.



Alinhamento	Átomo 1	Átomo 2	Átomo 3
1	C3	C12	N27
2	C3	C9	C23
3	C16	O19	C20

### 3.7 DEFINIÇÃO DOS TIPOS DE OCUPAÇÃO DAS CÉLULAS

Todas as conformações obtidas por simulação de DM de cada composto são inseridas na caixa virtual tridimensional de dimensão pré-definida, de acordo com os alinhamentos a serem testados. A frequência de ocupação das células dessa caixa por cada tipo de *IPE* selecionado é utilizada como descritor na metodologia de QSAR-4D e é computada para o conjunto de conformações selecionadas na etapa de DM para todas as estruturas do estudo (Hopfinger *et al.*, 1997). Três tipos de ocupação podem ser considerados para cada *IPE*: a *ocupação absoluta*, que independe de um composto de referência e a *ocupação comum* e a *auto-ocupação*, que dependem de um composto de referência.

A *ocupação absoluta* corresponde à frequência de ocupação das células pelos átomos das conformações dos compostos, independente de um composto de referência; na *ocupação comum* é computada a interseção com um composto de referência, que deve ser previamente selecionado; enquanto que na *auto-ocupação*, o valor de ocupação comum é subtraído do valor de ocupação absoluto.

Não existem “regras” para decidir qual tipo de ocupação deve ser utilizado em um determinado estudo. A utilização de um composto de referência influencia os modelos de QSAR-4D em relação às propriedades desse composto. Essa tendência pode ser útil quando o conjunto de treinamento é pequeno. Contudo, se o conjunto de treinamento é extenso, não há necessidade de introduzir nenhuma tendência na análise, e a utilização da ocupação absoluta se torna mais comum (Hopfinger *et al.*, 1997).

Alguns trabalhos descritos na literatura utilizaram entre 20 e 30 compostos para a construção dos modelos de QSAR-4D (Hopfinger *et al.*, 1997; Romeiro *et al.*, 2005; Senese & Hopfinger, 2003). Dessa forma, utilizou-se a ocupação absoluta porque o conjunto de treinamento do presente trabalho, constituído por 41 moléculas, apresenta um número razoável de compostos.

### 3.8 OBTENÇÃO DAS VARIÁVEIS INDEPENDENTES

As variáveis independentes da metodologia de QSAR-4D são os descritores de ocupação das células (*Grid Cell Occupancy Descriptors, GCODs*) que correspondem à frequência de ocupação atômica por célula da caixa 3D (Hopfinger *et al.*, 1997).

Os *GCODs* foram obtidos após a inserção de todas as conformações de cada um dos compostos dentro da caixa 3D, composta por células de tamanho previamente definido (item 3.5), de acordo com cada um dos três alinhamentos testados (item 3.6).

### 3.9 REDUÇÃO DOS DADOS

A análise de QSAR-4D gera inúmeros descritores, os *GCODs*, devido ao grande número de células da caixa tridimensional e aos sete tipos de elementos de interação farmacofórica (*IPEs*) (Hopfinger *et al.*, 1997), o que pode levar a diminuição da capacidade preditiva do modelo. Dessa forma, foram criadas estratégias de seleção de variáveis com o intuito de reduzir o número de dados.

Neste trabalho, foi utilizada uma redução padrão baseada na variância dos dados com valor igual a zero, significando que colunas com variância igual a zero são eliminadas do conjunto de variáveis independentes. Isto gera um conjunto de dados, denominado total, que sofreu outros tipos de redução de dados.

Para a redução adicional de dados, utilizaram-se critérios diferentes: valores de correlação dos descritores em relação à atividade biológica, valores de variância (seleção por um valor de variância maior que zero) e a quantidade de células desocupadas, com o objetivo de eliminar variáveis espúrias (Tabela 5) (Albuquerque *et al.*, 1998).



**Tabela 5.** Critérios para o pré-tratamento do banco de dados.

Alinhamento	Tamanho da célula	Correlação	Variância	Células Desocupadas
1	2Å	< 0,1	< 0,008	> 37
1	1Å	< 0,1	< 0,004	> 37
2	2Å	< 0,1	< 0,005	> 37
2	1Å	< 0,1	< 0,00015	> 36
3	2 Å	< 0,1	< 0,005	> 38
3	1 Å	< 0,1	< 0,00018	> 36

Após este pré-tratamento, os descritores (variáveis independentes) restantes foram utilizados na obtenção dos modelos de QSAR-4D.

### 3.10 OBTENÇÃO DOS MODELOS DE QSAR-4D

Os modelos (ou equações) são obtidos pelo uso combinado das técnicas de algoritmos genéticos (*GA, Genetic Algorithms*) (Rogers & Hopfinger, 1994) e de mínimos quadrados parciais (*PLS, Partial Least Squares*) (Glen *et al.*, 1989) que permite a avaliação de um grande número de expressões lineares contendo um conjunto mínimo de termos que reproduz a atividade (Rogers & Hopfinger, 1994; Hopfinger *et al.*, 1997). Nesta metodologia, os modelos iniciais são gerados pela seleção randômica de descritores (*GCODs*) do conjunto de treinamento, gerando uma população de modelos genéticos de QSAR através de seqüências aleatórias desses descritores. Modelos aprimorados são construídos por operações de cruzamento (*crossover*) e mutação (*mutation*) para recombinar os termos dos modelos de melhor avaliação, ordenados de acordo com os valores do Fator de Desajuste (*Lack of Fit, LOF*) de Friedman (Rogers & Hopfinger, 1994; Romeiro *et al.*, 2005).

O fator de desajuste, *LOF*, corresponde à medida dos mínimos quadrados penalizada, ou seja, quando dois modelos apresentam valores iguais de erro de mínimos quadrados (*Least Square Error, LSE*), aquele que possui o menor número de termos (variáveis independentes)

terá o menor, e conseqüentemente, o melhor valor de *LOF*, conforme descrito pela equação 4 (Rogers & Hopfinger, 1994).

$$LOF = \frac{LSE}{\left(1 - \left(\frac{c + dp}{M}\right)\right)^2} \quad \text{Equação (4)}$$

Na Equação (4), **LSE** corresponde ao erro dos mínimos quadrados, **c** corresponde ao número de termos do modelo, **d** corresponde ao fator de suavização (*smoothing factor*) que ajusta a quantidade de termos no modelo, **p** corresponde ao número de variáveis de cada termo (neste trabalho cada termo corresponde a uma variável independente) e **M** é o número de compostos do conjunto de treinamento.

Na etapa de *GA* foram testados valores de 6000 a 20000 para cruzamento, valor de 100% para mutação e valores de 0,5 a 3,0 para o fator de desajuste.

Na avaliação dos modelos finais de QSAR podem ser utilizados índices estatísticos que incluem o quadrado do coeficiente de correlação entre os descritores e a atividade biológica,  $R^2$ , e o teste F, que estima a significância do modelo.

O quadrado do coeficiente de correlação,  $R^2$ , indica a fração da variação total nas variáveis dependentes que é explicada pelo modelo.  $R^2$  pode adotar valores entre 0 e 1; o valor de 0 indica que a variação observada não é explicada pela variação das variáveis independentes, enquanto que o valor de 1 indica que a variação observada pode ser explicada. A desvantagem do valor de  $R^2$  é que o mesmo é dependente do número de variáveis independentes, com altos valores obtidos quando o conjunto de dados é grande (Leach, 2001).

O teste F mede a significância estatística de  $R^2$ , pois verifica o quanto da variabilidade da atividade biológica (variável dependente) pode ser explicada pela variação das variáveis independentes e o quanto pode ser atribuída ao erro aleatório. Para validar  $R^2$  pelo teste F, é

necessário comparar o valor de F obtido no modelo com o valor de referência, que em geral se refere ao nível de confiança de 95% e pode ser obtido em tabelas apropriadas (Gaudio & Zandonade, 2001).

### 3.11 VALIDAÇÃO INTERNA DOS MODELOS

Os dez melhores modelos obtidos pelo método *GA-PLS*, de acordo com seus valores de *LOF*, são submetidos à validação interna (ou validação cruzada). A forma mais comumente utilizada é pelo método *Leave One Out, LOO*, que gera valores de  $q^2$  (quadrado do coeficiente de correlação linear,  $R^2$ , após a validação cruzada), índice que traduz a capacidade preditiva do modelo. Para o cálculo deste coeficiente para um determinado modelo, um composto é excluído e os valores dos coeficientes para o modelo em avaliação são recalculados a partir dos  $(n - 1)$  compostos remanescentes, a nova equação assim obtida é utilizada para prever a atividade do composto que foi excluído. Quando todos os compostos forem excluídos pelo menos uma vez, o valor de  $q^2$  é calculado, como demonstrado pela Equação 5.

$$q^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{Equação (5)}$$

Na Equação (5),  $y_i$  é o valor da atividade experimental,  $\bar{y}$  é a média dos valores e  $\hat{y}$  é o valor da atividade predita pelo modelo (Migliavacca, 2003). Em geral, o valor de  $q^2$  será menor do que o valor de  $R^2$ , porém é mais representativo da qualidade do modelo (em termos de preditividade) do que o esperado pela regressão (Livingstone, 1995).

Para todos os alinhamentos testados, foram selecionadas as equações de QSAR com os maiores valores de  $q^2$ , contendo de 7 a 9 termos, pois deve haver cerca de cinco ou seis compostos para cada variável incluída no modelo (Kubinyi, 1993; Gaudio & Zandonade,

2001). Contudo, os valores de  $q^2$  devem ser transformados em  $q^2_{ajustado}$ , demonstrado pela Equação 6, que permite comparação entre modelos com número diferente de variáveis (Livingstone, 1995).

$$q^2_{ajustado} = 1 - \left(1 - q^2\right) \frac{n-1}{n-p} \quad \text{Equação (6)}$$

Na Equação (6),  $q^2$  corresponde ao quadrado do coeficiente de correlação após a validação cruzada,  $n$  corresponde ao número de compostos do conjunto de treinamento e  $p$  corresponde à quantidade de variáveis do modelo.

### 3.12 SELEÇÃO DOS MELHORES MODELOS

Os melhores modelos obtidos de QSAR-4D são selecionados por exclusão de modelos intercorrelacionados e de modelos com variáveis independentes intercorrelacionadas.

Os modelos intercorrelacionados são detectados pelo cálculo da matriz de correlação cruzada dos valores residuais, que correspondem às diferenças entre os valores de atividade biológica experimental ( $pIC_{50exp}$ ) e calculada ( $pIC_{50calc}$ ), eliminando modelos com o mesmo tipo de informação estrutura-atividade (Hopfinger *et al.*, 1997).

A matriz de correlação cruzada das diferentes variáveis independentes é calculada para verificar se existem termos altamente correlacionados num mesmo modelo, descartando aqueles que apresentem informação redundante.

### 3.13 VALIDAÇÃO EXTERNA DOS MODELOS E OBSERVAÇÃO DOS *OUTLIERS*

Para avaliação final dos modelos, foi realizada uma validação externa, utilizando o conjunto de teste composto por 13 compostos da série, mas que não foram incluídos no

processo de geração dos modelos de QSAR-4D. Isto é realizado pela obtenção das variáveis independentes para estes compostos conforme descrito nos itens 3.2 a 3.8. Esses valores são então aplicados nas equações obtidas de QSAR-4D para o cálculo dos valores de atividade biológica, sendo comparados aos valores de atividade biológica experimental.

O nível de preditividade dos modelos de QSAR pode ser verificado pelo cálculo do desvio-padrão dos valores residuais ( $DP_{res}$ ), para os conjuntos de treinamento e de teste. Os compostos denominados *outliers* não podem ser explicados pelo modelo (Kubinyi 1993) sendo definidos como aqueles cujos valores de resíduos são maiores que o dobro do  $DP_{res}$  (Albuquerque *et al.*, 1998). A análise cuidadosa dos *outliers* gera inúmeras informações sobre os limites da aplicação da metodologia e auxilia na compreensão do mecanismo de ação dos compostos em estudo (Pita, 2006).

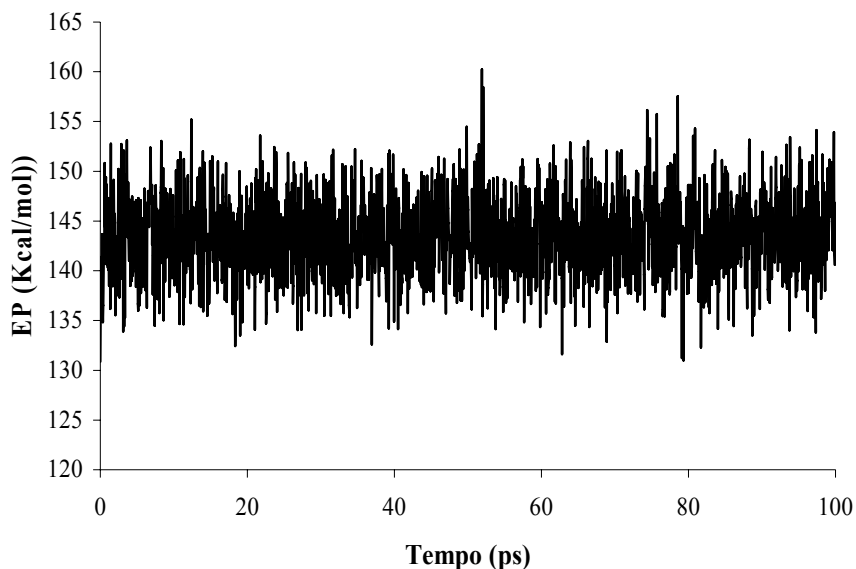
#### 3.14 SELEÇÃO DA CONFORMAÇÃO “BIOATIVA”

Na metodologia de QSAR-4D, a conformação “bioativa” é definida como a conformação que maximiza a potência do composto no modelo selecionado (Hopfinger *et al.*, 1997). Neste trabalho, selecionamos a conformação de mais baixa energia, até 10 kcal/mol da conformação de mínimo de energia, após a etapa de dinâmica molecular. A conformação obtida é, então, sobreposta no sítio ativo do *ER- $\alpha$*  (código *PDB* 1ERR), pelos respectivos alinhamentos testados, utilizando-se o programa *Hyperchem*<sup>®</sup> 5.0 (1997), com a intenção de observar possíveis interações com os resíduos de aminoácidos vizinhos (Hopfinger *et al.*, 1997; Romeiro *et al.*, 2005).

## 4 RESULTADOS E DISCUSSÃO

### 4.1 ANÁLISE DA SIMULAÇÃO POR DINÂMICA MOLECULAR

A análise do gráfico de valores de energia potencial (kcal/mol) *versus* o tempo de simulação (ps) demonstra que a etapa de simulação por dinâmica molecular (DM) atingiu o equilíbrio, visto que ocorreram flutuações constantes de energia potencial com o decorrer do tempo para todos os compostos. Como exemplo, a Figura 21 representa o gráfico do composto 1.



**Figura 21.** Gráfico dos valores de energia potencial (kcal/mol) *versus* tempo de simulação (ps) por simulação de dinâmica molecular em temperatura constante (310K) para o composto 1.

### 4.2 SELEÇÃO DO MELHOR ALINHAMENTO

Como descrito no item 3.6, três alinhamentos foram propostos. Os alinhamentos 1 e 2 foram baseados na importância farmacofórica dos átomos selecionados, enquanto que no alinhamento 3, os átomos selecionados estão próximos entre si, não englobando grupos

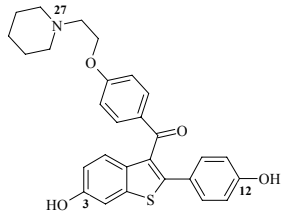
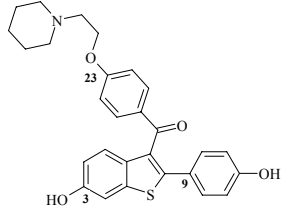
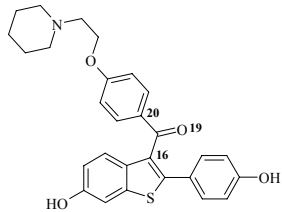
farmacofóricos. Utilizou-se células unitárias de 2Å de aresta e, posteriormente, os modelos foram refinados com células de 1Å de aresta.

Analisou-se os alinhamentos testados em função do número de variáveis independentes (*GCODs*). Para evitar um superajuste das equações, admite-se um número máximo de cinco termos por composto do conjunto de treinamento (Kubinyi H, 1993; Gáudio & Zandonade, 2001; Pita, 2006). Neste caso, o número máximo de termos seria 8,2 (para 41 compostos), o que levou à análise de equações com sete, oito e nove descritores. Modelos com mais de nove termos representam um superajuste (*overfitting*) dos dados, apesar dos valores de  $q^2$ .

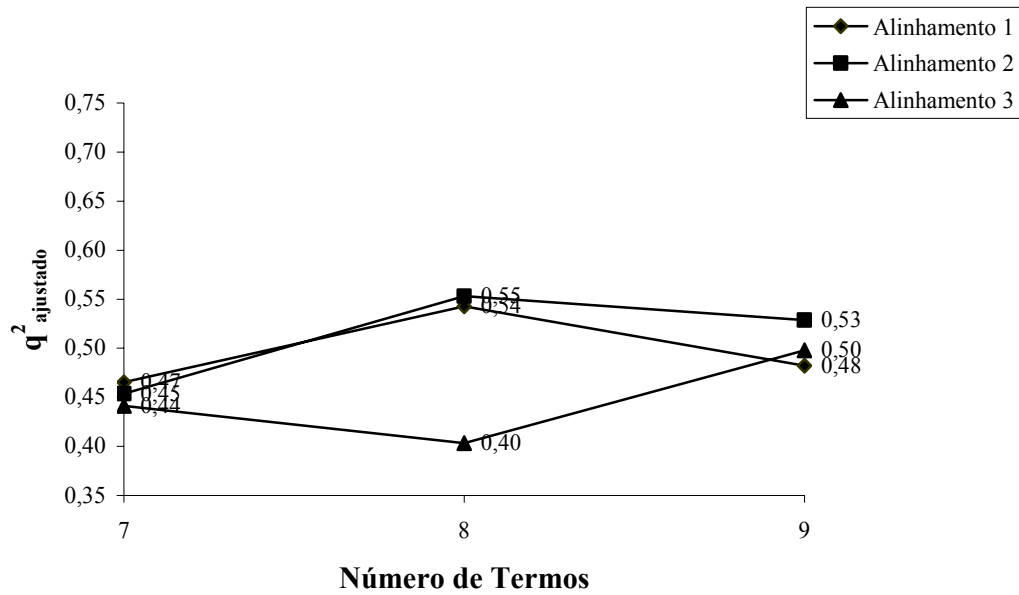
Na escolha do melhor alinhamento, também foram levados em consideração parâmetros estatísticos como o quadrado do coeficiente de correlação ajustado,  $R^2_{\text{ajustado}}$ , que reflete o ajuste dos dados à reta de regressão e, principalmente, o quadrado do coeficiente de correlação linear ajustado após a validação cruzada,  $q^2_{\text{ajustado}}$ , que reflete a capacidade preditiva do modelo (Pita, 2006). Para serem considerados com boa capacidade preditiva, os modelos devem apresentar valores de  $q^2$  maiores do que 0,5 (Kuninyi, 1993).

Na Tabela 6 são apresentados os valores de  $R^2$ ,  $R^2_{\text{ajustado}}$ ,  $q^2$  e  $q^2_{\text{ajustado}}$  para os modelos com 7 a 9 termos obtidos pelos alinhamentos 1 a 3 em células de 2 e 1Å. Os valores de  $R^2$  e  $q^2$  só podem ser usados para comparar equações com o mesmo número de termos, enquanto que os valores de  $R^2_{\text{ajustado}}$  e  $q^2_{\text{ajustado}}$  podem ser usados para comparar equações com número de termos iguais ou diferentes. Nas Figuras 22 e 23 são apresentados os gráficos dos valores de  $q^2_{\text{ajustado}}$  *versus* o número de termos das equações obtidas em células de 2 e 1Å, respectivamente, segundo os alinhamentos 1 a 3.

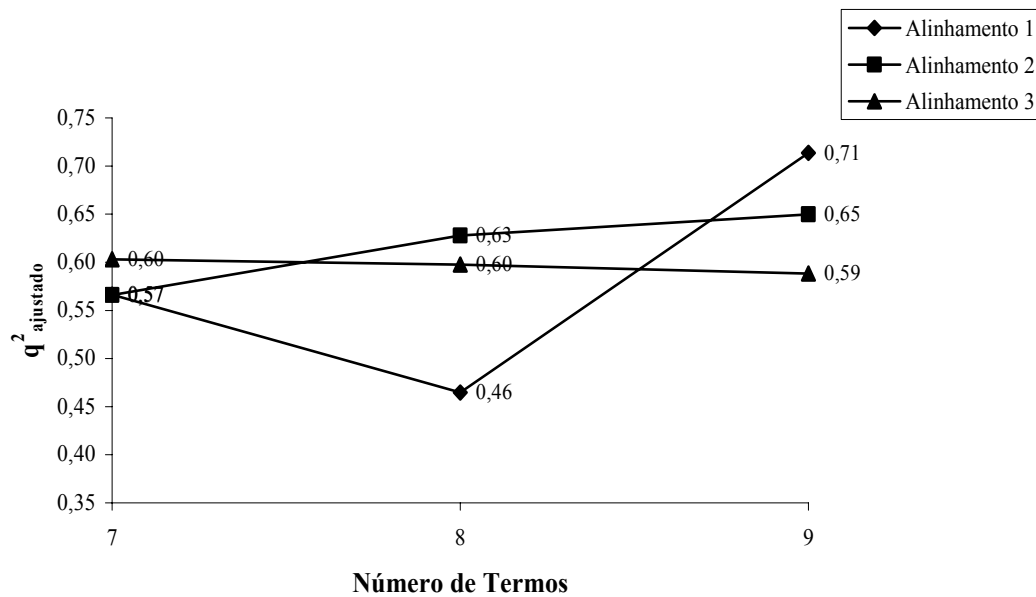
**Tabela 6.** Valores de  $R^2$ ,  $R^2_{\text{ajustado}}$ ,  $q^2$  e  $q^2_{\text{ajustado}}$  dos modelos com 7 a 9 termos obtidos pelos alinhamentos 1 a 3 em células de 2 e 1Å. A numeração do raloxifeno foi obtida automaticamente no programa Hyperchem e não é IUPAC.

Alinhamento	Célula	Modelo	Termos	$R^2$	$R^2_{\text{ajustado}}$	$q^2$	$q^2_{\text{ajustado}}$	
<b>1 (C3, C12, N27)</b> 		1A7	7	0,65	0,59	0,55	0,47	
	<b>2Å</b>	1A8	8	0,71	0,65	0,62	0,54	
		1A9	9	0,70	0,63	0,59	0,48	
	<b>1Å</b>	1B7	7	0,72	0,68	0,63	0,57	
		1B8	8	0,68	0,61	0,56	0,46	
		1B9	9	0,84	0,81	0,77	0,71	
	<b>2 (C3, C9, C23)</b> 		2A7	7	0,65	0,59	0,54	0,45
		<b>2Å</b>	2A8	8	0,73	0,67	0,63	0,55
			2A9	9	0,73	0,66	0,62	0,53
<b>1Å</b>		2B7	7	0,72	0,67	0,63	0,57	
		2B8	8	0,76	0,71	0,69	0,63	
		2B9	9	0,81	0,76	0,72	0,65	
<b>3 (C16, O19, C20)</b> 			3A7	7	0,62	0,55	0,52	0,44
		<b>2Å</b>	3A8	8	0,64	0,56	0,50	0,40
			3A9	9	0,71	0,65	0,59	0,50
	<b>1Å</b>	3B7	7	0,73	0,69	0,66	0,60	
		3B8	8	0,77	0,72	0,67	0,60	
		3B9	9	0,76	0,70	0,67	0,59	





**Figura 22.** Valores de  $q^2_{\text{ajustado}}$  versus número de termos dos modelos dos alinhamentos 1 a 3 em células de 2Å.



**Figura 23.** Valores de  $q^2_{\text{ajustado}}$  versus número de termos dos modelos dos alinhamentos 1 a 3 em células de 1Å.

Os dados da Tabela 6 indicam que, independente do alinhamento, houve uma melhora significativa quando se compara os modelos obtidos em células de 2 e 1Å, visto que, em células de 2Å, apenas quatro dos nove modelos apresentaram valores de  $q^2_{\text{ajustado}}$  maior do que

0,5, enquanto que, em células de 1Å, oito dos nove modelos apresentaram valores de  $q^2_{\text{ajustado}}$  maior do que 0,5. Além disso, quando se compara modelos com o mesmo número de termos, provenientes de um mesmo alinhamento, observa-se um aumento do valor do  $q^2_{\text{ajustado}}$  quando o tamanho da célula passa de 2Å para 1Å, com exceção do par de modelos 1A8 e 1B8 em que ocorre uma redução do valor de  $q^2_{\text{ajustado}}$ . Pode-se concluir que, no geral, os modelos provenientes de células de 1Å tem maior capacidade preditiva do que os modelos provenientes de células de 2Å.

Embora o alinhamento 3 tenha gerado quatro modelos com valores de  $q^2_{\text{ajustado}}$  satisfatórios, uma análise preliminar destes modelos (dados não mostrados) demonstrou que a localização espacial dos descritores selecionados pelos melhores modelos não estava coerente com o mecanismo de ação dos ligantes do *ER*. Por esta razão, os modelos provenientes do alinhamento 3 foram descartados, sendo analisados posteriormente apenas os modelos provenientes dos alinhamentos 1 e 2.

#### 4.3 ANÁLISE DOS MELHORES MODELOS DE QSAR-4D DO ALINHAMENTO 1

Os melhores modelos obtidos pelo alinhamento 1, em células de 1Å, apresentam valores de  $q^2$  e  $q^2_{\text{ajustado}}$  ligeiramente superiores àqueles obtidos em células de 2Å, com exceção dos modelos de 8 termos. Além disso, o modelo 1B8 demonstrou baixa capacidade preditiva, visto que o valor de  $q^2_{\text{ajustado}}$  encontrado é inferior a 0,5. Assim, nesse trabalho, foram analisados os modelos 1B7 e 1B9. A Tabela 7 apresenta as melhores equações de QSAR-4D, com 7 e 9 termos, obtidas pelo alinhamento 1 em células de 1Å. Na Tabela 7, os números entre parênteses indicam as coordenadas cartesianas x, y e z da célula e as letras entre parêntesis indicam o tipo de *IPE* dos átomos que ocupam a célula. O sinal do coeficiente indica se a ocupação da célula contribui para aumentar a potência, caso este sinal seja positivo, ou para diminuir a potência, caso este sinal seja negativo.

**Tabela 7.** Melhores modelos do alinhamento 1 em células de 1Å.

Modelo	Termos	Equação <sup>a</sup>
1B7	7	$\begin{aligned} \text{pIC}_{50} = & 7,89 - 13,99 (2,11,4)(\text{qa}) - 6,96 (-1,8,2)(\text{qa}) \\ & - 2,04 (0,-2,9)(\text{qa}) + 22,68 (1,13,1)(\text{qa}) \\ & + 12,16 (1,-1,9)(\text{qa}) + 1,84 (1,2,2)(\text{ar}) \\ & + 19,60 (1,11,-2)(\text{qa}) \end{aligned}$
1B9	9	$\begin{aligned} \text{pIC}_{50} = & 8,19 - 18,84 (2,11,4)(\text{qa}) - 6,13 (0,-2,-1)(\text{qa}) + 8,45 (0,1,-2)(\text{qa}) \\ & - 8,71 (0,11,5)(\text{qa}) - 7,54 (-1,8,2)(\text{qa}) - 20,98 (1,6,-2)(\text{ap}) \\ & + 15,96 (1,13,1)(\text{qa}) + 3,93 (1,2,0)(\text{qa}) \\ & + 1,61 (0,0,-2)(\text{qa}) \end{aligned}$

<sup>a</sup> Os números entre parênteses representam as coordenadas cartesianas x,y,z dos *IPEs*; as letras entre parênteses representam a classificação dos *IPEs*: qa = qualquer tipo de átomo, ar = aromático e ap = apolar.

Para verificar se os modelos apresentam o mesmo tipo de informação estrutural-atividade, foi calculada a matriz de correlação dos valores residuais dos modelos com 7 e 9 termos (1B7 e 1B9, respectivamente), como descrito no item 3.12. Os valores residuais são calculados como a diferença entre os valores de  $\text{pIC}_{50}$  calculados e os respectivos valores de  $\text{pIC}_{50}$  experimentais. Essa diferença é denominada de resíduo ou erro. Espera-se que modelos equivalentes apresentem uma distribuição dos resíduos idêntica e que modelos distintos tenham uma distribuição de resíduos não correlacionada (Albuquerque *et al.*, 1997; Romeiro, 2002; Pitta, 2006).

Entre os dois modelos, observa-se uma correlação (R) de 0,59, indicando uma correlação significativa. A razão desta semelhança pode ser atribuída ao fato de que os dois modelos apresentam, nas respectivas equações, células de coordenadas espaciais iguais e com o mesmo tipo de ocupação, ou seja, os modelos 1B7 e 1B9 apresentam os seguintes *GCODs* idênticos: (2,11,4)(qa), (-1,8,2)(qa) e (1,13,1)(qa).

Foi calculada também a matriz de correlação cruzada entre os valores de ocupação das células presentes nos modelos 1B7 e 1B9 (Anexo I). As maiores correlações ocorrem entre os

pares de células (0,1,-2)(qa) / (0,-2,-1)(qa) (R= 0,47) e (-1,8,2)(qa) / (1,13,1)(qa) (R= 0,52). O primeiro par de células está presente apenas no modelo 1B9, enquanto que o segundo par está presente nos dois modelos, o que corrobora para a tese de que os modelos contêm informações semelhantes entre si.

A partir dos valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) são calculados os valores de desvio padrão dos resíduos ( $DP_{res}$ ) e identificados os *outliers*, que correspondem aos compostos com valores de resíduo maiores do que o dobro do valor de  $DP_{res}$ . O  $DP_{res}$  é um critério de ajuste do modelo e, quanto maior a variabilidade dos valores que o modelo for capaz de explicar, maior será o valor de  $R^2$  e menor será o  $DP_{res}$  (Romeiro, 2002). Nota-se que o modelo 1B9 apresenta menor valor de  $DP_{res}$  e menor número de compostos considerados *outliers* (Tabela 9 e Anexo II).

A análise da matriz de correlação entre os valores de atividade biológica experimental e a frequência de ocupação das células dos modelos 1B7 e 1B9 (Anexo I) demonstra que, com exceção da célula (1,11,-2)(qa) (R= 0,57), as células que apresentam maior correlação (positiva ou negativa) com a atividade estão presentes no modelo 1B9: (2,11,4)(qa) (R= -0,46), (0,0,2)(qa) (R= 0,43), (1,2,0)(qa) (R= 0,36) e (0,11,5)(qa) (R= -0,35). Além disso, esse modelo também apresenta valores de  $R^2_{ajustado}$  e  $q^2_{ajustado}$  significativamente melhores. Dessa forma, o modelo 1B9 foi selecionado (com 9 termos) como o modelo representativo do alinhamento 1.

#### **4.3.1 Análise do Modelo Representativo do Alinhamento 1**

O modelo 1B9, conforme descrito na Equação (7), foi selecionado como o modelo representativo do alinhamento 1, devido aos parâmetros estatísticos adequados e à

comparação com os demais modelos obtidos pelo mesmo alinhamento, conforme descrito no item 4.2.1.

$$\begin{aligned}
 pIC_{50} = & 8,19 - 18,84 (2,11,4) (qa) - 6,13 (0,-2,-1) (qa) & \text{Equação (7)} \\
 & + 8,45 (0,1,-2) (qa) - 8,71 (0,11,5) (qa) - 7,54 (-1,8,2) (qa) \\
 & - 20,98 (1,6,-2) (ap) + 15,96 (1,13,1) (qa) + 3,93 (1,2,0) (qa) \\
 & + 1,61 (0,0,-2) (qa)
 \end{aligned}$$

$$N = 41 \quad R^2 = 0,84 \quad q^2 = 0,77 \quad F = 25,9 \quad SE = 0,42$$

A matriz de correlação cruzada dos valores de ocupação de todos os *GCODs* presentes no modelo foi calculada com o propósito de observar se eles indicam o mesmo tipo de informação, ou seja, se apresentam informações redundantes na equação. Na Tabela 8 é dada a matriz de correlação entre os *GCODs* do modelo 1B9.

**Tabela 8.** Matriz de correlação cruzada entre os *GCODs* do modelo 1B9 obtido pelo método de QSAR-4D em célula de 1Å a partir do alinhamento 1.

<i>GCODs</i>	(2,11,4) (qa)	(0,-2,-1) (qa)	(0,1,-2) (qa)	(0,11,5) (qa)	(-1,8,2) (qa)	(1,6,-2) (ap)	(1,13,1) (qa)	(1,2,0) (qa)	(0,0,-2) (qa)
(2,11,4) (qa)	1,00								
(0,-2,-1) (qa)	0,13	1,00							
(0,1,-2) (qa)	0,36	0,47	1,00						
(0,11,5) (qa)	0,30	0,01	0,13	1,00					
(-1,8,2) (qa)	0,20	0,10	0,12	-0,17	1,00				
(1,6,-2) (ap)	-0,21	-0,07	-0,13	-0,21	0,06	1,00			
(1,13,1) (qa)	0,13	0,00	-0,07	-0,37	0,52	0,01	1,00		
(1,2,0) (qa)	-0,01	0,18	0,18	-0,16	0,22	0,11	0,10	1,00	
(0,0,-2) (qa)	-0,31	0,26	-0,19	0,01	-0,11	0,07	-0,13	0,28	1,00

Pode-se observar que, com exceção de apenas dois pares de células, os demais descritores são pouco correlacionados, pois não apresentam valores de correlação significativa (ou seja, valor de R maior que 0,5). Isto significa que cada um desses descritores contribui de forma distinta para os modelos de QSAR-4D (Kubinyi, 1993; Pitta, 2006).

Os *GCODs* (1,13,1)(qa) e (-1,8,2)(qa), apesar de apresentarem contribuições opostas, visto que a ocupação do *GCOD* (1,13,1)(qa) aumenta a potência do composto e a ocupação do *GCOD* (-1,8,2)(qa) diminui a potência do composto, e estarem distante no espaço (5,48Å), parecem fornecer parcialmente o mesmo tipo de informação estrutura-atividade. Ambos os descritores estão localizados próximo à cadeia lateral dos compostos arilbenzotiofenos, sendo relacionados à flexibilidade desta cadeia lateral básica. Além disso, o fato do *GCOD* (1,13,1)(qa) ser o descritor que mais contribui para o aumento da potência (coeficiente igual a 15,96, Equação (4)), indica a relevância da conformação sinclinal da subestrutura aminoetóxi da cadeia lateral para a potência dos compostos.

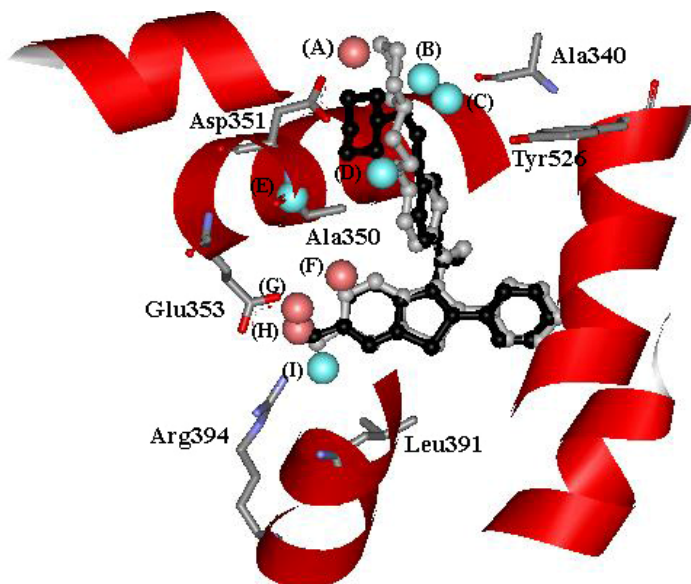
Como pode ser visto nas Figuras 24 e 25, o *GCOD* (1,13,1)(qa) está localizado mais próximo ao átomo de carbono 4 do anel piperidina na extremidade da cadeia lateral, enquanto que o *GCOD* (-1,8,2)(qa) está localizado mais próximo ao átomo de oxigênio do grupo piperidinil-etóxi-fenil no início da cadeia lateral. A correlação encontrada entre essas células se deve, provavelmente, à conformação sinclinal (*gauche*) da cadeia lateral que envolve os átomos O-C-C-N do grupo piperidinil-etóxi-fenil, devido à forte interação eletrostática atrativa entre o grupo etóxi e o grupo amina protonado, reforçada por uma interação por ligação hidrogênio do tipo NH...O (distância N...O = 3,02Å).

Além desses dois *GCODs*, as células (2,11,4)(qa) e (0,11,5)(qa), que estão próximas no espaço e apresentam coeficiente negativo, também estão relacionadas à cadeia lateral básica dos compostos, localizadas próximas aos átomos de carbono 2 e de nitrogênio do anel piperidina, respectivamente (Figura 24 e 25). Esses quatro *GCODs*, em conjunto, indicam

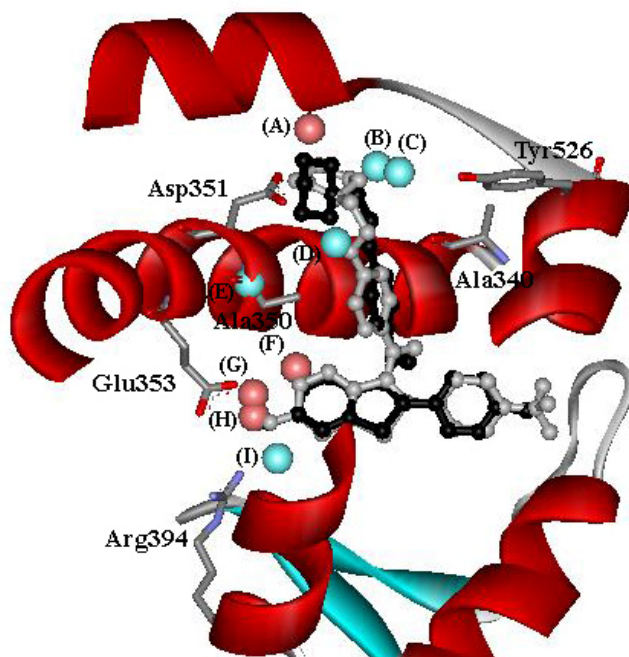
uma orientação preferencial do grupo piperidina que faz interação eletrostática intensificada por ligação hidrogênio com o resíduo Asp351, orientação essa que é essencial para o aumento ou diminuição da potência dos compostos derivados do raloxifeno. Além dessas interações, a cadeia lateral básica do raloxifeno também faz extensos contatos hidrofóbicos com as hélices H3, H5/6, H11 e a alça entre as hélices H11 e H12 (Brzozowski *et al.*, 1997), reforçando a importância da orientação e conformação da cadeia lateral.

As células (0,1,-2)(qa) e (0,-2,-1)(qa) também podem ser consideradas semelhantes, porque apresentam correlação muito próxima a 0,5 ( $R= 0,47$ ), apesar de representarem contribuições opostas, visto que a ocupação do *GCOD* (0,1,-2)(qa) aumenta a potência do composto enquanto que a ocupação do *GCOD* (0,-2,-1)(qa) diminui a potência, e por que estão próximas entre si no espaço (3,16Å), levando ao mesmo tipo de informação.

A representação tridimensional do modelo 1B9 é apresentada nas Figuras 24 e 25 para o composto **1**, o mais potente da série, e para o composto **54**, o menos potente da série, respectivamente. A conformação representada para cada composto é aquela obtida por DM que prediz a maior resposta biológica, dentro de uma diferença de energia de 10 kcal/mol da conformação de menor energia.



**Figura 24.** Representação da conformação “bioativa” do composto **1** (cinza), obtida pelo modelo 1B9 de QSAR-4D, sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 1. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).



**Figura 25.** Representação da conformação “bioativa” do composto **54** (cinza), obtida pelo modelo 1B9 de QSAR-4D, sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 1. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).



Os *GCODs* (0,1,-2)(qa), (1,2)(qa)(qa), (0,0,-2)(qa) e (0,-2,-1)(qa) indicam a relevância da ligação hidrogênio em torno do anel benzotiofeno para a atividade antagonista dos ligantes frente ao *ER-α*. A célula (0,0,-2)(qa) reflete a importância da ligação hidrogênio do grupo hidroxila na posição 6 do anel benzotiofeno com o resíduo Glu353, como descrito nos estudos de relação estrutura-atividade (Grese *et al.*, 1997a; Grese *et al.*, 1997b; Brzozowski, 1997), pois sua ocupação é drasticamente menor quando não há substituintes nesta posição, ou substituintes incapazes de promover esse tipo de interação (*e.g.* compostos **23**, **30**, **31**, **35**, **38** e **41**) (Anexo III). O mesmo ocorre com o *GCOD* (0,1,-2)(qa), pois compostos com substituintes aceptores de ligação hidrogênio na posição 6 ocupam bastante esta célula (*e.g.* compostos **37**, **49** e **50**) (Anexo III).

A estrutura de cristalografia de raios-X do complexo raloxifeno-*ER-α* mostra que o anel A do raloxifeno está envolvido por resíduos hidrofóbicos, como por exemplo, Ala350 e Leu387 (Brzozowski *et al.*, 1997). Por esta razão, a ocupação de qualquer tipo de átomo na célula (0,-2,-1) desfavorece a potência dos compostos devido a fatores estéricos. Isto pode ser explicado pelo fato de que compostos com substituintes mais volumosos que uma hidroxila, como por exemplo os compostos **21**, **28**, **37**, **39**, **43** e **52**), ocupam mais esta célula do que compostos com substituintes menos volumosos, como por exemplo, os compostos **19**, **23**, **30**, **34** e **38** (Anexo III).

Finalmente, a célula (1,6,-2)(ap) (Figura 24 e 25) está localizada numa região da caixa 3D que corresponde à posição do resíduo Ala350. O coeficiente negativo (Equação (4)) indica que a ocupação desta célula por átomos apolares, como os átomos de carbono e hidrogênio do anel piperidina, dependendo da conformação da cadeia lateral, diminui a potência dos compostos, pois se localiza muito próximo ao resíduo hidrofóbico de alanina, o que indica um impedimento estérico desta célula.

Embora as Figuras 24 e 25 não mostrem claramente nenhum átomo próximo a esta célula, determinadas conformações adotadas pelos compostos durante a etapa de DM podem ocupar esta região. As Figuras 24 e 25 demonstram apenas uma conformação, dentre as 2000 conformações de cada composto, selecionada como “bioativa” que apresenta potência máxima no modelo 1B9. Isto pode ser observado pela ocupação desse descritor (1,6,-2)(ap) (Anexo III), como exemplificado pelos compostos com substituintes na posição 2' da fenila que ocupam muito pouco esta célula (*e.g.* compostos **3**, **5**, **7** e **9**), assim como compostos com substituintes nas posições 4 e/ou 5 do anel benzotiofeno (*e.g.* compostos **13**, **35**, **43**, **44** e **45**). Isto indica que tais substituintes tentam manter uma certa orientação da cadeia lateral, levando-a a uma conformação propícia ao antagonismo sobre o *ER-α*. Esta é uma característica adicional não revelada pelo modelo de LIV-3D-QSAR proposto por Cunha e colaboradores (Cunha *et al.*, 2004).

A ausência de descritores ao redor da fenila, ligada à posição 2 do anel benzotiofeno, em particular, em relação à hidroxila da posição 4' que é responsável pela interação por ligação hidrogênio com o resíduo Phe524, corrobora o que estudos anteriores de relação estrutura-atividade (Grese *et al.*, 1997a, Grese *et al.*, 1997b) já demonstraram, *i.e.*, que a hidroxila da posição 6 do anel benzotiofeno é mais importante para a atividade biológica dos compostos do que a hidroxila na posição 4' da fenila.

A metodologia de LIV-3D-QSAR utiliza na construção dos modelos apenas uma conformação considerada “bioativa”, enquanto que a metodologia de QSAR-4D utiliza-se de múltiplas conformações, o que diminui a dificuldade em selecionar a conformação “bioativa”. Dessa forma, ao contrário do observado no modelo obtido pela metodologia de LIV-3D-QSAR por Cunha e colaboradores (2004), a conformação prevista como “bioativa” pelo modelo 1B9 de QSAR-4D para os compostos da série dos arilbenzotiofenos sintetizados por Grese e colaboradores (Grese *et al.*, 1997a) é muito semelhante àquela encontrada na

cristalografia por raios-X (Figura 23). O átomo de nitrogênio do anel piperidina se encontra numa posição muito próxima daquela observada no cristal, sendo a distância entre eles de 1,33Å para o composto mais potente da série (composto **1**), e de 0,29Å para o menos potente (composto **54**), distâncias estas calculadas após a sobreposição destes compostos com a estrutura de raios-X do raloxifeno, usando os átomos do alinhamento 1.

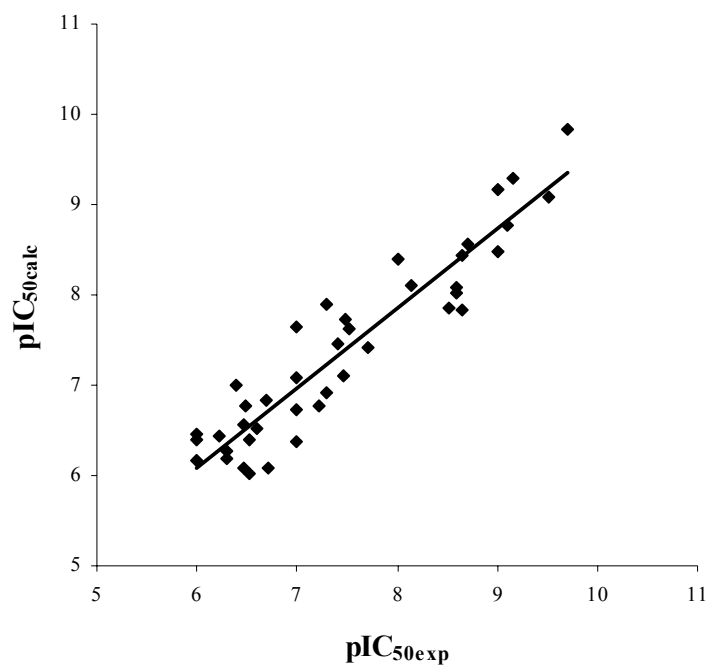
#### 4.3.1.1 Análise dos Valores Residuais e Identificação dos *Outliers* do Modelo Representativo do Alinhamento 1

A Tabela 9 contém os valores de atividade biológica experimentais ( $pIC_{50exp}$ ), os valores de atividade biológica calculados ( $pIC_{50calc}$ ) e os valores dos resíduos ( $pIC_{50calc} - pIC_{50exp}$ ) do modelo 1B9 de QSAR-4D. A Figura 26 apresenta o gráfico dos valores de  $pIC_{50exp}$  versus  $pIC_{50calc}$  e a Figura 27 o gráfico de barras dos valores dos resíduos das atividades dos 41 compostos da série em estudo.

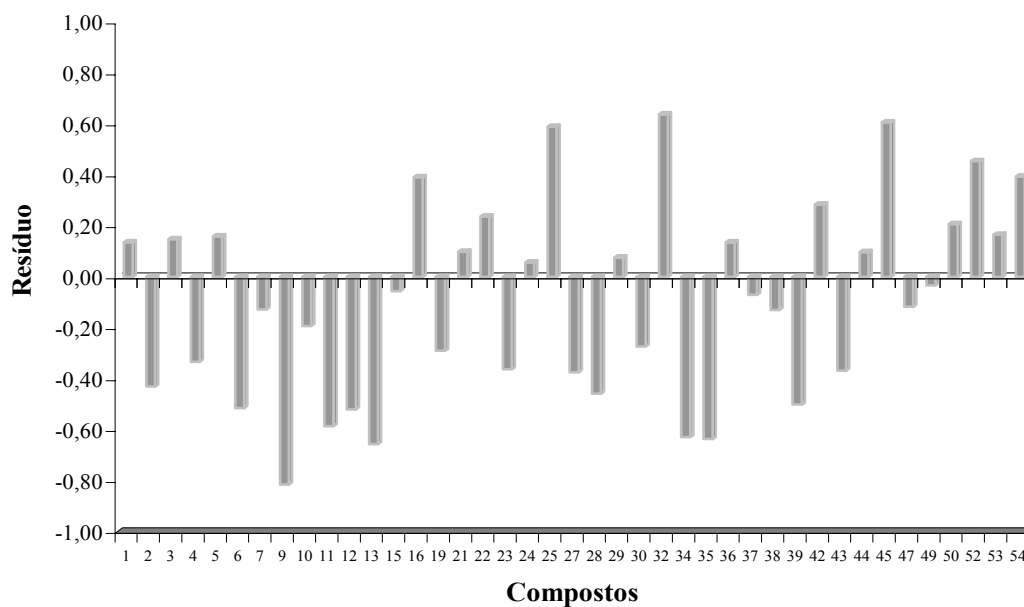
**Tabela 9.** Valores de atividade experimental ( $pIC_{50exp}$ ) e calculada ( $pIC_{50calc}$ ) e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) do modelo 1B9 de QSAR-4D para o conjunto de treinamento.

#	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>	#	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>
1	9,70	9,84	0,14	28	7,22	6,76	-0,46
2	9,52	9,09	-0,43	29	7,00	7,08	0,08
3	9,15	9,30	0,15	30	7,00	6,73	-0,27
4	9,10	8,77	-0,33	32	7,00	7,64	0,64
5	9,00	9,16	0,16	34	7,00	6,37	-0,63
6	9,00	8,49	-0,51	35	6,72	6,08	-0,64
7	8,70	8,57	-0,13	36	6,70	6,84	0,14
9	8,64	7,83	<b>-0,81</b>	37	6,60	6,53	-0,07
10	8,64	8,45	-0,19	38	6,52	6,39	-0,13
11	8,60	8,01	-0,59	39	6,52	6,02	-0,50
12	8,60	8,08	-0,52	42	6,49	6,77	0,28
13	8,52	7,86	-0,66	43	6,46	6,09	-0,37
15	8,15	8,09	-0,06	44	6,46	6,56	0,10
16	8,00	8,39	0,39	45	6,40	7,01	0,61
19	7,70	7,41	-0,29	47	6,30	6,18	-0,12
21	7,52	7,62	0,10	49	6,30	6,27	-0,03
22	7,49	7,73	0,24	50	6,22	6,43	0,21
23	7,46	7,10	-0,36	52	6,00	6,46	0,46
24	7,40	7,46	0,06	53	6,00	6,17	0,17
25	7,30	7,89	0,59	54	6,00	6,39	0,39
27	7,30	6,93	-0,37			$DP_{res} =$	0,38

<sup>a</sup> Os valores residuais em negrito correspondem aos compostos *outliers* que são os compostos com valor residual maior do que o dobro do  $DP_{res}$ .



**Figura 26.** Valores de atividade experimentais ( $pIC_{50exp}$ ) versus valores de atividade calculados ( $pIC_{50calc}$ ) dos 41 análogos do raloxifeno utilizados para a construção do modelo 1B9 de QSAR-4D.



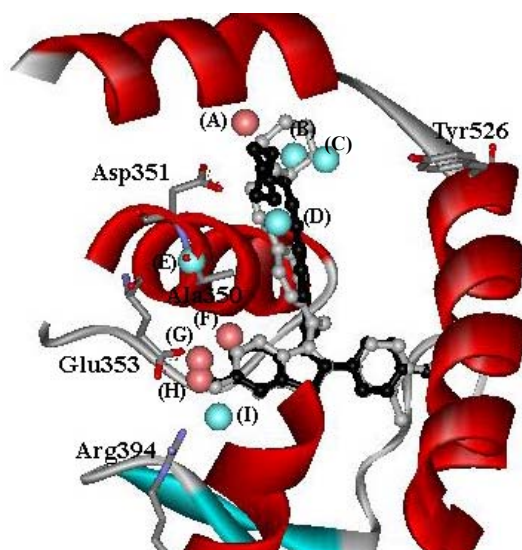
**Figura 27.** Valores dos resíduos das potências dos 41 análogos do raloxifeno utilizados para a construção do modelo 1B9 de QSAR-4D.

Como definido no item 3.13, os compostos denominados *outliers* são aqueles cujos valores de resíduos são maiores do que o dobro do valor de  $DP_{res}$  (Albuquerque *et al.*, 1998) e não podem ser explicados pelo modelo (Kubinyi 1993). Para o conjunto de treinamento utilizado, foi encontrado um valor de  $DP_{res}$  de 0,38, sendo o composto **9** (Figura 28) considerado um *outlier*. A única diferença entre este composto e o raloxifeno (composto **1**) é a presença de um átomo de cloro na posição 3' da fenila ligada ao anel benzotiofeno.

A atividade biológica calculada para o composto **9** é menor do que a experimental, ou seja, a equação proposta pelo modelo 1B9 prediz uma potência menor para o composto do que a real. Isto pode ser decorrente de uma limitação do modelo, que não apresenta descritores ao redor do grupamento fenila dos compostos da série em estudo ou da existência de poucos compostos com substituintes na posição 3', levando o modelo a subestimar a potência calculada.

Outra possível explicação para o composto **9** ser um *outlier* é devido a lipofilicidade da molécula, descritor que não é considerado na metodologia de QSAR-4D. A lipofilicidade é definida pelo logaritmo do coeficiente de partição ( $\log P$ ) de uma substância entre as fases orgânica e aquosa. Fármacos com valores elevados de  $\log P$  ultrapassam com maior facilidade as biomembranas hidrofóbicas, podendo apresentar um melhor perfil farmacológico (Barreiro & Fraga, 2001; Leo, 1993). Os valores de  $\log P$  calculados ( $ClogP$ ) para os compostos que apresentam um grupo hidroxila na posição 4' da fenila e um grupo hidroxila nas posições 4, 5 ou 7 do anel benzotiofeno não variam, como por exemplo, os compostos **31**, **35** e **41** cujo valor de  $ClogP$  corresponde a 5,96 (Gao *et al.*, 1999). Contudo, a inserção de um átomo de cloro na posição 3' da fenila aumenta a lipofilicidade do composto **9** ( $ClogP = 6,63$ ) (Gao *et al.*, 1999). Além disso, estudos de QSAR para compostos relacionados ao raloxifeno, incluindo a série em estudo, corroboram esta hipótese, pois demonstraram correlações

positivas significativas entre potência e hidrofobicidade, refletindo a capacidade de penetração das moléculas através das membranas celulares (Gao *et al.*, 1999).



**Figura 28.** Representação da conformação “bioativa” do composto 9 (cinza), identificado como *outlier*, obtida pelo modelo 1B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 1. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).

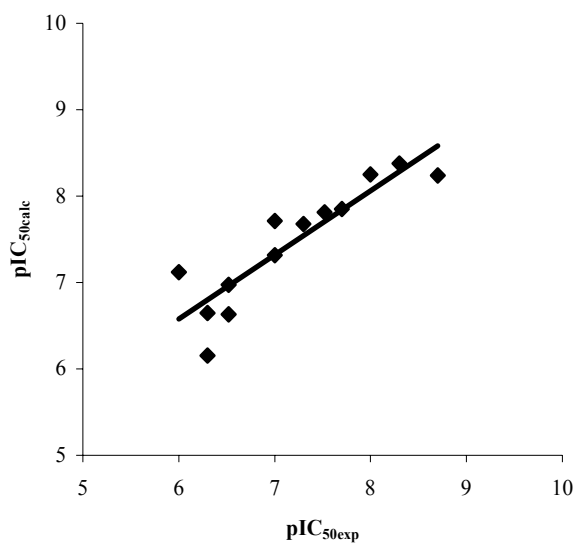
#### 4.3.1.2 Validação Externa do Modelo Representativo do Alinhamento 1

Os 13 compostos da série (Grese *et al.*, 1997a) que não foram incluídos na construção do modelo de QSAR-4D (conjunto de teste) foram utilizados para realizar uma predição “real” do modelo 1B9. A Tabela 10 lista os valores de atividade biológica experimentais ( $pIC_{50exp}$ ) e calculados ( $pIC_{50calc}$ ) e os valores dos resíduos do conjunto de teste do modelo 1B9 de QSAR-4D. A Figura 29 apresenta o gráfico dos valores de  $pIC_{50exp}$  *versus*  $pIC_{50calc}$  e a Figura 30, o gráfico de barras dos resíduos das atividades dos 13 compostos do conjunto de teste.

**Tabela 10.** Valores de atividade experimental e calculada e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ), do modelo 1B9 de QSAR-4D para o conjunto de teste.

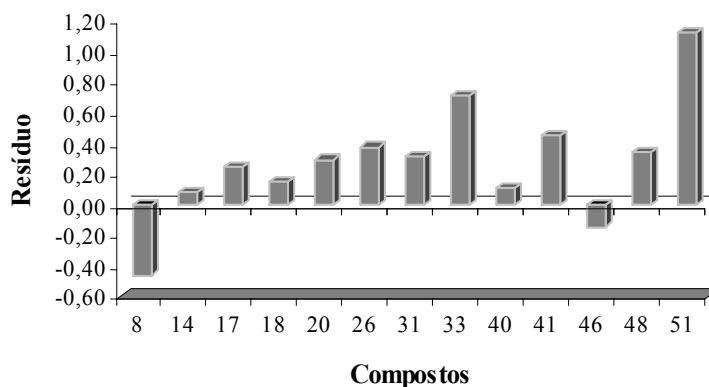
Composto	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>
8	8,70	8,24	-0,46
14	8,30	8,38	0,08
17	8,00	8,25	0,25
18	7,70	7,85	0,15
20	7,52	7,81	0,29
26	7,30	7,68	0,38
31	7,00	7,31	0,31
33	7,00	7,71	0,71
40	6,52	6,63	0,11
41	6,52	6,97	0,45
46	6,30	6,15	-0,15
48	6,30	6,65	0,35
51	6,00	7,12	<b>1,12</b>
			$DP_{res} =$ 0,38

<sup>a</sup> Os valores residuais em negrito correspondem aos compostos *outliers* que são os compostos com valor residual maior do que o dobro do  $DP_{res}$ .



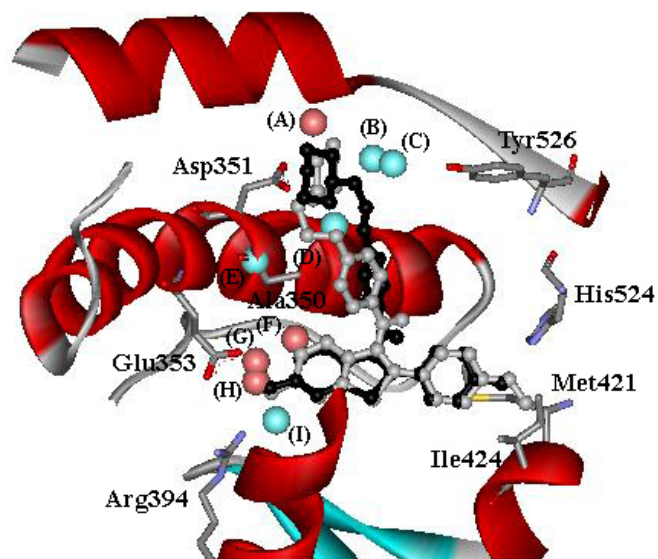
**Figura 29.** Valores de atividade experimentais ( $pIC_{50exp}$ ) versus valores de atividade calculados ( $pIC_{50calc}$ ) dos 13 compostos utilizados como conjunto de teste para a validação externa do modelo 1B9 de QSAR-4D.





**Figura 30.** Valores dos resíduos dos 13 compostos do conjunto de teste utilizados para a validação externa do modelo 1B9 de QSAR-4D.

O valor de  $DP_{res}$  encontrado para os 13 compostos do conjunto de teste foi de 0,38, indicando o composto **51** como *outlier*. A diferença da estrutura química entre o composto **51** para o composto **1** (raloxifeno) é a presença do grupo metoxila na posição 4' da fenila ligada ao anel benzotiofeno, enquanto que o raloxifeno apresenta um grupo hidroxila nesta posição. A atividade biológica do composto **51** foi superestimada pelo modelo 1B9, fato que pode ser decorrente de uma limitação do modelo, que não apresenta descritores ao redor do grupamento fenila ligado à posição 2 do anel benzotiofeno dos compostos da série em estudo. Dessa forma, não seria caracterizada alguma possível interação prejudicial do grupo metoxila com os resíduos ao seu redor (Figura 31), pois estudos de QSAR clássico demonstraram que há um efeito estérico negativo para substituintes na substituição 4' da fenila (Gao *et al.*, 1999).



**Figura 31.** Representação da conformação “bioativa” do composto **51** (cinza), identificado como *outlier*, obtida pelo modelo 1B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 1. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).

#### 4.4 ANÁLISE DOS MELHORES MODELOS DE QSAR-4D DO ALINHAMENTO 2

No alinhamento 2, assim como para o alinhamento 1, os melhores modelos foram obtidos em células de 1Å, visto que estes modelos apresentam valores de  $R^2_{\text{ajustado}}$  e  $q^2_{\text{ajustado}}$  maiores do que aqueles obtidos em células de 2Å. Na Tabela 11, encontram-se as melhores equações de QSAR-4D com 7 e 9 termos, obtidas pelo alinhamento 2 em células de 1Å.

**Tabela 11.** Melhores modelos do alinhamento 2 em células de 1Å.

Modelo	Termos	Equação <sup>a</sup>
2B7	7	$pIC_{50} = 7,70 + 2,47 (-1,9,4) (qa) - 10,58 (0,10,4) (qa) + 8,22 (0,7,-2) (qa) - 5,27 (-1,-2,11) (qa) - 35,20 (0,5,-3) (qa) + 21,70 (0,2,2) (ar) + 0,84 (0,-1,9) (ar)$
2B8	8	$pIC_{50} = 8,34 - 6,52 (-3,9,4) (qa) - 9,04 (0,10,4) (qa) - 7,64 (0,2,-1) (qa) - 5,08 (-1,-2,11) (qa) - 9,03 (0,1,-2) (qa) + 5,37 (-3,8,5) (qa) - 26,15 (0,5,-3) (qa) + 23,89 (0,2,2) (ar)$
2B9	9	$pIC_{50} = 6,66 - 30,48 (-1,12,6) (qa) + 10,92 (1,3,6) (alh) + 16,70 (0,11,3) (p-) - 15,21 (0,12,-1) (qa) + 14,18 (0,10,-2) (qa) + 0,57 (0,0,-2) (qa) - 4,01 (-1,-2,11) (qa) + 18,23 (0,2,2) (ar) - 25,70 (2,5)(qa) (qa)$

<sup>a</sup> Os números entre parênteses representam as coordenadas cartesianas x,y,z dos *IPEs*; as letras entre parênteses representam a classificação dos *IPEs*: qa = qualquer tipo de átomo, ar = aromático, alh = acceptor de ligação hidrogênio e ap = apolar.

A matriz de correlação dos resíduos das atividades dos modelos com 7, 8 e 9 termos (2B7, 2B8 e 2B9, respectivamente) foi calculada com o intuito de eliminar modelos com o mesmo tipo de informação sobre os compostos do conjunto de treinamento e está apresentada na Tabela 12.

**Tabela 12.** Matriz de correlação dos resíduos das atividades dos modelos 2B7, 2B8 e 2B9 de QSAR-4D, obtidos com célula de 1Å.

Modelos	2B7	2B8	2B9
2B7	1		
2B8	0,80	1	
2B9	0,73	0,58	1

Como pode ser observado na Tabela 12, existe alta correlação entre os resíduos de atividade dos três modelos, provavelmente em razão da presença de muitas células idênticas, em coordenadas espaciais e tipo de ocupação. As células (-1,-2,11)(qa) e (0,2,2)(ar) estão

presentes nos três modelos, enquanto que as células (0,10,4)(qa) e (0,5,-3)(qa) estão presentes nos modelos 2B7 e 2B8, i.e., o par de modelos 2B7/2B8 apresentam quatro células idênticas, enquanto que os pares de modelos 2B7/2B9 e 2B8/2B9 apresentam duas células idênticas.

O modelo 2B7 apresenta maior correlação com os outros modelos, em especial com o modelo 2B8 ( $R = 0,80$ ). Numa primeira análise, o modelo 2B7 incorpora quantitativamente os demais modelos, pois apresenta mais células idênticas com as outras equações. Entretanto, os modelos 2B8 e 2B8 estão menos correlacionados entre si ( $R = 0,58$ ) e, ao mesmo tempo, altamente correlacionados ao modelo 2B7.

Dessa forma, outros critérios foram adotados para seleção do modelo representativo do alinhamento 2, como a semelhança entre as células dos modelos, o valor de  $DP_{res}$  e o número de *outliers*, além dos parâmetros estatísticos já discutidos anteriormente.

A matriz de correlação cruzada entre os valores de ocupação das células dos modelos 2B7, 2B8 e 2B8 foi calculada (Anexo IV), observando-se alta correlação entre células de modelos diferentes, o que justifica mais uma vez a semelhança entre os modelos, mas não determina qual modelo melhor representaria o alinhamento 2.

A matriz de correlação cruzada entre os valores de atividade biológica experimental e a frequência de ocupação das células dos modelos 2B7, 2B8 e 2B9 (Anexo IV) indicam que as células que apresentam maior correlação com a atividade são aquelas encontradas no modelo 2B9. Estas células são (1,3,6)(alh) e (0,2,2)(ar), sendo que a célula (0,2,2)(ar) existe nos três modelos. Além disso, este par de *GCODs* apresenta alguma correlação entre si ( $R = 0,44$ ) (Anexo IV).

Comparando os valores de  $DP_{res}$  e o número de *outliers* (Tabelas 14 e Anexos V e VI) dos três modelos, observamos que o modelo 2B9, apesar de apresentar o mesmo número de compostos *outliers* considerando o conjunto de treinamento, apresenta menor valor de  $DP_{res}$

(Tabela 14). Além disso, este modelo também possui valores de  $R^2_{ajustado}$  e  $q^2_{ajustado}$  maiores, sendo selecionado como o modelo representativo do alinhamento 2.

#### 4.4.1 Análise do Modelo Representativo do Alinhamento 2

Como modelo representativo do alinhamento 2, selecionou-se o modelo 2B9, descrito na Equação (8).

$$\begin{aligned} pIC_{50} = & 6,66 - 30,48 (-1,12,6) (qa) + 10,92 (1,3,6) (alh) & \text{Equação (8)} \\ & + 16,70 (0,11,3) (p-) - 15,21 (0,12,-1) (qa) + 14,18 (0,10,-2) (qa) \\ & + 0,57 (0,0,-2) (qa) - 4,01 (-1,-2,11) (qa) + 18,23 (0,2,2) (ar) \\ & - 25,70 (2,5)(qa) (qa) \end{aligned}$$

$$N = 41 \quad R^2 = 0,81 \quad q^2 = 0,72 \quad F = 14,23 \quad SE = 0,52$$

A matriz de correlação cruzada das ocupações de todos os *GCODs* do modelo 2B9 foi calculada com o intuito de verificar se algum deles apresenta informações iguais, ou seja, informações redundantes na equação. A Tabela 13 contém a correlação entre os *GCODs* do modelo 2B9.

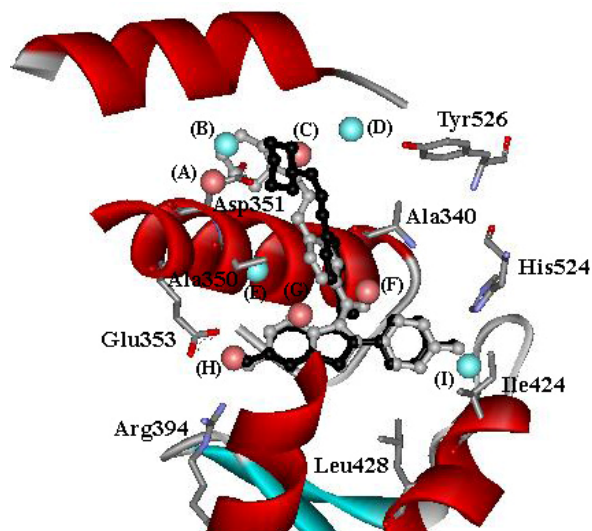
**Tabela 13.** Matriz de correlação cruzada entre os *GCODs* do modelo2B9 obtido pelo método de QSAR-4D em célula de 1Å.

<i>GCODs</i>	(-1,12,6) (qa)	(1,3,6) (alh)	(0,11,3) (p-)	(0,12,-1) (qa)	(0,10,-2) (qa)	(0,0,-2) (qa)	(-1,-2,11) (qa)	(0,2,2) (ar)	(2,5,0) (qa)
(-1,12,6) (qa)	1,00								
(1,3,6) (alh)	0,03	1,00							
(0,11,3) (p-)	0,13	-0,19	1,00						
(0,12,-1) (qa)	-0,11	-0,03	0,42	1,00					
(0,10,-2) (qa)	-0,06	0,28	0,29	0,73	1,00				
(0,0,-2) (qa)	-0,07	0,24	0,16	0,13	0,23	1,00			
(-1,-2,11) (qa)	0,13	0,02	-0,21	-0,16	-0,14	-0,30	1,00		
(0,2,2) (ar)	0,10	0,47	0,06	0,03	0,06	0,12	0,30	1,00	
(2,5,0) (qa)	-0,10	-0,01	-0,05	-0,30	-0,26	0,08	0,00	-0,07	1,00

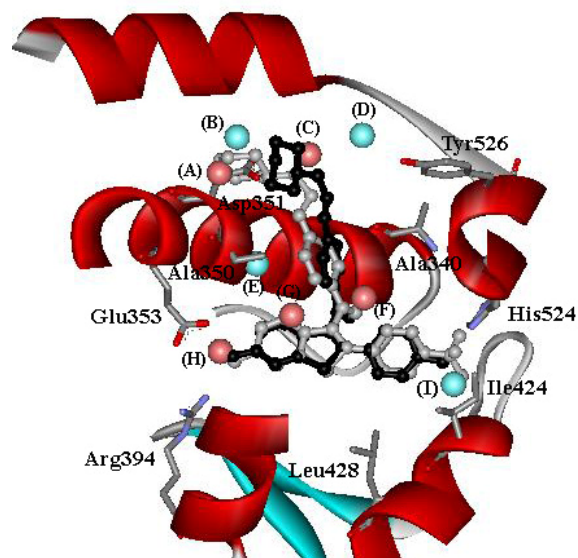
Observa-se na Tabela 13 que, com exceção de apenas dois pares de células, os demais descritores são praticamente ortogonais entre si e contribuem de forma distinta para os modelos de QSAR-4D, pois não apresentam valores de correlação significativa (ou seja, valor maior que 0,5) (Kubinyi, 1993; Pitta, 2006). O par de células (0,10,-2)(qa) e (0,12,-1)(qa) apresentam alta correlação entre si ( $R=0,73$ ). Embora a ocupação da célula (0,10,-2)(qa) favoreça o aumento da potência dos compostos e a ocupação da célula (0,12,-1)(qa) desfavoreça, estes *GCODs* estão espacialmente próximos, distantes 2,24Å, o que justificaria a correlação entre eles. A célula (0,10,-2)(qa) apresenta uma ambigüidade, pois se localiza numa região da caixa 3D que corresponde à posição ocupada pelo resíduo Asp351 (Figura 32 e 33). Logo, não é esperado que ocupação deste *GCOD* aumente a potência do composto. Este fato demonstra que o modelo 2B9 não é capaz de “predizer” a presença do resíduo Asp351 e a forte interação eletrostática atrativa que ocorre entre este resíduo ácido e o grupo piperidina básico do raloxifeno, como observado na estrutura 3D do raloxifeno complexado ao *ER- $\alpha$*  obtida por cristalografia de raios-X (Brzozowski, 1997). Isto pode ser uma explicação para o

fato do composto **1**, o mais potente da série, ter o valor de  $pIC_{50}$  calculado menor do que o valor experimental (Tabela 14).

O segundo par de *GCODs* relativamente correlacionado ( $R = 0,47$ ) corresponde às células (0,2,2)(ar) e (1,3,6)(alh). Ao contrário do caso anterior, esses descritores se encontram distantes entre si no espaço (4,24Å) e possuem o mesmo tipo de contribuição para o modelo, ou seja, ambos favorecem o aumento da potência de um composto quando ocupado. A célula (0,2,2)(ar) localiza-se próximo aos átomos de carbono e de hidrogênio da posição 4 do anel benzotiofeno, enquanto que a célula (1,3,6)(alh) localiza-se próximo ao átomo de oxigênio da carbonila presente nos compostos. Esses *GCODs* estão relacionados ao ângulo diedro formado pelo plano da carbonila e o plano do anel benzotiofeno, indicando a importância da coplanaridade da cadeia lateral, como descrito por Grese e colaboradores (Grese *et al.*, 1997a), pois as moléculas que mais ocupam estas células são, em geral, aquelas que possuem substituintes nas posições 2' e 3' da fenila, como por exemplo, os compostos **3**, **7** e **16** (Anexo VII). Substituintes nestas posições apresentam repulsão estérica com o átomo de oxigênio da carbonila, levando à não coplanaridade da cadeia lateral.



**Figura 32.** Representação da conformação “bioativa” do composto **1** (cinza) obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER- $\alpha$*  (código *PDB* 1ERR), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,11,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).



**Figura 33.** Representação da conformação “bioativa” do composto **54** (cinza) obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER- $\alpha$*  (código *PDB* 1ERR), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,11,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).



Os *GCODs* (-1,12,6)(qa) e (0,11,3)(p-), localizados próximos ao anel piperidina, demonstram a importância da orientação da cadeia lateral básica dos compostos da série em estudo. A ocupação da célula (-1,12,6) por qualquer tipo de átomo leva a uma diminuição da potência dos compostos, desviando o átomo de nitrogênio protonado do grupo piperidina da posição favorável em relação à formação da ligação hidrogênio com o resíduo Asp351.

A ocupação da célula (0,11,3)(p-) também se relaciona ao átomo de nitrogênio da cadeia lateral, pois compostos que possuem substituintes na posição 4 do anel benzotiofeno ou na posição 3' da fenila ocupam bastante esta célula (Anexo VII), como por exemplo os compostos **5**, **9**, **35** e **44**, indicando que estes substituintes são capazes de manter a cadeia lateral em uma orientação favorável para a atividade biológica antagonista no *ER- $\alpha$* .

A célula (2,5,0)(qa), localizada próxima ao resíduo de Ala350, como também observado para o modelo 1B9, indica que a cadeia lateral dos compostos deve apresentar uma determinada orientação que não ocupe esta região no espaço, o que provocaria um impedimento estérico entre a cadeia lateral e o resíduo Ala350, como também foi observado para o modelo 1B9 do alinhamento 1.

O *GCOD* (0,0,-2)(qa), localizado próximo ao resíduo Glu353, indica a importância da ligação hidrogênio deste resíduo com o substituinte da posição 6 do anel benzotiofeno, pois apresenta pequena ocupação quando não há substituintes nesta posição (*e.g.* compostos **23** e **30**) ou substituintes incapazes de promover esse tipo de interação (*e.g.* compostos **19** e **40**) (Anexo VII).

A célula (-1,-2,11)(qa), localizada próxima à posição 4' do anel fenila ligado ao carbono 2 do anel benzotiofeno, está relacionada aos contatos apolares em torno desta fenila dos compostos estudados. Como esta região está envolta por resíduos hidrofóbicos (*e.g.* os resíduos Ile424, Gly521 e Leu525), compostos com substituintes volumosos na posição 4' da fenila, como por exemplo os compostos **22**, **36** e **39**, apresentam impedimento estérico,

ocupando muito este *GCOD* (Anexo VII) e, conseqüentemente, diminuindo a potência do composto.

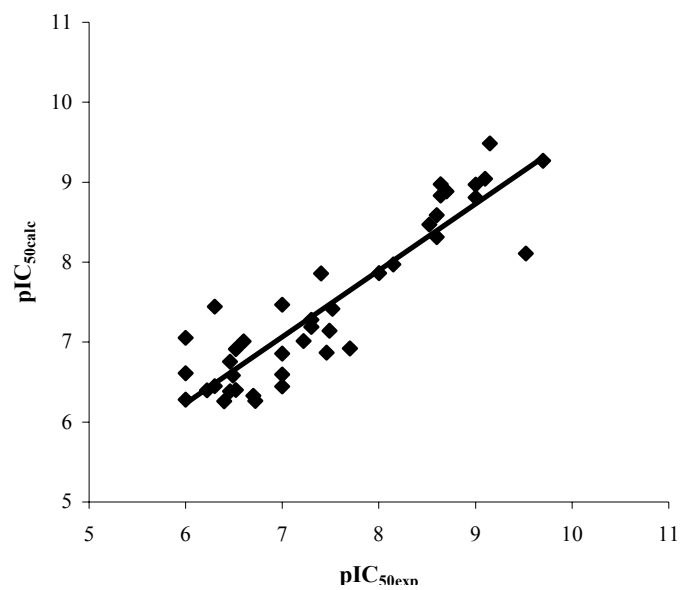
#### 4.4.1.1 Análise dos Valores Residuais e Identificação dos *Outliers* do Modelo Representativo do Alinhamento 2

Os valores de atividade biológica experimentais ( $pIC_{50exp}$ ) e calculados ( $pIC_{50calc}$ ) e os valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ), referentes ao modelo 2B8 de QSAR-4D, encontram-se na Tabela 14. A Figura 34 apresenta o gráfico dos valores de  $pIC_{50exp}$  *versus*  $pIC_{50calc}$  e a Figura 35 apresenta o gráfico de barras dos valores dos resíduos das atividades dos 41 compostos da série em estudo.

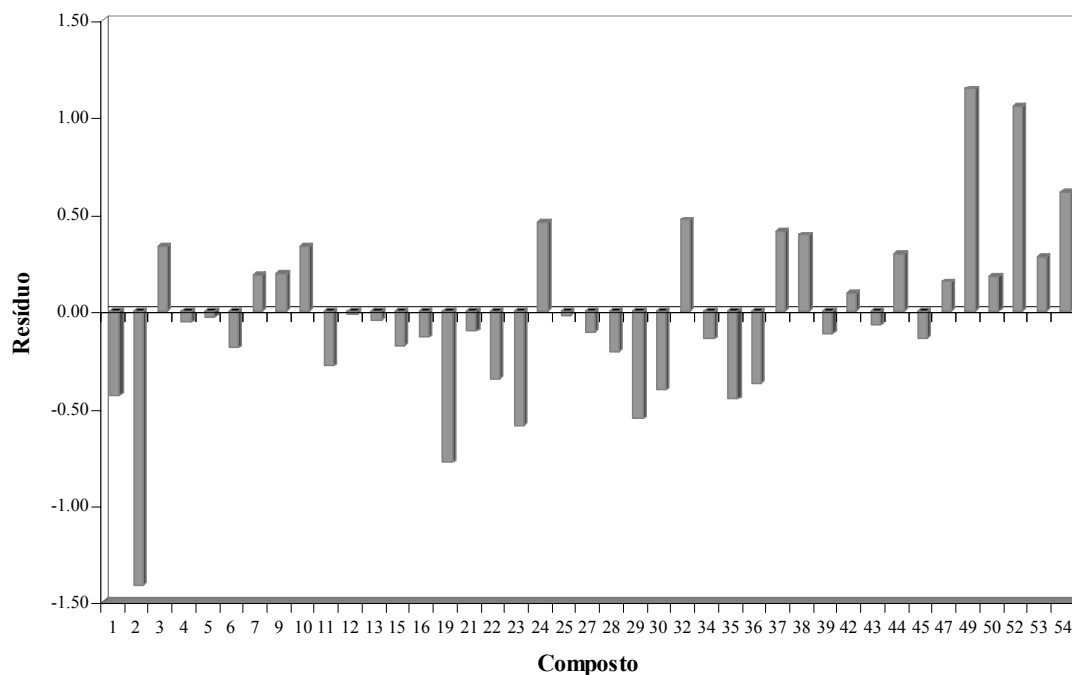
**Tabela 14.** Valores de atividade experimental ( $pIC_{50exp}$ ) e calculada ( $pIC_{50calc}$ ) e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) do modelo 2B9 de QSAR-4D para o conjunto de treinamento.

#	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>	#	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>
1	9,70	9,27	-0,43	28	7,22	7,01	-0,21
2	9,52	8,11	<b>-1,41</b>	29	7,00	6,45	-0,55
3	9,15	9,48	0,33	30	7,00	6,59	-0,41
4	9,10	9,04	-0,06	32	7,00	7,47	0,47
5	9,00	8,97	-0,03	34	7,00	6,86	-0,14
6	9,00	8,81	-0,19	35	6,72	6,27	-0,45
7	8,70	8,89	0,19	36	6,70	6,33	-0,37
9	8,64	8,83	0,19	37	6,60	7,01	0,41
10	8,64	8,97	0,33	38	6,52	6,91	0,39
11	8,60	8,32	-0,28	39	6,52	6,40	-0,12
12	8,60	8,59	-0,01	42	6,49	6,58	0,09
13	8,52	8,47	-0,05	43	6,46	6,39	-0,07
15	8,15	7,97	-0,18	44	6,46	6,76	0,30
16	8,00	7,86	-0,14	45	6,40	6,26	-0,14
19	7,70	6,92	-0,78	47	6,30	6,45	0,15
21	7,52	7,41	-0,11	49	6,30	7,44	<b>1,14</b>
22	7,49	7,14	-0,35	50	6,22	6,40	0,18
23	7,46	6,87	-0,59	52	6,00	7,05	<b>1,05</b>
24	7,40	7,86	0,46	53	6,00	6,28	0,28
25	7,30	7,28	-0,02	54	6,00	6,61	0,61
27	7,30	7,19	-0,11			$DP_{res} =$	0,46

<sup>a</sup> Os valores residuais em negrito correspondem aos compostos *outliers* que são os compostos com valor residual maior do que o dobro do  $DP_{res}$ .



**Figura 34.** Valores de atividade biológica experimentais ( $pIC_{50exp}$ ) versus calculados ( $pIC_{50calc}$ ) dos 41 derivados do raloxifeno utilizados para a construção do modelo 2B9 de QSAR-4D.



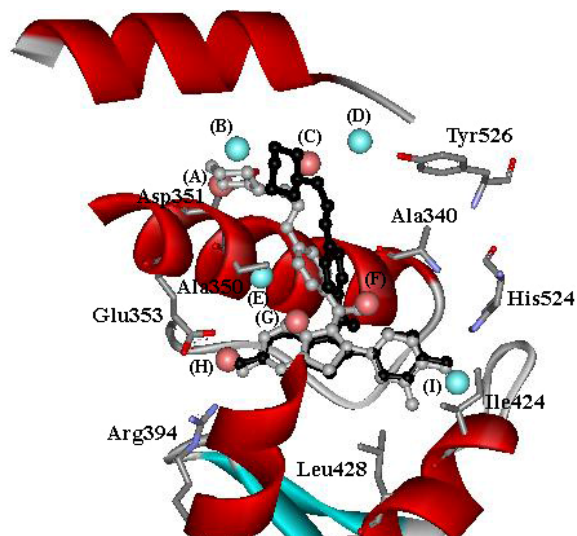
**Figura 35.** Valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) dos 41 derivados do raloxifeno utilizados para a construção do modelo 2B9 de QSAR-4D.

Considerando os compostos *outliers* como sendo aqueles que apresentam valores de resíduos maiores do que o dobro do  $DP_{res}$  (Albuquerque *et al.*, 1998), foram identificados como *outliers* três compostos para o conjunto de treinamento: **2**, **49** e **52** ( $DP_{res}=0,46$ ).

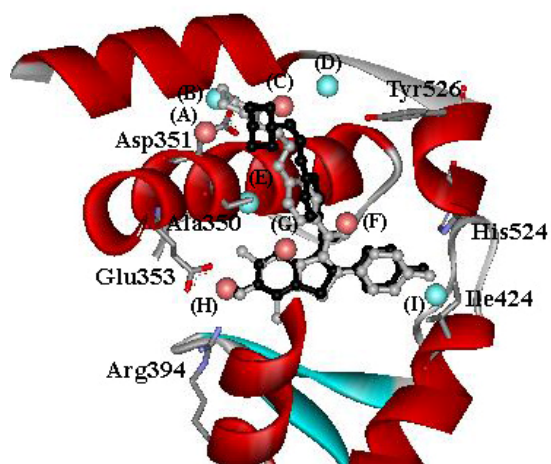
A atividade biológica calculada para o composto **2**, que difere do raloxifeno por apresentar um átomo de flúor na posição 3' da fenila ligada ao anel benzotiofeno (Figura 36), é menor do que a experimental, ou seja, a equação proposta pelo modelo 2B9 prediz uma potência menor para o composto do que a real. Da mesma forma que ocorreu com o modelo 1B9, este fato pode ser devido a existência de poucos compostos com substituintes na posição 3' da fenila, levando o modelo a não revelar a importância dos substituintes nesta posição. Além disso, a maior lipofilicidade do composto **2** ( $ClogP= 6,19$ ) (Gao *et al.*, 1999), frente ao raloxifeno, pode levar o modelo a subestimar sua potência.

Os compostos **49** e **52** apresentaram valores calculados de atividade biológica maiores do que os respectivos valores experimentais. O composto **49** se diferencia quimicamente do raloxifeno pela presença do grupo metila nas posições 5 e 7 do anel benzotiofeno. Isto pode ser devido ao fato que não existem células para descrever estes substituintes nas respectivas posições, visto que estes substituintes não são muito representados nos compostos do conjunto de treinamento.

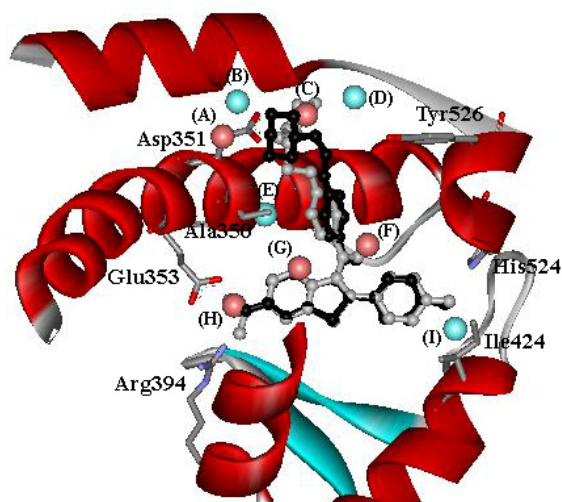
O composto **52** apresenta um grupo amida na posição 6 do anel benzotiofeno. Este grupo torna a célula (0,0,-2)(qa) muito ocupada, levando o modelo a favorecer a potência do composto.



**Figura 36.** Representação da conformação “bioativa” do composto 2 (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).



**Figura 37.** Representação da conformação “bioativa” do composto 49 (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).



**Figura 38.** Representação da conformação “bioativa” do composto **52** (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).

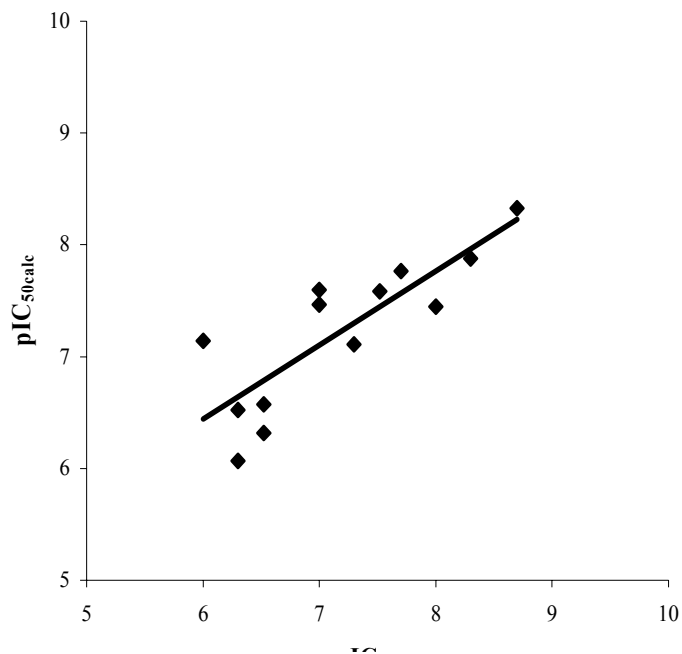
#### 4.4.1.2 Validação Externa do Modelo Representativo do Alinhamento 2

Foram usados 13 compostos da série em estudo (conjunto de teste) (Grese *et al.*, 1997a), não incluídos na construção do modelo 2B9 de QSAR-4D, para realizar a validação externa. A Tabela 15 contém os valores de atividade biológica experimentais ( $pIC_{50exp}$ ) e calculados ( $pIC_{50calc}$ ) e os valores dos resíduos para o conjunto de teste do modelo 2B9 de QSAR-4D. O gráfico dos valores de  $pIC_{50exp}$  versus  $pIC_{50calc}$  é apresentado na Figura 39 e os valores residuais dos 13 compostos do conjunto de teste são apresentados na forma de gráfico de barras na Figura 40.

**Tabela 15.** Valores de atividade experimental e calculada e valores residuais ( $pIC_{50calc} - pIC_{50exp}$ ) do modelo 2B9 de QSAR-4D para o conjunto de teste.

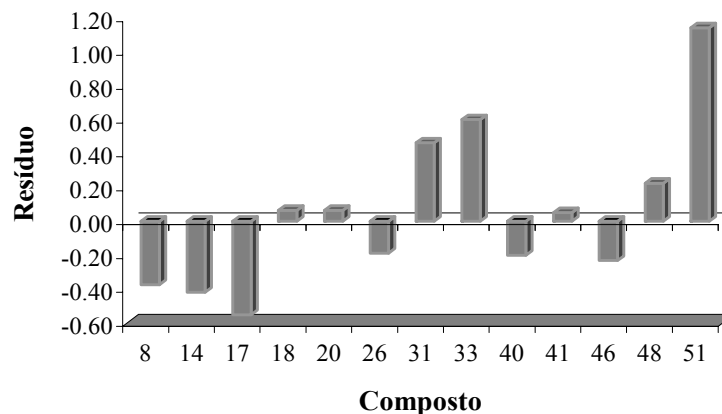
Composto	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>
8	8,70	8,32	-0,38
14	8,30	7,88	-0,42
17	8,00	7,45	-0,55
18	7,70	7,76	0,06
20	7,52	7,58	0,06
26	7,30	7,11	-0,19
31	7,00	7,46	0,46
33	7,00	7,60	0,60
40	6,52	6,32	-0,20
41	6,52	6,57	0,05
46	6,30	6,07	-0,23
48	6,30	6,52	0,22
51	6,00	7,14	<b>1,14</b>

<sup>a</sup> Os valores residuais em negrito correspondem aos compostos *outliers* que são os compostos com valor residual maior do que o dobro do  $DP_{res}$  ( $DP_{res} = 0,47$ )



**Figura 39.** Valores de atividade experimentais ( $pIC_{50exp}$ ) versus valores de atividade calculados ( $pIC_{50calc}$ ) dos 13 compostos utilizados como conjunto de teste para a validação externa do modelo 2B9 de QSAR-4D.

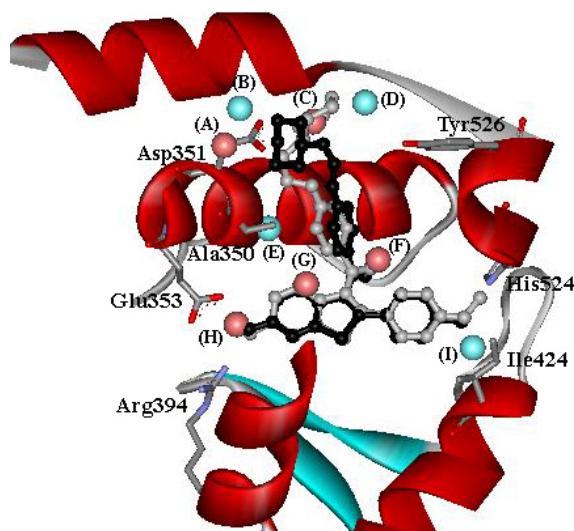




**Figura 40.** Valores residuais dos 13 compostos do conjunto de teste utilizados para a validação externa do modelo 2B9 de QSAR-4D.

O composto **51**, que apresenta um grupamento metóxi na posição 4' da fenila ligada ao anel benzotiofeno, foi identificado como *outlier* (resíduo de 1,14), sendo sua atividade biológica superestimada pelo modelo 2B9. Observa-se que este mesmo composto também comportou-se como *outlier* no modelo 1B9, no alinhamento 1.

De forma distinta do modelo representativo do alinhamento 1, o modelo 2B9 apresenta um descritor próximo à posição 4' da fenila, que confirma dados de QSAR sobre o efeito estérico negativo para substituintes nesta posição (Gao *et al.*, 1999). Este *GCOD* é bastante ocupado pelo composto **51** (Anexo VII), mas o mesmo contribui pouco para o modelo, devido ao pequeno valor de seu coeficiente quando comparado aos dos demais *GCODs* (Figura 41).



**Figura 41.** Representação da conformação “bioativa” do composto **51** (cinza), identificado como *outlier*, obtida pelo modelo 2B9 de QSAR-4D sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 2. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (0,10,-2)(qa); (B) (0,12,-1)(qa); (C) (0,13,3)(p-); (D) (-1,12,6)(qa); (E) (2,5)(qa)(qa); (F) (1,3,6)(alh); (G) (0,2,2)(ar); (H) (0,0,-2)(qa); (I) (-1,-2,11)(qa).

#### 4.5 NOVOS COMPOSTOS BASEADOS NA ANÁLISE DE QSAR-4D

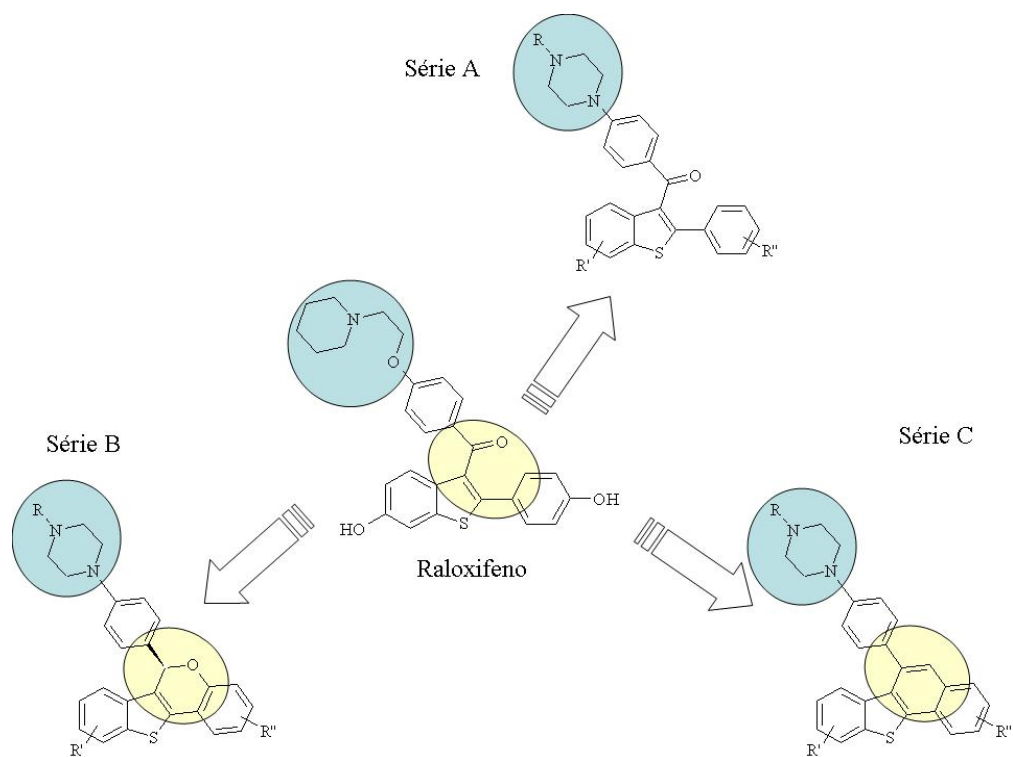
Embora o alinhamento 2 tenha sido realizado com átomos de regiões mais rígidas da molécula e apresente descritores espaciais mais distribuídos ao redor dos compostos, o alinhamento 1, além de apresentar índices estatísticos superiores, apresenta um modelo mais coerente com o mecanismo de ação proposto para o raloxifeno. Assim, baseado nos resultados obtidos pelo Modelo 1B9 de QSAR-4D e no modelo de LIV-QSAR-3D de por Cunha e colaboradores (2004), propõe-se a síntese de novos potenciais candidatos a *SERMs*.

Na química medicinal, a otimização da estrutura de compostos protótipos (*lead compounds*) ocorre por dois métodos principais (Becker *et al.*, 2000). O primeiro é baseado em modificações químicas da estrutura molecular, ou seja, modificações que alteram, essencialmente, suas propriedades químicas, e o segundo por aplicação de restrições conformacionais, que alteram a flexibilidade molecular.

As modificações químicas foram exaustivamente exploradas por Grese e colaboradores (Grese *et al.*, 1997a) na série de derivados do raloxifeno. Dessa forma, sugere-se neste trabalho a alteração da cadeia lateral do raloxifeno para a redução da sua flexibilidade. Propõe-se a substituição do grupo piperidinil-etóxi pelo grupo piperazina, que é capaz de interagir tanto por interação eletrostática como por ligação hidrogênio com o resíduo Asp351, pois este grupo também possui um átomo de nitrogênio básico numa distância aproximadamente igual à do raloxifeno. Esta modificação faz parte das três séries de compostos propostas, denominadas A, B e C.

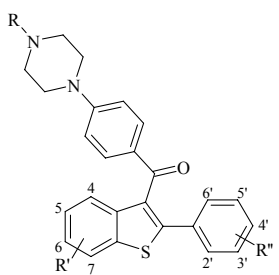
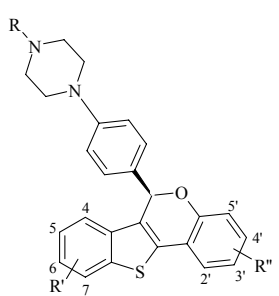
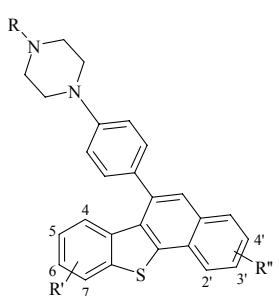
Ainda no intuito de tornar a orientação da cadeia lateral mais favorável ao antagonismo sobre o *ER- $\alpha$* , foram inseridos um grupo metila ou um átomo de cloro nas posições 2' e 6' ou 3' e 5' da fenila ligada ao anel benzotiofeno, na série A, como proposto neste trabalho e também por Cunha e colaboradores (Cunha *et al.*, 2004).

Nas séries B e C substituiu-se o grupo fenila pelos anéis benzopirano e nafto, respectivamente. A Figura 42 apresenta o perfil das modificações propostas para o raloxifeno e a Tabela 16 contém os compostos sugeridos para posterior síntese e avaliação farmacológica em *ER- $\alpha$* , além dos valores de  $pIC_{50}$  preditos pelo modelo 1B9 de QSAR-4D.



**Figura 42.** Representação esquemática das modificações estruturais propostas para o raloxifeno, baseadas no modelo 1B9 de QSAR-4D. As letras R, R' e R'' correspondem aos diferentes substituintes presentes nos análogos propostos.

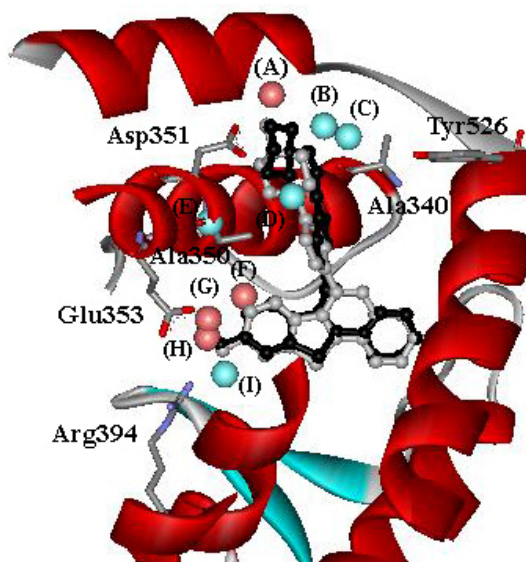
**Tabela 16.** Novos compostos sugeridos para síntese e avaliação como *SERMs*.

Série	R	R'	R''	#	pIC <sub>50</sub>
	-H	6-OH	4'-OH	A1	7,19
	-Me	6-OH	4'-OH	A2	6,60
	-H	6-OH	2'-Me, 4'-OH, 6'-Me	A3	7,67
	-Me	6-OH	2'-Me, 4'-OH, 6'-Me	A4	7,57
	-H	6-OH	2'-Cl, 4'-OH, 6'-Cl	A5	8,17
	-Me	6-OH	2'-Cl, 4'-OH, 6'-Cl	A6	7,95
	-H	6-OH	4'-OH	B1	8,49
	-Me	6-OH	4'-OH	B2	9,27
	-H	5-Me, 6-OH, 7-Me	4'-OH	B3	8,25
	-Me	5-Me, 6-OH, 7-Me	4'-OH	B4	8,99
	-H	6-OH	4'-OH	C1	9,40
	-Me	6-OH	4'-OH	C2	10,48

Analisando a Tabela 16, observa-se que o modelo 1B9 de QSAR-4D indica o análogo **A2** como o menos potente da nova série e o análogo **C2** como o mais potente. Além do análogo **C2**, destacam-se mais dois com potência elevada: **B2** e **C1**.

Segundo os dados de relação estrutura-atividade (Grese *et al.*, 1997a), como mencionado no item 1.5.4.2, o fato da cadeia lateral do raloxifeno não ser coplanar ao anel benzotiofeno torna esse composto favorável em relação ao seu perfil farmacológico. Contudo, é interessante notar que a molécula mais potente predita pelo modelo é o análogo **C2**, que

apresenta sua cadeia lateral coplanar ao anel benzotiofeno (Figura 43). Entretanto, este análogo possui a orientação da cadeia lateral oposta àquela encontrada para o composto VI, sendo semelhante ao raloxifeno e ao composto VII (Figura 17). Isto pode ser constatado pelo pequeno valor de desvio da raiz dos mínimos quadrados (*Root Mean Square, RMS* = 0,28Å) obtido pela sobreposição do composto C2 (usando os três átomos do alinhamento 1) com a conformação do raloxifeno da estrutura cristalográfica (código do *PDB*: 1ERR), indicando uma grande similaridade entre eles.



**Figura 43.** Representação da conformação inicial do análogo C2 (cinza) sobreposta à estrutura de raios-X do raloxifeno (preto) dentro do sítio ativo do *ER-α* (código *PDB* 1ERR), pelo alinhamento 1. A ocupação dos *GCODs* em rosa contribui para o aumento da potência dos compostos e, em azul, para a diminuição da potência. Os *GCODs* descritos são: (A) (1,13,1)(qa); (B) (2,11,4)(qa); (C) (0,11,5)(qa); (D) (-1,8,2)(qa); (E) (1,6,-2)(ap); (F) (1,2)(qa)(qa); (G) (0,1,-2)(qa); (H) (0,0,-2)(qa); (I) (0,-2,-1)(qa).

A análise das ocupações das células pelos análogos propostos (Anexo VIII) indica que a tentativa de aumentar a rigidez da cadeia lateral do raloxifeno e, conseqüentemente, gerar compostos com orientação mais favorável para realizar interações com os resíduos presentes no sítio de ligação, é bem sucedida. A célula (1,6,-2), cuja ocupação por átomos apolares desfavorece a potência dos compostos, não é ocupada pelos análogos em questão. Outro fato que fortalece esta afirmativa é a distância observada entre os átomos de oxigênio das

hidroxilas fenólicas e destes com o átomo de nitrogênio da cadeia lateral, cujos valores permanecem próximos àqueles observados na estrutura de raios-X do raloxifeno (Tabela 17).

**Tabela 17.** Distância interatômicas entre os heteroátomos das hidroxilas fenólicas e do grupo amino da cadeia lateral da conformação do raloxifeno observada por cristalografia de raios-X (código do *PDB*: 1ERR) e dos compostos das séries A, B e C.

Composto	6-OH····4'-OH (Å)	4'-OH····N (Å)	6-OH····N (Å)
Raloxifeno	11,72	13,37	10,91
A1	11,60	12,94	11,26
A2	11,58	13,20	11,12
A3	11,63	12,13	11,93
A4	11,62	13,24	10,94
A5	11,63	12,85	11,43
A6	11,64	13,36	10,84
B1	12,09	12,06	11,29
B2	11,65	11,76	11,29
B3	11,66	10,71	12,35
B4	11,66	10,76	12,35
C1	11,64	13,81	10,86
C2	11,58	13,81	10,96

O análogo **C2**, que apresentou a melhor potência predita, apresenta pouca ocupação da célula (0,11,5)(qa), que contribui de forma negativa para a potência dos análogos, enquanto que o *GCOD* (1,13,1)(qa), que favorece a potência, é muito ocupado.

Na Tabela 20 observa-se que os valores de  $pIC_{50}$  para a série A são consideravelmente inferiores em relação aos demais análogos. Isto se deve à presença do grupo carbonila que leva a um maior número de orientações da cadeia lateral, orientações essas que nem sempre são favoráveis às potências dos compostos. Dessa forma, a probabilidade de ocupação dos *GCODs* que contribuem negativamente para a ação antagonista no *ER- $\alpha$*  é maior.

## 5 CONCLUSÕES E PERSPECTIVAS

Uma série de 54 análogos do raloxifeno (Grese *et al.*, 1997a) foi selecionada para um estudo de QSAR-4D, sendo 41 compostos utilizados como conjunto de treinamento e 13 como conjunto de teste. Dentre os alinhamentos testados, os alinhamentos 1 e 2 apresentaram melhores resultados. Os modelos obtidos foram avaliados em termos de qualidade estatística e preditividade, além de validações baseadas no sistema biológico em questão.

Foi selecionado um modelo representativo para cada um dos alinhamentos 1 e 2 para análise. Ambos os modelos selecionados (modelos 1B9 e 2B9) revelam como a flexibilidade dos análogos do raloxifeno (Grese *et al.*, 1997a) influencia na atividade biológica das moléculas. O modelo 1B9, obtido pelo alinhamento 1, apresentou-se mais consistente com os dados de cristalografia de raios-X (Brzozowski *et al.*, 1997), principalmente em relação aos resíduos que interagem com os ligantes no sítio ativo.

Na proposta de novos compostos, foi utilizada uma estratégia de restrição conformacional da cadeia lateral dos compostos em estudo. O modelo 1B9 foi usado para predição da potência dos análogos propostos, sendo obtidos resultados de afinidade subnanomolar. É importante destacar que o composto predito como o mais potente pelo modelo foi aquele que apresenta a cadeia lateral no mesmo plano do anel benzotiofeno (análogo C2). Este resultado reflete que a rigidez imposta à cadeia lateral da molécula induz a uma conformação propícia para a formação de interações eletrostáticas e por ligação hidrogênio com o resíduo Asp351 presente no sítio de ligação do ligante (*LBD*) do *ER- $\alpha$* .

Espera-se que os análogos propostos como potenciais antagonistas do *ER- $\alpha$*  sejam sintetizados e avaliados farmacologicamente, para posterior continuação dos estudos sobre os *SERMs*.



## 6 REFERÊNCIAS BIBLIOGRÁFICAS

Albuquerque, M.G., Brito, M.A., Cunha, E.F.F., Alencastro, R.B., Antunes, O.A.C., Castro, H.C. & Rodrigues, C.R. Multidimensional-QSAR: Beyond the third-dimension in drug design. In: Taft, C.A. e Silva, C.H.T.P. (Ed.). **Current Methods in Medicinal Chemistry and Biological Physics**, 2007. *In Press*.

Albuquerque, M.G., Hopfinger, A.J., Barreiro, E.J. & Alencastro, R.B. (1998) Four-dimensional quantitative structure-activity relationship analysis of a series of interphenylene 7-oxabicycloheptane oxazole thromboxane A(2) receptor antagonists. **Journal of Chemical Information and Computer Sciences**, 38 (5): 925-938.

Barreiro, E.J. & Fraga, C.A.M. (2001) **Química Medicinal: As Bases Moleculares da Ação dos Fármacos**. Porto Alegre: Artmed Ed. 243 p.

Becker, O.M., Levy, Y. & Ravitz, O. (2000) Flexibility, conformation spaces, and bioactivity. **Journal of Physical Chemistry B**, 104 (9): 2123-2135.

Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F., Dinola, A. & Haak, J.R. (1984) Molecular-Dynamics with Coupling to an External Bath. **Journal of Chemical Physics**, 81 (8): 3684-3690.

Berman, H.M., Westbrook, J.F., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000) The Protein Data Bank. **Nucleic Acids Research**, 28 (1): 235-242.

Brueggemeier, R.W., Miller, D.D. & Witiak, D.T. Cholesterol, Adrenocorticoids, and Sex hormones. In: Foye, W.O., Lemke, D.A.W. e Willians, D.A. (Ed.). **Principles of Medicinal Chemistry**. Meier: Williams & Wilkins, 1995. p.444-498

Bryant, H.U. (2001) Mechanism of Action and Preclinical Profile of Raloxifene, a Selective Estrogen Receptor Modulator. **Reviews in Endocrine & Metabolic Disorders**, 2: 129-138.

Brzozowski, A.M., Pike, A.C.W., Dauter, Z., Hubbard, R.E., Bonn, T., Engstrom, O., Ohman, L., Greene, G.L., Gustafsson, J.A. & Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. **Nature**, 389 (6652): 753-758.

Cramer, R.D., Patterson, D.E. & Bunce, J.D. (1988) Comparative Molecular-Field Analysis (Comfa) .1. Effect of Shape on Binding of Steroids to Carrier Proteins. **Journal of the American Chemical Society**, 110 (18): 5959-5967.

Cunha, E., Martins, R., Albuquerque, M. & Alencastro, R. (2004) LIV-3D-QSAR model for estrogen receptor ligands **Journal of Molecular Modeling**, 10 (4): 297-304.

Cunha, E.F.F., Albuquerque, M.G., Antunes, O.A.C. & Alencastro, R.B. (2005) 4D-QSAR models of HOE/BAY-793 analogues as HIV-1 protease inhibitors. **QSAR & Combinatorial Science**, 24 (2): 240-253.

*4D-QSAR*. The Chem21 Group, Inc.; Lake Forest, IL 2000.

- Debnath, A.K. (2001) Quantitative Structure-Activity Relationship (QSAR) Paradigm – Hansch Era to New Millennium. **Mini Reviews in Medicinal Chemistry**, 1 (2): 187-195.
- Gao, H., Katzenellenbogen, J.A., Garg, R. & Hansch, C. (1999) Comparative QSAR analysis of estrogen receptor ligands. **Chemical Reviews**, 99 (3): 723-744.
- Gaudio, A.C. & Zandonade, E. (2001) Proposition, validation and analysis of QSAR models. **Quimica Nova**, 24 (5): 658-671.
- Glenn, W.G., Dunn, I.W.J. & Scott, D.R. (1989) Principal components analysis and partial least square. **Tetrahedron Computer Methodology**, 2 (6): 349-376.
- Grese, T.A., Cho, S., Finley, D.R., Godfrey, A.G., Jones, C.D., Lugar, C.W., Martin, M.J., Matsumoto, K., Pennington, L.D., Winter, M.A., Adrian, M.D., Cole, H.W., Magee, D.E., Phillips, D.L., Rowley, E.R., Short, L.L., Glasebrook, A.L. & Bryant, H.U. (1997a) Structure-activity relationships of selective estrogen receptor modulators: Modifications to the 2-arylbenzothiophene core of raloxifene. **Journal of Medicinal Chemistry**, 40 (2): 146-167.
- Grese, T.A., Sluka, J.P., Bryant, H.U., Cullinan, G.J., Glasebrook, A.L., Jones, C.D., Matsumoto, K., Palkowitz, A.D., Sato, M., Termine, J.D., Winter, M.A., Yang, N.N. & Dodge, J.A. (1997b) Molecular determinants of tissue selectivity in estrogen receptor modulators. **Proceedings of the National Academy of Sciences of the United States of America**, 94 (25): 14105-14110.

Hansch, C. (1969) A Quantitative Approach to Biochemical Structure-Activity Relationships. **Accounts of Chemical Research**, 2 (8): 232-239.

Hansch, C. & Fujita, T. (1964) Rho-Sigma-Pi Analysis. Method for Correlation of Biological Activity + Chemical Structure. **Journal of the American Chemical Society**, 86 (8): 1616-1626.

Hong, X. & Hopfinger, A. J. (2003) 3D-Pharmacophores of Flavonoid Binding at the Benzodiazepine GABA<sub>A</sub> Receptor Site using 4D-QSAR Analysis. **Journal of Chemical Information and Modeling**, 43 (1): 324 -336.

Hopfinger, A.J. (1980) A QSAR Investigation of Dihydrofolate-Reductase Inhibition by Baker Triazines Based Upon Molecular Shape-Analysis. **Journal of the American Chemical Society**, 102 (24): 7196-7206.

Hopfinger, A.J., Wang, S., Tokarski, J.S., Jin, B.Q., Albuquerque, M., Madhav, P.J. & Duraiswami, C. (1997) Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. **Journal of the American Chemical Society**, 119 (43): 10509-10524.

Hyperchem release 5 for Windows. Hypercube, Inc., *1115 NW 4th Street, Gainesville, FL 32601 USA.*

Hyperchem release 7 for Windows. Hypercube, Inc., *1115 NW 4th Street, Gainesville, FL 32601 USA.*

- Johnston, S.R.D. (2005) Endocrinology and hormone therapy in breast cancer - Selective Oestrogen Receptor Modulators and Downregulators for breast cancer - have they lost their way? **Breast Cancer Research**, 7 (3): 119-130.
- Jordan, V.C. (1999) Targeted antiestrogens to prevent breast cancer. **Trends in Endocrinology and Metabolism**, 10 (8): 312-317.
- Jordan, V. C. (2001) The past, present, and future of selective estrogen receptor modulation. **Annals of the New York Academy of Sciences**, 949: 72-79.
- Klinge, C. (2000) Estrogen receptor interaction with co-activators and co-repressors. **Steroids**, 65 227-251.
- Kraft, K.S., Ruenitz, P.C. & Bartlett, M.G. (1999) Carboxylic acid analogues of tamoxifen: (Z)-2-[p-(1,2-diphenyl-1-butenyl)phenoxy]-N,N-dimethylethylamine. Estrogen receptor affinity and estrogen antagonist effects in MCF-7 cells. **Journal of Medicinal Chemistry**, 42 (16): 3126-3133.
- Kubinyi, H. (1977) Quantitative Structure-Activity-Relationships .7. Bilinear Model, A New Model For Nonlinear Dependence Of Biological-Activity on Hydrophobic Character. **Journal of Medicinal Chemistry**, 20 (5): 625-629.
- Kubinyi, H. (1993) **Hansch Analysis and Related Approaches**. New York: VCH. 240 p.
- Kubinyi, H. (1997) QSAR and 3D QSAR in drug design .1. Methodology. **Drug Discovery Today**, 2 (11): 457-467.

- Kuiper, G., Carlsson, B., Grandien, K., Enmark, E., Haggblad, J., Nilsson, S. & Gustafsson, J.A. (1997) Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. **Endocrinology**, 138 (3): 863-870.
- Leach, A.R. (2001) **Molecular modelling: principles and applications**. Harlow: Pearson Education Limited. 744 p.
- Leo, A.J. (1993) Calculating Log P(Oct) from Structures. **Chemical Reviews**, 93 (4): 1281-1306.
- Lerner, L.J., Holthaus, F.J. & Thompson, C.R. (1958) Non-Steroidal Estrogen Antagonist 1-(P-2-Diethylaminoethoxyphenyl)-1-Phenyl-2-P-Methoxyphenyl Ethanol. **Endocrinology**, 63 (3): 295-318.
- Lewis, J.S. & Jordan, V.C. (2005) Selective estrogen receptor modulators (SERMs): Mechanisms of anticarcinogenesis and drug resistance. **Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis**, 591 (1-2): 247-263.
- Livingstone, D. (1995) **Analysis for Chemists: Applications to QSAR and Chemical Product Design**. New York: Oxford University Press. 256 p.
- Loose-Mitchell, D.S. & Stancel, G.M. Estrogênios e Progestênios. In: Hardman, J.G. e Limbird, L.E. (Ed.). **As Bases Farmacológicas da Terapêutica**. Rio de Janeiro: Mc Graw-Hill, 2003. p.1201-1229

McDonnell, D.P. (2000) Selective estrogen receptor modulators (SERMs): A first step in the development of perfect hormone replacement therapy regimen. **Journal of the Society for Gynecologic Investigation**, 7 (1): S10-S15.

McDonnell, D. P. (2004) The molecular determinants of estrogen receptor pharmacology. **Maturitas**, 48: S7-S12.

Migliavacca, E. (2003) Applied Introduction to Multivariate Methods Used in Drug Discovery. **Mini Reviews in Medicinal Chemistry**, 3 (8): 831-843.

Molsim 3.0. The Chem21 Group, Inc.; Lake Forest, IL 1994.

Nilsson, S. & Koehler, K.F. (2005) Oestrogen receptors and selective oestrogen receptor modulators: Molecular and cellular pharmacology. **Basic & Clinical Pharmacology & Toxicology**, 96 (1): 15-25.

Pappas, S.G. & Jordan, V.C. (2002) Chemoprevention of breast cancer: Current and future prospects. **Cancer and Metastasis Reviews**, 21 (3-4): 311-321.

Park, W.C. & Jordan, V.C. (2002) Selective estrogen receptor modulators (SERMS) and their roles in breast cancer prevention. **Trends in Molecular Medicine**, 8 (2): 82-88.

- Pita, S.S.R. (2006) **Modelos de QSAR-4D Dependente do Receptor de inibidores peptídicos da Tripanotona Redutase**. Instituto de Química / Departamento de Química Orgânica, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro. 94 p.
- Robertson, D.W., Katzenellenbogen, J.A., Hayes, J.R. & Katzenellenbogen, B.S. (1982) Anti-Estrogen Basicity Activity Relationships - A Comparison of the Estrogen-Receptor Binding and Anti-Uterotrophic Potencies of several analogs Of (Z)-1,2-Diphenyl-1-[4-[2-(Dimethylamino)Ethoxy]Phenyl]-1-Butene (Tamoxifen, Nolvadex®) having altered basicity. **Journal of Medicinal Chemistry**, 25 (2): 167-171.
- Rogers, D. & Hopfinger, A.J. (1994) Application of Genetic Function Approximation to Quantitative Structure-Activity-Relationships and Quantitative Structure-Property Relationships. **Journal of Chemical Information and Computer Sciences**, 34 (4): 854-866.
- Romeiro, N.C. (2002) **Estudos de QSAR Independentes e Dependente do Receptor aplicados a inibidores da proteína quinase p38**. Instituto de Química / Departamento de Química Orgânica, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro. 153 p.
- Romeiro, N.C., Albuquerque, M.G., Alencastro, R.B., Ravi, M. & Hopfinger, A.J. (2005) Construction of 4D-QSAR models for use in the design of novel p38-MAPK inhibitors. **Journal of Computer-Aided Molecular Design**, 19 (6): 385-400.
- Senese, C.L., Hopfinger, A.J. (2003) Receptor-Independent 4D-QSAR Analysis of a Set of Norstatine Derived Inhibitors of HIV-1 Protease. **Journal of Chemical Information and Computer Sciences**, 43 (4): 1297-1307.



Shiau, A.K., Barstad, D., Loria, P.M., Cheng, L., Kushner, P.J., Agard, D.A. & Greene, G.L. (1998) The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. **Cell**, 95 (7): 927-937.

Tavares, L.C. (2004) QSAR: A Abordagem de Hansch. **Química Nova**, 27 (4): 631-639.

Thompson, E.W., Reich, R., Shima, T.B., Albin, A., Graf, J., Martin, G.R., Dickson, R.B. & Lippman, M.E. (1988) Differential Regulation Of Growth And Invasiveness Of MCF-7 Breast-Cancer Cells By Antiestrogens. **Cancer Research**, 48 (23): 6764-6768.

van Gunsteren, W.F. & Berendsen, H.J.C. (1990) Computer-Simulation of Molecular-Dynamics - Methodology, Applications, and Perspectives in Chemistry. **Angewandte Chemie-International Edition in English**, 29 (9): 992-1023.

van Gunsteren, W.F. & Mark, A.E. (1992) On the Interpretation of Biochemical Data by Molecular-Dynamics Computer-Simulation. **European Journal of Biochemistry**, 204 (3): 947-961.

Vedani, A. & Dobler, M. (2002) 5D-QSAR: The key for simulating induced fit? **Journal of Medicinal Chemistry**, 45 (11): 2139-2149.

Vedani, A., Dobler, M. & Lill, M.A. (2005) Combining protein modeling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor. **Journal of Medicinal Chemistry**, 48 (11): 3700-3703.

Vedani, A., Dobler, M. & Lill, M.A. (2006) The Challenge of Predicting Drug Toxicity *in silico*. **Basic & Clinical Pharmacology & Toxicology**, 99 (3): 195-208.

Vogel, V.G., Costantino, J.P., Wickerham, D.L., Cronin, W.M., Cecchini, R.S., Atkins, J.N., Bevers, T.B., Fehrenbacher, L., Pajon, E.R., Wade, J.L., Robidoux, A., Margolese, R.G., James, J., Lippman, S.M., Runowicz, C.D., Ganz, P.A., Reis, S.E., Mccaskill-Stevens, W., Ford, L.G., Jordan, V.C. & Wolmark, N. (2006) Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes - The NSABP study of tamoxifen and raloxifene (STAR) P-2 trial. **JAMA-Journal of the American Medical Association**, 295 (23): 2727-2741.

#### **Sites consultados**

WHO, World Health Organization: [www.who.int/cancer/en/](http://www.who.int/cancer/en/) (Acesso em 22/11/2006)

INCA, Instituto do Câncer: [www.inca.gov.br/estimativa/2006/](http://www.inca.gov.br/estimativa/2006/) (Acesso em 22/11/2006)

## ANEXO I

Matriz de correlação cruzada entre os valores de ocupação das células presentes nos modelos 1B7 e 1B9 e entre os valores das potências dos compostos.

	Potência	(2,11,4) (qa)	(0,-2,-1) (qa)	(0,1,-2) (qa)	(0,11,5) (qa)	(-1,8,2) (qa)	(1,6,-2) (ap)	(1,13,1) (qa)	(1,2,0) (qa)	(0,0,-2) (qa)	(0,-2,9) (qa)	(1,-1,9) (qa)	(1,2,2) (ar)	(1,11,-2) (qa)
Potência	1,00													
(2,11,4) (qa)	-0,46	1,00												
(0,-2,-1) (qa)	-0,31	0,13	1,00											
(0,1,-2) (qa)	-0,26	0,36	0,47	1,00										
(0,11,5) (qa)	-0,35	0,30	0,01	0,13	1,00									
(-1,8,2) (qa)	-0,27	0,20	0,10	0,12	-0,17	1,00								
(1,6,-2) (ap)	-0,21	-0,21	-0,07	-0,13	-0,21	0,06	1,00							
(1,13,1) (qa)	0,12	0,13	0,00	-0,07	-0,37	0,52	0,01	1,00						
(1,2,0) (qa)	0,36	-0,01	0,18	0,18	-0,16	0,22	0,11	0,10	1,00					
(0,0,-2) (qa)	0,43	-0,31	0,26	-0,19	0,01	-0,11	0,07	-0,13	0,28	1,00				
(0,-2,9) (qa)	-0,30	0,05	0,20	0,16	-0,29	0,18	-0,05	0,12	-0,11	-0,35	1,00			
(1,-1,9) (qa)	0,36	-0,26	-0,11	-0,08	-0,11	-0,09	0,06	-0,27	0,41	0,35	-0,23	1,00		
(1,2,2) (ar)	0,33	-0,18	-0,27	0,21	0,14	-0,16	-0,09	-0,17	0,34	0,06	-0,18	0,35	1,00	
(1,11,-2) (qa)	0,60	-0,24	-0,10	-0,17	-0,41	-0,06	-0,27	0,06	0,08	0,31	0,07	0,02	-0,08	1,00

## ANEXO II

Valores de  $pIC_{50exp}$  e  $pIC_{50calc}$  e valores residuais, do modelo 1B7 de QSAR-4D para o conjunto de treinamento.

#	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>
1	9,70	9,27	-0,43
2	9,52	8,21	<b>-1,31</b>
3	9,15	9,03	-0,12
4	9,10	8,93	-0,17
5	9,00	9,23	0,23
6	9,00	7,92	<b>-1,08</b>
7	8,70	8,82	0,12
9	8,64	8,32	-0,32
10	8,64	8,40	-0,24
11	8,60	8,44	-0,16
12	8,60	8,87	0,27
13	8,52	7,44	<b>-1,08</b>
15	8,15	8,09	-0,06
16	8,00	8,33	0,33
19	7,70	6,70	-1,00
21	7,52	7,55	0,03
22	7,49	8,00	0,51
23	7,46	6,83	-0,63
24	7,40	7,27	-0,13
25	7,30	7,65	0,35
27	7,30	7,31	0,01
28	7,22	6,76	-0,46
29	7,00	7,15	0,15
30	7,00	6,25	-0,75
32	7,00	7,48	0,48
34	7,00	6,81	-0,19
35	6,72	6,55	-0,17
36	6,70	6,38	-0,32
37	6,60	6,79	0,19
38	6,52	6,14	-0,38
39	6,52	6,00	-0,52
42	6,49	6,95	0,46
43	6,46	6,70	0,24
44	6,46	6,09	-0,37
45	6,40	6,90	0,50
47	6,30	6,62	0,32
49	6,30	7,26	0,96
50	6,22	7,21	0,99
52	6,00	6,44	0,44
53	6,00	6,71	0,71
54	6,00	5,95	-0,05

<sup>a</sup> Os valores residuais em negrito correspondem aos compostos *outliers* que são os compostos com valor residual maior do que o dobro do  $DP_{res}$  ( $DP_{res} = 0,53$ )

### ANEXO III

Valores de ocupação das células do modelo 1B9 dos 54 derivados do raloxifeno (Grese *et al.*, 1997a).

#	(2,11,4)(qa)	(0,-2,-1)(qa)	(0,1,-2)(qa)	(0,11,5)(qa)	(-1,8,2)(qa)	(1,6,-2)(ap)	(1,13,1)(qa)	(1,2)(qa)(qa)	(0,0,-2)(qa)
1	0,028	0,0105	0,0005	0,0205	0,145	0	0,0865	0,3365	0,496
2	0,035	0,0035	0,0015	0,067	0,1935	0,001	0,0675	0,3225	0,7965
3	0,029	0	0,0065	0,069	0,184	0	0,0415	0,408	0,819
4	0,0275	0,0035	0,001	0,0225	0,178	0,008	0,033	0,2725	0,753
5	0,03	0,011	0,0005	0,0515	0,191	0	0,1045	0,285	0,4315
6	0,034	0,0125	0,001	0,043	0,1105	0,02	0,0385	0,2605	0,6145
7	0,0455	0	0,0065	0,0675	0,158	0,008	0,023	0,4165	0,6985
8	0,026	0,0095	0,0015	0,046	0,1435	0,009	0,0245	0,3315	0,3465
9	0,069	0,001	0,0025	0,1255	0,133	0	0,0335	0,2865	0,8395
10	0,066	0,017	0	0,0375	0,1315	0	0,0765	0,2845	0,3605
11	0,037	0,01	0,0005	0,0825	0,2095	0,0085	0,0725	0,31	0,42
12	0,027	0,0215	0,0015	0,0305	0,24	0,013	0,0745	0,264	0,3955
13	0,037	0	0,002	0,061	0,1535	0	0,026	0,051	0,883
14	0,038	0,0045	0,0005	0,0415	0,193	0	0,0555	0,287	0,451
15	0,0345	0,017	0,0005	0,0365	0,2895	0,001	0,095	0,269	0,3725
16	0,029	0,0075	0,0025	0,0605	0,141	0,016	0,034	0,352	0,478
17	0,0425	0,0165	0,0015	0,0595	0,093	0,01	0,0445	0,276	0,36
18	0,0235	0,0145	0,001	0,0755	0,1255	0,019	0,0335	0,272	0,359
19	0,05	0	0	0,0255	0,184	0,0025	0,052	0,2375	0,036
20	0,0665	0,0075	0,001	0,056	0,196	0,0095	0,073	0,2755	0,5135
21	0,0455	0,253	0,002	0,049	0,175	0,0185	0,084	0,28	0,9375
22	0,0615	0,0155	0,0005	0,0625	0,1775	0	0,0645	0,2735	0,349
23	0,041	0	0	0,05	0,1935	0,0005	0,042	0,232	0
24	0,0535	0,0115	0,0015	0,0605	0,1385	0,0025	0,0415	0,2615	0,1635
25	0,0495	0,0085	0	0,0555	0,18	0,0105	0,0455	0,2955	0,531
26	0,045	0,0015	0,0015	0,121	0,1755	0,0055	0,0355	0,2695	0,741

### ANEXO III (CONT.)

Valores de ocupação das células do modelo 1B9 dos 54 derivados do raloxifeno (Grese *et al.*, 1997a).

#	(2,11,4)(qa)	(0,-2,-1)(qa)	(0,1,-2)(qa)	(0,11,5)(qa)	(-1,8,2)(qa)	(1,6,-2)(ap)	(1,13,1)(qa)	(1,2)(qa)(qa)	(0,0,-2)(qa)
27	0,056	0,0155	0,0005	0,0305	0,143	0,036	0,0375	0,2785	0,177
28	0,049	0,311	0,006	0,081	0,1735	0,0105	0,0445	0,343	0,9455
29	0,054	0,0005	0,008	0,162	0,1265	0,005	0,048	0,1495	0,5905
30	0,074	0	0	0,082	0,145	0,0025	0,0515	0,246	0
31	0,0215	0	0	0,0195	0,1175	0	0,0345	0,008	0
32	0,031	0,0145	0,0015	0,0945	0,1045	0,006	0,0275	0,2355	0,2965
33	0,045	0,008	0,001	0,0875	0,173	0,0055	0,0485	0,2435	0,533
34	0,0465	0,017	0,0005	0,123	0,2425	0,0075	0,0405	0,2985	0,244
35	0,0995	0	0	0,064	0,2675	0,001	0,1025	0,1845	0
36	0,0445	0,0205	0	0,1005	0,2385	0	0,0405	0,3185	0,2385
37	0,0655	0,179	0,0935	0,0795	0,238	0	0,0495	0,3205	0,1965
38	0,064	0	0	0,048	0,325	0	0,0805	0,2515	0
39	0,0685	0,2165	0,068	0,082	0,197	0,004	0,035	0,3395	0,162
40	0,04	0,2475	0,015	0,0865	0,1775	0	0,03	0,327	0,565
41	0,061	0	0	0,071	0,192	0	0,066	0,2395	0
42	0,0335	0,0045	0,0005	0,0365	0,2075	0,0805	0,041	0,285	0,6415
43	0,0385	0,113	0,007	0,057	0,1445	0	0,048	0,017	0,007
44	0,1005	0,0045	0	0,0775	0,232	0	0,0635	0,205	0,553
45	0,033	0,014	0,0035	0,055	0,2295	0	0,059	0,0215	0,4185
46	0,008	0,0165	0,0005	0,011	0,17	0,114	0,0145	0,2805	0,4
47	0,045	0,0115	0,001	0,1175	0,045	0,004	0,0175	0	0,0415
48	0,098	0,14	0,0755	0,061	0,2055	0	0,068	0,3065	0,193
49	0,0475	0,0055	0,001	0,093	0,1735	0,007	0,0435	0,0155	0,3125
50	0,093	0,159	0,079	0,073	0,205	0,001	0,0655	0,3035	0,1605
51	0,061	0,0185	0	0,0475	0,1825	0	0,0345	0,2555	0,2635
52	0,0445	0,3295	0,003	0,0575	0,2435	0,003	0,05	0,3075	0,924
53	0,0395	0	0	0,0535	0,246	0,084	0,0865	0,2685	0,227
54	0,0705	0,01	0,0005	0,118	0,246	0,0095	0,0415	0,307	0,497

## ANEXO IV

Matriz de correlação cruzada entre os valores de ocupação das células presentes nos modelos 2B7, 2B8 e 2B9 e entre os valores das potências dos compostos.

Potência	(-1,12,6) (qa)	(1,3,6) (alh)	(0,11,3) (p-)	(0,12,-1) (qa)	(-1,9,4) (qa)	(0,10,-2) (qa)	(-3,9,4) (qa)	(0,10,4) (qa)	(0,7,-2) (qa)	(0,2,-1) (qa)	(0,0,-2) (qa)	(-1,-2,11) (qa)	(0,1,-2) (qa)	(-3,8,5) (qa)	(0,5,-3) (qa)	(0,2,2) (ar)	(2,5,0) (qa)	(0,-1,9) (ar)	
Potência (-1,12,6) (qa)	1,00																		
(1,3,6) (alh)	-0,30	1,00																	
(0,11,3) (p-)	0,40	0,03	1,00																
(0,12,-1) (qa)	0,37	0,13	-0,19	1,00															
(-1,9,4) (qa)	0,33	-0,11	-0,03	0,42	1,00														
(0,10,-2) (qa)	-0,15	0,03	-0,15	-0,13	-0,26	1,00													
(-3,9,4) (qa)	0,49	-0,06	0,28	0,29	0,73	-0,13	1,00												
(0,10,4) (qa)	-0,21	-0,27	-0,33	-0,28	-0,33	0,59	-0,27	1,00											
(0,7,-2) (qa)	-0,50	0,71	-0,29	-0,19	-0,28	0,31	-0,30	0,21	1,00										
(0,2,-1) (qa)	0,22	-0,14	0,26	0,14	0,22	-0,10	0,47	-0,24	-0,34	1,00									
(0,0,-2) (qa)	-0,16	0,12	0,07	-0,11	0,01	0,00	0,07	-0,09	0,08	-0,03	1,00								
(-1,-2,11) (qa)	0,41	-0,07	0,24	0,16	0,13	-0,24	0,23	-0,22	-0,38	0,29	-0,04	1,00							
(0,1,-2) (qa)	-0,34	0,13	0,02	-0,21	-0,16	-0,08	-0,14	-0,04	0,16	-0,07	-0,08	-0,30	1,00						
(-3,8,5) (qa)	-0,28	0,46	-0,03	-0,05	0,06	-0,19	-0,02	-0,14	0,24	-0,14	-0,09	-0,16	0,22	1,00					
(0,5,-3) (qa)	-0,21	-0,29	-0,33	-0,26	-0,29	0,69	-0,15	0,92	0,16	-0,15	-0,05	-0,26	-0,01	-0,15	1,00				
(0,2,2) (ar)	-0,22	-0,12	0,06	0,01	-0,25	-0,05	-0,18	-0,11	-0,21	0,59	-0,05	0,16	0,00	-0,12	-0,06	1,00			
(2,5,0) (qa)	0,45	0,10	0,44	0,06	0,03	-0,16	0,06	-0,16	-0,04	0,09	0,04	0,12	0,30	0,11	-0,20	-0,01	1,00		
(0,-1,9) (ar)	-0,30	-0,10	-0,01	-0,05	-0,30	0,06	-0,26	0,01	-0,09	0,47	-0,01	0,08	0,00	-0,14	0,07	0,94	-0,07	1,00	
	0,15	-0,09	0,02	-0,02	-0,25	-0,18	-0,24	-0,05	-0,01	0,08	-0,01	0,08	0,20	-0,11	-0,05	0,22	0,44	0,22	1,00

## ANEXO V

Valores de  $pIC_{50exp}$  e  $pIC_{50calc}$  e valores residuais, do modelo 2B7 de QSAR-4D para o conjunto de treinamento.

#	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>
1	9,70	9,04	-0,66
2	9,52	8,07	<b>-1,45</b>
3	9,15	8,97	-0,18
4	9,10	8,99	-0,11
5	9,00	8,17	-0,83
6	9,00	9,05	0,05
7	8,70	8,77	0,07
9	8,64	8,57	-0,07
10	8,64	8,87	0,23
11	8,60	7,79	-0,81
12	8,60	8,83	0,23
13	8,52	7,83	-0,69
15	8,15	7,89	-0,26
16	8,00	8,34	0,34
19	7,70	6,95	-0,75
21	7,52	6,94	-0,58
22	7,49	6,96	-0,53
23	7,46	7,21	-0,25
24	7,40	7,60	0,20
25	7,30	8,04	0,74
27	7,30	7,60	0,30
28	7,22	6,83	-0,39
29	7,00	6,93	-0,07
30	7,00	6,19	-0,81
32	7,00	7,18	0,18
34	7,00	6,99	-0,01
35	6,72	6,37	-0,35
36	6,70	6,50	-0,20
37	6,60	7,29	0,69
38	6,52	6,94	0,42
39	6,52	6,13	-0,39
42	6,49	6,34	-0,15
43	6,46	6,47	0,01
44	6,46	6,54	0,08
45	6,40	6,59	0,19
47	6,30	6,59	0,29
49	6,30	7,65	<b>1,35</b>
50	6,22	7,26	1,04
52	6,00	6,84	0,84
53	6,00	5,69	-0,31
54	6,00	6,01	0,01

<sup>a</sup> Os valores residuais em negrito correspondem aos compostos *outliers* que são os compostos com valor residual maior do que o dobro do  $DP_{res}$  ( $DP_{res} = 0,55$ ).



## ANEXO VI

Valores de  $pIC_{50exp}$  e  $pIC_{50calc}$  e valores residuais, do modelo 2B8 de QSAR-4D para o conjunto de treinamento.

#	$pIC_{50exp}$	$pIC_{50calc}$	Resíduo <sup>a</sup>
1	9,70	8,99	-0,71
2	9,52	8,15	<b>-1,37</b>
3	9,15	9,42	0,27
4	9,10	8,55	-0,55
5	9,00	8,15	-0,85
6	9,00	8,88	-0,12
7	8,70	8,66	-0,04
9	8,64	8,91	0,27
10	8,64	9,03	0,39
11	8,60	8,02	-0,58
12	8,60	8,95	0,35
13	8,52	7,46	<b>-1,06</b>
15	8,15	8,05	-0,10
16	8,00	8,42	0,42
19	7,70	7,21	-0,49
21	7,52	7,12	-0,40
22	7,49	7,15	-0,34
23	7,46	7,25	-0,21
24	7,40	7,45	0,05
25	7,30	7,92	0,62
27	7,30	7,56	0,26
28	7,22	6,71	-0,51
29	7,00	7,05	0,05
30	7,00	6,40	-0,60
32	7,00	6,86	-0,14
34	7,00	7,20	0,20
35	6,72	6,76	0,04
36	6,70	7,01	0,31
37	6,60	6,75	0,15
38	6,52	7,34	0,82
39	6,52	6,04	-0,48
42	6,49	6,51	0,02
43	6,46	6,35	-0,11
44	6,46	6,60	0,14
45	6,40	6,51	0,11
47	6,30	6,10	-0,20
49	6,30	6,34	0,04
50	6,22	6,85	0,63
52	6,00	6,97	0,97
53	6,00	5,82	-0,18
54	6,00	6,38	0,38

<sup>a</sup> Os valores residuais em negrito correspondem aos compostos *outliers* que são os compostos com valor residual maior do que o dobro do  $DP_{res}$  ( $DP_{res} = 0,49$ ).

## ANEXO VII

Valores de ocupâncias das células presentes no modelo 2B9 obtido pelo alinhamento 2 para todos os compostos da série (Grese *et al.*, 1997a).

#	(-1,12,6)(qa)	(1,3,6)(alh)	(0,11,3)(p-)	(0,12,-1)(qa)	(0,10,-2)(qa)	(-1,-2,11)(qa)	(0,1,-2)(qa)	(0,2,2)(ar)	(2,5)(qa)(qa)
1	0,0135	0,041	0,077	0,1145	0,1175	0,006	0,494	0,0605	0
2	0,027	0,049	0,0085	0,0215	0,035	0,021	0,7855	0,059	0,0005
3	0,0065	0,0645	0,0315	0,034	0,051	0	0,8015	0,062	0
4	0,014	0,0505	0,0395	0,048	0,0785	0	0,7605	0,045	0,0015
5	0,034	0,0405	0,1005	0,0245	0,044	0,002	0,425	0,0405	0
6	0,0065	0,074	0,007	0,032	0,038	0	0,6135	0,0715	0,011
7	0,031	0,0885	0,0155	0,017	0,0455	0,0015	0,69	0,0665	0,0015
8	0,0215	0,073	0,0055	0,0205	0,0395	0,004	0,34	0,0755	0,0145
9	0,018	0,0465	0,052	0,043	0,0425	0,0025	0,836	0,051	0
10	0,0165	0,0425	0,0655	0,0725	0,0745	0	0,3525	0,0605	0
11	0,034	0,041	0,057	0,082	0,0985	0	0,409	0,052	0,0015
12	0,0105	0,057	0,0095	0,045	0,0655	0	0,3985	0,0545	0
13	0,0155	0,0585	0,0275	0,0825	0,1435	0,024	0,884	0	0
14	0,002	0,053	0,004	0,022	0,0335	0,179	0,446	0,0525	0
15	0,003	0,04	0,024	0,0465	0,0495	0,1125	0,3715	0,045	0,0005
16	0,022	0,0555	0,0065	0,037	0,0335	0	0,4735	0,062	0,006
17	0,009	0,0485	0,017	0,0595	0,061	0,203	0,3485	0,0565	0,005
18	0,0125	0,0515	0,027	0,0235	0,0255	0,1705	0,35	0,0535	0,001
19	0,0035	0,04	0,0015	0,055	0,056	0,0105	0,0715	0	0,002
20	0,0185	0,043	0,0195	0,0245	0,036	0,167	0,5115	0,0525	0,001
21	0,018	0,042	0,023	0,036	0,0505	0,0135	0,9475	0	0,0075
22	0,0115	0,0445	0,037	0,053	0,0545	0,4245	0,3425	0,0695	0
23	0,02	0,0465	0,0135	0,0475	0,0575	0,0025	0	0	0
24	0,006	0,046	0,007	0,022	0,0385	0,107	0,1615	0,0515	0,002
25	0,0125	0,0325	0,016	0,068	0,055	0,0995	0,526	0,0485	0,006
26	0,0465	0,0355	0,0285	0,035	0,052	0,153	0,735	0,0545	0

## ANEXO VII (CONT.)

Valores de ocupâncias das células presentes no modelo 2B9 obtido pelo alinhamento 2 para todos os compostos da série (Grese *et al.*, 1997a).

#	(-1,12,6)(qa)	(1,3,6)(alh)	(0,11,3)(p-)	(0,12,-1)(qa)	(0,10,-2)(qa)	(-1,-2,11)(qa)	(0,1,-2)(qa)	(0,2,2)(ar)	(2,5)(qa)(qa)
27	0,005	0,0405	0,0085	0,0455	0,042	0,1185	0,1755	0,0475	0,0115
28	0,0175	0,0385	0,02	0,0445	0,0435	0,0205	0,9315	0	0,01
29	0,0355	0,001	0,0365	0,035	0,0355	0,007	0,5795	0	0,001
30	0,03	0,0515	0,0175	0,028	0,072	0,141	0	0	0,0015
31	0,0055	0,033	0,0255	0,15	0,176	0,008	0	0	0
32	0,0195	0,045	0,01	0,0315	0,063	0,1775	0,2885	0,053	0,0035
33	0,026	0,0485	0,0295	0,0175	0,037	0,13	0,53	0,0565	0,014
34	0,0365	0,044	0,0215	0,025	0,026	0,005	0,2375	0,0425	0,016
35	0,0265	0,011	0,041	0,03	0,0155	0,042	0	0,002	0,001
36	0,048	0,05	0,0115	0,037	0,036	0,2595	0,2375	0,074	0
37	0,0495	0,0395	0,0285	0,0665	0,0715	0,026	0,181	0,052	0
38	0,006	0,0415	0,014	0,0415	0,0265	0	0	0	0
39	0,046	0,0465	0,0085	0,0095	0,038	0,2545	0,1555	0,0565	0
40	0,041	0,034	0,01	0,0495	0,0605	0,0095	0,53	0	0
41	0,041	0,0405	0,0295	0,0585	0,082	0,0125	0	0	0
42	0,0115	0,042	0,011	0,015	0,027	0,157	0,6435	0,059	0,052
43	0,0155	0,051	0,0095	0,0295	0,036	0,146	0,0065	0	0
44	0,012	0,004	0,0445	0,052	0,015	0,015	0,548	0	0
45	0,0375	0,0565	0,0265	0,0575	0,0525	0,0185	0,411	0	0,0135
46	0,0055	0,0395	0,0005	0,011	0,0185	0,145	0,3945	0,0465	0,0565
47	0,006	0,009	0,001	0,0085	0,0235	0,0295	0,0495	0	0,01
48	0,028	0,0415	0,018	0,025	0,0295	0,2685	0,1735	0,0495	0
49	0,029	0,0375	0,0105	0,033	0,05	0,0035	0,3035	0,042	0,002
50	0,032	0,0445	0,0285	0,058	0,0295	0,183	0,1495	0,048	0,0005
51	0,01	0,0435	0,007	0,0335	0,031	0,2055	0,2615	0,0515	0
52	0,027	0,0465	0,009	0,046	0,0575	0,022	0,9355	0	0
53	0,021	0,043	0,0525	0,023	0,0315	0,0165	0,36	0	0,0515
54	0,0445	0,048	0,0205	0,026	0,024	0,1265	0,491	0,046	0,0045

## ANEXO VIII

Valores de ocupação das células presentes no modelo 1B9 para os análogos propostos do raloxifeno.

	<b>(2,11,4)(qa)</b>	<b>(0,-2,-1)(qa)</b>	<b>(0,1,-2)(qa)</b>	<b>(0,11,5)(qa)</b>	<b>(-1,8,2)(qa)</b>	<b>(1,6,-2)(ap)</b>	<b>(1,13,1)(qa)</b>	<b>(1,2)(qa)(qa)</b>	<b>(0,0,-2)(qa)</b>
<b>A1</b>	0	0,013	0,0015	0,003	0,3745	0	0,025	0,306	0,195
<b>A2</b>	0,0055	0,0205	0	0,0405	0,337	0	0	0,317	0,179
<b>A3</b>	0,006	0,0095	0,002	0,036	0,339	0	0,018	0,318	0,5995
<b>A4</b>	0,0025	0,017	0,001	0,01	0,3055	0	0	0,385	0,25
<b>A5</b>	0,01	0,0075	0,0025	0,0125	0,3545	0	0,031	0,437	0,4735
<b>A6</b>	0,0025	0,0095	0,0035	0,0205	0,3115	0	0	0,4255	0,426
<b>B1</b>	0,0005	0	0,0365	0,1375	0,0605	0	0	0,3185	0,251
<b>B2</b>	0,0135	0	0,0925	0,141	0,0535	0	0	0,267	0,701
<b>B3</b>	0,0035	0	0,1195	0,2535	0,0155	0	0	0,114	0,6145
<b>B4</b>	0,0165	0	0,109	0,1395	0,0185	0	0	0,141	0,612
<b>C1</b>	0,0015	0,002	0,0015	0,0095	0,304	0	0	0,6615	0,6265
<b>C2</b>	0,0025	0,002	0,0025	0,0055	0,301	0	0,076	0,652	0,5415

# 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](#)