



**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA**

DANIELA SANTOS CRUZ

**ESTUDOS DE QSAR 2D E 3D PARA DERIVADOS DE
AMINOIMIDAZÓIS, AMINOHIDANTOÍNAS E
AMINIPYRIDINAS COM ATIVIDADE INIBITÓRIA SOBRE
A ENZIMA BETA-SECRETASE HUMANA**

Feira de Santana, Ba

2012

DANIELA SANTOS CRUZ

**ESTUDOS DE QSAR 2D E 3D PARA DERIVADOS DE
AMINOIMIDAZÓIS, AMINOHIDANTOÍNAS E
AMINIPYRIDINAS COM ATIVIDADE INIBITÓRIA SOBRE
A ENZIMA BETA-SECRETASE HUMANA**

Dissertação apresentada ao Programa de Pós-graduação em Biotecnologia, da Universidade Estadual de Feira de Santana como requisito parcial para obtenção do título de Mestre em Biotecnologia.

Orientador: Prof. Dr. Marcelo Santos Castilho

Feira de Santana, Ba

2012

“Inteligência é a capacidade de se adaptar à mudança.”

Stephen Hawking

AGRADECIMENTOS

Ao Prof. Dr. Marcelo Santos Castilho, orientador desta dissertação, um anjo de asas cortadas, que sempre ajuda os seus alunos em tudo. Agradeço pelo acolhimento no seu laboratório, pela oportunidade de realização conjunta deste trabalho, pela valiosa orientação, pela confiança em mim depositada e por todo empenho e sabedoria.

Aos coordenadores do Programa de Pós-Graduação em Biotecnologia da UEFS, prof. Sandra Aparecida de Assis, prof. Aristóteles Góes Neto e ao secretário do programa, Helton Ricardo que, gentilmente, atenderam às solicitações e forneceram incentivos para que os resultados deste trabalho pudessem ser apresentados nos congressos.

Aos profs. Rafael V. C. Guido e Carlos Maurício Rabello de Sant'Anna-UFRRJ por aceitarem participar da banca de defesa desta dissertação.

A minha mãe e amigos do LABIMM, em especial Humberto, pela ajuda.

A todas as pessoas que, direta ou indiretamente, contribuíram para a execução dessa dissertação.

A FPESB pelo suporte financeiro essencial à realização deste trabalho.

LISTA DE ABREVIATURAS

2D	Bidimensional
3D	Tridimensional
ACh	Acetilcolina
AChE	Acetilcolinesterase
A β	Beta Amilóide (do inglês, <i>Amyloid -beta</i>)
AG	Algoritmo Genético
BHE	Barreira Hemato Encefálica
BuChE	Butirilcolinesterase
APP	Proteína Precursora de Amilóide (do inglês, <i>Amyloid Precursor Protein</i>)
BACE-1	Beta-secretase
CADD	Planejamento de Fármacos Auxiliado por Computador (do inglês, <i>Computer Aided Drug Design</i>)
CRC	Checagem Cíclica de Redundância (do inglês, <i>Cyclic Redundancy Check</i>)
CoMFA	Análise Comparativa dos Campos Moleculares (do inglês, <i>Comparative Molecular Field Analysis</i>)
DA	Doença de Alzheimer
DCNT	Doenças Crônicas Não Transmissíveis
HCA	Análise Hierárquica de Agrupamentos (do inglês, <i>Hierarchical Cluster Analysis</i>)
HEA	Hidroxietilamina
HQSAR	Holograma QSAR
HTS	Ensaio em Larga Escala (do inglês, <i>High-throughput Screening</i>),
LBDD	Planejamento de Fármacos Baseado no Ligante (do inglês, <i>Ligand-based Drug Design</i>)
LOO	Deixe um de fora (do inglês <i>Leave-one-out</i>)

LMO	Deixe muitos de fora (do inglês <i>Leave-many-out</i>)
MLR	Regressão Linear Múltipla (do inglês, <i>Multiple Linear Regression</i>)
PDB	Banco de Dados de Proteínas (do inglês, <i>Protein Data Bank</i>)
PLS	Mínimos Quadrados Parciais (do inglês, <i>Partial Least Squares</i>)
QSAR	Relações Estrutura Atividade Quantitativa (do inglês <i>Quatitative Structure Activity Relationship</i>)
RMSD	Desvio Médio da Raiz Quadrada (do inglês, <i>Root Mean Square Deviation</i>)
SBDD	Planejamento de Fármacos baseado na Estrutura (do inglês, <i>Structure-based Drug Design</i>)

LISTA DE FIGURAS

- Figura 1:** Alterações percentuais nas principais causas de morte por doença entre 2000 e 2008. 17
- Figura 2:** Risco estimado para a doença de Alzheimer por idade e sexo. 18
- Figura 3:** Imagem seccional de cérebro normal (à direita) e cérebro com doença de Alzheimer (à esquerda). À medida que a doença de Alzheimer progride, ocorre perda das células neuronais, sobretudo no hipocampo e no córtex. No quadro avançado da doença observa-se atrofia acentuada do encéfalo. 21
- Figura 4:** Achados histopatológicos de placas senis e emaranhados neurofibrilares em tecido neuronal de paciente com doença de Alzheimer. 22
- Figura 5:** Etiologia da Doença de Alzheimer segundo a hipótese da Cascata Amilóide. A cascata é iniciada através da geração de peptídeos A β , pelo processamento da APP. Os oligômeros insolúveis de A β -42 e placas amilóides podem induzir neurotoxicidade diretamente ou indiretamente, culminando em morte das células neuronais. 24
- Figura 6:** Agentes anticolinesterásicos em uso clínico para o tratamento de DA. 25
- Figura 7:** Mecanismo catalítico em aspartilproteases (à esquerda) e grupos não hidrolisáveis (à direita) presentes em compostos que mimetizam o estado de transição tetraédrico. 28
- Figura 8:** Inibidor de primeira geração de beta-secretase, baseado em um grupo de hidroxietileno (em destaque), isótero do estado de transição. 29
- Figura 9:** Exemplos de inibidores peptidomiméticos de BACE-1 planejados com base na estrutura do alvo macromolecular. 29
- Figura 10:** Desenvolvimento de inibidores não-peptidomiméticos de BACE-1. 30
- Figura 11:** Ciclização do grupo Acilguanidina presente em **1** (3,7 uM) e obtenção do inibidor de BACE-1 Aminiporidinico **2** (5,2uM). 31
- Figura 12:** Redução da dimensionalidade dos dados por PCA. 36
- Figura 13:** Etapas da construção de modelos de Holograma QSAR 39
- Figura 14:** Alinhamento molecular baseado na Similaridade Morfológica. 42
- Figura 15:** Busca Conformacional por Algoritmo Genético implementada no programa GOLD. 44
- Figura 16:** Etapas para a construção de modelos de QSAR 3D por CoMFA. 46
- Figura 17:** Esqueleto geral dos inibidores de BACE-1, onde estão destacados os átomos usados como referência para gerar o alinhamento I. 49
- Figura 18.** Esqueleto geral dos inibidores de BACE-1. Em negrito estão

- destacados os fragmentos usados para o realinhamento pela máxima estrutura comum. Na parte superior estão marcados os átomos usados para a sobreposição dos compostos quimicamente semelhantes ao ligante da estrutura do complexo da BACE-1 3INF (A) e na parte inferior, os átomos usados para a sobreposição dos compostos quimicamente semelhantes ao ligante da estrutura 3L38 (B). 51
- Figura 19:** Estrutura geral dos inibidores de BACE-1 utilizados nesse trabalho. 54
- Figura 20:** Dendograma do HCA. Linhas horizontais representam os compostos e as linhas verticais a similaridade entre eles. Pontos em vermelho do lado esquerdo indicam os compostos selecionados para o conjunto teste, usado na validação externa dos modelos de QSAR 2D e 3D. 55
- Figura 21:** Representação gráfica da distribuição dos valores de pIC_{50} para os compostos do conjunto treino e do conjunto teste. 55
- Figura 22.** Valores preditos versus experimentais de pIC_{50} para o conjunto treino (quadrados) e teste (triângulos) de acordo com o melhor modelo de HQSAR (distinção de fragmento: A/C/Ch; e tamanho de fragmento: 4–7). 59
- Figura 23:** Mapa de contribuição para inibidores de alta potência de BACE-1 (**20** $pIC_{50}= 7,50$ e **46** $pIC_{50}= 7,30$) de acordo com o melhor modelo de HQSAR. 60
- Figura 24:** Distribuição dos compostos na análise de PCA: quadrados laranja compostos de baixa potencia ($pIC_{50} < 5,65$), losangos, representam inibidores com potencia intermediária ($5,66 < pIC_{50} < 6,66$) ; quadrados brancos, moléculas com alta potencia ($pIC_{50} > 6,67$). 61
- Figura 25:** Gráfico de pesos dos descritores para PC1 e PC2 64
- Figura 26:** Valores preditos versus experimentais de pIC_{50} para o conjunto treino (quadrado) e teste (triângulos) de acordo com o melhor modelo de QSAR 2D baseado em descritores topológicos. 65
- Figura 27:** Contribuição dos descritores para o vetor de regressão obtido para o melhor modelo de QSAR 2D. 66
- Figura 28:** Distribuição dos descritores GGI5 e GGI6 versus valores de atividade biológica pra os compostos do conjunto treino. 67
- Figura 29:** Descritores B08 [N-N] e B09 [N-N] representados na estrutura de um composto de alta potência (**46**, $pIC_{50} = 7,30$) (A); e observação deste mesmo descritor no mapa de contribuição do HQSAR (B). 68

Figura 30: Resultado da superposição baseada na máxima estrutura comum Alinhamento I (acima); imagem comparativa que evidencia claramente as diferenças nas ‘conformações bioativas’ através da superposição dos ligantes do complexo 3L38 e 3INF (A) e das estruturas 3D desses ligantes minimizadas pelo método AM1 (B) (abaixo). 70

Figura 31: Alinhamento dos inibidores de BACE-1 obtido a partir do método baseado na similaridade de moléculas (Alinhamento II). Em amarelo estão indicadas as estruturas dos ligantes extraídos dos complexos da BACE-1 PDB 3INF e 3L38. 72

Figura 32: Representação esquemática das principais interações entre os inibidores da série em estudo e a BACE-1 (à esquerda) e resultados da superposição dos compostos no Alinhamento III (à direita). 72

Figura 33: Estratégia para o realinhamento molecular das conformações oriundas dos Alinhamentos II (acima) e alinhamento III (abaixo) pelo método da máxima estrutura comum, para a geração dos alinhamento IV e V respectivamente. Os ligantes complexados na estrutura cristalográfica de 3INF e 3L38 estão representados em amarelo. 73

Figura 34: Valores residuais (diferença entre pIC_{50} predito e observado experimentalmente) dos compostos do conjunto teste, para o melhor modelo de CoMFA obtido com o Alinhamento II (modelo B). 77

Figura 35: Valores preditos versus experimentais de pIC_{50} para o conjunto treino (quadrado) e teste (triângulos) de acordo com o melhor modelo obtido com o alinhamento IV (esquerda) e alinhamento V (direita). 78

Figura 36: Representação simplificada do sítio ativo da BACE-1 (PDB: 3INF). 80

Figura 37: Mapas de contorno para o composto **20** ($pIC_{50} = 7,5$) de acordo com os melhores modelos de CoMFA. Os mapas estéreos (esquerda) estão mostrados em verde e amarelo (contornos em 0,05 Kcal/mol e -0,017 Kcal/mol, respectivamente), enquanto que os mapas eletrostáticos (à esquerda) estão mostrados em vermelho e azul (contornos em 0,078 e -0,065 Kcal/mol). 81

Figura 38: Acoplamento do inibidor **20** ($pIC_{50} = 7,50$) no sítio ativo da BACE-1 (PDB: 3INF). A estrutura do ligante complexado está em rosa. Os mapas de contorno estéreos foram obtidos com o modelo E (verde: 0,05 Kcal/mol; amarelo: -0,017 Kcal/mol). Em destaque, o poliedro verde indica a substituição favorável em *para* no anel *m*-piridínico, possibilitando a formação de interações de hidrogênio adicionais com a carbonila dos resíduos Gly73 / Thr294 (setas tracejadas). A linha sólida em negrito destaca a interação do inibidor com o resíduo Ser 291, por meio de uma molécula de água conservada. 82

Figura 39: Acoplamento do inibidor **20** ($pIC_{50} = 7,50$) no sítio ativo da BACE-1 (PDB: 3INF). A estrutura do ligante complexado está em rosa. 83
Mapas de contorno estéreo e eletrostáticos foram obtidos com o modelo E (verde: 0,05 Kcal/mol; amarelo: -0,017 Kcal/mol; azul: 0,078 Kcal/mol; vermelho: 0,065 Kcal/mol). Em destaque a região FLAP e S2'.

Figura 40: Estratégias de modificações moleculares a partir da análise dos 84
mapas de contorno dos modelos de CoMFA para aumentar a atividade da série de compostos em estudo.

LISTA DE TABELAS

Tabela 1. Influencia da distinção dos fragmentos nos parâmetros estatísticos dos modelos de HQSAR usando o tamanho de fragmento padrão (4-7).	57
Tabela 2. Influencia do tamanho do fragmento nos parâmetros estatísticos do melhor modelo de HQSAR (A/C/Ch).	58
Tabela 3: Valores preditos e residuais de pIC_{50} para os 27 compostos do grupo teste, de acordo com o melhor modelo de HQSAR (A/C/Ch).	59
Tabela 4: Descritores selecionados para a modelagem do QSAR 2D baseado em descritores topológicos.	
Tabela 5. Valores preditos e residuais de pIC_{50} para os 27 compostos do grupo teste de acordo como melhor modelo de QSAR 2D clássico.	62
Tabela 6: Modelos preliminares de CoMFA obtidos com parâmetros padrão e espaçamento de 2Å.	65
Tabela 7: Influência da focagem da região e do espaçamento da grade nos valores de q^2 dos modelos de COMFA.	75
Tabela 8: Resultados dos melhores modelos de CoMFA obtido com os Alinhamentos I, II, III, IV e V.	77

RESUMO

A doença de Alzheimer (DA) é uma desordem progressiva, neurodegenerativa caracterizada fisiologicamente pela degeneração dos neurônios colinérgicos e formação de placas senis contendo peptídeo β -amilóide, que leva ao quadro de perda cognitiva, de memória e demência. O principal evento fisiopatológico na DA é a deposição extracelular do peptídeo β -amilóide (β A), que se origina da ação preteolítica da enzima beta-secretase humana (BACE-1) sobre a proteína precursora do amiloide (APP). Por essa razão, inibidores de BACE-1 podem atuar com maior eficácia terapêutica sobre o processo patológico da DA, visto que o mecanismo de ação dos fármacos atualmente disponíveis para o tratamento proporciona somente alívio temporário dos sintomas da doença, não alterando a sua progressão. O planejamento recente de fármacos inibidores de BACE-1 tem buscado o desenvolvimento de moléculas não peptídicas a fim de superar as limitações farmacocinéticas dos derivados miméticos do estado de transição. Contudo, a potência desses compostos ainda precisa ser otimizada. Visando contribuir para este objetivo modelos de QSAR 2D baseados em descritores topológicos e em fragmentos moleculares (Holograma QSAR), assim como modelos de QSAR 3D (CoMFA) foram desenvolvidos para uma série de 102 derivados de aminohidantoinas, aminoimidazóis e aminopiridinas que inibem BACE-1.

Os modelos de HQSAR apresentam boa qualidade estatística ($r^2 = 0,85/ q^2 = 0,84$ e $r^2_{pred} = 0,70$) e sugerem que o grupo amino ligado ao núcleo central aminoimidazol e aminohidantoinico, bem como o nitrogênio do anel piridínico tem contribuição positiva para a potência, todavia, foi a análise integrada desse resultado com aqueles oriundos dos modelos de QSAR baseados em descritores topológicos, ($r^2 = 0,87$, $q^2 = 0,85$ e $r^2_{pred} = 0,84$) que permitiu identificar fatores eletrônicos (GGI5 e GGI6) como os principais responsáveis por essa contribuição.

O melhor modelo CoMFA ($r^2 = 0,91/ q^2 = 0,73$ e $r^2_{pred} = 0,86$) foi obtido a partir do alinhamento molecular pela máxima subestrutura comum, utilizando conformações geradas através da comparação por similaridade morfológica de todos os inibidores com os ligantes encontrados nas estruturas cristalográficas 3INF ou 3L38 (Alinhamento IV). A análise dos mapas de contorno sugere por exemplo que existe uma interação estericamente desfavorável na região próxima ao núcleo aminoimidazol, enquanto não há restrição estérica significativa na posição *para* do anel *o*-cloro piridínico. A partir dessas observações estratégias de modificações estruturais foram propostas para a obtenção de derivados congêneres mais potentes.

PALAVRAS-CHAVES: Alzheimer, β -secretase, QSAR 2D, HQSAR, CoMFA

ABSTRACT

Alzheimer's disease (AD) is a neurodegenerative and progressive disorder, physiologically characterized by degeneration of cholinergic neurons and formation of senile plaques containing β -amyloid peptide, which leads to cognitive and memory loss as well as dementia. The main pathophysiological event in AD is the deposition of extracellular β -amyloid peptide (β A), which originates from proteolytic action of human beta-secretase (BACE-1) over the amyloid precursor protein (APP). Therefore, inhibitors of BACE-1 may have greater therapeutic efficacy on controlling pathological process of AD, once the mechanism of action of currently available drugs provides only temporary relief from AD symptoms, without altering its progression. Recent efforts to design BACE-1 inhibitors have focused on non-peptide molecules that might overcome pharmacokinetic limitations found in transition-state mimetic inhibitors. However, the potency of these compounds has yet to be optimized. In order to contribute for this purpose QSAR models based on 2D and topological descriptors molecular fragments (hologram QSAR) and 3D models QSAR (CoMFA) have been developed for 102 derivatives aminohydantoin, aminoimidazoles and aminopyridines as inhibitors of BACE-1.

HQSAR models exhibit good statistical values ($r^2 = 0.85/ q^2 = 0.84$ and $r^2_{pred} = 0.70$) and suggest that amine moieties bound to aminohydantoin or aminoimidazole rings, as well as nitrogen from the pyridine ring increase potency. However, only the additional information provided by QSAR models built with topological descriptors ($r^2 = 0.87/ q^2 = 0.85$ and $r^2_{pred} = 0.84$) indicated that electronic features (GGI5 and GGI6) are the major underlining factors to this result.

The best CoMFA model ($r^2 = 0.91/ q^2 = 0.73$ and $r^2_{pred} = 0.86$) was built on the basis of the maximum substructure alignment using conformation obtained by structural similarity comparison among all inhibitors and the ligands bound to 3INF or 3L38 crystallographic structures (alignment IV). Analysis of contour maps suggests, for instance, that steric hindrance near the aminoimidazole ring prevents the further addition of substituents at this point, whereas the presence of bulky moieties in the *para* position in the *o*-chloro pyridine ring would increase potency. Guided by such analysis, structural modification were proposed to develop inhibitors with increased potency.

KEYWORDS: Alzheimer, β -secretase, QSAR 2D, HQSAR, CoMFA

SUMÁRIO

1. INTRODUÇÃO	16
1.1. O ENVELHECIMENTO DA POPULAÇÃO E A DOENÇA DE ALZHEIMER	16
2. OBJETIVOS	20
2.1. OBJETIVOS GERAIS	20
2.2. OBJETIVOS ESPECÍFICOS	20
3. REVISÃO DA LITERATURA	21
3.1. ASPECTOS FISIOPATOLÓGICOS, BIOQUÍMICOS E MOLECULARES DA DOENÇA DE ALZHEIMER	21
3.1.1. VIA AMILOIDOGÊNICA DA APP E A HIPÓTESE DA CASCATA AMILÓIDE	23
3.2. TRATAMENTO DA DOENÇA DE ALZHEIMER	24
3.2.1. “HIPÓTESE COLINÉRGICA” E OS INIBIDORES COLINESTERÁSICOS	25
3.2.2. O ALVO TERAPÊUTICO BETA-SECRETASE (BACE-1)	27
3.2.3. DESENVOLVIMENTO DE INIBIDORES DE BACE-1	27
3.3. PLANEJAMENTO DE FÁRMACOS E AS FERRAMENTAS COMPUTACIONAIS	31
3.3.1. DELINEAMENTO DOS ESTUDOS DE QSAR	32
3.3.2. HOLOGRAMA QSAR	37
3.3.3. ANÁLISE COMPARATIVA DOS CAMPOS MOLECULARES (CoMFA)	39
3.3.3.1. Alinhamento Estrutural Rígido pela Máxima Estrutura Comum	41
3.3.3.2. Alinhamento Estrutural Baseado na Similaridade Morfológica	41
3.3.3.3. Alinhamento Estrutural Baseado em Acoplamento Molecular	43
3.3.3.4. A Construção dos Campos Moleculares e Geração dos Modelos	44

4. MÉTODOS COMPUTACIONAIS	46
4.1. CONJUNTO PADRÃO DE DADOS	46
4.2. QSAR 2D HOLOGRAMA QSAR	47
4.3. MODELOS QUIMIOMÉTRICOS 2D BASEADO EM DESCRITORES TOPOLÓGICOS	48
4.4. ALINHAMENTO ESTRUTURAL DOS COMPOSTOS	48
4.4.1. Alinhamento por Máxima Subestrutura Comum (Alinhamento I)	48
4.4.2. Alinhamento por Similaridade Morfológica (Alinhamento II)	49
4.4.3. Alinhamento por Acoplamento Molecular (Alinhamento III)	49
4.4.4. Realinhamento Rígido de II e III (Alinhamentos IV e V)	50
4.6. QSAR 3D CoMFA	51
4.7. VALIDAÇÃO DOS MODELOS DE QSAR	52
5. RESULTADOS E DISCUSSÃO	53
5.1. ANÁLISE DO CONJUNTO DE DADOS	53
5.2. ESTUDOS DE QSAR 2D	56
5.2.1. Holograma QSAR (HQSAR)	56
5.2.2. QSAR 2D Baseado em Descritores Topológicos	61
5.3. QSAR 3D CoMFA	69
5.3.1. O Alinhamento Estrutural dos Compostos	69
5.3.2. Desenvolvimento e Validação dos Modelos de CoMFA	74
5.3.3. Interpretação dos Mapas de Contorno do CoMFA	85
6. COLCLUSÃO	85
REFERÊNCIAS	87
APÊNDICES	
Apêndice A- Estrutura química e valores de pIC_{50} do conjunto de inibidores de bace-1 utilizados para os estudos de QSAR 2D e QSAR 3D (as moléculas do conjunto teste são estão indicadas por marcadores de fundo).	100
Apêndice B- Valores de pIC_{50} preditos e valores residuais para os 27 compostos do conjunto teste de acordo com os melhores modelos de CoMFA.	108

1. INTRODUÇÃO

1.1. O ENVELHECIMENTO DA POPULAÇÃO E A DOENÇA DE ALZHEIMER

Segundo dados da Organização Mundial de Saúde (OMS), até 2025 haverá no mundo cerca de 1,2 bilhões de habitantes com idade superior a 60 anos¹. Embora o aumento da expectativa de vida represente um triunfo, o desafio de proporcionar melhoria ou manutenção da saúde e qualidade de vida para as pessoas idosas instaura-se como o um novo desafio na sociedade moderna.

O envelhecimento é um processo biológico natural do organismo que traz modificações físicas (corporais), fisiológicas e psicológicas que estão intrinsecamente associados ao aparecimento de doenças crônicas não transmissíveis (DCNT)² tais como as doenças cardiovasculares, câncer, distúrbios musculoesqueléticos (ex: osteoporose e artrite), doenças neurodegenerativas (ex: doença de Parkinson, doença de Alzheimer), dentre outras³. Entre as DCNT destacam-se síndromes demenciais, que apresentam etiologias diversas, mas cujo aspecto fundamental é o declínio progressivo das funções cerebrais, como a memória, linguagem e discernimento comprometendo o funcionamento ocupacional ou social dos indivíduos acometidos⁴. De acordo com a OMS os portadores de demências com 60 anos ou mais passa cerca de 11,2% dos seus anos de vida incapacitados. Isto representa mais do que as limitações em idosos causadas por doenças cardiovasculares (5%) e todas as formas de câncer (2,4%)⁵.

A doença de Alzheimer (DA) é a mais comum das síndromes demenciais, representando de 50-60% de todos os casos, seguida pela demência por corpos de Lewy, que responde por 20% dos demais casos⁶. Trata-se de uma doença degenerativa, de início insidioso e que leva a um declínio gradativo e sequencial nas funções cognitivas, de linguagem, de memória, comportamentais e motoras, interferindo com as funções diárias do indivíduo e na sua qualidade de vida⁷.

A DA vem se tornando a principal causa de morte na população dos Estados Unidos e outros países com elevada expectativa de vida. Enquanto DCNT como o câncer de próstata e as doenças cardíacas tem experimentado declínios significativos como causa de morte, os números da doença de Alzheimer estão em ascensão. (Figura 1)⁸.

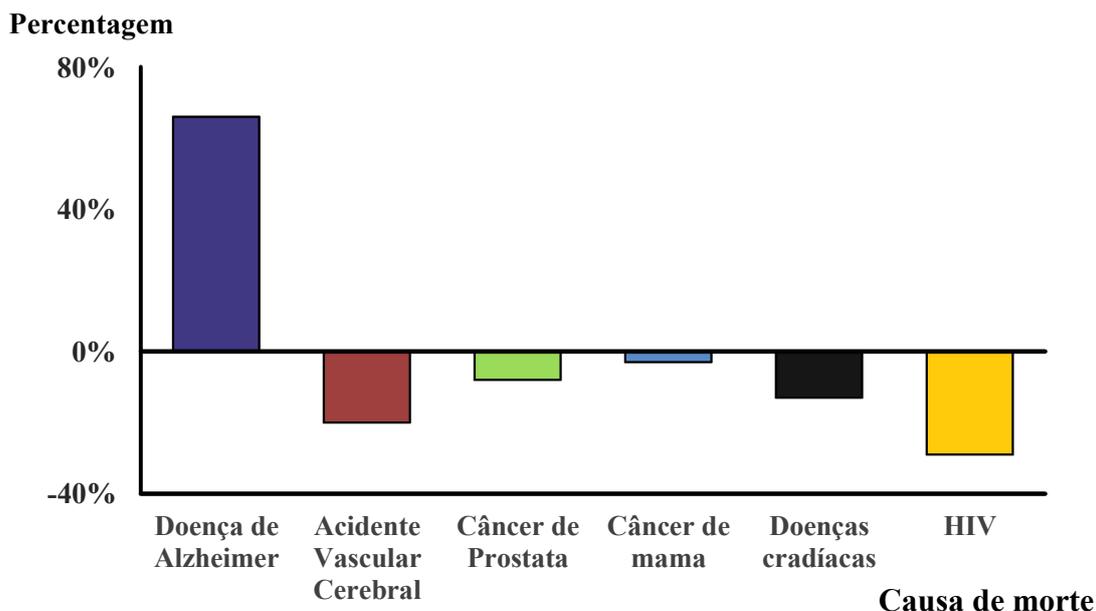


Figura 1: Alterações percentuais nas principais causas de morte por doença entre 2000 e 2008.⁸

Em 2006, havia cerca 26,6 milhões de casos de Alzheimer no mundo e a estimativa é de que até o ano de 2050, a prevalência da doença crescerá quatro vezes, chegando a 106,4 milhões de casos⁹. No Brasil, atualmente, são mais de um milhão de pessoas afetadas¹⁰, o que representa 55,1% dos casos de demência observados na população brasileira com mais de 65 anos¹¹.

A etiologia da DA ainda não está completamente esclarecida, porém sabe-se que a incidência é maior entre as mulheres¹² e embora um histórico familiar positivo¹³, bem como, níveis elevados de colesterol, diabetes¹⁴ e fumo¹⁵ possam estar relacionados ao aumento do risco para o desenvolvimento de Alzheimer, a idade é considerado o fator de maior relevância para a prevalência da doença¹⁶ (Figura 2). Por exemplo, a incidência anual estimada de aproximadamente 53 novos casos por 1.000 pessoas entre 65 a 74 anos sobe para 170 novos casos por 1.000 pessoas com idades entre 75-84 e para 231 novos casos por 1.000 pessoas com 85 anos ou mais¹⁷.

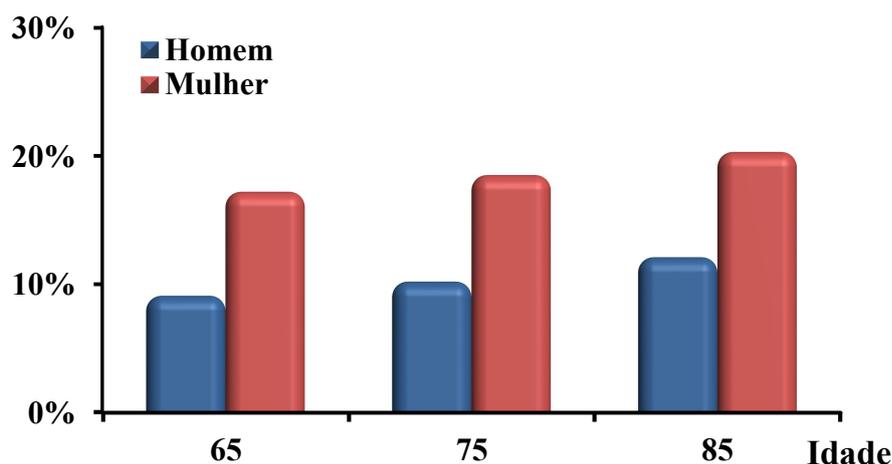


Figura 2: Risco estimado para a doença de Alzheimer por idade e sexo⁸.

Os sinais patológicos característicos da doença aparecem como resultado de eventos metabólicos que culminam na disfunção simpática e neuroquímica decorrentes da redução da atividade colinérgica central¹⁸. Como consequência desses eventos instala-se o quadro de redução da memória, deficiência das funções cognitivas, de atenção e fluência verbal, distúrbios comportamentais, etc¹⁹. O impacto socioeconômico da Doença de Alzheimer é reflexo da sua natureza progressiva da doença que amplia o risco de problemas de saúde física, tornando os portadores cada vez mais dependentes de serviços de saúde e assistência social e de outras pessoas. As mudanças nas atividades diárias, comportamento, hábitos e, sobretudo, os sintomas psicológicos dos portadores, tais como agitação, agressividade e ansiedade, representam uma causa comum para a institucionalização dos pacientes. Para familiares e amigos, que proporcionam os cuidados, o encargo é imensurável²⁰.

No tratamento farmacológico da DA, substâncias psicoativas são utilizadas para preservar ou restabelecer o comportamento, a cognição e as habilidades funcionais do paciente com a DA²¹. Contudo, os efeitos dessas substâncias limitam-se ao retardo na evolução natural da doença, permitindo apenas uma melhora temporária do estado funcional do paciente²².

O desenvolvimento de terapias para DA representa um grande desafio para acadêmicos e cientistas das áreas de biotecnologia, farmácia e afins. Uma das razões para isso reside no fato de que o cérebro humano, órgão que é o foco destas terapias é a estrutura mais complexa de toda a biologia. Apesar desta difícil tarefa, pesquisadores têm obtido progressos na compreensão da fisiopatologia da DA, bem como na identificação e

validação de novos alvos terapêuticos que são a base para novas estratégias de tratamento²³.

A patogênese neurodegenerativa característica da DA está relacionada com a formação do peptídeo β -amilóide que é produzido *in vivo* através da clivagem proteolítica da Proteína Precursora de Amilóide (APP) pelas enzimas β -secretase (BACE-1) e γ -secretase²⁴. Evidências sugerem que a clivagem da APP pela BACE-1 é um passo crítico na formação do peptídeo A β e que a prevenção da formação desse produto através da inibição dessa via, pode ter um impacto positivo sobre na progressão da DA. Por esta razão, a BACE-1 tem se mostrado um alvo terapêutico vantajoso e a busca por inibidores potentes e seletivos desta enzima tem se destacado nos campos de pesquisas de novas alternativas para o tratamento da doença^{25,26}.

A identificação dos requisitos estruturais moleculares que são essenciais para atividade biológica pode ser realizada através de modelos de QSAR²⁷, os quais se baseiam na premissa que a atividade de um fármaco é função de interações estabelecidas entre a estrutura molecular do composto e seu receptor no sistema biológico²⁸, é possível modificar a estrutura molecular de inibidores de BACE-1 a fim de planejar moléculas com melhor potência e/ou melhor perfil farmacológico. Visando contribuir para o planejamento de fármacos mais eficazes contra DA, modelos de QSAR podem ser desenvolvidos, buscando identificar propriedades físico-químicas e estruturais responsáveis pela inibição da BACE-1 e assim, fornecer novos dados que auxiliem na obtenção de inibidores mais potentes.

2. OBJETIVOS

2.1. OBJETIVOS GERAIS

Desenvolver modelos de QSAR robustos e preditivos que permitam a elucidação das relações entre a estrutura química e a atividade biológica para um conjunto de derivados não peptídeomiméticos ativos contra a enzima beta-secretase (BACE-1).

2.2. OBJETIVOS ESPECÍFICOS

- Organizar um conjunto de dados padrão de inibidores da enzima BACE-1;
- Desenvolver modelos de QSAR, empregando descritores bidimensionais (topológicos e/ ou fragmentos);
- Desenvolver modelos de QSAR, empregando a abordagem tridimensional de Análise Comparativa dos Campos Moleculares (CoMFA).
- Planejar, a partir das informações provenientes dos modelos de QSAR 2D e 3D, inibidores não peptídeomiméticos mais potentes da enzima BACE-1.

3. REVISÃO DA LITERATURA

3.1. ASPECTOS FISIOPATOLÓGICOS, BIOQUÍMICOS E MOLECULARES DA DOENÇA DE ALZHEIMER

Clinicamente os primeiros sinais da doença de Alzheimer se caracterizam pela perda de memória recente e, à medida que ocorre a progressão da doença o quadro agrava-se afetando a linguagem, compreensão e capacidade associativa do indivíduo. Adicionalmente, os pacientes apresentam distúrbios comportamentais, que interferem na razão, humor e capacidade de julgamento, sendo que na fase avançada da doença existem anormalidades motoras e sensoriais que levam a dificuldades de locomoção, realização de atividades básicas do cotidiano e de comunicação²⁹. Fisiologicamente, essas manifestações são o resultado do processo de degeneração no cérebro, sobretudo, nas camadas ricas em neurônios piramidais afetando, principalmente, as estruturas límbicas e os córtices associativos. O exame macroscópico do cérebro de indivíduos acometidos pela DA evidencia que, paralelamente à perda neuronal, ocorre intensa degeneração sináptica em regiões hipocámpais com atrofia cortical em graus variáveis e dilatação acentuada das fissuras do tecido cerebral³⁰ (Figura 3).

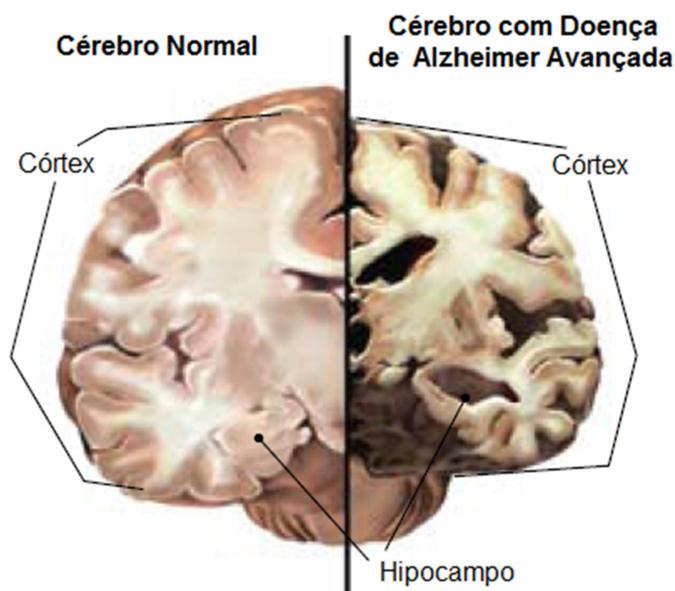


Figura 3: Imagem seccional de cérebro normal (à direita) e cérebro com doença de Alzheimer (à esquerda). À medida que a doença de Alzheimer progride, ocorre perda das células neuronais, sobretudo no hipocampo e no córtex. No quadro avançado da doença observa-se atrofia acentuada do encéfalo [Adaptado da ref. 30].

Os danos neurais observados na DA são causados por múltiplas alterações bioquímicas e moleculares. Nos tecidos cerebrais *post mortem* de indivíduos com DA, dois tipos de lesões histopatológicas características são identificadas (Figura 4): as placas senis (ou amilóides) formadas a partir da deposição do peptídeo beta-amiloide ($A\beta$); e emaranhados neurofibrilares intracelulares, compostos por depósitos de proteína Tau hiprefosforilada e componentes microtubulares³¹. Estudos sugerem que esses achados são os responsáveis pelo desencadeamento de eventos de toxicidade, como a ativação de células da micróglia e migração de astrócitos reativos acarretando em resposta inflamatória local. Esta cascata aumenta progressivamente, com reações oxidativas e produção de radicais livres, que comprometem a homeostasia celular levando à disfunção e morte das células neuronais³²⁻³³.

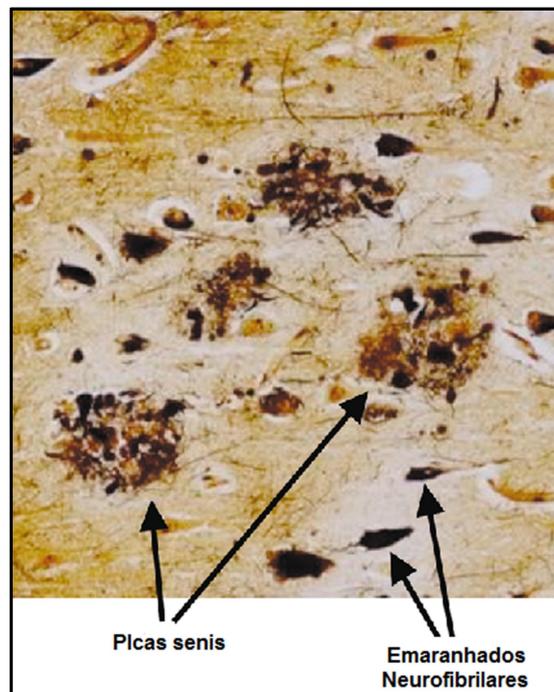


Figura 4: Achados histopatológicos de placas senis e emaranhados neurofibrilares em tecido neuronal de paciente com doença de Alzheimer. [Adaptado da ref. 33]

3.1.2. VIA AMILOIDOGÊNICA DA APP E A HIPÓTESE DA CASCATA AMILÓIDE

O principal mecanismo que explica a etiopatogenia da DA está relacionado ao metabolismo da Proteína Precursora de Amilóide (APP). Trata-se de uma glicoproteína transmembrana que apresenta diferentes isoformas, sendo expressa constitutivamente, sobretudo, na superfície de células do sistema nervoso central e da glia³⁴.

Na via amiloidogênica a APP sofre a ação proteolítica da enzima β -secretase (BACE-1), na porção transmembrana, liberando um fragmento solúvel e um polipeptídeo terminal C-99. Este último é clivado, na sequência, em várias posições pela ação da enzima γ -secretase, resultando na formação de fragmentos amilóides ($A\beta$)^{34,35}, constituídos de 42 aminoácidos ($A\beta$ -42) que apresentam maior capacidade de oligomerização, formando as placas de fibrilas que compõem o achado histopatológico característico da DA^{36,37}.

Outra vertente para explicar a fisiopatologia da DA está na identificação da proteína Tau como principal constituinte dos emaranhados neurofibrilares³⁸. A proteína Tau está ligada à superfície dos microtúbulos atuando no mecanismo de polimerização e despolimerização das estruturas de tubulina durante o processo de extensão axonal³⁹. A atividade dessa proteína é regulada, a nível pós-transducional, por fosforilação e desfosforilação decorrente da ação de enzimas cinases, tais como a GSK-3 β ⁴⁰. Em situações patogênicas, em que há a desregulação deste mecanismo, a proteína Tau pode ser hiperfosforilada alterando sua ligação aos microtúbulos e levando à desestabilização da tubulina. Por consequência, ocorre o desmonte das estruturas microtubulares, com desorganização do transporte axonal, aliado a deposição intracelular da proteína hiperfosforilada no interior da célula, o que resulta na formação de emaranhados neurofibrilares⁴¹.

Evidências mostram que os diferentes mecanismos patogênicos sugeridos estejam intimamente correlacionados, de forma que o mecanismo catalítico da APP pela via amiloidogênica pode promover a hiperfosforilação da proteína Tau⁴². Dessa maneira, a hipótese da Cascata Amilóide (Figura 5) inicia-se pela ação da enzima β -secretase, abrindo caminho para a formação do peptídeo $A\beta$ -42. Isso representa *per se* o princípio de uma sequência de eventos capazes de promover disfunção sináptica das células neuronais através de um mecanismo de toxicidade gerado diretamente pelos oligômeros

de A β -42 agregados ou, indiretamente, por modificações bioquímicas que favorecerem a ativação de enzimas cinases, com hiperfosforilação da proteína Tau e, conseqüentemente, uma maior formação dos emaranhados neurofibrilares. Como resultado de ambos mecanismos tem-se alterações morfológicas dos neurônios, culminando em última instância, em perda da função e morte neuronal⁴¹⁻⁴⁴.

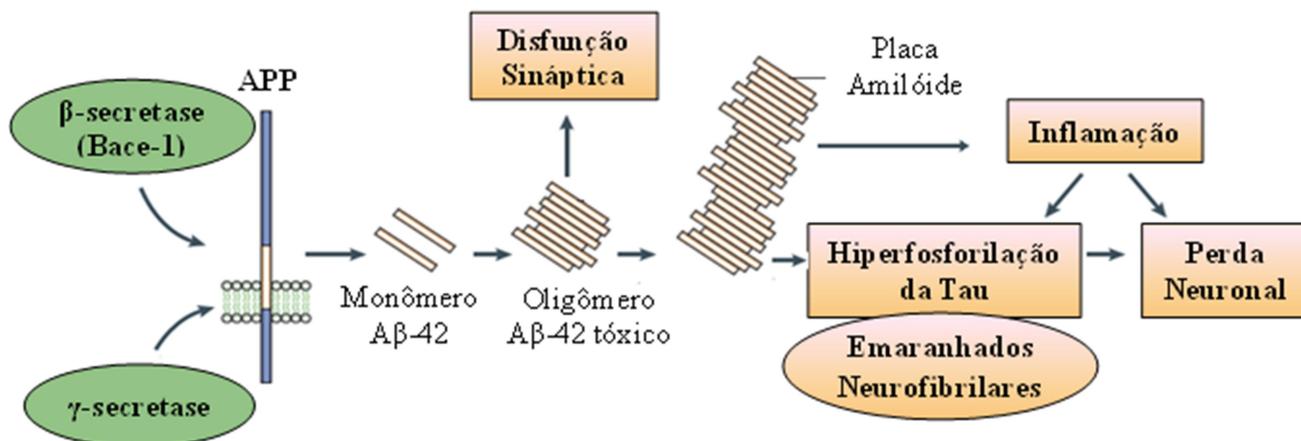


Figura 5: Etiologia da Doença de Alzheimer segundo a hipótese da Cascata Amilóide. A cascata é iniciada através da geração de peptídeos A β , pelo processamento da APP. Os oligômeros insolúveis de A β -42 e placas amilóides podem induzir neurotoxicidade diretamente ou indiretamente, culminando em morte das células neuronais. [Adaptado da ref. 44].

3.2. TRATAMENTO DA DOENÇA DE ALZHEIMER

Até o momento nenhuma intervenção eficaz, capaz de impedir, modificar ou reverter o declínio progressivo das funções cerebrais decorrentes da doença de Alzheimer foi desenvolvida. No entanto, há um número significativo de estratégias farmacológicas e não farmacológicas para o tratamento dos sintomas da doença^{22,45}.

Atualmente, a intervenção farmacológica para a DA baseia-se em duas estratégias principais: i) tratamento sintomático, que busca restauração parcial ou provisória das habilidades cognitivas e funcionais dos portadores; ii) terapêutica complementar, que está direcionada ao tratamento das manifestações não-cognitivas, principalmente as de natureza psíquica como depressão, agressividade, psicose, agitação, distúrbios do sono, etc., melhorando parcialmente as relações sociais e a qualidade de vida dos pacientes⁴⁶. A seguir serão apresentados maiores detalhes sobre fármacos que atuam pela primeira estratégia.

3.2.1. “HIPÓTESE COLINÉRGICA” E OS INIBIDORES COLINESTERÁSICOS

O sistema colinérgico origina-se no prosencéfalo e se projeta difusamente para o córtex cerebral, onde a Acetilcolina (ACh), e outros neurotransmissores, desempenham papel importante como mediadores químicos nos sistemas relacionados à cognição, memória e comportamento⁴⁷. A “hipótese colinérgica” sugere que a degeneração das vias colinérgicas no cérebro de pacientes com DA, por conta da perda acentuada de neurônios no núcleo do prosencéfalo, tem por consequência o déficit na produção e liberação da ACh⁴⁸. Dessa maneira, a utilização de fármacos potencializadores da função colinérgica central (Figura 6), capazes de proporcionar um incremento da transmissão neuronal, tem sido uma das abordagens mais aplicadas na tentativa de atenuar a deficiência cognitiva e de memória nos pacientes com Alzheimer⁴⁹.

O principal mecanismo de ação dos fármacos para restaurar o tônus colinérgico central baseia-se no aumento da disponibilidade de acetilcolina através da inibição das enzimas Acetilcolinesterase (AChE) e Butirilcolinesterase (BuChE)^{46, 50, 51}. Estas enzimas encontram-se presentes nas fendas sinápticas centrais, atuando na modulação da transmissão de impulsos nervosos, através da hidrólise rápida e efetiva da acetilcolina (ACh)⁵². Ao reduzir a degradação desse neurotransmissor, os fármacos anticolinesterásicos estabelecem um efeito compensatório sobre a deficiência na produção e liberação da ACh, decorrente do processo de neurodegeneração, o que garante uma atenuação dos sintomas clínicos observados⁵³.

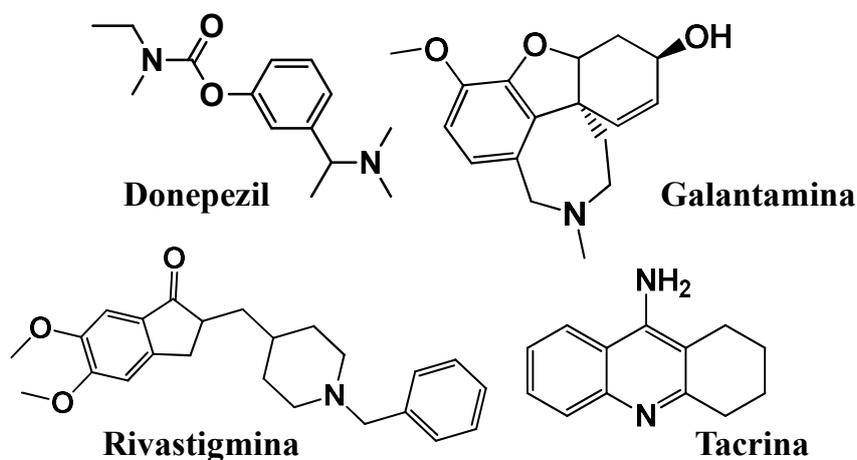


Figura 6: Agentes anticolinesterásicos em uso clínico para o tratamento de DA.

A galantamina (Nivalin®) é um alcalóide isolado de espécies vegetais da família *Amaryllidaceae*⁵⁴ e apresenta ação inibitória seletiva e reversível sobre a AChE. Adicionalmente esse fármaco exerce um efeito modulador sobre os receptores nicotínicos, o que confere propriedades neuroprotetoras vantajosas, comparado a outros inibidores de AChE^{55, 56}. A utilização do fármaco galantamina tem relevância na manutenção dos pacientes com Alzheimer, pois a sua inserção na terapêutica garante aos portadores em média um período de independência de aproximadamente um ano a mais, até o surgimento da necessidade de cuidados assistenciais em tempo integral ou institucionalização⁵⁷.

O fármaco tacrina (Cognex® e Nivalin®) é um derivado da acridina, sintetizado inicialmente como agente antimicrobiano e bloqueador da monoaminoxidase⁵⁸, foi o primeiro anticolinesterásico sintético aprovado na terapêutica da doença de Alzheimer. Apresenta eficácia significativa no alívio sintomático, sobretudo, nas manifestações iniciais⁵⁹. Porém, efeitos hepatotóxicos em aproximadamente 40% dos usuários limitam a sua utilização clínica^{60,61}. Já o Donepezil (Aricept®), um derivado das N-benzilpiperidinas, foi desenvolvido para superar as desvantagens da Tacrina e apesar dos efeitos colaterais decorrentes da sua utilização como náusea, diarreia, constipação, dor de cabeça, tonturas, distúrbios do sono e complicações estomacais⁶¹, oferece maior potência e é bem tolerado em comparação à Tacrina⁶².

O fármaco (S)-rivastigmina (Exelon®), um derivado de fenilcarbamato, possui um tempo de meia-vida maior que os outros inibidores de AChE (aproximadamente 10 horas) o que reduz o número de doses administradas⁶³. Enquanto donepezil e galantamina são seletivos para a AChE, a rivastigmina inibe tanto AChE quanto BuChE⁶⁴. Assim, este fármaco apresenta vantagens adicionais sobre os outros inibidores, pois se acredita que à medida que a DA evolui, a enzima BuChE tenha maior participação na degradação da ACh na fenda sináptica⁵⁰. Além disso, a presença de BuChE em estruturas como hipocampo e tálamo justifica os benefícios da Rivastigmina sobre sintomas comportamentais e psíquicos na DA⁶⁵.

Considerando o mecanismo de ação dos inibidores de colinesterase observa-se que o uso desses fármacos proporciona benefícios sobre a memória, cognição, sintomas psíquicos e comportamentais nos pacientes portadores de DA apenas a curto prazo. Os benefícios da terapêutica duram, em média, 2 a 5 anos⁶⁶, porém esses fármacos não são capazes de impedir a progressão da doença.

3.2.2. O ALVO TERAPÊUTICO β -SECRETASE (BACE-1)

A β -Secretase (BACE-1) é a primeira protease que processa a proteína precursora de amilóide (APP) levando à produção do peptídeo β -amilóide (A β). Trata-se de uma protease aspártica expressa na maioria das células e tecidos do corpo, mas sua atividade é principalmente desenvolvida nos tecidos e células neuronais^{31,34}.

Algumas características da BACE-1 permitem classificá-la como alvo terapêutico vantajoso para o tratamento da DA. A mais importante delas é ausência de outra enzima capaz de realizar uma atividade compensatória sobre a produção de A β , na ausência da BACE-1. Camundongos *knockout* do gene da BACE-1 apresentam redução drástica dos níveis de A β detectados no tecido cortical. Assim, a enzima *per se* é o principal requisito para a produção de A β e a sua inibição deve prevenir a acumulação de placas amiloides em pacientes com DA modificando, portanto, a patogênese da doença^{67,68}. Além disso, o fenótipo saudável observado nos animais *knockout*, para a BACE-1, sugere que a inibição da enzima é incapaz de desencadear efeitos adversos graves^{68, 69}. Adicionalmente, o sucesso precedente no desenvolvimento de fármacos inibidores de HIV protease, também uma protease aspártica, torna factível a obtenção de um inibidor de BACE-1 capaz de se tornar um fármaco clinicamente útil²⁶.

Tanto APP quanto a BACE-1 são proteínas transmembrana presentes na superfície da célula, contudo, o mecanismo catalítico ocorre intracelularmente após a endocitose de ambas as proteínas³¹. Assim, fármacos inibidores de BACE-1 devem apresentar propriedade que os permitam penetrar membranas subcelulares (Ex: LogP, pKa) bem como a barreira hematoencefálica já que o peptídeo A β é produzido no cérebro²⁶.

3.3.1. DESENVOLVIMENTO DE INIBIDORES DE BACE-1

Moléculas inibidoras de HIV protease e de renina (ambas aspartilproteases), que estão atualmente em uso clínico, são estruturas capazes de mimetizar quimicamente o estado de transição e concomitante impedem o processamento catalítico. A partir da elucidação desse mecanismo de ação o planejamento de inibidores de BACE-1, objetivou a descoberta de compostos capazes de atuar de maneira análoga⁷⁰. Assim, o

conhecimento acerca do mecanismo catalítico das aspartilproteases norteou a obtenção dos primeiros inibidores de BACE-1.

O processo de catálise se inicia com uma molécula de água (coordenada entre dois resíduos de aminoácidos aspárticos) sendo ativada pela abstração de um próton. Num segundo momento, esta molécula pode atuar como nucleófilo sobre o substrato. (Figura 7). Nesta reação observa-se a formação de um intermediário tetraédrico que desempenha um papel fundamental no mecanismo catalítico⁷¹. A partir desta informação idealizou-se que a substituição da ligação amídica do substrato por um grupamento semelhante, mas que apresente uma ligação não hidrolisável seria capaz de mimetizar o estado de transição e inibir a reação de hidrólise⁷².

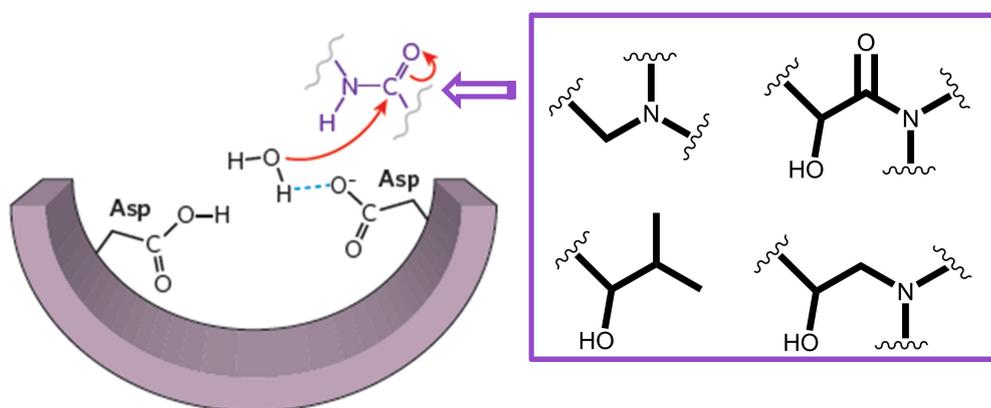


Figura 7: Mecanismo catalítico em aspartilproteases (à esquerda). Grupos não hidrolisáveis (à direita) presentes em compostos que mimetizam o estado de transição tetraédrico [Adaptado da ref. 71].

Um dos primeiros inibidores miméticos do estado de transição da BACE-1 foi a molécula OMM99-2 (Figura 8). Embora este composto seja muito potente ($K_i=1,6$ nM), ele apresenta baixa seletividade (ex: Catepsina D - $K_i=48$ nM), o que torna esta molécula propensa a facilmente desencadear efeitos colaterais indesejados⁷³. Por esta razão, o planejamento de moléculas subsequentes, a partir da estrutura OMM99-2 buscou elevar a seletividade frente a BACE-1⁷⁴.

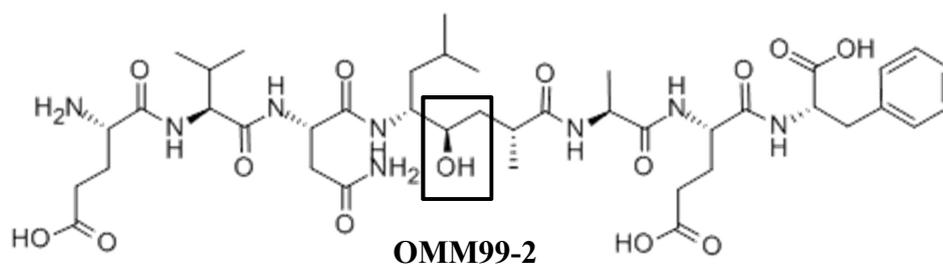


Figura 8: Inibidor de primeira geração de beta-secretase, baseado em um grupo de hidroxietileno (em destaque), isômero do estado de transição.

Para auxiliar no planejamento de novos inibidores estruturas cristalográficas com diferentes inibidores de BACE-1 foram determinadas. Atualmente são mais de 160 estruturas 3D de BACE-1 disponíveis no Banco de Dados de Proteínas (PDB, do inglês, *Protein Data Bank*), em complexo com moléculas de diversas classes, tais como derivados de isoftalamidas (1)⁷⁵, de hidroxietilaminas (2)^{76,77} (3)⁷⁸, (Figura 9) etc.

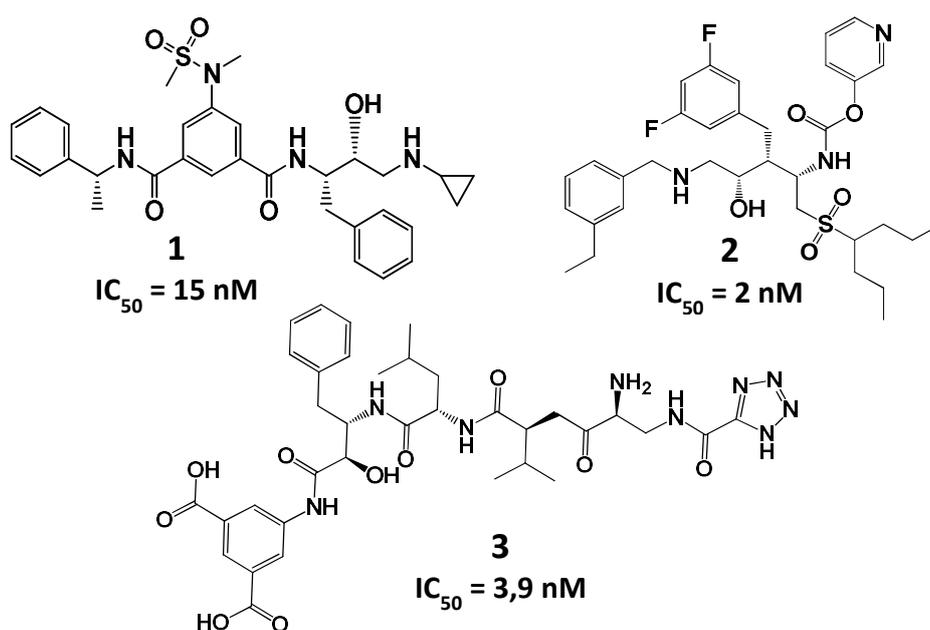


Figura 9: Exemplos de inibidores peptidomiméticos de BACE-1 planejados com base na estrutura do alvo macromolecular.

Embora os inibidores peptidomiméticos apresentem potência e eficácia na redução dos níveis do peptídeo A β nos ensaios *in vitro*, limitações relativas à sua permeabilidade pela BHE, suscetibilidade ao transporte pela glicoproteína P e sobretudo, baixa biodisponibilidade oral têm limitado o avanço desses compostos nas fases clínicas⁷⁹.

Visando superar esses problemas, moléculas com baixa massa molecular e apresentando caráter não peptidomimético têm sido propostas como inibidores de BACE-1. Nesse contexto, a caracterização do sítio de interação da BACE-1 permite, a partir dos estudos de relações estrutura-atividade, o desenho racional de moléculas capazes de realizar interações eficientes com a protease inibindo-a.

Utilizando ensaios em larga escala (HTS, do inglês, *High-throughput screening*), Malamas e colaboradores⁸⁰, identificaram uma molécula (**1**- Figura 10) com potência inibitória (IC_{50}) de 38 μ M. A partir da análise do composto em complexo com BACE-1 postulou-se que o anel tetrahidropirimidínico não representava pontos de contato favorável com o bolsão catalítico e assim, prosseguiu-se com o desenvolvimento de novas moléculas a partir da simplificação molecular e modificação bioisostérica do anel, levando aos compostos aminohidantoinínicos mais compactos. O ligante obtido (**2**) apresentou potência 10 vezes superior (IC_{50} = 3,4 μ M) e o acoplamento desta molécula no sítio da BACE-1 sugeriu pontos acessíveis para a substituição no ligante, capazes de melhorar a afinidade pela macromolécula e assim levar a obtenção de inibidores mais potentes (ex: **3** e **4**)⁸¹⁻⁸³.

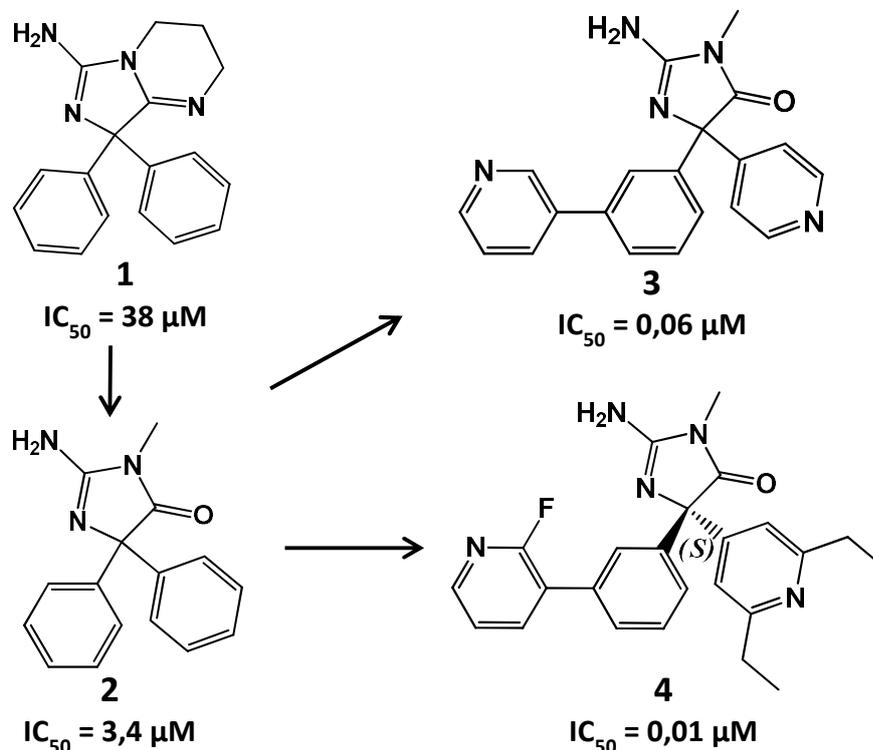


Figura 10: Desenvolvimento de inibidores não-peptidomiméticos de BACE-1.

Outra classe de inibidores sintetizados foram os derivados de acilguanidinas. Apesar da potência elevada desta classe de compostos (IC_{50} = ~50 nM), as estruturas

moleculares apresentam baixa permeabilidade pela BHE, provavelmente em função da hidrofilicidade do seu grupo central (acilguanidina)⁸⁴. Visando contornar este problema, Malamans e colaboradores empregaram como estratégia a ciclização do grupo acilguanidina visando com isso reduzir superfície polar da molécula. Posteriormente, o padrão de substituição no anel foi otimizado a fim melhor ocupar o sítio da BACE-1⁸⁵ (Figura 11).

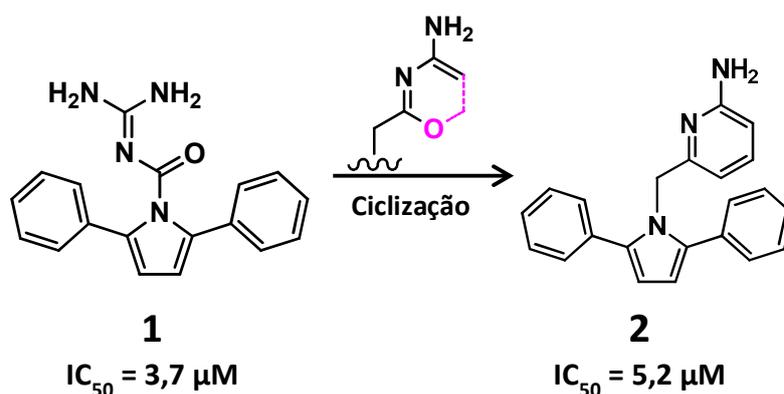


Figura 11: Ciclização do grupo Acilguanidino presente em **1** ($3,7 \mu M$) e obtenção do inibidor de BACE-1 Aminiporidinico **2** ($5,2 \mu M$)⁸⁵.

Apesar das várias estruturas cristalográficas fornecerem bases para estudos de relação estrutura-atividade, abordagens quantitativas para otimizar as características moleculares dos inibidores não peptidomiméticos BACE-1, objetivando aumentar a potência desses compostos, tem sido pouco explorada.

3.3. PLANEJAMENTO DE FÁRMACOS E AS FERRAMENTAS COMPUTACIONAIS

A química computacional e a bioinformática são ferramentas de química medicinal, cujas aplicações têm possibilitado, com relativo sucesso, a obtenção de moléculas com propriedades farmacológicas otimizadas, evitando os custos da síntese exploratória de um grande número de substâncias. O planejamento de fármacos auxiliado por computador (CADD do inglês, *Computer-aided drug design*) envolve todas as técnicas computacionais utilizadas para descobrir ou planejar novos compostos

biologicamente ativos ou ainda, aperfeiçoar ou melhorar as propriedades e características de moléculas que se tornarão fármacos⁸⁶.

Diversas estratégias dentro do campo da química medicinal podem ser empregadas no desenho e otimização de novos candidatos a agentes terapêuticos. Contudo, a aplicação de quaisquer desses métodos constitui etapa secundária no processo de planejamento, o qual deve estar fundamentado, irrefutavelmente, na definição prévia do alvo-terapêutico⁸⁷. Nesse contexto, os avanços expressivos da química e biologia, que tem contribuído para melhor compreensão de vias bioquímicas e mecanismos envolvidos nos processos fisiopatológicos permite também a identificação de alvos moleculares, cuja modulação é capaz de proporcionar efeitos benéficos no tratamento ou cura de doenças e disfunções em humanos⁸⁸.

A extensão das informações estruturais disponíveis sobre o alvo (enzima/receptor) e os ligantes determinam quais técnicas podem ser utilizadas nas etapas de planejamento de fármacos. Quando a estrutura do alvo é desconhecida, somente métodos de planejamento de fármacos baseado na estrutura do ligante (LBDD, do inglês, *ligand-based drug design*) podem ser utilizados, para explorar propriedades e características de séries de ligantes que são conhecidamente ativos. Por outro lado, no planejamento baseado na estrutura do receptor (SBDD, do inglês *structure-based drug design*) a forma e a natureza química do sítio de ligação do alvo macromolecular são conhecidos, e assim, aspectos espaciais e energéticos que afetam a afinidade podem ser considerados⁸⁹. Esta informação abre novas possibilidades para impulsionar a identificação de moléculas líderes através do acoplamento de ligantes (selecionados de bancos de dados de estruturas químicas) no sítio ativo da proteína ou pelo uso de modelos baseados em farmacóforos. Além disso, candidatos a fármacos podem ser desenhados *de novo* visando a melhora nas propriedades de complementariedade (físico-químicas e farmacocinéticas) de compostos líderes pela respectiva proteína-alvo⁹⁰.

3.3.1. DELINEAMENTO DOS ESTUDOS DE QSAR

Os estudos de relações estrutura-atividade quantitativas (QSAR do inglês *Quantitative Structure-Activity Relationships*) são baseados no princípio de que a

atividade biológica de um ligante está relacionada à sua estrutura molecular. Sendo uma abordagem de LBDD, os métodos de QSAR se constituem como uma das mais poderosas técnicas para prever a bioatividade de moléculas por meio de modelos teóricos⁹¹. Além disso, é possível propor racionalmente alterações estruturais que levem a obtenção de moléculas com propriedades farmacocinéticas e farmacodinâmicas superiores àquelas do conjunto de dados inicialmente disponível⁹².

O desenvolvimento dos modelos de QSAR envolve um processo sistemático com várias etapas. Inicialmente, deve-se fazer a seleção do conjunto de dados, onde um grupo de moléculas com atividade biológica contra um determinado alvo é conhecido e atuam por um mecanismo de ação comum. De preferência, os compostos são escolhidos a partir de um único laboratório ou fonte, a fim de evitar inconsistências e variabilidade dos bioensaios de propriedade biológica⁹³.

A segunda etapa dos estudos de QSAR objetiva o cálculo dos descritores, que representarão as propriedades moleculares (ex: eletrônicas, geométricas, topológicas, etc.) implícitas na estrutura química de cada composto. Posteriormente, é realizada uma busca da correlação desses descritores com a atividade biológica⁹⁴. Alguns programas computacionais como o DRAGON⁹⁵, TOPS-MODE⁹⁶ e CODESSA⁹⁷, incorporam conceitos teóricos e gráficos para caracterizar a estrutura química de compostos fornecendo uma grande variedade de descritores que podem ser utilizados para o desenvolvimento de modelos de QSAR. Algumas classes de descritores incluem: i) descritores físico-químicos (ex: logP, logD, solubilidade, pKa, etc); ii) parâmetros constitucionais (ex: número de ligações simples, duplas, triplas, anéis aromáticos); iii) descritores eletrônicos (ex: carga formal, carga parcial, momento e vetores dipolo, energia HOMO e LUMO); iv) parâmetros topológicos (ex: índice de conectividade molecular), etc²⁸.

Após a etapa de obtenção dos descritores recomenda-se o pré-tratamento dos dados a fim de excluir variáveis que não contém informações moleculares relevantes para descrever o fenômeno sob investigação, tais como aquelas com baixa variância e alta correlação⁹². Os descritores remanescentes desta etapa são reunidos aos valores de propriedade biológica para a construção dos modelos de QSAR. Como o número de variáveis descritivas é marcadamente superior ao número de variáveis experimentais, uma calibração multivariada deve ser utilizada para a correlação dos dados. O método mais simples, é o de regressão linear múltipla, RLM (MLR, do inglês *multiple linear regression*), na qual o dado a ser predito (y) é estimado a partir da combinação linear dos

termos independentes ($x_1, x_2, x_3 \dots$) contidos em uma matriz X, de acordo com a equação 1⁹⁸:

$$y = b_1x_1 + b_2x_2 + \dots + b_nx_n \quad \text{(Equação 1)}$$

Onde $b_1, b_2, b_3, \dots, b_n$ são coeficientes de ajuste da regressão.

Embora o sistema matricial de dados apresente um total de variáveis descritivas (k), o ideal é que apenas um subconjunto destas seja utilizado para que a equação de regressão linear tenha uma solução única. A incorporação excessiva de k ao modelo limita os graus de liberdade oferecidos aos desvios entre os valores reais e preditos fazendo com que os mesmos sejam artificialmente reduzidos. Com isto há um sobreajuste (do inglês, *overfitting*) dos dados, com a obtenção de valores elevados do coeficiente de correlação, que não correspondem aos do ajuste natural das variáveis em relação aos valores observados de atividade⁹⁹. O programa MOBYDIGS 1.0¹⁰⁰ por exemplo, emprega um subgrupo de descritores que são utilizados na construção de modelos de regressão linear múltipla (MLR, do inglês *multiple linear regression*), com algoritmo genético.

Algoritmos genéticos (AG) são inspirados na teoria da evolução das espécies, e amplamente empregados na otimização e buscas em sistemas complexos. Os métodos de AGs adotam terminologia similar à usada na teoria da evolução natural, onde cada indivíduo (solução), denominada de “*cromossomo*”, é formado por “*genes*” os quais correspondem às variáveis do problema. Operadores genéticos de “*mutação*” e “*cruzamento (crossover)*” são empregados para obter soluções que formam as “*populações*”¹⁰¹.

Uma vez estipulados os parâmetros de inicialização (população inicial) do algoritmo, os indivíduos são gerados aleatoriamente. A cada um destes, está associado um valor de aptidão (*fitness*) que representa a qualidade de tal solução tendo em vista a função objetivo a ser otimizada. A partir de um processo denominado avaliação, os modelos são classificados e ordenados de acordo com o valor do coeficiente de aptidão, sendo esta classificação utilizada como critério para a seleção de indivíduos que serão os “pais” da geração seguinte. Os indivíduos sofrem um evento reprodutivo em que seus cromossomos são recombinados entre si de acordo com a probabilidade de “*crossover*”. E, ainda, os indivíduos podem ter seus genes (descritores) aleatoriamente alterados

(mutação) ou não¹⁰¹. Se o “filho” gerado coincide com um dos indivíduos já presentes na própria população, este é rejeitado, caso contrário, a equação é avaliada e sendo o seu valor da função objetivo melhor do que o pior valor presente em outros modelos da população haverá uma reclassificação dos indivíduos. Todo o ciclo se repete e a cada nova geração, os indivíduos mais bem “adaptados” sobrevivem, sendo portanto, selecionados para compor a população de modelos que otimizam o valor da função objetivo¹⁰⁰.

Outro procedimento, que pode ser utilizado para a redução dos termos descritivos, consiste em encontrar combinações lineares das variáveis individuais que descrevam grande parte da informação original como no método de Análise dos Componentes Principais (PCA do inglês *Principal Component Analysis*)¹⁰².

A PCA é um método de análise multivariada não supervisionado utilizado para projetar dados n -dimensionais em um espaço de menor dimensão¹⁰³. Nesta abordagem o espaço original, ocupado pelos descritores em inúmeras dimensões é projetado linearmente em um subespaço com dimensão inferior que é representado por variáveis ortogonais denominadas Componentes Principais (PCs). Nesse processo, os descritores são analisados quanto a informação contida em cada um deles e quanto a correlação existente entre eles, sendo os grupos de descritores com comportamento análogo identificados e combinados para a construção das PCs (figura 12). Dessa maneira, preserva-se a maior informação possível do conjunto de dados inicial e minimiza-se a perda de informações estatisticamente significativas¹⁰⁴.

O novo sistema gerado terá os eixos representados pelas PCs, que são ordenadas da seguinte forma: a primeira componente (PC1) apresenta maior variância; a segunda componente (PC2) apresenta segunda maior variância, e assim por diante¹⁰⁵. De um modo geral, as combinações lineares das m -variáveis originais podem ser representadas pela equação 2:

$$PC_i = a_{i1}v_1 + a_{i2}v_2 + \dots + a_{im}v_m \quad \text{(Equação 2)}$$

Em que, v_1, v_2, \dots, v_m são os descritores originais e $a_{i1}, a_{i2}, \dots, a_{im}$ são os coeficientes que medem a importância de cada descritor na i -ésima componente principal (PC $_i$), ou seja, o peso que cada um deles tem naquela combinação linear¹⁰⁶.

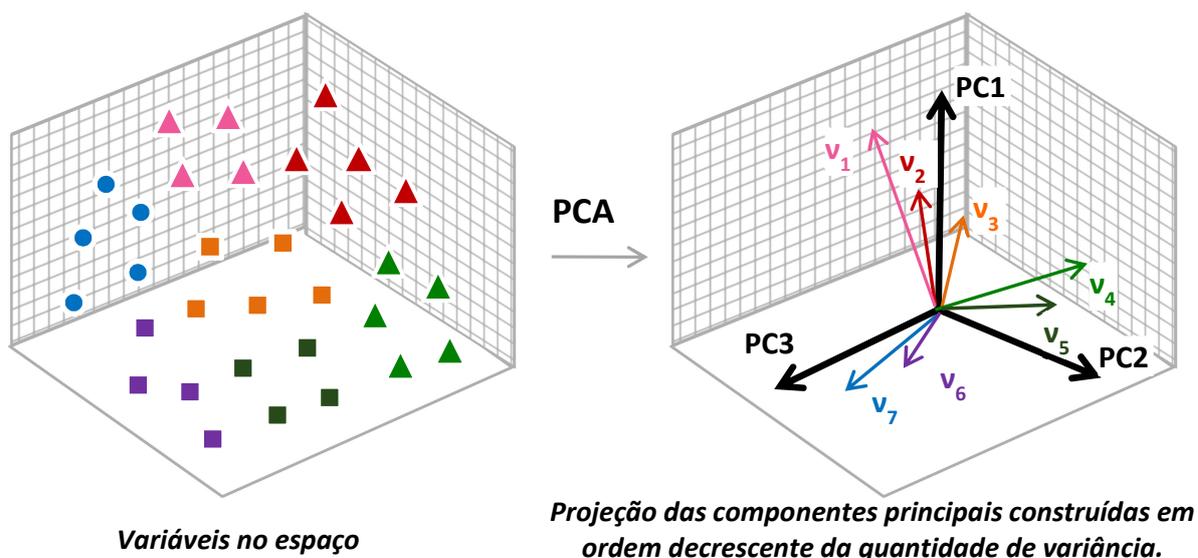


Figura 12: Redução da dimensionalidade dos dados por PCA.

Para encontrarmos a relação entre os descritores e os valores de propriedade, utiliza-se o método dos mínimos quadrados parciais (PLS, do inglês *partial least squares*). Ao contrário da técnica de PCA, o método de PLS é uma técnica de reconhecimento padrão supervisionada, pois a atividade é utilizada durante o cálculo das PCs. Nesta abordagem, uma PCA é aplicada às amostras, porém simultaneamente à obtenção das PCs, uma correlação matemática com as variáveis dependentes é procurada¹⁰⁷.

O método de PLS oferece vantagens sobre o método de MLR, pois a limitação para a razão entre o número de descritores e o número de compostos é reduzida, bem como não há sensibilidade às colinearidades entre as variáveis descritivas¹⁰⁴. Além disso, na técnica de MLR todas as variáveis selecionadas são incorporadas na matriz de correlação, sem quaisquer parâmetro de distinção, quanto ao grau de relevância das mesmas. Assim, uma quantidade significativa de informação irrelevante é incorporada, o que não ocorre quando se usa o método PLS^{98,99,104}.

Uma vez construído o modelo de QSAR, este deve ser validado. A quantidade de variância predita dos modelos de QSAR é determinada pelo processo de validação cruzada. Neste método, variáveis são sistematicamente removidas da matriz original de dados, que uma vez reduzida é então utilizada para a geração de um modelo de QSAR. A partir deste modelo, o valor de propriedades das respectivas variáveis removidas é

predito. Os resíduos, diferenças entre valores reais de atividade biológica observados e os valores estimados pelo processo, são empregados no cálculo do erro de predição. Esse processo é repetido até que todos os compostos tenham sido removidos ao menos uma vez¹⁰⁸. A validação cruzada é útil para determinar o número de componentes principais a ser utilizada, pois o número ótimo de componentes é aquele que forneceu a menor soma dos quadrados dos erros de predição¹⁰⁴.

A partir dos resultados obtidos na validação, são calculadas a soma dos quadrados dos resíduos (PRESS) e a soma dos quadrados dos desvios originados pela diferença entre os valores reais de atividade de cada composto e a média do grupo (SS). Esses dados são usados na obtenção do coeficiente de correlação por validação cruzada (Equação 3)¹⁰⁹:

$$Q^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{PRESS}{SS} \quad \text{(Equação 3)}$$

Onde: n é o número total de objetos da base de dados

\hat{y}_i é o valor predito de atividade biológica para o composto i do grupo;

y_i é o valor real de atividade biológica do composto i ;

\bar{y} é a média aritmética da atividade biológica do grupo.

Por fim, a robustez do modelo é verificada pela etapa de *validação externa*, na qual se analisa a sua capacidade de predizer a atividade de um conjunto de moléculas congêneres (conjunto teste) que não foi utilizado para a construção do modelo original. O critério estatístico usado neste caso é o coeficiente de correlação de predição (r^2_{pred}), que pode ser calculado de maneira idêntica à equação 3, porém neste caso, \hat{y}_i e y_i passam a ser respectivamente, o valor real e valor predito de atividade biológica para o composto i do conjunto teste¹¹⁰.

3.3.2. HOLOGRAMA QSAR

Um exemplo de modelo de QSAR baseado em fragmentos moleculares é o Holograma QSAR (HQSAR)¹¹¹. Esta técnica oferece vantagens sobre os métodos 2D clássico e de QSAR 3D pois não é necessário a seleção de descritores físico-químicos,

alinhamento molecular ou determinação da conformação bioativa e, portanto, permite a análise automatizada de grandes conjuntos de dados sem intervenção manual¹¹².

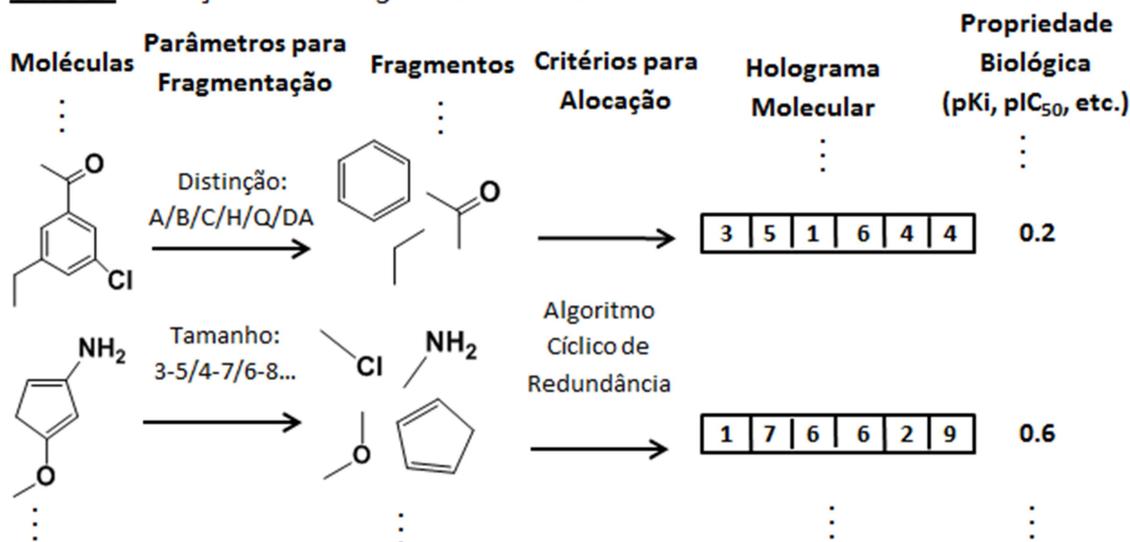
O primeiro passo no método de HQSAR é a decomposição das moléculas da base de dados em todos os fragmentos estruturais possíveis os quais são utilizados para a construção do holograma molecular. Cada holograma é formado por um conjunto de posições onde são contabilizados o tipo e frequência de cada fragmento. Dessa forma obtém-se uma codificação mais precisa que a impressão digital bidimensional tradicional no qual a frequência dos fragmentos não é levada em consideração¹¹³.

Durante o processo de construção dos modelos de HQSAR (Figura 14) pode-se testar de maneira interativa a combinação de vários parâmetros relevantes na obtenção dos hologramas, como a distinção e tamanho dos fragmentos, além do próprio comprimento, do holograma. Os fragmentos podem ser diferenciados com base nos tipos de átomos (A), na ligação (B), conectividade ou estado de hibridização atômica (C), presença de hidrogênios (H), quiralidade (Ch) e na propriedade dos átomos em doar ou receber ligações de hidrogênios (DA). Esses fragmentos deverão ter tamanho fixo, número máximo e mínimo de átomos determinado pelo usuário em cada análise¹¹¹⁻¹¹⁴. Como na maioria das vezes o tamanho do holograma é menor que a quantidade de fragmentos gerados, fragmentos idênticos ou semelhantes são alocados no mesmo espaço conforme os critérios designados pelo algoritmo cíclico de verificação de redundância (CRC, do inglês, *Cyclic Redundancy Check*)^{114,115}. Na segunda etapa do HQSAR, uma equação matemática de regressão é derivada para correlacionar os hologramas moleculares de cada composto com os valores correspondentes de atividades biológicas através do método dos mínimos quadrados parciais (PLS)¹¹¹⁻¹¹⁵ (Figura 13).

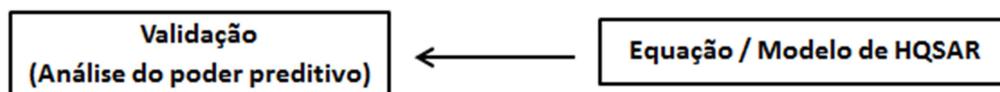
A técnica de HQSAR permite determinar através dos mapas de contribuição 2D, quais dos fragmentos moleculares contribuem positivamente ou negativamente para a atividade biológica dos compostos, guiando assim, a síntese de moléculas com propriedades otimizadas. A contribuição de cada átomo para a atividade é calculada da seguinte forma: inicialmente, os fragmentos são identificados nos hologramas e para cada um destes é designado o correspondente coeficiente de correlação gerado pelo método de regressão PLS (esse valor é tido como o valor da contribuição do fragmento para a atividade). Numa etapa posterior a contribuição é determinada a partir da razão entre o PLS e o número de átomos do fragmento. Assim, todos os átomos assumem contribuições equivalentes para atividade de um fragmento específico. A análise da contribuição de

cada grupo para a atividade é realizada através de um código de cores que descreve quantitativamente o resultado final¹¹⁴.

PASSO 1: Geração dos Hogramas Moleculares



PASSO 2: Geração e validação dos modelos



PASSO 3: Interpretação dos mapas de contribuição

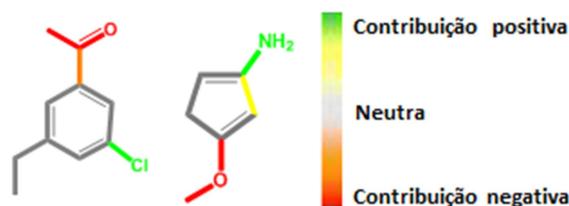


Figura 13: Etapas da construção de modelos de Holograma QSAR.

3.3.3. ANÁLISE COMPARATIVA DOS CAMPOS MOLECULARES (CoMFA)

A principal abordagem de QSAR-3D empregada atualmente é a Análise Comparativa dos Campos Moleculares (CoMFA do inglês *Comparative Molecular Field Analysis*)¹¹⁶. Os compostos selecionados para a construção de um modelo de QSAR devem ter um farmacóforo comum, porém não existe restrição quanto ao grau de similaridade estrutural dentro da série a ser analisada, como ocorre na maioria dos

métodos bidimensionais¹¹⁷. O fato de moléculas estruturalmente diferentes se ligarem de maneira semelhante no sítio de um mesmo alvo sugere que o receptor não reconhece o ligante pela sua estrutura química, mas reconhece as propriedades que estão envolvidas nas interações ligante-receptor. Partindo desse princípio, o método de CoMFA correlaciona as semelhanças na forma e arranjo espacial, associadas às propriedades intermoleculares, tais como interações de Van der Waals, interações eletrostáticas e hidrofóbicas, com as diferenças observadas na atividade biológica dos compostos sob investigação¹¹⁸.

O alinhamento conformacional dos ligantes é um dos fatores mais importantes na geração dos modelos de CoMFA, pois os descritores tridimensionais são amplamente sensíveis à posição e orientação das estruturas moleculares no espaço e assim, pequenas variações na orientação espacial das moléculas resultam em alterações significativas nos modelos gerados. Assim, a etapa de alinhamento representa um dos principais desafios do métodos de CoMFA^{119, 120}.

Para obter a comparação nas formas e arranjos das moléculas, a sobreposição estrutural dos compostos é requerida e assim, a identificação da ‘conformação bioativa’, é especialmente importante. Como, em geral, poucos ou nenhum dos compostos da série em estudo tem a sua conformação bioativa conhecida, torna-se necessário modelar a provável orientação dos ligantes¹¹⁸. Quando a informação tridimensional do alvo é desconhecida o alinhamento estrutural dos compostos pode ser feito considerando a conformação de menor energia. Neste caso, assume-se que as conformações mais estáveis são próximas da conformação bioativa e podem descrever as propriedades essenciais para a atividade biológica. Ainda que tais propriedades, na maioria das vezes, não sejam identificadas em uma única estrutura elas são exibidas quando outras moléculas ativas são adicionadas gerando um modelo de sobreposição¹²¹.

Por outro lado, quando há dados cristalográficos para o complexo ligante-receptor, o acoplamento do ligante no sítio fornece dados adicionais, uma vez que o inibidor ligado ao sítio ativo exibe o modo de interação que moléculas estruturalmente análogas devem adotar quando estão no sítio de ligação da macromolécula. Nesse sentido, pode-se utilizar a técnica de acoplamento molecular para guiar a escolha da conformação bioativa¹²².

3.3.3.1. Alinhamento Estrutural Rígido pela Máxima Estrutura Comum

A sobreposição de moléculas flexíveis envolvem problemas relacionados aos graus de liberdade rotacional e translacional das moléculas. Em muitos métodos, procedimentos de amostragem do mínimo local são realizados para a busca conformacional e na sequência as estruturas obtidas são sobrepostas como corpos rígidos¹²³.

O primeiro passo para o alinhamento rígido é delinear quais partes das moléculas formam pontos e/ou núcleos subestruturais comuns. Esta caracterização irá determinar os fragmentos e átomos que serão utilizados como base na superposição das estruturas. O ligante mais potente da série em estudo pode ser usado como modelo para a marcação dos pontos comuns, e a seguir, as demais moléculas nas suas respectivas conformações de baixa energia são então sobrepostas individualmente na molécula de referência usando-se um procedimento de forma a minimizar o desvio da raiz média quadrada (*RMSD*, do inglês *root-mean-square deviation*) da posição entre os átomos/ centros comuns dos ligantes¹²⁴.

3.3.3.2 Alinhamento Estrutural Baseado na Similaridade Morfológica

Outra estratégia de alinhamento estrutural, que dispensa a informação 3D da macromolécula prevê a comparação dos compostos do conjunto de dados, com base nas suas similaridades morfológicas. Para isto, são considerados pontos que levam em conta as distâncias mínimas para a superfície de Van der Waals e as distâncias mínimas para a superfície de átomos doadores e de aceptores de ligação de hidrogênio em uma determinada molécula tomada como objeto de referência. Esses pontos formam triângulos, cujas áreas são comparadas a de outras moléculas na busca por características correspondentes/ semelhantes. Ou seja, quando outra estrutura é sobreposta ao objeto de referência, são computadas as diferenças na morfologia e no volume de ambos, considerando várias posições de alocação. A sobreposição que fornece a máxima similaridade na forma das áreas e o mínimo volume total ocupado representa a melhor solução (Figura 14)^{125, 126}.

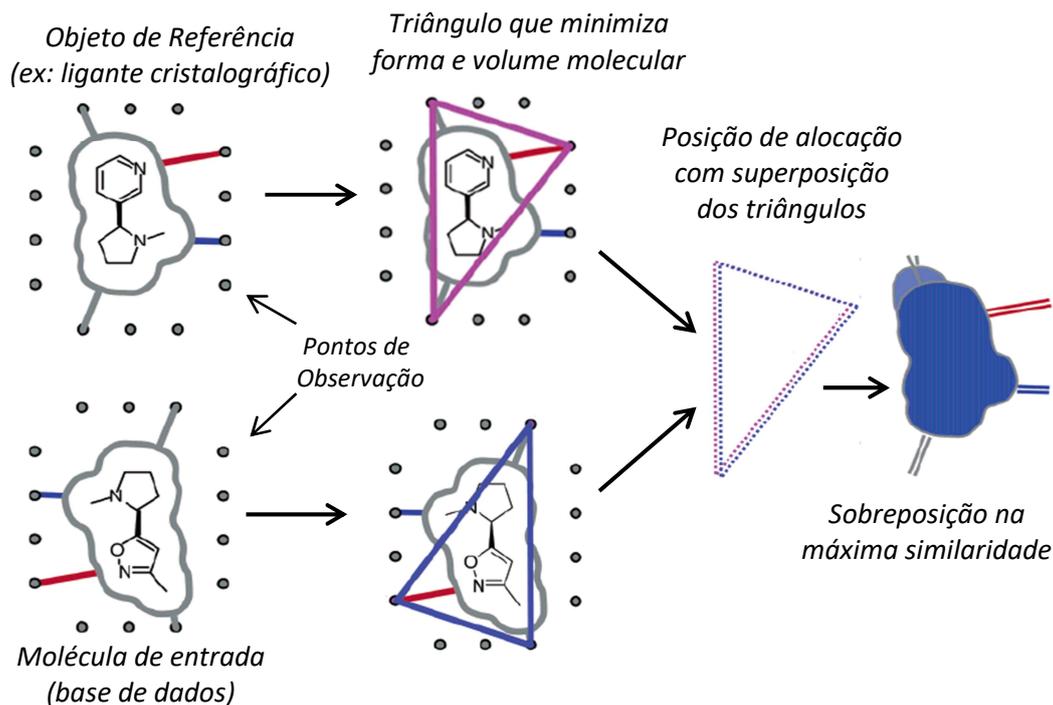


Figura 14: Alinhamento Molecular baseado na similaridade Morfológica (Adaptado da ref. 126).

O alinhamento estrutural construído com o auxílio do SURFLEXSIM (plataforma Sybyl X.1.1) é uma abordagem que combina o método de similaridade morfológica descrito com um algoritmo de construção incremental¹²⁷. Neste caso, os ligantes são decompostos em fragmentos, sendo gerado para cada um deles um conjunto de conformações, que são testadas quanto a sua semelhança tridimensional em relação ao objeto de referência (ligante cristalográfico). A sobreposição de cada conformação com o par relacionado no ligante gera uma hipótese do modo de interação com o alvo e assim, é realizado uma busca para encontrar, entre as hipóteses, as soluções que apresentam máxima semelhança entre si e mínimo volume total. A partir destes critérios, as conformações são ranqueadas e a de maior pontuação é classificada com líder. O processo é repetido para todos os fragmentos e a partir do fragmento líder preliminar uma construção incremental é iniciada, na qual os demais fragmentos são interativamente selecionados e ligados, até a obtenção das estruturas completas¹²⁶.

3.3.3.3. Alinhamento Estrutural Baseado em Acoplamento Molecular

Objetivando estabelecer de forma eficiente o modo de interação e a afinidade de pequenas moléculas no sítio de ligação de proteínas, diversos algoritmos de acoplamento molecular foram desenvolvidos e aprimorados nos últimos anos¹²⁸. Como exemplo, o uso do algoritmo genético (AG) tem sido empregado para explorar a variabilidade conformacional que ligantes flexíveis, podem adotar dentro do sítio de ligação da macromolécula^{129, 130}.

Conforme descrito previamente, os AG aplicam conceitos derivados da genética e da teoria da evolução para a otimização e busca de soluções complexas. Neste caso, o objetivo é a busca conformacional da forma bioativa dos ligantes, como implementado, por exemplo, pelos programas Autodock e GOLD (do inglês, *Genetic Optimisation for Ligand Docking*)¹³¹. O AG do GOLD parte de uma população inicial de diferentes conformações dos ligantes as quais são definidas por um conjunto de variáveis (*genes*) responsáveis pela descrição de aspectos relacionados aos graus de liberdade do ligante em relação ao receptor. Gerações sucessivas de indivíduos (conformações) são criadas através da troca de genes (recombinação ou cruzamento) e mudanças genéticas aleatórias (mutação), permitindo que as populações evoluam para melhores soluções (Figura 15), as quais são avaliadas em relação a uma função de ajuste pré-definida¹³².

Quando se utiliza o programa GOLD as funções comumente empregadas são ChemScore e Goldscore. A função Goldscore é baseada em campos de força padrões e compreende a energia de interação entre o receptor e o ligante (definida pelos termos de energia de ligação de hidrogênio do complexo receptor-ligante e energia de ligação de van der Waals) e a energia interna do ligante (definidas pelos termos de energia de ligação de hidrogênio intramolecular e de van der Waals do ligante)^{131,132}. Já a função Chemscore é empírica e assim, visa obter com mais precisão o cálculo da energia livre de interação do ligante com a macromolécula a partir do ajuste das equações provenientes dos campos de força, utilizando dados experimentais¹²⁸.

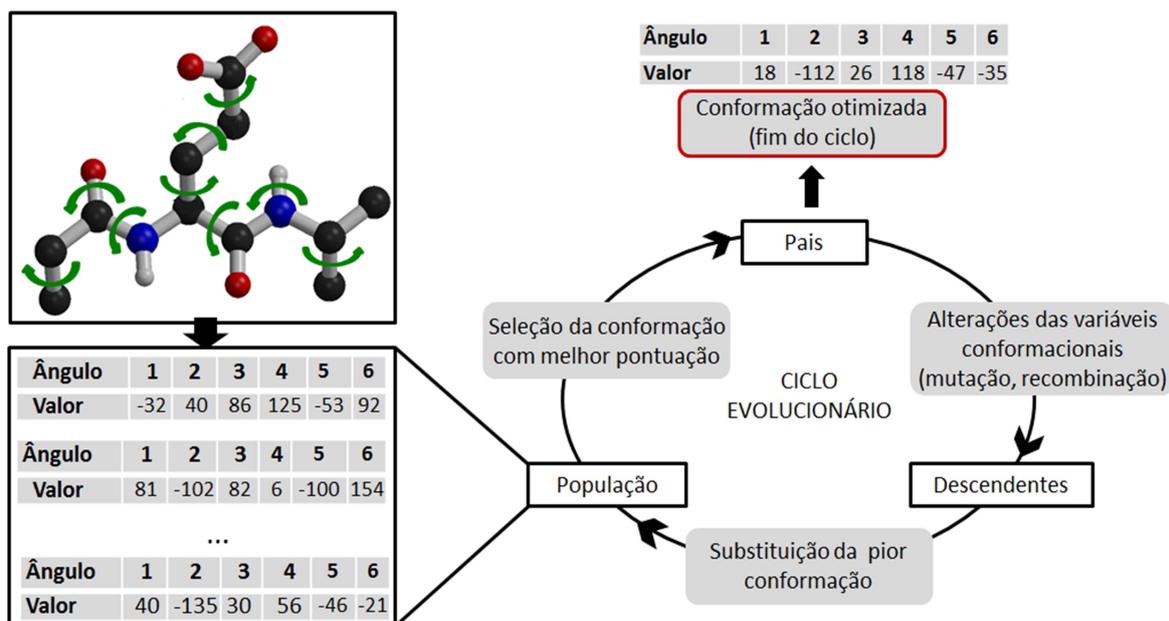


Figura 15: Busca Conformacional por Algoritmo Genético implementada no programa GOLD.

Os resultados do acoplamento são selecionados a partir da função escolhida e o número de soluções é definido pelo usuário. Em geral, após o acoplamento molecular é feita uma inspeção visual de cada solução levando-se em conta a complementaridade com o sítio, os quais as interações estão sendo feitas e, sobretudo, a semelhança da solução proposta com a conformação do ligante cristalográfico¹³³.

3.3.3.4. A Construção dos Campos Moleculares e Geração dos Modelos

O objetivo dos estudos de QSAR 3D CoMFA é determinar uma correlação entre os campos moleculares dos compostos e suas respectivas atividades biológicas. Para isto, uma caixa virtual que simula o receptor (do inglês *grid*) é construída ao redor das moléculas alinhadas. As paredes da caixa devem encobrir todo o volume ocupado pelas estruturas dos compostos e então, as energias de interação entre as moléculas dos ligantes e o *grid* serão calculadas nos pontos de intersecção usando-se átomos ou grupos de prova (ex.: oxigênio com hibridação sp^3 e carga -1, carbono com hibridação sp^3 e carga +1 e hidrogênio com carga +1). Os cálculos dos campos moleculares são feitos considerando-

se a função de energia potencial de Lennard-Jones (estéreo) e de Coulomb (eletrostático)¹²¹.

O potencial de Lennard-Jones (Equação 4) é um dos modelos matemáticos que representa o comportamento das forças que atuam sobre átomos e moléculas neutras: uma força atrativa (*i.e.*, força de Van der Waals) e uma força de repulsão de curto alcance, que ocorre entre os núcleos atômicos¹¹⁷:

$$V_{LJ} = \sum_{i=1}^n (A_{ij} r_{ij}^{-12} - C_{ij} r_{ij}^{-6}) \quad \text{(Equação 4)}$$

Onde, A_{ij} e C_{ij} são parâmetros específicos de V_{LJ} para cada átomo considerado; r_{ij} corresponde à distância entre o átomo i da molécula e o ponto j da grade onde a carga de prova é alocada; O termo r_{ij}^{-12} corresponde ao potencial repulsivo de curto alcance, decorrente de distorções das nuvens eletrônicas e prevalece quando a distância entre os átomos é pequena, enquanto que o termo r_{ij}^{-6} descreve o potencial atrativo de longo alcance, sendo predominante quando a separação entre os átomos aumenta em magnitude.

O potencial de Coulomb se relaciona com as forças de longo alcance entre as cargas elétricas dos átomos. Dessa forma, a energia potencial de interações eletrostática (V_C) no método de CoMFA é calculada como¹¹⁷:

$$V_C = \sum_{i=1}^n \frac{q_i q_j}{D r_{ij}} \quad \text{(Equação 5)}$$

Onde q_i e q_j correspondem respectivamente, às cargas dos átomos da molécula e do átomo de prova, que estão interagindo a uma distância r_{ij} ; e D é a constante dielétrica relativa do meio.

A energia de interação em cada ponto da grade 3D é armazenada, junto a atividade biológica de cada composto numa tabela virtual. Como nesta abordagem é gerado um grande número de variáveis, comparado ao número de compostos, as energias obtidas serão relacionadas às atividades biológicas pela técnica de mínimos quadrados parciais (PLS)¹²¹ (Figura 16).

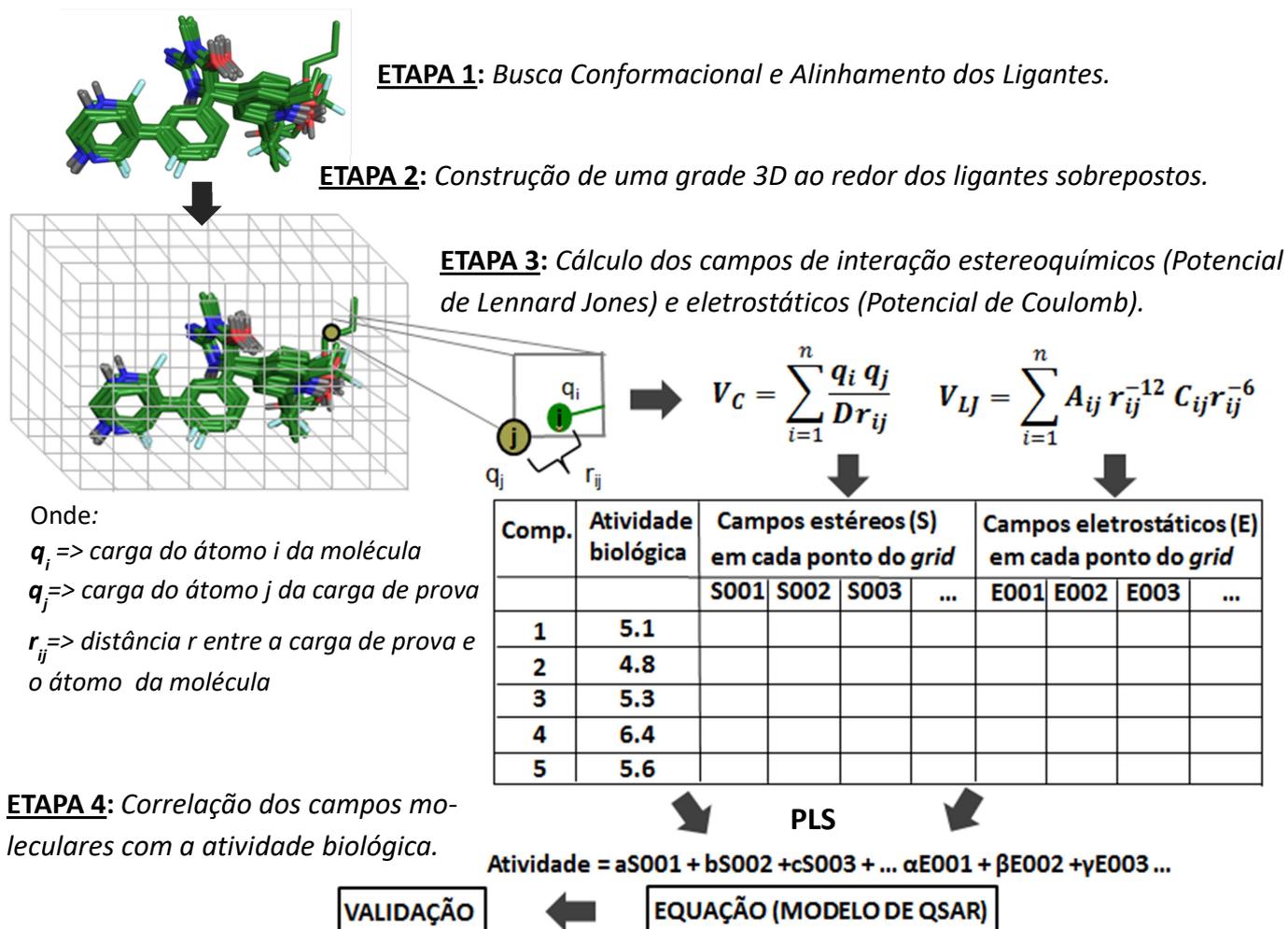


Figura 16: Etapas para a construção de modelos de QSAR 3D por CoMFA.

4. MÉTODOS COMPUTACIONAIS

4.1. CONJUNTO PADRÃO DE DADOS

Para o delineamento experimental, foi selecionado um grupo de 102 compostos com atividade sobre a enzima BACE-1 humana, publicado por *Michael Malamas* e colaboradores^{80-83,85} cuja potência inibitória relatada na forma de IC₅₀ (concentração requerida para a inibição de 50% da atividade enzimática) varia entre 0,01µM e 45 µM.

Os valores foram convertidos em pIC_{50} ($-\log IC_{50}$) e usados como variável dependente nas análises de QSAR. A estrutura química dos compostos e os seus respectivos valores de propriedade são apresentados no Apêndice A. Os compostos selecionados são derivados aminoimidazólicos (**1-20**), aminohidantoinícos (**21-74**) e aminopiridínicos (**75-102**).

As moléculas inicialmente tiveram suas estruturas desenhadas em formato 2D utilizando o programa ChemDraw 12.0 e, em seguida, foram convertidas para o formato 3D, sendo então submetidas aos procedimentos de otimização por mecânica molecular (campo de força Tripos) e as cargas fornecidas pelo método de AM1 (Palavras-chave: 1SCF XYZ ESP NOINTER SCALE=1.4 NSURF=2 SCINCR=0.4), como disponível na plataforma computacional Sybyl-X.1.1.

Após esta etapa preparatória, os compostos foram divididos em conjunto treino (75 compostos) e conjunto teste (27 compostos), aleatoriamente. O Método de Análise Hierárquica de Cluster (HCA), realizado com o auxílio do programa PIROUETTE 4.0, foi empregado a fim de garantir que a diversidade estrutural em ambos os conjuntos fosse aceitável/semelhante.

4.2. QSAR 2D HOLOGRAMA QSAR

O módulo de HQSAR em Sybyl-X.1.1 foi empregado na geração de modelos de relações estrutura-atividade quantitativas baseado em fragmentos. Inicialmente investigou-se a influência da distinção de fragmentos na qualidade estatística dos modelos. Para isto, foi conduzida uma combinação interativa dos parâmetros que determinam a distinção dos fragmentos: átomos (A), ligações (B), conectividade (C), átomos de hidrogênio (H), quiralidade (Ch), e átomo doador e receptor de ligação de hidrogênio (DA). O número de posições, que condicionam o tamanho do holograma foi determinado segundo uma estratégia tentativa e erro com diferentes valores padrão para o comprimento (53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353 e 401). Nas etapas preliminares foi considerado o tamanho de fragmento padrão (4-7), e para o melhor modelo obtido através da combinação dos parâmetros descritos acima, a influência do tamanho de fragmento (3-5, 6-8, 7-10) foi investigada.

4.3. MODELOS QUIMIOMÉTRICOS 2D BASEADO EM DESCRITORES TOPOLÓGICOS

O programa DRAGON 5.5 (Talette SRL) foi utilizado no cálculo de descritores 2D (geométricos, topológicos, constitucionais, autocorrelação 2D, etc). Para a pré-seleção das variáveis foi estabelecido como critério a eliminação dos dados que apresentaram valores constantes, baixa variância ($< 0,01$) ou alta correlação ($>0,97$). A etapa seguinte foi baseada na construção de modelos de regressão linear múltipla (MLR), com até 4 descritores através do método de algoritmo genético como disponível do programa MOBYDIGS 1.0. Essa abordagem computacional, conforme descrito na fundamentação teórica, busca identificar indivíduos (modelos matemáticos) que sejam bem adaptados à condição de evolução estabelecida. Para tanto, o algoritmo aplica os operadores de mutação e cruzamento no universo de descritores previamente calculados com auxílio do programa DRAGON 5.5 e indivíduos (modelos) evoluem a fim de maximizar os valores de q^2 . Devido a natureza estocástica desse processo, foi estabelecido que houvesse repetição do mesmo para um número de 10 populações diferentes, cada uma contendo 100 indivíduos, que evoluíram por pelo menos 10000 gerações. Os descritores presentes nos 10 melhores modelos foram reunidos, autoescalados e utilizados para análise exploratória dos dados através da análise de componentes principais (PCA) e para geração dos modelos de QSAR aplicando-se o método de regressão por mínimos quadrados parciais (PLS), como disponível no software PIROUETTE 4.0.

4.4. ALINHAMENTO ESTRUTURAL DOS COMPOSTOS

4.4.1. Alinhamento por Máxima Subestrutura Comum (Alinhamento I)

O alinhamento pela máxima estrutura comum foi realizado com as estruturas dos ligantes na sua forma minimizada, não sendo considerada a flexibilidade dos mesmos durante o processo. Para a obtenção desse alinhamento (alinhamento I) foram tomados como referência um anel benzênico e um grupamento amino ligado a um carbono sp^2

(Figura 17) presentes na estrutura do composto **60** ($pIC_{50} = 8,0$ – Apêndice A), sobre os qual os demais compostos foram alinhados, como disponível na ferramenta *Align_database*, da plataforma Sybyl X 1.1.

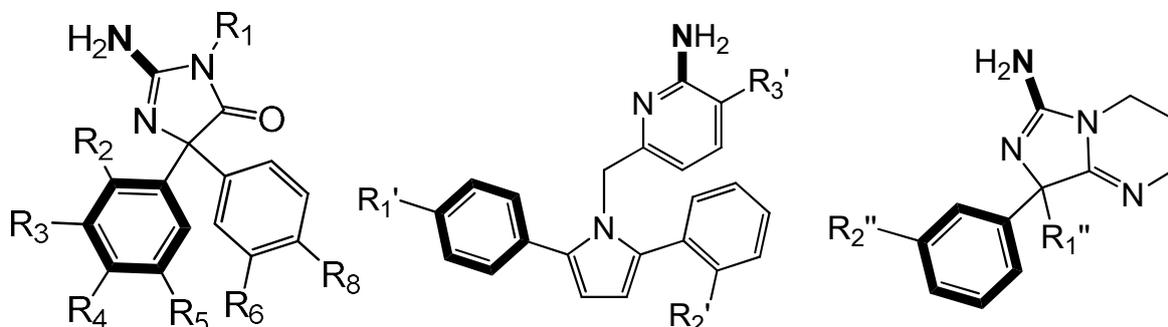


Figura 17: Esqueleto geral dos inibidores de BACE-1, onde estão destacados os átomos usados como referência para gerar o alinhamento I.

4.4.2. Alinhamento por Similaridade Morfológica (Alinhamento II)

Na segunda estratégia de alinhamento as moléculas foram sobrepostas com base na sua similaridade morfológica, utilizando o módulo SUERFLEX-SIM disponível na plataforma Sybyl X.1.1. Para este método, foi selecionado o modo flexível do algoritmo adotando o número máximo de 20 conformações por fragmento. Essas conformações foram testadas quanto a sua semelhança tridimensional em relação a objetos de referência (ligante complexado nas estruturas cristalográficas 3INF ou 3L38). A fim de encontrar a orientação que apresenta máxima semelhança e mínimo volume total. Após a obtenção das soluções as cargas atômicas AM1 foram atribuídas às moléculas utilizando as mesmas palavras-chave descritas no item 4.1

4.5.3. Alinhamento por Acoplamento Molecular (Alinhamento III)

As estruturas cristalográficas utilizadas para gerar o alinhamento por acoplamento molecular foram: 3INF e 3L38. A primeira corresponde ao complexo de BACE-1 com um derivado aminoimidazólico (composto 44 – Apêndice A) e segunda ao complexo de BACE-1 com um derivado aminopirimidínico (composto 92 – Apêndice A).

O processo de preparo das estruturas para o acoplamento envolveu as seguintes etapas: (i) identificação dos aminoácidos envolvidos na interação com o ligante; (ii) remoção do ligante do sítio ativo da enzima; (iii) eliminação das moléculas de água, exceto W27 e W470, presentes no sítio catalítico da estrutura 3INF e W473 e W531 presentes na estrutura de 3L38; (iv) adição de átomos de hidrogênio.

Para o acoplamento molecular dos inibidores no sítio ativo da BACE-1 foi utilizado o programa GOLD 4.2.1. (Cambridge Crystallographic Data Centre, Cambridge, UK). A região demarcada para o acoplamento foi definida em um raio de 10Å considerando as coordenadas X = 14,48, Y = -6,84, Z = 15,02 ao redor do átomo N4 do ligante do complexo de 3INF e coordenadas X = 15,44, Y = - 6,94 e Z = 15,58 ao redor do átomo N25 no ligante complexado em 3L38. Compostos estruturalmente semelhantes ao ligante encontrado na estrutura cristalográfica 3L38 (**1-74**, Apêndice A) foram acoplados no sítio dessa estrutura. Segundo esse mesmo critério, o acoplamento de inibidores semelhantes ao ligante encontrado na estrutura cristalográfica 3INF (**75-102**, Apêndice A) foi realizado nessa segunda macromolécula. Em todas as modelagens, duas moléculas de água cristalográficas (W27 e W470 para 3INF; W473 e W531 para 3L38) foram consideradas para os acoplamentos moleculares, nos quais os inibidores foram considerados como flexíveis, enquanto os resíduos da proteína foram considerados como fixos.

Os modelos de interação foram gerados adotando a opção *standard*, disponível no módulo GOLD 4.2.1 Resumidamente: 100 indivíduos (conformações), de cada um dos 102 inibidores do conjunto, evoluíram por 100.000 gerações, sofrendo taxa de mutação e cruzamento de 95%. As cinco melhores posições para molécula, segundo a função Goldscore, foram analisadas visualmente, com auxílio do programa PYMOL v. 1.2 na tentativa de identificar soluções que não apresentam perfil de interação adequado no sítio ativo. A solução de maior pontuação, não descartada na análise visual, foi utilizada para gerar o alinhamento molecular utilizado para o desenvolvimento dos modelos de QSAR 3D por CoMFA.

4.5.4. Realinhamento Rígido de II e III (Alinhamentos IV e V)

As conformações obtidas no Alinhamento II (similaridade morfológica) foram realinhadas segundo sua máxima subestrutura comum, assim moléculas similares ao

ligante presente na estrutura de 3INF (moléculas **1-74**, Apêndice A) foram sobrepostos sobre o mesmo utilizando como referência os átomos marcados em negrito na Figura 18A, enquanto os ligantes similares ao ligante complexado na estrutura 3L38 (moléculas **75-102**, Apêndice A) foram alinhados com base nos átomos marcados em negrito na Figura 18B. O alinhamento obtido foi denominado Alinhamento IV.

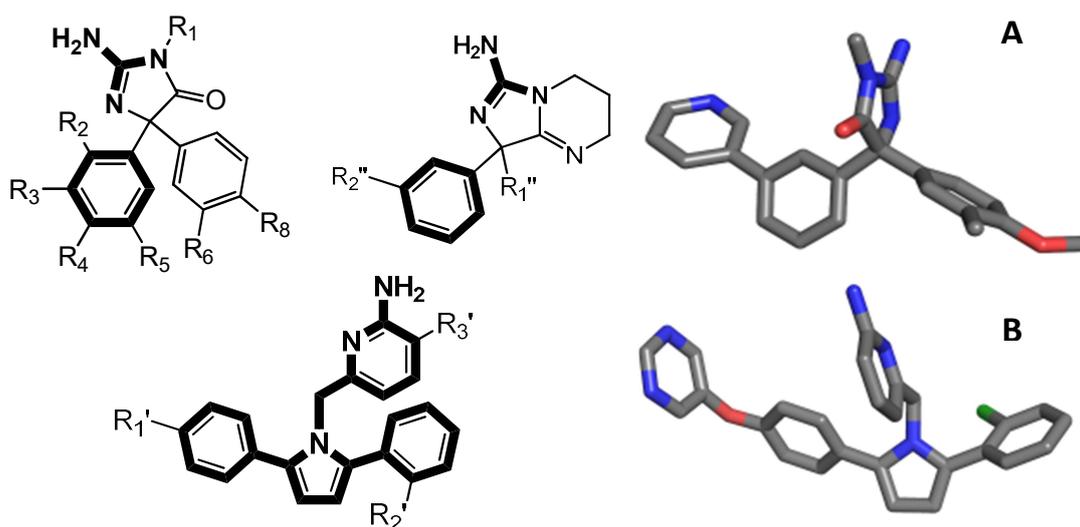


Figura 18. Esqueleto geral dos inibidores de BACE-1. Em negrito estão destacados os fragmentos usados para o realinhamento pela máxima estrutura comum. Na parte superior estão marcados os átomos usados para a sobreposição dos compostos quimicamente semelhantes ao ligante da estrutura do complexo da BACE-1 3INF (A) e na parte inferior, os átomos usados para a sobreposição dos compostos quimicamente semelhantes ao ligante da estrutura 3L38 (B).

Seguindo a mesma estratégia, as conformações oriundas do Alinhamento III (acoplamento molecular), após realinhamento com base na máxima subestrutura comum, deram origem ao Alinhamento V.

4.6. QSAR 3D CoMFA

O módulo CoMFA da plataforma SYBYLX1.1 foi empregado no desenvolvimento dos modelos de QSAR 3D. Campos estéricos e eletrostáticos (potenciais de Lennard–Jones e de Coulomb, respectivamente) foram calculados para cada molécula encontrada nos alinhamentos I, II, III, IV e V, usando como átomo de

prova um carbono de hibridização sp^3 (Csp^3) com carga +1,0. O valor de corte (*cutoff*) de energia padrão de 30 kcal/mol foi mantido no cálculo de ambos os potenciais. Os compostos alinhados foram inseridos numa caixa 3D virtual, regularmente dividida em grades de 2,0 Å de espaçamento. A distância mínima das moléculas até os limites externos da caixa foi de 4 Å.

Num segundo momento, o efeito do espaçamento da grade (2,0 Å, 1,5 Å, 1,0 Å, 0,5 Å) e do dimensionamento dos campos pelo desvio-padrão (0,3/ 0,6/ 0,9/ 1,2) (opção ‘Focus *CoMFA* >> standard deviation’) sobre os parâmetros estatísticos dos modelos foram investigados.

4.7. VALIDAÇÃO DOS MODELOS DE QSAR 2D E QSAR 3D

Os modelos estatísticos gerados nos estudos foram investigados usando o método PLS *Leave one-out* (LOO) que foi aplicado para estabelecer o número ideal de componentes principais (PCs). Os valores de coeficiente de regressão sem validação cruzada (r^2) e coeficientes de regressão por validação cruzada (q^2) forneceram uma estimativa da significância estatística e consistência interna, respectivamente, dos modelos gerados enquanto a validação interna (*Leave-many-out* - LMO) foi utilizada para avaliar a estabilidade do modelo final.

A capacidade preditiva real dos modelos foi analisada por validação externa utilizando um grupo teste de 27 compostos que não foram incluídos no desenvolvimento dos modelos de QSAR. A capacidade preditiva na validação externa é expressa pelo valor de r^2 predito (r^2_{pred}), como calculado abaixo¹¹⁰:

$$r^2_{pred} = 1 - \frac{PRESS}{SS}$$

Onde SS é a soma dos desvios quadrados entre o valor de atividade do conjunto de teste e o valor de atividade média do conjunto de treinamento e PRESS é a soma dos quadrados do erro de predição proveniente da validação cruzada.

5. RESULTADOS E DISCUSSÃO

5.1. ANÁLISE DO CONJUNTO DE DADOS

Modelos de QSAR têm sido usados para descrever a atividade de centenas de compostos biologicamente ativos em diferentes classes de moléculas. Nesses estudos, dados de inibição enzimática são correlacionados às propriedades físico-químicas, estéricas e eletrônicas de ligantes, visando com isso guiar o desenho de moléculas congêneras com características farmacodinâmicas e/ou farmacocinéticas otimizadas. No que se refere ao desenvolvimento de fármacos contra DA, essa técnica já foi empregada para investigar as exigências estruturais e eletrônicas de derivados de Tacrina^{134,135} e Fisiostigmina¹³⁶ que inibem as colinesterases responsáveis pela degradação de acetilcolina na fenda sináptica. De forma semelhante, estudos de QSAR têm sido utilizados para a elucidação das características estruturais e/ou químicas cruciais para a inibição da BACE-1 em derivados de estatina¹³⁷ e de hidroxietilaminas¹³⁸. Os dados obtidos nesses estudos fornecem indicativos de como melhorar a potência e o perfil farmacocinético dessas moléculas, porém a permeabilidade pela barreira hematoencefálica continua sendo um desafio para a utilização clínica de fármacos peptidomiméticos.

Por outro lado, inibidores não peptidomiméticos de BACE-1 apresentam melhor perfil de permeabilidade pela BHE, mas potência insuficiente para serem avaliados em ensaios clínicos. Considerando que ferramentas computacionais e técnicas quimiométricas podem ser guias úteis na otimização da potência desses inibidores, modelos de QSAR 2D e de QSAR 3D foram gerados para uma série de 102 inibidores não peptidomiméticos de BACE-1. Os valores de concentração suficiente para inibir 50% a atividade da enzima BACE-1 (IC_{50}), encontram-se na faixa de 0,01 a 45 μ M, um fator de potência de 4500 vezes, e foram medidos nas mesmas condições experimentais, o que possibilita o sucesso da análise de QSAR em termos de reprodutibilidade. Quimicamente os compostos do conjunto de dados são derivados aminoimidazólicos, aminohidantóinicos e aminopiridínicos (Figura 20), de baixa massa molecular e de caráter não peptídico, o que é uma vantagem, já que pequenas moléculas com tais características podem ter alta biodisponibilidade oral e permeabilidade pela barreira hematoencefálica¹³⁹.

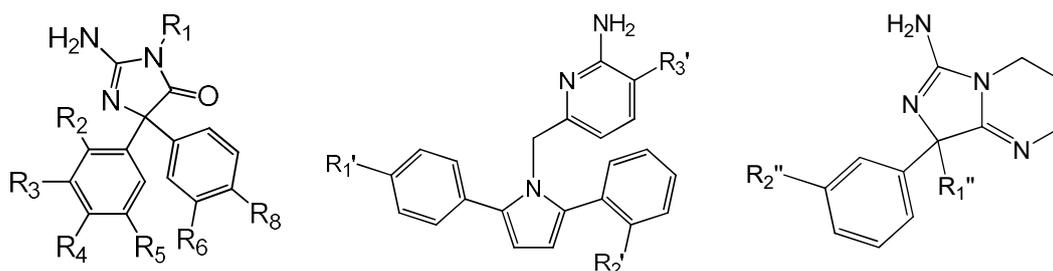


Figura 19: Estrutura geral dos inibidores de BACE-1 utilizados nesse trabalho.

Para a separação do conjunto em grupo treino e grupo teste inicialmente foi feita uma Análise Hierárquica de Agrupamentos (HCA), no programa PIROUETTE 4.0. A ferramenta de HCA auxilia na análise visual da distribuição dos compostos, representados num espaço químico bidimensional, em agrupamentos segundo suas propriedades físico-químicas e estruturais¹⁴⁰. Os resultados, mostrados na forma de um dendograma, evidenciaram a distribuição dos inibidores em 10 classes, considerando 50% de similaridade, o que sugere grande diversidade estrutural da base de dados em estudo. A partir da distribuição dos compostos no dendograma, 27 deles foram selecionados aleatoriamente para compor o conjunto teste, porém assegurando-se de que compostos de todas as classes e distribuídos ao longo das faixas de potência estivessem presentes. Os inibidores remanescentes (total de 75 moléculas) formaram o conjunto treino utilizado para a construção dos modelos de QSAR. Na tabela do apêndice A estão representadas as estruturas e os valores de pIC_{50} das moléculas do grupo treino e teste.

A distribuição das moléculas dos grupos treino e teste, de acordo com os valores de pIC_{50} , mostra uma distribuição normal em ambos os grupos (Figura 21). Este dado garante que além da diversidade estrutural, há uma boa distribuição em termos de valores de pIC_{50} para as moléculas em ambos os grupos treino e teste, o que assegura que os dados estão adequados para proposição e validação dos modelos de QSAR.

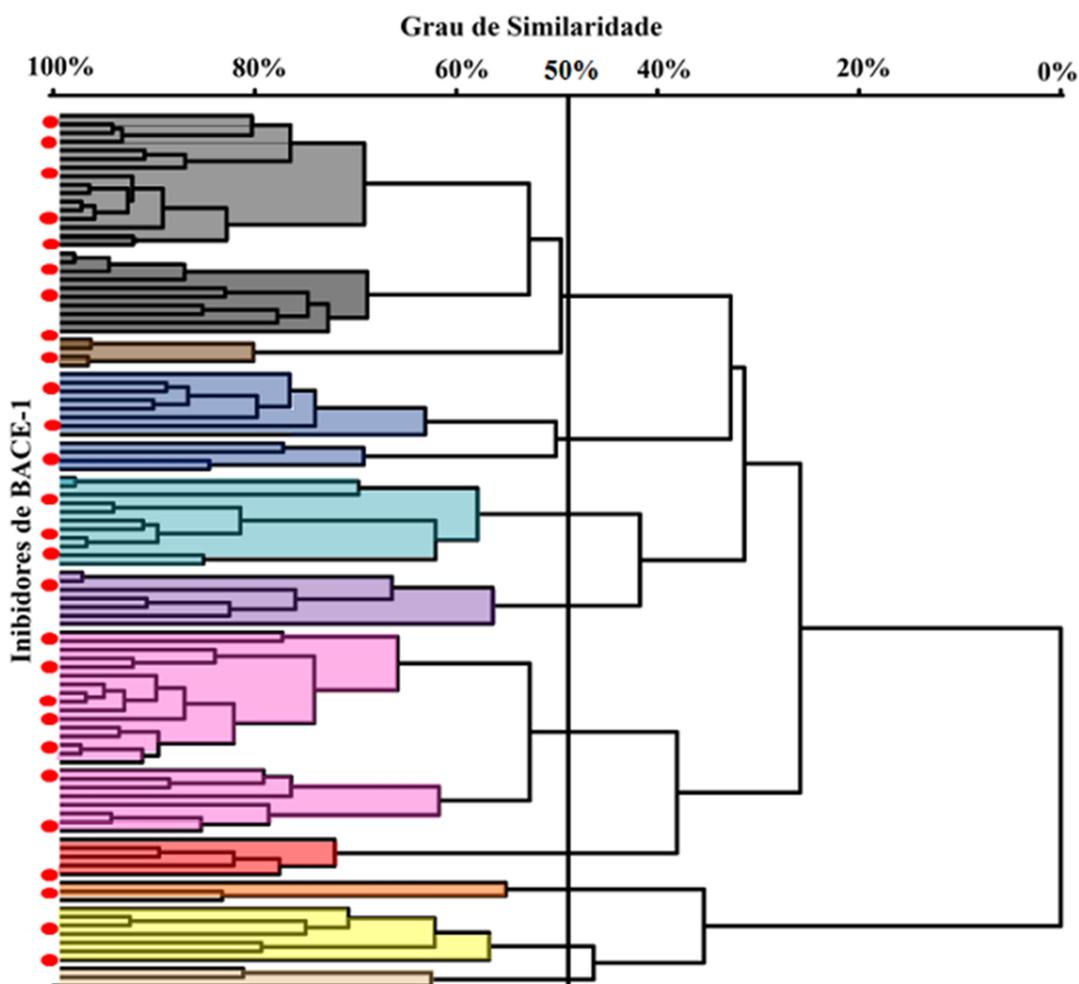


Figura 20: Dendrograma do HCA. Linhas horizontais representam os compostos e as linhas verticais a similaridade entre pares de compostos, entre classes de compostos ou entre um composto e uma classe. Pontos em vermelho do lado esquerdo indicam os compostos seleccionados para o conjunto teste, usados na validação externa dos modelos de QSAR 2D e 3D.

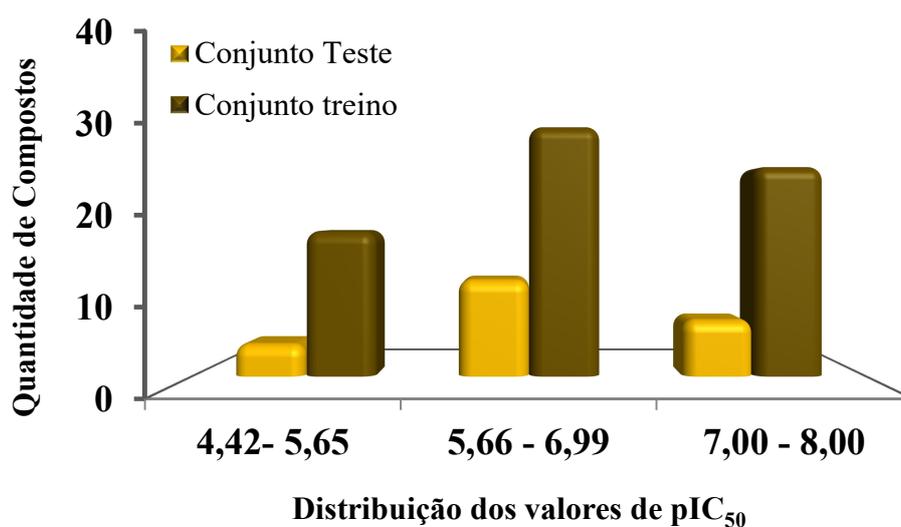


Figura 21: Representação gráfica da distribuição dos valores de pIC₅₀ entre os compostos do conjunto treino e do conjunto teste.

5.2. ESTUDOS DE QSAR 2D

A fundamentação dos estudos QSAR é baseada no princípio de que a variação de atividade pode ser quantitativamente correlacionada com as variações das características estruturais e das propriedades físico-químicas de uma série de moléculas, as quais diferem entre si, geralmente, pela presença de um ou mais grupos substituintes em posições definidas na estrutura química comum à série¹⁴¹. Uma das maneiras mais simples de estabelecer esta correlação é através da representação da estrutura dos ligantes somente por padrões estruturais bidimensionais. Métodos de QSAR 2D permitem a investigação de uma grande variedade de propriedades biológicas para um elevado número de compostos, incluindo os casos em que a informação estrutural 3D para o alvo biológico não está disponível¹⁴².

5.2.1. Holograma QSAR (HQSAR)

Entre os métodos de QSAR 2D baseados em fragmentos moleculares, o Holograma QSAR tem sido empregado para explorar o espaço químico e biológico dos compostos a partir da decomposição especializada de suas estruturas¹⁴³. Os fragmentos gerados são reunidos nos hologramas moleculares que, assim, armazenam todas as informações sobre o número e tipos de fragmentos (lineares, cíclicos, ramificados, etc.) presentes em cada estrutura¹⁴⁴. Um importante aspecto a ser considerado neste método é que nessa representação química, informações 3D como quiralidade e hibridização são codificadas implicitamente, mas nenhuma informação 3D explícita é requerida, sendo necessário somente a estrutura 2D e o valor da propriedade biológica, para a realização das etapas do QSAR¹⁴³.

Os fragmentos moleculares utilizados para construção dos hologramas foram obtidos a partir de uma combinação interativa de no mínimo três parâmetros de distinção (A, B, C, H, Ch e DA), considerando, inicialmente, o tamanho de fragmento padrão (4-7 átomos) e todos os comprimentos de holograma disponíveis (53 a 401). Os modelos de HQSAR são desenvolvidos empregando-se o método de regressões multivariadas por mínimos quadrados parciais (PLS) e técnicas de validação cruzada. Os resultados estatísticos dos modelos obtidos para os 75 compostos do conjunto treinamento utilizando as várias combinações de parâmetros estão na Tabela 1.

Analisando os modelos com base no valor do coeficiente de correlação, r^2 , e no valor q^2 obtido por validação cruzada (LOO), observa-se que a adição do parâmetro H à distinção de fragmento padrão (A/B/C) (2 vs 1) e a exclusão do termo C (4 vs 2), é prejudicial para os modelos. A mesma tendência é observada para a adição do termo DA, que reduz de forma significativa a qualidade estatística dos resultados (modelo 6 vs 7 ou modelo 9 vs 11). A exclusão do parâmetro B é capaz de fornecer modelos melhores (7 vs 8 ou 2 vs 12) e a sua eliminação da combinação (A/B/C/Ch) (3) levou ao modelo com o mais alto valor estatístico (**11** - $q^2 = 0,84 / r^2 = 0,91$ e 6 PCs).

Tabela 1. Influência da distinção dos fragmentos nos parâmetros estatísticos dos modelos de HQSAR usando o tamanho de fragmento padrão (4-7).

Modelo	Parâmetros de Fragmentos	Parâmetros Estatísticos do HQSAR			
		q^2	r^2	HL	PC
1	A/B/C	0,83	0,90	97	6
2	A/B/C/H	0,80	0,89	353	6
3	A/B/C/Ch	0,83	0,90	97	6
4	A/B/H	0,62	0,94	61	5
5	A/B/C/DA	0,82	0,88	353	6
6	A/B/C/H/Ch	0,81	0,88	97	6
7	A/B/C/H/Ch/DA	0,69	0,84	199	6
8	A/C/H/Ch/DA	0,75	0,86	257	6
9	A/C/Ch/DA	0,78	0,89	257	6
10	A/C/H/Ch	0,83	0,89	151	6
11	A/C/Ch	0,84	0,91	53	6
12	A/C/H	0,82	0,89	53	6

* q^2 , coeficiente de correlação por validação cruzada; r^2 , coeficiente de correlação sem validação cruzada; HL, tamanho do holograma; PC, número ótimo de componentes.

Num segundo momento, foi investigada a influência da variação no tamanho dos fragmentos estruturais sobre os parâmetros estatísticos do melhor modelo obtido até esse ponto. Nessa etapa, novos modelos foram gerados variando-se a quantidade mínima e máxima dos átomos a ser considerada na composição dos fragmentos, preservando-se o melhor tamanho de holograma (53). Conforme pode ser observado na Tabela 2, essa estratégia não produziu modelos com parâmetros estatísticos superiores àqueles encontrados para o modelo 11.

Tabela 2. Influência do tamanho do fragmento nos parâmetros estatísticos do melhor modelo de HQSAR (A/C/Ch).

Modelo	Tamanho do fragmento	Prâmetros Estatísticos do HQSAR			
		q^2	r^2	HL	PC
13	2-5	0,80	0,86	53	5
14	3-6	0,81	0,87	53	5
15	5-8	0,80	0,89	53	6
16	6-9	0,80	0,88	53	6
17	7-10	0,79	0,88	53	6

Para avaliar melhor a consistência interna do modelo 11, protocolos de validação cruzada mais rigorosos foram adotados e através da separação sistemática do conjunto treino em grupos aleatórios de 5, 10, 15 e 25 moléculas novos modelos foram gerados. Os resultados não evidenciaram alterações sobre os valores estatísticos do modelo 11, indicando assim sua alta estabilidade. Embora o elevado valor de q^2 seja uma condição necessária para que um modelo de QSAR tenha boa capacidade preditiva, estudos mostram que este resultado estatístico não qualifica automaticamente o modelo como adequado para a predição da atividade biológica de novos compostos¹⁴⁴⁻¹⁴⁵. Por esta razão é aconselhável a utilização de protocolos de validação externa como uma alternativa para investigar a capacidade preditiva dos modelos gerados¹⁴⁶. Considerando que a estrutura química das moléculas, codificada nos hologramas está relacionada diretamente aos respectivos valores de potência inibitória, pode-se dizer que os modelos de HQSAR reúnem informações suficientes para prever a atividade biológica em moléculas estruturalmente relacionadas.

Dessa maneira, avaliou-se a capacidade preditiva do modelo 11 para as 27 moléculas que foram completamente excluídas da calibração do modelo (conjunto teste). Os valores residuais, estabelecidos pelas diferenças entre os valores de pIC_{50} reais e os valores preditos pelo modelo, não desviam mais do que 0,51 unidades logarítmicas para 90% dos compostos (Tabela 3 e Figura 22) sugerindo que o modelo mostra-se útil para prever a atividade biológica de moléculas congênicas. Essa afirmativa é embasada no valor do coeficiente de correlação de preditividade encontrado para o melhor modelo de HQSAR ($r^2_{pred} = 0,70$).

Tabela 3: Valores preditos e residuais de pIC_{50} para os 27 compostos do grupo teste de acordo com o melhor modelo de HQSAR (A/C/Ch).

Comp.	Exp. ^a	Pred. ^b	Resid ^c	Comp.	Exp. ^a	Pred. ^b	Resid ^c	Comp.	Exp. ^a	Pred. ^b	Resid ^c
02	5,35	4,66	0,69	40	6,68	6,33	0,35	67	7,27	7,00	0,27
07	5,31	5,09	0,22	42	6,54	6,55	-0,01	71	7,47	6,85	0,62
10	5,57	6,68	-1,11	48	7,28	7,30	-0,02	77	5,91	5,94	-0,03
15	7,05	6,72	0,33	50	7,52	7,15	0,37	79	5,97	6,26	-0,29
23	5,98	5,87	0,11	53	7,11	7,44	-0,33	85	5,73	5,75	-0,02
25	5,69	5,86	-0,17	55	7,16	6,82	0,34	90	6,70	6,39	0,31
32	4,75	5,46	-0,71	57	7,11	6,60	0,51	91	6,41	6,09	0,32
35	5,48	5,51	-0,03	58	7,25	7,05	0,20	99	6,45	6,77	-0,32
37	6,15	5,82	0,33	59	7,42	7,70	-0,28	102	7,42	7,15	0,27

^a Valor Experimental / ^b Valor predito / ^c Diferença entre valor experimental e predito.

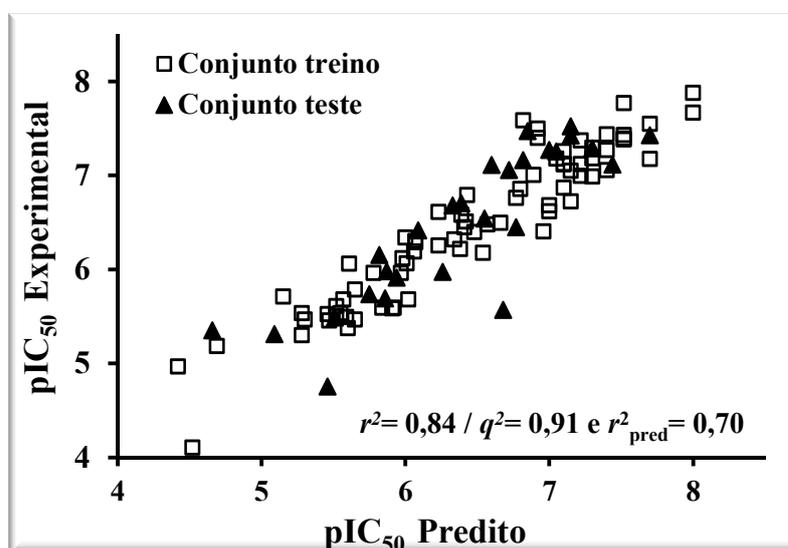


Figura 22: Valores preditos versus experimentais de pIC_{50} para o conjunto treino (quadrados) e teste (triângulos) de acordo com o melhor modelo de HQSAR (distinção de fragmento: A/C/Ch; e tamanho de fragmento: 4–7).

Apesar da qualidade estatística ser imprescindível para um bom modelo de QSAR, é sua contribuição para elucidar as relações entre a estrutura química e a atividade biológica que garantem sua aplicabilidade nos projetos de desenvolvimento de fármacos. Nessa perspectiva, os mapas de contribuição do HQSAR podem ser de grande valia uma vez, que fornecem dados sobre a magnitude da contribuição de átomos/grupos e fragmentos moleculares para a propriedade sob investigação. Para isto, o módulo do HQSAR utiliza um código de cores, segundo o qual a região espectral do vermelho, laranja e tons próximos a estas cores são usadas para indicar as contribuições desfavoráveis enquanto que tons verde e amarelo refletem as contribuições favoráveis. Átomos e fragmentos nas cores branco/cinza

indicam contribuição neutra ou intermediária, ou átomos presentes em todas as moléculas da série. É válido destacar que fragmentos coloridos em branco podem pertencer ao farmacóforo da molécula, porém devido a soma das contribuições atômicas positivas e negativas decorrentes de fragmentos que se sobrepõe, o fragmento é considerado neutro.

Mapas de contribuição para compostos de alta potência evidenciam a importância positiva de átomos de nitrogênio presentes no núcleo central aminoimidazol/aminohidantoinico e no grupo amino ligado a estes núcleos (Figura 23). Esta informação é compatível com dados observados na estrutura cristalográfica de 3INF, na qual o inibidor complexado (cuja estrutura é quimicamente semelhante à do composto 46) tem esses átomos posicionados para a realização de interações com resíduos aspárticos na porção catalítica da enzima. Adicionalmente, o complexo cristalográfico mostra que átomo de nitrogênio presente no anel *m*-piridínico do ligante interage com um resíduo de Serina através de uma molécula de água, aumentando a afinidade do composto⁸⁰⁻⁸³. Esta informação é destacada nos mapas de contribuição do HQSAR, que mostra a contribuição positiva do átomo de nitrogênio do respectivo anel piridínico na estrutura de alguns compostos, como nos inibidores **20** e **46**.

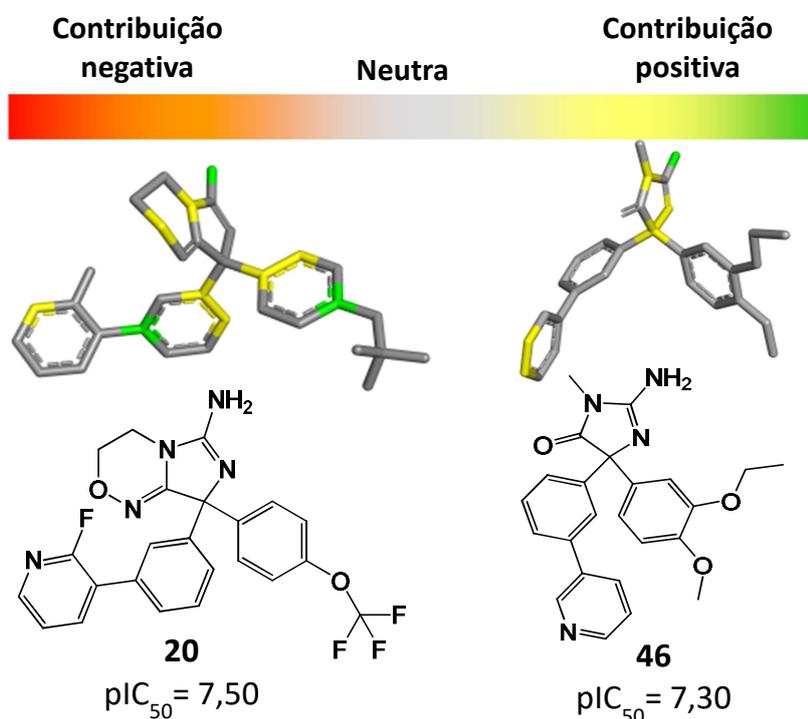


Figura 23: Mapa de contribuição para inibidores de alta potência de BACE-1 (**20** $pIC_{50} = 7.50$ e **46** $pIC_{50} = 7.30$) de acordo com o melhor modelo de HQSAR.

Uma das limitações do Holograma QSAR é que o algoritmo não é capaz de fornecer dados explícitos que permitam explicar as razões químicas para o esquema de cores exibidos nos mapas de contribuição. O mapa de contribuição do composto **20**, por exemplo, destaca positivamente a contribuição do átomo de oxigênio no anel conjugado ao grupo aminoimidazol. De fato, a introdução desse átomo nesta posição proporciona melhora significativa na potência se compararmos com o composto que não apresenta o átomo de oxigênio substituído (**13** – $pIC_{50} = 6,43$, Apêndice A). Contudo, os espectros obtidos são a soma das contribuições atômicas individuais das moléculas, que por sua vez, resultam de uma relação numérica estabelecida entre o coeficiente de PLS do modelo e os átomos de cada fragmento, ou seja, não se pode afirmar que uma determinada contribuição é resultante de componentes estéreo e/ou eletrônicos. Dessa maneira, visando complementar a informação químico-medicinal proveniente dos modelos de HQSAR, modelos de QSAR empregando descritores topológicos foram obtidos.

5.2.2. QSAR 2D Baseado em Descritores Topológicos

Descritores topológicos são construções matemáticas que não têm significado químico único, descrevendo informações sobre o tamanho, forma, ramificações, presença de heteroátomos, ligações múltiplas moleculares, etc¹⁴⁷. Tais descritores tem sido amplamente utilizados na predição de diferentes propriedades como atividade biológica, propriedades físico-químicas, farmacocinéticas, de toxicidade, dentre outras, permitindo, através de estudos quantitativos das relações entre a estrutura química e a atividade biológica, o desenvolvimento de moléculas com propriedades farmacocinéticas e farmacodinâmicas otimizadas¹⁴⁸.

Seguindo as etapas descritas na seção experimental, um total de 2489 descritores (ex.: topológicos, átomos centrados em fragmentos, autocorrelação 2D, etc.) foram calculados através do programa DRAGON 5.5, entretanto somente 425 deles permaneceram após a etapa de pré-seleção, realizada para a eliminação dos descritores que não contém informações relevantes para descrever as propriedades moleculares que influenciam potência inibitória dos compostos, bem como para a eliminação dos descritores que contém essencialmente a mesma informação, minimizando-se o ‘ruído’ causado por dados redundantes⁹². Objetivando o desenvolvimento de modelos de QSAR

preliminares, esses descritores foram empregados na construção de modelos por regressão linear múltipla (MLR), com até quatro descritores através da abordagem de algoritmo genético. Apesar do melhor modelo obtido nesta etapa apresentar bons resultados estatísticos ($r^2 = 0,85$, $q^2 = 0,84$) a capacidade preditiva foi considerada insatisfatória ($r^2_{pred} = 0,62$), para guiar o planejamento de derivados mais potentes. O baixo poder de preditividade dos modelos pode ser decorrente do ajuste forçado dos dados decorrentes da colinearidade dos descritores nos modelos, os quais não conseguem descrever o espaço químico de forma apropriada¹⁰².

Tabela 4: Descritores selecionados para a modelagem do QSAR 2D baseados em descritores topológicos.

Tipo de Descritor	Símbolo	Descrição
Constitutional	nR09	Número de anéis de 9 membros
Topológico	DELS	Variação molecular eletrotológica
Índice de Conectividade	X3sol	Índice de conectividade e solvatação chi-3
Índice de Informação	IC1	Índice de informação de conteúdo (simetria da vizinhança de ordem 1)
Descritores de autovalor de cargas	BEHm7	Maior autovalor n. 7 da matriz de carga / ajustada pelas massas atômicas
	BEHm8	Maior autovalor n. 8 da matriz de carga / ajustada pelas massas atômicas
	BEHv8	Maior autovalor n. 8 da matriz de carga / ajustada pelo volume atômico de Van der Waals
Índice de carga topológica	GGI2	Índice de carga topológica de ordem 2
	GGI5	Índice de carga topológica de ordem 5
	GGI6	Índice de carga topológica de ordem 6
	JGI5	Média do índice de carga topológica de ordem 5
Quantidade de grupos funcionais	nArX	Número X de anéis aromáticos
Átomo centrado em fragmentos	C-027	R--CH--X
	H-053	H ligado a C sp ³ com 2 halogênios ligados ao C adjacente
	N-075	R--N--R ^e / R--N—X (^e é um grupo piridina)
Impressão digital binária 2D	B02[C-O]	presença/ausência de C-O a uma distancia topológica 2
	B02[N-F]	presença/ausência de N-F a uma distancia topológica 2
	B03[N-N]	presença/ausência de N-N a uma distancia topológica 3
	B08[C-O]	presença/ausência de C-O a uma distancia topológica 8
	B08[N-N]	presença/ausência de N-N a uma distancia topológica 8
	B09[N-N]	presença/ausência de N-N a uma distancia topológica 9
	B10[C-C]	presença/ausência de C-C a uma distancia topológica 10

Como alternativa, foi utilizado o método de PCA para verificar a capacidade dos descritores em descrever as propriedades físico-químicas e estruturais da série. A técnica de PCA agrupa os descritores altamente correlacionados, substituindo-os por um número menor de novos descritores, não correlacionados, que são denominados Componentes Principais (PCs). Para esta etapa os 22 descritores presentes nos 10 melhores modelos de MLR (Tabela 4) foram reunidos e autoescalonados e análise exploratória realizada no programa PIROUETTE 4.0

Os resultados da análise de PCA (Figura 24) mostrou que 65,9% da variância total está contida nas três primeiras componentes principais (PCs). Contudo, a PC1 individualmente é capaz de separar os compostos do grupo treino de acordo com a sua potência: inibidores de baixa potência ($pIC_{50} < 5,65$) apresentam valores negativos para a PC1; enquanto que os de alta potência ($pIC_{50} > 6,67$) exibem valores positivos para a PC1. Essa análise mostra que os descritores selecionados são capazes de discriminar os inibidores do conjunto em função dos seus valores de propriedade biológica.

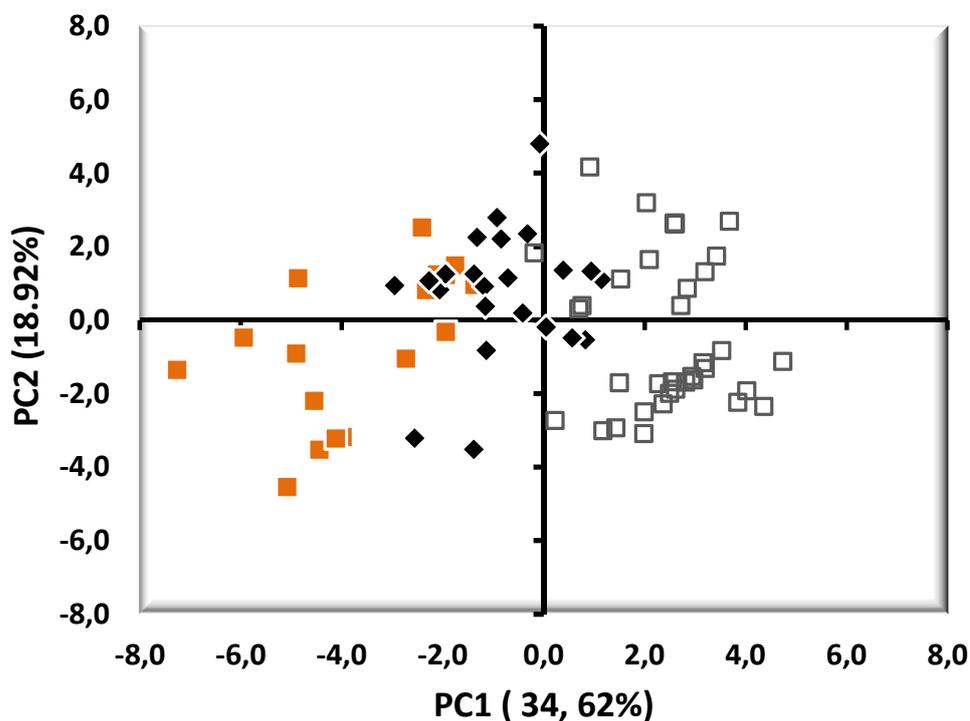


Figura 24: Distribuição dos compostos na análise de PCA: quadrados laranja compostos de baixa potência ($pIC_{50} < 5,65$), losangos, representam inibidores com potência intermediária ($5,66 < pIC_{50} < 6,66$); quadrados brancos, moléculas com alta potencia ($pIC_{50} > 6,67$).

Adicionalmente, o gráfico de peso (contribuição) (Figura 25) mostra que propriedades eletrônicas (descritas por GGI5 e GGI6) e estruturais (descritas por B08[N-N] e B09[N-N]) tem papel de destaque para PC1 e, portanto, são úteis para separar as moléculas de acordo com sua potência. Isso mostra que os descritores selecionados na etapa anterior devem ser úteis para explicar a atividade biológica. Dessa forma, a fim de estabelecer a correlação dos descritores com a propriedade dos compostos da série, modelos de QSAR foram gerados utilizando PLS, como disponível no PIROUETTE 4.0.

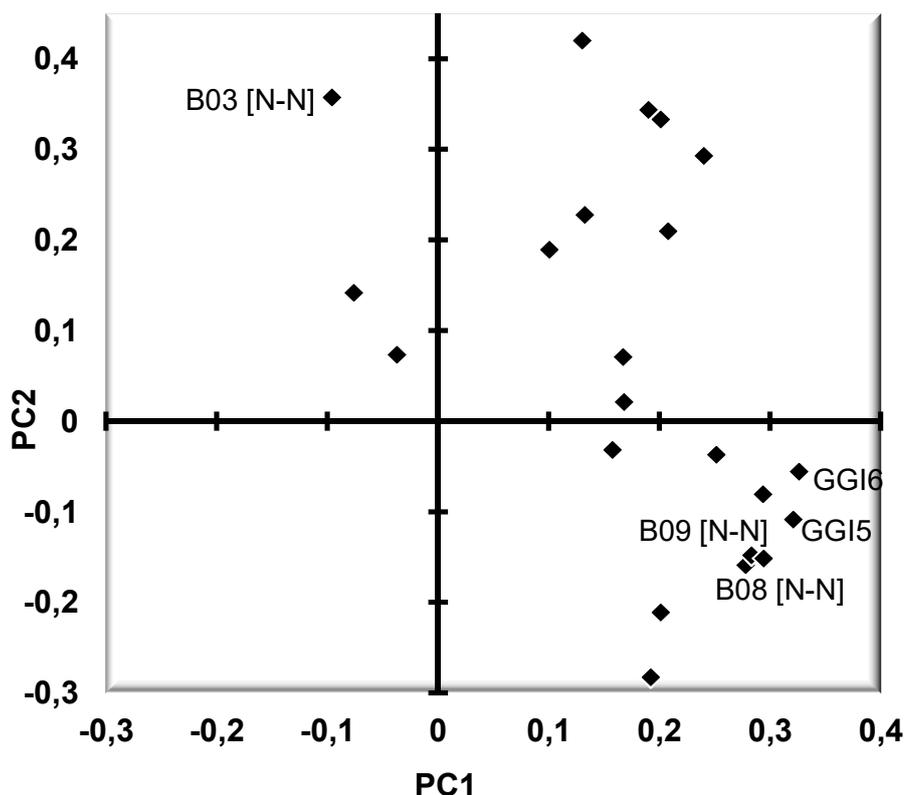


Figura 25: Gráfico de pesos dos descritores para PC1 e PC2.

O melhor modelo de QSAR obtido apresentou altos valores estatísticos ($r^2= 0.87$ e $q^2= 0.85$), evidenciando elevada correlação entre os descritores e a propriedade biológica, assim como considerável consistência interna. Técnicas de validação cruzada mais exigentes (LMO) não produziram mudanças significativas nos parâmetros estatísticos, confirmando assim significância e estabilidade do modelo. A fim de investigar o poder preditivo foram estimados os valores de pIC_{50} dos 27 compostos do

conjunto teste. O resultado desta análise (Tabela 5 e Figura 26) mostra que a capacidade preditiva do modelo ($r^2_{\text{pred}} = 0,84$) é maior do que aquela obtida nos modelos de HQSAR, sugerindo que a utilização de informações estruturais e fisico-químicas, codificadas em descritores topológicos é crucial para descrever as relações entre a estrutura química e a atividade biológica da série de inibidores de BACE-1 em estudo.

Tabela 5. Valores preditos e residuais de pIC_{50} para os 27 compostos do grupo teste de acordo com o melhor modelo de QSAR 2D clássico.

Comp.	Exp. ^a	Pred. ^b	Resid ^c	Comp.	Exp. ^a	Pred. ^b	Resid ^c	Comp.	Exp. ^a	Pred. ^b	Resid ^c
02	5,35	5,34	0,01	40	6,68	6,55	0,13	67	7,27	6,88	0,39
07	5,31	5,29	0,02	42	6,54	6,69	-0,15	71	7,47	6,99	0,48
10	5,57	6,12	-0,55	48	7,28	7,34	-0,06	77	5,91	5,97	-0,06
15	7,05	7,06	-0,01	50	7,52	7,30	0,22	79	5,97	6,09	-0,12
23	5,98	6,39	-0,41	53	7,11	7,23	-0,12	85	5,73	5,89	-0,16
25	5,69	5,63	0,06	55	7,16	7,26	-0,10	90	6,70	6,31	0,39
32	4,75	4,93	-0,18	57	7,11	6,92	0,19	91	6,41	6,23	0,18
35	5,48	5,69	-0,21	58	7,25	7,03	0,22	99	6,45	6,61	-0,16
37	6,15	5,64	0,51	59	7,42	7,40	0,02	102	7,42	7,00	0,42

^a Valor Experimental / ^b Valor predito / ^c Diferença entre valor experimental e predito.

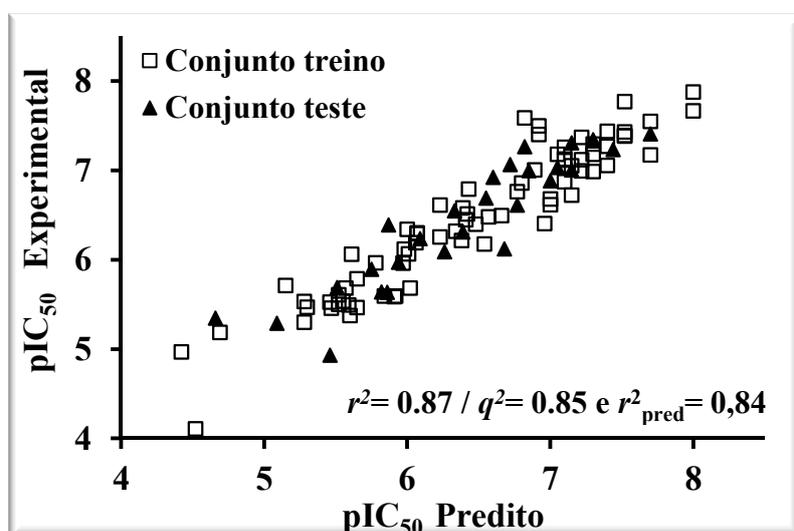


Figura 26: Valores preditos versus experimentais de pIC_{50} para o conjunto treino (quadrado) e teste (triângulos) de acordo com o melhor modelo de QSAR 2D baseado em descritores topológicos.

A análise da contribuição dos descritores para o vetor de regressão (Figura 28) permite verificar quantitativamente a importância de cada um deles para o modelo final de QSAR. Os descritores mais relevantes corroboram com aqueles descritos no gráfico de

contribuição das PCs (Figura 27) e assim, a interpretação das informações codificadas neles devem fornecer elementos importantes, em termos de características químicas e estruturais que influenciam na potência das moléculas.

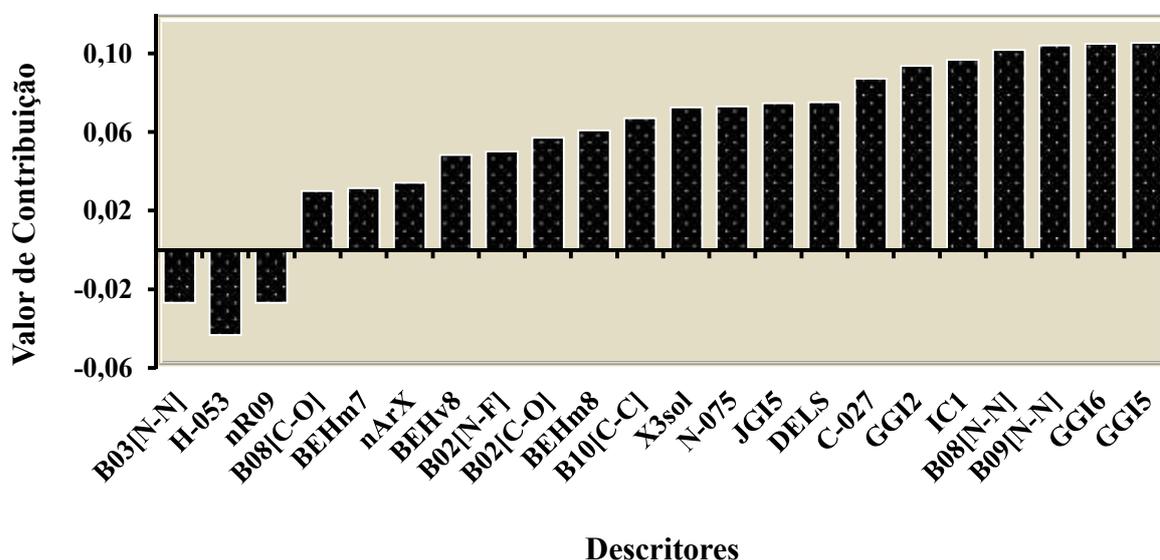


Figura 27: Contribuição dos descritores para o vetor de regressão obtido para o melhor modelo de QSAR 2D.

Descritores de índice de carga topológica são caracterizados pela sua capacidade de avaliar a transferência de cargas entre pares de átomos, levando em conta a carga global da molécula¹⁴⁹. Dentre os descritores selecionados, GGI5 e GGI6 descrevem a distribuição de cargas entre átomos separados a uma distância topológica igual a 5 e 6, respectivamente. A análise da distribuição dos valores de GGI6 em função da potência mostra uma relação direta entre esses valores (Figura 28). De fato, a análise das moléculas mais potentes mostra um número maior de átomos capazes de participar da transferência de carga, se comparado ao número de moléculas de baixa potência, as quais apresentam estruturas mais simples. A correlação não é tão evidente para GGI5, porém deve-se salientar que, na análise, os valores desse descritor são multiplicados pela sua influência (*loading*) para as PCs, portanto a falta de correlação observada não invalida ou contradiz a interpretação do modelo de QSAR.

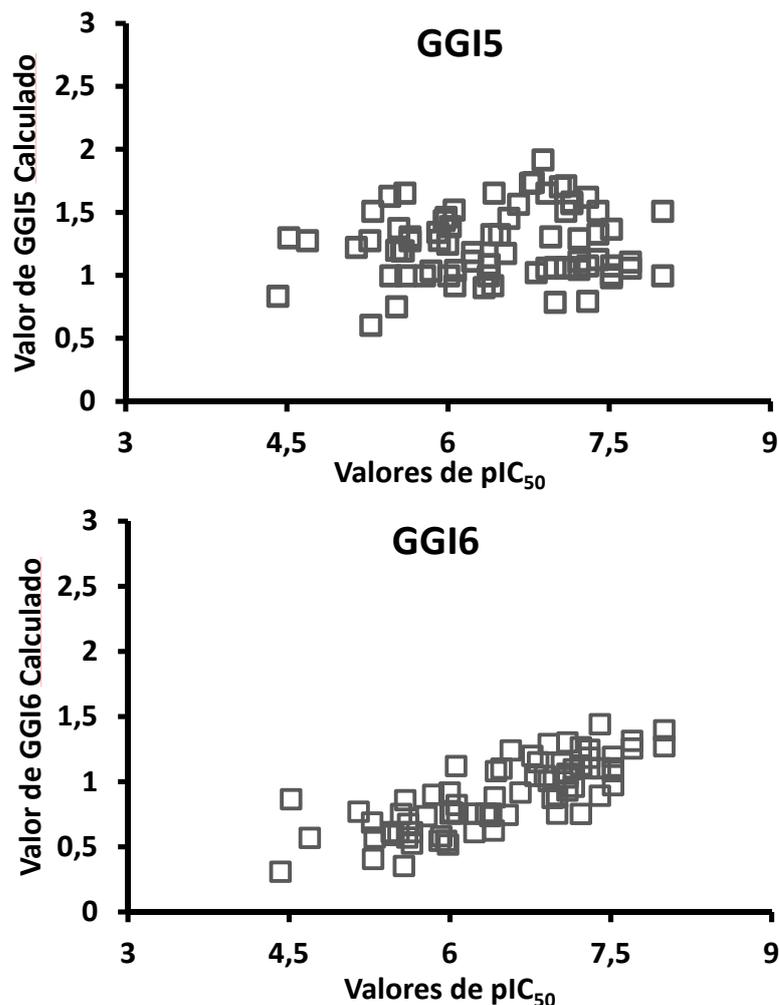


Figura 28: Distribuição dos descritores GGI5 e GGI6 versus valores de atividade biológica para os compostos do conjunto treino.

O gráfico do vetor de regressão também destaca a contribuição positiva para os descritores B08 [N-N] e B09 [N-N] (Figura 29 -A), que representam a presença de fragmentos estruturais moleculares de pares de átomos de nitrogênio localizados a uma distância topológica 8 ou 9, respectivamente. Analisando o conjunto de dados, foi possível observar que mais de 90% dos compostos com ambos os descritores na sua estrutura exibem alta potência ($pIC_{50} > 6,80$). Este resultado está de acordo com os mapas de contribuição dos modelos de HQSAR, que também destacam os mesmos átomos como tendo contribuições positivas para a potência (Figura 29 – B). Por outro lado, B03[N-N], que descreve o número de pares de átomos de nitrogênio separados por 3 ligações parece ter efeitos prejudiciais sobre a atividade inibitória. Considerando que uma mudança na distância entre os átomos de nitrogênio inverte a contribuição, é razoável assumir que

átomos de nitrogênio posicionados incorretamente não são capazes de interagir de maneira eficiente com os resíduos polares encontrados no sítio ativo da BACE-1, reduzindo assim a potência do composto.

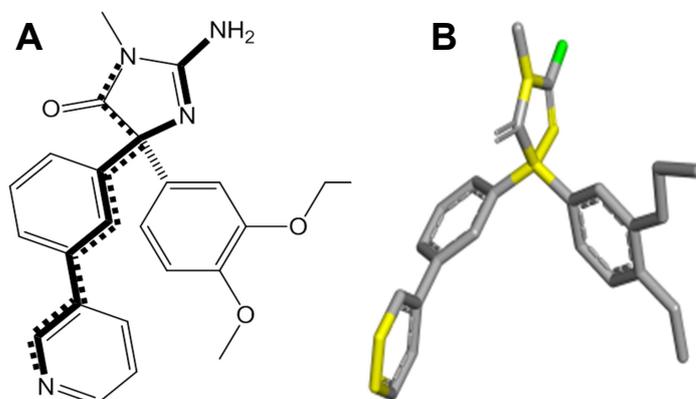


Figura 29: Descritores B08 [N-N] e B09 [N-N] representados na estrutura de um composto de alta potência (**46**, $pIC_{50} = 7,3$) (A); e observação deste mesmo descritor no mapa de contribuição do HQSAR (B).

Embora a utilização de descritores 2D na geração de modelos de QSAR apresente vantagens, já que eles podem ser calculados para todas as moléculas existentes e a obtenção dos seus valores seja relativamente rápida, uma das principais críticas na sua utilização está na dificuldade em se atribuir um significado físico-químico claro para muitos deles¹⁵⁰. De fato, diversos descritores topológicos presentes no modelo final são de difícil interpretação e, portanto, tem sua aplicabilidade limitada para o planejamento de inibidores.

Dessa maneira, considerando que a resposta biológica é decorrente de um evento tridimensional (encaixe do ligante no sítio ativo da BACE-1) e visando aprofundar o estudo sobre parâmetros eletrônicos e estéreos que influenciam na atividade das moléculas em estudo, decidimos realizar uma análise comparativa dos campos moleculares como alternativa para construção de novos modelos de QSAR.

5.3. ESTUDOS DE QSAR 3D

5.3.1. QSAR 3D CoMFA

5.3.1.1. O Alinhamento Estrutural dos Compostos

A geração de modelos de QSAR a partir do método de CoMFA se baseia no postulado de que as propriedades que condicionam a afinidade de ligação estão relacionadas com aquelas representadas pelos campos estereoquímicos e eletrostáticos¹¹⁹. Conforme descrito na metodologia, um dos requisitos primários do método de COMFA é o alinhamento relativo de todos os compostos, para que eles tenham conformação e orientação comparável no sítio ativo.

Entre os 102 inibidores selecionados, somente 6 tem as suas respectivas estruturas cristalográficas disponíveis na base de dados do PDB (3IGB, 3IND, 3INE, 3INF, 3L3A e 3L38). Por esta razão, é necessário encontrar a provável conformação bioativa dos demais compostos, o que foi feito por quatro estratégias diferentes:

O primeiro método de alinhamento usado (Alinhamento I - Figura 31) é baseado na premissa de que as moléculas devem se ligar de modo semelhante no alvo macromolecular e portanto, fragmentos idênticos devem ser encontrados na mesma região do espaço. Por isso, grupos funcionais e/ou átomos comuns a todas as moléculas foram identificados e utilizados para sobrepor as estruturas minimizadas dos inibidores de BACE-1 utilizados nesse trabalho. Embora a orientação gerada não forneça exatamente a ‘conformação bioativa’, se comparada aos ligantes do complexo da BACE-1 (Figura 31, A), as estruturas minimizadas podem representar pontos acessíveis do espaço conformacional dos inibidores no sistema biológico, fornecendo dados de relação estrutura-atividade úteis¹²¹.

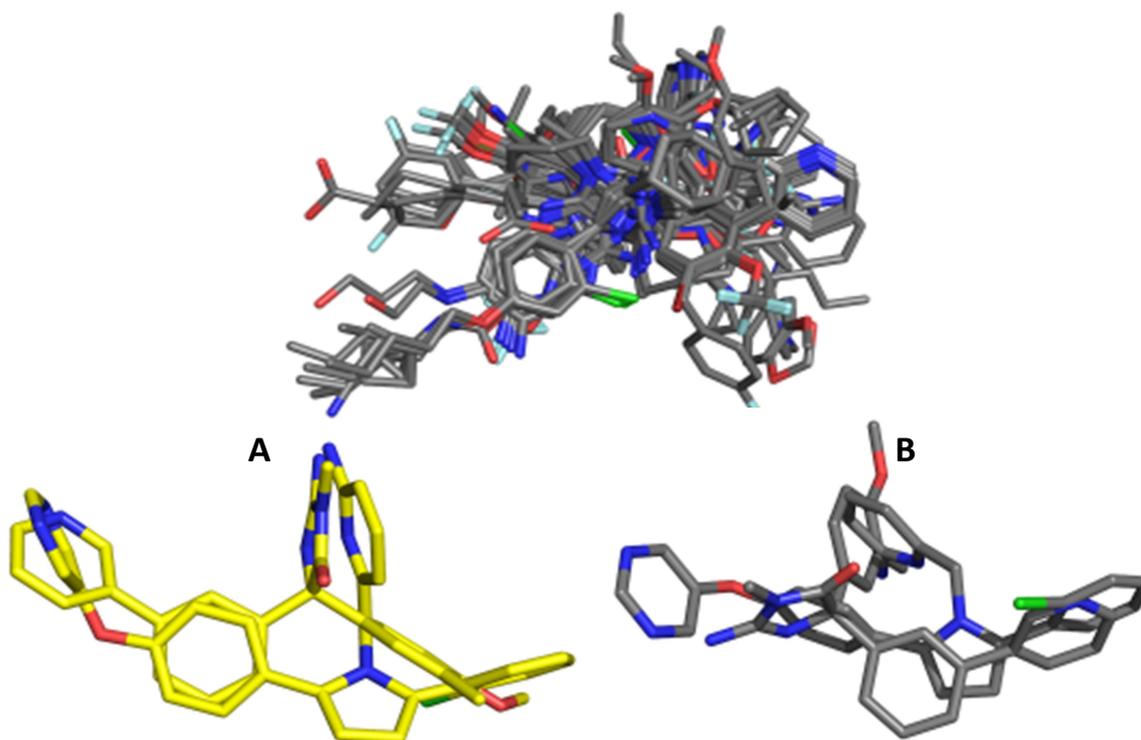


Figura 30: Resultado da sobreposição baseada na máxima estrutura comum Alinhamento I (acima); imagem comparativa que evidencia claramente as diferenças nas ‘conformações bioativas’ através da sobreposição dos ligantes do complexo 3L38 e 3INF (A) e das estruturas 3D desses ligantes minimizadas pelo método AM1 (B) (abaixo).

A segunda estratégia de alinhamento adotada (Alinhamento II) parte da mesma hipótese anterior (modo de interação semelhante no sítio ativo), mas agora estima a semelhança dos inibidores com base na similaridade morfológica calculada com o auxílio do programa SURFLEX-SIM. Este método baseia-se na comparação das moléculas em relação as suas superfícies de Van der Waals, átomos doadores e aceptores de ligação de hidrogênios e átomos carregados negativamente a fim de obter conformações próximas das bioativas, visando a máxima sobreposição entre elas, sem contudo ocupar um volume muito maior que o da molécula referência, que no nosso caso foram os ligantes presentes nos complexos cristalográficos da BACE-1 (PDB: 3L38 e 3INF). A partir desses critérios, 74 inibidores foram alinhados sobre as coordenadas cristalográficas do inibidor de BACE-1 encontrado no complexo 3INF e 28 inibidores sobre o ligante do complexo 3L38, as conformações obtidas foram reunidas para compor o alinhamento II (Figura 32).

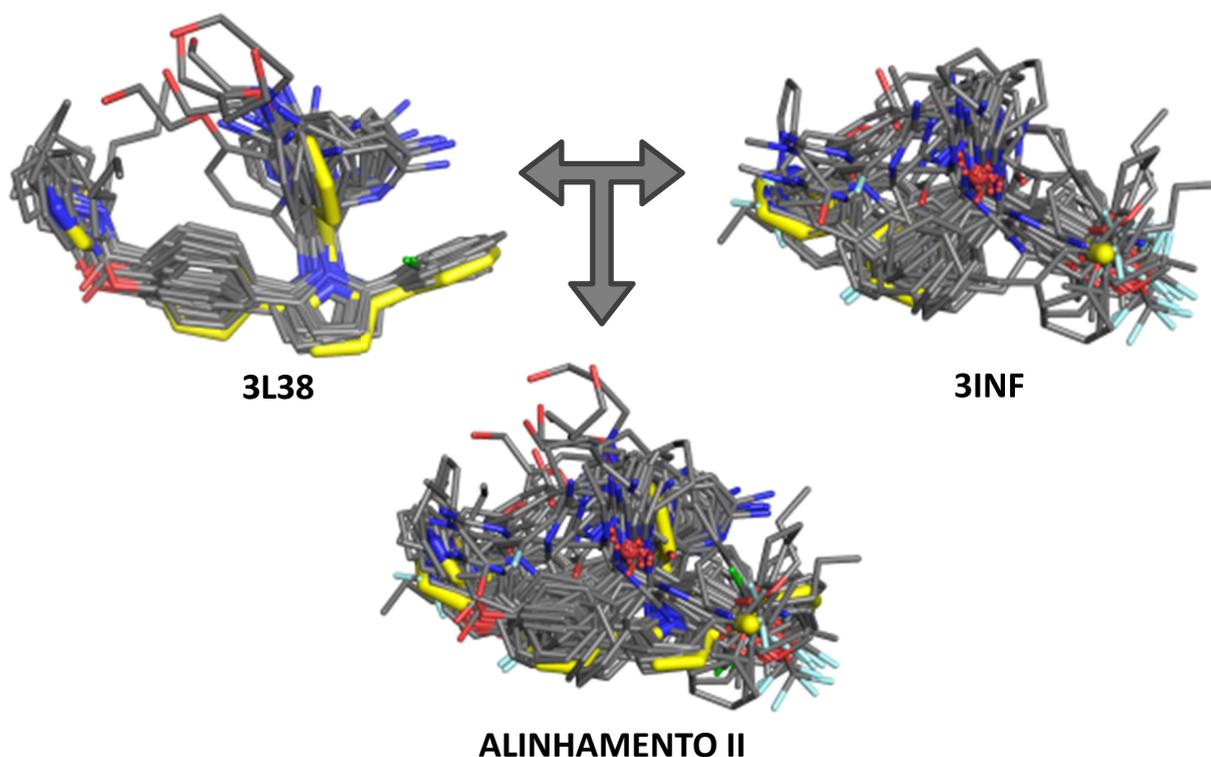


Figura 31: Alinhamento dos inibidores de BACE-1 obtido a partir do método baseado na similaridade de moléculas (Alinhamento II). Em amarelo estão indicadas as estruturas dos ligantes extraídos dos complexos da BACE-1 PDB 3INF e 3L38.

A terceira estratégia de alinhamento baseou-se no acoplamento das estruturas dos ligantes no sítio da BACE-1 com o auxílio programa GOLD 4.2.1. Para a obtenção do Alinhamento III, foram tomadas as cinco melhores soluções fornecidas pelo algoritmo, de acordo com o protocolo descrito na metodologia. A qualidade das soluções foi avaliada de acordo com o posicionamento de cada conformação em relação aos resíduos no sítio ativo, baseando-se no modo de interação observado para o ligante cristalográfico. As soluções que melhor se subrepunham ao ligante, obedecendo, por exemplo, ao posicionamento mais adequado para a realização das interações com os resíduos aspárticos na região catalítica e no bolsão hidrofílico S3 (Figura 33) foram escolhidas para compor o alinhamento III.

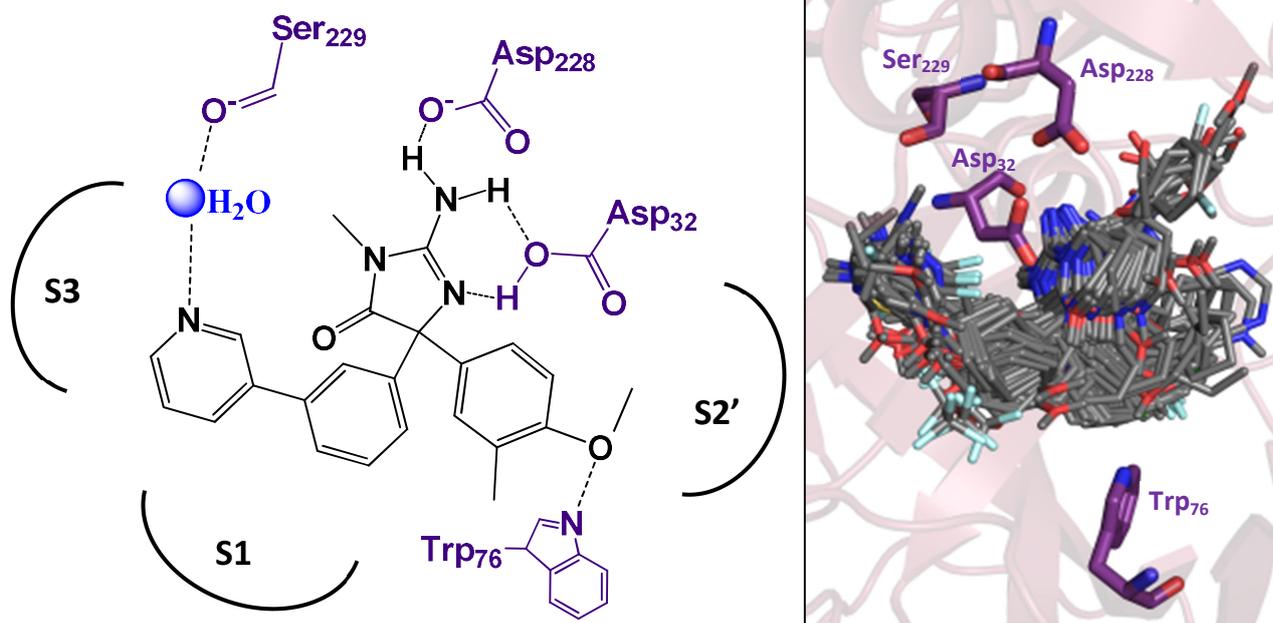


Figura 32: Representação esquemática das principais interações entre os inibidores da série em estudo e a BACE-1 (à esquerda) e resultados da sobreposição dos compostos no Alinhamento III (à direita).

Finalmente, a quarta estratégia de alinhamento objetivou unir as vantagens das informações sobre as conformações dos ligantes cristalográficos com a sobreposição espacial pelo método baseado na máxima estrutura comum. Tomadas juntas, estas estratégias podem proporcionar a aproximação de grupos semelhantes no espaço atenuando as diferenças na sobreposição relativa entre as moléculas nos Alinhamentos II e III, levando a uma melhora na qualidade estatística dos modelos de CoMFA gerados. Desta forma, as conformações obtidas após a análise por similaridade morfológica, bem como aquelas obtidas por acoplamento molecular foram realinhadas espacialmente com base em grupos funcionais e átomos comuns, conforme descrito na seção experimental, gerando os Alinhamentos IV e V (Figura 32).

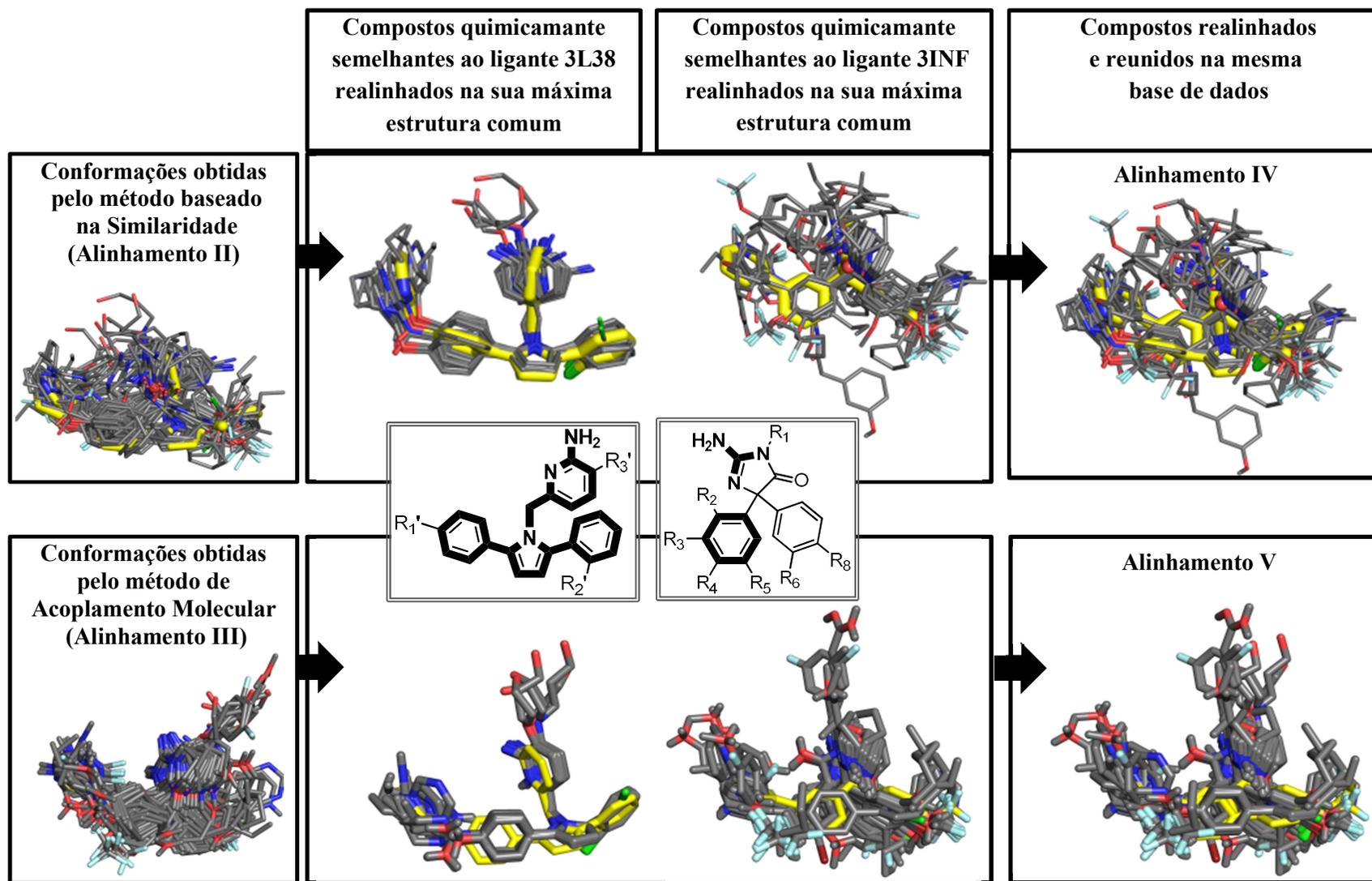


Figura 33: Estratégia para o realinhamento molecular das conformações oriundas dos Alinhamentos II (acima) e alinhamento III (abaixo), pelo método da máxima estrutura comum, para a geração dos alinhamentos IV e V respectivamente. Os ligantes complexados na estrutura da BACE-1 e 3INF e 3L38 estão representados em amarelo.

5.3.1.2. Desenvolvimento e Validação dos Modelos de CoMFA

Os descritores (campos estéricos e eletrostáticos) foram calculados inicialmente, para cada molécula do conjunto treino, nas conformações e orientações espaciais definidas nos alinhamentos I, II e III usando os parâmetros padrões dos campos moleculares e espaçamento da grade de 2,0 Å. Embora modelos preliminares (Tabela 6) apresentem valores estatisticamente aceitáveis (q^2 próximos a 0,70), observa-se uma discrepância entre os valores q^2 e r^2 ($> 0,25$). Essas diferenças podem ser devido a um sobreajuste dos modelos, seja por valores de r^2 superestimados, ou por conta da baixa preditividade estimada para o conjunto durante a validação (baixo valor de q^2)¹⁵¹.

Tabela 6: Modelos de CoMFA obtidos com parâmetros padrão e espaçamento de 2Å.

Alinhamento	Parâmetros Estatísticos					
	q^2	r^2	SEE	PCs	FS (%)	FE(%)
I (Modelo 1)	0,66	0,92	0,250	5	46	54
II (Modelo 2)	0,68	0,94	0,208	4	38	62
III (Modelo 3)	0,69	0,97	0,161	6	48	52

q^2 : Coeficiente de Correlação por Validação Cruzada; r^2 : Coeficiente de Correlação sem Validação Cruzada; SEE: Erro Padrão de Estimativa; PC: Número Ótimo de Componentes; FS: Fração de Contribuição do Campo Estérico; FE: Fração de Contribuição do Campo Eletrostático

A fim de garantir resultados de maior confiabilidade estatística novos modelos foram gerados usando o procedimento de focagem da região, o qual permite a modificação do espaçamento do *grid* para o cálculo dos campos moleculares, bem como a utilização de um coeficiente discriminante como alternativa para destacar somente os componentes do CoMFA, gerados na análise prévia de PLS, que contribuem para a predição da potência inibitória. Nesta etapa, diferentes tamanhos de espaçamentos foram considerados (0,5, 1,0, 1,5 e 2Å) e os campos moleculares foram multiplicados por um peso fixo (0,3, 0,6, 0,9 ou 1,2) utilizando-se a opção '*Stdev_Coefficients*' SDC. A combinação de ambos os parâmetros deu origem a 16 modelos de CoMFA para cada alinhamento (Tabela 7).

A análise da influência desses dois fatores mostra que para o mesmo espaçamento (Å) os valores de q^2 não variam significativamente, obtendo-se em alguns casos, valores estatísticos idênticos. Neste caso, o melhor modelo foi selecionado pela capacidade de predição da validação externa. A maioria dos modelos apresentaram valores elevados de q^2 com o espaçamento de 1,0 Å e SDC 0,6. Adicionalmente, é possível observar que a discrepância entre os valores de q^2 e r^2 foi marcadamente reduzida, se comparado aos resultados obtidos com a análise previamente discutida (Tabela 6).

Visando comparar os modelos obtidos, segue abaixo uma breve descrição das principais características e resultados provenientes de cada um deles.

Tabela 7: Influência da focagem da região e do espaçamento da grade nos valores de q^2 dos modelos de COMFA.

Alinhamento	GS (Å)	Valor de SDC (Focus COMFA)			
		0.3	0.6	0.9	1.2
		q^2 / PCs			
I	0.5	0,75/5	0,75/5	0,75/5	0,73/6
	1.0	0,77/5	0,77/5	0,76/6	0,74/6
	1.5	0,71/5	0,72/5	0,75/5	0,72/5
	2.0	0,55/3	0,56/3	0,58/3	0,60/6
II	0.5	0,72/4	0,73/5	0,74/6	0,74/6
	1.0	0,75/4	0,76/4	0,76/5	0,75/5
	1.5	0,53/4	0,55/5	0,54/5	0,54/5
	2.0	0,57/5	0,58/6	0,60/6	0,58/6
III	0.5	0,73/6	0,75/6	0,75/5	0,74/6
	1.0	0,76/5	0,78/5	0,78/6	0,77/6
	1.5	0,56/6	0,55/6	0,53/6	0,52/6
	2.0	0,59/3	0,61/3	0,62/3	0,61/3
IV	0.5	0,69/4	0,70/4	0,70/4	0,69/4
	1.0	0,72/4	0,73/4	0,73/5	0,73/6
	1.5	0,70/4	0,72/4	0,71/4	0,73/5
	2.0	0,58/5	0,60/4	0,59/3	0,59/5
V	0.5	0,73/6	0,77/6	0,79/6	0,79/5
	1.0	0,79/5	0,81/4	0,82/3	0,81/3
	1.5	0,64/4	0,63/5	0,62/5	0,60/4
	2.0	0,58/5	0,61/4	0,59/5	0,57/5

O melhor modelo de CoMFA obtido aplicando-se a focagem da região para o Alinhamento I, apresentou elevada significância estatística ($r^2 = 0,90$, $q^2 = 0,77$ / 5 PCs – Modelo A, Tabela 8) e capacidade preditiva ($r^2_{pred} = 0,73$) indicando que esse alinhamento pode produzir modelos preditivos e com consistência interna razoável. Já o melhor resultado de CoMFA obtido a partir do Alinhamento II (modelo B- Tabela 8), forneceu um modelo mais simples, (4PCs), comparado ao modelo A e com valor de $q^2 = 0,76$. Contudo, é possível observar que 3 moléculas do conjunto teste (**2**, **25** e **77** – Apêndice B) exibem diferenças significativas ($> 0,75$) entre os valores residuais, contribuindo para a redução da capacidade de predição global do modelo ($r^2_{pred} = 0,64$).

Os resíduos de um modelo de regressão contêm as informações que justificam as razões pelas quais o próprio modelo não consegue explicar 100% da variabilidade dos valores observados de propriedade biológica. Em geral, isso ocorre devido à presença de erros aleatórios relativos à determinação quantitativa do ensaio experimental ou devido à especificação imprópria do modelo (falta de ajuste)¹⁵². Nos casos em que o modelo é bem ajustado, o conjunto de resíduos descreve apenas os erros experimentais e, portanto, a distribuição dos resíduos deve revelar um padrão estritamente aleatório. Porém, o gráfico dos valores residuais obtidos com o modelo B (Figura 34) mostra ao contrário, um padrão não aleatório da distribuição dos pontos, evidenciando que os resíduos contêm erros sistemáticos devidos à especificação incorreta do modelo.

Devido à baixa confiabilidade estatística do modelo B, este foi excluído das análises. Na sequência, os mesmos protocolos descritos para os modelos anteriores foram empregados para construir modelos CoMFA para as moléculas do Alinhamento III (Figura 31). O melhor modelo obtido apresenta elevada capacidade de correlação ($r^2 = 0,95$ e $q^2 = 0,78$, com 5PCs modelo C - Tabela 8) e se mostrou robusto quanto a sua capacidade de predição externa ($r^2_{pred} = 0,78$).

Tabela 8: Resultados dos melhores modelos de CoMFA obtido com os Alihamentos I, II, III, IV e V.

Parâmetros Estatísticos	ALINHAMENTO				
	I Modelo A	II Modelo B	III Modelo C	IV Modelo D	V Modelo E
GS (Å) ^a	1.0	1.0	1.0	1.0	1.0
SDC ^b	0.6	0.6	0.6	0.6	0.9
q^2 ^c	0,77	0,76	0,78	0.73	0.82
PC ^d	5	4	5	4	3
r^2 ^e	0,90	0,94	0,95	0.91	0,89
SEE ^f	0.271	0.213	0.193	0.259	0.285
F ^g	128.9	271.3	268.4	178.6	192.1
r^2_{pred} ^h	0,73	0,64	0.78	0.86	0.76
FS (%) ⁱ	54	41	47	46	49
FE (%) ^j	46	59	53	54	51

^aEspaçamento da Grade; ^bValor do coeficiente para focagem da região por *Standard Deviation Coefficient*; ^cCoefficiente de correlação por validação cruzada; ^dNúmero ótimo de Componentes Principais; ^eCoefficiente de correlação sem validação cruzada; ^fErro padrão estimado.; ^gValor do teste F. ^hCoefficiente de correlação para a predição do conjunto teste; ⁱFração de Contribuição do Campo Estérico; ^jFração de Contribuição do Campo Eletrostático

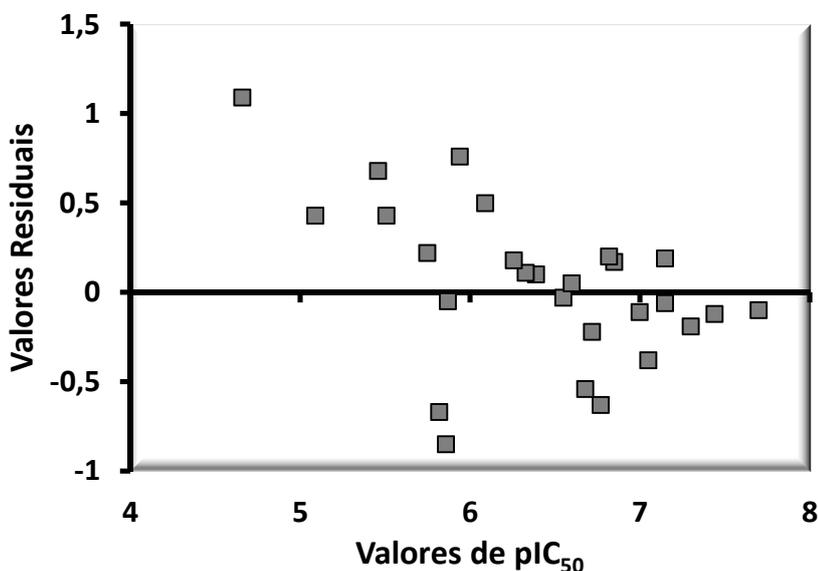


Figura 34: Valores residuais (diferença entre pIC_{50} predito e observado experimentalmente) dos compostos do conjunto teste, para o melhor modelo de CoMFA obtido com o Alinhamento II (modelo B).

Finalmente, na quarta estratégia de sobreposição, considerando as conformações oriundas dos alinhamentos II e III, foi realizado um realinhamento rígido das conformações para obtenção dos alinhamentos IV e V. O modelo gerado a partir do alinhamento IV (Modelo D – Tabela 8) apresentou valores estatísticos ligeiramente inferiores ($r^2=0,91$ $q^2=0,73$ e 4PCs) aos resultados obtidos no modelo B. Porém, a capacidade de predição externa desse modelo mostrou-se significativamente maior ($r^2_{pred} = 0,86$). Já o alinhamento V (modelo E) forneceu resultados de $r^2=0,89$, $q^2=0,82$ e 3PCs, sendo que, diferentes dos demais modelos discutidos, o melhor resultado foi obtido utilizando peso fixo atribuído na focagem por SDC de 0.9 (Tabela 6). Neste caso, a preditividade externa evidenciou valores semelhantes ao modelo inicial baseado no acoplamento molecular ($r^2_{pred}=0,76$). A pequena diferença numérica entre o pIC₅₀ obtido experimentalmente e o pIC₅₀ predito para os compostos, de acordo com ambos os modelos D e E (Figura 35), evidencia boa a confiabilidade e consistência para os dois modelos. Esses dados sugerem que o realinhamento gerado proporcionou uma melhora na posição relativa dos compostos comparado aos alinhamentos iniciais baseados na similaridade e no acoplamento molecular.

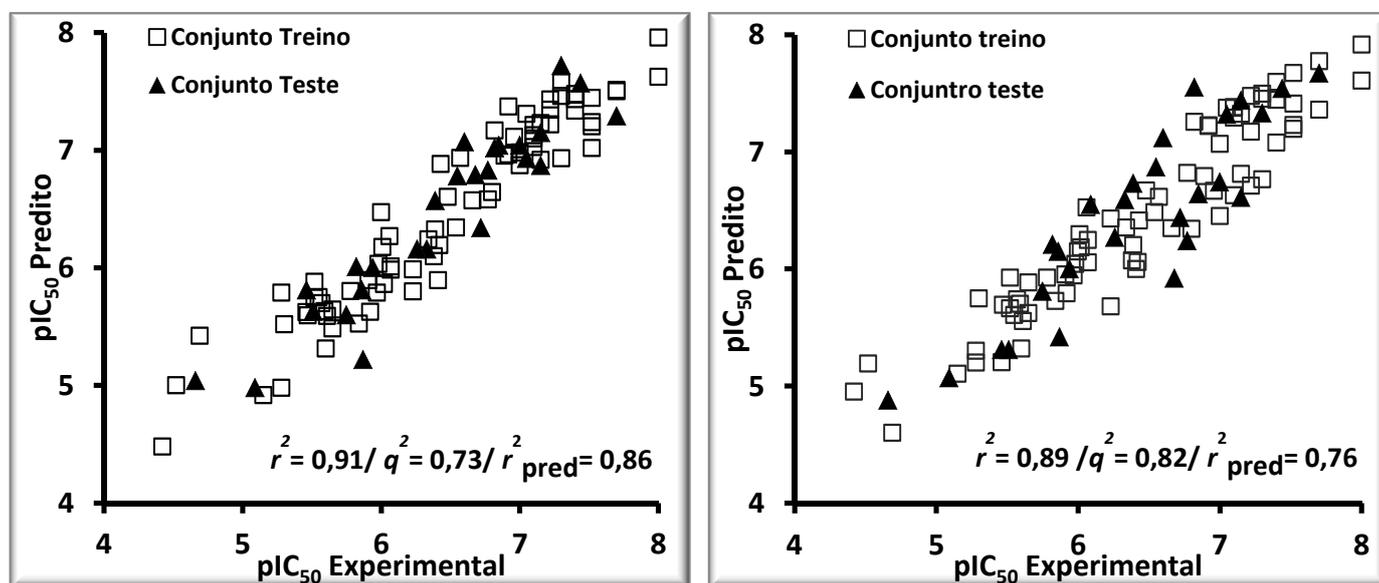


Figura 35: Valores preditos versus experimentais de pIC₅₀ para o conjunto treino (quadrado) e teste (triângulos) de acordo com o melhor modelo obtido com o alinhamento IV (esquerda) e alinhamento V (direita).

De forma geral, os modelos de CoMFA apresentaram qualidade estatística comparável e portanto, foi feita uma análise das informações oriundas de cada modelo a fim de identificar pontos consensuais que pudessem guiar o planejamento de inibidores mais potentes.

5.3.1.3. Interpretação dos Mapas de Contorno do CoMFA

Os mapas de contorno gerados através método de CoMFA são poliedros gerados no espaço por interpolação dos respectivos termos de energia estérea e eletrostática calculadas e são importantes para explicar os diferentes valores de potência apresentado pelos compostos da série em estudo. Esses mapas são projetados nas regiões cujos pontos estão relacionados à variância das propriedades estéreas e eletrostáticas dos ligantes. Dessa maneira, a ausência de poliedros em determinadas áreas não implica exatamente que um dado grupo ou fragmento molecular não tenha importância, apenas que todos os compostos examinados exercem naquela região a mesma influência¹⁵³.

Os campos 3D representados por poliedros nas cores verdes e amarelas referem-se a características estereoquímicas, enquanto aqueles em azul e vermelho se referem a características eletrostáticas. Os campos em verde sugerem substituições com grupos mais volumosos, indicando que há pouca ou nenhuma restrição estérea em determinada região e os campos amarelos, em oposto, referem-se a regiões do espaço as quais são favoráveis grupos menos volumosos. A presença de poliedros azuis em torno da molécula é sugestiva para a introdução de grupos com maior carga positiva enquanto que poliedros vermelhos sugerem substituintes mais eletronegativos. Para a melhor compreensão dos mapas de contorno, uma descrição prévia do sítio ativo da BACE-1 é mostrada.

A estrutura tridimensional da BACE-1 evidencia que o domínio catalítico da enzima possui uma longa fenda para o reconhecimento do substrato, com dois resíduos aspárticos posicionados no sítio ativo, no qual bolsões laterais apresentam resíduos de aminoácidos que interagem com o substrato para corretamente posicioná-los durante o mecanismo de catálise¹⁵⁴. A estrutura cristalográfica 3INF (Figura 36) mostra o núcleo aminoimidazol do ligante complexado (**66** - Apêndice A) orientado frente aos resíduos aspárticos catalíticos (Asp²⁹⁰ e Asp⁹⁴). Moléculas de água se distribuem ao longo da fenda catalítica até o bolsão S2', que é rico em grupos polares/carregados. Adjacente à

esta região uma cavidade estreita (FLAP) é delimitada por um resíduo tirosina (Tyr133) que altera a sua conformação para melhor acomodar o substrato. O bolsão S1 é o mais hidrofóbico sendo rico em aminoácidos aromáticos, enquanto que S3 constitui uma cavidade mais larga com característica mista, de hidrofobicidade e hidrofiliidade⁸⁰.

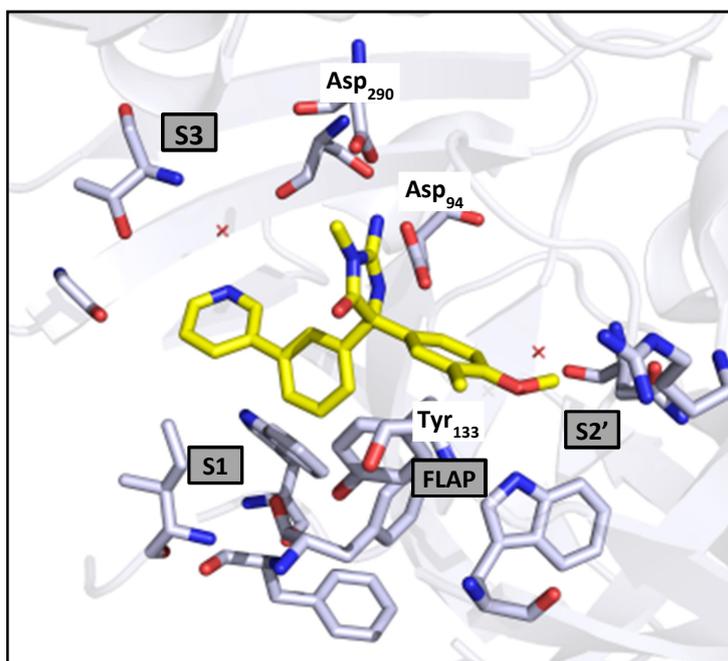


Figura 36: Representação simplificada do sítio ativo da BACE-1 (PDB: 3INF).

Mapas de contorno estéreo e eletrostáticos foram gerados para um inibidor de alta potência ($pIC_{50} = 7,5$ - Figura 37) segundo os modelos C, D e E os quais apresentaram maior capacidade de preditiva. Os descritores de campos eletrostáticos para os modelos C, D e E (Tabela 8) explicam 53% e 54%, 51%, respectivamente, da variância total dos dados, enquanto que os estereoquímicos explicam 47%, 46% e 49% indicando que as características eletrostáticas são mais importantes para diferenciar a atividade inibitória dos compostos. Este dado corrobora os modelos de QSAR 2D, que ressaltam a importância de descritores que descrevem propriedades eletrônicas para a potência das moléculas.

Algumas informações evidenciadas são comuns a todos os mapas, como por exemplo, a interação desfavorável na região próxima ao núcleo aminoimidazol, indicada por poliedros amarelos projetados em direção a área acessível ao solvente.

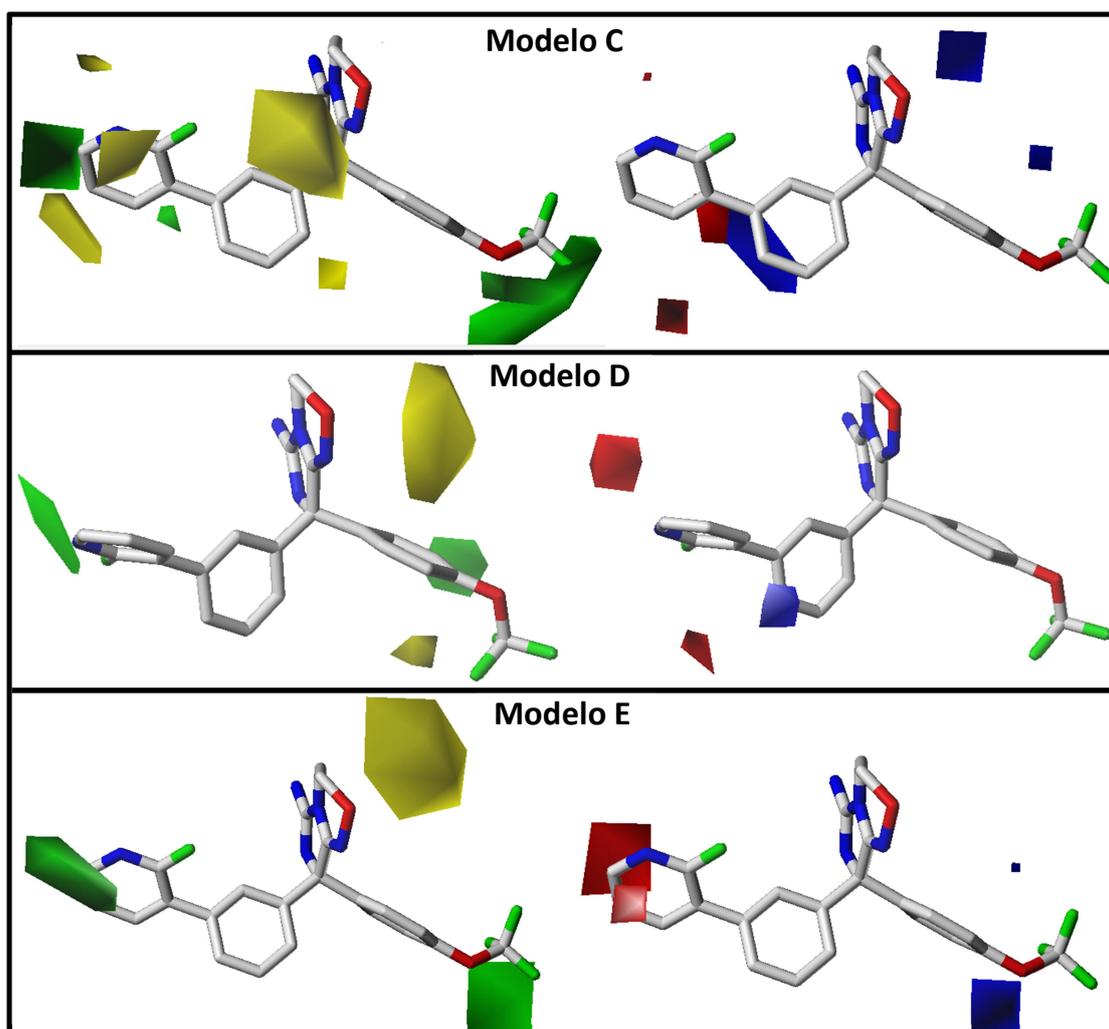


Figura 37: Mapas de contorno para o composto **20** ($pIC_{50} = 7,5$) de acordo com os melhores modelos de CoMFA. Os mapas estéreo (esquerda) estão mostrados em verde e amarelo (contornos em 0,05 Kcal/mol e -0,017 Kcal/mol, respectivamente), enquanto que e os mapas eletrostáticos (à esquerda) estão mostrados em vermelho e azul (contornos em 0,078 e -0,065 Kcal/mol).

Os mapas de contorno mostram ainda, que não há restrição estérea significativa na posição *para* do anel *o*-cloro piridínico, sendo favorável para a potência a introdução de um átomo ou grupo eletronegativo. No acoplamento do composto **20** no sítio ativo da BACE-1 (PDB: 3INF), observa-se de forma mais clara que um substituinte em *para* é capaz de realizar uma interação com um resíduo de glicina (Gly₇₃) no bolsão **S3** da macromolécula (Figura 38). É válido referir que o anel *m*-piridínico foi estrategicamente usado para aumentar a afinidade de inibidores pela enzima por meio de interações de hidrogênio de uma molécula de água (W470 – PDB 3INF) conservada e um resíduo de

serina (Ser291 – PDB 3INF), próximo à cavidade catalítica da BACE-1⁸⁰⁻⁸³ (Figura 38). Visto que a maioria dos compostos com pIC_{50} superior a 6,5 apresentam o anel *m*-piridínico (ex: **44** - **55** , Apêndice A), considera-se que a manutenção deste grupo é crucial e assim, derivados 3,4-dissubstituídos devem ter potência melhorada.

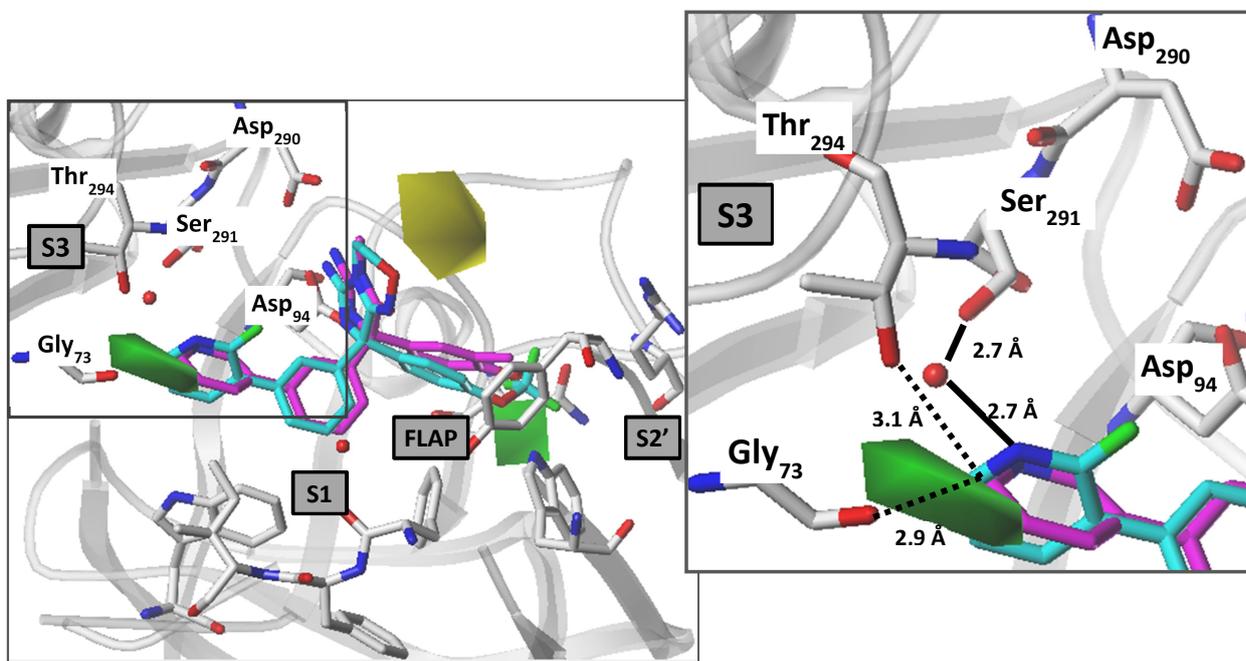


Figura 38: Acoplamento do inibidor **20** ($pIC_{50} = 7,50$) no sítio ativo da BACE-1(PDB: 3INF). A estrutura do ligante complexado está em rosa. Os mapas de contorno estéreo foram obtidos com o modelo E (verde: 0,05 Kcal/mol; amarelo: -0,017 Kcal/mol). Em destaque, o poliedro verde indica a substituição favorável em *para* no anel *m*-piridínico, possibilitando a formação de interações de hidrogênio adicionais com a carbonila dos resíduos Gly73 / Thr294 (setas tracejadas). A linha sólida em negrito destaca a interação do inibidor com o resíduo Ser 291, por meio de uma molécula de água conservada.

Mapas de contorno eletrostáticos obtidos com os modelos C e D exibem poliedros azuis em torno do anel benzênico ligado diretamente ao núcleo aminoimidazol, indicando que anéis eletrodeficientes nessa região são favoráveis para a potência. Neste caso, é possível sugerir a introdução de pequenos grupos/ átomos elétron-sacadores no benzeno (ex: -CN, NO₂, -F, -Cl, etc). A importância desta substituição pode ser observada entre os compostos com potência mais elevada da série (ex: **60** – $pIC_{50} = 8,0$, **63** e **64** – $pIC_{50} = 7,7$) os quais apresentam na sua estrutura um anel *p*-flúor- benzênico.

Adicionalmente, campos estéreos e eletrostáticos ressaltam modificações direcionadas à região S2' (Figura 39). O poliedro em azul obtido com o modelo E mostra que é favorável a introdução de grupos eletrodeficientes na cavidade imediatamente adjacente à região FLAP, enquanto que poliedros em verde nesta mesma região, exibidos por todos os mapas de contorno, indicam que é possível a introdução de grupos volumosos. É válido referir que a área em questão apresenta restrição estérea parcial, especialmente devido a presença de um resíduo de triptofano (Trp138 em 3INF), o qual é capaz de realizar interações adicionais com os ligantes. Como prova disso, o átomo de oxigênio do grupo trifluormetoxi no composto **20**, interage com o N de Trp138 contribuindo positivamente para a atividade da série^{81,82}.

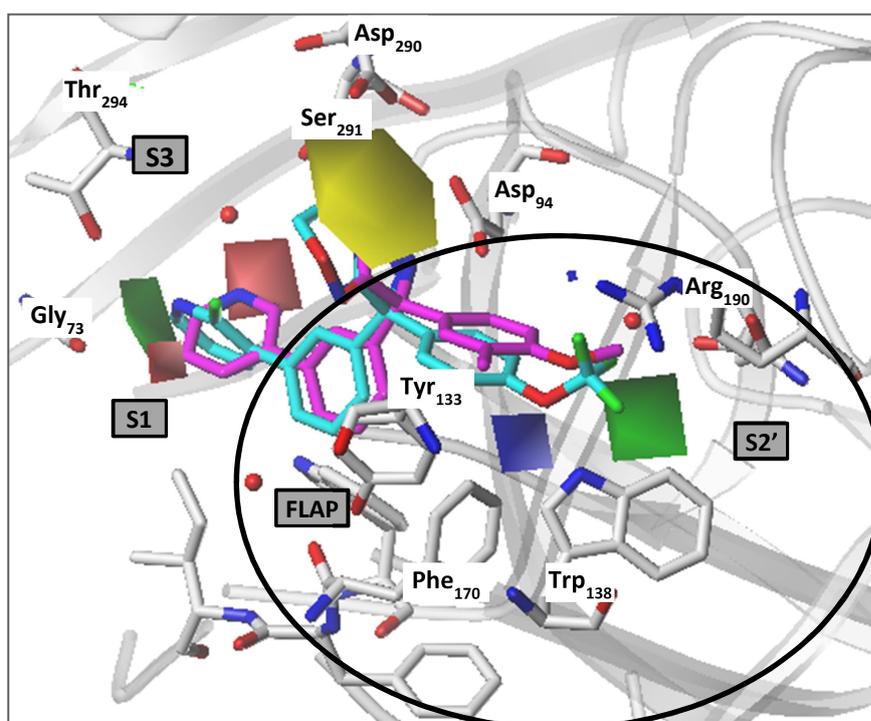


Figura 39: Acoplamento do inibidor **20** ($pIC_{50} = 7,50$) no sítio ativo da BACE-1(PDB: 3INF). A estrutura do ligante complexado está em rosa. Mapas de contorno estéreos e eletrostáticos foram obtidos com o modelo E (verde: 0,05 Kcal/mol; amarelo: -0,017 Kcal/mol; azul: 0,078 Kcal/mol; vermelho: 0,065 Kcal/mol). Em destaque a região FLAP e S2'.

Por fim, a partir do melhor modelo (modelo D), as estratégias de modificações moleculares sugeridas foram reunidas e espera-se que estas informações possam auxiliar na obtenção de derivados congêneres, inibidores de BACE-1 mais potentes (Figura 40).

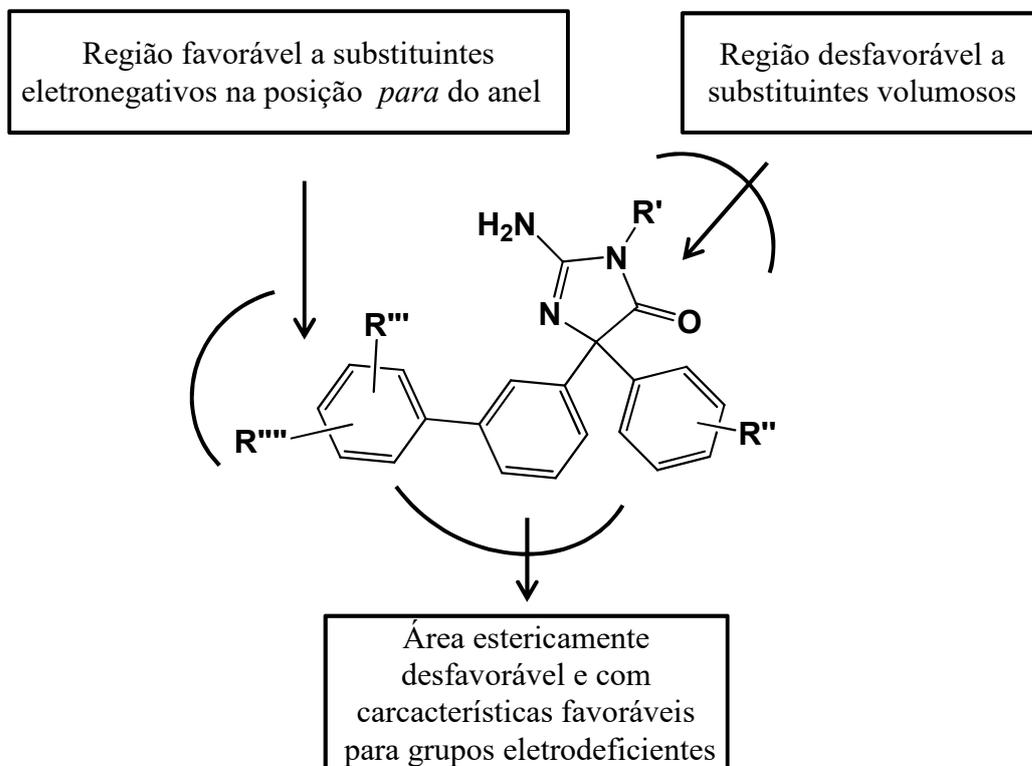


Figura 40: Estratégias de modificações moleculares a partir da análise dos mapas de contorno dos modelos de CoMFA para aumentar a atividade da série de compostos em estudo.

6. CONCLUSÃO

A diversidade química e a distribuição dos valores de propriedade farmacológica (potência inibitória) no conjunto de dados organizado neste trabalho é apropriado, permitindo considerável cobertura do espaço químico biológico. Os modelos finais de HQSAR para os inibidores de BACE-1 apresentaram bons coeficientes de correlação e capacidade preditiva satisfatória. Paralelamente, os modelos quimiométricos de QSAR 2D baseados em descritores topológicos apresentaram consistência interna e externa elevada, indicada pelos valores dos coeficientes de correlação e de preditividade externa, os quais foram superiores aos obtidos com hologramas moleculares (modelos de HQSAR). Os descritores GGI5 e GGI6, de maior contribuição para o vetor de regressão do melhor modelo de QSAR 2D, se referem transferência de cargas entre pares de átomos separados a uma ordem de 5 e 6 átomos, respectivamente, indicando que propriedades eletrônicas são importantes na caracterização de inibidores não peptídeomiméticos de BACE-1 com elevada potência. Já os descritores B08[N-N] e B09[N-N], que também exibiram relevante contribuição para o vetor de regressão, descrevem a presença de fragmentos de cadeia que possuem átomos de N separados a uma distância topológica de 8 e 9 átomos. A visualização desse fragmento na estrutura de moléculas de alta potência da série pôde ser feita em concordância com os mapas de contorno do HQSAR, evidenciando que modelos de QSAR 2D baseado em descritores complementam informações sugeridas pelo método baseado em fragmentos.

Os modelos de QSAR 3D gerados para o mesmo conjunto treino usado nas análises de QSAR 2D, utilizando diferentes estratégias alinhamentos, destacam a importância da sobreposição relativa na robustez dos modelos. Nesse sentido, o realinhamento rígido das conformações obtidas pelo método de sobreposição baseado na similaridade forneceu um modelo com capacidade preditiva marcadamente superior ao modelo previamente obtido. Ao passo que, a mesma estratégia aplicada à sobreposição baseada no acoplamento molecular permitiu a obtenção de um novo modelo com melhor interpretação (menor fragmentação) dos mapas de contorno. Adicionalmente, os mapas de contorno do CoMFA, obtidos com técnicas distintas de alinhamento, forneceram modelos robustos cuja interpretação evidenciaram características comuns, que reunidas, são guias úteis para guiar o planejamento de inibidores congêneres mais potentes.

A aplicação de estratégias integradas em química medicinal é importante no planejamento de fármacos. O uso de métodos baseados na estrutura do receptor em conjunto com o emprego de técnicas de QSAR 3D e QSAR 2D confirma mais uma vez que é válido associar estratégias procurando uma sinergia útil no planejamento de novas moléculas. Os modelos finais de QSAR devem ser úteis para guiar futuros trabalhos em química medicinal no planejamento de novos inibidores não peptidomiméticos de BACE-1 com potência otimizada.

REFERÊNCIAS

- 1- WHO. Active Ageing, A Policy Framework. A contribution of the WHO to the Second United Nations World Assembly on Ageing. Madrid, Spain, April, 2002.
- 2- SILVA, F. M. L.; NEVES, R.P.C.N.; SILVA, M. de M. M. LOBONATI. Processo de Envelhecimento e suas Alterações nos Diversos Sistemas Orgânicos. **Jornal Brasileiro de Medicina**. Edição especial, n.3, p.43-48, 2005.
- 3- NAUGHTON, C.; BBENNETT, K., FEELY, J. Prevalence of chronic disease in the elderly based on a national pharmacy claims database. **Age Ageing**, v.36 p.633-636, 2006.
- 4- DE PAULA, J. A.; ROQUE, F. P.; ARAUJO, F. S. Qualidade de vida em cuidadores de idosos portadores de demência de Alzheimer. **Jornal Brasileiro de Psiquiatria**, Rio de Janeiro, v. 57, n. 4, 2008.
- 5- WHO. World Health Report 2003 – shaping the future. Genebra; 2003.
- 6- GALLUCCI , N. J.; TAMELINI, M. G.; FORLENZA. O.V. Diagnóstico diferencial das demências. **Revista de Psiquiatria Clínica** v.32 n.3 p. 119-130, 2005.
- 7- LYKETSOS C. G. , CARRILLO, M. C. RYAN, J. M. , KHACHATURIAN A. S. et al, Neuropsychiatric symptoms in Alzheimer’s disease **Alzheimer's and Dementia** v.7, p. 532-539, 2011.
- 8- ALZHEIMER’S ASSOCIATION, Alzheimer’s disease facts and figures. **Alzheimer's and Dementia** v. 7, n. 2, p 208-244, 2011.
- 9- BROOKMEYER, R. JOHNSON, E.; ZIEGLER-GRAHAM, K.; ARRIGHI, M. H. Forecasting the global burden of Alzheimer’s disease **Alzheimer's and Dementia** v.3, p. 186-191, 2007.
- 10- TATSCH M. F., BOTTINO C. M, AZEVEDO D., HOTOTIAN S. R., et al Neuropsychiatry symptoms in Alzheimer disease and cognitively impaired, non demented elderly from a community-based sample in Brazil: prevalence and relationship with dementia severity., **American Journal of Geriatric Psychiatry**, v. 14, p.438-45.,2006.
- 11- HERRERA, E.; CARAMELLI, P.; SILVEIRA, A. S. B.; NITRINI, R. Epidemiologic survey of dementia in a community-dwelling Brazilian population. **Alzheimer Disease & Associated Disorders**, v. 16, p. 103-108, 2002.
- 12- CRAIG, M. C.; MURPHY, D. G.M. Alzheimer’s disease in women **Best Practice & Research Clinical Obstetrics & Gynaecology** v. 23, p. 53-61, 2009

- 13- FLORENCE, R.; PHILIPPE, A. Genetic susceptibility factors for Alzheimer's disease. **European Journal of Pharmacology** v. 412, n.1, p. 1-12, 2001.
- 14- MARTINS, I. J.; HONE, E.; FOSTER, J. K.; SUNRAM-LEA, S. I.; GNJEC, A.; et al. Apolipoprotein E, cholesterol metabolism, diabetes, and the convergence of risk factors for Alzheimer's disease and cardiovascular disease. **Molecular Psychiatry**, v. 11, p. 721-736, 2006
- 15- ANSTEY, K. J.; VON SANDEN, C.; SALIM, A.; O'KEARNEY, R. Smoking as a risk factor for dementia and cognitive decline: a meta-analysis of prospective studies. **American Journal of Epidemiology**, v. 166, p. 367-378, 2007.
- 16- SESHADRI, S.; BEISER, A.; KELLY-HAYES, M.; KASE, C.S.; AU, R; et al. "The lifetime risk of stroke: Estimates from the Framingham Study." **Stroke** v.37 n.2 p.345–350. 2006
- 17- HEBERT, L. E.; BECKETT, L. A.; SCHERR, P. A.; EVANS, D. A. "Annual incidence of Alzheimer disease in the United States projected to the years 2000 through 2050." **Alzheimer Disease & Associated Disorders** v.15 n.4 p. 169–173, 2001.
- 18- MESULAM, M. The Cholinergic Lesion of Alzheimer's Disease: Pivotal Factor or Side Show? **Learning Memory** v.11 p. 43-49; 2004.
- 19- SERENIKI, A.; VITAL, M. A. B. F.. A doença de Alzheimer: aspectos fisiopatológicos e farmacológicos. **Revista de Psiquiatria do Rio Grande Sul**, v. 30, n. 1, 2008.
- 20- CAMPOS, M. O.; RODRIGUES, J. F. N. Doenças crônicas não transmissíveis: fatores de risco e repercussão na qualidade de vida Revista Baiana de Saúde Pública v.33, n.4, p.561-581 out./dez. 2009
- 21- LOPES, M. A.; BOTTINO, C. M. C. Prevalência de demência em diversas regiões do mundo: Análise dos estudos epidemiológicos de 1994 a 2000. **Arquivos de Neuropsiquiatria**, v. 60, n. 1, 2002.
- 22- ENGELHARDT, E.; BRUCKI, S. M.T.; CAVALCANTI, J. L. S.; FORLENZA, O.V. et al. Tratamento da doença de Alzheimer: recomendações e sugestões do Departamento Científico de Neurologia Cognitiva e do Envelhecimento da Academia Brasileira de Neurologia. **Arquivos de Neuropsiquiatria.**, v.63, n.4, p. 1104-1112, 2005.
- 23- REFOLO, L. M; FILLIT, H. M. Drug discovery for Alzheimer's disease: the end of the beginning. **Journal of molecular neuroscience**, v. 24, n.1, p.1-8, 2004
- 24- CAI, H. WANG; Y. MCCARTHY, D.; WEN, H.. BACE1 is the major beta-secretase for generation of a beta peptides by neurons. **Nature Neuroscience** v. 4 p. 233–234, 2001.

- 25- DINGWALL, C.. Spotlight on BACE: the secretases as targets for treatment in Alzheimer disease. **Journal of Clinical Investigation**. v.108, n.9, p. 1243-1246, 2001.
- 26- GHOSH, A. K.; GEMMA, S.; TANG J.. β -Secretase as a Therapeutic Target for Alzheimer's Disease. **Neurotherapeutics** v. 5, n. 3, p. 399-408, 2008.
- 27- CHAVATTE, P.; YOUS, S.; MAROT C.; BAURIN, N.; LESIEUR, D. Three-dimensional Quantitative Structure-Activity Relationships of Cyclo-Oxygenase-2 (COX-2) Inhibitors: A Comparative Molecular Field Analysis, **Journal Medicinal Chemistry**., v.44, p. 3223-3230, 2001.
- 28- WATERBEEMD, H. V.; ROSE, S. Quantitative approaches to structure-activity relationships. In: WERMUTH, C. G. **The Practice of Medicinal Chemistry**, 2 ed., Elsevier, cap. 22, p. 351-369, 2003.
- 29- MOHS, R. C. The clinical syndrome of Alzheimer's disease: aspects particularly relevant to clinical trials. **Genes Brain and Behavior**, v. 4, p. 129-133, 2005.
- 30- SOCIETY FOR NEUROSCIENCE. Unraveling Alzheimer's Disease, 2010 http://www.sfn.org/index.aspx?pagename=brainBriefings_10_unravelingalzheimers. (Acessado em 18/10/2011).
- 31- LAFERLA, F. M.; GREEN, K. N.; ODDO, S. Intracellular amyloid- β in Alzheimer's disease. **Nature Review Neuroscience**, v.8, p.499–509, 2007
- 32- MARCELLO, E.; EPIS, R.; DI LUCA, M. Amyloid flirting with synaptic failure: Towards a comprehensive view of Alzheimer's disease pathogenesis. **European Journal of Pharmacology**, v. 585 p. 109–118, 2008.
- 33- KAJ BLENNOW. Alzheimer's disease . **The Lancet**, v.368, n. 9533, p. 387-403, 2006.
- 34- LING, Y.; MORGAN, K.; KALSHEKER, N.. Amyloid precursor protein (APP) and the biology of proteolytic processing: relevance to Alzheimer's disease. **The International Journal of Biochemistry & Cell Biology** v. 35, p. 1505–1535, 2003.
- 35- HAASS, C., SELKOE, D. J., 2007. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β -peptide. **Nature** v. 8, p. 101-112, 2007.
- 36- OMAR, M. A.; DEVINDER S. M.; BHRAMA P. P.; AUSTEN. B. M. Oligomerization and Toxicity of β -Amyloid-42 Implicated in Alzheimer's Disease **Biochemical and Biophysical Research Communications**, v. 273, n.3, p. 1003-1007, 2000.
- 37- FINDEIS, M. A. The role of amyloid β peptide 42 in Alzheimer's disease Review Article **Pharmacology & Therapeutics**, v. 116, n. 2, p. 266-286, 2007.

- 38- GOLDE, T. E., Disease modifying therapy for AD? **Journal of Neurochemistry**. v. 99, n. 3, p. 689-707, 2006.
- 39- AVILA, J. LUCAS, J. J., PÉREZ M. HERNÁNDEZ, F. Role of Tau Protein in Both Physiological and Pathological Conditions. **Physiological Review**, v. 84 n. 2 p. 361-384, 2004.
- 40- HERNÁNDEZ, F.; DE BARREDA, E. G.; FUSTER-MATANZO, Almudena; LUCAS, José J., AVILA, Jesús GSK3: A possible link between beta amyloid peptide and tau protein **Experimental Neurology** v. 223, n.2, p. 322-325, 2010.
- 41- IQBAL, K., ALONSO, Alejandra del C., CHEN, She, CHOCHAN, M. O., EL-AKKAD, Ezzat , et al , Tau Pathology in Alzheimer disease and other tauopathies. **Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease** v.1739, p.198-210, 2005.
- 42- SOMMER, B. Alzheimer's disease and the amyloid cascade hypothesis: ten years on **Current Opinion in Pharmacology**, v. 2, n. 1, p. 87-92. 2002.
- 43- COLE, S. L.; VASSAR, R.. BACE1 Structure and Function in Health and Alzheimer's Disease. **Current Alzheimer Research**, v.5, n. 2, p. 100-120, 2008.
- 44- CITRON, M. Alzheimer's disease: strategies for disease modification, **Drug Discovery**, v.9, p. 387-398, 2010.
- 45- FORLENZA, O. V. Tratamento farmacológico da doença de Alzheimer. **Revista de Psiquiatria Clínica**, v. 32, p. 137-148, 2005.
- 46- ROBLES, Alfredo. Pharmacological Treatment of Alzheimer's Disease: Is it Progressing Adequately? **The Open Neurology Journal**, v. 3, p. 27-44, 2009
- 47- MESULAM, M. M. The Cholinergic Lesion of Alzheimer's Disease: Pivotal Factor or Side Show? **Learning and Memory**. v.11, p.43-49, 2004.
- 48- FRANCIS, P. T.; PALMER, A. M.; SNAPE, M.; WILCOCK, G. K. The cholinergic hypothesis of Alzheimer's disease: a review of progress. **Journal of Neurology Neurosurgery and Psychiatry**, v. 66, p. 137-147, 1999.
- 49- LANCTÔT, K. L.; HERRMANN, N.; LOULOU, M. M. Correlates of response to acetylcholinesterase inhibitor therapy in Alzheimer's disease. **Journal Psychiatry Neuroscience**; v.28, n.1,p.13-26, 2003.
- 50- DARVESH, S., WALSH, R, KUMAR, R., CAINES, A., et al Inhibition of human cholinesterases by drugs used to treat Alzheimer disease, **Alzheimer Disease and Associated Disorders**.v.17, n.2, p.117-26. 2003.
- 51- DARVESH S., HOPKINS D.A. Differential distribution of butyrylcholinesterase and acetylcholinesterase in the human thalamus. **Journal of Computational Neuroscience**. v. 463, p.25-43, 2003.

- 52- MESULAM, M.M., GUILLOZET, A., SHAW, P., LEVEY, A. et al. Acetylcholinesterase knockouts establish central cholinergic pathways and can use butyrylcholinesterase to hydrolyze acetylcholine. **Neuroscience** v. 110 p. 627–639, 2002.
- 53- BALLARD C. G. Advances in the treatment of Alzheimer's disease: Benefits of dual cholinesterase inhibition. **European Neurology**, v. 47, p. 64-70, 2002.
- 54- HARVEY, A. L. The pharmacology of galanthamine and its analogues **Pharmacology & Therapeutics** v. 68, n.1, p.113-128., 1995.
- 55- LILIENFELD S., PARYS W Galantamine: Additional benefits to patients with Alzheimer's disease. **Dementia and Geriatric Cognitive Disorders.**; v.1, p. 19-27. 2000.
- 56- GEERTS, H. Indicators of neuroprotection with galantamine **Brain Research Bulletin** v. 64, n.6, 2005, p. 519-524.
- 57- MIGLIACCIO-WALLE, K.; GETSIOS, D.; CARO, J.J.; ISHAK K.J., O'BRIEN JA, et al. Economic evaluation of galantamine in the treatment of mild to moderate Alzheimer's disease in the United States. **Clinical Therapeutics** v. 25: p.1806-1825, 2003.
- 58- PROCTOR, G.R.; HARVEY, A.L. Synthesis of Tacrine Analogues and Their Structure-Activity Relationships **Current Medicinal Chemistry**, v.7, p.295-302, 2000.
- 59- HENKE, C. J.; BURCHMORE M. J. The economic impact of tacrine in the treatment of Alzheimer's disease **Clinical Therapeutics** v.19, n. 2, p. 330-345, 1997.
- 60- CUMMINGS J. L,. Use of Cholinesterase Inhibitors in Clinical Practice. **American Journal of Geriatric Psychiatry** v.11, p.131-145, 2003.
- 61- RACCHI, M., MAZZUCHELLI, M., PORRELLO, E., et al. Acetylcholinesterase inhibitors: novel activities of old molecules **Pharmacological Research**. v .50, p.441-451, 2004.
- 62- SNAPE, M.F.; MISRA, A.; MURRAY, T.K.; DE SOUZA, R.J., et al A comparative study in rats of the in vitro and in vivo pharmacology of the acetylcholinesterase inhibitors tacrine, donepezil and NXX-066 **Original Research Neuropharmacology**, v. 38, n. 1, 1999, p. 181-193.
- 63- BRADLEY R. W.; ARLETTE N.; MARK A. G. A review of rivastigmine: A reversible cholinesterase inhibitor **Clinical Therapeutics**, v. 25, n. 6, p.1634-1653, 2003.
- 64- WEINSTOCK M. Selectivity of cholinesterase inhibition. Clinical implications for the treatment of Alzheimer's Disease. **CNS Drugs**. n. 12, p.307-323, 1999.

- 65- FINKEL, S. I. Effects of rivastigmine on behavioral and psychological symptoms of dementia in Alzheimer's disease **Clinical Therapeutics** v.26, n7, , p. 980-990, 2004.
- 66- FARLOW, M. R. & CUMMINGS, J. L. Effective pharmacologic management of Alzheimer's disease. **American Journal of Medicine**, v. 120, p. 388-397, 2007.
- 67- LUO. Y.; BOLON, B.; KAHN, S.; et al. Mice deficient in BACE1, the Alzheimer's β -secretase, have normal phenotype and abolished β -amyloid generation. **Nature Neuroscience** v.4 p.231–232, 2001
- 68- ROBERDS, S. L., ANDERSON. J. ; BASI, G.; et al. BACE knockout mice are healthy despite lacking the primary β -secretase activity in brain: implications for Alzheimer's disease therapeutics. **Human Molecular Genetics** v.10, p.1317–1324. 2001.
- 69- HARRISON, S.M.; HARPER, A.J.; HAWKINS J, et al. BACE1 (β -secretase) transgenic and knockout mice: identification of neurochemical deficits and behavioral changes. **Molecular and Cellular Neuroscience** v.24, p.646–655, 2003.
- 70- NGUYEN, J-T; HAMADA Y., KIMURA, T. KISO, Y.. Design of potent aspartic protease inhibitors to treat various diseases. **Archiv der Pharmazie** v. 341, n. 9, p.523–535, 2008.
- 71- EREZ, E.; FASS, D.; BIBI E. How intramembrane proteases bury hydrolytic reactions in the membrane. **Nature** v. 459, p. 371-378, 2009.
- 72- VERN L. S. Enzymatic transition states and transition state analog design **Annual Review of Biochemistry** v. 67 p.693-720, 1998.
- 73- GHOSH A.K.; SHIN, D.; Downs, B.; KOELSCH, G.; et al. Design of potent inhibitors for human brain memapsin 2 (β -secretase). **Journal American Chemical Society**, v.122 p.3522–3523, 2000.
- 74- STACHEL, S. J.; COBURN C. A.; STEELE, T. G.; JONES K. G., et al Structure based design of potent and selective cell-permeable inhibitors of human beta-secretase (BACE-1). **Journal Medicinal Chemistry** v.47, p.6447–6450, 2004.
- 75- STACHEL S. J.; COBURN, C.A.; STEELE, T.G.; CROUTHAMEL, M.C. et al Conformationally biased P3 amide replacements of β -secretase inhibitors. **Bioorganic & Medicinal Chemistry Letters** v. 16, p. 641–644, 2006.
- 76- FRESKOS J. N.; FOBIAN Y. M.; BENSON T.E.; BIENKOWSKI M.J., et al . Design of potent inhibitors of human β -secretase. Part 1. Bioorg **Bioorganic & Medicinal Chemistry Letters** v.17 p.73–77, 2007.

- 77- FRESKOS, J. N.; FOBIAN, Y. M.; BENSON, T. E.; MOON, J. B.; BIENKOWSKI, M. J. et al. Design of potent inhibitors of human β -secretase. Part 2. **Bioorganic & Medicinal Chemistry Letters** v.17 p. 78–81, 2007.
- 78- KIMURA T., SHUTO, D., HAMADA Y. et al. Design and synthesis of highly active Alzheimer's b-secretase (BACE1) inhibitors, KMI- 420 and KMI-429, with enhanced chemical stability. **Bioorganic & Medicinal Chemistry Letters**. v. 15, P. 211–215, 2005.
- 79- SILVESTRI R. Boom in the development of non-peptidic β -Secretase (BACE1) Inhibitors for the treatment of Alzheimer's disease. **Medicinal Research Reviews**, v. 29, n. 2, p. 295 -338, 2009.
- 80- MALAMAS M.S.; ERDEI J.; GUNAWAN, I. et al. Aminoimidazoles as potent and selective human beta-secretase (BACE1) inhibitors. **Journal of Medicinal Chemistry** v.52, n.20., p.6314-23, 2009.
- 81- MALAMAS M.S.; ERDEI J.; GUNAWAN, I.; et al. Design and synthesis of 5,5'-disubstituted aminohydantoinas as potent and selective human beta-secretase (BACE1) inhibitors. **Journal of Medicinal Chemistry**. v.53 n.3, p. 1146-1158, 2010.
- 82- MALAMAS M.S.; BARNES, K.; JOHNSON, M. et al. Di-substituted pyridinyl aminohydantoinas as potent and highly selective human β -secretase (BACE1) **Bioorganic & Medicinal Chemistry**, v.18 n.2 p. 630-639, 2010.
- 83- NOWAK P., COLE, D. C.; AULBAUGH, A. et al Discovery and initial optimization of 5,50-disubstituted aminohydantoinas as potent b-secretase (BACE1) inhibitors. **Bioorganic & Medicinal Chemistry Letters** v.20, p 632-635, 2010.
- 84- COLE, D.C.; MANAS, E.S., STOCK, J.R., et al. Acylguanidines as small-molecule beta -secretase inhibitors. **Journal Medicinal Chemistry**. v. 49, p. 6158-61, 2006.
- 85- MALAMAS M.S.; BARNES, K. HUI Y. et al Novel pyrrolyl 2-aminopyridines as potent and selective human β -secretase (BACE1) inhibitors **Bioorganic & Medicinal chemistry Letters** v. 20, n.7 p. 2068-2073, 2010.
- 86- SANT'ANNA C. M. R. Métodos de modelagem molecular para estudo e planejamento de compostos bioativos: uma introdução **Revista Virtual de Quimica**, v.1, n.1, p. 49-57, 2009.
- 87- BARREIRO, E. J.. Estratégia de simplificação molecular no planejamento racional de fármacos: a descoberta de novo agente cardioativo. **Química Nova**, v. 25, n. 6, 2002
- 88- GUIDO, R. V. C.; ANDRICOPULO, A. D.; OLIVA, G. Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas. **Estudos Avançados** v. 24, n. 70, 2010.

- 89- GUIDO, R. V. C.; ANDRICOPULO, A. D. Modelagem molecular de fármacos. **Revista Processos Químicos**, v.2, n.4, p.24-26, 2008.
- 90- ANDRICOPULO A. D.; SALUM, L.B.; ABRAHAM, D.J. Structure-based drug design strategies in medicinal chemistry. **Current Topics in Medicinal Chemistry** v.9, n.9, p.771-90, 2009.
- 91- ESPOSITO, E.X.; HOPFINGER, A.J.; MADURA, J.D. Methods for applying the quantitative structure-activity relationship paradigm. **Methods in Molecular Biology** v. 275, p.131–213, 2004.
- 92- WINKLER, D. A. The role of Quatitative Structure Activity relationships (QSAR) in Bimolecular discovery. **Briefings in Bioinformatics** v.3, n.1 73-86, 2002.
- 93- PERKINS, R.; FANG, H.; TONG, W.; WELSH, W. Quantitative structure-activity relationship methods: perspectives on drug discovery and toxicology. **Environmental Toxicology and Chemistry**, v. 22, p.1666–1679, 2003.
- 94- CONSONNI V., TODESCHINI R., PAVAN M. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. **Journal of Chemical Information and Computer Sciences.**; v.42, n.3, p.682-92, 2002.
- 95- TODESCHINI, R.; CONSONNI, V.; MAURI A.; PAVAN, M. DRAGON 5.5 for Windows, 2008, Talete srl, Milan, Italy.
- 96- GONZALEZ, M. P.; HELGUERA, A. M. TOPS-MODE *versus* DRAGON descriptors to predict permeability coefficients through low-density polyethylene **Journal of Computer-Aided Molecular Design.**, v.17, n.10, p. 665-78, 2003.
- 97- KATRITZKY, A.R; LOBNOV, V; KARELSON, M. CODESSA (*Comprehensive Descriptors for Structural and Statistical Analysis*) University of Florida: Gainesville, FL, 1994;
- 98- SENA, M. M. de; POPPI, Ronei J.; FRIGHETTO, R. T. S.; VALARINI, P. J.. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova** , v.23, n.4, p. 547-556, 2000.
- 99- FERREIRA, M. M. C.; MONTANARI, C. A.; GAUDIO, A. C.. Seleção de variáveis em QSAR. **Química Nova**, v. 25, n. 3, p. 349-448, 2002 .
- 100- TODESCHINI, R, CONSONNI, V, MAURI, A, PAVAN, M. MobyDigs: software for regression and classification models by genetic algorithms, in: ‘Nature-inspired Methods in Chemometrics: Genetic Algorithms and Artificial Neural Networks’. (Leardi R., Ed.), Chapter 5, Elsevier, 2004 p 141-67.
- 101- JUDSON, R. Genetic algorithms and their use in chemistry. **Reviews in Computational Chemistry**.v. 10, p.1–66, 1997.

- 102- WOLD, S.; ESBENSEN, K.; GELADI, P. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems.*; v.2 n. 1-3, p.37-52. 1987.
- 103- BORGES, E. C. L.; MOZETO, A. A.; NEVES, E. F. A. et al. Estudo da capacidade de complexação e sua relação com algumas variáveis ambientais em cinco represas do Rio Tietê/Brasil. **Química Nova**, v.30, n.7, p. 1505-1511, 2007.
- 104- FERREIRA, M. M. C., ANTUNES, A. M., MELGO, M. S., VOLPE, P. L. O. “Quimiometria I: calibração multivariada, um tutorial”, **Química Nova** v. 22, n. 5, pp. 724-731, 1999.
- 105- NETO, J. M. M.; MOITA, G. C. Uma introdução a análise exploratória de dados multivariados. **Química Nova**, v. 21, n. 2, p. 467-469, 1998.
- 106- MOTA, S.G.R. Ensaio *in vitro* e análise quimiométrica de inibidores da enzima lanosterol 14 α -desmetilase de *moniliophthora perniciosa*. (Mestrado em Biotecnologia). Instituto de química,UEFS, Feira de Santana, 2009. 79p.
- 107- CRAMER, R. D. Partial least squares (PLS): its strength and limitations **Perspectives in Drug Discovery and Design**, v. 1 , p.269-278, 1993.
- 108- KUBINYI, H. Variable Selection in QSAR Studies. II. A Highly Efficient Combination of Systematic Search and Evolution. **Quantitative structure-activity relationships**. p. 393-401, 1994.
- 109- CONSONNI, V.; BALLABIO, D.; TODESCHINI, R. Comments on the Definition of the Q² Parameter for QSAR. **Validation Journal of Chemical Information and Modeling** v.49 , n.7, p.1669-1678, 2009.
- 110- ROY, P. P. ; LEONARD, J. T.; ROY, K. Exploring the impact of size of training sets for the development of predictive QSAR models. **Chemometrics and Intelligent Laboratory Systems**, v. 90, n.1, p. 31-42, 2008.
- 111- LOWIS, D. R. HQSAR: A New, Highly Predictive QSAR Technique. Tripos Technical Notes, 1, 1997.
- 112- DODDAREDDY, M.; LEE, Y. Hologram Quantitative Structure Activity Relationship Studies on 5-HT₆ Antagonists. **Bioorganic & Medicinal Chemistry.**, v. 12, p. 3815-3824, 2004.
- 113- CHEN, Da; YIN, Chunsheng; WANG, Xiaodong; WANG, Liansheng. Holographic QSAR of selected esters **Chemosphere** v. 57, n.11, p.1739–1745, 2004.
- 114- HQSARTM Manual, Tripos Inc., St. Louis, 2006.
- 115- WANG, X.; TANG, S.; LIU, S.; CUI, S.; WANG, L. Molecular hologram derived quantitative structure-property relationships to predict physico-chemical

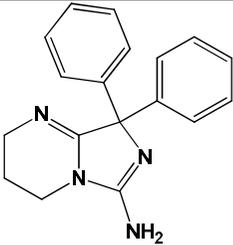
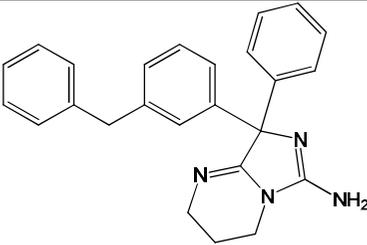
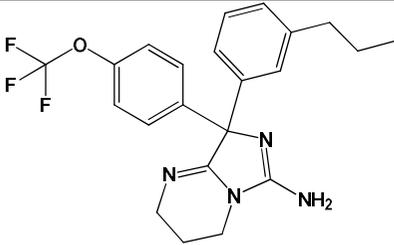
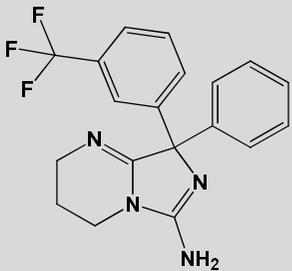
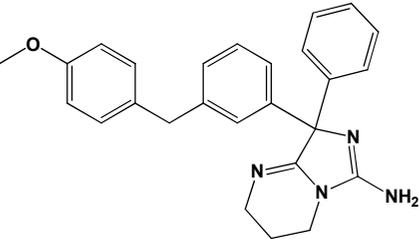
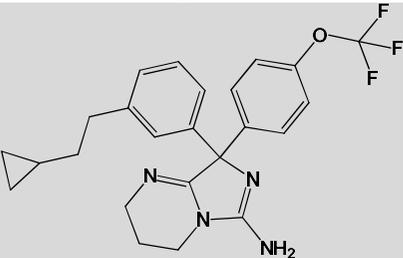
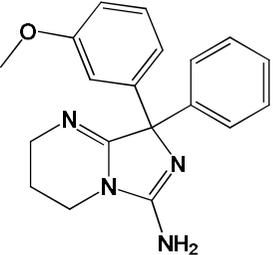
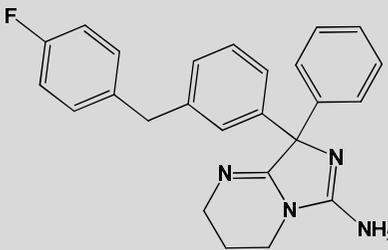
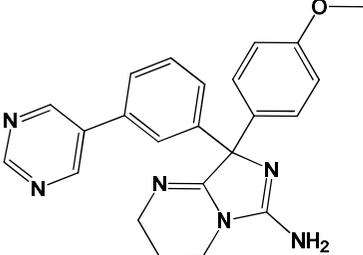
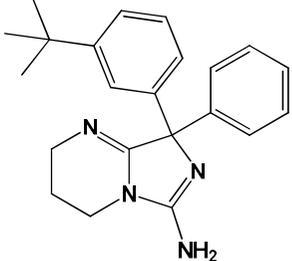
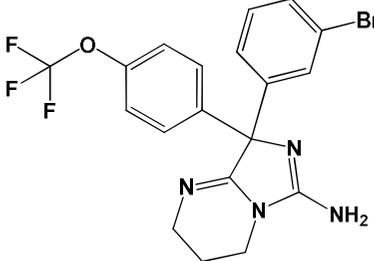
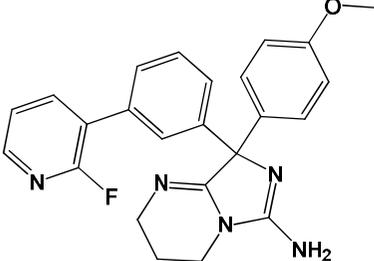
- properties of polychlorinated biphenyls. **Chemosphere**. v.51, n.7, p.617-32. 2003.
- 116- CRAMER III, R. D.; PATTERSON, D. E.; BUNCE, J. D. Comparative molecular field analysis (CoMFA) 1. Effect of shape on binding of steroids to carrier proteins. **Journal American Chemical Society**, v. 110, p. 5959–5967. 1988.
- 117- KUBINYI, H. QSAR and 3D-QSAR in drug design. Part 1: methodology. **Drug Discovery Today**, v.11, p.457-467, 1997.
- 118- LANGER, T.; BRYANT, S.D. 3D Quantitative Structure-Property Relationships. In: WERMUTH, C.G., ed. *The Practice of Medicinal Chemistry*. 3rd ed. Oxford: Elsevier, 2008. p. 587-604.
- 119- BOHM, M.; STURZEBECHER, J.; KLEBE, G.. Three-dimensional quantitative structure–activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. **Journal of Medicinal Chemistry**. v. 42, p. 458–477, 1999.
- 120- PETERSON, S.D.; SCHAAL, W.; KARLE'N, A. Improved CoMFA modeling by optimization of settings. **Journal of Chemical Information and Modeling**, v. 46, p. 355-364, 2006.
- 121- KUBINYI, H. **3D QSAR in drug design: theory, methods and applications**. Leiden: ESCOM, 1993.
- 122- CHO, S. J.; GARSIA, M. L. S.; BIER, J.; TROPSHA , A. Structure-based alignment and comparative molecular field analysis of acetylcholinesterase inhibitors. **Journal of Medicinal Chemistry**, v. 39, p. 5064-5071, 1996.
- 123- CHAE, C. H.; GWEON O. D.; SHIN, W. Flexible Molecular Superposition: Development of a Combined Similarity Index and Application of the Constrained Optimization Technique **Journal of Computational Chemistry**, v. 22, n. 8, p.888–900, 2001.
- 124- SALUM, L. B. ; DIAS, L. C. ; ANDRICOPULO, A. D. . Structural and chemical basis for anticancer activity of a series of -tubulin ligands: molecular modeling and 3D QSAR studies. **Journal of the Brazilian Chemical Society**, v.20, p. 693-703, 2009.
- 125- CLEVES , A. E.; JAIN, A .N Robust ligand-based modeling of the biological of known drugs, **Journal of Medicinal Chemistry**, p. 49, p.2921-38, 2006.
- 126- JAIN, A .N. Ligand-based structural hypotheses for virtual screening **Journal of Medicinal Chemistry**, v. 47, p 947-961, 2004.

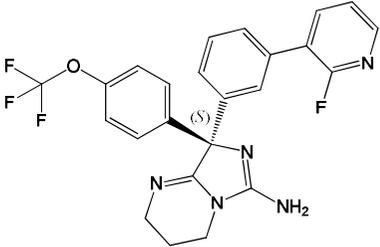
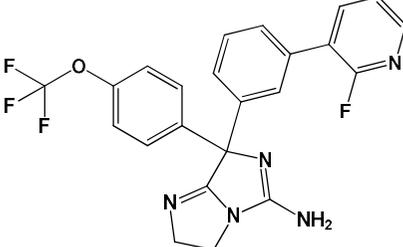
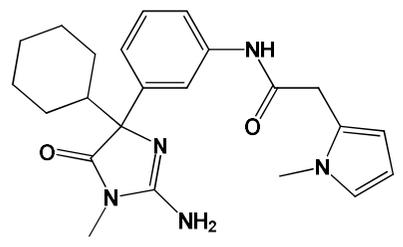
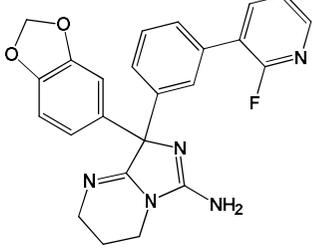
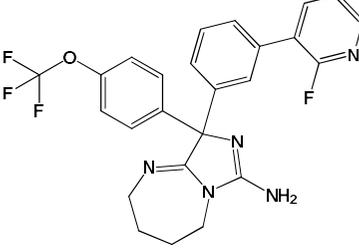
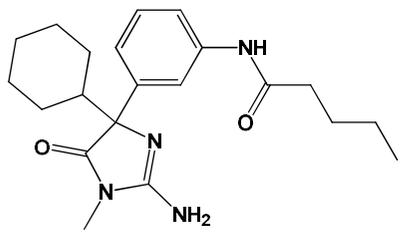
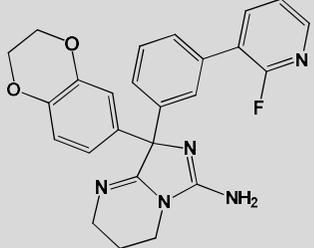
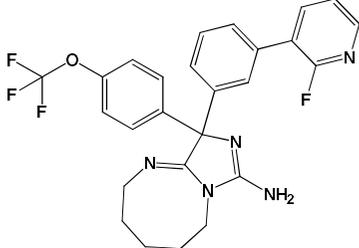
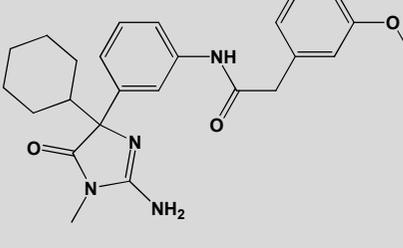
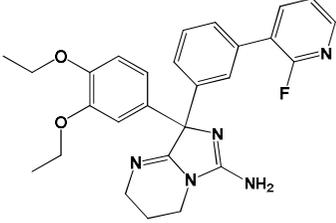
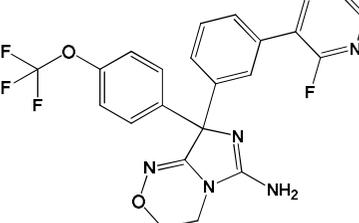
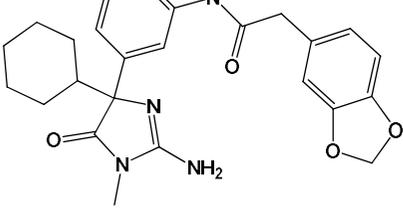
- 127- WELCH, W.; RUPPERT, J.; JAIN, A .N. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. **Chemistry & Biology** v.3 n.6 p. 449-462, 1996.
- 128- KITCHEN, D. B.; DECORNEZ, H.; FURR, J. R.; BAJORATH, J., Docking and scoring in virtual screening for drug discovery: Methods and applications. **Nature Reviews Drug Discovery**, v.3, n.11, p.935-949, 2004.
- 129- BRODMEIER, T; PRETSCH, E. Application of genetic algorithms in molecular modelling. **Journal of Computational Chemistry**. v.15, p.588-595, 1994.
- 130- JONES, G.; WILLETT, P.; GLEN, R. C. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. **Journal of Computer-Aided Molecular Design**. v. 9, p.532 -549, 1995.
- 131- JONES, G.; WILLETT, P.; GLEN, R. C.; LEACH, A. R.; TAYLOR, R. Development and validation of a genetic algorithm for flexible docking. **Journal of Molecular Biology**, v. 267 p.727-748, 1997.
- 132- VERDONK, M. L.; COLE, J. C.; HARTSHORN, M. J.; MURRAY, C. W.; TAYLOR, R. D. Improved protein-ligand docking using GOLD. **Proteins-Structure Function and Genetics**, v. 52, p. 609-623, 2003.
- 133- TITMUSS, S. J.; KELLER, P. A., GRIFFITH, R. Docking experiments in the flexible non-nucleoside inhibitor binding pocket of HIV-1 reverse transcriptase **Bioorganic & Medicinal Chemistry** v.7, n. 6, p.1163-1170, 1999.
- 134- FREITAS, H. F.; PAZ, O. S.; CASTILHO, M. S.. Estudos de QSAR 3D para um conjunto de inibidores de butirilcolinesterase humana. **Química Nova**, v. 32, n. 8, p.2114-2121, 2009.
- 135- CASTILHO, C.; GUIDO, R. V.; ANDRICOPULO, A.D. Classical and Hologram QSAR Studies on a Series of Tacrine Derivatives as Butyrylcholinesterase Inhibitors. **Letters in Drug Design & Discovery**, v. 4, p. 106-113, 2007.
- 136- RECANATINI, M.; CAVALLI, A.; HANSCH, C.A. Comparative QSAR analysis of Acetylcholinesterase Inhibitors currently studied for the Treatment of Alzheimer's disease. **Chemico-biological Interactions** v.105, p.199-228, 1997.
- 137- ZUO, Z.; LUO, X; ZHU, W.; SHEN, J. et al. Molecular docking and 3D-QSAR studies on the binding mechanism of statine-based peptidomimetics with b-secretase. **Bioorganic & Medicinal Chemistry**, v. 13, p.2121-2131, 2005.
- 138- PANDEY, A., MUNGALPARA, J., MOHAN, C. G. Comparative molecular field analysis and comparative molecular similarity indices analysis of hydroxyethylamine derivatives as selective human BACE-1 inhibitor. **Molecular Diversity**. v.14, n.1, p. 39-49, 2010.

- 139- COBURN C. A.; STACHEL S. J.; LI Y. M.; RUSH D. M., et al. Identification of a small molecule nonpeptide active site beta-secretase inhibitor that displays a nontraditional binding mode for aspartyl proteases. **Journal of Medicinal Chemistry** v.47, n. 25, p. 6117-9. 2004.
- 140- FERREIRA, M. M. C. Multivariate QSAR. **Journal of Brazilian Chemical Society**, v. 13, n. 6, p. 742-753, 2002.
- 141- ALMEIDA, V. L.; LOPES, J. C. D.; OLIVEIRA, S. R.; DONNICI, C. L.; MONTANARI, C. A. Estudos de relações estrutura-atividade quantitativas (QSAR) de Bis-benzamidas com atividade antifúngica **Química Nova**, v. 33, n.7, p. 1482-1489, 2010.
- 142- LILL, M. A. Multi-dimensional QSAR in drug discovery. **Drug Discovery Today**, v. 12, n. 23-24, p. 1013-1017, 2007.
- 143- SALUM, L. B.; ANDRICOPULO, A. D. Fragment-based QSAR: perspectives in drug design. **Molecular Diversity** v. 13, n.3, p.277-285, 2009.
- 144- GOLBRAIKH, A.; TROPSHA A. Beware of q²! **Journal of Molecular Graphics and Modelling** , v. 20, n. 4, p.269-276, 2002
- 145- GOLBRAIKH, A.; TROPSHA, A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. **Journal of Computer-Aided Molecular Design** v. 16, n. 5-6, p. 357-69, 2002.
- 146- CHIRICO, N.; GRAMATICA P. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. **Journal of Chemical Information and Modeling**, v. 51, n. 9, p. 2320–2335, 2011.
- 147- GOZALBES, R.; DOUCET J. P.; DEROUIN, F. Application of topological descriptors in QSAR and drug design: history and new trends. **Current Drug Targets - Infectious Disorders**, v.2, n. 1 p.93-102, 2002.
- 148- ROY, K. Topological descriptors in drug design and modeling studies. **Molecular Diversity**, v. 8, p. 321–323, 2004.
- 149- CABALLERO, J.; GARRIGA, M.; FERNANDÉZ, M. Genetic neural network modeling of the selective inhibition of the intermediate conductance Ca²⁺ activated K⁺ channel by some triarylmethanes using topological charge indexes descriptors. **Journal of Computer-Aided Molecular Design** v.19, p. 771–789, 2005.
- 150- FERNANDES, M. B.; SCOTTI, M. T.; FERREIRA, M. J. P.; EMERENCIANO, V. P. Use of selforganizing maps and molecular descriptors to predict the cytotoxic activity of sesquiterpene lactones. **European Journal of Medicinal Chemistry**, v.43, n.10, p.2197- 205, 2008.

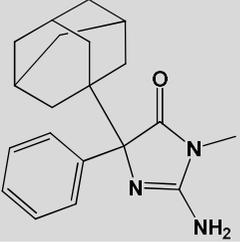
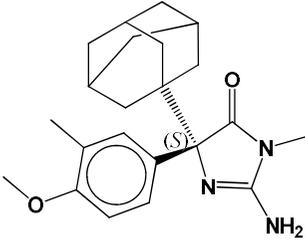
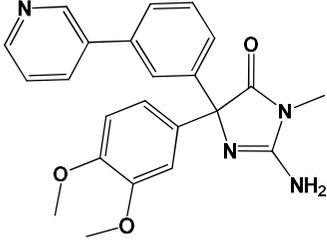
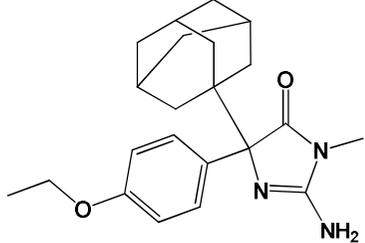
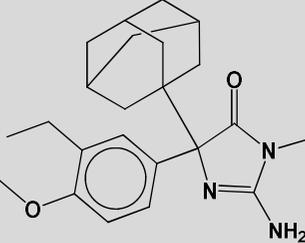
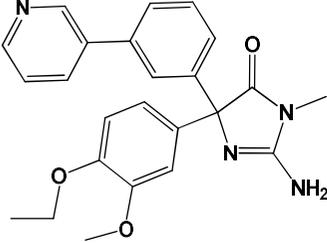
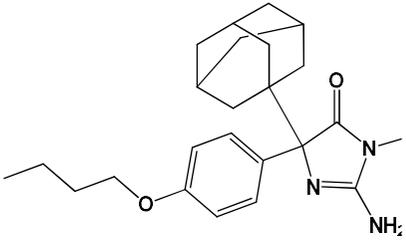
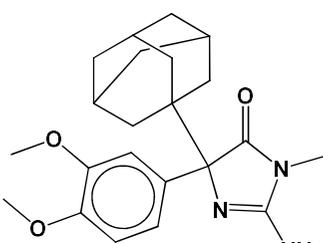
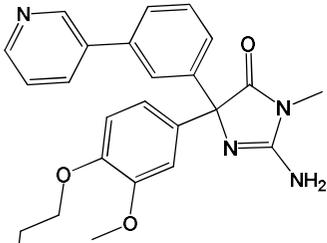
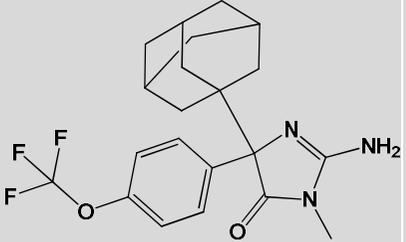
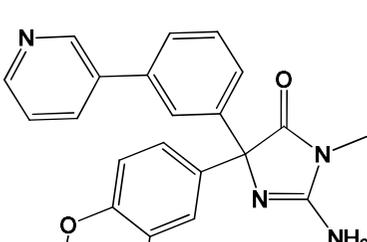
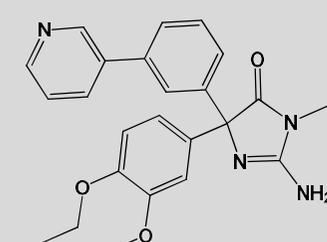
- 151- TODESCHINI, R.; CONSONNI, V.; MAURI, A., PAVAN, M. Detecting “bad” regression models: multicriteria fitness functions in regression analysis. **Analytica Chimica Acta** v. 515 p.199–208, 2004.
- 152- GAUDIO, A. C.; ZANDONADE, E. Proposição, validação e análise dos modelos que correlacionam estrutura química e atividade biológica. **Química Nova**, v. 24, p. 658-671. 2001.
- 153- RODIGUES, C. R., CASTRO, H. C. BRITO, M. A., et al, Métodos de Correlação Quantitativa entre Estrutura Molecular Tridimensional e Atividade Farmacológica (3D-QSAR) em: MONTANARI, CARLOS A. Química Medicinal: Métodos e Fundamentos em planejamento de Fármacos. São Paulo, Edusp, cap. 13, 2011, p. 455-484.
- 154- HONG, L.; KOELSCH, G.; LIN, X.; et al. Structure of the protease domain of memapsin 2 (β -secretase) complexed with inhibitor. **Science**, v. 290, p.150 – 153, 2000.

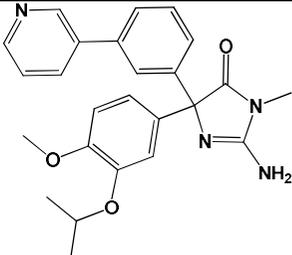
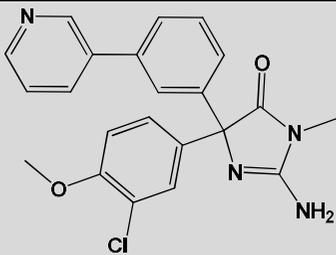
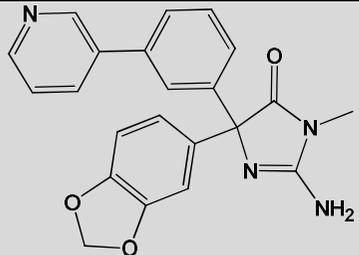
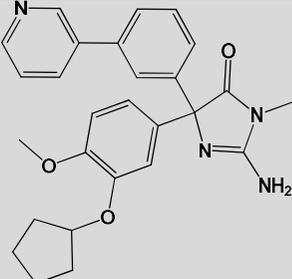
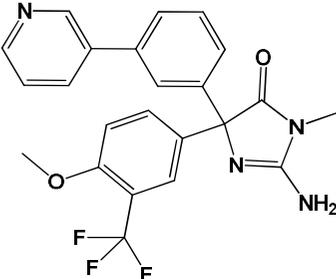
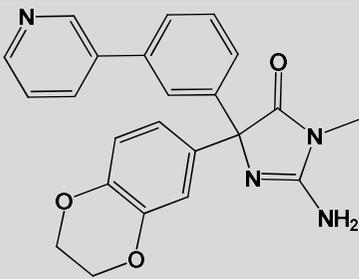
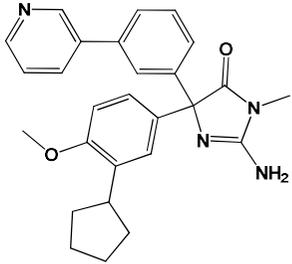
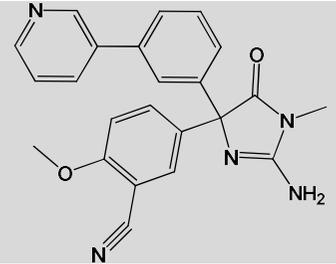
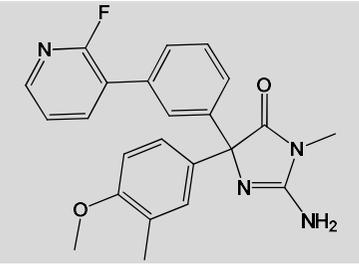
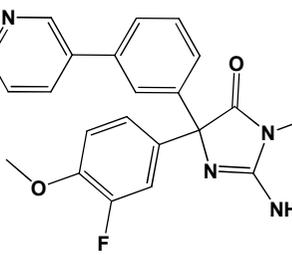
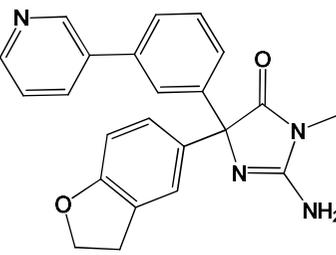
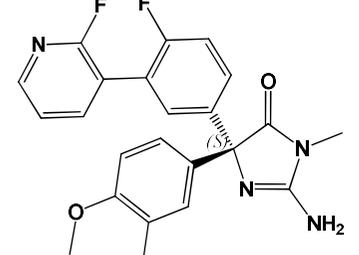
Apêndice A- Estrutura química e valores de pIC_{50} do conjunto de inibidores de bace-1 utilizados para os estudos de QSAR 2D e QSAR 3D.
(As moléculas do conjunto teste são estão indicadas por marcadores de fundo).

Comp	Estrutura	pIC_{50}	Comp	Estrutura	pIC_{50}	Comp	Estrutura	pIC_{50}
01		4,42	05		5,28	09		5,84
02		4,63	06		5,15	10		6,68
03		4,69	07		5,09	11		7,00
04		4,52	08		5,59	12		7,15

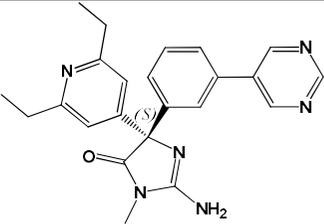
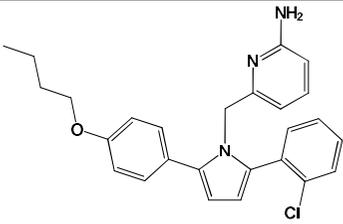
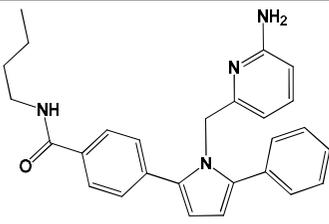
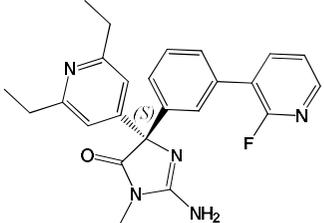
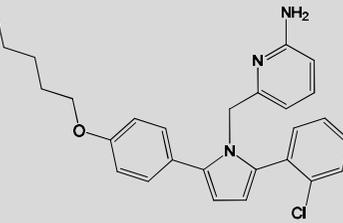
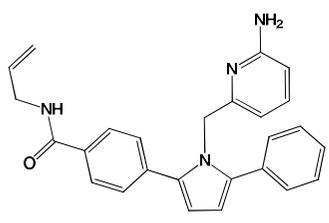
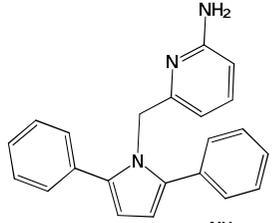
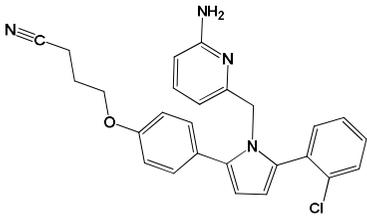
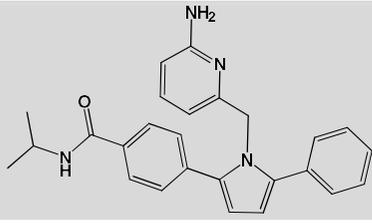
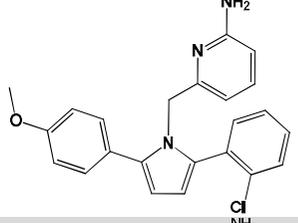
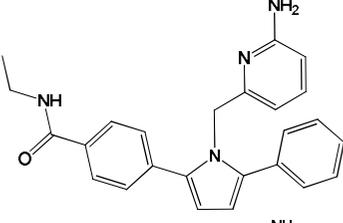
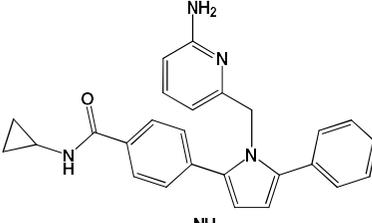
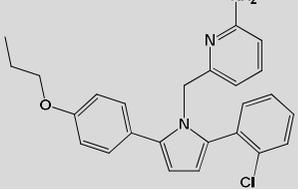
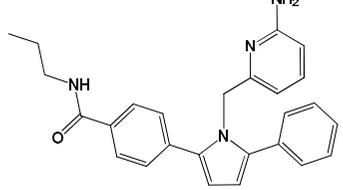
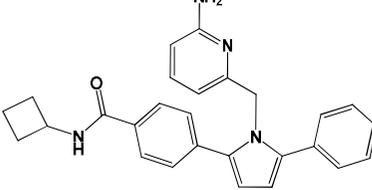
Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀
13		6,43	17		7,10	21		6,07
14		6,80	18		7,10	22		5,55
15		6,72	19		6,89	23		5,87
16		7,10	20		7,52	24		6,42

Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀
25		5,86	29		5,97	33		5,46
26		6,02	30		5,98	34		5,60
27		6,41	31		5,30	35		5,51
28		5,91	32		5,46	36		5,65

Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀
37		5,82	41		6,77	45		7,05
38		6,48	42		6,55	46		7,30
39		6,06	43		6,57	47		7,30
40		6,33	44		7,40	48		7,30

Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀
49		7,22	53		7,44	57		6,60
50		7,15	54		7,40	58		7,05
51		6,92	55		6,82	59		7,70
52		7,30	56		6,92	60		8,00

Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀
61		7,22	65		5.57	69		7,52
62		7,52	66		7.22	70		7,15
63		7,70	67		7.00	71		6,85
64		7,70	68		7.10	72		6,82

Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀
73		7,52	78		5,78	83		5,52
74		8,00	79		6,26	84		6,23
75		5,28	80		6,38	85		5,75
76		5,61	81		5,65	86		5,92
77		5,94	82		5,52	87		5,47

Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀	Comp	Estrutura	pIC ₅₀
88		6,07	93		6,34	98		6,00
89		6,23	94		6,39	99		6,77
90		6,39	95		6,54	100		6,66
91		6,09	96		6,01	101		7,40
92		7,00	97		6,96	102		7,15

Apêndice B - Valores de pIC₅₀ preditos e valores residuais para os 27 compostos do conjunto teste de acordo com os melhores modelos de CoMFA.

Conjunto Teste		Modelo A		Modelo B		Modelo C		Modelo D		Modelo E	
Comp.	pIC ₅₀	Pred. ^a	Res. ^b								
02	4,66	5,25	0,59	5,75	1,09	5,55	0,89	5,04	0,38	4,88	0,22
07	5,09	5,16	0,07	5,52	0,43	5,07	-0,02	4,98	-0,11	5,07	-0,02
10	6,68	5,5	-1,18	6,14	-0,54	6,16	-0,52	6,79	0,11	5,92	-0,76
15	6,72	6,17	-0,55	6,50	-0,22	6,95	0,23	6,34	-0,38	6,44	-0,28
23	5,87	6,22	0,35	5,82	-0,05	5,42	-0,45	5,22	-0,65	5,42	-0,45
25	5,86	5,81	-0,05	5,01	-0,85	6,11	0,25	5,81	-0,05	6,15	0,29
32	5,46	5,23	-0,23	6,14	0,68	4,85	-0,61	5,81	0,35	5,31	-0,15
35	5,51	5,73	0,22	5,94	0,43	5,78	0,27	5,63	0,12	5,31	-0,20
37	5,82	5,43	-0,39	5,15	-0,67	6,00	0,18	6,01	0,19	6,21	0,39
40	6,33	6,51	0,18	6,44	0,11	6,44	0,11	6,16	-0,17	6,59	0,26
42	6,55	6,98	0,43	6,52	-0,03	6,62	0,07	6,78	0,23	6,87	0,32
48	7,30	7,46	0,16	7,11	-0,19	7,16	-0,14	7,72	0,42	7,33	0,03
50	7,15	7,19	0,04	7,34	0,19	7,51	0,36	7,15	0,00	7,44	0,29
53	7,44	7,36	-0,08	7,32	-0,12	7,25	-0,19	7,57	0,13	7,54	0,10
55	6,82	7,33	0,51	7,02	0,2	7,28	0,46	7,02	0,20	7,55	0,73
58	7,05	6,95	-0,10	6,67	-0,38	7,33	0,28	6,93	-0,12	7,32	0,27
57	6,6	6,47	-0,13	6,65	0,05	7,11	0,51	7,07	0,47	7,12	0,52
59	7,70	7,53	-0,17	7,60	-0,1	7,38	-0,32	7,29	-0,41	7,67	-0,03
67	7,00	7,32	0,32	6,89	-0,11	6,79	-0,21	7,04	0,04	6,74	-0,26
71	6,85	7,41	0,56	7,02	0,17	7,34	0,49	7,04	0,19	6,64	-0,21
77	5,94	6,03	0,09	6,70	0,76	5,88	-0,06	6,00	0,06	6,00	0,06
79	6,26	6,06	-0,20	6,44	0,18	6,20	-0,06	6,16	-0,10	6,27	0,01
85	5,75	5,76	0,01	5,97	0,22	5,98	0,23	5,60	-0,15	5,81	0,06
90	6,39	6,86	0,47	6,49	0,1	6,14	-0,25	6,57	0,18	6,73	0,34
91	6,09	6,09	0,00	6,59	0,5	6,13	0,04	6,46	0,37	6,55	0,46
99	6,77	6,29	-0,48	6,14	-0,63	6,64	-0,13	6,83	0,06	6,24	-0,53
102	7,15	7,29	0,14	7,09	-0,06	6,88	-0,27	6,87	-0,28	6,61	-0,54

^a Valores de pIC₅₀ preditos / ^b diferença entre valor observado experimentalmente e predito pelo modelo.