



**ELISÂNGELA ALVES DE FREITAS**

**ESTUDOS DE CORRELAÇÕES  
QUANTITATIVAS ESTRUTURA-ATIVIDADE  
ENTRE INIBIDORES NÃO ESTEROIDAIIS DA  
AROMATASE HUMANA**

**LAVRAS - MG**

**2011**

**ELISÂNGELA ALVES DE FREITAS**

**ESTUDOS DE CORRELAÇÕES QUANTITATIVAS ESTRUTURA-  
ATIVIDADE ENTRE INIBIDORES NÃO ESTEROIDAIIS DA  
AROMATASE HUMANA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, para obtenção de título de Mestre.

Orientadora

Dra. Elaine Fontes Ferreira da Cunha

**LAVRAS - MG**

**2011**

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da  
Biblioteca da UFLA**

Freitas, Elisângela Alves de.

Estudos de correlações quantitativas estrutura-atividade entre  
inibidores não esteroidais da aromatase humana / Elisângela Alves  
de Freitas. – Lavras : UFLA, 2011.

74 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2011.

Orientador: Elaine Fontes Ferreira da Cunha.

Bibliografia.

1. Química medicinal computacional. 2. Modelagem molecular.  
3. Câncer de mama. 4. Compostos não esteroidais. 5. QSAR 5D/6D.  
I. Universidade Federal de Lavras. II. Título.

CDD – 615.1901

**ELISANGELA ALVES DE FREITAS**

**ESTUDOS DE CORRELAÇÕES QUANTITATIVAS ESTRUTURA-  
ATIVIDADE ENTRE INIBIDORES NÃO ESTEROIDAIIS DA  
AROMATASE HUMANA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, para obtenção de título de Mestre.

APROVADA em 22 de fevereiro de 2011.

Dr. Teodorico de Castro Ramalho UFLA

Dra. Suzan Kelly V. Bertolucci UFLA

Dra. Elaine Fontes Ferreira da Cunha  
Orientadora

**LAVRAS - MG**

**2011**

*Dedico este trabalho às pessoas que fazem uso do conhecimento para encontrar uma cura para o câncer, dedicando suas vidas ao estudo, com objetivo de proporcionar qualidade de vida as mulheres que perderam a esperança de viver devido ao câncer de mama.*

## AGRADECIMENTOS

A Deus, o Autor da vida, por me permitir admirar cada dia mais as suas obras e descobrir a quão maravilhosa e dinâmica é a vida, iluminando os meus passos e por proporcionar o fim de mais uma etapa em minha vida.

Aos meus pais, José Antônio e Neuza, pelo amor, confiança e dedicação.

Aos meus amados irmãos Wellington, Edneide e Milene, pela força, carinho e ajuda que sempre me deram.

Ao meu amado noivo Luís que me ajudou desde o início desta etapa até o fim.

A toda a minha família, pela credibilidade, amizade e por participar nos momentos mais importantes.

A minha amiga Aline que me auxiliou em tudo neste trabalho, pela paciência e dedicação que teve comigo.

A galera da Química Computacional, ou melhor, o famoso “Bloquinho” que me acolheram com muito carinho.

Aos professores Dr. Teo por acreditar em mim e me ajudar nas minhas dificuldades e a Dra. Elaine por me orientar, serei eternamente grata por tudo que fizeram por mim.

À Universidade Federal de Lavras e ao Departamento de Pós-Graduação em Agroquímica, pela oportunidade de realização deste trabalho.

A CAPES, pela bolsa concedida.

Agradeço imensamente.

## RESUMO

O câncer é uma doença com alto índice de mortalidade atingindo pessoas de diferentes raças, costumes, condições financeiras e sociais. Entre as mulheres o câncer de mama é o tipo mais comum. No Brasil, o número de novos casos deste tipo de câncer esperados para 2010 e 2011 é de 49.240, com um risco de 49 casos a cada 100 mil mulheres, consoante estudos realizados pelo Instituto Nacional de Câncer (INCA). Compostos não esteroidais têm sido estudados para o tratamento de câncer de mama hormônio-dependente em mulheres na pós-menopausa. Esses compostos são capazes de inibir a ação da enzima aromatase, que produz estrogênio, um hormônio altamente relacionado com o crescimento do tumor. Estudos de correlações quantitativas estrutura-atividade (QSAR) foram realizados usando QSAR-5D (*RAPTOR, Receptor as Poly Tier Object Representation*) e QSAR-6D (*QUASAR, The quase-atomistic receptor modeling*), permitindo sugerir características relacionadas ao sítio de ligação da aromatase para o desenvolvimento e/ou aperfeiçoamento de fármacos para tratamento do câncer de mama. Foram selecionados 43 compostos não esteroidais, divididos em 35 compostos para o grupo de treinamento e 08 compostos para o grupo de teste. O *software Symposar* gerou um conjunto de conformações para cada ligante e seu alinhamento. Do *software Raptor* obteve-se do melhor modelo um  $R^2$  (desvio médio) de 0,750 e capacidade preditiva de 72,1% ( $q^2$ ) para o grupo de treinamento. A partir do grupo de teste validou-se externamente o modelo obtendo-se um  $r^2$  preditivo de 0,791. Do *software Quasar* obteve-se do melhor modelo  $R^2$  (desvio médio) de 0,772 e capacidade preditiva de 73,60% ( $q^2$ ) para o grupo de treinamento. A partir do grupo de teste validou-se externamente o modelo obtendo-se um  $r^2$  preditivo de 0,791. Os estudos de QSAR 5D e 6D demonstraram consenso sobre as regiões hidrofóbicas do sítio ativo da enzima. Substituições polares na posição R3 de compostos análogos ao YM511 e a substituintes polares na posição R2 de compostos análogos ao Letrozol podem potencializá-los, mantendo assim o grupo triazol, benzil e ciano. De acordo com os resultados os modelos obtidos abrangem importantes aspectos os quais podem ser aplicados na modificação estrutural ou no desenvolvimento de novos inibidores da enzima aromatase.

Palavras-chave: QSAR-5D. QSAR-6D. Enzima Aromatase. Câncer de mama. Química medicinal.

## ABSTRACT

Cancer is a disease with a high mortality rate affecting people of different races, customs and social classes. Among women, breast cancer is the most common type. In Brazil the number of new cases of cancer expected in 2010 and 2011 is 49,240, with a risk of 49 cases per 100,000 women, according to studies by the National Cancer Institute (INCA). Non-steroidal compounds have been studied for the treatment of hormone-dependent breast cancer in postmenopausal women. These compounds are able to inhibit the enzyme aromatase which produces estrogen, a hormone that is highly correlated with tumor growth. Quantitative structure-activity relationship studies (QSAR) were performed using 5D-QSAR (*RAPTOR*, Receptor As Poly Tier Object Representation) and 6D-QSAR (*QUASAR*, The quasi-atomistic receptor modeling), suggesting features that are related to the binding site for the enzyme aromatase development and/or enhancement of medicines for the treatment of breast cancer. Forty three non-steroidal compounds were selected, divided into thirty five compounds for the training group and eight compounds for the test group. The *Symposar* software generated a group of conformations for each ligand and its alignment. From the *Raptor* software it was concluded from the best model a  $R^2$  (mean deviation) of 0.750 and a predictive capacity of 72.1% ( $q^2$ ) for the training group. From the test group, the model was externally validated by obtaining a predictive  $r^2$  of 0.791. The *Quasar* software presented from the best model a  $R^2$  (mean deviation) of 0.772 and a predictive ability of 73.60% ( $q^2$ ) for the training group. In regards to the testing group, the model was externally validated by obtaining a predictive  $r^2$  0.791. The QSAR studies 5D and 6D showed consensus on the hydrophobic regions of the active site of the enzyme c. Polar substitutions at the R3 position of compounds analogous to YM511 and polar substituents at the R2 position of compounds, analogous to Letrozole can enhance them, thus keeping the triazole group, benzyl and cyano. According to the results obtained, the models cover important aspects which can be applied to structural change or development of new inhibitors of the enzyme aromatase.

Keywords: 5D-QSAR. 6D-QSAR. Enzyme Aromatase. Breast cancer. Medicinal chemistry.

## LISTA DE ILUSTRAÇÃO

Figura 1	Anatomia da mama, visão em corte lateral.....	23
Figura 2	Estrutura química dos estrogênios: (1) Estrutura básica dos estrogênios, (2) 17 $\beta$ - estradiol, (3) Estrona e (4) Sulfato de estrona.....	25
Figura 3	Biossíntese dos estrogênios, nos ovários, a partir de uma molécula de colesterol.....	26
Figura 4	Conversão de androstenediona a estrona, catalisada pela enzima aromatase.....	28
Figura 5	Imagem da estrutura cristalográfica do citocromo P450 aromatase da placenta humana em complexo com androstenediona.....	29
Figura 6	Estrutura química do Tamoxifeno.....	30
Figura 7	Estruturas químicas de inibidores da enzima aromatase.....	31
Figura 8	Esquema das duas etapas no processo de alinhamento.....	36
Figura 9	Alinhamento e conformações do modelo A utilizando o software Symposar.....	53
Figura 10	Representação de dupla camada do modelo A obtido por estudos de 5D-QSAR, utilizando o composto E43 como referência. À esquerda, regiões hidrofóbicas e à direita, regiões hidrofílicas doadoras de ligação de hidrogênio.....	59
Figura 11	Representação da superfície mapeada pelas propriedades "quasi atomística" do modelo do receptor, gerada pelo Quasar utilizando, à esquerda a estrutura tridimensional do composto E43 e à direita todos os compostos. As cores significam: castanho = muito hidrofóbico; amarelo = hidrofóbico; cinza = neutro ou hidrofóbico; vermelho = acceptor de ligação de hidrogênio; azul = doadora de	63

	ligação de hidrogênio	
Figura 12	Estereorepresentação do composto E43. À esquerda de acordo com a contribuição do grupo funcional ao $\Delta G$ (azul = muito bom, verde=bom, branco=neutro, amarelo=fraco, vermelho=ruim, cinza=não resolvido), à direita representação da superfície mapeado do receptor.....	64
Figura 13	Visualização gráfica das contribuições à energia livre calculada dos compostos análogos ao YM511 (azul = muito bom, verde=bom, branco=neutro, amarelo=fraco, vermelho=ruim, cinza=não resolvido).....	65
Gráfico 1	Tipos de cânceres mais incidentes estimados para 2010 na população brasileira, exceto o de pele não melanoma.....	21
Gráfico 2	Disposição dos compostos dos grupos de treinamento e teste de acordo com pIC50.....	50
Gráfico 3	Valores de pIC50 experimental e calculado dos 43 inibidores não esteroidais da enzima aromatase.....	58
Gráfico 4	Valores da atividade biológica experimental e calculada.....	62

## LISTA DE TABELAS

Tabela 1	Estimativas de casos novos de câncer para o ano 2010 por região*- Brasil- 2009.....	22
Tabela 2	Dimensões usadas em QSAR.....	33
Tabela 3	Análogos ao Letrozol (composto33).....	44
Tabela 4	Análogos ao YM511 (composto 40).....	45
Tabela 5	Anastrozol.....	48
Tabela 6	Distribuição dos compostos de acordo com sua atividade biológica em ordem crescente de pIC50.....	49
Tabela 7	Números de confôrmeros para cada modelo.....	52
Tabela 8	Resultados estatísticos dos modelos de QSAR obtidos com compostos não esteroidais.....	54
Tabela 9	Valores de atividade e energia livre (kcal mol-1) de compostos do grupo de treinamento e de teste referentes ao melhor modelo obtido a partir de compostos não esteroidais.....	55
Tabela 10	Valores de atividade e energia livre (kcal mol-1) de compostos do grupo de treinamento e teste.....	60

## LISTA DE ESTRUTURAS QUÍMICAS

(1) Estrutura básica dos estrogênios	25
(2) 17 $\beta$ - estradiol	25,26
(3) Estrona	25,26,28
(4) Sulfato de Estrona	25
(5) Colesterol	26
(6) Pregnenolona	26
(7) Hidroxipregnenolona	26
(8) Progesterona	26
(9) Desidroepiandrosterona	26
(10) Hidroxiprogesteronona	26
(11) Androstenediona	26,28
(12) Testosterona	26
(13) 19 Hidroxiandrostenediona	28
(14) 19,19 Dihidroxiandrostenediona	28
(15) 19 Oxiandrostenediona	28
(16) Intermediário peróxi-enzima	28
(17) Tamoxifeno (Z)-2-[4-(1,2-difenilbut-1-enil) penoxi]-N, N-dimetil-etanamina	30
(18) Aminoglutetimida	31
(19) Formestano	31
(20) Fadrozol	31
(21) Exemestano	31
(22) Anastrozol (2-[3-(ciano-dimetil-metil)-5[1, 2,4] triazol-1-ilmetil-fenil]2-metil-propionitrila) - Composto E29	31,48

(23)Letrozol	(4-[4-(cianofenil)-(1,2,4-triazol-1-ila)	metil]	
	benzonnitrila)-	Composto E33	31, 44
(24)Vorozol			31
(25)Composto E01			44
(26)Composto E02			44
(27)Composto E03			44
(28)Composto E04			45
(29)Composto E05			44
(30)Composto E06			44
(31)Composto E07			45
(32)Composto E08			45
(33)Composto E09			45
(34)Composto E10			45
(35)Composto E11			45
(36)Composto E12			45
(37)Composto E13			45
(38)Composto E14			44
(39)Composto E15			44
(40)Composto E16			45
(41)Composto E17			44
(42)Composto E18			44
(43)Composto E19			45
(44)Composto E20			44
(45)Composto E21			46
(46)Composto E22			46
(47)Composto E23			46
(48)Composto E24			46

(49)Composto E25	46
(50)Composto E26	46
(51)Composto E27	44
(52)Composto E28	46
(53)Composto E30	46
(54)Composto E31	46
(55)Composto E32	46
(56)Composto E34	47
(57)Composto E35	47
(58)Composto E36	47
(59)Composto E37	47
(60)Composto E38	47
(61)Composto E39	47
(62)YM511(4-[(4-bromobenzil)-(1,2,4-triazol-4-ilamino)] trila)- Composto E40	47
(63)Composto E41	47
(64)Composto E42	47
(65)Composto E43	47

## LISTA DE ABREVIATURAS

UFLA	Universidade Federal de Lavras
AIDS/SIDA	<i>Acquired Immunodeficiency Syndrome/Síndromeda</i> Imunodeficiência Adquirida
RNA	Ácido ribonucleico
IARC	<i>International Agency for Research on Cancer</i>
WHO/OMS	<i>World Health Organization/Organização Mundial da Saúde</i>
INCA	Instituto Nacional de Câncer
RE	Receptor de estrogênio
CYP450	Citocromo P450
IUBMB	International Union of Biochemistry and Molecular Biology
EC	Enzyme Commission (Comissão sobre Enzimas)
NADPH	Nicotinamida Adenina Dinucleotídeo-P
NADH	Nicotinamida Adenina Dinucleotídeo
PDB	<i>Protein Data Bank</i> (Banco de dados de proteínas)
QSAR	Quantitative Structure-Activity Relationship (Relação quantitativa estrutura-atividade)
KI	Constante de equilíbrio de inibição
IC50	Concentração de inibição máxima em 50%
pKa	Potencial de dissociação de um ácido
logP	Coefficiente de partição (octanol/água)
3D	Tridimensional
OECD	Organisation for Economic Co-operation and Development
RAPTOR	<i>Receptor as Poly Tier Object Representation</i> (Representação de receptor por camada)
JEG-3	<i>Human Placental Choriocarcinoma Cell Line</i> (Linha Celular Coriocarcinoma Humana Placentária)
LoF	<i>Lack of fit</i> (falta de ajuste)

## LISTA DE SÍMBOLOS

$\sigma$	Sigma
$\Delta G^\circ$	Varição de energia livre de Gibbs padrão
R	Constante universal dos gases perfeitos – $1,987 \cdot 10^{-3} \text{kcal. K}^1 \cdot \text{mol}^{-1}$
T	Temperatura – 298K (25 °C)
$\Delta S$	Varição de entropia
$\alpha$	Alfa
$R^2$	Coefficiente de determinação
$r^2$	Coefficiente de determinação (preditividade)
$q^2$	Validação cruzada por <i>Leave-one-out</i>

## SUMÁRIO

1	INTRODUÇÃO.....	17
2	REFERENCIAL TEÓRICO.....	19
2.1	Câncer.....	19
2.2	Estimativas de câncer.....	20
2.3	Câncer de mama.....	22
2.4	A enzima Aromatase.....	27
2.5	Inibidores da Aromatase.....	29
2.6	Correlação quantitativa estrutura atividade e a contribuição no desenvolvimento de novos medicamentos.....	32
2.7	Estudos de QSAR.....	34
2.8	<i>Software BiografX 1.1</i> .....	35
2.9	<i>Software Symposar</i> .....	35
2.10	<i>Software Raptor</i> .....	36
2.11	<i>Software Quasar</i> .....	38
2.12	Validação.....	41
3	MATERIAL E MÉTODOS.....	43
3.1	Inibidores da Aromatase.....	43
3.2	Construções dos modelos tridimensionais e cálculo das cargas parciais atômicas.....	50
3.3	Validação dos modelos.....	51
4	RESULTADOS E DISCUSSÕES.....	52
4.1	Resultados do <i>Symposar</i> .....	52
4.2	Resultados do <i>RAPTOR (5D)</i> .....	53
4.3	Resultados do <i>QUASAR (6D)</i> .....	59
5	CONCLUSÃO.....	66
6	PERSPECTIVAS.....	67
	REFERÊNCIAS.....	68

## 1 INTRODUÇÃO

Atualmente o câncer é a doença que apresenta o maior índice de mortalidade superando a AIDS, a tuberculose e a malária em conjunto no número de mortes (GARCIA et al., 2007).

A origem do câncer ocorre por meio de alterações, geralmente somáticas, nos genes supressores de tumores e genes de microRNA, oncogenes e também pela ocorrência de ação de agentes carcinogênicos (químicos, físicos ou biológicos) (CROCE, 2008).

Diversos métodos para o tratamento do câncer foram desenvolvidos como: radioterapia, quimioterapia, hormonioterapia e a cirurgia. A hormonioterapia tem sido amplamente utilizada no tratamento de câncer de mama em mulheres pós-menopausa hormônio dependente (RUGO, 2008).

A enzima aromatase tem sido alvo no desenvolvimento de novos fármacos, por ter um papel importante na biossíntese dos estrogênios, um hormônio que está intrinsecamente relacionado ao crescimento do tumor de mama (HISCOX; DAVIES; BARRET-LEE, 2009).

Métodos computacionais em química têm desempenhado um papel importante no processo de descoberta e desenvolvimento de novos fármacos. Um método que tem sido amplamente utilizado é o estudo quantitativo da correlação entre a estrutura e a atividade de fármacos. Ele tem sido usado para considerar qualidades e potencialidades *in vitro* de um composto que apresenta atividade biológica, auxiliando na orientação de qual característica (grupos farmacofóricos) da molécula contribui para a sua atividade e, também modificar compostos melhorando suas propriedades físico-químicas (MYINT; QUN-XIE, 2010).

Neste trabalho, procura-se sugerir características importantes relacionadas ao sítio de ligação enzimática através da técnica QSAR 5D e 6D

usando 43 compostos não esteroidais, das quais possam contribuir na elaboração de novos medicamentos para tratamento do câncer de mama hormônio dependente em mulheres pós menopáusicas.

## 2 REFERENCIAL TEÓRICO

### 2.1 Câncer

Todos os organismos são formados por células, desde as bactérias até o homem (SPENCE, 1991). São pequenas unidades de membranas delimitadas, preenchidas com uma concentrada solução aquosa de substâncias químicas, com capacidade extraordinária de criar cópias de si mesmas (ALBERTS; BRAY, 2006). As células eucariontes (grego *eu* que significa verdadeiro e *karyon* (kernel) núcleo) são constituídas por: membrana plasmática, citoplasma, organelas e núcleo (SPENCE, 1991).

As células saudáveis estão sob um estrito controle bioquímico para crescerem e se diferenciarem. Isto acontece sob a influência de vários estimuladores de crescimento e de morte programada (apoptose). Nos processos anormais em que as células crescem e dividem-se incontrolavelmente, perdem a estrutura e função devido à incapacidade de se diferenciarem adequadamente (FOYE et al., 2008). Assim, o câncer é uma doença genética, resultante de mutações que afetam genes que controlam as funções das células (proto-oncogenes e genes supressores de tumores) ou da atividade de genes polimórficos que regem com sistemas enzimáticos (fase I e II das reações enzimáticas) (LENHARD; OSTEEEN; GANSLER, 2001).

A palavra câncer vem do grego que significa caranguejo, o termo mais adequado à doença, visto que as projeções de propagação do tumor, os quais podem invadir tecidos próximos, lembram a forma de um caranguejo (AMERICAN CANCER SOCIETY, 2010).

As células cancerosas também podem se espalhar para outras partes do corpo através do sangue e sistema linfático, ocorrendo à metástase. O câncer é classificado de acordo com o local de sua ocorrência, a saber:

- a) Carcinoma: tumor que tem início na pele ou nos tecidos que cobrem os órgãos internos;
- b) Sarcoma: tumor que começa no osso, cartilagem, gordura, músculo, vasos sanguíneos, ou do tecido conjuntivo ou de suporte;
- c) Linfoma e mieloma múltiplos: tumores que se iniciam nas células do sistema imunológico e;
- d) Tumores do sistema nervoso central: aqueles que começam nos tecidos do cérebro e da medula espinhal (NATIONAL CANCER INSTITUTE, 2010).

## **2.2 Estimativas de câncer**

Segundo recente relatório da Agência Internacional para Pesquisa em Câncer - IARC, o impacto global do câncer mais que dobrou em 30 anos (WORLD HEALTH ORGANIZATION, 2005). A ocorrência de novos casos, para o ano de 2008, estimado pela IARC/WHO (2005) foi de aproximadamente 12,4 milhões de novos casos e 7,6 milhões de óbitos devido ao câncer no mundo. Destes, os mais incidentes foram os cânceres de pulmão (1,52 milhões de casos novos), mama (1,29 milhões), cólon e reto (1,15 milhões) (INSTITUTO NACIONAL DE CÂNCER, 2009).

O tipo de câncer mais comum entre os homens é o câncer de próstata, seguido por pulmão, estômago, cólon e reto. O tipo mais incidente entre as mulheres é o câncer de mama, antecedendo os de colo do útero, cólon, reto, estômago e pulmão.

As estimativas de acordo com INCA (2009) para o ano de 2010 no Brasil, válidas também para o ano de 2011, apontam à ocorrência de 489.270 novos casos de câncer, sendo os tipos mais incidentes os cânceres de próstata e

de pulmão no sexo masculino e os cânceres de mama e do colo do útero no sexo feminino (Gráfico 1) seguindo o mesmo perfil da magnitude observada para a América Latina (INCA, 2009).

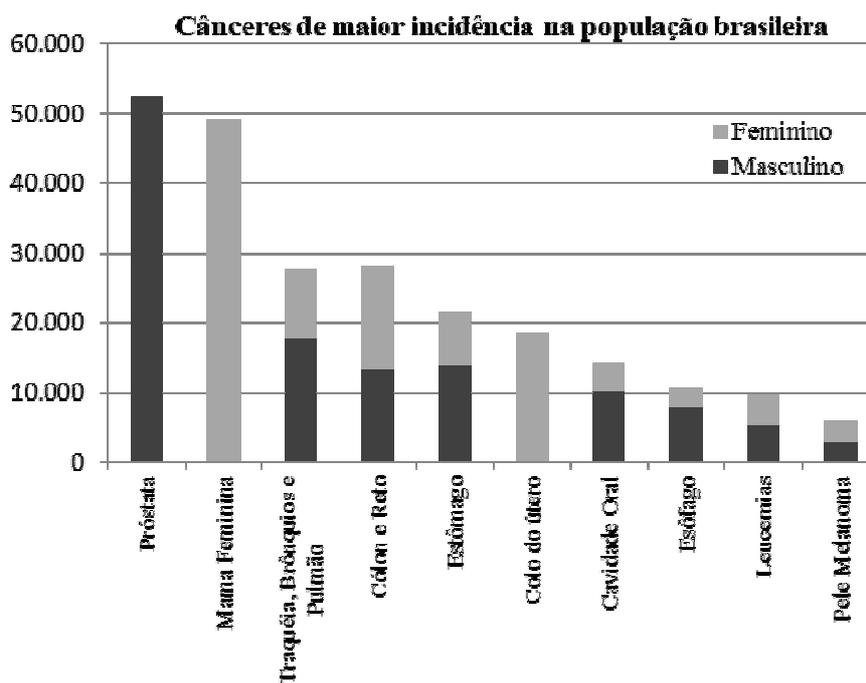


Gráfico 1 Tipos de cânceres mais incidentes estimados para 2010 na população brasileira, exceto o de pele não melanoma

Fonte: Adaptada de INCA (2009)

Dados do INCA (2009) de distribuição regional de novos casos de cânceres entre os brasileiros (Tabela 1) evidenciam o número de maior probabilidade de novos casos de câncer de mama na região Sudeste.

Tabela 1 Estimativas de casos novos de câncer para o ano 2010 por região\*  
Brasil- 2009.

Localização Primária	Norte	Nordeste	Centro-Oeste	Sul	Sudeste
Mama Feminina	1,350	8,270	2,690	9,310	27,620
Traquéia, Brônquio e Pulmão	1,080	3,950	1,760	7,230	13,610
Estômago	1,300	4,280	1,270	4,090	10,560
Próstata	1,960	11,570	3,430	9,820	25,570
Colo do Útero	1,820	5,050	1,410	3,110	7,040
Cólon e Reto	620	3,040	1,580	6,150	16,720
Esôfago	260	1,530	580	3,040	5,220
Leucemias	560	2,070	650	1,790	4,510
Cavidade Oral	410	2,810	800	2,510	7,590
Pele Melanoma	180	540	250	2,020	2,940
Outras Localizações	5,260	14,780	8,090	28,810	80,960
Subtotal	14,800	57,890	22,510	77,880	202,340
Pele não Melanoma	4,320	31,460	7,830	24,600	45,640
<b>TOTAL</b>	<b>19,120</b>	<b>89,350</b>	<b>30,340</b>	<b>102,480</b>	<b>247,980</b>

\* Números arredondados para 10 ou múltiplos de 10

Fonte: Adaptada de INCA (2009)

### 2.3 Câncer de mama

As glândulas mamárias são elevações hemisféricas cobertas de pele localizadas na superfície dos músculos peitorais maiores. Internamente, cada glândula é constituída de tecido adiposo mantido por um estroma conjuntivo, os quais separam a gordura do tecido glandular e o contorno liso das mamas (SPENCE, 1991, p.43). Existem de 15 a 20 lobos centrais, cada lobo é drenado por um único ducto lactífero que se abre no mamilo. O tecido glandular estende-se também para trás da mama, nas axilas. A Figura 1 ilustra a forma de uma glândula mamária. Após a puberdade, sob estimulações de estrógenos e

progesterona, as glândulas mamárias começam seu desenvolvimento (SPENCE, 1991).

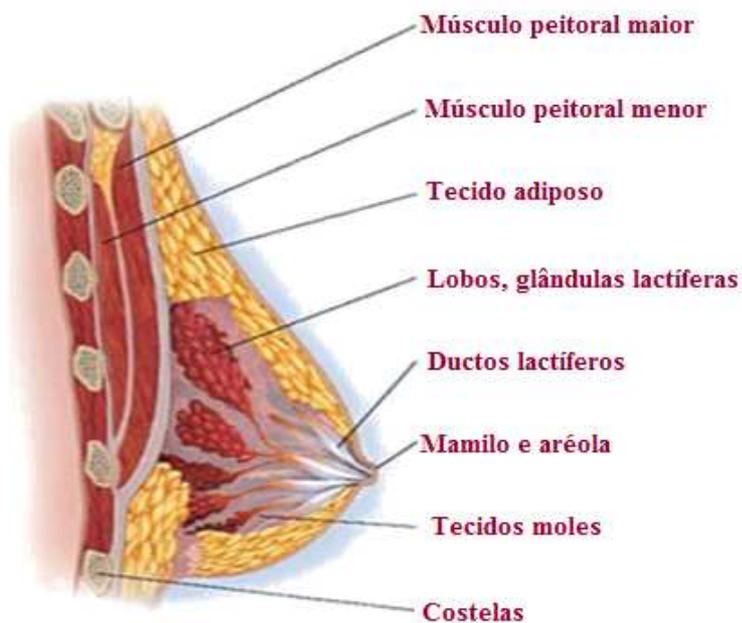


Figura 1 Anatomia da mama, visão em corte lateral

Fonte: Adaptada de Casa de amparo a mãe e ao bebê e incentivo a amamentação (2010)

O estudo desenvolvido por Lotti et al. (2008) relata o impacto que o câncer de mama exerce na mulher, as quais passam por momentos de angústia, sofrimento e ansiedade desde quando diagnosticadas e durante o tratamento (LOTTI et al., 2008).

Segundo Bijus et. al (2008), o câncer de mama é a principal causa de morte entre as mulheres, porém com diagnóstico precoce e melhorias de tratamento há uma redução na redução da taxa de mortalidade (BIJUS et.al, 2008).

Diversos são os fatores que aumentam o risco de desenvolvimento do câncer de mama. Dentre eles o histórico familiar e a predisposição genética, correspondem aos 5% dos casos. Outros fatores relacionados à vida reprodutiva da mulher, tais como: menarca precoce, nuliparidade, idade da primeira gestação acima dos 30 anos, anticoncepcionais orais, menopausa tardia e terapia de reposição hormonal. Porém, a idade continua sendo um dos fatores de risco mais importante, aumentando rapidamente a taxa de incidência até aos 50 anos e após diminuindo lentamente, sendo atribuída à menopausa esta mudança na taxa de incidência (INCA, 2007).

Dois tipos estão relacionados com a idade: o tipo que ocorre na pré-menopausa caracterizado por ser mais agressivo e receptor estrogênico negativo (RE<sup>-</sup>); e o que ocorre na pós-menopausa associado com características insensíveis a dor e, principalmente, por ser receptor estrogênico positivo (RE<sup>+</sup>).

Aproximadamente dois terços das pacientes com câncer de mama pós-menopausa expressam RE<sup>+</sup>, o que contribui para o crescimento do tumor devido ao aumento de estrogênio (JONG et al., 1997; BRUEGGEMEIER; HACKETT; DIAZ-CRUZ, 2005).

Os estrogênios estão envolvidos em muitos processos fisiológicos como o desenvolvimento e manutenção de órgãos sexuais femininos, ciclo reprodutivo, reprodução e funções neuroendócrinas (TURGEON et.al, 2004). Nas células, os estrogênios atuam na ativação de genes: que interagem com receptores no citoplasma (JONG et al., 1997). Este complexo é transportado para dentro do núcleo celular, ativando genes específicos, o que leva a síntese de RNA<sub>mensageiro</sub> e, conseqüentemente produzindo proteínas (ex. as enzimas) que influenciam em diversas reações e processos celulares. Os estrogênios podem agir de duas formas na carcinogênese: aumentando o número de células suscetíveis ou parcialmente transformadas por carcinogêneos, ou estimulando o crescimento seletivo de células malignas (SPENCE, 1991).

Simpson e Dowsett (2002), em seus estudos, fazem a diferenciação entre a principal fonte de estrogênio em mulheres na pré-menopausa, que são os ovários e as mulheres na pós-menopausa, os quais atuam em nível local, produzidos em vários lugares como tecido adiposo, osteoblastos, músculos lisos e em inúmeros lugares no cérebro (SIMPSON; DOWSETT, 2002).

Os estrogênios que ocorrem naturalmente possuem quatro anéis fundidos, no qual o anel A é aromático, com o grupo hidroxila no átomo de carbono C-3, uma ramificação metila nos átomos de carbono C-13 e C-10 conforme exemplificado na Figura 2 (PARL, 2000).

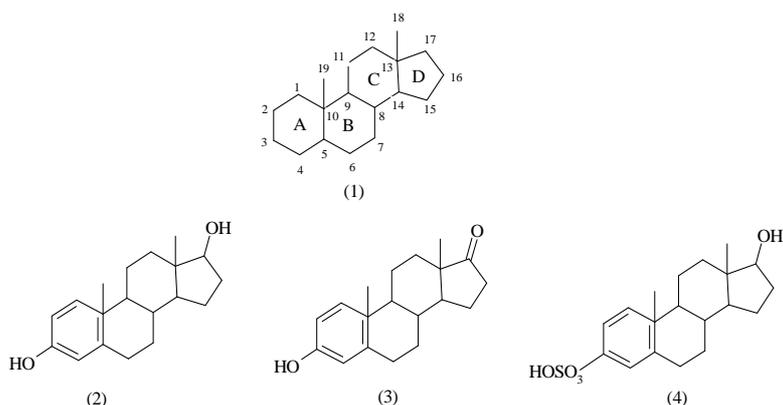


Figura 2 Estrutura química dos estrogênios: (1) Estrutura básica dos estrogênios, (2)17β- estradiol, (3) Estrona e (4) Sulfato de estrona  
Fonte: Adaptada de Parl (2000)

A biossíntese dos estrogênios envolve uma longa série de passos enzimáticos (PARL, 2000). Em resumo, a partir da molécula de colesterol ocorre a clivagem parcial da cadeia lateral do anel D, originando os esteróides de 21 átomos de carbonos os corticóides e progestágenos. Posteriormente, a cadeia lateral do anel D é totalmente removida formando os andrógenos (androstenediona e testosterona) e, por fim, com a atuação da enzima aromatase, o anel A passa a ser aromático, formando os estrogênios

(SIMPSON;DOWSETT, 2002). A Figura 3 esquematiza a biossíntese dos estrogênios a partir de uma molécula de colesterol no corpo lúteo no ovário.

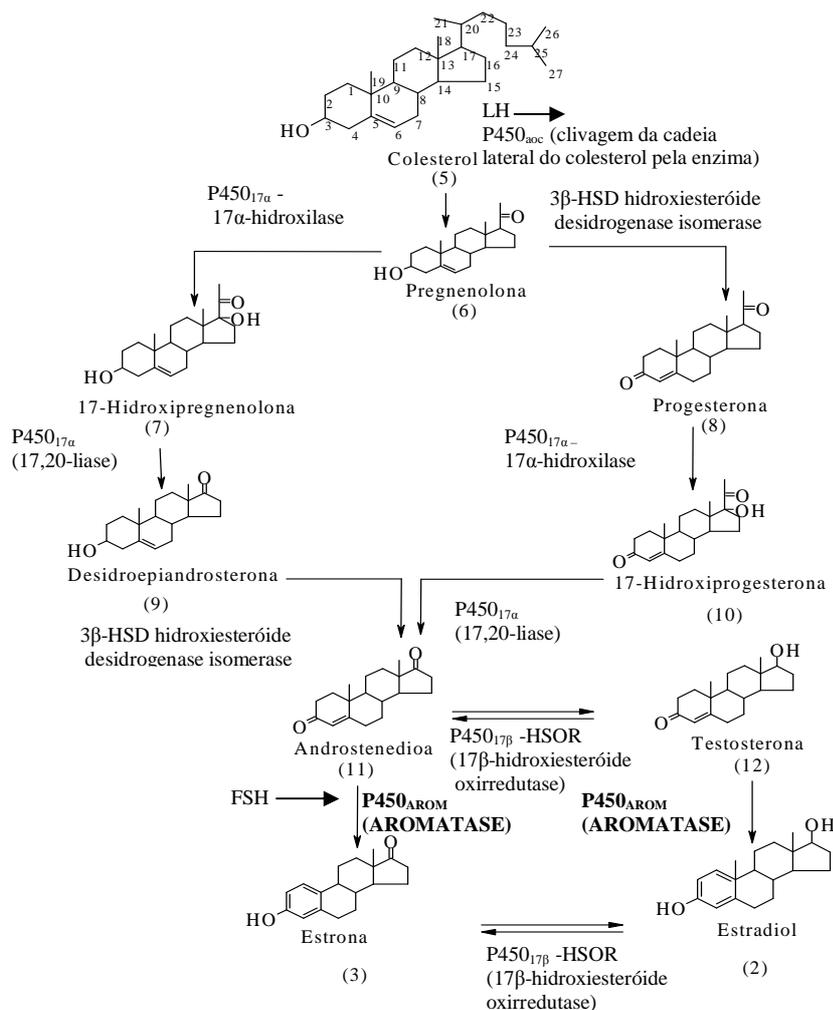


Figura 3 Biossíntese dos estrogênios, nos ovários, a partir de uma molécula de colesterol

Fonte: Adaptada de Universidade Federal de São Paulo- UNIFESP (2010)

## 2.4 A enzima Aromatase

De acordo com a classificação do *International Union Of Biochemistry And Molecular Biology* (IUBMB) a enzima aromatase (EC 1.14.14.1), corresponde à classe de enzimas óxido-redutases. É uma heme proteína, localizada no retículo endoplasmático liso, pertence ao grupo do citocromo P450 (CYP450), uma superfamília de mais de 480 genes distintos compondo 74 famílias, na qual a aromatase pertence à família 19 (SIMPSON; DOWSETT, 2002).

A aromatase está intimamente ligada à nicotinamida adenina dinucleotídeo fosfato citocromo redutase P-450(NADPH-CPR), que atua como doador de elétrons, o que é essencial para a reação de aromatização do anel A dos androgênios (Figura 2). Este processo pode ser dividido em três etapas principais (Figura 4), em que cada etapa exige um mol de oxigênio O<sub>2</sub> e um mol de NADPH-CPR redutase, as duas primeiras etapas consiste em hidroxilação do grupo metil do átomo de carbono C-19, e a terceira etapa é a aromatização do anel A, exclusivo a aromatase (GHOSH et al., 2009), produzindo apartir dos androgênios (androstenediona e testosterona) estrogênios (estrona e estradiol) (HISCOX; DAVIES; BARRET-LEE, 2009).

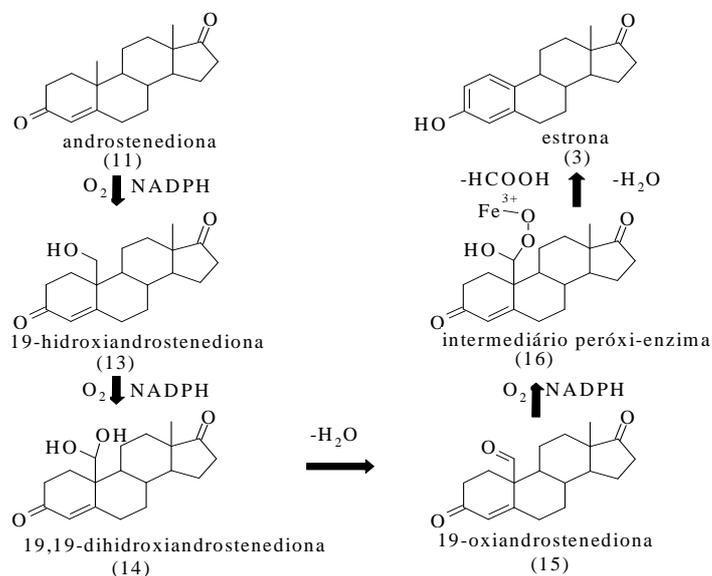


Figura 4 Conversão de androstenediona a estrona, catalisada pela enzima aromatase

Fonte: Adaptada de Hiscox, Davies e Barret-lee (2009)

A estrutura cristalográfica da aromatase foi elucidada no início de 2009 por Ghosh et al. (2009), a qual é constituída de uma cadeia polipeptídica de 503 resíduos de aminoácidos e uma porfirina. Possui 12 grandes  $\alpha$ -hélices e 10 folhas  $\beta$ -pregueadas distribuídas em três folhas menores e uma maior, característica do citocromo P450. A aromatase possui um sítio de ligação de caráter hidrofóbico específico a androgênios. A Figura 5 ilustra a estrutura da aromatase humana purificada da placenta, a uma resolução de 2,90 Å complexada com o seu substrato natural a androstenediona (GHOSH et al., 2009), disponível no *Protein Data Bank* (PDB) sob o código 3EQM (RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS PROTEIN DATA BANK, 2009).

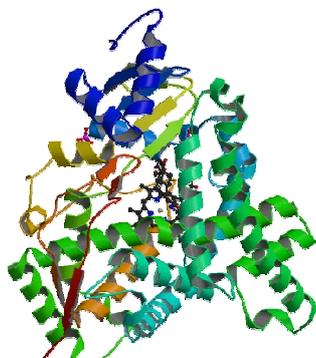


Figura 5 Imagem da estrutura cristalográfica do citocromo P450 aromatase da placenta humana em complexo com androstenediona

Fonte: RCSB PDB (2009)

Portanto, com o conhecimento das características da enzima aromatase, pode-se elaborar novos fármacos com maior eficácia atuando na produção de estrogênios (GHOSH et al., 2009).

## 2.5 Inibidores da Aromatase

As terapias endócrinas amplamente utilizadas no tratamento de câncer de mama em mulheres pós-menopausa hormônio dependente incluem moduladores seletivos do receptor estrogênio, tal como o (Z)-2-[4-(1,2-difenilbut-1-enil) penoxi]-N,N-dimetil-etanamina (Tamoxifeno), que impede a ligação do estrogênio ao receptor (Figura 6) (LEONETTI et al., 2004) e, mais recentemente, o uso de inibidores da aromatase que inibem a biossíntese destes hormônios (MILLER et al., 2008).

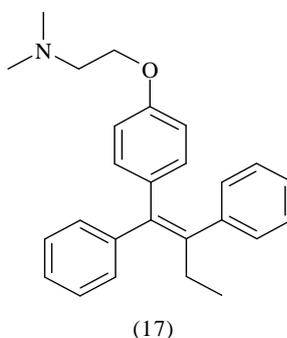


Figura 6 Estrutura química do Tamoxifeno  
 Fonte: Adaptada de Bermudez-Pirela et.al (2005)

O primeiro inibidor da aromatase desenvolvido em 1970 foi a aminoglutetimida (Figura 7), um inibidor não esteroidal que constituiu a primeira geração de inibidores da aromatase (ALTUNDAG; IBRAHIM, 2006). A segunda geração de inibidores compreendeu o inibidor esteroidal Formestano e o não esteroidal Fadrozol (Figura 7). Posteriormente, surgiu a terceira geração de inibidores que abrange o inibidor esteroidal Exemestano e os não esteroidais Anastrozol, Letrozol e Vorozol (BUZDAR, 2001) (Figura 7).

Os compostos esteroidais, também conhecidos como Tipo 1 (e.g. Exemestano) têm uma estrutura análoga a androstenediona –substrato natural da enzima (CAMPOS, 2004). Segundo Simpson (2009), o inibidor é convertido em um reativo intermediário, ligando-se de forma competitiva e irreversível ao sítio ativo da enzima. Estes tipos de inibidores referidos como inibidores suicidas possuem alta afinidade a enzima e eficácia prolongada, uma vez que a recuperação da atividade enzimática depende da re-síntese da enzima e a farmacocinética do fármaco (SIMPSON; DOWSETT, 2002).

Os compostos não esteroidais (Tipo 2), possuem um grupo triazol em sua estrutura básica, que lhes permite interagir com o átomo de ferro do grupo prostético da enzima (FAVIA et al., 2006). A alta afinidade destes compostos

com a enzima aromatase é determinada pelos outros aspectos estruturais e da forma que estes se ajustam ao sítio ativo. A aminoglutetimida, letrozol e vorozol, aparentemente possuem um efeito de preenchimento mais completo no sítio de ligação (SIMPSON; DOWSETT, 2002). Diversos estudos têm consolidado que o uso de inibidores da aromatase como alternativa no tratamento de câncer de mama em mulheres pós menopáusicas hormônio dependente, entretanto, o seu uso é restrito a mulheres que atingiram a menopausa ou que passaram por cirurgia de remoção dos ovários (BUZDAR, 2001; BUIJS et al., 2008). Geralmente são bem tolerados pelas pacientes, porém, é possível observar alguns efeitos colaterais como: enxaqueca, fadiga e um aumento em desenvolver osteoporose e fraturas ósseas (BUIJS et al., 2008).

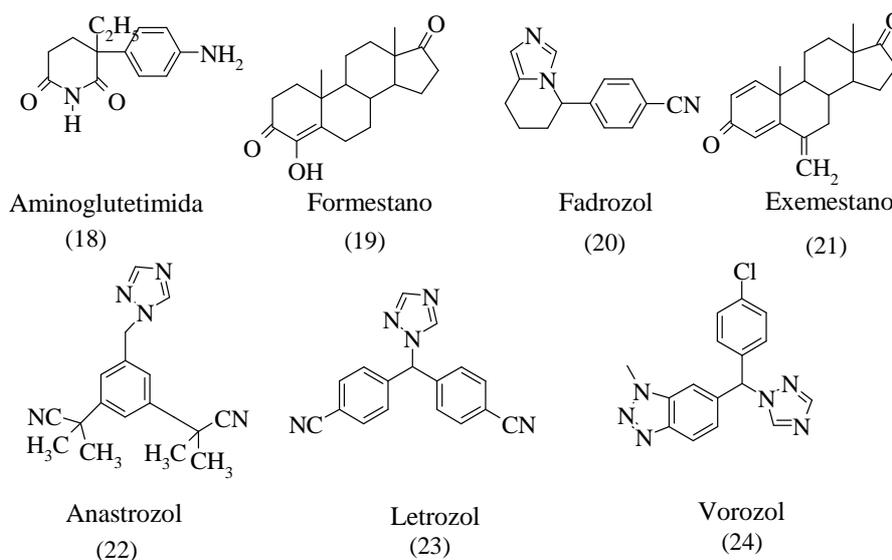


Figura 7 Estruturas químicas de inibidores da enzima aromatase  
 Fonte: Adaptada de Buzdar (2001), Altundag e Ibrahim (2006)

## **2.6 Correlação quantitativa estrutura atividade e a contribuição no desenvolvimento de novos medicamentos**

A correlação quantitativa estrutura atividade (QSAR) é um modelo matemático que relaciona as propriedades físico-químicas das moléculas à suas atividades biológicas, feito através de cálculos de descritores como: peso molecular, números de rotações de ligação, impedimento estérico, energia de interações eletrostáticas e log P (GAD, 2008), com a finalidade de identificar compostos ao planejamento de análogos mais específicos com maior atividade, assim aumentando as chances de sucesso de inserção dos novos fármacos no arsenal terapêutico (TAVARES, 2004).

Estudos de QSAR são aplicados em:

- a) Identificar novos compostos com atividades farmacológicas, pesticidas e biocidas;
- b) Aperfeiçoar as atividades farmacológicas, biocidas e pesticidas;
- c) Modelar de forma racional agentes surfactantes, perfumes, corantes e química fina;
- d) Identificar compostos perigosos em estágios iniciais do desenvolvimento ou na triagem de compostos existentes;
- e) Compreender efeitos colaterais e toxicidade em novos compostos;
- f) Predizer a toxicidade à exposição ocasional ou ocupacional em humanos;
- g) Predizer a toxicidade em ambientes;
- h) Selecionar compostos com propriedades farmacocinéticas desejáveis;
- i) Predizer as propriedades físico-químicas das moléculas;
- j) Predizer moléculas que são liberadas no ambiente e,

- k) Racionalizar e prever os efeitos das moléculas que estão em misturas ou em formulações (PUZYN; LESZCZYNSKI; CRONIN, 2009).

Desde 1940, com os trabalhos de Hansch e Fujita, o QSAR tem se consolidado como uma ferramenta de grande importância na descoberta de novos fármacos (LILL, 2007). Da idéia original do QSAR, em que se baseia numa propriedade físico-química, por exemplo, o coeficiente de partição (logP), influenciar a atividade biológica da molécula, hoje diversas aproximações são propostas levando em consideração a estrutura tridimensional, os vários conformeros, a simulação do encaixe induzido e solvatação, aproximando o modelo hipotético ao receptor biológico (LILL, 2007).

Na Tabela 2, encontram-se várias aproximações utilizadas no estudo de QSAR.

Tabela 2 Dimensões usadas em QSAR.

Dimensão do QSAR	Relação
1D	Afinidade correlacionada a $pK_a$ , logP etc.
2D	Afinidade correlacionada a uma estrutura referência
3D	Afinidade correlacionada à estrutura tridimensional do ligante
4D	Afinidade correlacionada a múltiplas representações conformacionais do ligante
5D	Afinidade correlacionada a múltiplas representações de cenários de encaixe induzido
6D	Afinidade correlacionada a múltiplas representações de modelos de solvatação

Fonte: Adaptada de Gad (2008)

## 2.7 Estudos de QSAR

Para desenvolver um modelo significativo no estudo de QSAR os dados biológicos devem ser exatos e precisos (SELASSIE, 2003, p.25). A atividade biológica de um inibidor é determinada pela constante de equilíbrio da afinidade ente a enzima e o inibidor ( $K_i$ ) ou pela concentração de inibição máxima em 50% ( $IC_{50}$ ), quanto menor for a  $K_i$  ou  $IC_{50}$  maior é a afinidade da enzima ao inibidor. Esta interação é estabilizada, frequentemente, por ligações não covalentes e, às vezes, aumentada por poucas ligações covalentes (NELSON; COX, 2005). A formação espontânea da ligação entre átomos resulta na diminuição da energia livre. Quando o inibidor liga-se a enzima em temperatura 25 °C e pressão a 1atm (condição padrão), o sistema entra em desordem e a entropia aumenta com possível desprendimento de calor. Os sistemas têm propensões ao menor nível de energia, indicando que ocorrerão de forma espontânea, são favoráveis termodinamicamente. Quando as variações de energia livre de Gibbs são positivas indica que não ocorrerão de forma espontânea. No sistema em equilíbrio entre o substrato (inibidor) e produto,  $\Delta G=0$ . A relação de energia livre  $\Delta G$  e  $\Delta G^\circ$ , está descrita nas Equações 1 e 2 (WENG; VAJDA; DELIS, 1996).

$$\Delta G = \Delta G^\circ + RT \ln K \quad (1)$$

$$\Delta G^\circ = -RT \ln K \quad (2)$$

Em que:

$\Delta G^\circ$  = Energia livre de Gibbs padrão;

R = Constante universal dos gases perfeitos –  $1,987 \cdot 10^{-3} \text{ kcal} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$

T = Temperatura – 298K (25 °C)

K = Constante de equilíbrio (WENG; VAJDA; DELIS, 1996).

Assim, pequenas mudanças na variação de energia livre  $\Delta G^\circ$  podem ter profundo efeito nas constantes de equilíbrio (SELASSIE, 2003).

### **2.8 Software BiografX 1.1**

O programa BiografX é uma plataforma baseada em Java para hospedagem multidimensional de QSAR e interface entre as tecnologias Symposar, Raptor e Quasar (DOBLER; VEDANI; LILL, 2009).

### **2.9 Software Symposar**

Oliveira, Ramalho e Cunha (2009) ressalta que o alinhamento é um dos passos mais importante no método QSAR 3D. *Software Symposar*, fornece um conjunto de dados com alta similaridade de propriedades físico-químicas no espaço 3D entre os ligantes, e permite configurações alternativas destes (4D) (OLIVEIRA; RAMALHO; CUNHA, 2009). A molécula em primeiro passo é alinhada no espaço 2D de uma ou mais moléculas referência (normalmente a com maior atividade). Cada modelo irá representar diferentes estados de protonação, estereoisômeros, conformações e orientações, retratando o modo de ligação dos compostos modelos (Figura 8), posteriormente, baseado no alinhamento da molécula referência no espaço 3D, a subestrutura é minimizada (LILL; VEDANI, 2006). A fim de identificar combinações ótimas entre as características farmacofóricas tridimensionais e as moléculas referências, é feita uma minimização pelo método de Monte Carlo com objetivo de refinar as soluções obtidas. A melhor solução em 3D será utilizada como modelo para o posterior alinhamento de outras moléculas, as quais serão alinhadas de acordo com as suas afinidades experimentais à proteína alvo. A seleção da conformação

e o alinhamento irão representar a conformação bioativa (LILL; VEDANI, 2006).

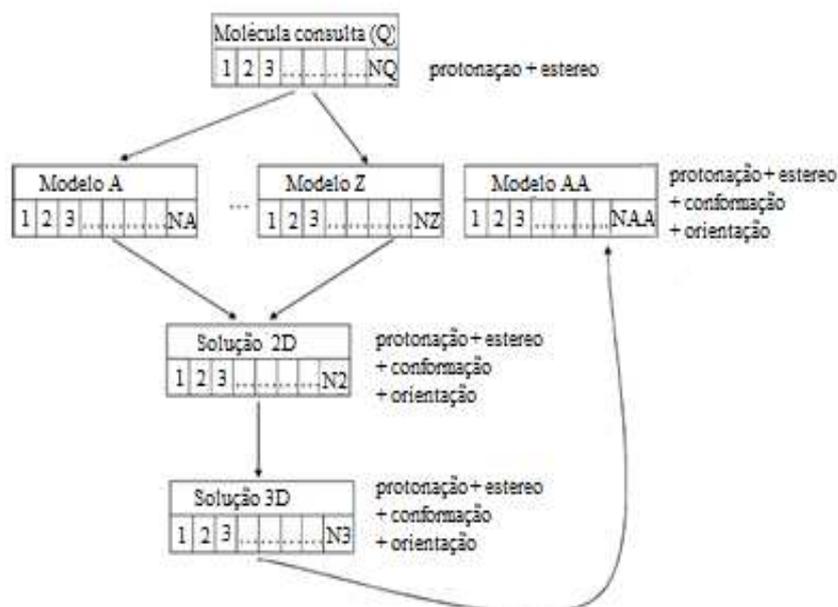


Figura 8 Esquema das duas etapas no processo de alinhamento  
Fonte: Adaptada de Lill e Vedani (2006)

### 2.10 Software Raptor

A tecnologia *Raptor* visa obter um modelo interpretável de um sítio de ligação e prever a energia livre do ligante, que é de extrema importância para a simulação do encaixe induzido (LILL; VEDANI; DOBLER, 2004). O encaixe induzido não é limitado ao efeito estérico, mas inclui a variação do campo físico-químico juntamente com ele. É representado por dupla camada tridimensional do modelo do sítio de ligação, mapeado com propriedades físico-químicas (caráter hidrofóbico e inclinações para ligações de hidrogênio) (LILL; VEDANI; DOBLER, 2004). A camada interior baseia-se no ligante mais potente

do grupo de treinamento, enquanto que a camada exterior representa todas as moléculas do grupo de treinamento. A adaptação entre ligante e receptor hipotético depende de suas interações (VEDANI; DOBLER, 2002).

A variação da energia livre ( $\Delta G$ ) de interação entre receptor-ligante inclui os termos representados na Equação 3 (LILL; VEDANI; DOBLER, 2004).

$$\Delta G = \Delta G_{\text{constante}} + f_{\text{LH}} \Delta G_{\text{LH}} + f_{\text{Hb}} \Delta G_{\text{Hb}} + f_{\text{EI}} \Delta G_{\text{EI}} + f_{\text{TAS}} T\Delta S \quad (3)$$

A  $\Delta G$  de cada ligante é determinada por uma função pontual, que considera hidrofobicidade ( $\Delta G_{\text{Hb}}$ ), ligações de hidrogênio ( $\Delta G_{\text{LH}}$ ), adaptações topológicas ( $\Delta G_{\text{EI}}$ ) e alterações entrópicas ( $T\Delta S$ ).  $\Delta G_{\text{constante}}$  representa uma perda ou ganho global de entropia translacional e rotacional da molécula ligante devido à dessolvatação do sítio ativo. Os coeficientes  $f_{\text{LH}}$ ,  $f_{\text{Hb}}$ ,  $f_{\text{EI}}$  e  $f_{\text{TAS}}$  são gerados para cada molécula específica e aperfeiçoados durante a simulação (LILL et al., 2005). A topologia de uma molécula ligante no *RAPTOR* é representada pela sua superfície acessível ao solvente. As ligações de H e hidrofobicidade produzidos por átomos do ligante são projetados em sua superfície. As adaptações destas propriedades na topologia do receptor hipotético variam, dependendo do ajuste estérico devido à atração ou repulsão entre o ligante e o receptor hipotético (OLIVEIRA; RAMALHO; CUNHA, 2009).

A fim de se obter uma melhor concordância entre a afinidade da ligação calculada e a experimental para os ligantes do grupo de treinamento, o algoritmo aperfeiçoa a natureza e distribuição das propriedades mapeadas dentro das duas camadas, tal processo de otimização emprega a Equação 4:

$$Q = \begin{cases} 0 & |\Delta G_{calc} - \Delta G_{exp}| < \alpha/2 \\ \sum_{Ligands} \left( 2\sigma \frac{|\Delta G_{calc} - \Delta G_{exp}| - \alpha/2}{\alpha} \right)^4 & |\Delta G_{calc} - \Delta G_{exp}| > \alpha/2 \end{cases} \quad (4)$$

Em RAPTOR tem-se a opção para acoplar  $\alpha$  (estimativa da incerteza experimental) a energia livre de ligação, visto que as afinidades das ligações possuem incertezas experimentais. As constantes de ligação dos compostos com alta afinidade são medidas com uma maior exatidão do que as moléculas com baixa afinidade. Quanto maior for o fator escolhido de acoplamento, mais forte é o acoplamento a energia de ligação. O parâmetro  $\sigma$  é a nitidez da função de penalização e determina quão abruptos são os aumentos da função de qualidade com o desvio previsto a partir da afinidade de ligação experimental. A quarta potência refere-se a pequenos desvios que normalmente se aproximam de incertezas experimentais (LILL et al., 2005; DOBLER; VEDANI; LILL, 2009).

### 2.11 Software Quasar

A tecnologia *Quasar* gera um receptor hipotético semi-atomístico, o qual é aperfeiçoado por meio de um algoritmo genético, levando em consideração diversos modelos de solvatação. A construção do modelo no *Quasar* inclui os seguintes passos (DOBLER; VEDANI; LILL, 2009):

- a) Geração da superfície do receptor: o encaixe induzido pode ser simulado por adaptação de uma superfície gerada acerca de todos ligantes do grupo de treinamento, o qual é alcançado para cada

molécula ligante. Este processo imita o local do encaixe induzido (VEDANI; DOBLER; LILL, 2005);

- b) Geração de uma família inicial de estruturas principais: propriedades semi-atomísticas são mapeadas para pontos distribuídos igualmente nas superfícies individuais. Sítios potenciais de ligações de H são restritos a posições na superfície do receptor, os quais estão localizados dentro de uma razoável distância e em uma orientação favorável com respeito a qualquer ligação de H (doador ou aceptor) da porção da molécula ligante do grupo de treinamento (VEDANI; DOBLER; LILL, 2005);
- c) Evolução de uma família modelo: a família inicial de modelos de receptor progride em eventos de *crossover* e mutação usando um algoritmo genético. Durante um *crossover*, existe uma pequena probabilidade de um erro na transcrição, em que é expresso por uma mutação aleatória. Depois disto, dois indivíduos da população com o maior *lack-of-fit* (LOF) são descartados, este processo é repetido até o objetivo do valor  $q^2$ , o máximo de números de *crossover* ou limite de erro experimental é alcançado (DOBLER; VEDANI; LILL, 2009);
- d) Estimativa da relação de energia livre de ligação do ligante: a energia de solvatação e a diminuição da entropia sobre a ligação ao receptor é determinada pela Equação 5:

$$E_{\text{bdg}} = E_{\text{lig-ree}} - E_{\text{solv,lig}} - T\Delta S_{\text{bdg}} + \Delta E_{\text{int,lig}} + E_{\text{env.adapt,lig}} \quad (5)$$

Em que:

$E_{\text{lig-ree}}$  é o termo que corresponde à contribuição entálpica da interação ligante receptor;

$E_{\text{solv,lig}}$  é o termo que corresponde à energia requerida para a dessolvatação, a partir de um ambiente aquoso em uma cavidade hidrofóbica do receptor;

$T\Delta S_{\text{bdg}}$  é estimado pela atribuição do montante de 0,7 kcal/mol para cada rotação de ligação, excluindo o terminal  $-\text{CH}_3$ ;

$\Delta E_{\text{int,lig}}$ : é o termo que corresponde à melhora potencial da energia interna do ligante enquanto ligado ao receptor hipotético;

$E_{\text{env.adapt,lig}}$  é o termo que está associado com o consumo de energia sobre como modificar a superfície média do receptor para a superfície individual (VEDANI; DOBLER; LILL, 2005).

A energia livre dos ligantes,  $\Delta G_{\text{pred}}$ , são preditos pela média de uma regressão linear entre  $\Delta G_{\text{exp}}$  e  $E_{\text{bdg}}$  usando as moléculas ligantes do grupo de treinamento (DOBLER; VEDANI; LILL, 2009), conforme a Equação 6:

$$\Delta G_{\text{pred}} = |a| \cdot E_{\text{bdg}} + b \quad (6)$$

- e) Análise da família modelo: um critério obrigatório para validar uma família de modelos de receptores é a sua capacidade de prever a relação de energia livre dos ligantes com um grupo teste externo de moléculas, os quais não usados durante a construção do modelo. Outros critérios de validação incluem: validação cruzada  $q^2$ , lack-of-fit do grupo de treinamento, a variação de  $\Delta G_{\text{pred}}$ . No entanto, o teste final é a verificação experimental e predição da atividade de novos compostos (VEDANI; DOBLER; LILL, 2005).

## 2.12 Validação

O termo validação segundo Organisation for Economic Co-operation and Development – OECD (2007) é: “o processo pelo qual a confiança e relevância de uma aproximação, método, processo ou avaliação é estabelecido para um propósito definido” (ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT - OECD, 2007). Os modelos devem ser cientificamente válidos ou validados se são usados em avaliação química (PUZYN; LESZCZYNSKI; CRONIN, 2009).

Segue abaixo a descrição de alguns índices estatísticos usados na validação de um QSAR:

- a) Coeficiente de determinação linear quadrático ( $R^2$ ): mede o grau de ajuste da equação de regressão múltipla aos dados amostrais. Representa a proporção da divergência de uma variável que pode ser predita da outra variável (SOUZA, 2009);
- b) Validação cruzada: segundo OECD (2007) é um método que utiliza uma ou mais técnicas estatísticas, em que compostos do grupo de treinamento são omitidos e, em seguida, o QSAR é desenvolvido com base nos dados dos compostos restantes, os quais serão utilizados para fazer previsões dos compostos que foram omitidos. Este processo é repetido várias vezes de forma que um valor estatístico seja obtido a partir da comparação dos dados preditos com os dados conhecidos (OECD, 2007). Uma técnica muito utilizada é a validação cruzada por *Leave-One Out* (LOO) ( $q^2$ ), processo que é empregado na formação de **n** grupo de treinamento, em que apenas um composto é excluído do grupo de treinamento original. São gerados **n** modelos para cada grupo de treinamento **n-1**

e, o valor do composto excluído de cada modelo é previsto (OECD, 2007).

- c) Validação externa: é um critério obrigatório em um estudo de QSAR (DOBLER; VEDANI; LILL, 2009), refere-se à validação em que os compostos químicos do grupo de teste são diferentes daqueles incluídos no grupo de treinamento, porém representam a mesma predominância dos compostos químicos. O modelo gerado, utilizando o grupo de treinamento, é aplicado ao grupo de teste, com o objetivo de verificar a capacidade preditiva do modelo. OECD (2007) ressalta que é a forma mais rigorosa de validação, desde que os dados experimentais sejam disponíveis e as estruturas químicas do grupo de teste sejam selecionadas com critério, a fim de permitir uma cobertura suficiente à aplicabilidade do modelo (OECD, 2007).
- d) *Scramble test* (Teste de randomização): é um critério de validação em que testa a sensibilidade do modelo obtido aos dados biológicos. No processo, as atividades experimentais dos compostos do grupo de treinamento são distribuídas aleatoriamente ao acaso e, posteriormente, a simulação é repetida de forma idêntica ao processo de geração do modelo original. Se  $r^2$  (o valor preditivo) obtido for semelhante ao modelo sem o *scramble test*, o modelo gerado não é sensível aos dados biológicos (DOBLER; VEDANI; LILL, 2009).

Conforme Vedani et.al (2002, 2005) , os valores mínimos de  $q^2$  e  $r^2_{\text{preditivo}}$  aceitáveis em trabalhos de QSAR são respectivamente, 0,5 e 0,7 (VEDANI; DOBLER, 2002; VEDANI; DOBLER; LILL, 2005).

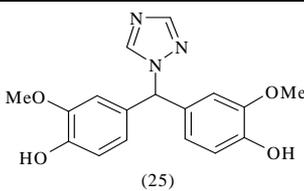
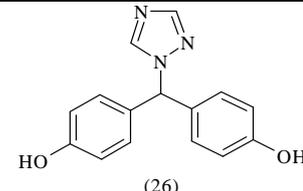
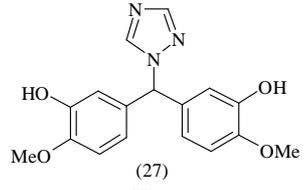
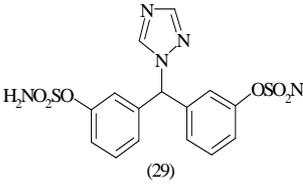
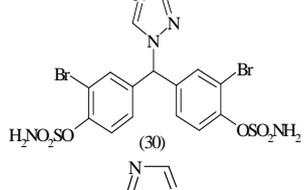
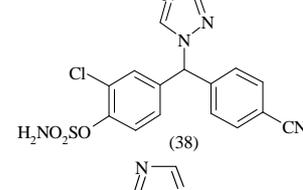
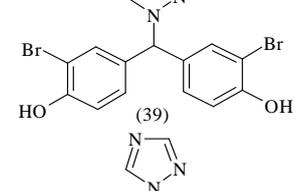
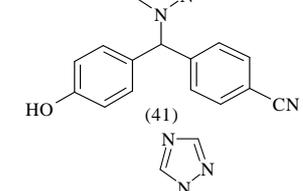
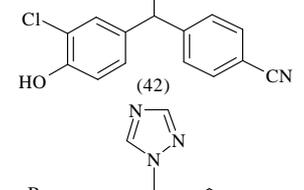
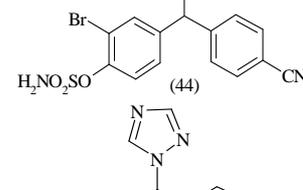
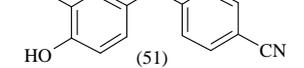
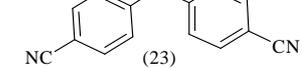
### 3 MATERIAL E MÉTODOS

Desenvolveu-se este estudo no Laboratório de Química Computacional, do departamento de Química da Universidade Federal de Lavras- UFLA. Os computadores utilizados neste trabalho foram um Mac OX e um PC. Utilizou-se dos *softwares* PC *Spartan Pro*, *MacBio* versão 3.1, *Biograf<sup>x</sup>* versão 1.1, *Symposar* versão 2.0, *RAPTOR* versão 2.0 e *Quasar* versão 5.2.

#### 3.1 Inibidores da Aromatase

O conjunto de dados das estruturas químicas dos quarenta e três inibidores não esteroidais da enzima aromatase utilizados nos estudos de QSAR, foram obtidas a partir dos resultados publicados dos trabalhos de Wood (2008), Woo (2003, 2007) e Jackson; et al. (2007) submetidas ao mesmo ensaio de inibição *in vitro* em células JEG-3(WOO et al., 2003; JACKSON et al., 2007; WOO et al., 2007; WOOD et al., 2008). As Tabela 3,4 e 5 apresenta as estruturas químicas dos inibidores não esteroidais análogos ao Letrozol (4-[4-cianofenil)-(1,2,4-triazol-1-ila)metil]benzotriila), YM511(4-[(4-bromobenzil)-(1,2,4-triazol-4-ilamino)]benzotriila) e Anastrozol (2-[3-(ciano-dimetil-metil)-5[1,2,4]triazol-1-ilmetil-fenil]2-metil-propionitrila) respectivamente.

Tabela 3 Análogos ao Letrozol (composto33)

Compostos	Estruturas	Compostos	Estruturas
E01 <sup>1</sup>	 (25)	E02 <sup>1</sup>	 (26)
E03 <sup>1</sup>	 (27)	E05 <sup>1</sup>	 (29)
E06 <sup>1</sup>	 (30)	E14 <sup>1</sup>	 (38)
E15 <sup>1</sup>	 (39)	E17 <sup>1</sup>	 (41)
E18 <sup>1</sup>	 (42)	E20 <sup>1</sup>	 (44)
E27 <sup>1</sup>	 (51)	E33 <sup>1</sup>	 (23)

<sup>1</sup>Dados obtidos a partir dos estudos de Wood et al. (2008); -OMe= grupo funcional metóxi; H<sub>2</sub>NO<sub>2</sub>SO- = grupo funcional sulfamato

Tabela 4 Análogos ao YM511 (composto 40)

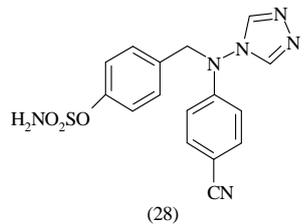
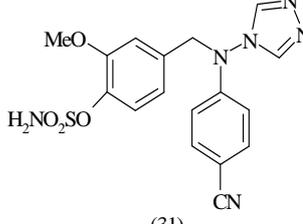
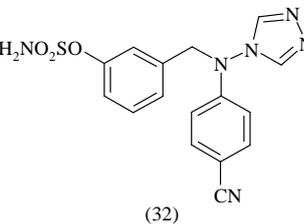
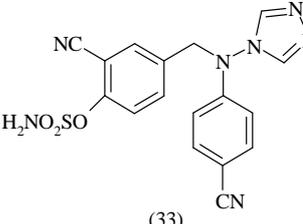
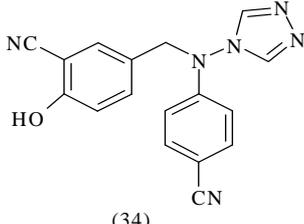
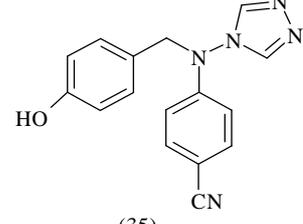
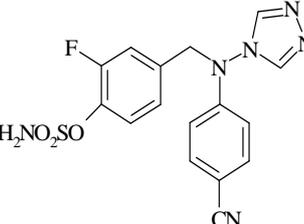
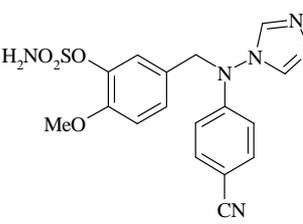
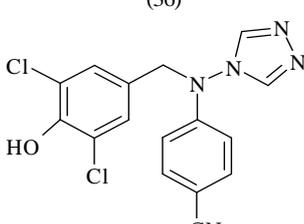
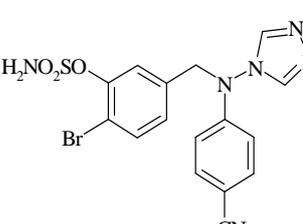
Compostos	Estruturas	Compostos	Estruturas
E04 <sup>2</sup>	 <p>(28)</p>	E07 <sup>2</sup>	 <p>(31)</p>
E08 <sup>2</sup>	 <p>(32)</p>	E09 <sup>2</sup>	 <p>(33)</p>
E10 <sup>2</sup>	 <p>(34)</p>	E11 <sup>2</sup>	 <p>(35)</p>
E12 <sup>2</sup>	 <p>(36)</p>	E13 <sup>2</sup>	 <p>(37)</p>
E16 <sup>2</sup>	 <p>(40)</p>	E19 <sup>2</sup>	 <p>(43)</p>

Tabela 4. continua

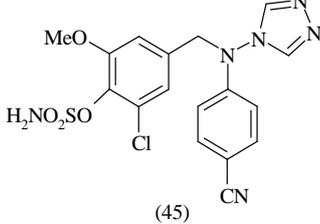
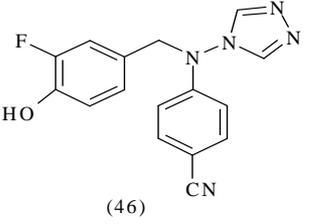
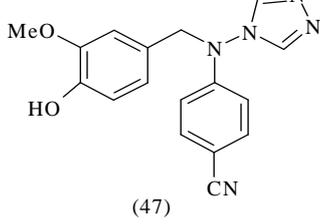
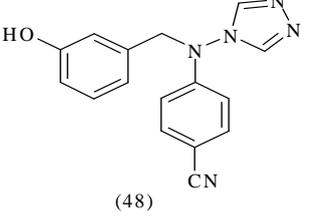
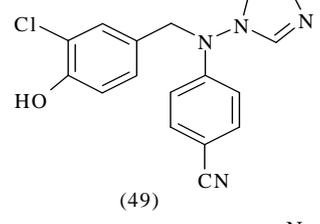
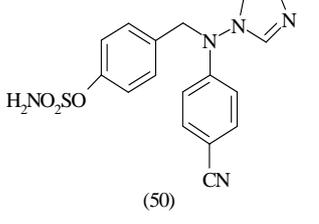
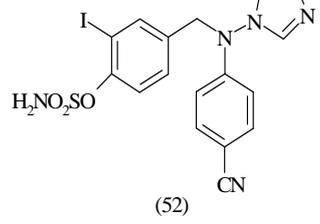
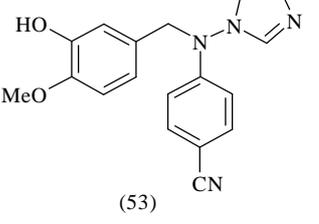
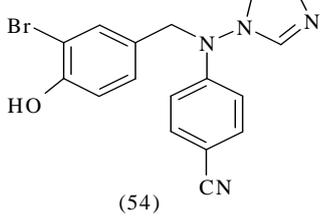
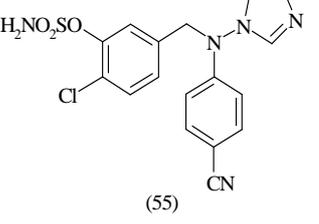
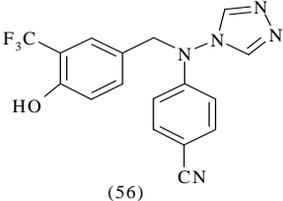
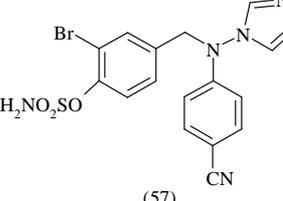
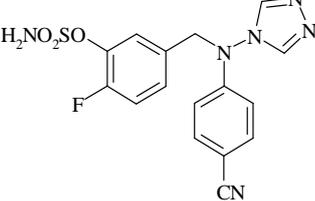
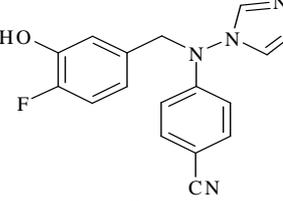
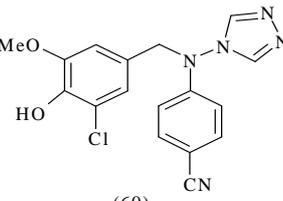
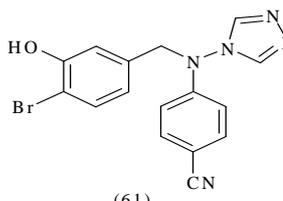
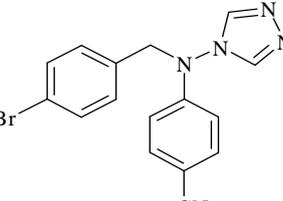
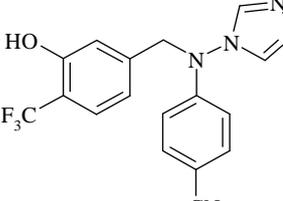
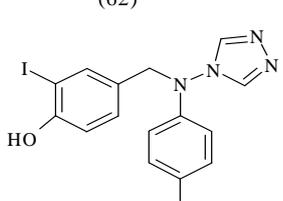
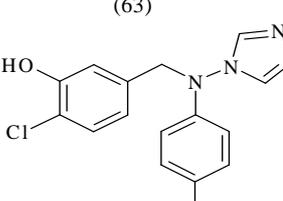
Compostos	Estruturas	Compostos	Estruturas
E21 <sup>2</sup>	 <p>(45)</p>	E22 <sup>2</sup>	 <p>(46)</p>
E23 <sup>2</sup>	 <p>(47)</p>	E24 <sup>2</sup>	 <p>(48)</p>
E25 <sup>2</sup>	 <p>(49)</p>	E26 <sup>2</sup>	 <p>(50)</p>
E28 <sup>2</sup>	 <p>(52)</p>	E30 <sup>2</sup>	 <p>(53)</p>
E31 <sup>2</sup>	 <p>(54)</p>	E32 <sup>2</sup>	 <p>(55)</p>

Tabela 4, continua

Compostos	Estruturas	Compostos	Estruturas
E34 <sup>2</sup>	 (56)	E35 <sup>2</sup>	 (57)
E36 <sup>2</sup>	 (58)	E37 <sup>2</sup>	 (59)
E38 <sup>2</sup>	 (60)	E39 <sup>2</sup>	 (61)
E40 <sup>3</sup>	 (62)	E41 <sup>2</sup>	 (63)
E42 <sup>2</sup>	 (64)	E43 <sup>2</sup>	 (65)

<sup>2</sup> Dados obtidos a partir dos estudos de Woo et.al (2007)<sup>3</sup> Dados obtidos a partir dos estudos de Woo et.al (2003)-OMe= grupo funcional metóxi; H<sub>2</sub>NO<sub>2</sub>SO- = grupo funcional sulfamato

Composto	Estrutura
E29 <sup>4</sup>	<p style="text-align: center;">(53)</p>

<sup>4</sup>Dados obtidos a partir dos estudos de Jackson et al. (2007)

Os 43 inibidores foram divididos em dois grupos: treinamento e teste, e subdivididos em sete subgrupos em ordem crescente de atividade biológica (Tabela, 6), no qual o  $IC_{50}$  foi convertido em unidade molar e, então, expressos em  $pIC_{50}$  ( $-\log IC_{50}$ ) e em energia livre  $\Delta G$  ( $-RT \ln IC_{50}$ ) (OLIVEIRA, 2009). A composição do grupo de teste fez-se pela separação de compostos relacionados estruturalmente, retirando um composto dos subgrupos I, II, III e IV e dois do subgrupo V e VI, formando um grupo teste com oito compostos (**E02, E07, E11, E18, E24, E30, E35 e E40**), correspondendo aproximadamente 20% dos compostos totais. O grupo de treinamento é formado pelos trinta e cinco compostos **E01, E03- E06, E08-E10, E12-E17, E19-E23, E25-E29, E31-E34, E36-E39, E41-E43**. As disposições dos compostos de treinamento e teste de acordo com as suas atividades biológicas, podem ser observadas no Gráfico 2, mostrando uma homogeneidade nos dados.

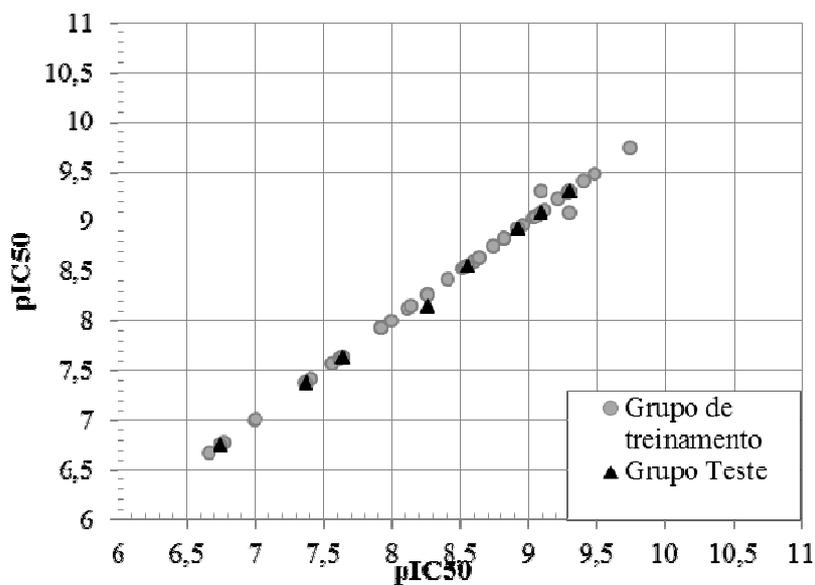
Tabela 6 Distribuição dos compostos de acordo com sua atividade biológica em ordem crescente de pIC<sub>50</sub>.

Subgrupos	Compostos	IC <sub>50</sub> (nM)	pIC <sub>50</sub>	ΔG (-RTlnIC <sub>50</sub> )
I (6,5-7,0)	E01	215±16	6,668	-8,938
	E02*	178	6,750	-9,048
	E03	169±50	6,772	-9,078
	E04	100 ±7.8	7,000	-9,383
II (7,0-7,5)	E05	99±0,01	7,004	-9,389
	E06	43±3,2	7,367	-9,875
	E07*	42±1	7,376	-9,888
	E08	39±4	7,408	-9,931
III (7,5-8,0)	E09	27±10	7,569	-10,146
	E10	24±5.5	7,620	-10,214
	E11*	23±1	7,638	-10,239
	E12	12±1.8	7,921	-10,618
	E13	12±1.9	7,921	-10,618
	E14	12±3,4	7,921	-10,618
	E15	10±2,7	8,000	-10,724
IV (8,0-8,5)	E16	7.6±2	8,119	-10,884
	E17	7,2±1,1	8,142	-10,915
	E18*	5,5±3,7	8,260	-11,072
	E19	3,9±0,9	8,409	-11,272
V (8,5-9,0)	E20	3,0±0,1	8,523	-11,425
	E21	2,9±0,2	8,537	-11,444
	E22	2,9±0,4	8,537	-11,444
	E23	2,8±0,5	8,553	-11,465
	E24*	2,8±0,8	8,553	-11,465
	E25	2,5±0,3	8,602	-11,531
	E26	2,3±0,3	8,638	-11,579
	E27	1,8±0,3	8,744	-11,722
	E28	1,5±0,1	8,824	-11,828
	E29	1,5±0,5	8,824	-11,828
	E30*	1,2±0,8	8,920	-11,958
VI (9,0-9,5)	E31	1,1±0,1	8,959	-12,009
	E32	0.92±0.03	9,036	-12,113
	E33	0,89±0,13	9,050	-12,132
	E34	0.88±0.1	9,056	-12,139
	E35*	0.82±0.3	9,086	-12,180
	E36	0.77±0.03	9,113	-12,216
	E37	0.6±0.1	9,222	-12,362

Tabela 6, continua

Subgrupos	Compostos	IC <sub>50</sub> (nM)	pIC <sub>50</sub>	ΔG (-RTlnIC <sub>50</sub> )
VII (9,5-10,0)	E38	0.51±0.06	9,292	-12,456
	E39	0.5±0.1	9,301	-12,468
	E40*	0.5±0.03	9,301	-12,468
	E41	0.4±0.02	9,398	-12,598
	E42	0.33±0.05	9,481	-12,710
	E43	0.18±0.04	9,745	-13,063

\*Compostos do grupo Teste

Gráfico 2 Disposição dos compostos dos grupos de treinamento e teste de acordo com pIC<sub>50</sub>

### 3.2 Construções dos modelos tridimensionais e cálculo das cargas parciais atômicas

Através do *software* PC *Spartan Pro* (WAVEFUNCTION, 1999) foram obtidas as estruturas tridimensionais dos compostos. Através do *software* *MacBio* 3.1 foram realizados os cálculos de cargas parciais atômicas, bem como o aperfeiçoamento da geometria das moléculas no vácuo.

### 3.3 Validação dos modelos

Os modelos gerados de QSAR foram avaliados usando quatro critérios:

- a) Coeficiente de correlação da validação cruzada “leave-one-out” (LOO-cv) ou  $q^2$ ;
- b) Desvio da raiz quadrada média ou  $R^2$ ;
- c) Predições reais utilizando compostos do grupo teste (validação externa ou  $r^2$  preditivo) e,
- d) *Scramble test* – sensibilidade do modelo aos dados experimentais.

## 4 RESULTADOS E DISCUSSÕES

### 4.1 Resultados do *Symposar*

Do *software Symposar*, obteve-se três conjuntos de dados com números distintos de confôrmeros (3D/4D) dos 43 inibidores não esteroidais. Os modelos A, B e C representam as diferentes soluções encontradas pelo programa na formação do conjunto de dados com similaridade das propriedades físico-químicas, conforme descrito no tópico 3.3. O Modelo A fornece o menor número de confôrmeros, enquanto o Modelo B fornece um número médio e o Modelo C o número grande de confôrmeros. Observa-se que vários compostos admitem um maior número de conformações (Tabela 7), podendo ser explicado pelo efeito estérico dos átomos das moléculas impedindo as rotações livres das ligações. As conformações bioativas adotadas para o modelo A estão apresentadas na Figura 9.

Tabela 7 Números de confôrmeros para cada modelo

Composto	Modelo A	Modelo B	Modelo C	Composto	Modelo A	Modelo B	Modelo C
E01	07	10	10	E02	03	05	05
E03	07	10	10	E04	01	01	03
E05	02	04	04	E06	03	04	10
E07	01	02	02	E08	01	02	03
E09	01	02	04	E10	02	03	06
E11	01	02	06	E12	01	01	03
E13	01	02	02	E14	02	03	07
E15	07	10	10	E16	01	02	03
E17	04	07	09	E18	06	08	09
E19	01	02	04	E20	06	10	10
E21	01	02	03	E22	02	02	04

Tabela 6, continua

Composto	Modelo	Modelo	Modelo	Composto	Modelo	Modelo	Modelo
	A	B	C		A	B	C
E23	01	02	04	E24	01	01	03
E25	03	03	07	E26	01	02	05
E27	10	10	10	E28	01	02	03
E29	10	10	10	E30	01	01	03
E31	02	04	05	E32	01	02	03
E33	03	06	09	E34	02	02	04
E35	02	03	08	E36	01	02	04
E37	01	02	02	E38	03	04	07
E39	02	02	03	E40	01	03	03
E41	01	01	04	E42	02	07	09
E43	01	01	01				

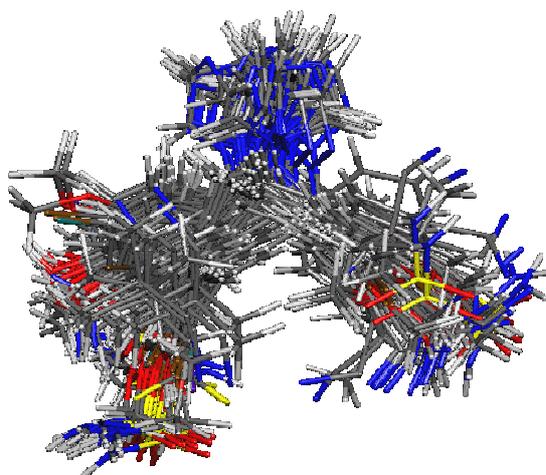


Figura 9 Alinhamento e conformações do modelo A utilizando o *software Symposar*

#### 4.2 Resultados do *RAPTOR* (5D)

Foram gerados três modelos de superfície do receptor para a simulação do encaixe induzido (5D). Na Tabela 8 estão apresentados os resultados obtidos.

Tabela 8 Resultados estatísticos dos modelos de QSAR obtidos com compostos não esteroidais.

Modelo	R <sup>2</sup>	r <sup>2</sup> preditivo	q <sup>2</sup>	Scramble Test	Equação
A	0,750	0,791	0,721	-1,321	$\Delta G = - 1,991 + 1,248\Delta G_{LH} + 1,249\Delta G_{Hb} + 1,000\Delta G_{EI} - 0,306T\Delta S$
B	0,754	0,671	0,741	-1,099	$\Delta G = -1,991 + 1,210\Delta G_{LH} + 1,249\Delta G_{Hb} + 1,026\Delta G_{EI} - 0,351T\Delta S$
C	0,719	0,722	0,749	-1,121	$\Delta G = - 1, 991 + 1,244\Delta G_{LH} + 1,249\Delta G_{Hb} + 1,001\Delta G_{EI} - 0, 301T\Delta S$

Os modelos obtidos apresentaram LOO-cv > 0,5 e os Modelos A e B apresentaram R<sup>2</sup> > 0,7. Todos os modelos apresentaram valores de r<sup>2</sup> preditivo > 0,5. Utilizando o método de validação *scramble test* foi verificado que todos os modelos foram sensíveis aos dados biológicos, ou seja, os compostos não admitem valores diferentes a suas atividades biológicas. O modelo A apresentou os melhores resultados estatísticos, assim ele foi selecionado como o melhor modelo.

O coeficiente de determinação, R<sup>2</sup>, igual a 0,750 do Modelo A, permite inferir que existe uma correlação aceitável entre os resultados analisados e os biológicos. Na determinação da relação entre a resposta biológica e estrutura química, utilizou-se do método estatístico de validação cruzada LOO-cv. O modelo A apresentou um valor de q<sup>2</sup> 0,721 para o grupo de treinamento, ou seja, apresenta capacidade preditiva de 72,1%, demonstrando sua utilidade à predição da atividade de novos compostos. A validação externa do Modelo A foi realizada por meio do grupo de teste, em que se obteve um r<sup>2</sup> preditivo de 0,791, ou seja, o modelo gerado a partir do grupo de treinamento possui características em comum com os compostos do grupo teste, sendo capaz de encontrá-las em uma nova molécula. Segue abaixo a equação de QSAR do modelo A obtida a partir dos 43 compostos:

$$\Delta G = -1,991 + 1,248\Delta G_{LH} + 1,249\Delta G_{Hb} + 1,000\Delta G_{EI} - 0,306T\Delta S \quad (1)$$

Para identificar características potenciais que permite um aumento da afinidade de ligação, é analisada a contribuição à energia livre ao sítio de ligação. Na Tabela 9 estão apresentados os valores de atividade biológica  $pIC_{50}$  experimental e calculado, variação na energia livre experimental e calculada, variações nas ligações de H e hidrofobicidade, adaptações topológicas e variações entrópicas dos compostos dos grupos de treinamento e teste .

Tabela 9 Valores de atividade e energia livre ( $\text{kcal mol}^{-1}$ ) de compostos do grupo de treinamento e de teste referentes ao melhor modelo obtido a partir de compostos não esteroidais.

Compostos	$pIC_{50}$ (exp)	$pIC_{50}$ (calc)	Fator off*	$\Delta G$ (exp)	$\Delta G$ (calc)	$\Delta G$ (EI)	$T\Delta S$	$\Delta G$ (Hb)	$\Delta G$ (LH)
Grupo de Treinamento									
E01	6,668	6,422	1,761	-8,938	-8,608	2,404	0,662	-7,790	-1,895
E03	6,772	6,520	1,786	-9,078	-8,740	2,324	0,662	-7,275	-2,463
E04	7,000	7,780	6,035	-9,383	-10,430	1,237	0,662	-8,504	-1,836
E05	7,004	7,124	1,319	-9,389	-9,550	2,679	1,104	-7,400	-3,944
E06	7,367	8,028	4,591	-9,875	-10,762	2,618	1,104	-10,215	-2,280
E08	7,408	7,949	3,474	-9,931	-10,656	1,176	0,662	-8,694	-1,811
E09	7,569	7,626	1,141	-10,146	-10,223	1,465	0,662	-8,492	-1,868
E10	7,620	7,522	1,250	-10,214	-10,084	0,805	0,442	-8,076	-1,265
E12	7,921	8,123	3,051	-10,618	-10,587	1,271	0,662	-8,757	-1,774
E13	7,921	7,813	1,281	-10,618	-10,474	1,567	0,662	-8,587	-2,127
E14	7,921	7,509	2,583	-10,618	-10,065	2,102	0,883	-8,696	-2,367
E15	8,000	8,548	3,531	-10,724	-11,458	1,992	0,662	-10,089	-2,035
E16	8,119	8,882	5,790	-10,884	-11,906	0,803	0,442	-10,120	-1,041
E17	8,142	7,873	1,860	-10,915	-10,554	1,762	0,662	-8,241	-2,748
E19	8,409	8,742	2,152	-11,272	-11,718	1,239	0,662	-9,258	-2,372
E20	8,523	8,221	2,003	-11,425	-11,021	2,263	0,883	-9,201	-2,976
E21	8,537	8,067	2,949	-11,444	-10,814	1,683	0,662	-9,171	-2,000
E22	8,537	8,433	1,270	-11,444	-11,305	0,654	0,442	-9,100	-1,311

Tabela 9, continua

Compostos	pIC <sub>50</sub> (exp)	pIC <sub>50</sub> (calc)	Fator off*	ΔG (exp)	ΔG (calc)	ΔG (EI)	TΔS	ΔG (Hb)	ΔG (LH)
E23	8,553	8,262	1,952	-11,465	-11,076	1,019	0,442	-8,625	-1,922
E25	8,602	8,379	1,814	-11,531	-11,231	0,753	0,442	-9,270	-1,166
E26	8,638	8,378	1,675	-11,579	-11,232	1,230	0,662	-9,271	-1,865
E27	8,744	8,678	1,165	-11,722	-11,633	1,728	0,662	-9,433	-2,601
E28	8,824	8,473	2,240	-11,828	-11,359	1,318	0,662	-9,446	-1,904
E29	8,824	8,569	1,797	-11,828	-11,487	2,970	0,883	-11,245	-2,106
E31	8,959	8,705	1,793	-12,009	-11,669	0,705	0,442	-9,496	-1,331
E32	9,036	8,792	1,755	-12,113	-11,785	1,158	0,662	-9,222	-2,394
E33	9,050	8,491	3,628	-12,132	-11,382	1,737	0,662	-9,041	-2,751
E34	9,056	9,345	1,947	-12,139	-12,527	1,013	0,442	-10,678	-1,314
E36	9,113	8,152	9,135	-12,216	-10,928	1,103	0,662	-8,727	-1,978
E37	9,222	8,793	2,686	-12,362	-11,787	0,562	0,442	-8,804	-1,997
E38	9,292	8,736	3,600	-12,456	-11,710	1,176	0,442	-9,569	-1,769
E39	9,301	9,269	1,076	-12,468	-12,425	0,596	0,442	-9,672	-1,801
E41	9,398	9,558	1,446	-12,598	-12,813	0,833	0,442	-10,340	-1,758
E42	9,481	9,180	2,001	-12,710	-12,306	0,786	0,442	-9,640	-1,906
E43	9,745	10,193	2,808	-13,063	-13,664	0,035	0,442	-9,946	-2,206
Grupo Teste									
E02	6,750	7,368	4,154	-9,048	-9,877	1,934	0,662	-8,304	-2,180
E07	7,376	7,394	1,042	-9,888	-9,912	1,699	0,662	-8,288	-1,996
E11	7,638	7,898	1,055	-10,239	-10,888	0,657	0,442	-8,637	-1,361
E18	8,260	7,966	1,967	-11,072	-10,678	1,693	0,662	-8,946	-2,099
E24	8,553	8,474	1,197	-11,465	-11,360	0,701	0,442	-8,547	-1,967
E30	8,920	8,524	2,490	-11,958	-11,427	0,827	0,442	-8,975	-1,731
E35	9,086	9,454	2,334	-12,180	-12,673	1,464	0,662	-9,882	-2,929
E40	9,301	8,858	2,772	-12,468	-11,874	0,598	0,442	-9,878	-1,047

pIC<sub>50</sub> = -log(IC<sub>50</sub>); Fator *off*: desvio máximo individual e rms entre o valor experimental e o valor calculado de IC<sub>50</sub>, calculam-se como IC<sub>50</sub>(exp)/IC<sub>50</sub>(calc) para IC<sub>50</sub>(exp) > IC<sub>50</sub>(calc) e IC<sub>50</sub>(calc)/IC<sub>50</sub>(exp) para IC<sub>50</sub>(calc) > IC<sub>50</sub>(exp); ΔG<sub>exp</sub>: energia livre de Gibbs experimental; ΔG<sub>calc</sub>: energia livre de Gibbs calculada; ΔG<sub>EI</sub>: energia livre de Gibbs adaptações topológicas; TΔS: alterações entrópicas; ΔG<sub>Hb</sub>: energia livre de Gibbs hidrofobicidade; ΔG<sub>LH</sub>: energia livre de Gibbs ligações de hidrogênio

O desvio máximo encontrado entre o valor experimental e o valor calculado do grupo de treinamento está atribuído ao composto **E36** (Gráfico 3)

correspondendo ao fator *off* de 9,135 unidades, ou seja, as características físico-químicas calculadas para o composto **E36** têm o maior desvio em relação a sua atividade biológica experimental. Para o grupo de teste o desvio máximo confere ao composto **E02** (Gráfico 3) correspondendo ao fator *off* de 4,154 unidades (Tabela 9).

Observa-se que o maior valor obtido para a energia de hidrofobicidade foi para o composto **E29** (Anastrozol) (Tabela 9), pode-se inferir que a estrutura com os grupos benzil e ciano potencializa a interação hidrofóbica com o receptor hipotético. O menor valor confere ao composto **E05** (Tabela 9), provavelmente, devido aos grupos sulfamatos nas posições R1/R3 com características polares. Em relação à adaptação topológica o melhor resultado confere ao composto **E43** (Tabela 9), ou seja, a energia necessária para que ocorra o encaixe induzido é menor em comparação ao composto **E29**, o qual obteve o maior valor de energia. O melhor valor obtido na energia de ligação de hidrogênio confere ao composto **E05** (Tabela 9), como mencionado acima pode ser explicado pela presença dos grupos sulfamatos nas posições R1/R3, favorecendo as ligações de hidrogênio. Segundo estudo de Roy e Roy (2010), uma ou mais interações de hidrogênio com o resíduo MET374 é um dos requerimentos essencial para que o ligante seja um ótimo inibidor (ROY;ROY, 2010). O menor valor configura ao composto **E16** (Tabela 9), por possuir nas posições R2/R4 dois átomos de cloro, o que possivelmente ocorra o efeito estérico, influenciando nas ligações de hidrogênio.

Para uma melhor visualização, no Gráfico 3 estão dispostos os valores de atividade biológica experimental e calculado, dos compostos não esteroidais em estudo.

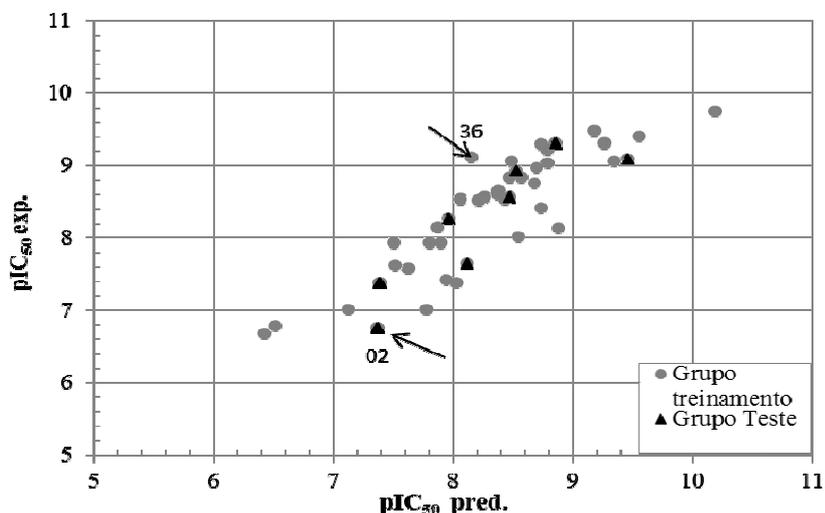


Gráfico 3 Valores de  $pIC_{50}$  experimental e calculado dos 43 inibidores não esteroidais da enzima aromatase

A Figura 10 é a representação gráfica da dupla camada do Modelo A utilizando como referência a estrutura tridimensional do composto **E43**, pois apresentou o melhor resultado de  $pIC_{50}$  experimental e calculado, levando a crer características de um potente inibidor. As camadas que representam o modelo do sítio de ligação do receptor podem ser coloridas de acordo com a mudança de hidrofobicidade na superfície. A cor marrom indica que a região é muito hidrofóbica, a cor amarela a região é hidrofóbica, e a cor cinza indica que a região é neutra (nem hidrofílico ou hidrofóbico). Regiões hidrofílicas são ambas aceptoras de interações do tipo ligação de hidrogênio (vermelho) ou doadoras de ligação de hidrogênio (azul). A região hidrofóbica dominante foi observada em torno dos anéis de cor marrom (Figura 10). Regiões doadoras de ligação de hidrogênio (cor azul) são observadas próximas ao grupo triazol, na posição R3 dos compostos análogos ao YM511 e na posição R2 de compostos análogos ao Letrozol. Portanto, pode-se inferir que substituintes aceptores de ligação de

hidrogênio próximo a camada azul (Figura 10) contribuem positivamente na potência do composto.

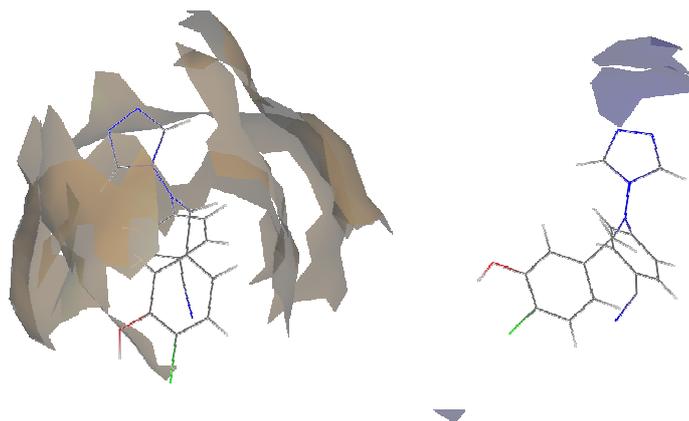


Figura 10 Representação de dupla camada do modelo A obtido por estudos de 5D-QSAR, utilizando o composto **E43** como referência. À esquerda, regiões hidrofóbicas e à direita, regiões hidrofílicas doadoras de ligação de hidrogênio

#### 4.3 Resultados do *QUASAR* (6D)

No *Quasar*, foi formada uma família composta por 200 modelos que evoluíram ao longo dos 5000 *crossovers*, correspondendo a 200 gerações. A superfície hipotética do receptor foi gerada a partir do Modelo A. Obteve-se um valor de coeficiente de determinação ( $R^2$ ) igual a 0,772, a qual é uma relação aceitável com os dados experimentais. Para verificar a relação entre a resposta biológica e estrutura química utilizou-se o método de validação cruzada LOO-cv (*leave-one-out*) com valor de 0,736, ou seja, uma capacidade preditiva de 73,6%, demonstrando sua utilidade à predição da atividade de novos compostos. A validação externa do Modelo A realizada por meio do grupo de teste, apresentou um  $r^2$  preditivo de 0,791, ou seja, o modelo gerado a partir do grupo

de treinamento possui características em comum com os compostos do grupo teste, sendo capaz de encontrá-las em uma nova molécula.

Para identificar características potenciais que permite um aumento da afinidade de ligação, é analisada a contribuição à energia livre ao sítio de ligação. Na Tabela 10 estão apresentados os valores de atividade biológica  $pIC_{50}$  experimental e calculado, variação na energia livre experimental e calculada.

Tabela 10 Valores de atividade e energia livre ( $kcal\ mol^{-1}$ ) de compostos do grupo de treinamento e teste

Compostos	$pIC_{50}$ exp.	$pIC_{50}$ calc.	Fator off*	$\Delta G$ (exp)	$\Delta G$ (calc)
Grupo Treinamento					
E01	6,668	6,621	0,1	-8,938	-8,875
E03	6,772	6,705	0,2	-9,078	-8,988
E04	7,000	7,938	7,7	-9,383	-10,641
E05	7,004	7,029	0,1	-9,389	-9,423
E06	7,367	7,394	0,1	-9,875	-9,912
E08	7,408	8,461	10,3	-9,931	-11,342
E09	7,569	7,685	0,3	-10,146	-10,302
E10	7,620	8,197	2,8	-10,214	-10,988
E12	7,638	8,090	0,5	-10,618	-10,845
E13	7,921	7,867	0,1	-10,618	-10,546
E14	7,921	8,032	0,3	-10,618	-10,766
E15	8,000	7,969	0,1	-10,724	-10,682
E16	8,119	8,854	4,4	-10,884	-11,868
E17	8,142	7,711	1,7	-10,915	-10,336
E19	8,409	8,863	1,8	-11,272	-11,880
E20	8,523	8,222	1,0	-11,425	-11,021
E21	8,537	8,423	0,3	-11,444	-11,291
E22	8,537	8,156	1,4	-11,444	-10,933
E23	8,553	8,727	0,5	-11,465	-11,698
E25	8,638	8,606	0,0	-11,531	-11,536
E26	8,602	8,415	0,7	-11,579	-11,281
E27	8,744	8,778	0,1	-11,722	-11,766

Tabela 10, continua

Compostos	pIC <sub>50</sub> exp.	pIC <sub>50</sub> calc.	Fator off*	ΔG (exp)	ΔG (calc)
E28	8,824	8,924	0,3	-11,828	-11,962
E29	8,824	8,679	0,4	-11,828	-11,634
E31	8,959	8,893	0,2	-12,009	-11,921
E32	9,036	8,738	1,0	-12,113	-11,712
E33	9,050	8,960	0,2	-12,132	-12,011
E34	9,056	8,677	1,4	-12,139	-11,631
E36	9,113	8,427	3,9	-12,216	-11,297
E37	9,222	8,765	1,9	-12,362	-11,749
E38	9,292	8,898	1,5	-12,456	-11,928
E39	9,301	9,295	0,0	-12,468	-12,460
E41	9,398	9,396	0,0	-12,598	-12,595
E42	9,481	9,108	1,4	-12,710	-12,209
E43	9,745	9,779	0,1	-13,063	-13,108
Grupo Teste					
E02	6,750	6,711	0,1	-9,048	-8,996
E07	7,376	7,979	3,0	-9,888	-10,696
E11	7,921	8,040	1,5	-10,239	-10,778
E18	8,260	7,990	0,9	-11,072	-10,711
E24	8,553	8,413	0,4	-11,465	-11,277
E30	8,920	8,523	1,5	-11,958	-11,425
E35	9,086	8,685	1,5	-12,180	-11,642
E40	9,301	8,856	1,8	-12,468	-11,871

pIC<sub>50</sub> = -log(IC<sub>50</sub>); Fator *off*: desvio máximo individual e rms entre o valor experimental e o valor calculado de IC<sub>50</sub>, calculam-se como IC<sub>50</sub>(exp)/IC<sub>50</sub>(calc) - 1 para IC<sub>50</sub>(exp) > IC<sub>50</sub>(calc) e IC<sub>50</sub>(calc)/IC<sub>50</sub>(exp) - 1 para IC<sub>50</sub>(calc) > IC<sub>50</sub>(exp); ΔG<sub>exp</sub>: energia livre de Gibbs experimental; ΔG<sub>calc</sub>: energia livre de Gibbs calculada.

O ligante **E08** apresentou um desvio máximo individual, correspondendo ao *fator off* de 10,3 unidades (Tabela 10), portanto, as características físico-químicas calculadas para o composto **E08** têm o maior desvio em relação a sua atividade biológica experimental. No grupo de teste, o ligante **E07** apresentou um desvio máximo individual, correspondendo ao *fator off* de 3,0 unidade (Tabela 10). Para uma melhor visualização, o Gráfico 4 está

dispostos os valores de atividade biológica experimental e calculado, evidenciando os compostos com o maior desvio encontrado.

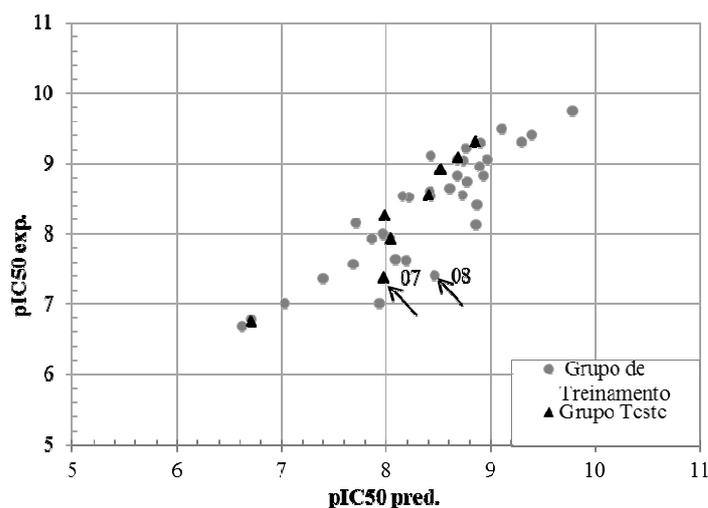


Gráfico 4 Valores da atividade biológica experimental e calculada

A Figura 11 representa a superfície do receptor hipotético, com predominância hidrofóbica, e pequenas regiões hidrofílicas. Notam-se regiões doadoras de ligações de hidrogênio próximas ao grupo triazol, em uma pequena região na posição R2 de compostos análogos ao YM511, e a substituintes polares na posição R2 de compostos análogos ao Letrozol. Substituintes polares na posição R3 de compostos análogos ao YM511 observam-se regiõesceptoras de ligações de hidrogênio. As regiões com as cores marrom claro apresentam características hidrofóbicas positivamente carregadas, na qual em comparação ao receptor biológico pode se inferir a presença de aminoácidos como a Lisina, Histidina ou Arginina. Nas regiões de cor cinza apresentam característica hidrofóbica, na qual se observa a presença dos anéis aromáticos.

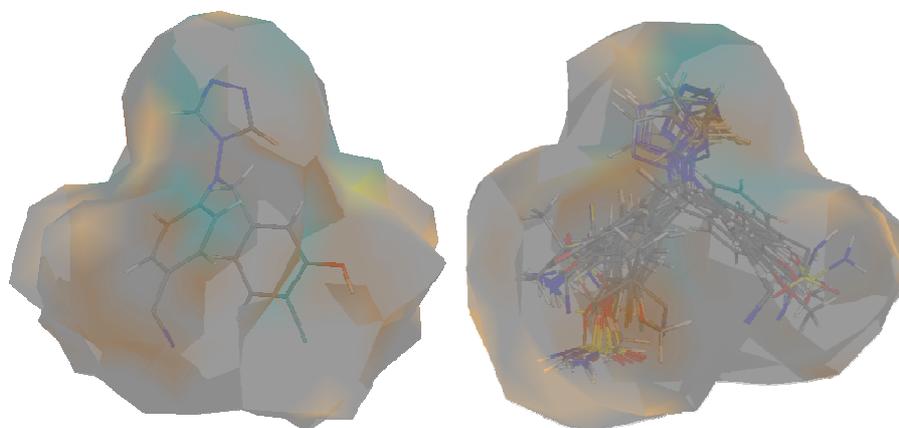


Figura 11 Representação da superfície mapeada pelas propriedades “quasi atomística” do modelo do receptor, gerada pelo *Quasar* utilizando, à esquerda a estrutura tridimensional do composto E43 e à direita todos os compostos. As cores significam: castanho = muito hidrofóbico; amarelo = hidrofóbico; cinza = neutro ou hidrofóbico; vermelho = acceptor de ligação de hidrogênio; azul = doadora de ligação de hidrogênio

Pode-se dizer que, um inibidor que preenche os requisitos necessários à superfície do receptor hipotético gerado, devendo manter os grupos triazol, benzil e ciano, os análogos ao YM511 possuir substituinte acceptor de ligações de hidrogênio em R3 substituinte, corroborando com o estudo 5D.

No *Quasar*, as interações ligante-receptor podem ser analisadas por meio das contribuições dos grupos funcionais à energia livre. A idéia por de trás desta opção, é visualizar os efeitos de melhora ou atenuantes de um grupo funcional com relação à afinidade de ligação e, assim, facilitar o processo do *design* de fármacos.

Na Figura 12 observa-se que o grupo triazol e substituintes em R3 do composto **E43**, possui maior contribuição na energia livre,  $\Delta G_{\text{calc.}}$ , enquanto o grupo ciano contribui de forma satisfatória e a hidroxila do grupo fenol muito pouco à energia livre.

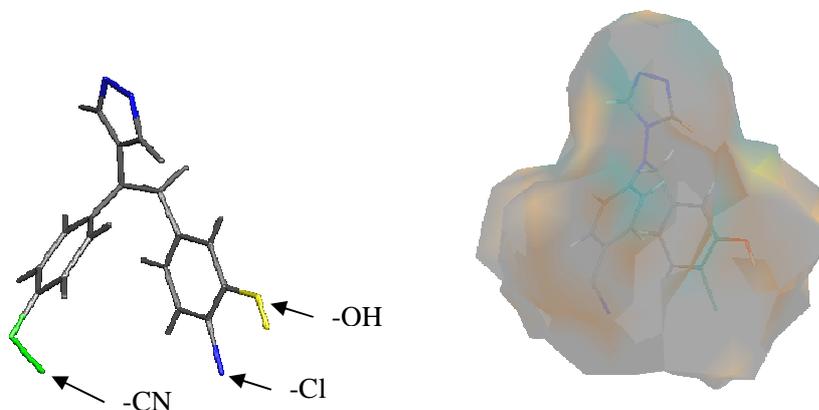


Figura 12 Estereorepresentação do composto E43. À esquerda de acordo com a contribuição do grupo funcional ao  $\Delta G$  (azul = muito bom, verde=bom, branco=neutro, amarelo=fraco, vermelho=ruim, cinza=não resolvido), à direita representação da superfície mapeado do receptor

Os compostos **E33** (Letrozol), **E40** (4-[(4-Bromobenzilo)[1,2,4] Triazol-4-Ylamino] Benzonitrila) e seus análogos: *p*-fenóis: **E16**, **E31** e **E42**; os *m*-fenóis: **E39**, **E41** e **E43**; *m*-sulfamato **E19** e *p*-sulfamato **E28**, apresentaram bons valores de  $IC_{50}$  calculados e  $\Delta G_{\text{calculado}}$  tanto na simulação de 5D quanto no 6D, a partir destes resultados obtidos supõe que modificações do tipo polares nas posições R3 dos *p*-fenóis, nos *m*-fenóis nas posições R3 com substituintes volumosos, nos *m*-sulfamato e *p*-sulfamato substituir o átomo de nitrogênio do grupo sulfamato, nestes compostos possam melhorar os requisitos do receptor hipotético gerado, e posteriormente calcular sua energia ( $\Delta G$ ) e prever sua potência inibitória pela relação entre energia e  $IC_{50}$ .

Na Figura 13, tem-se a visualização gráfica dos grupos funcionais que contribuem de forma satisfatória ou dos que pouco contribui a energia livre,  $\Delta G_{\text{calc}}$ , dos compostos **E16**,**E19**,**E28**,**E31**,**E33**,**E39**,**E40**,**E41** e **E42**.

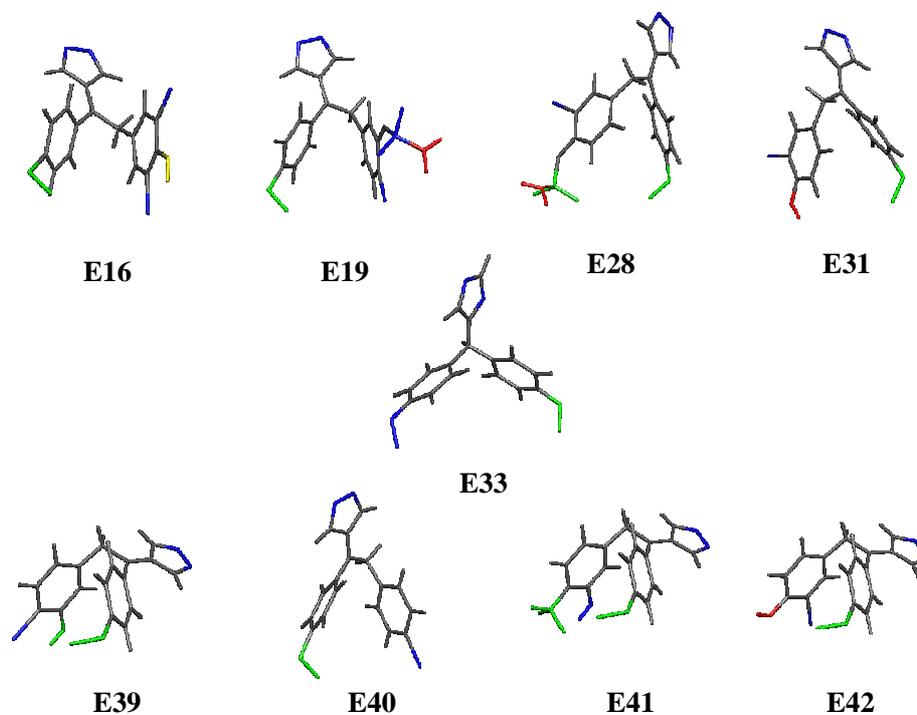


Figura 13 Visualização gráfica das contribuições à energia livre calculada dos compostos análogos ao YM511 (azul = muito bom, verde=bom, branco=neutro, amarelo=fraco, vermelho=ruim, cinza=não resolvido)

Observa-se em todos os compostos que o grupo triazol contribuiu de forma excelente a energia livre, enquanto os átomos de nitrogênio dos grupos sulfamato dos compostos **19** e **28**, e as hidroxilas dos compostos **31** e **42** contribui muito pouco a energia livre, influenciando em suas atividades biológicas.

## 5 CONCLUSÃO

No estudo de QSAR, os modelos gerados fornecem importantes informações sobre o sítio ativo de uma proteína alvo, as quais podem ser aplicadas no planejamento de novos inibidores da aromatase a partir das propriedades físico-químicas.

Os modelos gerados em QSAR 5D e 6D demonstraram consenso sobre as regiões hidrofóbicas do sítio ativo da enzima, correspondendo aos anéis aromáticos presentes nos compostos deste estudo. Pelo modelo QSAR-5D para melhorar os compostos devem possuir substituintes polares na posição R3 de compostos análogos ao YM511 e a substituintes polares na posição R2 de compostos análogos ao Letrozol. Utilizando o QSAR-6D, que considera diversos modelos de solvatação, os resultados corroboraram com os resultados obtidos no 5D.

Os compostos **33** (Letrozol), **40** (4-[(4-Bromobenzilo)[1,2,4] Triazol-4-Ylamino] Benzonitrila) e seus análogos: *p*-fenois: **16**, **31** e **42**; os *m*-fenois: **39**, **41** e **43**; *m*-sulfamato **19** e *p*-sulfamato **28**, apresentaram os melhores valores  $IC_{50\text{calculado}}$  e  $\Delta G_{\text{calculado}}$ , nos quais modificações propostas podem resultar em inibidores mais potentes, mantendo assim o grupo triazol, benzil e ciano. De acordo como os resultados obtidos estes modelos são passíveis de serem aplicados na formulação de novos compostos para o tratamento de câncer de mama.

## **6 PERSPECTIVAS**

Como perspectivas, o estudo desenvolvido nos motiva a dar continuidade à busca pelo conhecimento a cerca destes novos compostos hipotéticos e estudar os inibidores da enzima sulfatase, a qual é responsável pela hidrólise de esteroides alquil e aril sulfatos formando estrona no tecido com câncer de mama, em que foi demonstrado que a formação deste estrogênio é 10 vezes mais originada pelo sulfato de estrona do que androstenediona.

Outro objetivo é desenvolver um estudo pré clínico, a fim de avaliar a potencialidade terapêutica destes compostos, validando o modelo adquirido pelo QSAR-5D/6D, a fim de que estes possam ser utilizados no tratamento de câncer de mama.

## REFERÊNCIAS

ALBERTS, B. et al. **INTRODUCCIÓN LA BIOLOGÍA CELULAR**. 2th. ed. Buenos Aires: Médica Panamericana, 2006. 864 p. Disponível em: <[http://books.google.com.br/books?id=qrrYZJhrRm4C&printsec=frontcover&dq=bruce+alberts&hl=pt-BR&ei=LBmnTaLiGIfy0gHV8H5CA&sa=X&oi=book\\_result&ct=result&resnum=2&ved=0CCwQ6AEwATgK#v=onepage&q&f=false](http://books.google.com.br/books?id=qrrYZJhrRm4C&printsec=frontcover&dq=bruce+alberts&hl=pt-BR&ei=LBmnTaLiGIfy0gHV8H5CA&sa=X&oi=book_result&ct=result&resnum=2&ved=0CCwQ6AEwATgK#v=onepage&q&f=false)>. Acesso em: 01 jun. 2009.

ALTUNDAG, K.; IBRAHIM, N. K. Inhibitors in breast cancer: an overview. **The Oncologist**, Durham, v. 11, p. 553-562, June 2006.

AMERICAN CANCER SOCIETY. **The history of cancer**. Disponível em: <[http://www.cancer.org/docroot/CRI/content/CR\\_I26x\\_the\\_history\\_of\\_cancer72.asp](http://www.cancer.org/docroot/CRI/content/CR_I26x_the_history_of_cancer72.asp)>. Acesso em: 29 mar. 2010.

BERMUDEZ-PIRELA, V. et al. Quimioprevenção del cáncer de mama: fronteras y horizontes. **Avft**, Caracas, v. 24, n. 1, p. 32-41, 2005.

BRUEGGEMEIER, R. W.; HACKETT, J. C.; DIAZ-CRUZ, E. S. Aromatase Inhibitors in the treatment of breast cancer. **Endocrine Reviews**, Baltimore, v. 26, n.3, p. 331-345, Jan. 2005.

BUIJS, C. et al. The influence of endocrine treatments for breast cancer on health-related quality of life. **Cancer Treatment Reviews**, London, v. 34, p. 640-655, June 2008.

BUZDAR, A. U. Advances in endocrine treatments for postmenopausal women with metastatic and early breast cancer. **The Oncologist**, Durham, v.7, p. 335-341, Aug. 2001.

CASA DE AMPARO A MÃE E AO BEBÊ E INCENTIVO A AMAMENTAÇÃO. **Anatomia da mama**. Disponível em: <[http://www.cambia.blogspot.com/2009\\_02\\_27\\_archive.html](http://www.cambia.blogspot.com/2009_02_27_archive.html)>. Acesso em: 15 fev. 2010.

CAMPOS, S. M. Aromatase inhibitors for breast cancer in postmenopausal women. **The Oncologist: Breast Cancer**, Durham, v. 9, p. 126-136, Apr. 2004.

CROCE, C. M. Molecular origins of cancer: oncogenes and cancer. **The New England Journal Of Medicine**, Waltham, v. 358, p. 502-511, Jan. 2008.

DOBLER, M.; VEDANI, A.; LILL, M. A. **User and reference manual BiografX 1.5 -consensus scoring using the Quasar, Raptor and Symposar technologies**. Basel: Biographics Laboratory 3R, 2009. 30 p.

FAVIA, A. D. et al. Three-dimensional model of the human aromatase enzyme and density functional parameterization of the iron-containing protoporphyrin IX for a molecular dynamics study of heme-cysteinato cytochromes. **Proteins: Structure, Function, and Bioinformatics**, New York, v. 62, p. 1074-1087, Mar. 2006.

FOYE, W. O. et al. **Foye's principles of medicinal chemistry**. 6th ed. Baltimore: Lippincott Williams & Wilkins, 2008. 1377 p.

GAD, S. C. **Preclinical development handbook: ADM**, biopharmaceutical properties. 2th. ed. New Jersey: J. Wiley, 2008. cap. 1, p. 1-46.

GARCIA, M. et al. **Global cancer facts & figures 2007**. Atlanta: American Cancer Society, 2007. 52 p.

GHOSH, D. et al. Structural basis for androgen specificity and oestrogen synthesis in human aromatase. **Nature**, London, v. 457, p. 219-223, Jan. 2009.

HISCOX, S.; DAVIES, E. L.; BARRETT-LEE, P. Aromatase inhibitors in breast cancer. **Maturitas**, Amsterdam, v. 63, p. 275-279, May 2009.

INSTITUTO NACIONAL DE CANCER. **Estimativa 2008: incidência de câncer no Brasil**. Rio de Janeiro: Coordenação De Educação (CEDC) Serviço de Edição e Informação Técnico-científica, 2007. 98 p.

INSTITUTO NACIONAL DE CANCER. **Estimativa 2010: incidência de câncer no Brasil**. Rio de Janeiro: Coordenação De Educação (CEDC) Serviço de Edição e Informação Técnico-científica, 2009. 98 p.

JACKSON, T. et al. Dual aromatase-sulfatase inhibitors based on the anastrozole template: synthesis, in vitro SAR, molecular modelling and in vivo activity. **Organic & Biomolecular Chemistry**, Cambridge, v. 5, p. 2940-2952, July 2007.

JONG, P. C. de et al. Inhibition of breast cancer tissue aromatase activity and estrogen concentrations by the third-generation aromatase inhibitor Vorozole. **Cancer Research**, Baltimore, v. 57, p. 2109-2111, June 1997.

LENHARD, R. E.; OSTEEEN, R. T.; GANSLER, T. S. **Clinical oncology**. Atlanta: American Cancer Society, 2001. 919 p.

LEONETTI, F. et al. Design, synthesis, and 3D QSAR of novel potent and selective aromatase inhibitors. **Journal of Medicinal Chemistry**, Washington, v. 47, p. 6792-6803. June 2004.

LILL, M. A.; VEDANI, A.; DOBLER, M. Raptor: Combining dual-shell representation, induced-fit simulation, and hydrophobicity scoring in receptor modeling: application toward the simulation of structurally diverse ligand sets. **Journal of Medicinal Chemistry**, Washington, v. 47, n. 25, p. 6174-6186, Apr. 2004.

LILL, M. A. et al. Impact of induced fit on ligand binding to the androgen receptor: a multidimensional QSAR study to predict endocrine-disrupting effects of environmental chemicals. **Journal of Medicinal Chemistry**, Washington, v. 48, p. 5666-5674. Apr. 2005.

LILL, M. A.; VEDANI, A. Combining 4D pharmacophore generation and multidimensional QSAR: modeling ligand binding to the bradykinin b2 receptor. **Journal of Chemical Information and Modeling**, Washington, v. 46, n. 5, p. 2135-214518, Aug. 2006.

LILL, M. A. Multi-dimensional QSAR in drug discovery. **Drug Discovery Today**, Kidlington, v. 12, n. 23/24, p. 1013-1017, Dec. 2007.

LOTTI, R. C. B. et al. Impacto do tratamento de cancer de mama na qualidade de vida. **Revista Brasileira de Cancerlogia**, Rio de Janeiro, v. 54, n. 4, p. 367-371, maio 2008.

MILLER, W. R. et al. Aromatase inhibitors: are there differences between steroidal and nonsteroidal aromatase inhibitors and do they matter? **The Oncologist**, Durham, v. 13, p. 829-837, Dec. 2008.

MYINT, K. Z.; XIE, X. Recent advances in fragment-based QSAR and multi-dimensional QSAR methods. **Zeolites: the International Journal of Molecular Sieves**, Guildford, v. 11, p. 3846-3866, Oct. 2010.

NATIONAL CANCER INSTITUTE. **Dictionary of cancer** terms. Disponível em: <<http://www.cancer.gov/dictionary/?CdrID=45333>>. Acesso em: 20 mar. 2010.

NELSON, D. L.; COX, M. M. **Lehninger principles of biochemistry**. 4th. ed.. New York: Worth Publishers, 2005. 1100 p.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Guidance Document On The Validation Of (Quantitative) Structure-Activity Relationship [(Q)Sar] Models.** 69th ed. Paris: Inter-organisation Programme for the Sound Management of Chemicals, 2007. 154 p.

OLIVEIRA, A. A.; RAMALHO, T. C.; CUNHA, E. F. F. QSAR study of androstenedione analogs as aromatase inhibitors. **Letters In Drug Design & Discovery**, [S.l.], v. 6, n. 8, p. 554-562, Dec. 2009.

PARL, F. F. **Estrogens, estrogen receptor and breast cancer:** estrogen synthesis and metabolism. Amsterdam: Ios Press, 2000. 263 p. Disponível em: <[http://books.google.com.br/books?id=v7ai5Mz9TZQC&printsec=frontcover&dq=Estrogens,+Estrogen+Receptor+and+Breast+Cancer&hl=pt-BR&ei=vaGDTYmCHcycgQfXxuXfCA&sa=X&oi=book\\_result&ct=result&resnum=1&ved=0CDQQ6AEwAA#v=onepage&q&f=false](http://books.google.com.br/books?id=v7ai5Mz9TZQC&printsec=frontcover&dq=Estrogens,+Estrogen+Receptor+and+Breast+Cancer&hl=pt-BR&ei=vaGDTYmCHcycgQfXxuXfCA&sa=X&oi=book_result&ct=result&resnum=1&ved=0CDQQ6AEwAA#v=onepage&q&f=false)>. Acesso em: 2 maio 2009.

PUZYN, T.; LESZCZYNSKI, J.; CRONIN, M. T. D. **Recent advances in QSAR studies methods and applications.** New York: Springer Dordrecht Heidelberg, 2009. cap. 1, p. 3-11.

RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS. **Protein Data Bank.** Disponível em: <<http://www.rcsb.org>>. Acesso em: 4 mar. 2009.

ROY, P. P.; ROY, K. Docking and 3D-QSAR studies of diverse classes of human aromatase (CYP19) inhibitors. **Journal Molecular Model**, Berlin, v. 16, p. 1597-1616, Mar. 2010.

RUGO, H.S. The breast cancer continuum in hormone-receptor-positive breast cancer in postmenopausal women: evolving management options focusing on aromatase inhibitors. **Annals of Oncology**, Dordrecht, v. 19, n. 1, p. 16-27, Jan. 2008.

SELASSIE, C. D. History of quantitative structure-activity relationships. In: BURGER, A.; ABRAHAM, D. J. **Burger's Medicinal Chemistry and Drug Discovery**. 6th. ed. London: J. Wiley, 2003. cap. 1, p. 1-48.

SIMPSON, E. R.; DOWSETT, M. Aromatase and its inhibitors: significance for breast cancer therapy. **Recent Progress in Hormone Research: Proceedings of the Laurentian Hormone Conference**, San Diego, v. 57, p. 317-338, Mar. 2002.

SOUZA, T. C. S. **Estudos quantitativos multidimensionais de correlação estrutura-atividade de inibidores peptídicos da HIV-1 protease**. 2009. 114 p. Dissertação (Mestrado em Agroquímica)-Universidade Federal de Lavras, Lavras, 2009.

SPENCE, A. P. **Anatomia humana básica**. 2. ed. São Paulo: Barueri, 1991. 713 p.

TAVARES, L.C. QSAR: a abordagem de Hansch. **Química Nova**, São Paulo, v. 27, n. 4, p. 631-639, 2004.

TURGEON, J. L. et al. Hormone therapy: physiological complexity belies therapeutic simplicity. **Science**, Hanover, v. 304, p. 1269-1273, July 2004.

UNIVERSIDADE FEDERAL DE SÃO PAULO. **Ovário**. Disponível em: <<http://www.unifesp.br/dmorfo/histologia/ensino/ovario/img/021.jpg>>. Acesso em: 25 mar. 2010.

VEDANI, A.; DOBLER, M. 5D-QSAR: The key for simulating induced fit?. **Journal of Medicinal Chemistry**, Washington, p. 2139-2149, Oct. 2002.

VEDANI, A.; DOBLER, M.; LILL, M. A. combining protein modeling and 6D-QSAR. simulating the binding of structurally diverse ligands to the estrogen receptor. **Journal of Medicinal Chemistry**, Washington, v. 48, p. 3700-3703, June 2005.

WENG, Z.; VAJDA, S.; DELIS, C. Prediction of protein complexes using empirical free energy functions. **Protein Science**, Cambridge, v. 5, p. 614-626, Sept. 1996.

WORLD HEALTH ORGANIZATION - WHO. **Global action against cancer - updated version, 2005**. Disponível em: < <http://www.who.int/cancer/media/en/GlobalActionCancerEnglfull.pdf> >. Acesso em: 02 dez. 2010.

WOOD, P. M. et al. Chiral aromatase and dual aromatase - steroid sulfatase inhibitors from the letrozole template: synthesis, absolute configuration, and in vitro activity. **Journal Medicinal Chemistry**, Washington, v. 51, p. 4226-4238, Feb. 2008.

WOO, L. W. L. et al. Dual aromatase - steroid sulfatase inhibitors. **Journal Medicinal Chemistry**, Washington, v. 50, p. 3540-3560, Jan. 2007.