



DANIEL GEDDER SILVA

**MODELAGEM RACIONAL DE NOVOS
DIPEPTÍDEOS MODIFICADOS COMO
INIBIDORES DA ENZIMA CONVERSORA DE
ANGIOTENSINA**

LAVRAS – MG

2013

DANIEL GEDDER SILVA

**MODELAGEM RACIONAL DE NOVOS DIPEPTÍDEOS
MODIFICADOS COMO INIBIDORES DA ENZIMA CONVERSORA DE
ANGIOTENSINA**

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

Orientador

Dr. Matheus Puggina de Freitas

Coorientadora

Dra. Elaine Fontes Ferreira da Cunha

LAVRAS – MG

2013

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

Silva, Daniel Gedder.

Modelagem racional de novos dipeptídeos modificados como inibidores da enzima conversora de angiotensina / Daniel Gedder Silva. – Lavras : UFLA, 2013.

120 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2013.

Orientador: Matheus Puggina de Freitas.

Bibliografia.

1. Ancoramento molecular. 2. Anti-hipertensivos. 3. MIA-QSAR. 4. QSAR-TRADICIONAL. 5. Hipertensão arterial. 6. Métodos computacionais. I. Universidade Federal de Lavras. II. Título.

CDD – 547.756

DANIEL GEDDER SILVA

**MODELAGEM RACIONAL DE NOVOS DIPEPTÍDEOS
MODIFICADOS COMO INIBIDORES DA ENZIMA CONVERSORA DE
ANGIOTENSINA**

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

Aprovada em 28 de fevereiro de 2013.

Dra. Maria das Graças Cardoso UFLA

Dr. Carlos Alberto Montanari IQSC/USP

Dr. Matheus Puggina de Freitas
Orientador

**LAVRAS – MG
2013**

Àos que me engrandeceram.

Dedico...

AGRADECIMENTOS

A cada vitória o reconhecimento devido ao meu **Deus**, pois só Ele é digno de toda honra, glória e louvor.

Senhor, obrigado por permitir que eu continuasse nessa caminhada para o fim de mais uma etapa e o início de muitas outras.

E o que seria de mim sem o amor e apoio de minha **família**, sempre presente, a qual é responsável por grande parte desta vitória. Obrigado, meu superpai Geraldo e às mulheres da minha vida: minha mãe, Zulmira; minha irmã, Marielle e meu grande amor, Renata.

Agradeço ao meu **orientador**, Matheus e aos demais professores (Elaine, Teodorico, Cleiton e Silvana), pela imensa contribuição para a construção do saber e a concretização deste presente trabalho. À Capes, pela bolsa de estudos e à UFLA e ao DQI, pelo suporte.

Obrigado, também, aos amigos e colegas que sempre me auxiliaram e estiveram do meu lado, em especial ao Edilson.

*Quem não compreende um olhar,
tampouco compreenderá uma longa explicação...*

Mario Quintana

RESUMO

Em razão da elevada taxa de mortalidade decorrente das doenças cardíacas, a hipertensão arterial tornou-se um grave problema de saúde pública. Os medicamentos empregados para o controle dessa doença exibem maior eficácia quando ministrados em conjunto com outros fármacos, o que potencializa os efeitos colaterais. Dentre as classes de anti-hipertensivos utilizadas no tratamento, destacam-se os responsáveis por inibir a enzima conversora de angiotensina (ECA), enzima chave no processo de regulação da pressão arterial. O dipeptídeo modificado captopril, um dos fármacos mais utilizados no tratamento, foi o composto de referência para este presente estudo. Com o objetivo de propor novos dipeptídeos modificados, técnicas computacionais (MIA-QSAR, QSAR-tradicional e *docking*) foram aplicadas, associadas à estratégia de combinar subestruturas químicas de moléculas altamente bioativas, como inibidores da ECA. O modelo MIA-QSAR foi o escolhido para prever a atividade biológica dos novos inibidores propostos, devido aos resultados estatísticos satisfatórios obtidos das análises, quais sejam: $r^2_{Cal.}$ de 0,909, q^2 de 0,607 e r^2_{teste} de 0,779. O QSAR-tradicional com 5 parâmetros indicou algumas propriedades atômicas e moleculares que seriam importantes para aumentar a atividade biológica da série de compostos estudada, o que foi confirmado pelos estudos de ancoramento molecular (*docking*) e pelos dados experimentais disponíveis para duas estruturas propostas. A inserção de uma nova série de dipeptídeos e tripeptídeos ao conjunto de dados resultou nas propostas de novos compostos derivados de selênio. Um novo composto, **F**, exibiu um elevado valor de $pIC_{50pred.}$ (8,40), bem como ligação de hidrogênio adicional entre o átomo de selênio e Glu385, de acordo com os resultados de *docking*. Modificações estruturais (ligação dupla adicionada no anel de cinco membros) nos compostos **A** e **F** resultaram em melhorias nos valores de *docking score* (**J** = -93,2 kcal mol⁻¹ e **K** = -93,6 kcal mol⁻¹). Sugere-se que tais modificações estruturais e a presença de átomos, como selênio ou enxofre, determinaram o aumento nas bioatividades dos compostos propostos em relação às moléculas existentes. Além disso, os compostos propostos exibiram parâmetros farmacocinéticos calculados comparáveis ao captopril.

Palavras chave: Hipertensão arterial. Anti-hipertensivos. Enzima conversora de angiotensina. MIA-QSAR. QSAR-tradicional. Ancoramento molecular.

ABSTRACT

Because of the high mortality caused by cardiac diseases, blood hypertension has become alarming. Medicines usually employed to control this disease are more efficient when administered together with other drugs, thus enhancing side effects. Inhibitors of the angiotensin-converting enzyme (ACE), the key enzyme responsible for the regulation of blood pressure, are the main class of antihypertensives used for the treatment. The dipeptide-like Captopril is widely used for the treatment of hypertension and was considered the standard compound for comparison with the results of this work. Computational techniques, namely MIA-QSAR, traditional QSAR and docking studies, together with the strategy of combining substructures of the main compounds in the ACE inhibitors series, were used to propose new drug candidates with improved profiles. The MIA-QSAR model was chosen to predict the biological activities of the proposed compounds, given the satisfactory statistical results obtained from the analysis, that is $r^2_{Cal.}$ of 0.909, q^2 of 0.607 and r^2_{test} of 0.779. The traditional QSAR using five parameters indicated some atomic and molecular properties as key roles of the biological activities for the series of compounds analyzed, which was confirmed by docking studies and also by experimental values for two proposed structures already available. Introduction of a new series of di- and tripeptide-like compounds into the data set resulted in new proposals, which are selenium derivatives. A given proposed compound (**F**) exhibited a high predicted $pIC_{50pred.}$ (8.40), as well as an additional hydrogen bond between the selenium atom and the Glu385 amino acid residue, according to the docking studies. Structural modification in compounds **A** and **F** (replacement of a single bond by a double bond in the five-membered ring to give an indolyl moiety) gave compounds **J** and **K** with improved docking scores ($-93.2 \text{ kcal mol}^{-1}$ and $-93.6 \text{ kcal mol}^{-1}$, respectively). Such structural changes and the presence of atoms like selenium and sulfur are suggested to determine the increase in bioactivities of the proposed compounds relative to the existing inhibitors. Moreover, the proposed compounds were calculated to exhibit pharmacokinetic parameters comparable to those of Captopril.

Keywords: Blood hypertension. Antihypertensives. Angiotensin-converting enzyme. MIA-QSAR. Traditional QSAR. Molecular docking.

LISTA DE ILUSTRAÇÕES

Figura 1	Estrutura química da classe de compostos pertencentes ao grupo denominado IECAs. As diferenças e as similaridades encontram-se destacadas na cor azul e vermelha, respectivamente. O captopril é o fármaco de referência neste estudo.....	28
Figura 2	Estrutura tridimensional da ECA, extraída do PDB, com os respectivos códigos: captopril (1UZF), lisinopril (1086) e enalapril (1UZE).....	31
Figura 3	Representação do sistema renina-angiotensina.....	33
Figura 4	Processo de clivagem dos aminoácidos (destacado em vermelho) realizado pelas enzimas renina, ECA e angiotensinase.....	34
Figura 5	Principais razões de insucesso de novas entidades químicas no processo de desenvolvimento.....	36
Figura 6	Etapas do processo de descoberta e planejamento racional de fármacos.....	39
Figura 7	Cronologia da construção da metodologia QSAR.....	42
Figura 8	Distinção entre os formalismos da metodologia QSAR....	44
Figura 9	Ilustração das diferenças nas representações das moléculas presentes no conjunto de dados deste trabalho (moléculas extraídas de GRUENFELD et al., 1983). A representação de pixels pretos e espaços em brancos são as diferenças determinantes que explicam a variação no bloco das variáveis dependentes (no caso, as atividades biológicas).....	48

Figura 10	Diagrama das etapas envolvidas na realização do presente estudo.....	56
Figura 11	Primeiras etapas para a construção do modelo MIA-QSAR, que consiste na preparação e na organização do conjunto de dados. Dois programas são utilizados: ChemSketch (para desenhar as estruturas químicas) e o aplicativo Paint do Microsoft Windows (para alinhar e converter as estruturas químicas em imagens).....	63
Figura 12	Resultado da sobreposição de todas as imagens, formando um arranjo tridimensional. A parte da imagem em negrito corresponde à parte similar da série, os substituintes são as porções variáveis nas moléculas e, portanto, responsáveis pelas alterações nos valores de atividade biológica.....	65
Figura 13	Construção do arranjo tridimensional e desdobramento para a matriz X. Na parte superior da figura, a interface do programa MatLab; na parte inferior, a demonstração do processo de desdobramento.....	66
Figura 14	Processo de validação cruzada leave-one-out. Uma molécula de cada vez (na Matriz X) e seu respectivo valor de pIC_{50exp} . (Bloco Y) são deixados de fora (em vermelho) e procede-se com a regressão; utilizam-se os parâmetros de regressão para calcular o valor de pIC_{50} da amostra deixada de fora. Repete-se o procedimento até que todas as amostras tenham sido deixadas de fora...	69
Figura 15	Processo de validação externa. Um conjunto de moléculas que não faz parte do conjunto de treinamento é usado para validar o modelo.....	70

Figura 16	Randomização do Bloco Y, seguida da regressão por meio de PLS. O processo é repetido algumas vezes, a fim de avaliar se a verdadeira correlação é obra do acaso ou não.....	71
Figura 17	Moléculas mais biologicamente ativas de cada série congênere (15, 27 e 34), escolhidas por meio dos valores de pIC50exp.(Tabela 7). A parte em vermelho de uma molécula foi combinada com a parte preta de outra molécula.....	90
Figura 18	Estruturas propostas usando a estratégia de combinar subestruturas dos compostos mais ativos experimentalmente da série. Os valores de pIC50 preditos obtidos a partir do modelo MIA-QSAR são apresentados entre parêntesis.....	91
Figura 19	Moléculas mais biologicamente ativas da nova série congênere (88 e 98), escolhidas por meio dos valores de pIC50exp. (Tabela 8). A parte em vermelho de uma molécula é combinada com a parte preta de outra molécula.....	98
Figura 20	Novas estruturas propostas, utilizando-se a estratégia de combinar subestruturas dos compostos mais ativos experimentalmente da série. Os valores de pIC50pred. obtidos a partir do modelo MIA-QSAR são apresentados entre parêntesis.....	99
Figura 21	Principais interações (ligações de hidrogênio) de F (sobreposto ao captopril) com o sítio ativo da ECA.....	101
Figura 22	Principais interações (ligações de hidrogênio) de F com o sítio ativo da ECA. Legenda: Compd.= Composto.....	102

Figura 23	Novas estruturas propostas J e K; uma dupla ligação foi acrescentada nos compostos A e F (destacada em vermelho), resultando em um indol modificado. Os valores de docking score estão descritos abaixo das moléculas.....	103
Gráfico 1	Ocorrência de casos de hipertensão arterial nas regiões de atuação da OMS.....	23
Gráfico 2	Determinação do número ótimo de variáveis latentes PLS, para o modelo MIA-QSAR.....	75
Gráfico 3	pIC50 experimentais vs. ajustados/preditos utilizando-se o método MIA-QSAR para a série de peptídeos modificados (inibidores da ECA).....	77
Gráfico 4	Cálculo do número ótimo de variáveis latentes PLS, para o modelo QSAR-tradicional com 1.459 descritores físico-químicos Dragon.....	79
Gráfico 5	Representação gráfica dos resultados do PCA. Gráfico de pesos da componente principal 2 (CP2) eixo Y versus componente principal 1 (CP1) eixo X.....	81
Gráfico 6	Representação gráfica dos resultados do PCA para a etapa 'd'. A CP1 contém 98,41% e a CP2 contém 1,49% de informação.....	83
Gráfico 7	Cálculo do número ótimo de variáveis latentes PLS, para o novo modelo MIA-QSAR.....	96
Gráfico 8	pIC50 experimentais vs. ajustados/preditos, utilizando-se o método MIA-QSAR, para a nova série de 102 compostos.....	97
Gráfico 9	Correlação entre valores de pIC50 preditos de A-I usando MIA-QSAR e os scores de docking.....	100

LISTA DE TABELAS

Tabela 1	Classificação da pressão arterial de acordo com a medida casual no consultório (> 18 anos).....	21
Tabela 2	Anti-hipertensivos comercialmente disponíveis no Brasil..	24
Tabela 3	Série de compostos usados nas análises QSAR.....	58
Tabela 4	Resultados estatísticos para o modelo MIA-QSAR, utilizando 6 variáveis latentes PLS.....	75
Tabela 5	Resultados estatísticos do modelo QSAR-TRADICIONAL com 1459 descritores físico-químicos Dragon, utilizando 2 e 3 VL.....	80
Tabela 6	Resultados estatísticos dos modelos QSAR-TRADICIONAL e MIA-QSAR (incluído para comparação).....	86
Tabela 7	Valores de pIC50exp. usados para construção do modelo; pIC50pred. e pIC50LOOCV, obtidos pelos modelos MIA-QSAR e QSAR-TRADICIONAL construído.....	87
Tabela 8	Nova série de compostos usados nas análises do novo modelo MIA-QSAR.....	94
Tabela 9	Resultados estatísticos do novo modelo MIA-QSAR, utilizando 5 variáveis latentes PLS.....	96
Tabela 10	Parâmetros calculados da "regra dos cinco" de Lipinski e TPSA para os compostos propostos (A-K), e para o composto de referência Captopril (27).....	104
Tabela 11	Parâmetros farmacocinéticos computados para os compostos propostos (A-K), e para o composto de referência Captopril (27).....	105

LISTA DE ABREVIATURAS

Ala	Alanina
Asp	Aspartato
bmp	Bitmaps
Calib.	Calibrado
Gln	Glutamina
Glu	Glutamato
His	Histidina
Leu	Leucina
Log	Logaritmo
Lys	Lisina
Me	Metil
Param.	Parâmetro
px	Pixel
Tyr	Tirosina

LISTA DE SIGLAS

3D	Terceira dimensão
4D	Quarta dimensão
5D	Quinta dimensão
5-HT _{2C}	Antagonistas dos receptores
6D	Sexta dimensão
ADME	Absorção, distribuição, metabolismo e excreção
ADMET	Absorção, distribuição, metabolismo, excreção e toxicidade
ANVISA	Agência Nacional de Vigilância Sanitária
ARAs	Antagonistas dos receptores de angiotensina II
BBB	<i>Blood brain barrier penetration</i>
CoMFA	<i>Comparative molecular field analysis</i>
CoMSIA	<i>Comparative molecular similarity indices analysis</i>
CP1	Componente principal 1
CP2	Componente principal 2
D2	Subtipo do receptor de dopamina
DQI	Departamento de Química
ECA	Enzima conversora de angiotensina
EMEA	<i>European Medicines Agency</i>

F	<i>Oral bioavailability</i>
FDA	<i>Food and Drug Administration</i>
GA	<i>Genetic algorithm</i>
HIA	<i>Human intestinal absorption</i>
HIV	<i>Human immunodeficiency virus</i>
HQSAR	<i>Holograma QSAR</i>
IECAs	<i>Inibidores da enzima conversora de angiotensina</i>
LOOCV	<i>Leave one out cross validation</i>
LQTA-QSAR	<i>Laboratório de Quimiometria Teórica e Aplicada-QSAR</i>
MIA-QSAR	<i>Multivariate image analysis - QSAR”</i>
MVD	<i>Molegro virtual docker</i>
MW	<i>Molecular weight</i>
NDA	<i>New drug application</i>
OMS	<i>Organização Mundial da Saúde</i>
PAD	<i>Pressão arterial diastólica</i>
PAS	<i>Pressão arterial sistólica</i>
PCA	<i>Principal component analysis</i>
PDB	<i>Protein data bank</i>
PhRMA	<i>Pharmaceutical Research and Manufacturers of America</i>
PIB	<i>Produto Interno Bruto</i>
PLS	<i>Partial least squares</i>
QSAR	<i>Quantitative structure activity relationship</i>
RGB	<i>Red green blue</i>
RLM	<i>Regressão linear múltipla</i>
RMSE	<i>Root mean square error</i>
RMSECV	<i>Root mean square error cross validation</i>
SAR	<i>Structure activity relationship</i>
S-RAA	<i>Sistema renina-angiotensina-aldosterona</i>
SSS	<i>Sistema Suplementar de Saúde</i>
SUS	<i>Sistema Único de Saúde</i>
TPSA	<i>Topological polar surface area</i>
UFLA	<i>Universidade Federal de Lavras</i>
VL	<i>Variáveis latentes</i>
WEE1	<i>Enzima tirosina quinase</i>

LISTA DE SÍMBOLOS

KDa	Unidade de massa atômica
P_h	Coefficiente de partição de um composto de origem
P_x	Termo derivante
Log P	Coefficiente de partição octanol/água
r^2	Coefficiente de correlação
q^2	Coefficiente de correlação com validação cruzada
pIC ₅₀	Log de IC ₅₀
IC ₅₀	50% de concentração inibitória para a ECA, em mol L ⁻¹
$r^2_{Cal.}$	Coefficiente de correlação da calibração
pIC _{50exp.}	Log de IC ₅₀ experimental
pIC _{50LOOCV}	Log de IC ₅₀ LOOCV
pIC _{50pred}	Log de IC ₅₀ predito
r^2_{teste}	Coefficiente de correlação da validação externa
r^2_m	Coefficiente de correlação modificado
R^2_{pe}	Diferença entre coeficiente de correlação randomizado calibrado
E_{score}	Função score
E_{Inter}	Energia de interação ligante-proteína
E_{Intra}	Energia interna do ligante
n_{ON}	Número de aceptores de ligação de hidrogênio
n_{OHNH}	Número de doadores de ligação de hidrogênio
y_i	Valores de pIC ₅₀ experimentais
\hat{y}_i	Valores estimados/preditos de pIC ₅₀
\bar{y}	Valores médios de pIC ₅₀
r_0^2	Coefficiente de correlação ao quadrado sem intercepto
R^2_{rand}	Coefficiente de correlação do <i>Y-randomization</i>
kcal mol ⁻¹	Quilocaloria por mol
n_{viol}	Número de violações a regra de Lipinsk
n_{rotb}	Valores de flexibilidade molecular

SUMÁRIO

1	INTRODUÇÃO.....	18
2	REFERENCIAL TEÓRICO.....	20
2.1	Hipertensão arterial: situação global.....	20
2.1.1	Anti-hipertensivos.....	24
2.1.2	Inibidores diretos da renina, IECAs e ARAs.....	27
2.2	Enzima conversora de angiotensina (ECA).....	30
2.2.1	O sistema renina-angiotensina-aldosterona (S-RAA).....	32
2.3	O desenvolvimento de novos fármacos.....	35
2.4	QSAR, metodologia de auxílio no processo de elaboração de novos fármacos.....	39
2.4.1	QSAR, inserção de parâmetros moleculares tridimensionais.....	43
2.4.2	Descritores físico-químicos na construção de modelos QSAR.....	45
2.4.3	Análise multivariada de imagem aplicada ao QSAR.....	46
2.4.3.1	Sistematização da metodologia qsar: critérios e regras.....	52
3	OBJETIVO GERAL.....	55
3.1	Objetivos específicos.....	55
4	METODOLOGIA.....	56
4.1	Construção do modelo MIA-QSAR.....	57
4.2	Construção do modelo QSAR-tradicional.....	67
4.3	Métodos de validação e construção dos modelos QSAR....	68
4.4	<i>Docking</i>	71
4.5	Avaliações da biodisponibilidade.....	72
5	RESULTADOS E DISCUSSÃO.....	74
5.1	MIA-QSAR.....	74
5.2	QSAR-tradicional.....	78
5.3	Proposta de novos inibidores para ECA.....	90
5.3.1	Inserção de um novo conjunto de dados.....	92
5.3.2	Ancoramento molecular (<i>Docking</i>).....	100
5.3.3	Avaliações da biodisponibilidade.....	103
6	CONCLUSÃO.....	107
	REFERÊNCIAS.....	108

1 INTRODUÇÃO

A pressão arterial elevada constitui um dos principais problemas de saúde pública, sendo as doenças cardiovasculares as principais causas de morte no mundo, segundo a Organização Mundial da Saúde – OMS (2012). No Brasil, a incidência de casos desta doença é alarmante.

Embora não exista cura disponível para hipertensão, é possível fazer um controle eficaz, por meio de mudanças saudáveis no estilo de vida e pelo uso de anti-hipertensivos. Os inibidores da enzima conversora de angiotensina (ECA) estão entre as classes mais amplamente utilizadas para esse fim (ERIKSSON; DANILCZYK; PENNINGER, 2002). Esses compostos inibem a atividade da ECA, uma enzima responsável pela conversão de angiotensina I em angiotensina II, um potente vasoconstritor. O ácido (2S)-1-[(2S)-2-metil-3-sulfonilpropanil] pirrolidino 2-carboxílico (captopril) foi o primeiro inibidor da ECA desenvolvido e ainda é amplamente comercializado. Contudo, o uso dessa classe de medicamento acarreta vários efeitos adversos, tendo em vista que o sistema para a regulação da pressão arterial é muito mais complexo do que inicialmente se pensava.

Os medicamentos existentes que atuam sobre a ECA exibem problemas farmacocinéticos que podem ser superados, o que oferece uma oportunidade para busca de novos compostos, que minimizem os efeitos colaterais e exerçam maior potência de inibição. Contudo, tradicionalmente, o complexo desenvolvimento de uma nova droga exige altos investimentos e oferece poucas garantias de êxito. Entretanto, os avanços ocorridos nas áreas tecnológicas, biológicas e químicas descrevem, atualmente, um percurso mais seguro para propostas de novas entidades químicas, com baixo custo e em menor tempo, acelerando o processo.

Alguns dipeptídeos modificados podem apresentar propriedades inibitórias para a ECA e têm sido desenvolvidos para aumentar a potência relativa, como, por exemplo, em relação ao captopril, bem como para melhorar a farmacocinética e reduzir os efeitos colaterais. Uma série de estudos tem sido realizada em busca de novos dipeptídeos modificados como potenciais inibidores da ECA (BHUYAN; MUGESH, 2011, 2012; CUSHMAN et al., 1977; GRUENFELD et al., 1983; OLEG et al., 2004). Nesse sentido, a estratégia de se combinarem subestruturas químicas mostra-se uma aproximação interessante para gerar novos protótipos de compostos bioativos (FREITAS, 2006; GOODARZI et al. 2010). Portanto, estudos da relação quantitativa entre estrutura e atividade (QSAR) podem ser utilizados para construir modelos a partir dessas classes de dipeptídeos modificados e a atividade de novos inibidores da ECA (resultante da combinação de subestruturas químicas) pode ser prevista usando o modelo QSAR obtido. Além disso, estudos de *docking* (ancoragem molecular), quando efetuados, podem validar e contribuir para a interpretação dos dados bioquímicos, enquanto estudos preliminares de parâmetros ADME (absorção, distribuição, metabolismo e excreção) podem ser realizados para avaliar o perfil farmacológico dos compostos propostos.

O presente trabalho foi realizado com o objetivo de elaborar novas estruturas químicas potencialmente ativas como inibidoras da enzima conversora de angiotensina (anti-hipertensivos) por meio de técnicas computacionais, quais sejam, QSAR e ancoragem molecular, além de realizar avaliação preliminar sobre o perfil farmacocinético das estruturas propostas.

2 REFERENCIAL TEÓRICO

2.1 Hipertensão arterial: situação global

Por muitos anos, as atenções estiveram voltadas, principalmente, para as doenças sexualmente transmissíveis, em razão da grande ocorrência de óbitos, da falta de informações e do reduzido número de tratamentos disponíveis. Dados referentes a 2010, da Organização Mundial de Saúde (OMS, 2010), demonstram que aproximadamente 30 milhões de pessoas morreram infectadas pelo vírus da imunodeficiência humana (HIV, da sigla em inglês para *human immunodeficiency virus*), desde o início da epidemia, por volta de 1981.

Entretanto, as grandes vilãs, atualmente, são as doenças não transmissíveis (doenças cardíacas, câncer, diabetes e doenças respiratórias), responsáveis por quase dois terços de todas as mortes no mundo (WHO, 2012). Dentre essas, destacam-se as doenças cardiovasculares, silenciosas e com múltiplas causas, como, por exemplo, tabagismo, diabetes, excesso de peso, sedentarismo e ingestão excessiva de sal ou álcool. Esses fatores de risco contribuem para o surgimento da hipertensão arterial, que se caracteriza pelo aumento mantido nos valores da pressão sanguínea (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, 2010).

Os níveis pressóricos classificados pela OMS determinam que o indivíduo é considerado hipertenso quando a pressão arterial sistólica PAS \geq 140 mmHg e/ou a pressão arterial diastólica PAD \geq 90 mmHg. Esses valores são classificados de acordo com a medida casual no consultório e são apresentados na Tabela 1.

O III Consenso Brasileiro de Hipertensão Arterial, publicado em agosto de 1999, chama a atenção para o diagnóstico de pacientes hipertensos; a confirmação de um quadro positivo deve ser feita com bastante cautela, pois a

aferição da pressão arterial envolve uma série de procedimentos. Classificar um paciente como hipertenso exige múltiplas repetições da aferição da pressão arterial, em diferentes horários, condições alimentares e com o corpo em diferentes posições verticais e horizontais (KOHLMANN, 1999).

Tabela 1 Classificação da pressão arterial de acordo com a medida casual no consultório (>18 anos).

Classificação	PAS (mmHg)	PAD (mmHg)
Ótima	< 120	< 80
Normal	< 130	< 85
Limítrofe	130-139	85-89
Hipertensão estágio 1	140-159	90-99
Hipertensão estágio 2	160-179	100-109
Hipertensão estágio 3	≥ 180	≥ 110

*Quando a pressão sistólica e a diastólica situam-se em categorias diferentes, a maior deve ser utilizada para classificação da pressão arterial.

Fonte: adaptado das DBH VI (SOCIEDADE BRASILEIRA DE CARDIOLOGIA, SBC, 2010).

Qualquer indivíduo, em qualquer época de sua vida, inclusive na infância, está sujeito à hipertensão arterial, situação esta que pode levar a problemas cardiovasculares futuros, como infarto agudo do miocárdio e acidente vascular do tipo cerebral. Com isso, é fundamental detectar precocemente alterações, tratar e controlar os níveis pressóricos.

A grande preocupação que gira em torno das doenças cardiovasculares não é por acaso. A OMS afirma que, anualmente, morrem mais pessoas vítimas de problemas cardíacos do que de qualquer outra causa. Estima-se que cerca de 17,3 milhões de pessoas morreram em 2008, 80% dessas mortes ocorreram em

países de baixa e média renda e, provavelmente, quase 23,6 milhões de pessoas morrerão de doenças do coração até 2030. Isso representa 57 milhões de anos de vida perdidos. Representa, ainda, em nível macroeconômico, uma redução do Produto Interno Bruto (PIB) entre 1% a 5%, porque atinge os indivíduos em seus anos mais produtivos, o que reflete no crescimento econômico de países em desenvolvimento.

Esses números podem ser modificados levando em consideração os dados publicados pela OMS em documento denominado “Saúde mundial: estatísticas 2012” (WORLD HEALTH STATISTICS, WHS, 2012). Os estudos demonstraram que, em países de renda per capita maior, mesmo tendo índices de obesidade elevados – um alto fator de risco para o desenvolvimento dessas doenças –, os casos de mortes por doenças do coração diminuíram, e em países subdesenvolvidos isso não ocorreu. Essas diferenças refletem na forma como é tratada a hipertensão arterial. Em países como os Estados Unidos realiza-se um eficiente programa de prevenção, tratamento com diagnóstico precoce e uso de medicamentos; 29,9% dos americanos, em 2008, apresentaram hipertensão arterial. Em Moçambique, no entanto, em torno de 50,9% da população tem pressão sanguínea elevada.

No Brasil essa situação é preocupante, pois cerca de 42,3% dos brasileiros têm hipertensão arterial. Em 2005, o custo com o tratamento representou 0,08% do PIB brasileiro, totalizando aproximadamente R\$1.631.878.386,00, no Sistema Único de Saúde (SUS) e no Sistema Suplementar de Saúde (SSS) (DATASUS; 2005). As internações registradas pelo SUS, em 2007, foram de 1.157.509 e, apenas em novembro de 2009, houve 91.970 internações, resultando em um custo de R\$165.461.644,33.

Outros índices podem ser verificados no Gráfico 1, os quais se referem às taxas de prevalência da pressão arterial elevada em ambos os sexos, em

pessoas acima de 25 anos, com níveis pressóricos de PAS \geq 140 mmHg e/ou PAD \geq 90 mmHg, de acordo com a região de atuação da OMS.

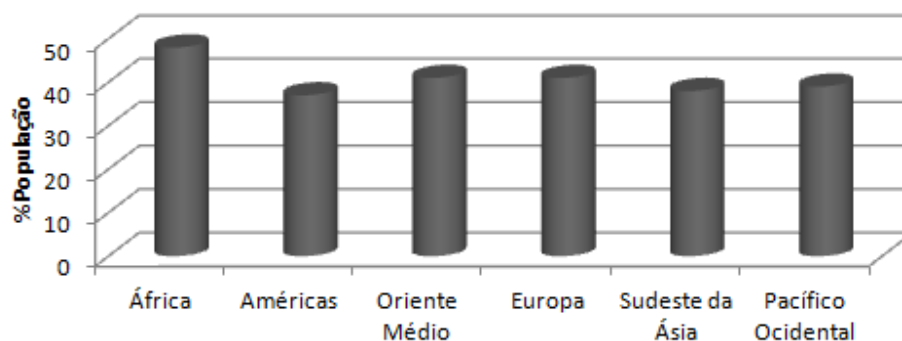


Gráfico 1 Ocorrência de casos de hipertensão arterial nas regiões de atuação da OMS

Fonte: OMS (2012)

A necessidade da implementação de um programa eficaz, que visa à prevenção, o diagnóstico precoce e um tratamento efetivo, requer certa urgência, não somente no Brasil, mas na maioria dos países em desenvolvimento. A consequência dessa ação está diretamente ligada à diminuição dos gastos públicos, à demanda em centros de saúde (consultas, cirurgias, internações, entre outros) e à taxa de mortalidade, causa principal.

Por ser uma doença sem sintomas e de cura não disponível, é preciso evitá-la. A reformulação de hábitos de vida se faz necessária. Uma alimentação saudável e a manutenção do peso ideal, o consumo controlado de sódio e álcool, a ingestão de potássio, o combate ao sedentarismo e ao tabagismo e a administração do estresse são meios de reduzir os riscos (SBC, 2010).

Uma vez diagnosticada a elevação da pressão arterial (Tabela 1), um controle mais eficaz é imprescindível. Isso é possível com o uso de uma classe de medicamentos, denominados anti-hipertensivos, responsáveis por manter os níveis pressóricos dentro do padrão normal ou reduzir a pressão arterial a um nível aceitável.

2.1.1 Anti-hipertensivos

Uma ampla classe de medicamentos destinados ao controle da pressão arterial é comercializada, atualmente, no Brasil (Tabela 2). Esses fármacos que constituem a opção terapêutica são classificados de acordo com o local de ação.

Tabela 2 Anti-hipertensivos comercialmente disponíveis no Brasil.

Diuréticos			
Tiazídicos	Alça	Poupadores de potássio	
Clortalidona	Bumetamida	Amilorida	
Hidroclorotiazida	Furosemida	Espironolactona	
Indapamida	Piretanida	Triantereno	
		Amilorida	
Inibidores adrenérgicos			
Ação central	Betabloqueadores		Alfabloqueadores
Alfametildopa	Atenolol	Nadolol	Doxazosina
Clonidina	Bisoprolol	Nebivolol	Prazosina
Guanabenzol	Carvedilol	Propranolol	Prazosina
Moxonidina	Metoprolol	Pindolol	Terazosina
Rilmenidina			
Reserpina			

...continua.

Tabela 2, continuação

Bloqueadores dos canais de cálcio		
Fenilalquilaminas	Di-idropiridinas	Benzodiazepinas
Verapamil Retard	Anlodipino	Manidipino
	Felodipino	Nifedipino Oros
	Isradipina	Nifedipino Retard
	Lacidipina	Nisoldipino
	Lercarnidipino	Nitrendipino
Vasodilatadores diretos		Inibidor direto da renina
Hidralazina		Alisquireno
Minoxidil		
IECAs (inibidores da enzima conversora de angiotensina)		
Benazepril	Captopril	Cilazapril
Delapril	Enalapril	Fosinopril
Lisinopril	Perindopril	Quinapril
Ramipril	Trandolapril	Ceranapril
ARAs (antagonistas dos receptores de angiotensina II)		
Candesartana	Irbersartana	Losartana
Olmesartana	Telmisartana	Valsartana

Fonte: adaptado SBC (2010).

Os diuréticos receberam esse nome devido à sua função de promover a eliminação de líquidos (contendo sódio e cloro), por meio do aumento do fluxo urinário. Existem três tipos (PIMENTA, 2008), que são:

- a) tiazídicos: indicados para o controle da pressão arterial;
- b) alça: apropriado para situações de emergência, como hipertensão associada à insuficiência renal;
- c) poupadores de potássio: úteis na prevenção e no tratamento de hipopotassemia¹, quando administrado em conjunto com outros diuréticos.

¹ Baixos níveis de potássio no sangue.

Além da hipopotassemia, o uso de diuréticos pode provocar alguns efeitos adversos, como a intolerância à glicose, o risco de desenvolver diabetes e o aumento de triglicérides.

Por meio da estimulação dos receptores alfa-2 adrenérgicos pré-sinápticos no sistema nervoso central, os inibidores adrenérgicos de ação central reduzem o tônus sináptico. Os efeitos são discretos e, por isso, sua eficiência aumenta quando associados a outros anti-hipertensivos. O mesmo ocorre com os alfabloqueadores. Em contrapartida, os betabloqueadores são mais promissores para o tratamento da hipertensão arterial, atuam diminuindo o débito cardíaco e catecolaminas nas sinapses nervosas, reduzem a secreção de renina e auxiliam na readaptação dos barorreceptores. Entre os efeitos adversos decorrentes do uso dos inibidores adrenérgicos, destacam-se a disfunção sexual, a hipotensão postural² e a boca seca (TAVARES; PLAVNIK, 1998).

Os três subgrupos dos bloqueadores dos canais de cálcio têm ação anti-hipertensiva por meio da redução da resistência vascular periférica, resultante da diminuição da concentração de cálcio nas células musculares lisas. Cefaleia³, tontura e rubor facial são as reações mais comuns (POULTER, 2011).

O tratamento da pressão alta (nome comum para hipertensão arterial) consiste no uso de combinações de classes medicamentosas. Raramente, o paciente diagnosticado com esse problema fará uso de apenas um tipo de medicamento. A utilização de vasodilatadores diretos promove relaxamento muscular com conseqüente vasodilatação e eles podem ser administrados com diuréticos (PIMENTA, 2008) e/ou betabloqueadores; por terem funções distintas, essas associações aumentam a eficácia do tratamento. Diversas formas de tratamento estão disponíveis; este é apenas um dos muitos exemplos. A

² Queda súbita de pressão sanguínea quando um indivíduo assume a posição ereta.

³ Dor de cabeça.

indicação do uso de uma classe de anti-hipertensivos ou outra depende da relação entre causa e aumento dos níveis pressóricos.

Uma ampla gama de remédios é comercializada para esse fim. Um grupo de medicamentos importantes para o tratamento é responsável por inibir a enzima que atua na regulação da pressão arterial, os quais se classificam como: inibidores diretos da renina, IECAs e ARAs. Devida à sua importância e por serem tema central do presente trabalho, serão discutidos nos próximos tópicos.

2.1.2 Inibidores diretos da renina, IECAs e ARAs

Os inibidores diretos da renina não compõem uma subclasse de anti-hipertensivos responsáveis por inibir a angiotensina (enzima responsável por regular a pressão arterial), mas sua ação se relaciona com essa via de regulação (Figura 3). O único representante atual dessa classe, o alisquireno, promove inibição direta da ação da renina, com conseqüente diminuição da angiotensina II (GOESSLER; POLITO, 2012; RIBEIRO, 2006), entre outras ações ainda não confirmadas por estudos confiáveis, não cabendo, aqui, serem citadas.

Os IECAs são agentes anti-hipertensivos que inibem a enzima conversora de angiotensina que converte a angiotensina I em angiotensina II. Ao inibir essa enzima, os IECAs produzem vasodilatação periférica, diminuindo a pressão arterial. Com eficácia no tratamento, existem relatos de diminuição da morbidade e da mortalidade cardiovascular, associada ao uso dessa classe medicamentosa (PAUL; MEHR; KREUTZ, 2006).

Vários efeitos adversos podem surgir, como dor de cabeça, fadiga, problemas renais e tosse seca. Alguns dos efeitos não desejados dos IECAs, como a tosse, devem-se ao acúmulo de bradicininas no sangue, as quais estimulariam a via respiratória a ficar hipertensiva (BAKRIS, 2009).

Esse grupo de anti-hipertensivos pode ser separado em subgrupos, quando baseados na estrutura química de cada fármaco (PAULA et al., 2005). O captopril é o único que contém um grupo sulfidrilo; fosinopril e ceranapril apresentam um átomo de fósforo. Benazepril, enalapril, lisinopril e quinapril, entre outros citados na Tabela 2, constituem a maior classe com um grupo carboxila. Basicamente, existe, em todos os compostos pertencentes a essa classe de anti-hipertensivos, certa similaridade estrutural [o grupo ácido (*S*)-pirrolidino-2-carboxílico], como observado na Figura 1, que ilustra a estrutura química de um composto de cada subgrupo.

Algumas peculiaridades estão relacionadas à estrutura química do captopril. Destacam-se a simplicidade estrutural (poucos átomos, não contém anéis benzênicos e cadeias laterais longas) e a presença de um grupo tiólico, capaz de fazer ligações de hidrogênio (pode ter relação na ação e na biodisponibilidade). Esse é o único inibidor da ECA comercializado nos Estados Unidos da América que contém um grupo tiólico (NATESH et al., 2004).

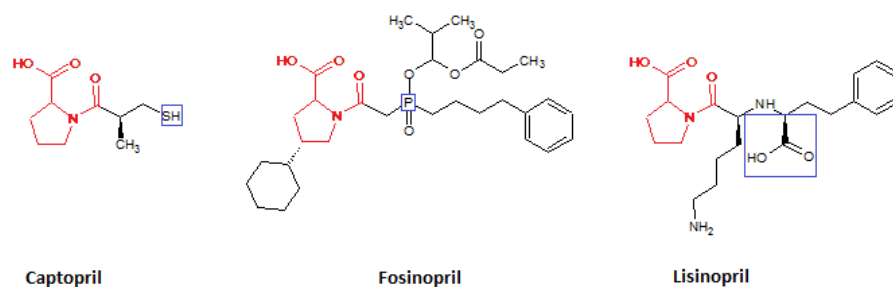


Figura 1 Estrutura química da classe de compostos pertencentes ao grupo denominado IECAs. As diferenças e as similaridades encontram-se destacadas na cor azul e vermelha, respectivamente. O captopril é o fármaco de referência neste estudo

A descoberta desse fármaco (CUSHMAN et al., 1977) representou, na época, uma importante inovação terapêutica para o tratamento da hipertensão, por meio de estudos farmacológicos sobre os efeitos do veneno de serpente brasileira da espécie *Bothrops jararaca*. Os princípios ativos, de natureza peptídica, foram isolados, sintetizados e testados, destacando-se como mais ativo o nonapeptídeo, denominado teprotide, segundo resultado de pesquisas do grupo brasileiro Maurício R. e Silva (SILVA; BERALDO; ROSENFELD, 1949). Posteriormente, na Faculdade de Medicina da Universidade de São Paulo, em Ribeirão Preto, por Sérgio H. Ferreira. Sir John Vane beneficiou-se desses estudos, propondo modificações estruturais e obteve os primeiros anti-hipertensivos naturais utilizados como modelo pela indústria farmacêutica Squibb, para a síntese do captopril (LIMA, 2007).

O exato mecanismo de ação do fármaco ainda não foi completamente elucidado, mas sabe-se que o uso regular desse medicamento resulta na diminuição de concentrações de angiotensina II (uma das mais potentes substâncias vasoconstritoras endógenas) e aldosterona (hormônio responsável pela regulação do balanço eletrolítico), tendo efeitos benéficos na hipertensão e na insuficiência cardíaca.

A eficiência do captopril relaciona-se à função que desempenha, tendo em vista que pequenas alterações no sistema renina-angiotensina tornam-se suficientes para desencadear respostas máximas de produção de aldosterona. Assim, a redução da angiotensina II leva a uma secreção diminuída de aldosterona e, como resultados, podem ocorrer pequenos aumentos de potássio sérico⁴, juntamente com perda de sódio e fluidos, com conseqüente regulação da pressão arterial (NATESH et al., 2004).

⁴ Nível sérico é um termo utilizado por profissionais de saúde para se referirem à quantidade de uma determinada substância no sangue.

Ainda é esperado outro efeito. Como a enzima bradicininase é idêntica à ECA, o captopril pode interferir na degradação da bradicinina. Essa ação contribui para a eficácia do tratamento, devido ao aumento das concentrações de bradicinina ou prostaglandina E2 (substâncias vasodilatadoras endógenas) (SILVA; BERALDO; ROSENFELD, 1949; PAUL; MEHR; KREUTZ, 2006; NATESH et al., 2004).

Dos compostos listados na Tabela 2 como pertencentes ao grupo dos IECAs, apenas o captopril (conforme discorrido) e o lisinopril (possui funções carboxílicas e não requer hidrólise prévia) são drogas ativas (FERNANDEZ et al., 2003), ou seja, dispensam metabolização, o que aumenta a ação desses dois fármacos.

Outra via de regulação da pressão arterial é obtida com o uso dos ARAs, responsáveis por ligarem-se à Angiotensina II, com conseqüente diminuição da vasoconstrição das arteríolas promovida por essa proteína. Obtém-se eficácia maior no tratamento com a combinação de grupos de medicamentos de classes diferentes (SBC, 2010; POULTER, 2011).

2.2 Enzima conversora de angiotensina (ECA)

Localizada na membrana plasmática como um terminal transmembrânico, foi denominada enzima conversora de angiotensina (ACE; EC 3.4.15.1), quando isolada pela primeira vez, em 1956 (SKEGGS; KAHN; SHUMWAY, 1956). Desempenha duas atividades principais, que são:

- a) age como uma peptidildipeptidase que remove os dipeptídeos carboxi-terminais de seu substrato;
- b) age como endopeptidase em substratos que são amidas na carboxila terminal (TSCHOPE; SCHULTHEISS; WALTHER, 2002);

Trata-se, portanto, de uma enzima chave no processo de regulação da pressão arterial (S-RAA: sistema renina-angiotensina-aldosterona), por apresentar especificidade pela angiotensina I. Classifica-se como metaloproteína, da classe das carboxipeptidases, pesando, aproximadamente, 130 KDa (BARREIRO; FRAGA, 2008). Os braços homólogos (carboxi-terminal e amino-terminal) encontram-se ligados por um íon zinco, essencial à atividade catalítica. Portanto, os inibidores precisam competir com o substrato natural pelo zinco (II) no sítio ativo, bem como por meio de ligações de hidrogênio e interações hidrofóbicas (PAULA et al., 2005).

A estrutura quaternária tridimensional e alguns de seus respectivos inibidores estão representados na Figura 2, disponível no banco de dados de proteínas *Protein Data Bank* ou PDB.

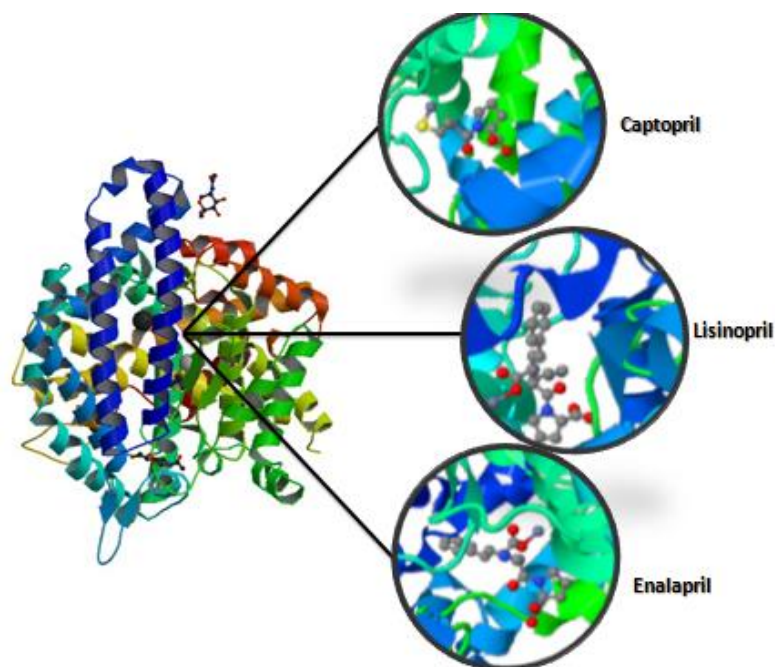


Figura 2 Estrutura tridimensional da ECA, extraída do PDB, com os respectivos códigos: captopril (1UZF), lisinopril (1086) e enalapril (1UZE).

2.2.1 O sistema renina-angiotensina-aldosterona (S-RAA)

O S-RAA é um complexo processo de regulação (Figura 3) que desempenha papel fundamental na manutenção da homeostase da pressão arterial, por meio do relaxamento (vasodilatação) e da contração (vasoconstrição) dos vasos sanguíneos, assim como no equilíbrio de líquidos e sais em mamíferos. Esse mecanismo de regulação tem início com a liberação de uma enzima proteolítica de origem renal, denominada renina (**a**), na corrente sanguínea. A renina age sobre o angiotensinogênio (**b**), uma proteína de origem hepática inativa que se converte em um decapeptídeo (angiotensina I), que migra pela circulação. Uma reação de hidrólise catalisada pela ECA ocorre quando a angiotensina I passa pelos vasos pulmonares (**c**), transformando-se em um octapeptídeo (angiotensina II) capaz de provocar intensa vasoconstrição. Via corrente sanguínea, a angiotensina II é transportada para os órgãos alvos: os rins (**d**) e o córtex das glândulas suprarrenais (**e**). Nos rins, estimula a reabsorção de sódio e água, e no córtex das glândulas suprarrenais libera o hormônio aldosterona, que também se dirige aos rins, aumentando ainda mais a absorção de sódio e água, e excreção de potássio. Ambos os caminhos resultam em aumento da quantidade de sangue circulando no corpo, com consequente elevação da pressão arterial (TSCHOPE; SCHULTHEISS; WALTHER, 2002, PAUL; MEHR; KREUTZ, 2006).

A etapa principal do S-RAA que contribui para a desregulação da pressão arterial ocorre com a formação da angiotensina II, responsável por ativar negativamente um ciclo que sobrecarrega os órgãos envolvidos na estabilidade hemodinâmica. Após realizar sua função, o processo finaliza-se quando a angiotensina II é inativada. Isso ocorre por meio da ação de enzimas angiotensinas e, como resultado, forma-se uma nova proteína: a angiotensina III (ERIKSSON; DANILCZYK; PENNINGER, 2002).

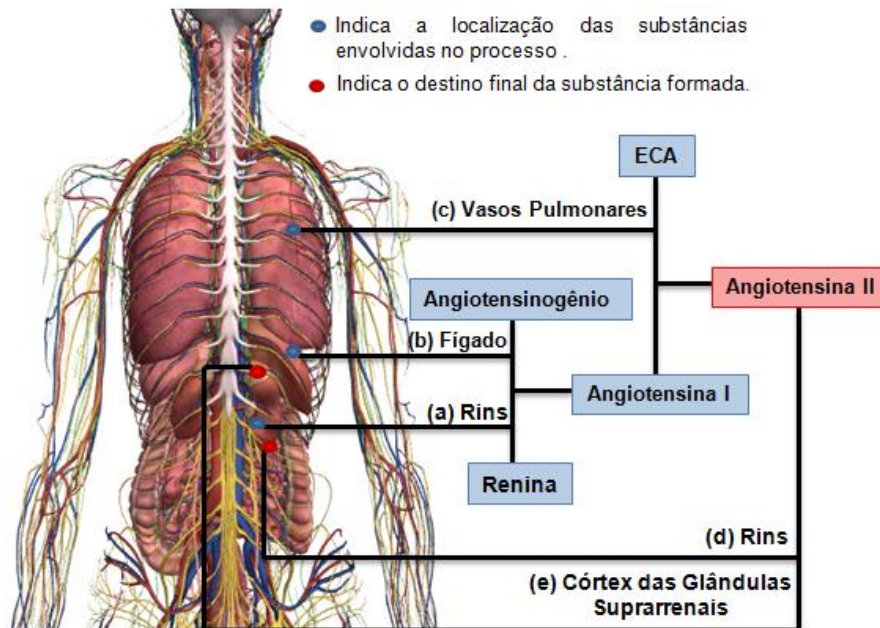


Figura 3 Representação do sistema renina-angiotensina
(Fonte: Google Body Browser (2012)).

As reações que levam ao desenvolvimento da hipertensão arterial iniciam-se com a conversão do angiotensinogênio, uma proteína de 411 aminoácidos em outra de 10 aminoácidos. Em seguida, ocorre uma clivagem de dois aminoácidos, histidina (His) e leucina (Leu), formando uma substância de oito aminoácidos. A fase final acontece com a clivagem de aspartato (Asp), de acordo com a Figura 4.

Como já foi discutido, observa-se, nas Figuras 3 e 4, que existem duas enzimas envolvidas na ação de desregulação da pressão arterial, a renina e a ECA, ficando evidente que, para interromper esse processo, basta inativar ou regular o funcionamento dessas enzimas. Assim, caso existisse uma nova entidade química capaz de interagir com o sítio ativo de ambas as enzimas,

ocorreria uma regulação mais eficaz do S-RAA. Até o momento, nenhum dos fármacos comercializados desempenha essa ação.

A descoberta de drogas específicas que regulam o S-RAA representou um grande passo à frente no tratamento da hipertensão arterial e da insuficiência cardíaca. Nos últimos anos, entretanto, o sistema renina-angiotensina tem se mostrado um sistema muito mais complexo do que inicialmente se pensava. Isso não só amplia o conhecimento sobre a evolução e a biologia de regulação da pressão arterial, mas também oferece uma oportunidade para desenvolver novas estratégias terapêuticas.

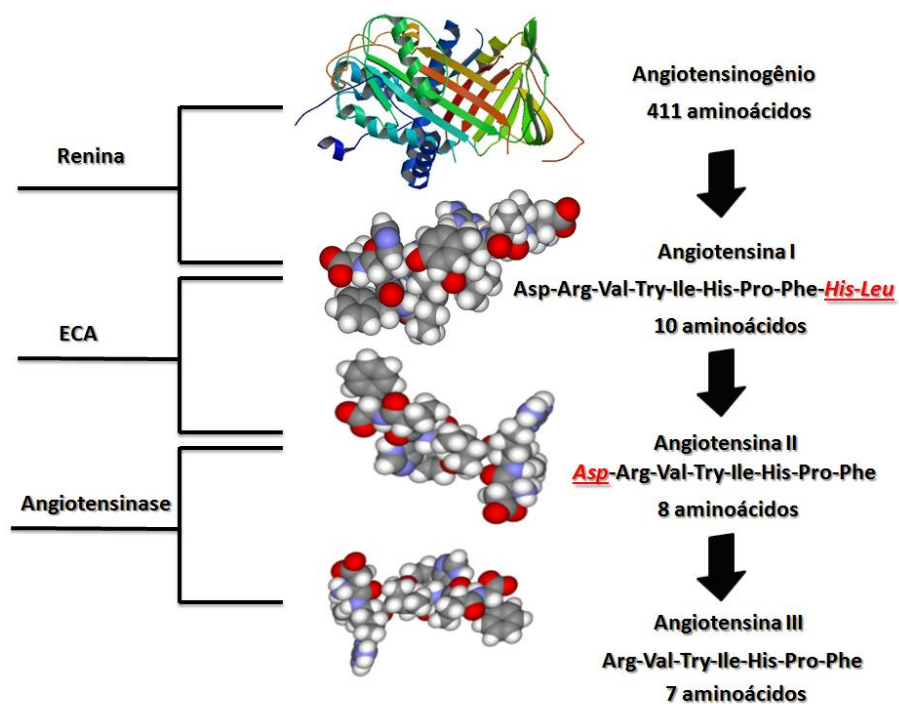


Figura 4 Processo de clivagem dos aminoácidos (destacado em vermelho) realizado pelas enzimas renina, ECA e angiotensinase
Fonte: adaptado de Eriksson, Danilczyk e Penninger (2002)

2.3 O desenvolvimento de novos fármacos

Pesquisas realizadas pela *The Pharmaceutical Research and Manufacturers of America*, PhRMA, apontam que o custo médio de desenvolvimento de uma nova droga passou de US\$ 138 milhões de dólares, em 1975, para US\$ 802 milhões de dólares, atualmente, nos Estados Unidos. Estudos realizados por diversos pesquisadores da Universidade de Tufts deixam claro que esse custo elevado se deve ao grande número de falhas no início do processo; apenas um em cada 5.000 selecionados a candidatos à nova droga chega ao mercado como medicamento. Mesmo assim, isso não garante lucro, uma vez que apenas três dos dez medicamentos comercializados produzem receitas que correspondem ou excedem os custos de desenvolvimento. Sabe-se que leva cerca de dez a quinze anos para um novo medicamento ser aprovado e comercializado (REES, 2003).

As etapas que caracterizam o processo de descobrimento de novas entidades químicas abrangem uma gama de fatores que contribuem, muitas vezes, para o insucesso da busca por novos medicamentos. A cargo da Química Medicinal encontra-se esse trabalho laborioso, que consiste em elaborar novas drogas ou aprimorar as existentes, por meio de modificações estruturais, com o objetivo de gerar compostos capazes de atingir um determinado alvo terapêutico, sem causar danos à saúde. Assim, é necessário que uma nova droga exiba biodisponibilidade, metabolismo adequado e baixa toxicidade, assegurando sua administração (BARREIRO, 2002).

No desenvolvimento de fármacos, estima-se que cerca de 79% dos protótipos sejam eliminados nas etapas de avaliações clínicas (DIMASI; HANSEN; GRABOWSKI, 2003). Segundo Waterbeemd e Gifford (2003), em torno de 39% dos insucessos encontrados durante o desenvolvimento de um novo fármaco se devem ao fato de os candidatos não apresentarem as

propriedades farmacocinéticas requeridas. A perda de eficácia (30%), outro fator importante, provém de falhas inesperadas de um medicamento em produzir o efeito planejado, determinado por investigação científica prévia. Outros fatores, de menor efeito, juntam-se aos citados, atribuindo a esse processo um alto risco, conforme apresentado na Figura 5.

As modificações e o amadurecimento na indústria farmacêutica têm minimizado os erros apresentados na Figura 5, acelerando todo o processo de elaboração de novos medicamentos, mas a porcentagem de sucesso ainda é pequena. Contudo, quando se comparam as estratégias empregadas na década de 1950 com o modelo atual, percebe-se uma evolução considerável.

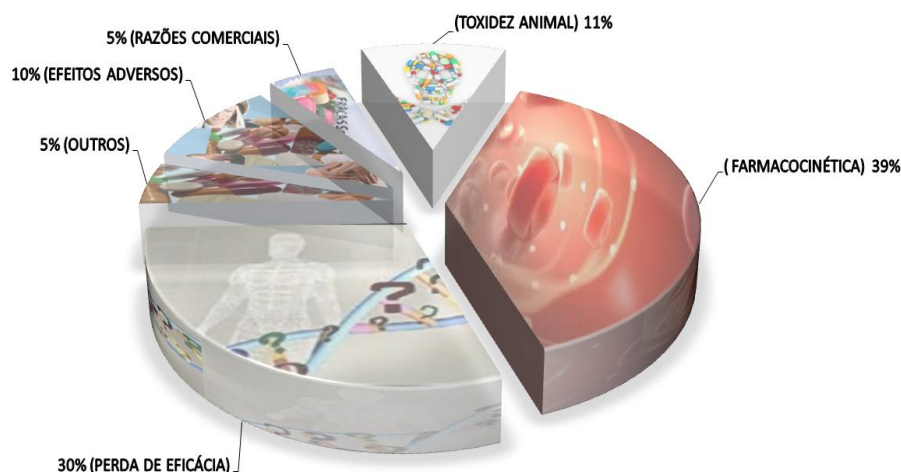


Figura 5 Principais razões de insucesso de novas entidades químicas no processo de desenvolvimento
Fonte: adaptado de Waterbeemd e Gifford (2003).

Ainda naquela década, e até mesmo nos dias atuais, o acaso contribuiu de forma significativa para o descobrimento de novas drogas. Isto é ilustrado por exemplos como o Viagra, concebido como uma droga para o tratamento cardíaco, que está rendendo bilhões, sendo vendido para a disfunção erétil. A talidomida, desenvolvida como um sedativo, causou terríveis problemas teratogênicos, quando utilizada por gestantes visando à diminuição de enjoos e, hoje, é uma promissora droga no tratamento da hanseníase. Dietilestilbestrol, desenvolvido para impedir aborto em mulheres grávidas, apresentou-se como precursor no desenvolvimento de câncer nos recém-nascidos e, atualmente, é utilizado como antineoplásico. Finalmente, sabemos que uma das melhores drogas antitumorais do século XX (cisplatina) foi desenvolvida por um físico estudando o crescimento de bactérias em um campo elétrico (KUBINYI, 2003).

Nos últimos 40 anos, diversas técnicas foram criadas para auxiliar no desenvolvimento de novas drogas e ajudar a simplificar a busca por novos alvos terapêuticos. Com essas técnicas tornou-se possível racionalizar a criação de novos fármacos, antes descobertos, basicamente, ao acaso.

A finalização da elaboração de uma droga acontece quando se cumprem todas as etapas de um dispendioso sistema (Figura 6), criado com o objetivo de assegurar que a utilização do novo medicamento exerça seu papel. O desafio da Química Medicinal conta com o apoio dos avanços científicos e tecnológicos observados em diversas áreas, como a Química Orgânica e Computacional, a Biologia Estrutural e Molecular, e, ainda, a Farmacologia. Assim, uma interdisciplinaridade entre essas áreas é imprescindível na realização dessa árdua tarefa (BARREIRO, 2002).

O complexo desenvolvimento e a descoberta de uma nova droga, ilustrados na Figura 6, exigem longos períodos e envolve altos investimentos (BERNDT et al., 2005). Inicialmente, escolhe-se um alvo terapêutico. Nesta etapa, estudos computacionais auxiliam na diminuição do tempo, do

investimento e de sínteses exploratórias desnecessárias. Os compostos promissores são selecionados para síntese. Uma avaliação toxicológica básica é realizada, por meio de estudos *in silico* e *in vivo* (em animais), para posterior tentativa de produção em escala industrial, entre outros testes. Caso o novo composto desempenhe a função desejada, um pedido de proteção de propriedade intelectual é encaminhado aos órgãos responsáveis e realiza-se uma pesquisa de mercado farmacêutico. Estas etapas compreendem a fase pré-clínica (MODA, 2007).

Ao vencer as etapas iniciais, testa-se o novo fármaco em seres humanos, verifica-se a dosagem dos princípios ativos, os níveis de toxidez, a segurança e a eficácia, incluindo monitoramento favorável das propriedades farmacocinéticas (ADMET). Concluída a fase clínica, uma solicitação de nova aplicação de fármaco (*new drug application*, ou NDA) é realizada e, após aprovação pelas agências regulatórias (ANVISA, no Brasil; FDA, nos Estados Unidos e EMEA, na Europa) e, eventualmente, o fármaco entrará no mercado (LIMA, 2007).

O sucesso comercial de um fármaco depende de alguns fatores, tais como simplicidade estrutural – com possibilidade de modificações, visando à otimização de seu perfil farmacoterapêutico –, ser membro de uma série congênere com relação estrutura/atividade (SAR) bem estabelecida e apresentar boas propriedades farmacocinéticas (OPREA et al., 2001).

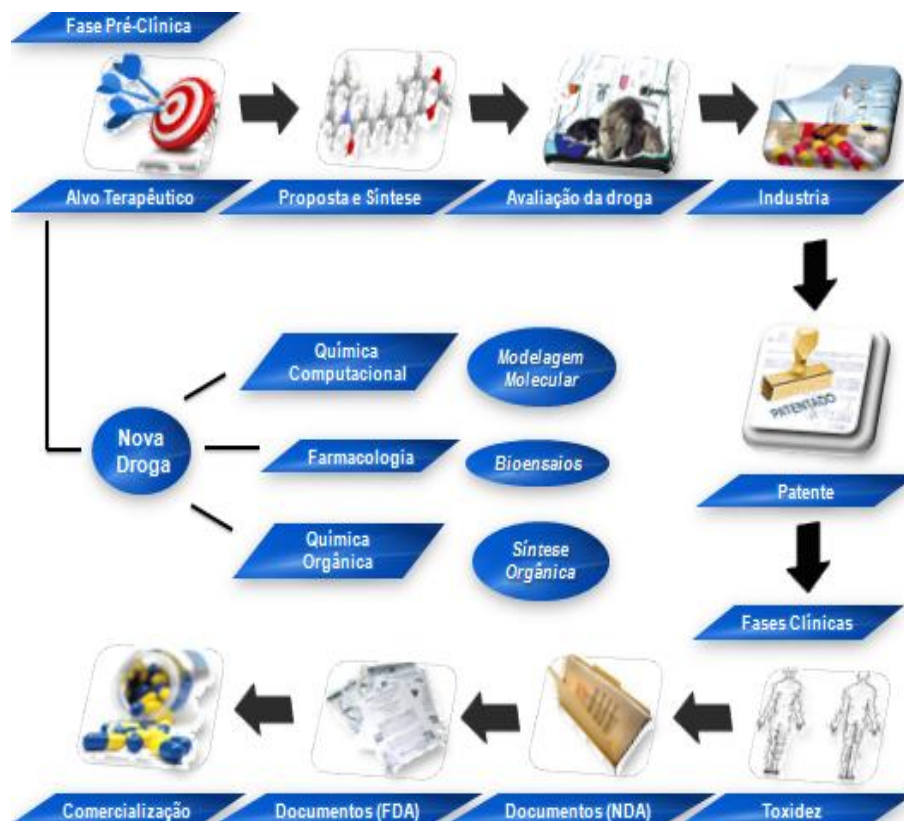


Figura 6 Etapas do processo de descoberta e planejamento racional de fármacos
 Fonte: adaptado de Barreiro (2002); Lombardino e Lowe (2004).

2.4 QSAR, metodologia de auxílio no processo de elaboração de novos fármacos

Ao passo que a tecnologia desenvolveu-se com a criação de computadores, softwares e hardwares, foram criados programas específicos para serem utilizados em problemas químicos. Criou-se, assim, um novo ramo da química para auxiliar no planejamento de fármacos, a Química Computacional. Uma técnica que tem se destacado neste meio é a Relações quantitativas entre

estrutura-atividade, ou QSAR, do inglês *Quantitative structure-activity relationships*.

O método baseia-se na possibilidade de a atividade e/ou propriedade ser uma função da estrutura molecular, ou seja, correlaciona a estrutura com determinada propriedade química, com o objetivo de fundamentar o planejamento de novas substâncias que possuam perfil terapêutico mais adequado às necessidades (NEVES, et al., 1998; TAVARES, 2004).

A área desperta grande interesse, visto que dispensa sínteses experimentais exploratórias de grande número de compostos, pois consegue projetar moléculas com propriedades bem definidas, o que diminui os custos e o tempo (HANSCH et al., 2002; MARTINS et al., 2009; MARTINS, 2010).

A construção de modelos QSAR envolve uma série de procedimentos e regras, adotados na medida em que a metodologia evoluiu (Figura 7). Tratamentos estatísticos, criação de novos descritores e implementação de novos métodos caracterizam esse processo, na tentativa de garantir maior capacidade preditiva do modelo. Contudo, a organização do método ocorreu em 1964 e nos anos seguintes. O trabalho de Hansch e Fujita (1964) representa essa sistematização e, para tanto, foram utilizadas as constantes σ e π para os substituintes, definindo o modelo como

$$\pi = \text{Log } P_x - \text{Log } P_h \quad (1)$$

em que P_h representa o coeficiente de partição do composto de origem não substituído (com hidrogênio) e P_x é um termo correspondente para o derivado substituído por X. Neste trabalho, correlacionaram-se os efeitos dos substituintes com a atividade biológica de ácidos benzoicos.

Mais tarde, ao utilizar compostos derivados N,N-dimetil- α -bromo-feniletilamina, foi demonstrado que os efeitos lipofílico e eletrônico são capazes

de descrever a atividade antiadrenérgica, partindo da estrutura do composto (HANSCH; LIEN, 1968).

Em outro artigo, Gargiulo, Giotta e Wang (1973) contribuíram também na elaboração de regras para a construção de modelos QSAR. O objetivo central, na época, girava em torno da preocupação da utilização equivocada da metodologia, pois a matemática envolvida na construção de modelos é relativamente simples, mas a validação e a interpretação apropriada dos resultados, certamente, exigiam maiores esforços e conhecimentos (GAUDIO; ZANDONADE, 2001).

Na verdade, para cada sistema, há uma ou mais possibilidades de se estabelecerem relações entre estrutura química e atividade biológica. Talvez, o mais importante seja elaborar um modelo da melhor maneira possível, de forma que os resultados sejam obtidos confiavelmente e, se possível, de maneira simples e interpretável.

Os coeficientes de partição octanol/água ($\log P$), a constante de Hammett, o parâmetro de lipofilicidade e os descritores empíricos, entre outros, permaneceram, por anos, como os mais empregados em análises QSAR (MARTINS, 2010).

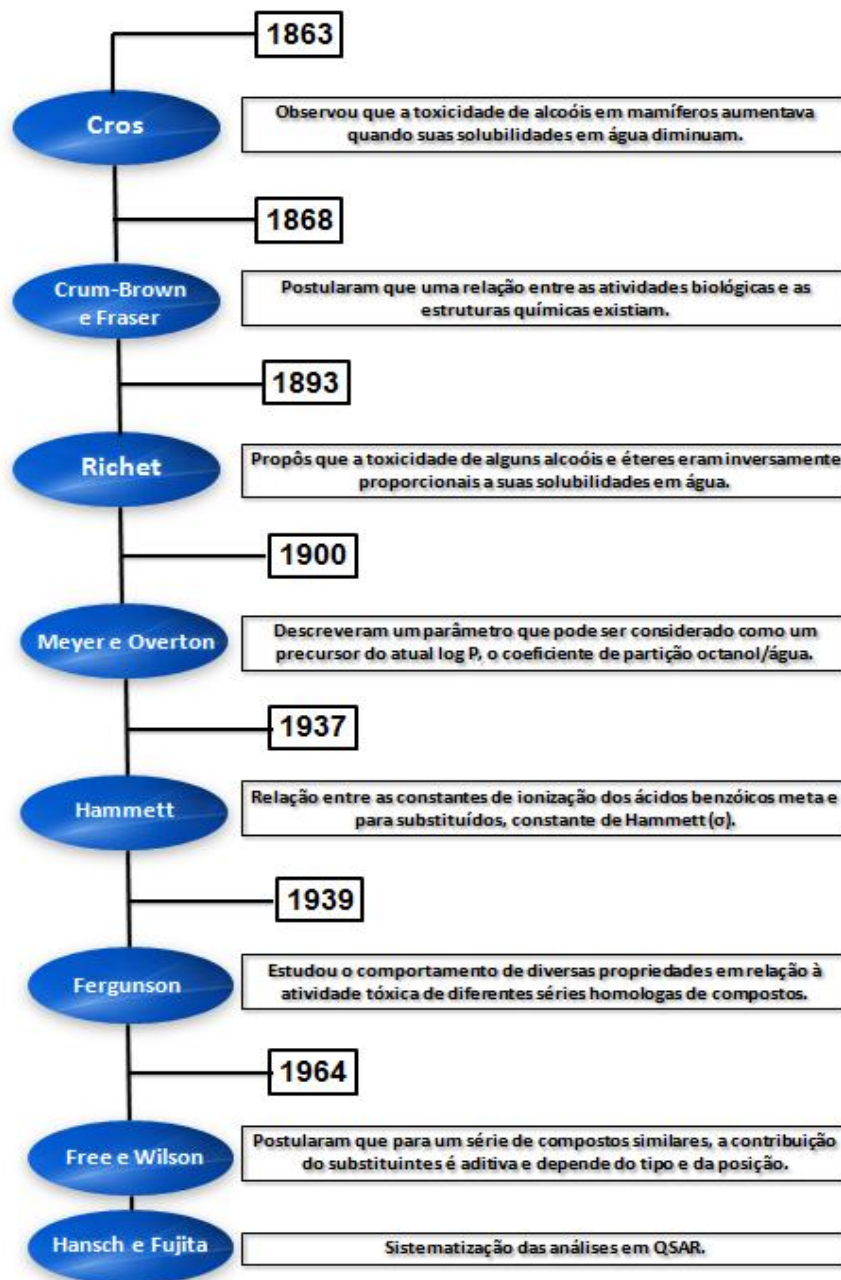


Figura 7 Cronologia da construção da metodologia QSAR
 Fonte: Hammett (1937); Hansch et al. (2002); MARTINS (2010).

2.4.1 QSAR, inserção de parâmetros moleculares tridimensionais

Em 1988, com a inserção de parâmetros moleculares tridimensionais, as técnicas sofreram profunda reestruturação. Uma nova metodologia, denominada análise comparativa do campo molecular, ou CoMFA, sigla do inglês *comparative molecular field analysis*, foi a responsável por essas modificações. O método baseia-se na representação das moléculas em sua forma tridimensional, alinhadas dentro de uma caixa 3D virtual, diferindo entre confôrmeros, estereoisômeros ou enantiômeros, por meio de descritores específicos, atribuindo a terceira dimensão - QSAR-3D (CRAMER; PATTERSON; BUNCE, 1988).

Alguns dos problemas envolvidos no uso de QSAR-3D (identificação das conformações ativa, alinhamento e geração de grande número de descritores) podem ser superados com a metodologia proposta por Hopfinger et al. (1997), que incorpora liberdade conformacional, constituindo a quarta dimensão (QSAR-4D), e ainda subsidia uma equação mais simples para o desenvolvimento de modelos QSAR-3D, pois reduz o número de descritores.

A criação da metodologia QSAR-5D tornou possível a representação do receptor, com a simulação do encaixe induzido, e de uma adaptação topológica do modelo da superfície do sítio de ligação, para moléculas individuais do ligante (VEDANI; DOBLER, 2002).

Uma avaliação simultânea de diferentes modelos de solvatação, desenvolvida por Vedani Dobler e Lill (2005), possibilita uma simulação *in silico* de efeitos adversos provocados pelas drogas e produtos químicos e foi denominada QSAR-6D.

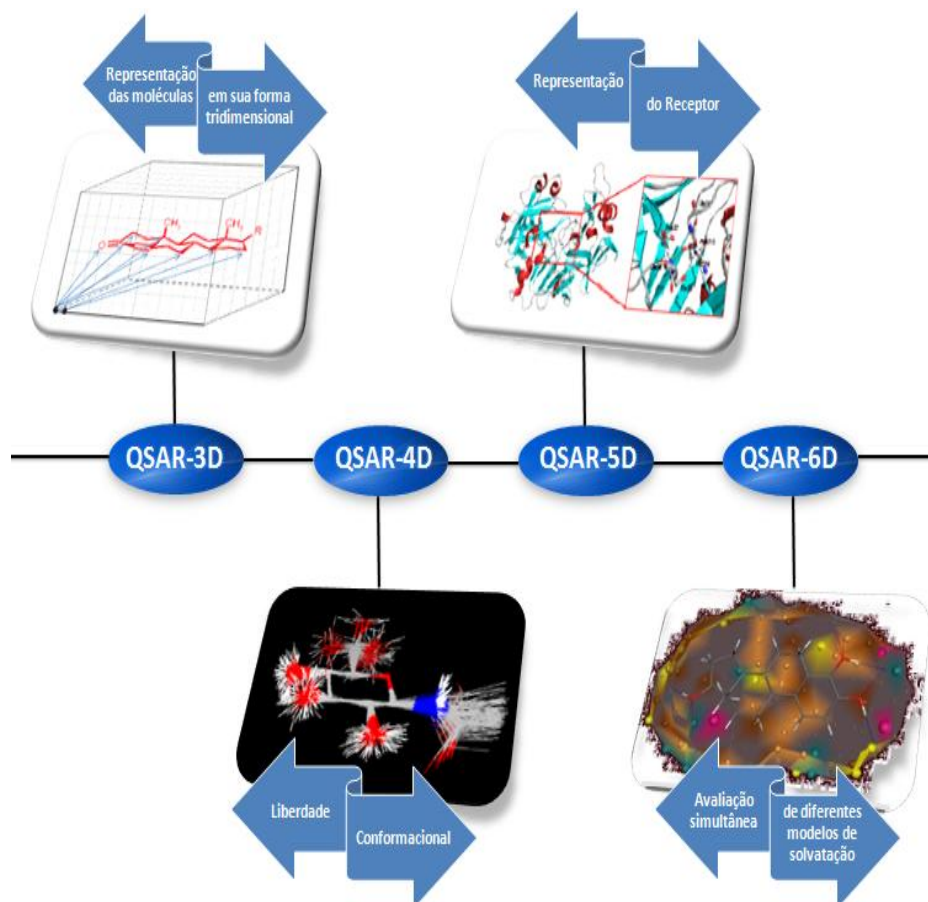


Figura 8 Distinção entre os formalismos da metodologia QSAR

Existem ainda outras metodologias de QSAR, bem como diversas outras abordagens que vêm sendo utilizadas nesses estudos. Dentre elas, destacam-se CoMSIA (KLEBE; ABRAHAM; MEITZNER, 1994), HQSAR (LOWIS, 1997), LQTA-QSAR (MARTINS et al., 2009) e MIA-QSAR (FREITAS; BROWN; MARTINS, 2005). Todas as metodologias QSAR exibem diferenças, vantagens

e desvantagens, mas buscam descrever um sistema de dados da maneira mais semelhante ao real possível. Na Figura 8 estão sintetizadas as diferenças básicas entre as metodologias QSAR 3D, 4D, 5D e 6D, resumindo o que foi discutido.

2.4.2 Descritores físico-químicos na construção de modelos QSAR

A metodologia tradicional de QSAR, aqui denominada QSAR-tradicional, também merece destaque. A mesma se baseia no uso de descritores físico-químicos e pode oferecer informações úteis para a compreensão de propriedades biológicas, químicas ou físicas que estejam relacionadas à estrutura química.

Um fator problemático nessa metodologia consiste na geração de um elevado número de descritores e, para ter êxito na interpretação físico-química biológica do sistema, é necessário que o número de descritores do modelo não seja exacerbado. Idealmente, do total de descritores encontrados, cerca de seis ou menos devem ser capazes de correlacionar os dados, pois, por convenção, para cada grupo de cinco ou seis moléculas incluídas no modelo QSAR, apenas uma variável deve ser utilizada (KUBINYI; 2003; TOPLISS; COSTELLO, 1972; UNGER; HANSCH, 1973). Uma vez que a relação quantitativa entre estrutura e atividade, segundo essa metodologia, é usualmente feita por meio de regressão linear múltipla (RLM), quanto menos descritores forem utilizados, melhor a interpretação físico-químico-biológica alcançada.

O ponto crucial é quais descritores são capazes de atingir a melhor interpretação do modelo. Com tantos descritores, essa tarefa torna-se difícil ou, até mesmo, impossível. Entretanto, com o uso de técnicas de seleção de variáveis, como a análise dos componentes principais (*principal component analysis*, ou PCA) (PEARSON, 1901) e algoritmo genético (*genetic algorithm*, ou GA) (HOLLAND, 1975), obtém-se uma redução do número de descritores.

O PCA transforma dados complexos e apresenta-os na forma de gráficos que possibilitam a visualização de quais descritores contribuem mais para o sistema, ou aqueles irrelevantes, passíveis de deleção. Assim, espera-se que as informações mais importantes tornem-se evidentes. O GA baseia-se na evolução da vida na Terra (cruzamento, seleção natural e mutação genética), avalia o sistema e apresenta os descritores que melhor evoluíram durante o processo. Dessa maneira, aqueles de maior importância para descrever a propriedade em questão podem ser escolhidos (FERREIRA; MONTANARI; GAUDIO, 2002).

Em ambas as metodologias (PCA e GA), o processo é aleatório; então, a repetição do mesmo procedimento resulta na seleção de diferentes descritores, o que gera várias equações RLM, devendo ser escolhida aquela que melhor explica o sistema.

2.4.3 Análise multivariada de imagem aplicada ao QSAR

O método de análise de imagem multivariada aplicada às relações quantitativas entre estrutura/atividade (*multivariate image analysis applied to quantitative structure-activity relationships*, ou MIA-QSAR) (FREITAS; BROWN; MARTINS, 2005) consiste numa técnica simples, bastante utilizada para auxiliar nas propostas de novas drogas. Essa abordagem inicia-se com a criação de descritores binários que são obtidos a partir de *pixel* de desenhos de estruturas químicas 2D. Essa técnica pode envolver o uso de programas de livre acesso, facilitando o uso da metodologia. Por ser uma técnica com utilização de imagens 2D para gerar descritores, eliminam-se muitas complicações envolvidas em QSAR 3D, tais como varredura conformacional e regras de alinhamento tridimensional dos ligantes. Dessa forma, pode-se alcançar um avanço no custo computacional (FREITAS, 2007).

Não raramente, encontram-se, na literatura, afirmações deletérias acerca do uso de QSAR-2D, cujos autores ressaltam que, apesar de ser uma metodologia bastante eficaz na investigação das relações quantitativas entre a estrutura química e a atividade de compostos bioativos, apresenta limitação em relação à abordagem do aspecto tridimensional (TAVARES, 2004). Por outro lado, outros autores (ESTRADA; MOLINA; PERDOMO-LÓPEZ, 2001, BROWN; MARTIN, 1997, TIAN; ZHOU; LI, 2007) afirmam que descritores bidimensionais não são inferiores aos tridimensionais, pelo menos em muitos exemplos práticos.

Segundo Freitas, Brown e Martins (2005), a imagem 2D de uma molécula não representa efeitos particulares, como, por exemplo, interações estéricas ou de qualquer parâmetro físico-químico, mas é uma forma geral para representar toda a estrutura química. Pequenas diferenças no desenho de uma estrutura podem causar mudanças significativas nas suas propriedades, ou seja, cada pixel atua como um código, os descritores são binários, os pixels pretos correspondem ao dígito 0 e a cor branca corresponde ao número 765, de acordo com o sistema de composição de cores *red-green-blue*, ou RGB (vermelho-verde-azul, em que cada componente corresponde ao valor 255). Como exemplo, os enantiômeros *R* e *S* de uma molécula quiral podem ser distinguidos pelo desenho de cunha ou tracejados das ligações nesse centro e isso fará com que, se for o caso, diferenças substanciais na propriedade estudada sejam notadas (por exemplo, para as atividades biológicas), ou seja, este método é altamente sensível ao desenho. Tal fato é confirmado por Tropsha (2010), ao ressaltar que o resultado não depende do nível de representação da estrutura química, 2D ou 3D, mas da capacidade do método em distinguir pares de enantiômeros.

Na Figura 9 está representado um par de enantiômeros presente no conjunto de dados deste estudo. A cunha é totalmente preenchida por pixel

preto, enquanto a ligação tracejada é mesclada por pixel preto e espaços em branco. Esta simples diferença determina sequências de descritores diferentes, o que discrimina entre isômeros *R* e *S*.

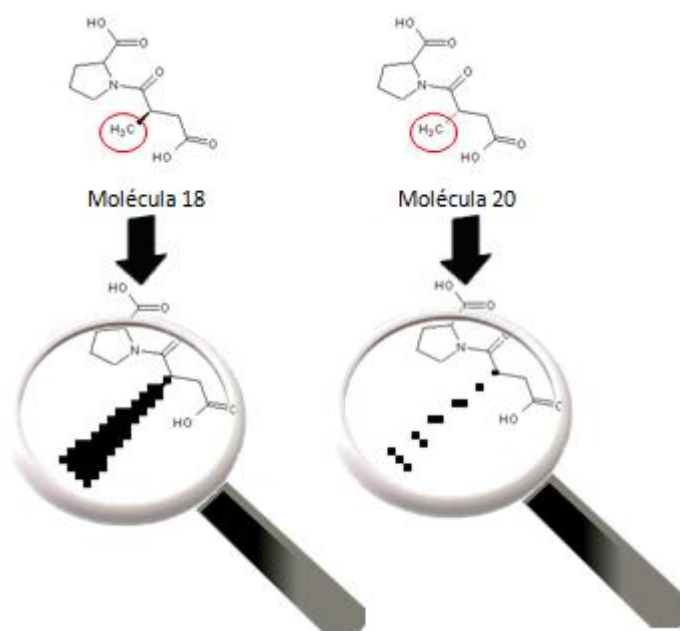


Figura 9 Ilustração das diferenças nas representações das moléculas presentes no conjunto de dados deste trabalho (moléculas extraídas de GRUENFELD et al., 1983). A representação de pixels pretos e espaços em brancos são as diferenças determinantes que explicam a variação no bloco das variáveis dependentes (no caso, as atividades biológicas)

Nesse sentido, o desenho das estruturas químicas compreende a etapa mais importante na construção de um modelo MIA-QSAR. As moléculas devem ser desenhadas cuidadosamente, mantendo a mesma disposição dos átomos,

grupos substituintes e direção da cadeia carbônica para todas as estruturas. Contudo, diferenças entre representar um grupo substituinte metil, como -Me ou -CH₃, não alteram os resultados, desde que todos os compostos sejam desenhados da mesma forma (GOODARZI; FREITAS; FERREIRA, 2009), conforme circulado em vermelho nas estruturas químicas da Figura 9.

A afirmação de que descritores MIA podem codificar propriedades químicas, físicas e biológicas, bem como o exemplo ilustrado na Figura 9 (isômeros), depende de como os substituintes estão representados, pois a série de descritores altera-se conforme a representação 2D da estrutura química, obtendo-se a correlação. A comprovação de que tal correlação não é obra do acaso é ilustrada pelo trabalho divulgado recentemente por Cormanich, Nunes e Freitas (2012). Um modelo MIA-QSAR foi construído a partir de uma série com 70 derivados benzimidazólicos e atividade tricomonocida conhecida, e comparado a outro modelo MIA-QSAR, cujas imagens correspondem às letras do alfabeto, que foram correlacionadas aos respectivos números no ordenamento do alfabeto (a letra A corresponde ao número 1, a letra B ao número 2, e assim sucessivamente). O primeiro modelo resultou em uma boa correlação; já o segundo modelo não resultou em correlação alguma, demonstrando que a metodologia MIA-QSAR é uma ferramenta útil para modelagem molecular.

Na literatura encontram-se diversos estudos envolvendo a utilização do método MIA-QSAR. Grande parte desses estudos pode ser de grande relevância para o avanço nas descobertas de novas drogas, como o trabalho desenvolvido por Cormanich, Goodarzi e Freitas (2009), em que foi avaliada a atividade biológica de uma série de inibidores da WEE1 (enzima tirosina quinase), os quais foram previamente modelados por meio dos métodos CoMFA e CoMSIA (KURUP; GARG; HANSCH, 2001; VERMA; HANSCH, 2007; 2009). Resultados estatísticos semelhantes foram encontrados entre as técnicas,

confirmando que MIA-QSAR explica as mesmas variações observadas por meio de métodos tridimensionais e é tão seguro quanto os mesmos.

Estudos utilizando o método MIA-QSAR também foram desenvolvidos para modelar as atividades anti-HIV de uma série de compostos, alcançando resultados satisfatórios e coerentes com aqueles obtidos por meio de outras metodologias (FREITAS, 2006; GOODARZI; FREITAS, 2008; KURUP et al., 2003; PINHEIRO et al., 2008). Recentemente, pela primeira vez, a modelagem MIA-QSAR foi utilizada para a previsão das citotoxicidade de derivados do ritonavir (inibidores da HIV-1 protease), alcançando-se bons resultados. Esse estudo ratifica o fato de que o método é aplicável a diferentes propriedades de compostos (SILVA; FREITAS, 2012).

Freitas (2008) desenvolveu outros estudos com alguns compostos com atividades ansiolíticas. A calibração do conjunto composto por 47 moléculas (antagonistas dos receptores 5-HT_{2C}) foi construída e um teste de *leave-one-out* (ou “deixe uma amostra fora”), bem como uma validação externa, foi realizada, a fim de testar a capacidade preditiva do modelo. Os resultados estatísticos $r^2 = 0,866$ e $r^2_{\text{teste}} = 0,681$ (utilizando seis variáveis latentes) foram comparáveis aos obtidos por outros pesquisadores (BROMIDGE et al. 1998), em que um modelo CoMFA foi utilizado como abordagem 3D.

A metodologia é aplicável a estudos voltados para problemas químicos mais simples, como, por exemplo, no trabalho realizado por Goodarzi e Freitas (2008), em que os pontos de ebulição de um conjunto de 58 alcoóis alifáticos foram modelados por meio da técnica MIA-QSAR. Isso deixa claro que a técnica não é somente aplicada à Química Medicinal, mas às diversas áreas da ciência.

A combinação de problemas químicos com o conhecimento matemático torna a metodologia MIA-QSAR aplicável aos mais diversos estudos, nos quais agrupar cálculos matemáticos com propriedades físico-químicas amplia a gama

de possibilidades de estudo na área. Por exemplo, a implementação de lógica fuzzy para analisar dados obtidos por MIA-QSAR apresentou melhores resultados do que simples regressão dos descritores com suas variáveis dependentes por mínimos quadrados parciais PLS (GOODARZI; FREITAS; JENSEN, 2009).

Como em toda metodologia, é necessário ter cuidado ao interpretar os dados; muitas vezes, encontram-se falsos resultados obtidos por meio de uma interpretação equivocada. Assim, são necessários uma análise criteriosa e um profundo conhecimento para que o estudo seja realmente válido, tendo em vista que, dependendo do objetivo, outras técnicas podem ser empregadas.

Um estudo de caso com a finalidade de comparar a metodologia MIA-QSAR (FREITAS; BROWN; MARTINS, 2005) com uma metodologia tridimensional baseada no uso do programa GRID (NILSSON et al., 1998), em que mesmo conjunto de dados (58 compostos com afinidade para o subtipo do receptor D2 de dopamina) foi utilizado para a construção dos dois modelos, demonstrou que os resultados são compatíveis em ambas as metodologias. Curiosamente, uma molécula contendo um grupo nitro (NO_2) como substituinte comportou-se de maneira bastante peculiar, em ambas as análises. Essa molécula apresentou comportamento diferente do das restantes no conjunto de dados, permanecendo afastada da reta onde concentra os compostos usados na calibração do modelo, provavelmente por influência do grupo nitro. Essa comparação é ideal para concluir que o modelo MIA-QSAR é robusto e consegue distinguir as diferenças estruturais nos compostos químicos e parece responder igualmente aos métodos cuja descrição é baseada puramente em informação química (como efeitos estéricos e eletrostáticos).

O MIA-QSAR envolve diversas possibilidades de modelagem e aprimoramento, mas, sobretudo, apresenta-se útil na busca por novas entidades químicas. A viabilidade do uso dessa ferramenta é fundamentada pela

simplicidade envolvida no processo de construção do modelo, enquanto os resultados alcançados sugerem o uso dessa metodologia para modelar as atividades biológicas de moléculas pertencentes a uma série congênere.

2.4.3.1 Sistematização da metodologia QSAR: critérios e regras

As metodologias QSAR, desde seu surgimento e sistematização, por volta de 1964, com o trabalho de Hansch e Fujita, até os dias atuais, passaram por diversas transformações. O constante crescimento de seu uso para auxiliar na proposta de novas drogas é uma preocupação, visto que facilmente se constróem modelos baseados na correlação entre a estrutura química e a atividade biológica. Contudo, a confiabilidade do modelo não pode ser garantida apenas por valores ótimos de r^2 (coeficiente de correlação ao quadrado) $> 0,8$ ou q^2 (coeficiente de correlação ao quadrado na validação cruzada *leave-one-out*) $> 0,5$, mas de um conjunto de testes estatísticos, destacando-se a validação externa que, segundo Golbraikh e Tropsha (2002), é a única maneira de alcançar um modelo QSAR confiável.

Ainda sim, outros cuidados mais complexos devem ser tomados, como a combinação exploratória de todos os possíveis conjuntos de descritores e selecionar modelos caracterizados pela alta precisão na previsão, por exemplo, da atividade biológica do conjunto de dados utilizado. Deve-se atentar para os erros estruturais ou experimentais presentes no conjunto de amostras. Enquanto os primeiros podem ser notados facilmente e eliminados, os equívocos experimentais não podem ser sanados durante a etapa de modelagem, uma vez que não existem métodos capazes de identificá-los. A remoção de compostos inorgânicos e organometálicos, sais e misturas, dentre outros compostos que se apresentam fora do limite de correlação do método empregado, ajuda a minimizar os erros envolvidos no processo de modelagem molecular. Outro

fator importante é o tamanho do conjunto de dados. Tratando-se da atividade biológica, é aconselhável que, no mínimo, 40 compostos façam parte do modelo, divididos entre conjunto treino e teste (TROP SHA, 2010).

O início da construção de qualquer modelo QSAR acontece com uma busca, na literatura, por um conjunto de moléculas com atividade biológica conhecida e que sejam similares estruturalmente, ou seja, pertencentes a uma série congênere.

Essa busca não garante que todas as moléculas encontradas devam fazer parte do modelo, pois a identificação e a deleção de *outliers* (moléculas que apresentam um comportamento não análogo ao das outras) constituem etapas importantes empregadas nessa metodologia. Caracteriza-se por processo laborioso, pois, muitas vezes, a retirada de um composto do modelo pode prejudicar a capacidade de predição e a não remoção pode influenciar a estabilidade do modelo, o que dificulta a escolha correta, mas são passos que garantem ao modelo maior potencial para a previsão da propriedade de novas amostras.

Verma e Hansch (2005) discorrem que *outliers* podem agir por um mecanismo diferente ou interagir com o receptor em diferentes modos, sendo, portanto, indispensáveis na interpretação de quais compostos atuam por um mecanismo comum. Uma avaliação atenta pode resultar em informações preciosas, como, por exemplo, a retirada de um único composto que apresenta um anel benzênico na estrutura e, conseqüentemente, sua atividade biológica é a mais elevada. A remoção da molécula do modelo QSAR pode prejudicar a previsão da atividade biológica de novos compostos ou a razão para que esse composto seja um *outlier* pode ser decorrente de erros experimentais, visto que outros compostos pertencentes ao conjunto de dados não têm anel benzênico e apresentam atividade biológica mensurada similar, o que justificaria sua deleção.

Em conjunto com as metodologias QSAR, outras técnicas podem ser aplicadas, a fim de melhorar o entendimento e o conhecimento de um determinado grupo de moléculas, bem como as interações e as afinidades desses ligantes com seus respectivos receptores, de forma a validar e interpretar os dados bioquímicos. Um método computacional que busca prever geometrica e energeticamente o encaixe de um composto no sítio ativo de uma proteína, denominado *docking* (ancoramento molecular), é útil para esse objetivo (GONÇALVES, 2008). Essa técnica, inclusive, é útil para validar as estimativas de bioatividades de novas moléculas concebidas por meio de análises QSAR.

A busca por um conhecimento mais específico acerca do comportamento de um novo candidato a fármaco desvia do escopo da metodologia QSAR empregada no presente trabalho. Para isso, existem ferramentas computacionais, baseadas em modelos construídos a partir de grandes bibliotecas de dados, úteis para avaliar, ainda que preliminarmente, o perfil farmacológico de novos compostos. Avaliações da biodisponibilidade (porcentagem do composto que alcança a corrente sanguínea em um determinado tempo), uma das principais propriedades farmacocinéticas, são desejáveis, uma vez que estudos preliminares de **absorção**, **distribuição**, **metabolismo** e **excreção**, ou ADME, podem oferecer uma noção se o fármaco exercerá, de fato, seu efeito terapêutico.

Finalmente, confrontar os dados obtidos com dados experimentais atribui ao estudo fundamentos necessários para garantir a veracidade das informações, validando, dessa forma, o modelo QSAR construído.

3 OBJETIVO GERAL

Este trabalho foi realizado com o objetivo geral de elaborar novas estruturas químicas potencialmente ativas como inibidoras da enzima conversora de angiotensina (anti-hipertensivos), por meio das técnicas computacionais QSAR e *docking*, além de proceder à avaliação preliminar do perfil farmacocinético das estruturas propostas.

3.1 Objetivos específicos

Especificamente, buscou-se:

1. modelar as atividades biológicas de uma série de anti-hipertensivos, utilizando o método MIA-QSAR e propor estruturas químicas potencialmente ativas, empregando-se o modelo MIA-QSAR construído;
2. modelar as atividades biológicas de uma série de anti-hipertensivos utilizando descritores físico-químicos obtidos pelo programa Dragon, bem como extrair informações físicas, químicas e/ou biológicas úteis para propor estruturas químicas potencialmente ativas, utilizando o modelo construído baseado em descritores físico-químicos;
3. realizar estudos de *docking* utilizando o programa Molegro Virtual Docker, para avaliação bioquímica das interações enzima-substrato que respondem pela atividade biológica das moléculas propostas;
4. calcular parâmetros ADME das moléculas propostas para avaliação farmacocinética e potencialidades dos novos fármacos.

4 METODOLOGIA

O desenvolvimento do trabalho ocorreu no Laboratório de Modelagem Molecular, no Departamento de Química da Universidade Federal de Lavras (UFLA), por meio de uma estação de trabalho Dell Precision T7500 (processador Intel Quad-Core Xeon X3430 2,4 GHz, 8 GB DDR3 de memória e HD de 500 GB, sistema operacional *Windows 7 Professional*).

Na Figura 10 é mostrado um diagrama com as sequências das etapas que foram realizadas, além da metodologia aplicada.

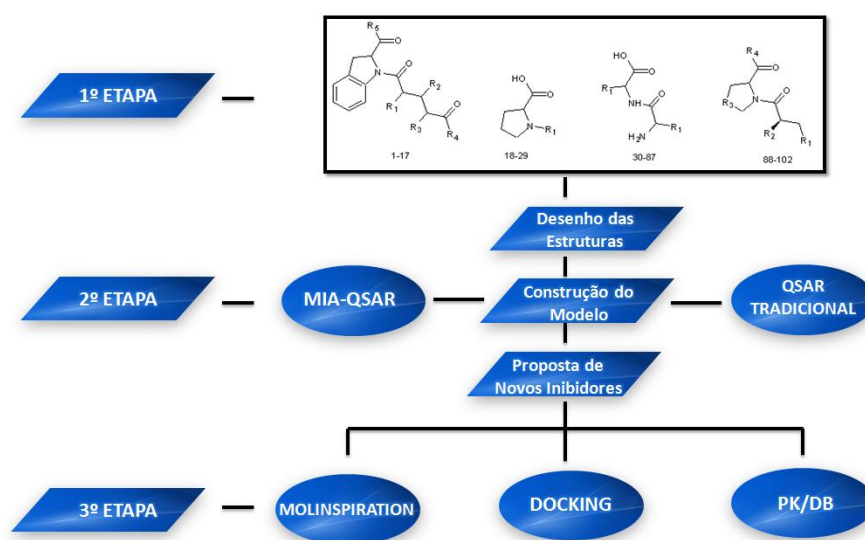


Figura 10 Diagrama das etapas envolvidas na realização do presente estudo.

A primeira etapa consistiu em uma busca, na literatura, por um determinado grupo de moléculas pertencentes a uma série congênere, pois um mínimo de similaridade estrutural é necessário para o alinhamento 2D (bidimensional) na metodologia MIA-QSAR. A série de compostos tem atividade biológica experimental conhecida, ou seja, atividade anti-hipertensiva. O conjunto de dados encontrado é apresentado na Tabela 3.

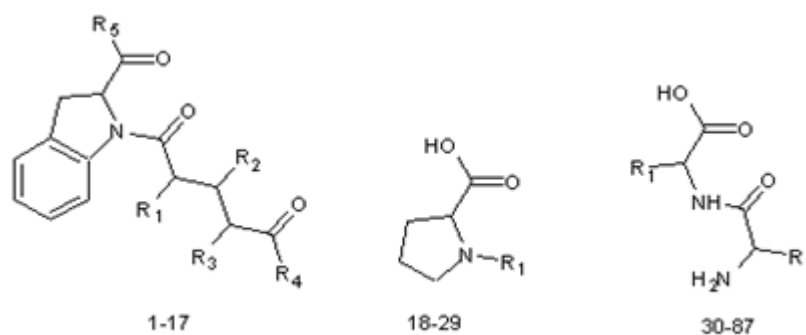
Dois métodos QSAR foram empregados na segunda etapa: MIA-QSAR e QSAR tradicional baseado em descritores físico-químicos. Estes descritores foram selecionados utilizando-se duas técnicas, PCA e GA. Em seguida, potenciais inibidores para ECA foram propostos e suas atividades biológicas foram previstas empregando-se os modelos construídos.

Para a avaliação da afinidade ligante-receptor, aplicou-se o *docking*. Efetuaram-se, também, avaliações preliminares acerca de alguns parâmetros farmacocinéticos dos inibidores propostos, com os programas Molinspiration e PK/DB, disponíveis nos sites: <http://www.molinspiration.com> e <http://miro.ifsc.usp.br/pkdb/>, respectivamente, sendo esta a terceira etapa.

4.1 Construção do modelo MIA-QSAR

No presente trabalho, uma série com 87 dipeptídeos modificados com atividade anti-hipertensiva (inibidores de ECA) foi obtida da literatura (CUSHMAN et al., 1977; GRUENFELD et al., 1983; OLEG et al., 2004), os quais são apresentados na Tabela 3. Dentro desse conjunto de dados encontra-se o captopril (composto 27), fármaco de referência para o estudo.

Tabela 3 Série de compostos usados nas análises QSAR.



Composto	R ₁	R ₂	R ₃	R ₄	R ₅	pIC ₅₀ exp.
1^a	H	H	H	OH	OH	6,59
2	Me	H	H	OH	OH	7,19
3	Me	H	H	OMe	OH	5,16
4	Me	H	H	NH ₂	OH	4,80
5	(<i>R</i>)-Me	H	H	OH	OH	7,36
6^a	(<i>R</i>)-Me	H	H	OEt	OH	4,96
7	<i>i</i> -Pr	H	H	OH	OH	7,30
8	H	Me	H	OH	OH	4,43
9	H	Me ₂	H	OH	OH	3,54
10	H	H	Me	OH	OH	6,04
11^a	H	H	Me	OH	OEt	3,77
12	Me	H	Me	OH	OEt	5,06
13	Me	H	Me	OH	OH	6,72
14	Me	H	Me	OH	OH	7,27
15	(<i>R</i>)-Me	H	(<i>R</i>)-Me	OH	OH	7,55

...continua

Tabela 3, continuação...

Composto	R ₁	R ₂	R ₃	R ₄	R ₅	pIC ₅₀ exp.
16 ^a	(<i>R</i>)-Me	H	(<i>R</i>)-Me	OEt	OH	5,26
17	(<i>R</i>)-Me	H	(<i>R</i>)-Me	OH	OEt	4,55
18	$\text{HOOCCH}_2\overset{\text{CH}_3}{\underset{\downarrow}{\text{C}}}\text{HCO}^-$	-	-	-	-	4,66
19	$\text{HOOC}\overset{\text{CH}_3}{\underset{\cdot\cdot}{\text{C}}}\text{HCH}_2\text{CO}^-$	-	-	-	-	3,21
20	$\text{HOOCCH}_2\overset{\text{CH}_3}{\underset{\cdot\cdot}{\text{C}}}\text{HCO}^-$	-	-	-	-	2,83
21 ^a	$\text{HOOC}\overset{\text{CH}_3}{\underset{\downarrow}{\text{C}}}\text{HCH}_2\text{CO}^-$	-	-	-	-	2,59
22	$\text{HOOCCH}_2\text{CH}_2\overset{\text{CH}_3}{\underset{\downarrow}{\text{C}}}\text{HCO}^-$	-	-	-	-	5,31
23	$\text{HOOCCH}_2\overset{\text{CH}_3}{\underset{ }{\text{C}}}\text{HCH}_2\text{CO}^-$	-	-	-	-	2,92
24	$\text{HOOC}\overset{\text{CH}_3}{\underset{ }{\text{C}}}\text{HCH}_2\text{CH}_2\text{CO}^-$	-	-	-	-	3,59
25	$\text{HOOCCH}_2\text{CH}_2\overset{\text{CH}_3}{\underset{\cdot\cdot}{\text{C}}}\text{HCO}^-$	-	-	-	-	3,02
26 ^a	$\text{HS}\overset{\text{CH}_3}{\underset{ }{\text{C}}}\text{HCO}^-$	-	-	-	-	5,96
27 ^b	$\text{HSCH}_2\overset{\text{CH}_3}{\underset{\downarrow}{\text{C}}}\text{HCO}^-$	-	-	-	-	7,64
28	$\text{HS}\overset{\text{CH}_3}{\underset{ }{\text{C}}}\text{HCH}_2\text{CO}^-$	-	-	-	-	5,96

...continua

Tabela 3, continuação...

Composto	R ₁	R ₂	R ₃	R ₄	R ₅	pIC ₅₀ exp.
29	$\begin{array}{c} \text{CH}_3 \\ \\ \text{HSCH}_2\text{CHCO}^- \end{array}$	-	-	-	-	5,62
30	YG ^c	-	-	-	-	2,70
31^a	YA ^c	-	-	-	-	3,34
32	WG ^c	-	-	-	-	2,23
33	VY ^c	-	-	-	-	4,66
34^b	VW ^c	-	-	-	-	5,80
35	VP ^c	-	-	-	-	3,38
36^a	VG ^c	-	-	-	-	2,96
37	VF ^c	-	-	-	-	4,28
38	TG ^c	-	-	-	-	2,00
39	SG ^c	-	-	-	-	2,07
40	RW ^c	-	-	-	-	4,80
41^a	RP ^c	-	-	-	-	3,74
42	RF ^c	-	-	-	-	3,64
43	RA ^c	-	-	-	-	3,34
44	QG ^c	-	-	-	-	2,13
45	PG ^c	-	-	-	-	1,77
46^a	MG ^c	-	-	-	-	2,32
47	LG ^c	-	-	-	-	2,06
48	LA ^c	-	-	-	-	3,51
49	KG ^c	-	-	-	-	2,49
50	KA ^c	-	-	-	-	3,42
51^a	IY ^c	-	-	-	-	5,43
52^b	IW ^c	-	-	-	-	5,70

...continua

Tabela 3, continuação...

Composto	R ₁	R ₂	R ₃	R ₄	R ₅	pIC ₅₀ exp.
53	IP ^c	-	-	-	-	3,89
54	IG ^c	-	-	-	-	2,92
55	IF ^c	-	-	-	-	3,03
56^a	HL ^c	-	-	-	-	2,49
57	HG ^c	-	-	-	-	2,20
58	GY ^c	-	-	-	-	3,68
59	GW ^c	-	-	-	-	4,52
60	GV ^c	-	-	-	-	2,34
61^a	GT ^c	-	-	-	-	2,24
62	GS ^c	-	-	-	-	2,42
63	GR ^c	-	-	-	-	2,49
64	GQ ^c	-	-	-	-	2,15
65	GP ^c	-	-	-	-	3,35
66^a	GM ^c	-	-	-	-	2,85
67	GL ^c	-	-	-	-	2,60
68	GK ^c	-	-	-	-	2,27
69	GI ^c	-	-	-	-	2,92
70	GH ^c	-	-	-	-	2,51
71^a	GG ^c	-	-	-	-	2,14
72	GF ^c	-	-	-	-	3,20
73	GE ^c	-	-	-	-	2,27
74	GD ^c	-	-	-	-	2,04
75	GA ^c	-	-	-	-	2,70
76^a	FR ^c	-	-	-	-	3,04
77	FG ^c	-	-	-	-	2,43

...continua

Tabela 3, continuação...

Composto	R ₁	R ₂	R ₃	R ₄	R ₅	pIC ₅₀ exp.
78	EG ^c	-	-	-	-	2,00
79	EA ^c	-	-	-	-	2,00
80	DG ^c	-	-	-	-	1,85
81^a	DA ^c	-	-	-	-	2,42
82	AY ^c	-	-	-	-	4,06
83	AW ^c	-	-	-	-	5,00
84	AP ^c	-	-	-	-	3,64
85	AG ^c	-	-	-	-	2,60
86^a	AF ^c	-	-	-	-	3,72
87	AA ^c	-	-	-	-	3,21

^a Compostos de validação externa. ^b Os compostos mais ativos de cada uma das três séries (**27**, **34** e **52**), usados para combinar as subestruturas e, então, fornecer os compostos propostos **A-E**. ^c As denominações dos dipeptídeos **30-87** correspondem às letras utilizadas para designar aminoácidos.

Fonte: Cushman et al. (1977); Gruenfeld et al. (1983); Oleg et al. (2004)

As estruturas químicas foram desenhadas sistematicamente, utilizando-se o programa ChemSketch (ADVANCED CHEMISTRY DEVELOPMENT ACD, 2009), considerando os cuidados citados no referencial teórico, pois, neste momento, tem-se a etapa mais importante na construção do modelo e qualquer erro pode ter influência nos resultados. Esse processo é ilustrado na Figura 11.

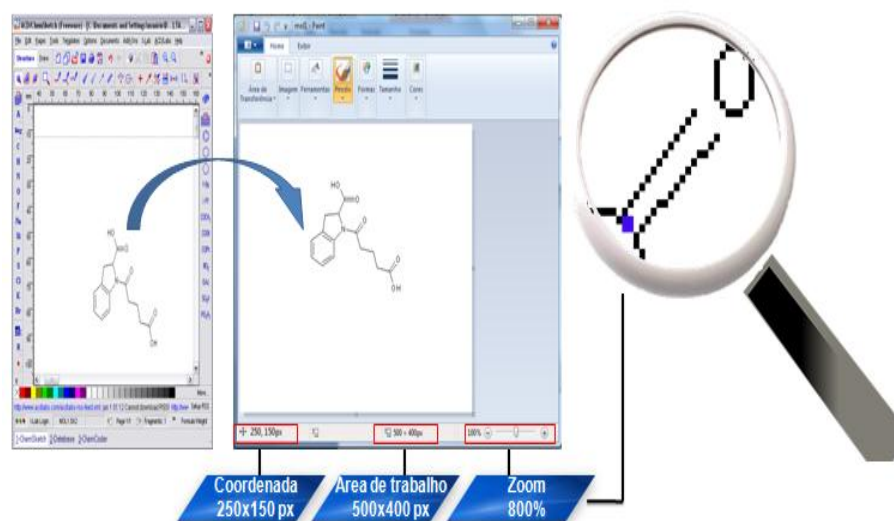


Figura 11 Primeiras etapas para a construção do modelo MIA-QSAR, que consiste na preparação e na organização do conjunto de dados. Dois programas são utilizados: ChemSketch (para desenhar as estruturas químicas) e o aplicativo Paint do Microsoft Windows (para alinhar e converter as estruturas químicas em imagens)

Cada estrutura química foi transferida para uma área de trabalho do aplicativo *Paint*, do Microsoft Windows. A definição do tamanho da área de trabalho é obrigatória, pois, posteriormente, todas as matrizes devem apresentar a mesma dimensão que, nesse caso, foi definida como 500 x 400 px, conforme destacado na Figura 11. Em seguida, um pixel (px) comum a todas as estruturas químicas, localizado no carbono da carbonila adjacente ao anel pirrolidínico (destacado por um px azul na Figura 11), foi selecionado. Daí a importância de que a série de compostos escolhidos apresentasse certa similaridade estrutural. Uma a uma, as estruturas químicas foram movidas de tal forma que o pixel selecionado fosse ajustado à coordenada 250×150 px da área de trabalho, de

acordo com o demonstrado na Figura 11. Esse último passo corresponde ao alinhamento 2D e é feito manualmente (com o auxílio do *mouse*). Com o uso da ferramenta de *ZOOM*, do aplicativo *Paint* (aumentando a imagem de 100% para 800%, por exemplo), essa tarefa é facilitada, pois se consegue identificar o pixel escolhido e arrastar a estrutura química na coordenada fixada. Após conferir se a estrutura química encontrava-se na coordenada correta, a imagem foi salva como bitmaps (.bmp). Portanto, essa é uma etapa fundamental na análise, pois cada imagem (um plano bidimensional) será sobreposta à outra, formando um arranjo tridimensional em que as partes comuns entre as estruturas da série congênere (o esqueleto básico) estejam congruentes. O resultado da sobreposição de todas as imagens é ilustrado na Figura 12. Sendo assim, os substituintes e as porções das moléculas que variam entre si explicam a variância dos dados biológicos.

Em MIA-QSAR, as imagens correspondem a desenhos de estruturas químicas e o fundamento do método está na utilização de pixels de imagens como descritores. Como os pixels podem ser tratados numericamente como binários, a cor branca equivale ao dígito 765 e os pixels pretos ao dígito 0, de acordo com o sistema de cores RGB (a soma das cores vermelha, verde e azul, cada uma correspondente a 255, equivale à cor branca; a ausência das três componentes equivale à cor preta). As modificações estruturais ou mudança na posição dos substituintes em uma série congênere de moléculas correspondem a alterações nas coordenadas dos pixels da imagem, e essas alterações explicam a variância no bloco **Y**, o bloco correspondente às variáveis dependentes (atividades biológicas).

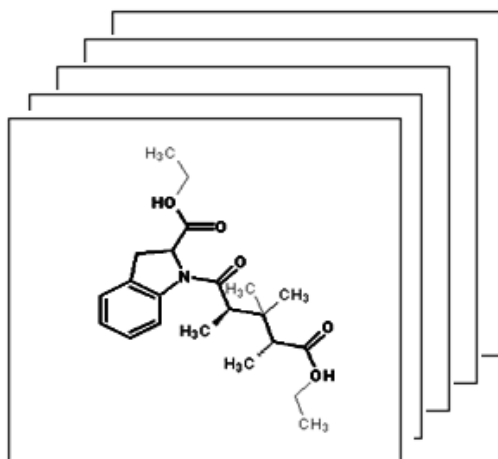


Figura 12 Resultado da sobreposição de todas as imagens, formando um arranjo tridimensional. A parte da imagem em negrito corresponde à parte similar da série, os substituintes são as porções variáveis nas moléculas e, portanto, responsáveis pelas alterações nos valores de atividade biológica

O arranjo tridimensional pode ser desdobrado para um arranjo bidimensional, o que permite a regressão dessa matriz com o bloco **Y** por meio de mínimos quadrados parciais (PLS bilinear, do inglês *partial least squares*), o qual é um dos mais indicados, pois o número de descritores excede o número de amostras.

Na Figura 13 observa-se o procedimento para tratamento e análise das imagens, e os comandos a seguir mostram como as imagens podem ser carregadas e convertidas em binários por meio do programa Matlab:

```
[filename,MAP]=imread('filename.bmp','bmp');
filename=double(filename);
filename=(filename(:,:,1)+filename(:,:,2)+filename(:,:,3));
```

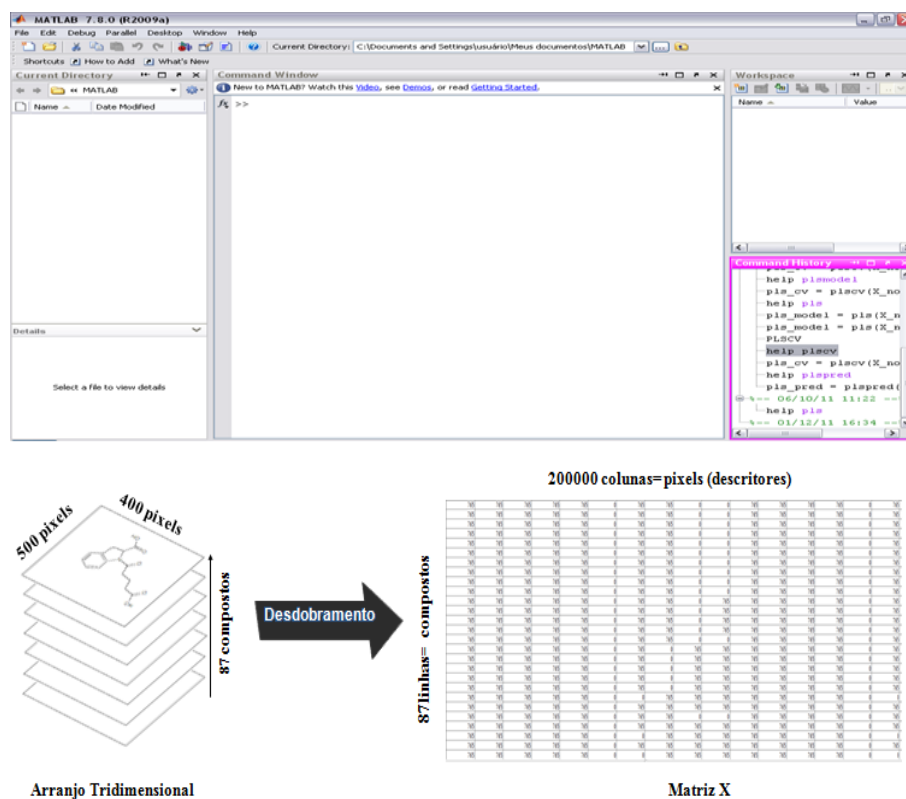


Figura 13 Construção do arranjo tridimensional e desdobramento para a matriz **X**. Na parte superior da figura, a interface do programa MatLab; na parte inferior, a demonstração do processo de desdobramento

A área de trabalho definida foi de 500x400 px, em que cada px corresponde a um descritor, o que resultou em 200.000 descritores, após o desdobramento. O conjunto de dados contém 87 compostos e cada linha na **Matriz X** representa uma estrutura química, descrita por uma sequência de descritores que varia entre 0 ou 765. Assim, a **Matriz X** tem uma dimensão de 87 linhas (compostos) por 200.000 colunas (descritores). Uma analogia ao método MIA-QSAR é o funcionamento dos computadores que, com o uso de

códigos binários (0 e 1), formam sequências que são traduzidas em letras, comandos, figuras, entre outros, ou seja, cada estrutura química no método MIA-QSAR transforma-se em sequências numéricas.

Para agilizar os cálculos, as colunas com variância zero (por exemplo, os espaços em branco comuns para todas as imagens ou as partes congruentes das estruturas químicas) foram removidas. Essa ação foi desempenhada pelo programa *Matlab* (MATHWORKS, 2005). A redução da **Matriz X** foi significativa, passando para uma dimensão de 87×4533, o que diminui consideravelmente a demanda computacional.

4.2 Construção do modelo QSAR-tradicional

O modelo QSAR-tradicional baseia-se em descritores físico-químicos e apresenta grande relevância para uma melhor interpretação físico-químico-biológica do conjunto de amostras. Para sua construção, os 87 compostos foram desenhados e otimizados no programa HyperChem (HYPERCUBE, 2007), utilizando-se o método do Gradiente Conjugado – Versão Polak-Ribiere. Os descritores físico-químicos foram gerados pelo programa Dragon (TODESCHINI; PAVAN; CONSONNI, 2002). Geralmente, grandes quantidades de descritores resultam desse processo, o que dificulta a interpretação desejada. No presente estudo, 1.459 descritores foram gerados.

O emprego de métodos de seleção de variáveis, como o PCA e o GA, pode resolver esse problema.

A estratégia adotada foi tentar distinguir qual método de seleção de variáveis resultaria em descritores que melhor descrevessem o modelo, ou a combinação dos dois métodos, da seguinte maneira:

- a) utilizou-se apenas PCA;
- b) PCA seguido de GA;

- c) apenas GA;
- d) GA seguido de PCA;
- e) GA, repetindo o processo consecutivas vezes.

Os mesmos procedimentos seguidos para a construção do modelo MIA-QSAR foi seguido para a construção do modelo QSAR-tradicional, ou seja, mesmo conjunto de treinamento e teste (**Tabela 3**) foi utilizado. A única diferença foi que, quando o número de descritores excedeu o número de amostras, PLS foi utilizado e, quando o número de descritores foi menor que o número de amostras, MLR foi utilizado, para correlacionar as atividades biológicas com os descritores selecionados. As validações estatísticas são explicadas no próximo tópico.

4.3 Métodos de validação e construção dos modelos QSAR

Com o objetivo de construir os modelos QSAR, as 87 amostras com valores experimentais de pIC_{50} ($IC_{50} = 50\%$ de concentração inibitória para a ECA, em mol L^{-1}) disponíveis foram divididas em grupos de treinamento (69 moléculas, aproximadamente 80% do conjunto de dados) e teste (cerca de 20%, 18 compostos). As amostras do grupo teste foram escolhidas aleatoriamente para compor a validação externa, de forma que os compostos com valores extremos de pIC_{50} foram evitados, mas uma distribuição regular de valores de pIC_{50} ao longo do conjunto foi mantida (Tabela 3).

A utilização de validações estatísticas se faz necessária, para que o modelo QSAR construído seja considerado robusto e preditivo. Duas validações foram empregadas no modelo de calibração: a validação cruzada *leave-one-out* (LOOCV, deixe uma amostra fora) e validação externa (conjunto de teste).

Na validação cruzada (Figura 14) retira-se uma molécula do conjunto de treinamento; em seguida, realiza-se a regressão e, após esse procedimento, a

atividade biológica é prevista para a molécula que permaneceu de fora do modelo de calibração. Repete-se o processo até que todas as moléculas tenham ficado de fora do conjunto de calibração pelo menos uma vez. Utiliza-se essa validação interna para calcular um número ótimo de variáveis latentes (ou componentes PLS) e construir o modelo QSAR. Dessa forma, obtém-se uma ideia inicial do poder de previsibilidade e robustez. Bons modelos QSAR devem apresentar valores de $q^2 > 0,5$ (r^2_{LOOCV}) e $r^2_{\text{Cal.}} > 0,8$ (calibrado).

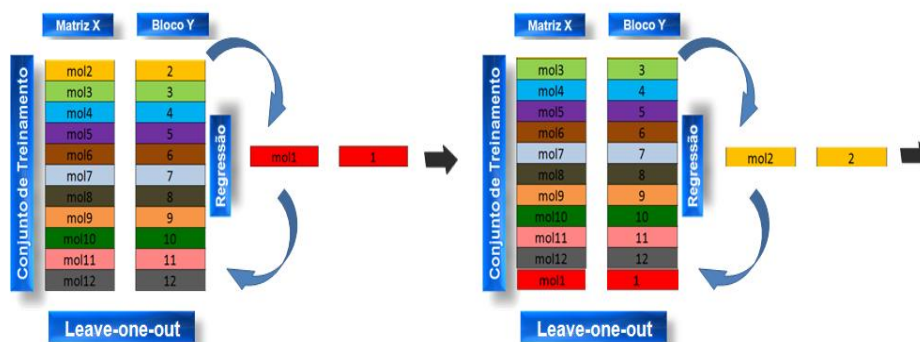


Figura 14 Processo de validação cruzada *leave-one-out*. Uma molécula de cada vez (na Matriz X) e seu respectivo valor de $\text{pIC}_{50\text{exp}}$ (Bloco Y) são deixados de fora (em vermelho) e procede-se com a regressão; utilizam-se os parâmetros de regressão para calcular o valor de pIC_{50} da amostra deixada de fora. Repete-se o procedimento até que todas as amostras tenham sido deixadas de fora

A validação externa (Figura 15) é o procedimento mais seguro para atestar um modelo QSAR (GOLBRAIKH; TROPSHA, 2002). Esse método de validação consiste em utilizar o modelo construído para prever o $\text{pIC}_{50\text{pred}}$ (predito) do conjunto de teste, composto por moléculas que em nenhum momento fizeram parte da construção do modelo. Como o $\text{pIC}_{50\text{exp}}$

(experimental) dessas moléculas é conhecido, pode-se comparar o $pIC_{50pred.}$ com o $pIC_{50exp.}$ Valores de $r^2_{teste} > 0,5$ validam o modelo QSAR.



Figura 15 Processo de validação externa. Um conjunto de moléculas que não faz parte do conjunto de treinamento é usado para validar o modelo

Outra validação importante baseia-se em desorganizar o bloco **Y**, denominado *Y-randomization*. Então, com os valores de $pIC_{50exp.}$ não correspondentes às suas verdadeiras amostras, espera-se que o $r^2_{rand.}$ (randomizado) resulte em valores menores que o $r^2_{Cal.}$, o que demonstra que o modelo QSAR construído não é obra do acaso. Este processo está esquematizado na Figura 16.

Algumas validações adicionais podem ser empregadas. Recentemente, Roy e colaboradores desenvolveram dois cálculos simples e úteis (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2009; 2012). O primeiro refere-se à determinação do coeficiente de correlação ao quadrado modificado, r^2_m , que é empregado para comprovar que os resultados da validação externa não somente exibem boa correlação entre pIC_{50} experimentais e preditos, mas que

também têm valores absolutos similares. O segundo, R^2_p , determina a diferença estatística entre os valores de $r^2_{Cal.}$ e $r^2_{rand.}$ (*Y-randomization*). Valores de r^2_m e R^2_p acima de 0,5 são considerados aceitáveis (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2009; 2012).



Figura 16 Randomização do Bloco Y, seguida da regressão por meio de PLS. O processo é repetido algumas vezes, a fim de avaliar se a verdadeira correlação é obra do acaso ou não

Após construção dos modelos QSAR, novos inibidores para ECA foram propostos a partir da estratégia de combinar as subestruturas mais ativas de cada série, e as atividades biológicas foram previstas utilizando-se os modelos QSAR.

4.4 Docking

Inicialmente, foi feita uma busca pela estrutura cristalográfica da ECA no banco de dados de proteínas (PDB) e utilizou-se, neste estudo, a estrutura

cristalográfica da ECA de código 1UZF, que foi cristalizada e armazenada no banco de dados com a estrutura do ligante, o captopril, fármaco de referência para este estudo.

Os estudos de *docking* foram realizados para as estruturas propostas utilizando-se o programa Molegro Virtual Docker® (MVD) (THOMSEN; CHRISTENSEN, 2006), que é capaz de prever a conformação mais provável de como um ligante se liga a uma macromolécula. A função *score* MolDock (E_{score}) utilizada na MVD® é definida por

$$E_{Score} = E_{Inter} + E_{Intra} \quad (2)$$

em que E_{Inter} é a energia de interação ligante-proteína e E_{Intra} é a energia interna do ligante.

4.5 Avaliações da biodisponibilidade

A "regra dos cinco" de Lipinski: $\log P \leq 5$, peso molecular ≤ 500 , número de aceptores de ligação de hidrogênio (n_{ON}) ≤ 10 e número de doadores de ligação de hidrogênio (n_{OHNH}) ≤ 5 , é útil para avaliar o potencial farmacocinético de candidatos a fármaco; moléculas que violam mais do que uma dessas regras podem ter problemas com biodisponibilidade. Essa regra recebeu esse nome porque os valores de fronteira são 5, 500, 2*5 e 5. Esses parâmetros, em conjunto com a área superficial polar da molécula (TPSA, relacionada com a biodisponibilidade oral), foram avaliados utilizando-se o programa Molinspiration (MOLINSPIRATION CHEMINFORMATICS, 1986). Essas propriedades, principalmente hidrofobicidade, ligações de hidrogênio, tamanho da molécula e flexibilidade e, claro, presença de características

farmacofóricas, influenciam o comportamento da molécula em um organismo vivo e devem ser levadas em consideração para propostas de novas drogas.

Introspecções similares acerca disso podem ser alcançadas utilizando-se o programa Hologram QSAR Technique (www.pkdb.ifsc.usp.br), desenvolvido por Moda et al. (2008). A base de dados consegue avaliar a absorção intestinal humana, a biodisponibilidade oral, a penetração à barreira hemato-encefálica e a solubilidade de novos fármacos.

5 RESULTADOS E DISCUSSÃO

5.1 MIA-QSAR

O MIA-QSAR (FREITAS; BROWN; MARTINS, 2005) foi o método utilizado para obter o modelo QSAR e os descritores foram correlacionados aos dados biológicos correspondentes utilizando-se PLS. Como apontado anteriormente, em MIA-QSAR, os descritores são binários, ou seja, pixels das imagens que correspondem às estruturas químicas bidimensionais. A parte variável entre o conjunto de estruturas químicas (geralmente os substituintes) explica a variância no bloco das bioatividades, enquanto a parte congruente da série é utilizada para o alinhamento 2D.

A qualidade do modelo MIA-QSAR foi avaliada pelos valores de r^2 e q^2 , definidos como

$$1 - [\Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2] \quad (3)$$

em que y_i corresponde aos valores de pIC_{50} experimentais, \hat{y}_i são os valores estimados/preditos de pIC_{50} e \bar{y} corresponde aos valores médios de pIC_{50} ; r^2 superiores a 0,8 e q^2 superiores 0,5 são valores considerados bons para atestar a qualidade de um modelo QSAR.

A etapa de calibração realizada com os 69 compostos resultou em r^2_{Cal} de 0,909 (erro quadrático médio, RMSE, de 0,492), utilizando-se 6 variáveis latentes PLS (Gráfico 2). O modelo QSAR foi validado por validação cruzada *leave-one-out*, resultando em q^2 de 0,607 (RMSE = 1,044). Golbraikh e Tropsha (2002) concluíram que a validação externa é a única maneira de alcançar um modelo QSAR confiável; portanto, as atividades dos compostos do conjunto teste foram previstas utilizando-se o modelo QSAR construído. Os resultados

foram: r^2_{teste} de 0,779 (RMSE = 0,687) e valida o modelo ($r^2_{\text{teste}} > 0,5$ é considerado aceitável). Os valores estão representados na Tabela 4.

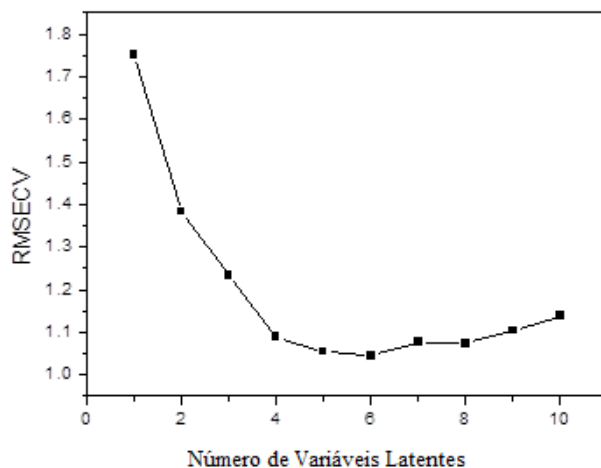


Gráfico 2 Determinação do número ótimo de variáveis latentes PLS, para o modelo MIA-QSAR

Tabela 4 Resultados estatísticos para o modelo MIA-QSAR, utilizando seis variáveis latentes PLS.

Parâmetro	Calibração	LOOCV	Conjunto de teste	Y-randomization ^a
r^2	0,909	0,607	0,779	0,600
RMSE	0,492	1,044	0,687	1,028

^a média de 10 repetições.

Para atribuir ao modelo MIA-QSAR construído maior confiabilidade, uma validação adicional foi utilizada, para avaliar o valor alcançado na validação externa (r^2_{teste}). O r_m^2 (r^2 modificado) foi obtido como definido na Equação 4, proposta por Roy e colaboradores (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2009; 2012),

$$r_m^2 = r^2 [1 - (r^2 - r_0^2)^{1/2}] \quad (4)$$

em que r^2 e r_0^2 correspondem aos valores do coeficiente de correlação ao quadrado entre valores observados e preditos de pIC_{50} para o conjunto de teste com e sem intercepto, respectivamente. O r_m^2 obtido é útil para testar se os valores experimentais e preditos de pIC_{50} são congruentes. O r_m^2 encontrado foi 0,748, confirmando a capacidade do modelo MIA-QSAR em prever novas amostras.

Para garantir que o modelo construído não foi obtido por sorte, aplicou-se o teste *Y-randomization* (Tabela 4), o qual, para uma média de 10 repetições, resultou em $r^2_{rand.}$ de 0,600 (RMSE=1,028). Para um modelo confiável, é desejável que $r^2_{Cal.}$ e $r^2_{rand.}$ sejam significativamente diferentes. Assim, para descartar a hipótese de que o modelo real foi obtido ao acaso e confirmar que $r^2_{Cal.}$ e $r^2_{rand.}$ diferem estatisticamente, utilizou-se um teste de R^2_p , calculado de acordo com a Equação 5 (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2009; 2012),

$$R^2_p = R^2 (R^2 - R^2_{rand})^{1/2} \quad (5)$$

O valor de R^2_p encontrou-se dentro do estabelecido ($R^2_p = 0,505$). Portanto, as bioatividades de novos compostos podem ser confiavelmente

estimadas utilizando-se o modelo MIA-QSAR obtido. A representação gráfica dos resultados do modelo pode ser observada no Gráfico 3.

Após construção do modelo MIA-QSAR, outras informações importantes podem ser extraídas, confirmando o que foi apresentado anteriormente. Na Figura 9 ilustra-se um exemplo de que desenhos de enantiômeros, diferenciados pelas ligações *tracejadas* ou em *cunha* relativas a um centro quiral, podem ser capazes de explicar a variância nas atividades biológicas. Com essa finalidade, duas moléculas do conjunto de dados (Tabela 3) foram utilizadas, os compostos **18** (isômero *R*) e **20** (isômero *S*), com pIC_{50exp} de 4,66 e 2,83, respectivamente.

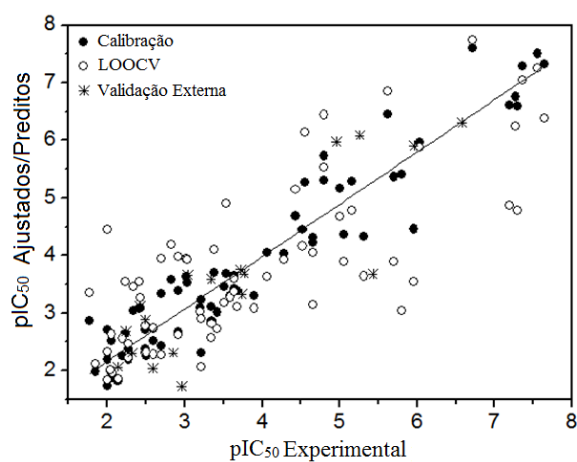


Gráfico 3 pIC_{50} experimentais vs. ajustados/preditos utilizando-se o método MIA-QSAR para a série de peptídeos modificados (inibidores da ECA)

Na Tabela 7 apresentam-se todos os valores de $pIC_{50pred.}$ obtidos pelos modelos construídos no presente trabalho, tornando possível a comparação entre os dois modelos empregados, eleitos como os melhores: MIA-QSAR e QSAR-tradicional com cinco descritores físico-químicos, que serão apresentados nos próximos tópicos. Observa-se que o modelo MIA-QSAR conseguiu diferenciar entre os isômeros *R* e *S*, visto que previu valores de $pIC_{50pred.}$ diferentes para as moléculas **18** ($pIC_{50pred.}=4,31$) e **20** ($pIC_{50pred.}=3,57$), o que não ocorreu com o modelo QSAR-tradicional, que atribuiu os mesmos valores de $pIC_{50pred.}$ para os enantiômeros ($pIC_{50pred.}=3,19$ para os compostos **18** e **20**), uma vez que não considera a estrutura tridimensional da molécula para gerar seus descritores. Portanto, esse exemplo, bem como relatos na literatura, demonstra a qualidade do método MIA-QSAR, que tem apresentado melhores capacidades de modelagem do que análises QSAR tradicionais baseadas em descritores físico-químicos clássicos (ANTUNES; FREITAS; RITTNER, 2008; BITENCOUR; FREITAS, 2009), elegendo-o um método mais confiável para a previsão de novos compostos.

5.2 QSAR-tradicional

O discorrido no tópico anterior demonstra um pouco das limitações da aplicação da metodologia QSAR-tradicional. Contudo, esse fato não invalida as outras vantagens na sua utilização, como uma interpretação química mais profunda sobre os resultados. Então, descritores físico-químicos foram calculados para a série de 87 compostos desenhados e otimizados (conforme descrito na metodologia), utilizando-se o programa Dragon (CONSONNI; PAVAN; TODESCHINI, 2002). Desse processo, 1.459 descritores foram gerados. Como o número de descritores excede o número de amostras, o método

indicado para realizar a correlação dos descritores com os valores de pIC_{50exp} . é o PLS (FERREIRA; MONTANARI; GAUDIO, 2002; GELADI; KOWALSKI, 1986; GOODARZI; FREITAS, 2009). Os mesmos compostos do conjunto de treino e teste (Tabela 3) empregados na construção do modelo MIA-QSAR foram mantidos. O primeiro passo é calcular o número de variáveis latentes. O resultado desse processo está representado no Gráfico 4, no qual se observa-se que existe uma diferença mínima entre os valores de RMSECV (erro quadrático médio validação cruzada), utilizando-se duas ou três variáveis latentes PLS (VL). Então, dois modelos QSAR foram construídos, utilizando-se 2 e 3 VL, resultando nos valores apresentados na Tabela 5.

Os modelos QSAR-tradicional construídos com todos os descritores físico-químicos gerados pelo programa Dragon não apresentaram bons valores de r^2 , com 2 e 3 VL, conforme demonstrado na Tabela 5. Portanto, não são suficientes para atribuir uma grande confiabilidade ao modelo em prever bioatividades de novas amostras. Ainda, utilizando todos os descritores, a interpretação físico-química torna-se impossível, deixando clara a necessidade de selecionar os melhores descritores.

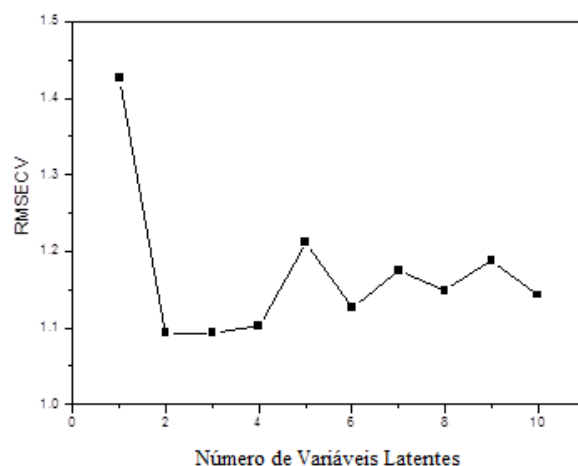


Gráfico 4 Cálculo do número ótimo de variáveis latentes PLS, para o modelo QSAR-tradicional com 1.459 descritores físico-químicos Dragon

Tabela 5 Resultados estatísticos do modelo QSAR-tradicional com 1.459 descritores físico-químicos Dragon, utilizando-se 2 e 3 VL.

VL	Parâmetro	Calibração	LOOCV	Conjunto de teste	<i>Y-randomization</i> ^a
2	r^2	0,607	0,573	0,536	0,014
	RMSE	1,041	1,093	0,999	1,879
3	r^2	0,603	0,559	0,518	0,014
	RMSE	1,032	1,093	0,998	1,776

^a média de 10 repetições.

Para melhorar a interpretabilidade, descritores colineares e aqueles irrelevantes foram removidos, empregando-se métodos de seleção de variáveis, PCA e GA, que foram divididos em etapas, que diferem uma da outra, conforme descrito:

- a) utilizou-se apenas PCA;
- b) PCA seguido de GA;
- c) apenas GA;
- d) GA seguido de PCA;
- e) GA, repetindo o processo consecutivas vezes.

Na etapa ‘a’, apenas PCA foi utilizado. Essa metodologia consegue transformar dados complexos, apresentando-os na forma de gráficos, o que possibilita extrair quais informações são mais importantes para o sistema (FERREIRA; MONTANARI; GAUDIO, 2002). Tratando-se de descritores, conforme apresentado no Gráfico 5, o PCA atribui um melhor entendimento de como os descritores se correlacionam e quais informações contêm.

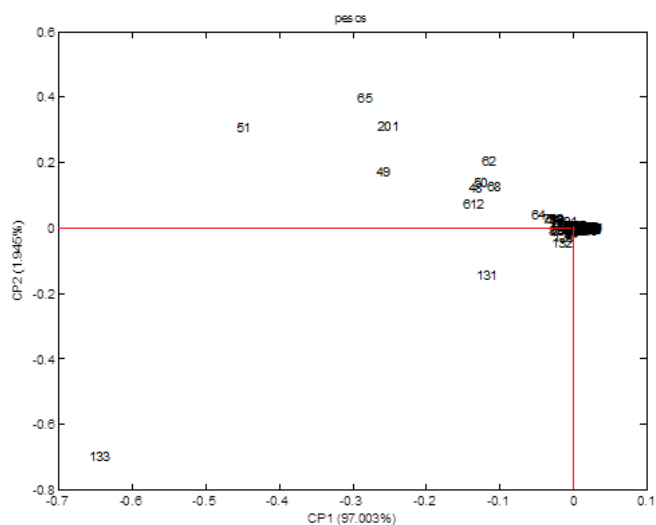


Gráfico 5 Representação gráfica dos resultados do PCA. Gráfico de pesos da componente principal 2 (CP2) eixo Y *versus* componente principal 1 (CP1) eixo X

A CP1 explica 97% do sistema e a CP2 explica 1,94% do sistema, ou seja, quase 100% do comportamento dos descritores físico-químicos obtidos pelo programa Dragon podem ser explicados por essas componentes. No Gráfico 5, a linha vermelha que corta o eixo Y em direção ao eixo X, encontra-se ligada no centro do gráfico, onde se concentra a maior quantidade de descritores. Neste ponto, tanto a CP1 quanto a CP2 não conseguem explicar nada sobre o sistema, visto que o valor onde a linha vermelha teve início corresponde ao valor 0% de explicação da CP1 e CP2. Assim, ou esses descritores explicam a mesma coisa e, portanto, são redundantes (colineares), ou nada explicam. Descritores que se encontram afastados dos pontos 0% podem ser úteis para alcançar uma boa correlação. Nesse sentido, os melhores encontrados foram os descritores localizados nas colunas 48, 49, 50, 51, 62, 65, 68, 131, 133, 201 e 612, correspondentes a SMTI, SMTIV, GMTI, GMTIV, HyDp, ww, Wap, SRW08,

SRW10, IDMT e W3D, respectivamente, totalizando 11 descritores (Matriz **X**), que foram correlacionados com o Bloco **Y** por meio de MLR que, para esta situação, pode ser usado, pois o número de amostras é maior que o número de descritores. Todavia, com 11 descritores, a interpretação físico-química pode ser complicada, pois é desejável que haja menos descritores para a obtenção de uma explicação mais acurada. Os resultados estatísticos alcançados para as validações não foram bons, o que descarta a utilização do modelo QSAR-tradicional construído com 11 descritores Dragon selecionados por PCA. Os resultados dessa etapa (**a**), bem como das próximas etapas, discutidos adiante, estão expressos na Tabela 6.

No intuito de diminuir a quantidade de descritores encontrados na etapa anterior, na etapa '**b**', utilizaram-se os onze descritores selecionados por PCA, para o processo seguinte de seleção de variáveis empregando-se o método GA, o qual, por sua vez, excluiu sete descritores, restando apenas quatro (SMTI, HyDp, SRW08 e SRW10), um bom número para a elaboração de uma equação MLR. Contudo, após regressão, a calibração do modelo obteve um r^2 não satisfatório ($r^2_{Cal.}=0.622$).

Na etapa '**c**', uma nova tentativa foi feita: aplicou-se GA para todos os descritores e, ao final do processo dos 1.459 descritores, 407 foram selecionados pelo método empregado. PLS foi usado para a regressão dos descritores selecionados (Matriz **X**) com as atividades biológicas (Bloco **Y**) e as mesmas validações estatísticas foram utilizadas, resultando nos seguintes valores: $r^2_{Cal.} = 0,873$, $RMSE = 0,582$; $q^2 = 0,729$, $RMSE = 0,862$; $r^2_{test} = 0,674$, $RMSE = 0,832$, $r^2_{rand.}=0,377$, $RMSE=1,286$. O modelo QSAR-tradicional obtido apresentou bons resultados de validação cruzada e validação externa, bem como *Y-randomization*. Esse modelo pode ser comparado com os obtidos anteriormente, por meio da Tabela 6.

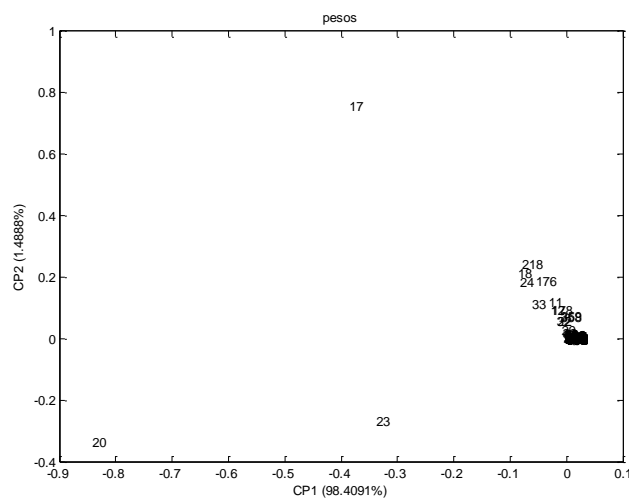


Gráfico 6 Representação gráfica dos resultados do PCA para a etapa ‘d’. A CP1 contém 98,41% e a CP2 contém 1,49% de informação.

A utilização de GA, apenas, não diminuiu o número de descritores a uma quantidade ideal para uma melhor interpretação. Por isso, na etapa ‘d’, os 407 descritores gerados foram submetidos à análise de PCA, que selecionou nove descritores (ZM1V, SMTI, QW, ww, Wap, Whetm, CENT, QZZe e Mor01u), conforme mostrado no Gráfico 6, os quais se encontram identificados pelos números 11, 17, 18, 20, 23, 24, 33, 176 e 218, respectivamente.

Utilizando-se os nove descritores encontrados nessa etapa, efetuou-se uma regressão por MLR com o Bloco Y, após execução das validações cruzada e externa. Os resultados apresentaram-se fora do estabelecido (Tabela 6), sendo esse o pior modelo construído.

Ao que parece, a metodologia PCA não conseguiu extrair as informações mais precisas contidas no conjunto de descritores. Então, uma nova estratégia foi adotada: execuções sucessivas utilizando GA, que constitui um

processo mais lento que PCA, porém, possibilita uma melhor avaliação da evolução de cada descritor.

Na última etapa (**e**), novamente, todos os 1.459 descritores foram utilizados. Na primeira execução de GA, 114 descritores foram selecionados, diferentemente do número de descritores encontrados na etapa '**c**' (407), que em nada diferiu da etapa '**e**'. O fundamento para o funcionamento de GA é o mesmo para as duas etapas; contudo, o processo envolvido na seleção de descritores nesse método é aleatório na primeira geração, etapa em que o método escolhe um conjunto de descritores para iniciar o processo evolutivo, seguido do cruzamento e da mutação. Os descritores selecionados são resultado de um processo baseado na evolução na Terra, em que diversas equações foram construídas avaliando-se os melhores coeficientes de correlação, e selecionando-se os melhores descritores. Por esse motivo, diferentes descritores podem ser obtidos realizando-se o mesmo procedimento, pois cada processo evolutivo em GA inicia-se de maneira diferente, o que explica a diferença entre os números de descritores encontrados nas etapas '**c**' e '**e**'.

No passo seguinte, foi realizada uma segunda execução de GA, com os 114 descritores selecionados, resultando em 46 descritores. Esse processo se repetiu por mais quatro vezes, totalizando seis execuções seguidas de GA e restando apenas cinco descritores físico-químicos, que foram utilizados na construção do modelo MLR, cuja capacidade preditiva foi razoável. Os resultados encontram-se na Tabela 6.

O modelo QSAR-tradicional de 5 parâmetros pode ser descrito como

$$\text{pIC}_{50} = 2,741 - 0,254 \times \text{H-047} - 0,081 \times \text{TPSA}[\text{NO}] + 0,081 \times \text{TPSA}[\text{Tot}] + 0,262 \times \text{F04}[\text{C-C}] - 0,339 \times \text{F05}[\text{N-O}]$$

em que H-47 é um descritor relacionado à composição molecular e conectividades atômicas; TPSA[NO] e TPSA[Tot] são obtidos a partir de fragmentos polares, como o nitrogênio e oxigênio (TPSA[NO]), e de fragmentos polares com nitrogênio e oxigênio mais fragmentos “ligeiramente polares” contendo fósforo e enxofre (TPSA[Tot]); e F04[C-C] e F05[N-O] estão relacionados com as distâncias topológicas entre pares de átomos (C-C e N-O, respectivamente).

Na etapa ‘e’, pode ser observado que o $r^2_{Cal.}$ não atingiu o valor recomendado de 0,8; contudo, os resultados dessa análise podem ser informativos, uma vez que o $r^2_{Cal.}$ obtido foi apenas marginalmente inferior a 0,8, e os resultados de validação foram satisfatórios. Além disso, o $r^2_{Cal.}$ de 0,787 não pareceu ser devido à correlação ao acaso. Isso pode ser confirmado por teste *Y-randomization*, em que o bloco das atividades é aleatorizado e uma regressão subsequente com a matriz de descritores intacta é realizada. Devido ao baixo valor de $r^2_{rand.}$ de 0,068 (média de dez repetições) e mais as validações adicionais que resultaram em valores maiores que 0,5 ($r_m^2 = 0,510$; $R_p^2 = 0,667$), a robustez do modelo original pode ser confirmada, indicando que a correlação não foi fortuita.

Tabela 6 Resultados estatísticos dos modelos QSAR-tradicional e MIA-QSAR (incluído para comparação).

Modelos QSAR	Param.	Calib.	LOOCV	Conjunto de teste	Y-rand. ^a
MIA-QSAR	r^2	0,909	0,604	0,779	0,600
	RMSE	0,492	1,044	0,687	1,028
a) PCA (11 descritores)	r^2	0,602	0,559	0,518	0,014
	RMSE	1,034	1,094	1,000	1,784
b) PCA/AG (4 descritores)	r^2	0,622	0,562	0,552	0,061
	RMSE	1,003	1,082	0,938	1,580
c) AG (407 descritores)	r^2	0,873	0,729	0,674	0,377
	RMSE	0,582	0,862	0,832	1,286
d) AG/PCA (9 descritores)	r^2	0,537	0,463	0,390	0,033
	RMSE	1,113	1,217	1,148	1,650
e) AG/6X (5 descritores)	r^2	0,787	0,742	0,650	0,068
	RMSE	0,753	0,831	0,884	1,574

^a média de 10 repetições. Legenda: Param.= Parâmetro, Calib. = Calibrado.

Além dos índices de conectividade, os valores de pIC_{50} da série de compostos relacionam-se com TPSA que, por sua vez, tem relação com biodisponibilidade oral (ERTL; ROHDE; SELZER, 2000); fragmentos com superfície polar elevada, como aqueles que contêm N e O, diminuem a atividade biológica (sinal negativo na equação MLR), enquanto os fragmentos menos polares contribuem para aumentar os valores de pIC_{50} , tais como aqueles contendo enxofre, que também está presente no captopril. Os descritores de impressões digitais 2D sugerem que distâncias curtas N-O (repulsão dipolar alta ou a possibilidade de ligação de hidrogênio intramolecular $NH \cdots O$) diminuem a bioatividade, enquanto interações hidrofóbicas C-C, ou a presença mais/maiores grupos alquila, aumentam a atividade inibidora da ECA.

Esses resultados validam o modelo QSAR-tradicional construído e os descritores físico-químicos selecionados sugerem uma tendência para o aumento e a diminuição da atividade biológica para os compostos dessa série, demonstrando a utilidade da metodologia empregada e evidenciando a importância do uso de métodos capazes de selecionar variáveis importantes para o modelo QSAR.

Tabela 7 Valores de $pIC_{50exp.}$ usados para a construção do modelo; $pIC_{50pred.}$ e $pIC_{50LOOCV}$, obtidos pelos modelos MIA-QSAR e QSAR-tradicional construído.

Compostos	$pIC_{50exp.}$	MIA-QSAR		Compostos	$pIC_{50exp.}$	MIA-QSAR		GA/6X		GA/6X	
		$pIC_{50pred.}$	LOOCV 6VL			$pIC_{50pred.}$	LOOCV 6VL	$pIC_{50pred.}$	LOOCV	$pIC_{50pred.}$	LOOCV
1^a	6,59	6,31	-	45	1,77	2,87	3,35	5,19	-	2,30	2,35
2	7,19	6,61	4,87	46^a	2,32	2,31	-	5,98	5,90	3,61	-
3	5,16	5,28	4,78	47	2,06	2,52	2,63	5,47	5,51	3,08	3,13
4	4,80	5,73	6,44	48	3,51	3,46	3,19	5,98	6,06	3,60	3,60
5	7,36	7,29	7,05	49	2,49	2,37	2,31	5,98	5,89	2,59	2,59
6^a	4,96	5,98	-	50	3,42	3,01	2,72	5,99	-	3,10	3,09
7	7,30	6,59	4,78	51^a	5,43	3,67	-	6,50	6,43	3,90	-
8	4,43	4,68	5,14	52	5,70	5,36	3,90	5,19	5,22	4,95	4,92
9	3,54	3,69	4,90	53	3,89	3,30	3,08	5,19	5,25	3,34	3,32
10	6,04	5,96	5,88	54	2,92	2,67	2,62	5,45	5,43	2,82	2,81
11^a	3,77	3,68	-	55	3,03	3,53	3,93	5,46	-	3,64	3,68
12	5,06	4,36	3,90	56^a	2,49	2,89	-	6,51	6,68	3,20	-
13	6,72	7,60	7,75	57	2,20	2,25	2,56	6,50	6,48	2,68	2,73
14	7,27	6,76	6,24	58	3,68	3,38	3,11	6,50	6,43	3,12	3,09

...continua

Tabela 7, continuação...

Compostos	pIC ₅₀ exp.	MIA-QSAR		Compostos	pIC ₅₀ exp.	MIA-QSAR		GA/6X		GA/6X	
		pIC ₅₀ pred.	LOOCV 6VL			pIC ₅₀ pred.	LOOCV 6VL	pIC ₅₀ pred.	LOOCV	pIC ₅₀ pred.	LOOCV
15	7,55	7,51	7,26	59	4,52	4,44	4,16	6,50	6,40	4,17	4,15
16 ^a	5,26	6,08	-	60	2,34	3,04	3,46	6,77	-	2,56	2,57
17	4,55	5,26	6,13	61 ^a	2,24	2,68	-	6,51	6,74	2,31	-
18	4,66	4,31	4,05	62	2,42	3,08	3,54	3,19	2,98	2,06	2,04
19	3,21	2,31	2,07	63	2,49	2,70	2,79	2,67	2,59	2,61	2,63
20	2,83	3,57	4,19	64	2,15	1,82	1,85	3,19	3,24	2,50	2,53
21 ^a	2,59	2,04	-	65	3,35	2,87	2,82	2,67	-	2,30	2,21
22	5,31	4,33	3,64	66 ^a	2,85	2,31	-	4,39	4,34	3,61	-
23	2,92	3,39	3,98	67	2,60	2,52	2,28	3,61	3,64	3,08	3,10
24	3,59	3,30	3,28	68	2,27	2,37	2,46	3,87	3,88	2,59	2,60
25	3,02	3,64	3,95	69	2,92	2,67	2,62	4,39	4,47	2,82	2,81
26 ^a	5,96	5,90	-	70	2,51	2,25	2,32	6,23	-	2,68	2,70
27	7,64	7,32	6,38	71 ^a	2,14	2,06	-	6,49	5,92	2,04	-
28	5,96	4,46	3,55	72	3,20	3,10	3,03	6,23	6,36	2,86	2,80
29	5,62	6,45	6,85	73	2,27	2,20	2,22	6,49	6,93	2,16	2,15
30	2,70	3,34	3,95	74	2,04	1,99	2,01	3,12	3,15	1,90	1,88
31 ^a	3,34	3,59	-	75	2,70	2,43	2,28	3,64	-	2,56	2,55
32	2,23	2,65	3,55	76 ^a	3,04	3,65	-	4,17	4,25	3,42	-
33	4,66	4,23	3,14	77	2,43	3,10	3,27	3,64	3,61	2,86	2,92
34	5,80	5,40	3,04	78	2,00	2,20	2,32	4,68	4,65	2,16	2,18
35	3,38	3,70	4,10	79	2,00	2,70	4,45	3,08	3,07	2,68	2,76
36 ^a	2,96	1,72	-	80	1,85	1,99	2,12	3,08	-	1,90	1,91
37	4,28	4,04	3,92	81 ^a	2,42	3,12	-	3,37	3,30	2,42	-
38	2,00	1,74	1,84	82	4,06	4,04	3,63	2,58	2,60	3,64	3,63
39	2,07	1,87	1,96	83	5,00	5,17	4,68	2,06	2,06	4,68	4,68
40	4,80	5,30	5,53	84	3,64	3,42	3,37	4,74	4,71	3,08	3,05

...continua

Tabela 7, continuação...

Compostos	pIC ₅₀ exp.	MIA-QSAR		Compostos	pIC ₅₀ exp.	MIA-QSAR		GA/6X		GA/6X	
		pIC ₅₀ pred.	LOOCV 6VL			pIC ₅₀ pred.	LOOCV 6VL	pIC ₅₀ pred.	LOOCV	pIC ₅₀ pred.	LOOCV
41^a	3,74	3,32	-	85	2,60	2,72	2,74	3,13	-	2,56	2,55
42	3,64	3,65	3,60	86^a	3,72	3,76	-	3,42	3,37	3,37	-
43	3,34	3,11	2,57	87	3,21	3,24	2,90	3,12	3,08	3,07	3,06
44	2,13	1,82	1,86	-	-	-	-	2,50	2,53	-	-

^a Compostos de validação externa. Destacados em vermelho, os valores experimentais e preditos de pIC₅₀ das moléculas 18 e 20, citadas como exemplo na Figura 9.

Na Tabela 6 exibem-se todos os resultados para os modelos QSAR construídos, exceto para o QSAR-tradicional, com todos os descritores físico-químicos calculados utilizando-se o programa Dragon. Comparando-se os resultados obtidos para todos os modelos, constata-se que os melhores foram para o MIA-QSAR e para o QSAR-tradicional com 5 descritores físico-químicos. Contudo, o MIA-QSAR apresentou melhores resultados de r^2 para calibração e validação externa, demonstrando sua melhor capacidade para prever a atividade biológica de novas moléculas. De fato, uma comprovação pode ser feita pela comparação dos valores de pIC₅₀ predito pelos dois modelos QSAR, principalmente pela incapacidade do modelo QSAR-tradicional em diferenciar entre pares de enantiômeros, conforme discutido no texto **5.1** para as moléculas **18** e **20**, que têm seus valores apresentados na Tabela 7, destacados em vermelho.

5.3 Proposta de novos inibidores para ECA

A proposta de novos inibidores foi feita a partir da combinação de subestruturas dos compostos mais ativos biologicamente, presentes em cada série congênere utilizada para a construção do modelo. As moléculas mais ativas de cada série são 15, 27 e 34 (Figura 17) e foram combinadas para construir os compostos **A-E** da Figura 18.

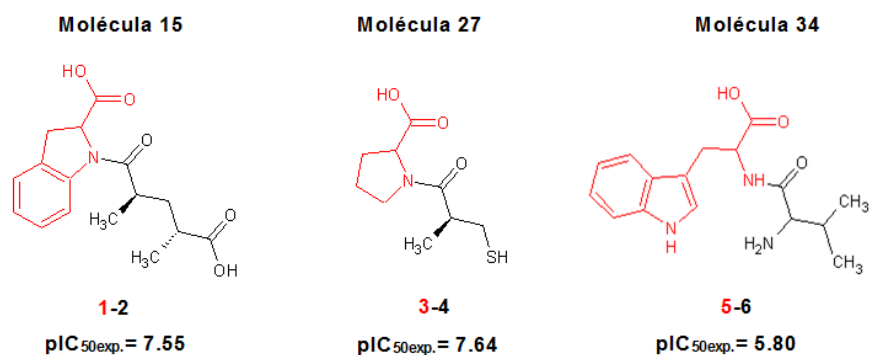


Figura 17 Moléculas mais biologicamente ativas de cada série congênere (**15**, **27** e **34**), escolhidas por meio dos valores de pIC_{50exp.} (Tabela 7). A parte em vermelho de uma molécula foi combinada com a parte preta de outra molécula.

A combinação das moléculas mais biologicamente ativas frente à ECA ocorreu da seguinte forma: **1-4** (molécula **A**), **1-6** (molécula **B**), **3-2** (molécula **C**), **3-6** (presente no conjunto de dados), **5-2** (molécula **D**) e **5-4** (molécula **E**), ou seja, a parte em vermelho da molécula **15** correspondente ao número **1** foi combinada com a parte preta da molécula **27**, que corresponde ao número **4**, e assim por diante, conforme foi descrito. A combinação **3-6** resultou em um

composto presente no conjunto de dados (molécula **35**), por isso não foi usada. O mais importante ao se combinar as subestruturas mais ativas é manter o esqueleto básico comum a todas as estruturas, o que foi respeitado na proposta de novos inibidores, conforme Figura 18.

De fato, novos análogos de dipeptídeos com bioatividades potencialmente altas podem ser concebidos pela combinação de subestruturas de dois ou mais compostos altamente ativos de diferentes séries, mantendo o esqueleto comum inalterado. Surpreendentemente, esta combinação tem mostrado sinergismo, ou seja, alguns compostos obtidos a partir dessa metodologia têm mostrado bioatividades calculadas maiores do que aqueles que os originaram (PINHEIRO et al., 2008; ANTUNES et al., 2008), já que a melhor porção de cada composto foi utilizada.

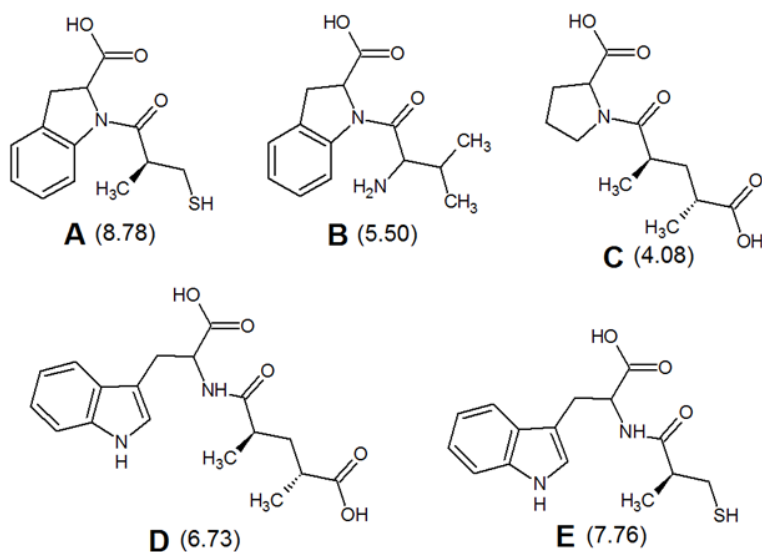


Figura 18 Estruturas propostas usando a estratégia de combinar subestruturas dos compostos mais ativos experimentalmente da série. Os valores de pIC_{50} preditos obtidos a partir do modelo MIA-QSAR são apresentados entre parêntesis

Os valores de pIC_{50} desses análogos de dipeptídeos foram previstos utilizando-se o modelo MIA-QSAR, que foi considerado mais confiável para essa finalidade. Os compostos **A** e **E** exibiram valores de pIC_{50pred} (Figura 18) mais elevados do que o maior valor encontrado dentro dos dados experimentais (Figura 17), indicando o efeito sinérgico mencionado acima. É importante mencionar, também, que a presença de um átomo de enxofre, nas melhores estruturas propostas, tais como no captopril, pode desempenhar um papel importante como uma possível contribuição para a elevada atividade inibidora da ECA.

A abordagem MIA-QSAR afirma que os compostos **A** ($pIC_{50pred}=8,78$) e **E** ($pIC_{50pred}=7,76$) são mais potentes do que o captopril ($pIC_{50pred}=7,32$ e $pIC_{50exp}=7,64$). Na verdade, **A** e **E** já foram sintetizados e testados (KIM et al., 1981; OYA, et al., 1981) e os resultados obtidos confirmam os resultados computacionais. No estudo realizado por Oya et al. (1981), o composto **E** apresentou um valor de pIC_{50exp} de 7,46, comparável com o valor previsto pelo modelo MIA-QSAR.

Em última análise, os compostos propostos podem ser úteis para conduzir a síntese de moléculas semelhantes de dipeptídeos contendo enxofre, ou algum outro grupo capaz de desempenhar o mesmo papel biológico.

5.3.1 Inserção de um novo conjunto de dados

Os estudos realizados até este ponto demonstraram uma tendência do aumento da atividade biológica decorrente da presença de grupos menos polares que o nitrogênio e o oxigênio, como o enxofre, bem como interações hidrofóbicas C-C ou a presença mais/maiores grupos alquila. A correlação entre aumento da atividade biológica e essas características foi prevista pelos descritores físico-químicos Dragon e reafirmada pelas previsões das

bioatividades das novas moléculas propostas, utilizando-se o modelo MIA-QSAR, pelo fato de os compostos mais ativos (**A** e **E**) exibirem essas características estruturais e os estudos experimentais existentes confirmaram essa hipótese.

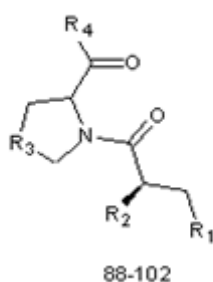
Os melhores modelos QSAR propostos foram validados por esses estudos experimentais. Porém, os melhores inibidores não são novas drogas; por isso, uma nova busca na literatura foi feita, levando-se em consideração as características estruturais da série utilizada neste trabalho. Considerando o fato de que fragmentos ligeiramente polares, tais como calcogênios, podem aumentar a atividade anti-hipertensiva de pequenos peptídeos modificados, uma série de 15 compostos (di e tripeptídeos modificados) derivados de selênio e enxofre recentemente sintetizados (BHUYAN; MUGESH, 2011; 2012) foi adicionada ao modelo, totalizando 102 amostras. A nova série encontra-se na Tabela 8.

O novo conjunto de dados inserido contém dipeptídeos modificados, também derivados de enxofre, mas algumas moléculas são derivadas de selênio. Esse átomo exibe características químicas parecidas com o enxofre (pertence a mesma família química), portanto, espera-se que exerça um papel biológico semelhante, ou, ainda, que apresente um melhor desempenho na inibição da ECA, quando comparado com os dipeptídeos derivados de enxofre. Movidos por essa hipótese, os pesquisadores Bhuyan e Mugesh propuseram e testaram dipeptídeos e tripeptídeos derivados de selênio, análogos aos derivados de enxofre. Com isso, pode-se observar quais desempenham melhor sua função, isto é, inibir a ECA.

Na Tabela 8, os resultados experimentais para esses derivados de selênio comprovam que esses compostos são promissores, visto que apenas a troca de um átomo de enxofre por um átomo de selênio melhorou significativamente o poder de inibição.

Bhuyan e Mugesh (2012) concluíram que a porção necessária para aumentar significativamente a bioatividade desses compostos corresponde ao aminoácido prolina (esqueleto básico exibido no topo da Tabela 8) e que modificações (representadas pelo R₃, Tabela 8) no anel de cinco membros correspondente à pirrolidina nesse aminoácido, junto com um átomo de selênio (representado pelo R₁, Tabela 8), são a chave para a busca por um inibidor mais potente. Concluíram também que aminoácidos como a valina ou a alanina (aminoácidos alifáticos) diminuem a potência de inibição. Entretanto, eles não observaram que a inserção do benzeno à pirrolidina (molécula **A**), bem como a substituição do aminoácido prolina por um aminoácido triptofano (molécula **E**), também contribui para o aumento da atividade biológica.

Tabela 8 Nova série de compostos utilizados nas análises do novo modelo MIA-QSAR.



Composto	R ₁	R ₂	R ₃	R ₄	R ₅	pIC ₅₀ exp.
88 ^b	SeH	Me	-	OH	-	7,44
89	SeH	NH ₂	-	OMe	-	6,47
90	SH	NH ₂	-	OMe	-	5,19
91 ^a	SeH	Me	-	$\begin{array}{c} \text{CH}_2\text{C}_6\text{H}_5 \\ \\ \text{CH}_3\text{O}_2\text{CCHNH}^- \end{array}$	-	5,39

...continua

Tabela 8, continuação

Composto	R ₁	R ₂	R ₃	R ₄	R ₅	pIC ₅₀ exp.
92	SeH	NH ₂	-	$\begin{array}{c} \text{CH}_2\text{C}_6\text{H}_5 \\ \\ \text{CH}_3\text{O}_2\text{CCHNH}^- \end{array}$	-	6,74
93	SH	NH ₂	-	$\begin{array}{c} \text{CH}_2\text{C}_6\text{H}_5 \\ \\ \text{CH}_3\text{O}_2\text{CCHNH}^- \end{array}$	-	5,46
	SeH	NH ₂	-		-	5,06
94				$\begin{array}{c} \text{CH}(\text{CH}_3)_2 \\ \\ \text{CH}_3\text{O}_2\text{CCHNH}^- \end{array}$		
95	SeH	NH ₂	-	$\begin{array}{c} \text{CH}_3 \\ \\ \text{CH}_3\text{O}_2\text{CCHNH}^- \end{array}$	-	4,90
96 ^a	SeH	NH ₂	CH ₂	OMe	-	5,16
97	SH	Me	-	$\begin{array}{c} \text{CH}_2\text{C}_6\text{H}_5 \\ \\ \text{HO}_2\text{CCHNH}^- \end{array}$	-	6,16
98 ^b	SeH	Me	-	$\begin{array}{c} \text{CH}_2\text{C}_6\text{H}_5 \\ \\ \text{HO}_2\text{CCHNH}^- \end{array}$	-	7,12
99	SeH	Me	-	$\begin{array}{c} \text{CH}_2(\text{C}_6\text{H}_4)\text{-4-OH} \\ \\ \text{HO}_2\text{CCHNH}^- \end{array}$	-	5,27
100	SeH	Me	-	$\begin{array}{c} \text{CH}(\text{CH}_3)_2 \\ \\ \text{HO}_2\text{CCHNH}^- \end{array}$	-	5,29
101 ^a	SeH	NH ₂	-	OH	-	6,68
102	SH	NH ₂	-	OH	-	5,64

^a Compostos de validação externa. ^b Os compostos mais ativos de cada uma das duas séries (**88** e **98**) usados para combinar as subestruturas para fornecer os compostos propostos **F-I**.

Fonte: Bhuyan e Mugesh (2011; 2012)

Com essas novas conclusões, um novo modelo MIA-QSAR foi elaborado, de acordo com os passos citados nos tópicos anteriores. A única diferença foi o conjunto teste (21 compostos) e treino (81 compostos) que aumentou. O modelo foi calibrado utilizando 5 VL, de acordo com o Gráfico 7.

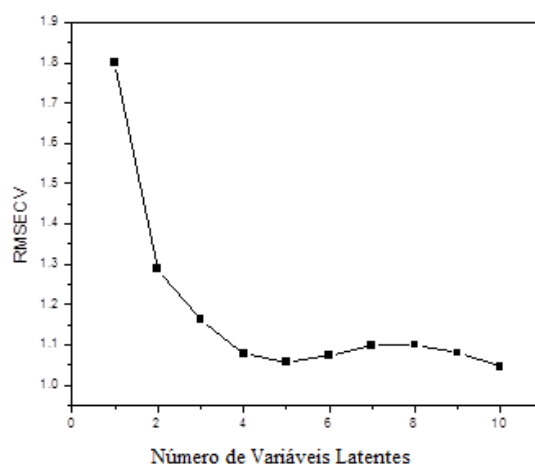


Gráfico 7 Cálculo do número ótimo de variáveis latentes PLS, para o novo modelo MIA-QSAR

Os resultados estatísticos das análises encontram-se na Tabela 9 e atestam a habilidade do novo modelo MIA-QSAR em prever novas amostras.

Tabela 9 Resultados estatísticos do novo modelo MIA-QSAR, utilizando cinco variáveis latentes PLS

Parâmetro	Calibração	LOOCV	Conjunto de teste	<i>Y-randomization</i> ^a
r^2	0,860	0,630	0,548	0,489
RMSE	0,641	1,057	1,102	1,220

^a média de 10 repetições.

No gráfico a seguir exibem-se os resultados dos valores preditos *versus* experimentais de pIC_{50} do novo modelo MIA-QSAR construído.

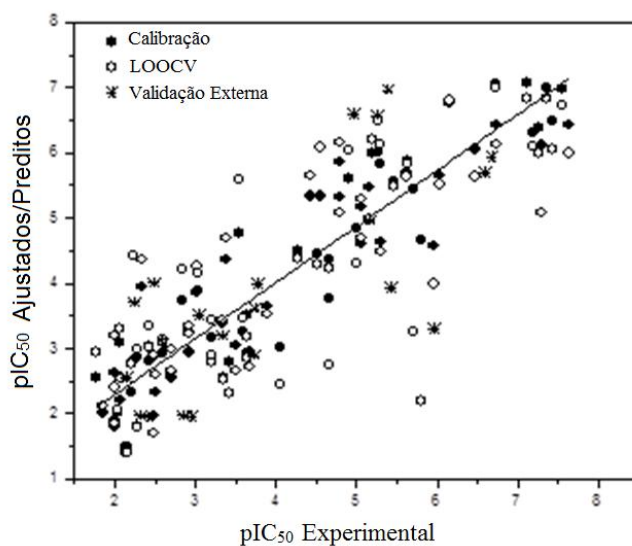


Gráfico 8 pIC_{50} experimentais vs. ajustados/preditos, utilizando-se o método MIA-QSAR, para a nova série de 102 compostos

A próxima etapa foi combinar as subestruturas mais ativas de cada série para propor novos inibidores. As mesmas moléculas (**15**, **27** e **34**) do primeiro conjunto de dados permaneceram e as novas moléculas mais bioativas da série congênere adicionada são **88** e **98**, apresentadas na Figura 19.

De todas as possíveis combinações entre as cinco moléculas, apenas quatro são novas moléculas, resultado da combinação entre: **1-8** (molécula **F**), **5-8** (molécula **G**), **9-2** (molécula **H**) e **9-6** (molécula **I**), totalizando nove moléculas propostas.

O novo modelo MIA-QSAR construído foi empregado para prever a atividade biológica das nove moléculas propostas e o resultado é apresentado na Figura 20.

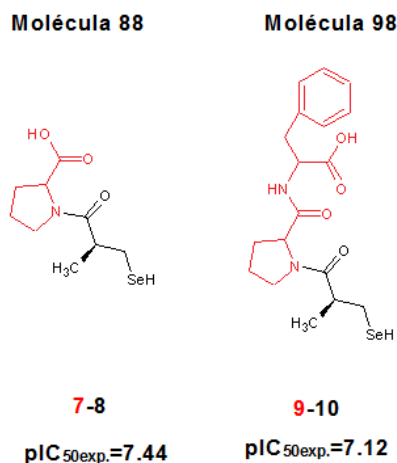


Figura 19 Moléculas mais biologicamente ativas da nova série congênere (**88** e **98**), escolhidas por meio dos valores de pIC_{50exp.} (Tabela 8). A parte em vermelho de uma molécula é combinada com a parte preta de outra molécula

As quatro novas moléculas propostas apresentaram pIC_{50pred.} razoáveis; contudo, uma nova candidata, a molécula **F** (derivado de selênio), exibiu alto valor de pIC_{50pred.} (8,40), sendo este um novo promissor inibidor para ECA.

A molécula **G** não alcançou pIC_{50pred.} muito alto, mas, levando em consideração que o captopril não é o composto mais ativo biologicamente, nem o menos ativo, e que somente elevada atividade biológica não é suficiente para inferir sobre o desempenho que um fármaco, a molécula **G** também poderá ser uma candidata líder na busca por novos inibidores. Por isso, uma investigação com outras metodologias pode formar um conjunto de resultados que esclareça o correto perfil farmacológico das novas estruturas propostas.

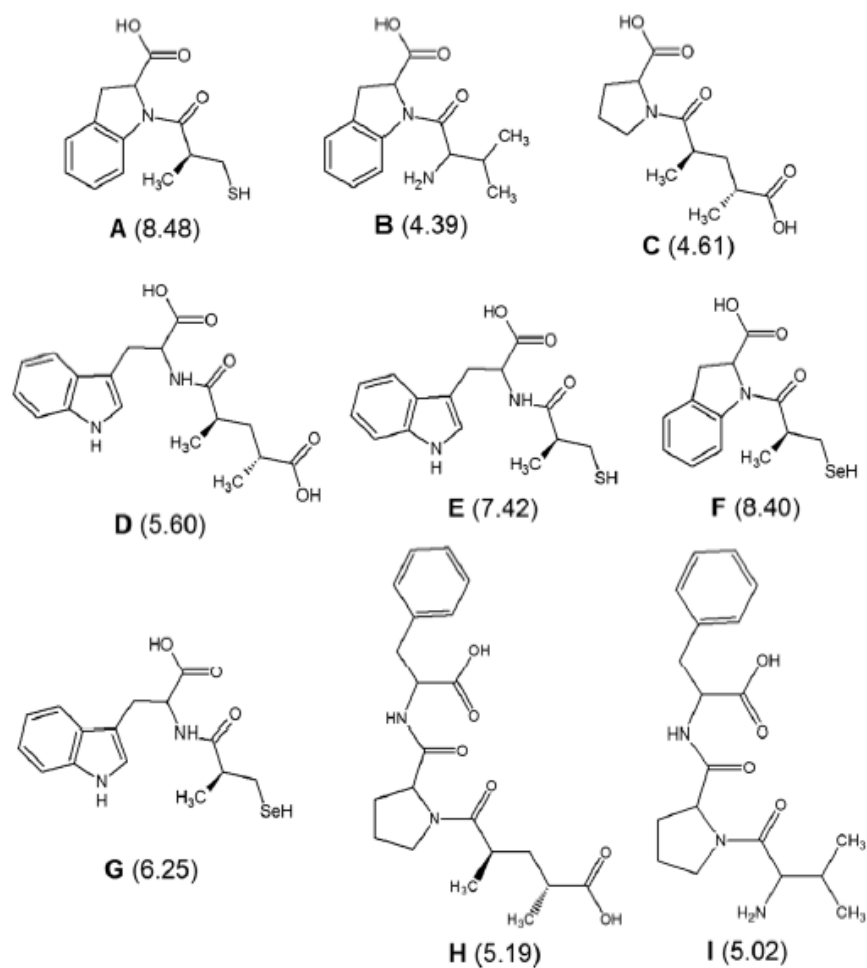


Figura 20 Novas estruturas propostas, utilizando-se a estratégia de combinar subestruturas dos compostos mais ativos experimentalmente da série. Os valores de $pIC_{50pred.}$ obtidos a partir do modelo MIA-QSAR são apresentados entre parêntesis

5.3.2 Ancoramento molecular (*docking*)

Para validar as previsões obtidas pelas análises de QSAR, estudos de *docking* foram realizados para **A-I**. A estrutura cristalina do captopril dentro do sítio ativo da ECA foi obtida a partir do PDB (código PDB: 1UZF) e utilizada para o encaixe e o alinhamento das estruturas propostas **A-I**, usando o software MVD.

Os valores de E_{score} para os compostos propostos correlacionam-se muito bem com os valores de pIC_{50} preditos pelo modelo MIA-QSAR ($r^2=0.823$, Gráfico 9).

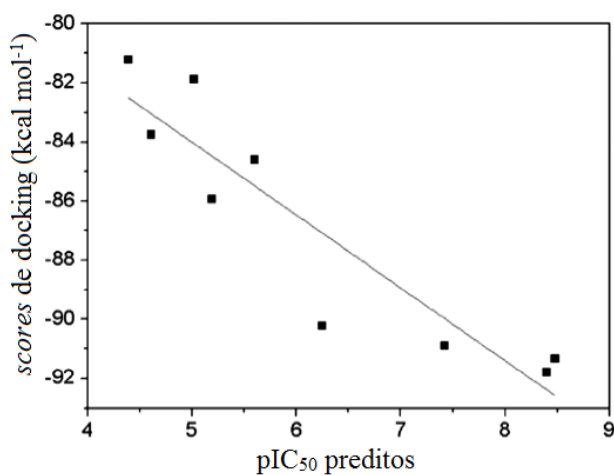


Gráfico 9 Correlação entre valores de pIC_{50} preditos de **A-I** usando MIA-QSAR e os *scores* de docking

Todos os compostos propostos estabelecem ligações de hidrogênio com os resíduos de aminoácido Gln281, His353, Ala354, Lys511 e Tyr520, mas apenas os melhores (**A**, **E**, **F** e **G**) estabeleceram tais interações entre o átomo de enxofre (**A** e **E**) e Glu385, assim como aqueles contendo o átomo de selênio (**F** e **G**). Na Figura 21 observa-se a sobreposição da melhor conformação do composto **F** no sítio ativo da ECA, juntamente com a estrutura do captopril obtido a partir da estrutura cristalina (PDB código 1UZF; *docking score* do captopril= $-65,5 \text{ kcal mol}^{-1}$). Na Figura 22, apenas o composto **F** é mostrado.

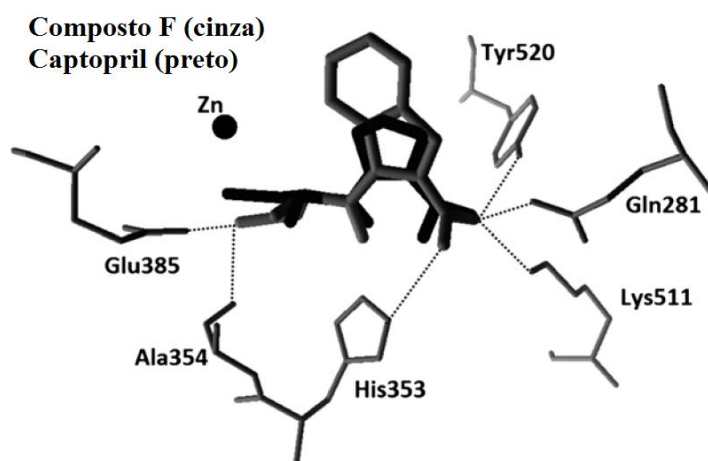


Figura 21 Principais interações (ligações de hidrogênio) de **F** (sobreposto ao captopril) com o sítio ativo da ECA

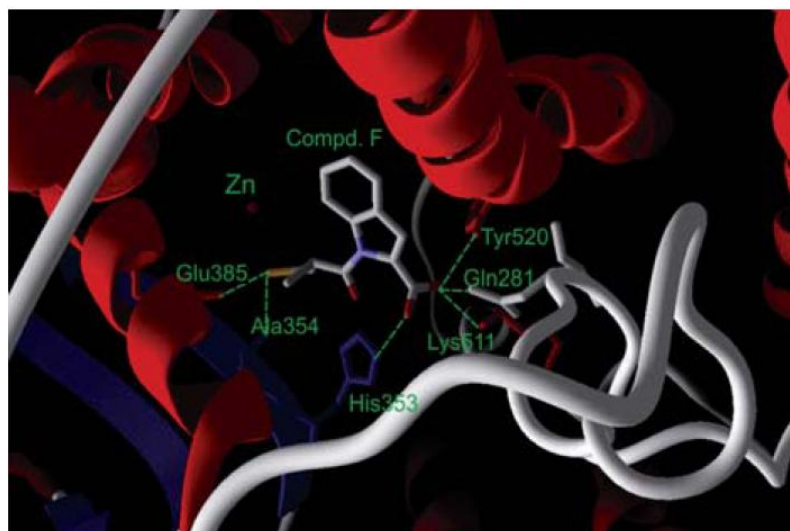


Figura 22 Principais interações (ligações de hidrogênio) de **F** com o sítio ativo da ECA. Legenda: Compd.= Composto

Para testar se pequenas alterações nos melhores compostos (**A** e **F**) implicaria em melhoras nas bioatividades, uma ligação dupla no anel de cinco membros foi acrescentada na porção correspondente ao aminoácido prolina (esqueleto básico exibido no topo da Tabela 8), resultando em duas novas moléculas **J** e **K**, que contêm uma porção indólica (Figura 23).

As estruturas modificadas (**J** e **K**) foram docadas dentro do sítio ativo da ECA (mesmo procedimento realizado para as outras moléculas) e foi observado um aumento nos valores de *docking score* (**J** = -93,2 kcal mol⁻¹ e **K**= -93,6 kcal mol⁻¹).

A análise de *docking* demonstrou a influência dos átomos de enxofre e selênio, que aumentam o poder de inibição, pois realizam uma ligação de hidrogênio a mais (com a Glu385) do que as outras moléculas, validando as metodologias QSAR empregadas. A última modificação nas estruturas

comprovou a hipótese investigada por Bhuyan e Mugesh, (2012): de fato, a parte indólica dessas moléculas pode exercer influência na atividade biológica em conjunto com os átomos de enxofre/selênio, validando os estudos computacionais utilizados neste trabalho.

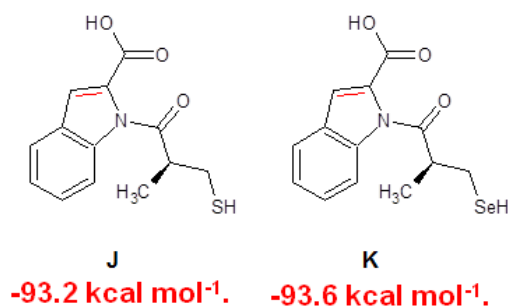


Figura 23 Novas estruturas propostas **J** e **K**; uma dupla ligação foi acrescentada nos compostos **A** e **F** (destacada em vermelho), resultando em um indol modificado. Os valores de *docking score* estão descritos abaixo das moléculas

5.3.3 Avaliações da biodisponibilidade

Ao que tudo indica, as novas moléculas propostas são potenciais inibidores da ECA, de acordo com os resultados alcançados. No entanto, além de fármacos altamente potentes, é desejável que os inibidores propostos sejam absorvidos adequadamente. Essa propriedade foi avaliada utilizando-se o programa Molinspiration (www.molinspiration.com) que, por meio da "regra dos cinco" de Lipinski, permite avaliar os parâmetros capazes de prever se novas drogas apresentariam biodisponibilidade adequada.

Tabela 10 Parâmetros calculados da "regra dos cinco" de Lipinski e TPSA para os compostos propostos (**A-K**), e para o composto de referência captopril (**27**).

Compostos	pIC ₅₀	logP	TPSA	MW	n _{ON}	n _{OHNH}	n _{rotb}	n _{viol}
Captopril (27)	7,64 ^a	-1,09	57,6	217,3	4	1	3	0
A	8,48 ^b	0,20	57,6	265,3	4	1	3	0
B	4,39 ^b	-0,88	83,6	262,3	5	3	3	0
C	4,61 ^b	-1,15	94,9	257,3	6	2	5	0
D	5,60 ^b	0,16	119,5	346,4	7	4	8	0
E	7,42 ^b	0,22	82,2	306,4	5	3	6	0
F	8,40 ^b	-0,09	57,6	312,2	4	1	3	0
G	6,25 ^b	-0,07	82,2	353,3	5	3	6	0
H	5,19 ^b	-0,25	124,0	404,5	8	3	9	0
I	5,02 ^b	-1,27	112,7	361,4	7	4	7	0
J	-	2,98	59,3	263,3	4	1	3	0
K	-	2,69	59,3	310,2	4	1	3	0

^a Dados experimentais. ^b Valores preditos usando MIA-QSAR.

Conforme se observa na Tabela 10, nenhum dos compostos propostos viola a "regra dos cinco" de Lipinski (n_{viol}=0, última coluna), indicando não ter qualquer problema com absorção.

Valores de TPSA iguais ao captopril foram encontrados para a nova estrutura proposta (**F**) e pouca discrepância para **J** e **K**, demonstrando provável habilidade de absorção, comparável ao captopril.

Tabela 11 Parâmetros farmacocinéticos computados para os compostos propostos (**A-K**), e para o composto de referência captopril (**27**).^a

Compostos	HIA (%)	F (%)	BBB (logBB)	Solubilidade (logS) ^b
Captopril (27)	78,73 (84 ^c)	57,94 (70 ^d)	-0,18	-1,02
A	83,23	61,04	-0,36	-2,80
B	86,33	59,14	-0,18	-2,84
C	82,76	73,19	-0,52	-0,97
D	75,85	75,62	-1,21	-2,99
E	75,91	75,64	-1,04	-3,04
F	75,75	59,92	0,54	-1,81
G	91,24	79,10	0,45	-1,99
H	97,38	68,39	0,56	-1,79
I	99,14	69,44	0,58	-1,33
J	98,71	66,85	-0,10	-2,94
K	100	64,51	0,17	-3,01
Error	+/- 13	+/- 15	+/- 0,30	+/- 0,82

^a HIA = absorção intestinal humana; F = biodisponibilidade oral; BBB = penetração à barreira hemato-encefálica. ^b Para obter a solubilidade em g L⁻¹, multiplicar logS pelo peso molecular do composto. ^c (GUNTURI; NARAYANAN, 2007). ^d (DUCHIN, et al., 1988).

Outro descritor, que avalia a biodisponibilidade usando uma medida de flexibilidade molecular (n_{rotb}), atribuiu às melhores moléculas propostas (**F**, **J** e **K**) os mesmos valores do captopril.

Avaliações similares às descritas acima, porém mais diretas, foram feitas utilizando-se o banco de dados PK/DB disponível no site: www.pkdb.ifsc.usp.br, cujos dados farmacocinéticos são retratados na Tabela 11.

As estruturas propostas exibiram propriedades farmacocinéticas similares ou melhores que as do captopril, de acordo com os resultados da Tabela 11. Além disso, dados experimentais do captopril (DUCHIN et al., 1988; GUNTURI; NARAYANAN, 2007) para HIA (84) e *F* (70) são comparáveis aos obtidos pela metodologia usada no PK/DB (HQSAR), validando os resultados teóricos e, mais uma vez, reafirmando que as moléculas propostas apresentam potenciais propriedades farmacocinéticas para uma adequada absorção, distribuição, metabolismo e excreção.

6 CONCLUSÃO

Dipeptídeos modificados foram propostos utilizando-se uma série congênere de moléculas com o auxílio de técnicas computacionais, em que as moléculas que contêm um grupo SH ou SeH foram estimadas serem mais promissoras do que as outras concebidas a partir da combinação de subestruturas químicas sem esses grupos. Modificações estruturais no anel de cinco membros (adição da dupla ligação) exibiram melhoras nos compostos propostos, juntamente com os átomos de S/Se. Isso foi confirmado por estudos de *docking* e pelos dados experimentais existentes para os compostos mais ativos **A** e **E**.

Este suporte teórico, em conjunto com a capacidade de modelagem e informações químicas fornecidas pelos modelos QSAR, pode conduzir à síntese de novos (e mais potentes) inibidores da ECA.

REFERÊNCIAS

ADVANCED CHEMISTRY DEVELOPMENT. **ACD/ChemSketch** - version 12.01. Toronto, 2009. Disponível em: <http://alainsheratan.com/chemsk_t12.pdf>. Acesso em: 4 ago. 2012.

ANTUNES, J. E.; FREITAS, M. P.; RITTNER, R. Bioactivities of a series of phosphodiesterase type 5 (PDE-5) inhibitors as modelled by MIA-QSAR. **European Journal of Medicinal Chemistry**, v. 43, n. 8, p. 1632-1638, Aug. 2008.

BAKRIS, G. L. Is blockade of the renin-angiotensin system appropriate for all patients with diabetes? **Journal of the American Society of Hypertension**, v. 3, n. 5, p. 288-290, Oct. 2009.

BARREIRO, E. J. Estratégia de simplificação molecular no planejamento racional de fármacos: a descoberta de novo agente cardioativo. **Química Nova**, São Paulo, v. 25, n. 6b, p. 1172-1180, dec. 2002.

BARREIRO, E. J.; FRAGA, A. M. F. **Química medicinal**: as bases moleculares da ação dos fármacos. Porto Alegre: Artmed, 2008. n. 2, p. 150-151.

BERNDT, E. R. et al. Industry funding of the FDA: effects of PDUFA on approval times and withdrawal rates. **Nature Reviews Drug Discovery**, v. 4, n. 7, p. 545-554, July 2005.

BHUYAN, B. J.; MUGESH, G. Antioxidant activity of peptide-based angiotensin converting enzyme inhibitors. **Organic & Biomolecular Chemistry**, v. 10, n. 11, p. 2237-2247, Feb. 2012.

BHUYAN, B. J.; MUGESH, G. Effect of peptide-based captopril analogues on angiotensin converting enzyme activity and peroxynitrite-mediated tyrosine nitration. **Organic & Biomolecular Chemistry**, v. 9, n. 14, p. 5185-5192, May 2011.

BITENCOURT, M.; FREITAS, M. P. Bi- and multilinear PLS Coupled to MIA-QSAR in the prediction of antifungal activities of some Benzothiazole Derivatives. **Medicinal Chemistry**, v. 5, n. 1, p. 79-86, Jan. 2009.

BROMIDGE, S. M. et al. Novel and selective 5-HT_{2C/2B} receptor antagonists as potential anxiolytic agents: Synthesis, quantitative structure-activity relationships, and molecular modeling of substituted 1-(3-pyridylcarbamoyl)indolines. **Journal of Medicinal Chemistry**, v. 41, n. 10, p. 1598-1612, May 1998.

BROWN, R. D.; MARTIN, Y. C. The information content of 2D and 3D structural descriptors relevant to ligand-receptor binding. **Journal of Chemical Information and Computational Science**, v. 37, n. 1, p. 1-9, Jan. 1997.

CORMANICH, R. A.; GOODARZI, M.; FREITAS, M. P. Improvement of multivariate image analysis applied to quantitative structure-activity relationship (QSAR) analysis by using wavelet-principal component analysis ranking variable selection and least-squares support vector machine regression: QSAR study of checkpoint kinase WEE1 inhibitors. **Chemical Biology & Drug Design**, v. 73, n. 2, p. 244-252, Feb. 2009.

CORMANICH, R. A.; NUNES, C. A. FREITAS, M. P. Desenhos de estruturas químicas correlacionam-se com propriedades biológicas: MIA-QSAR. **Química Nova**, v. 35, n. 6, p. 1157-1163, Apr. 2012.

CRAMER, R. D.; PARTTERSON, D. E.; BUNCE, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. **Journal of the American Chemical Society**, v. 110, n.18, p. 5959-5967, Aug. 1988.

CUSHMAN, D. W. et al. Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and Mercaptoalkanoyl Amino Acids. **Biochemistry**, v. 16, n. 25, p. 5484-5491, May 1977.

DATASUS - **Sistema Único de Saúde**. 2005. Disponível em: <www.datasus.gov.br/>. Acesso em: 28 jan. 2013.

DIMASI, J. A.; HANSEN, R. W.; GRABOWSKI, H. G. The price of innovation: new estimates of drug development costs. **Journal of Health Economics**, v. 22, n. 2, p. 151-185, Mar. 2003.

DUCHIN, K. L. et al. Pharmacokinetics of captopril in healthy-subjects and in patients with cardiovascular-diseases. **Clinical Pharmacokinetics**, v. 14, n. 4, p. 241-259, Apr. 1988.

ERIKSSON, U.; DANILCZYK, U.; PENNINGER, J. M. Just the beginning: Novel functions for angiotensin-converting enzymes. **Current Biology**, v. 12, n. 21, p. 745-752, Oct. 2002.

ERTL, P.; ROHDE, B.; SELZER, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. **Journal of Medicinal Chemistry**, v. 43, n. 20, p. 3714-3717, Oct 2000.

ESTRADA, E.; MOLINA, E.; PERDOMO-LÓPEZ, I. Can 3D Structural Parameters Be Predicted from 2D (Topological) Molecular Descriptors? **Journal of Chemical Information and Computational Science**, v. 41, n. 4, p. 1015-1021, Jun. 2001.

FERNANDEZ, J. H. et al. Structural basis of the lisinopril-binding specificity in N- and C-domains of human somatic ACE. **Biochemical and Biophysical Research Communications**, v. 308, n. 2, p. 219-226, Aug. 2003.

FERREIRA, M. M. C.; MONTANARI, C. A.; GAUDIO, A. C. Seleção de variáveis em QSAR. **Química Nova**, São Paulo, v. 25, n. 3, p. 439-448, May 2002.

FREITAS, M. P.; BROWN, S. D.; MARTINS, J. A. MIA-QSAR: a simple 2D image-based approach for quantitative structure-activity relationship analysis. **Journal of Molecular Structure**, v. 738, n. 1-3, p. 149-154, Mar. 2005.

FREITAS, M. P. MIA-QSAR modelling of anti-HIV-1 activities of some 2-amino-6-arylsulfonylbenzotrioles and their thio and sulfinyl congeners. **Organic & Biomolecular Chemistry**, v. 4, n. 6, p. 1154-1159, Mar. 2006.

FREITAS, M. P.; Multivariate image analysis applied to QSAR: Evaluation to a series of potential anxiolytic agents. **Chemometrics and Intelligent Laboratory Systems**, v. 91, n. 2, p. 173-176, Apr. 2008.

FREITAS, M. P. Multivariate QSAR: From classical descriptors to new perspectives. **Current Computer-Aided Drug Design**, v. 3, n. 4, p. 235-239, Apr. 2007.

GARGIULO, R. J.; GIOTTA, G. J.; WANG, H. H. Spin-labeled analogs of local anesthetics. **Journal of Medicinal Chemistry**, v. 16, n. 6, p. 707-708, Jun. 1973.

GAUDIO, A. C.; ZANDONADE, E. Proposição, validação e análise dos modelos que correlacionam estrutura química e atividade biológica. **Química Nova**, São Paulo, v. 24, n. 5, Oct. 2001.

GELADI, P.; KOWALSKI, B. R. Partial least-squares regression - a tutorial. **Analytica Chimica Acta**, v. 185, p. 1-17, Jul. 1986.

GOESSLER, K. F.; POLITO, M. D. Inibidor direto da renina e hipertensão arterial: uma revisão. **Revista Brasileira de Cardiologia**, v. 25, n. 3, p. 241-255, jun. 2012.

GOLBRAIKH, A.; TROPSHA, A. Beware of $q(2)$! **Journal of Molecular Graphics & Modelling**, v. 20, n. 4, p. 269-276, Jan. 2002.

GONÇALVES, R. B. Laboratório Nacional de Computação. 2008. Disponível em: <http://www.gmmsb.lncc.br/pdf/T_Goncalves2008.pdf> Acesso em: 4 fev. 2013.

GOODARZI, M. et al. QSAR and docking studies of novel antileishmanial diaryl sulfides and sulfonamides. **European Journal of Medicinal Chemistry**, v. 45, n. 11, p. 4879-4889, Aug. 2010.

GOODARZI, M.; FREITAS, M. P. Augmented three-mode MIA-QSAR Modeling for a series of anti-HIV-1 compounds. **Qsar & Combinatorial Science**, v. 27, n. 9, p. 1092-1097, Sept. 2008.

GOODARZI, M.; FREITAS, M. P.; FERREIRA, E. B. Influence of changes in 2-D chemical structure drawings and image formats on the prediction of biological properties using MIA-QSAR. **QSAR & Combinatorial Science**, v. 28, n. 4, p. 458-464, Feb. 2009.

GOODARZI, M.; FREITAS, M. P.; JENSEN, R. Feature selection and linear/nonlinear regression methods for the accurate prediction of glycogen synthase kinase-3 beta inhibitory activities. **Journal of Chemical Information and Modeling**, v. 49, n. 4, p. 824-832, Apr. 2009.

GOODARZI, M.; FREITAS, M. P. On the use of PLS and N-PLS in MIA-QSAR: Azole antifungals. **Chemometrics and Intelligent Laboratory Systems**, v. 96, n. 1, p. 59-62, Mar. 2009.

GOODARZI, M.; FREITAS, M. P. Predicting boiling points of aliphatic alcohols through multivariate image analysis applied to quantitative structure-property relationships. **Journal of Physical Chemistry A**, v. 112, n. 44, p. 11263-11265, Nov. 2008.

GOOGLE BODY BROWSER: e-books e cultura. **Corpo humano**. Disponível em: <<http://www.zygotebody.com>>. Acesso em: 12 ago. 2012.

GRUENFELD, N. et al. Angiotensin converting enzyme inhibitors: 1-glutarylindoline-2-carboxylic acid derivatives. **Journal of Medicinal Chemistry**, v. 26, n. 9, p. 1277-1282, Nov. 1983.

GUNTURI, S. B.; NARAYANAN, R. In silico ADME modeling 3: computational models to predict human intestinal absorption using sphere exclusion and kNN QSAR methods. **Qsar & Combinatorial Science**, v. 26, n. 5, p. 653-668, May. 2007.

HAMMETT, L. P. The effect of structure upon the reactions of organic compounds. benzene derivatives. **Journal of the American Chemical Society**, v. 59, p. 96-103, Jan. 1937.

HANSCH, C. et al. Chem-bioinformatics: Comparative QSAR at the interface between chemistry and biology. **Chemical Reviews**, v. 102, n. 3, p. 783-812, Mar. 2002.

HANSCH, C.; FUJITA, T. rho-sigma-pi-analysis - A method for the correlation of biological-activity and chemical-structure. **Journal of the American Chemical Society**, v. 86, n. 47, p. 1616-1626, Nov. 1964.

HANSCH, C.; LIEN, E. J. An analysis of structure-activity relationship in adrenergic blocking activity of beta-haloalkylamines. **Biochemical Pharmacology**, v. 17, n. 5, p. 709-720, May 1968.

HOLLAND, J. H. **Adaptation in natural and artificial systems**. University of Michigan, 1975. Disponível em: <http://pt.wikipedia.org/wiki/John_Henry_Holland>. Acesso em: 4 fev. 2013.

HOPFINGER, A. J. et al. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. **Journal of the American Chemical Society**, v. 119, n. 43, p. 10509-10524, Oct. 1997.

HYPERCUBE. HyperChem 8.0.8. Gainesville, 2007. Disponível em: <<http://www.hyper.com/Products/HyperChemProfessional/75Overview/tabid/361/Default.aspx>> Acesso em: 5 de nov. 2011.

KIM, D. H. et al. **US Patent Number US4303583, 1981**. Disponível em: <<http://www.google.com/patents/US4303583>>. Acesso em: 20 jul. 2012.

KLEBE, G.; ABRAHAM, U.; MEITZNER, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. **Journal of Medicinal Chemistry**, v. 37, n. 24, p. 4130-4146, Nov. 1994.

KOHLMANN, J. R. et al. III Consenso Brasileiro de Hipertensão Arterial. **Arquivos Brasileiros de Endocrinologia & Metabologia**, v. 43, n. 4, p. 257-286, Aug. 1999.

KUBINYI, H. Drug research: myths, hype and reality. **Nature Reviews Drug Discovery**, v. 2, n. 8, p. 665-668, Aug. 2003.

KURUP, A. et al. HIV-1 protease inhibitors: a comparative QSAR analysis. **Current Medicinal Chemistry**, v. 10, n. 17, p. 1679-1688, Sept. 2003.

KURUP, A.; GARG, R.; HANSCH, C. Comparative QSAR study of tyrosine kinase inhibitors. **Chemical Reviews**, v. 101, n. 8, p. 2573-2600, Aug. 2001.

LIMA, L. M. Química medicinal moderna: desafios e contribuição brasileira. **Química Nova**, São Paulo, v. 30, n. 6, Dec. 2007.

LOMBARDINO, J. G.; LOWE, J. A. The role of the medicinal chemist in drug discovery - then and now. **Nature Reviews Drug Discovery**, v. 3, n. 10, p. 853-862, Oct. 2004.

LOWIS, D. R. HQSAR: A new, highly predictive QSAR technique. **Tripes Technical Notes**, v. 1, n. 5, p. 1-17. Oct. 1997.

MARTINS, J. P. A. et al. LQTA-QSAR: A New 4D-QSAR Methodology. **Journal of Chemical Information and Modeling**, v. 49, n. 6, p. 1428-1436, Jun. 2009.

MARTINS, J. P. A. **Química computacional aplicada a QSAR**. mar. 2010
Disponível em:
<http://www.cenapad.unicamp.br/servicos/treinamentos/apostilas/apostila_QuimComp.pdf>. Acesso em: 4 nov. 2011.

MATHWORKS. **MatLab, Version 7.5**. Natick, MA, 2005. Disponível em:
<<http://www.mathworks.com/products/matlab/whatsnew.html>>. Acesso em: 20 jul. 2011.

MITRA, I.; SAHA, A.; ROY, K. Exploring quantitative structure–activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. **Molecular Simulation**, v. 36, n. 36, p. 1067-1079, Apr. 2010.

MODA, T. L. **Desenvolvimento de modelos *In Silico* de propriedades de ADME para a triagem de novos candidatos a fármacos**. Universidade de São Paulo, São Carlos, Fev. 2007. <Disponível em:
<<http://www.teses.usp.br/teses/disponiveis/76/76132/tde-22032007-112055/en.php>>. Acesso em: 7 nov. 2011.

MODA, T. L. et al. PK/DB: database for pharmacokinetic properties and predictive in silico ADME models. **Bioinformatics**, v. 24, n. 19, p. 2270-2271, Oct. 2008.

MOLINSPIRATION CHEMINFORMATICS. **Molinspiration**. Eslováquia, Nova Úlica, 1986. Disponível em: <<http://www.molinspiration.com/>>. Acesso em: 4 ago. 2012.

NATESH, R. et al. Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. **Biochemistry**, v. 43, n. 27, p. 8718-8724, Jul. 2004.

NEVES, P. J. et al. TOP - um programa de cálculo de descritores topológicos para uso em correlações entre estrutura e atividade. **Química Nova**, São Paulo, v. 21, n. 6, nov. 1998.

NILSSON, J. et al. A multiway 3D QSAR analysis of a series of (S)-N-(1-ethyl-2-pyrrolidinyl)methyl -6-methoxybenzamides. **Journal of Computer-Aided Molecular Design**, v. 12, n. 1, p. 81-93, Jan. 1998.

OJHA, P. K. et al. Further exploring r_m^2 metrics for validation of QSPR models. **Chemometrics and Intelligent Laboratory Systems**, v. 107, p. 194-205, Jan. 2011.

OLEG, U. et al. Qsar study on dipeptide ace inhibitors. **Carpathian Journal of Mathematics**, v. 20, n. 2, p. 275-280, Jan. 2004.

OPREA, T. I. et al. Is there a difference between leads and drugs? A historical perspective. **Journal of Chemical Information and Computer Sciences**, v. 41, n. 5, p. 1308-1315, Oct. 2001.

ORGANIZAÇÃO MUNDIAL DE SAÚDE. **HIV/AIDS**. 2010. Disponível em: <http://www.who.int/cardiovascular_diseases/en/index.html>. Acesso em: 28 jan. 2013.

ORGANIZAÇÃO MUNDIAL DE SAÚDE. **World Heart Day 2012**. 2012. Disponível em: <http://www.who.int/cardiovascular_diseases/en/index.html>. Acesso em: 28 Jan. 2013.

OYA, M. et al. Thiol compounds. II.1) Synthesis and antihypertensive activity of mercaptoacylamino acids. **Chemical and Pharmaceutical Bulletin**, v. 29, n. 4, p. 940-947, Apr. 1981.

PAUL, M.; MEHR, A. P.; KREUTZ, R. Physiology of local renin-angiotensin systems. **Physiological Reviews**, v. 86, n. 3, p. 747-803, Jul. 2006.

PAULA, W. X. A química inorgânica no planejamento de fármacos usados no controle da hipertensão. **Cadernos Temáticos de Química Nova na Escola**, v. 1, n. 6, p. 19-23, jul. 2005.

PEARSON, K. On lines and planes of closest fit to systems of points in space. **Philosophical Magazine**, v. 2, n. 11, p. 559-572, 1901.

PIMENTA, E. Hidroclorotiazida x clortalidona: os diuréticos tiazídicos são todos iguais? **Revista Brasileira de Hipertensão**, v.15, n. 3, p. 166-167, Jun. 2008.

PINHEIRO, J. R. et al. Novel anti-HIV cyclotriazadisulfonamide derivatives as modeled by ligand- and receptor-based approaches. **Bioorganic & Medicinal Chemistry**, v. 16, n. 4, p. 1683-1690, Feb. 2008.

POULTER, N. R. A new dimension in hypertension management with the amlodipine/perindopril combination. **Journal of Hypertension**, v. 29, p. 15-21, Dec. 2011.

REES, P. Big pharma learns how to love information technology. **Scientific Computing World**, July/Aug. p. 16-18, 2003. Disponível em: <http://www.scientific-computing.com/features/feature.php?feature_id=111>. Acesso em: 28 jan. 2013.

RIBEIRO, A. B. Inibidores da renina – uma nova classe de hipertensivos. **Revista Brasileira de Hipertensão**, v. 13, n. 3, p. 219-220, jul. 2006.

ROY, K. et al. Comparative Studies on Some Metrics for External Validation of QSPR Models. **Journal of Chemical Information and Modeling**, v. 52, n. 2, p. 396-408, Feb. 2012.

ROY, P. P. et al. On two novel parameters for validation of predictive QSAR Models. *Molecules*, 2009, 14, 1660-170. **Molecules**, v. 15, p. 604-605, Jan. 2010.

ILVA, D. G.; FREITAS, M. P. MIA-QSAR modeling of the anti-HIV-1 protease activities and cytotoxicities of Ritonavir analogues. **Current Drug Safety**, v. 7, n. 4, p. 269-281, Sept. 2012.

SILVA, M. R. E.; BERALDO, W. T.; ROSENFELD, G. Bradykinin, a hypotensive and smooth muscle stimulating factor released from plasma globulin by snake venoms and by trypsin. **American Journal of Physiology**, v. 156, n. 2, p. 261-273, Feb. 1949.

SKEGGS Jr., L. T.; KAHN, J. R.; SHUMWAY, N. P. The preparation and function of the hypertensin-converting enzyme. **The Journal of Experimental Medicine**, v. 103, n. 3, p. 295-299, Mar. 1956.

SOCIEDADE BRASILEIRA DE CARDIOLOGIA. SOCIEDADE BRASILEIRA DE HIPERTENSÃO. SOCIEDADE BRASILEIRA DE NEFROLOGIA. VI Diretrizes Brasileiras de Hipertensão: DBH VI. **Arquivos Brasileiros de Cardiologia**, 2010. Disponível em: <http://publicacoes.cardiol.br/consenso/2010/Diretriz_hipertensao_associados.pdf>. Acesso em: 28 jan. 2013.

TAVARES, A.; PLAVNIK, F. L. Inibidores do sistema simpático. **HiperAtivo**, v. 5, n. 2, p. 91-96, jun. 1998.

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

THOMSEN, R.; CHRISTENSEN, M. H. MolDock: a new technique for high-accuracy molecular docking. **Journal of Medicinal Chemistry**, v. 49, n. 11, p. 3315-3332, June 2006.

TIAN, F.; ZHOU, P.; LI, Z. A novel atom-pair hologram (APH) and its application in peptide QSARs. **Journal of Molecular Structure**, v. 871, n. 1-3, p. 140-148, Dec. 2007.

TODESCHINI, R.; CONSONNI, V.; PAVAN, M.; **Dragon Software**. Milano, 2002. Disponível em: <<http://www.vcclab.org/lab/edragon/>>. Acesso em: 25 nov. 2011.

TOPLISS, J. G.; COSTELLO, R. J. Chance correlations in structure-activity studies using multiple regression-analysis. **Journal of Medicinal Chemistry**, v. 15, n. 10, p. 1066-1068, Oct. 1972.

TROPSHA, A. Best practices for QSAR model development, validation, and exploration. **Molecular Informatics**, v. 29, n. 6-7, p. 476-488, July 2010.

TSCHOPE, C.; SCHULTHEISS, H. P.; WALTHER, T. Multiple interactions between the renin-angiotensin and the kallikrein-kinin systems: Role of ACE inhibition and AT1 receptor blockade. **Journal of Cardiovascular Pharmacology**, v. 39, n. 4, p. 478-487, Apr. 2002.

UNGER, S. H.; HANSCH, C. Model-building in structure-activity relationships - re-examination of adrenergic blocking activity of beta-halo-beta-arylalkylamines. **Journal of Medicinal Chemistry**, v. 16, n. 7, p. 745-749, July 1973.

VEDANI, A.; DOBLER, M. 5D-QSAR: The key for simulating induced fit? **Journal Medicinal Chemistry**. v. 45, n. 11, p. 2139-2149, May 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**, v. 48, n. 11, p. 3700-3703, June 2005.

VERMA, R. P.; HANSCH, C. An approach toward the problem of outliers in QSAR. **Bioorganic & Medicinal Chemistry**, v. 13, n. 15, p. 4597-4621, Aug. 2005.

VERMA, R. P.; HANSCH, C. Camptothecins: A SAR/QSAR Study. **Chemical Reviews**, v. 109, n. 1, p. 213-235, Jan. 2009.

VERMA, R. P.; HANSCH, C. Matrix metalloproteinases (MMPs): Chemical-biological functions and (Q)SARs. **Bioorganic & Medicinal Chemistry**, v. 15, n. 6, p. 2223-2268, Mar. 2007.

WATERBEEMD, H. van de; GIFFORD, E. ADMET in silico modelling: Towards prediction paradise? **Nature Reviews Drug Discovery**, v. 2, n. 3, p. 192-204, Mar. 2003.

WORLD HEALTH STATISTICS. 2012. Disponível em:
<http://www.who.int/healthinfo/EN_WHS2012_Full.pdf>. Acesso em 28 Jan. 2013.