

**ESTUDOS QUANTITATIVOS
MULTIDIMENSIONAIS DE CORRELAÇÃO
ESTRUTURA-ATIVIDADE DE INIBIDORES
PEPTÍDICOS DA HIV-1 PROTEASE**

THAIS CRISTINA SILVA DE SOUZA

2009

THAIS CRISTINA SILVA DE SOUZA

**ESTUDOS QUANTITATIVOS MULTIDIMENSIONAIS DE
CORRELAÇÃO ESTRUTURA-ATIVIDADE DE INIBIDORES
PEPTÍDICOS DA HIV-1 PROTEASE**

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

Orientadora

Prof^a. Elaine Fontes Ferreira da Cunha

LAVRAS
MINAS GERAIS - BRASIL
2009

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

Souza, Thaís Cristina Silva de.

Estudos quantitativos multidimensionais de correlação
estrutura-atividade de inibidores peptídicos da HIV-1 protease
/ Thaís Cristina Silva de Souza. – Lavras : UFLA, 2009.

100 p. : il.

Dissertação (Mestrado) – Universidade Federal de Lavras, 2009.

Orientador: Elaine Fontes Ferreira da Cunha.

Bibliografia.

1. Aids. 2. HIV-1. 3. Protease. 4. QSAR-5/6D. 5. Inibidores
peptídicos. I. Universidade Federal de Lavras. II. Título.

CDD – 616.9792

THAIS CRISTINA SILVA DE SOUZA

**ESTUDOS QUANTITATIVOS MULTIDIMENSIONAIS DE
CORRELAÇÃO ESTRUTURA-ATIVIDADE DE INIBIDORES
PEPTÍDICOS DA HIV-1 PROTEASE**

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

APROVADA em 09 de Julho de 2009

Prof. Matheus Puggina de Freitas

UFLA

Prof. Luciano Vilela Paiva

UFLA

Profa. Elaine Fontes Ferreira da Cunha
UFLA
(Orientadora)

LAVRAS
MINAS GERAIS - BRASIL

Esta dissertação foi realizada sob a orientação da Prof^ª. Elaine Fontes Ferreira da Cunha, com o apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) no Laboratório de Química Computacional da Universidade Federal de Lavras.

DEDICATÓRIA

Dedico este trabalho aos meus pais, Hermenegildo e Maria de Fátima, os quais acompanharam de perto meus passos em todos esses anos de caminhada, sempre tentando guiar-me da melhor forma possível, pelos esforços, dedicação e compreensão que me foram dispensados e pelo amor incondicional e recíproco que me foram dedicados; ao meu namorado, Victor, pela sua amizade, companhia, dedicação, imensa paciência, incentivo e amor, que me deram forças para terminar essa etapa de minha vida; às minhas amigas e companheiras, que estiveram presentes e contribuíram, cada uma a sua maneira, para minimizar tristezas e aumentar alegrias.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me guiar e me proteger em todos os dias de minha vida, iluminando-me para sempre seguir o melhor caminho.

Aos meus pais, Hermenegildo José de Souza e Maria de Fátima Silva de Souza, pelo amor incondicional, paciência, amizade, compreensão, ao apoio sentimental e financeiro que, apesar de todas as dificuldades, sempre fizeram de tudo para que eu tivesse o melhor, e ao meu irmão Thiago, por todo o apoio nesses anos.

A minha família, em especial a meus avós Nair e Hermínio, minha tia Maria Olívia e sua filha Aryadne, por toda a colaboração e apoio nesses anos de estudo.

Aos meus orientadores, Elaine e Teodorico pela amizade, paciência, compreensão, apoio, carinho e confiança depositada em mim durante este trabalho e toda a minha estadia no grupo.

Ao Victor pela presença, apoio, confiança, dedicação, amor e carinho em todos os maravilhosos momentos que temos passado juntos.

À Roberta, minha fiel amiga, companheira, confidente, que sempre esteve presente em todos os momentos, de alegria, das risadas, das tristezas e dos sucessos e fracassos compartilhados nesses anos.

Aos colegas de turma, Milene Aparecida Andrade, Emanuella Maria Barreto Fonseca, Marcus Vinícius Juliaci Rocha, Jeancarlo Pereira dos Anjos e companheiros do Laboratório de Química Computacional, e amigos, os quais estiveram junto a mim em momentos felizes e especiais, momentos de conquistas, de crescimento, mas estavam presentes também em momentos tristes, de impaciência, onde a amizade foi essencial.

Aos professores Matheus Puggina de Freitas e Zuy Maria Magriotis, por toda amizade e confiança e pelos longos bate papos em todos esses anos de convivência.

SUMÁRIO

	Página
LISTA DE TABELAS.....	i
LISTA DE FIGURAS.....	ii
RESUMO.....	vii
ABSTRACT	viii
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	2
2.1 Aids (HIV).....	2
2.1.1 Síndrome da imunodeficiência adquirida	2
2.1.2 Vírus	6
2.1.3 O Vírus HIV	9
2.1.4 Ciclo viral	13
2.1.5 Estrutura da HIV-protease	17
2.1.6 Inibidores da HIV-1 protease.....	21
2.2 Inibidores peptídicos com simetria C-2	22
2.2.1 Grupo diol vicinal	22
2.3 QSAR.....	27
2.3.1 História do desenvolvimento do QSAR.....	27
2.3.2 Relações quantitativas entre estrutura e atividade/propriedade (QSAR/QSPR).....	28
2.4 Métodos e índices estatísticos.....	29
2.4.1 Coeficiente de correlação linear (r).....	29
2.4.2 Coeficiente de correlação linear quadrático (r^2 ou R^2).....	30
2.4.3 Análise de regressão pelo método dos mínimos quadrados parciais (PLS)	30
2.4.4 Validação cruzada.....	31

2.4.5 Índice de falta de ajuste	32
2.4.6 Validação externa dos modelos	32
2.4.7 Teste de <i>Scramble</i>	32
3 OBJETIVOS	34
4 MATERIAL E MÉTODOS	36
4.1 Dados biológicos.....	36
4.2 QSAR-4D	46
4.2.1 Módulo <i>Symposar</i>	46
4.2.2 QSAR-5D	49
4.2.3 QSAR-5D – adição da quinta dimensão	49
4.2.4 O Conceito raptor.....	50
4.2.5 Score function	51
4.2.6 Algoritmo de otimização	53
4.3 QSAR-6D	54
4.3.1 O conceito Quasar.....	54
4.3.2 Geração de uma família inicial de estruturas parentais.....	55
4.3.3 Estimativa da energia livre relativa do sítio de ligação	57
4.3.4 Análise da família modelo	59
5 RESULTADOS E DISCUSSÃO.....	60
5.1 QSAR-4D – módulo <i>symposar</i>	60
5.2 QSAR-5D	61
5.2.1 Análise dos modelos de QSAR-5D.....	61
5.3 QSAR-6D	74
5.3.1 Análise dos modelos de QSAR-6D.....	74
6 CONCLUSÕES	84
REFERÊNCIAS BIBLIOGRÁFICAS	86
ANEXOS	97

LISTA DE TABELAS

	Página
TABELA 1	Dados estimados sobre os casos de HIV no mundo e no Brasil..... 05
TABELA 2	Resumo das funções das proteínas codificadas pelos nove genes do HIV..... 11
TABELA 3	Estrutura química dos análogos diol baseada na simetria C-2..... 23
TABELA 4	Estrutura dos 46 inibidores peptídicos com carbono vicinal central da HIV-1 protease com os respectivos valores de K_i (nM)..... 38
TABELA 5	Propriedades Atomísticas Mapeadas para a superfície do Receptor no Quasar..... 55
TABELA 6	Valores de pK_i experimental e predito para o conjunto de treinamento e conjunto de testes obtidos no QSAR-5D.... 62
TABELA 7	Detalhes entálpicos, entrópicos, solvatação, e contribuições do encaixe-induzido para o cálculo da afinidade do sítio de ligação..... 69
TABELA 8	Coeficientes das equações geradas para os modelos de QSAR-5D pelo software Raptor..... 71
TABELA 9	Valores de pK_i experimental e predito para o conjunto de treinamento obtidos no QSAR-6D..... 79

LISTA DE FIGURAS

	Página
FIGURA 1	Visão global da infecção por HIV, no ano de 2007, segundo o último informe da epidemia mundial de HIV em 2008, publicado pela ONUSIDA (2009)..... 03
FIGURA 2	Visão da America Latina da infecção por HIV, no ano de 2007, segundo o último informe da epidemia mundial de HIV em 2008, publicado pela ONUSIDA..... 04
FIGURA 3	Estrutura esquemática do vírus do HIV..... 08
FIGURA 4	Desenho do vírion do HIV maduro, circundado pela representação em fitas das proteínas virais e dos fragmentos de proteína, caracterizados estruturalmente. As estruturas das proteínas foram representadas na mesma escala em que SU= glicoproteínas de superfície, TM=proteínas da transmembrana, MA=matriz protéica, CA= proteína do capsídio, NC= proteína do nucleocapsídio, PR=protease, RT=transcriptase reversa, IN=integrase, Nef=proteínas acessórias..... 12
FIGURA 5	Ciclo de vida do HIV. Os textos em vermelho se referem aos alvos inibidos por agentes existentes aprovados pela FDA. Os textos em azul referem-se aos alvos que são inibidos por agentes que poderão estar disponíveis no futuro..... 16
FIGURA 6	Representação esquemática do sítio catalítico da HIV-Pr associado com um substrato polipeptídico. S e S' identificam os subsítios (resíduos de aminoácidos) da

	enzima e P e P' identificam os subsítios (resíduos de aminoácidos) do substrato polipeptídico. O polipeptídeo representado é Pro-Ala-Val-Ser-Leu-Met-Thr. O círculo em negrito representa o local de clivagem	18
FIGURA 7	Mudança conformacional provocada pela presença do inibidor no sítio ativo. (A) HIV-1 Protease em sua forma livre e em (B) o complexo com o inibidor. A ligação do inibidor induz uma mudança conformacional na qual os <i>flaps</i> se fecham sobre o sítio ativo.....	19
FIGURA 8	Ilustração em 3D da HIV-1protease complexada com o inibidor Saquinavir (código 1HXB no PDB). A cor vermelha identifica as hélices α , ciano as folhas β e cinza as voltas sem estrutura secundária fixa. Estão em destaque na proteína os aminoácidos da tríade catalítica fazendo ligações com o ligante. O Saquinavir segue a mesma representação: vermelho para átomos de oxigênio, cinza para carbono e azul para nitrogênio (átomos de hidrogênio foram omitidos).....	21
FIGURA 9	Estrutura dos dois peptídeos idênticos, unidos por um grupo diol vicinal (R1-R2-diol-R2-R1), como pode ser exemplificado pelo composto derivados de (1S,2R)-(-) cis-1-amino-2-indanol.....	24
FIGURA 10	Estado de transição e compostos gerados posteriormente..	25
FIGURA 11	Esquema geral de ligações de hidrogênio entre HIV-1 PR e um substrato representativo.....	27
FIGURA 12	Estrutura do esqueleto principal dos inibidores peptídicos da HIV-1 protease com grupamento diol vicinal central....	35
FIGURA 13	Gráfico da distribuição dos valores de constante de	

	inibição (pKi) dos 46 compostos em estudo. Os compostos de treinamento estão representados por (■) e os do conjunto de teste por estrela vermelha.....	46
FIGURA 14	Procedimento geral do processo de alinhamento combinando buscas 2D e 3D de similaridades entre as estruturas no <i>Symposar</i>	48
FIGURA 15	Campo de Força Direcional implementado no pacote de programas Birograf ^x	49
FIGURA 16	Esboço da molécula do ligante (representado seu SAS; superfície cinza) na representação da dupla camada do receptor (linhas sólidas). Durante o processo de adaptação estérica, os campos gerados pelo sítio de ligação da proteína para o ligante SAS são calculados por interpolação linear entre interior e exterior da camada, se o ligante SAS ficar entre as duas camadas (setas tracejadas). Para pontos localizados na superfície dentro da camada interna, este último pode se adaptar apenas em parte para a topologia do ligante (flechas sólidas; linha pontilhada).....	53
FIGURA 17	Estereorepresentação do alinhamento obtido para a série de compostos em estudo no <i>Symposar</i>	61
FIGURA 18	Gráfico dos valores de pKi experimental e predito para os valores obtidos no QSAR-5D.....	64
FIGURA 19	Estereorepresentação dos modelos do receptor obtidos no encaixe induzido pela representação da dupla camada mapeada de acordo com as propriedades físico-químicas. As cores significam: castanho = muito hidrofóbico; amarelo = hidrofóbico; cinza = neutro ou hidrofóbico;	

	vermelho = acceptor de ligação de hidrogênio; azul = doadora de ligação de hidrogênio.....	65
FIGURA 20	Estereorepresentação da dupla camada para os inibidores 8 (A) e 42 (B). A camada interna representada em <i>surface</i> e a externa em pontos.....	67
FIGURA 21	Gráfico obtido no <i>Scramble</i> Teste dos valores de pKi experimental e predito para os valores obtidos no QSAR- 5D.....	72
FIGURA 22	Estrutura do inibidor que mais se adequa à superfície hipotética do receptor gerado no QSAR-5D. Em vermelho, o substituinte R1 e, em verde, o grupo hidrofóbico substituinte em R2.....	73
FIGURA 23	Superfície mapeada pelas propriedades “quase atomísticas” dos inibidores peptídicos obtida no QSAR- 6D. As cores significam: verde=ligação hidrogênio doadora, amarelo=ligação hidrogênio acceptora, marrom claro= região hidrofóbica carregadas positivamente, marrom escuro= região hidrofóbica carregadas negativamente e cinza= região hidrofóbicas neutra.....	75
FIGURA 24	Superfície mapeada pelas propriedades “quase atomísticas” dos inibidores peptídicos 8 e 42, que apresentam o maior e o menor valor de inibição experimental (valor de Ki), obtida no QSAR-6D.....	77
FIGURA 25	Gráfico dos valores de pKi experimental e predito para os valores obtidos no QSAR-6D.....	78
FIGURA 26	Gráfico obtido no <i>Scramble</i> Teste dos valores de pKi experimental e predito para os valores obtidos no QSAR- 6D.....	81

FIGURA 27 Estrutura hipotética, proposta para estudos posteriores, baseados nos resultados obtidos no estudo de QSAR de inibidores peptídicos com diol vicinal central..... 83

RESUMO

SOUZA, Thais Cristina Silva de. **Estudos quantitativos multidimensionais de correlação estrutura-atividade de inibidores peptídicos da HIV-1 protease.** 2009. 100 p. Dissertação (Mestrado em Agroquímica) – Universidade federal de Lavras, Lavras, MG.¹

A AIDS é uma doença provocada por um vírus específico, que afeta o sistema de defesa imunológico. HIV-1 protease é um homodímero, que desempenha o papel de “tesoura molecular” ao clivar diversas poliproteínas não funcionais em proteínas menores e funcionais, de forma a produzir proteínas maduras. Segundo o último informativo da The Joint United Nations Programme on HIV/AIDS - UNAIDS (2009), estima-se que vivem com HIV/AIDS 34.000.000 de pessoas em todo o mundo. Estudos quantitativos de relação estrutura-atividade em 5D e 6D (QSAR-5D e 6D), de 46 inibidores peptídicos com grupamento diol vicinal central da HIV-1 protease, foram realizados usando QSAR-5D (*RAPTOR, Receptor as Poly Tier Object Representation*) e QSAR-6D (*QUASAR, The quase-atomistic receptor modeling*). Doze compostos, que não fizeram parte dos cálculos na etapa de geração dos modelos, foram usados para validar os métodos QSAR-5D e 6D. Os modelos de QSAR-5D e 6D foram gerados apenas utilizando a metodologia dos mínimos quadrados parciais. O melhor modelo do método QSAR-5D gerou valores de $R^2_{\text{Treinamento}}$ e R^2_{Teste} iguais a 0,944 e 0,793, respectivamente. O melhor modelo do método QSAR-6D gerou valores de $R^2_{\text{Treinamento}}$ e R^2_{Teste} iguais a 0,001 e 0,-0,040, respectivamente. Os métodos servirão, portanto, para mapear os sítios de interação da série de compostos em estudo e propor modificações estruturais com o objetivo de torná-los mais potentes e, conseqüentemente, melhores inibidores da HIV-1 protease.

¹ Orientadora: Profa. Elaine Fontes Ferreira da Cunha - UFLA

ABSTRACT

SOUZA, Thais Cristina Silva de. **Quantitative studies of correlation structure activity of peptides inhibitors of the HIV-1 protease.** 2009. 100 p. Dissertation (Master in Agrochemistry) – Federal University of Lavras, Lavras, MG.²

AIDS is a disease provoked by a specific virus, which affects the immunologic defense system. HIV-1 protease is a homodimer, which plays the role of “molecular scissors” in cleaving several non-functional polyproteins into smaller and functional proteins in order to generate mature proteins. According to The Joint United Nations Programme on HIV/AIDS - UNAIDS (2009), latest informative, it is estimated that 34,000,000 persons live with HIV/AIDS all over the world. Quantitative studies of structure-activity relationship on 5D and 6D (QSAR-5D and 6D) of 46 peptide inhibitors with central vicinal diol grouping of HIV-1 protease, were performed by using QSAR-5D (*RAPTOR, Receptor as Poly Tier Object Representation*) and QSAR-6D (*QUASAR, The quase-atomistic receptor modeling*). Twelve compounds, which did not make part of the calculations at the step of generating the models, were used to validate the QSAR-5D and 6D methods. The QSAR-5D and 6D models were generated only by utilizing the partial least square methodology. The best model of the QSAR-5D method generated values of R^2_{Training} and R^2_{Test} equal to 0.944 and 0.793, respectively. The best model of the QSAR-6D method generated values of R^2_{Training} and R^2_{Test} equal to 0.001 and 0.040, respectively. Then, the methods will be useful to map the interaction sites of compounds under study and propose structural modifications with the objective of making them more potent and, consequently, better HIV-1 protease inhibitors.

² Adviser: Profa. Elaine Fontes Ferreira da Cunha - UFLA

1 INTRODUÇÃO

A química medicinal, química biomédica ou química biológica, compreende a síntese ou isolamento de compostos com atividade biológica, a elucidação ou confirmação da estrutura, a caracterização das propriedades físico-químicas, a determinação da atividade biológica, a exploração desta atividade ao nível molecular e o estudo das relações entre estrutura química e atividade biológica.

Um projeto de Química Medicinal compreende as etapas de descoberta, otimização e desenvolvimento do protótipo. Entende-se por descoberta a etapa destinada à eleição do alvo terapêutico, útil para o tratamento de uma determinada fisiopatologia, a aplicação de estratégias de planejamento molecular para desenho de ligantes do alvo selecionado; a etapa de otimização compreende o melhoramento da estrutura do protótipo, através de modificações planejadas com auxílio de técnicas definidas – que podem identificar os diferentes níveis de contribuição farmacofórica, fornecer informações sobre a conformação bioativa e antecipar o índice de reconhecimento molecular pelo alvo selecionado; e por fim a etapa de desenvolvimento do protótipo objetiva a otimização de suas propriedades de absorção, distribuição, metabolismo, excreção e toxicidade e farmacêuticas (solubilidade, odor, sabor), de modo a viabilizar seu uso clínico, através da preparação eventual de pró-fármacos, complexos de inclusão, microemulsões, formulações de liberação lenta, entre outras.

Dessa forma, as pesquisas na área de química medicinal computacional têm como objetivo desenvolver novos fármacos no combate a doenças como o HIV, tuberculose, Alzheimer, entre outras. Esta área de pesquisa é de grande importância à sociedade e ao mercado, devido a uma economia de tempo e dinheiro investido na produção e introdução de um novo fármaco no mercado.

2 REFERENCIAL TEÓRICO

2.1 Aids (HIV)

2.1.1 Síndrome da Imunodeficiência Adquirida

Especula-se que entre as décadas de 20 e 30, o vírus da imunodeficiência símia (SIV, do inglês, *Simian Immunodeficiency Virus*), responsável pela destruição das células do sistema imunológico de macacos, tenha saltado a barreira das espécies e infectado seres humanos, quando um caçador no oeste da África Central foi mordido, manuseou ou ingeriu carne de uma subespécie de chimpanzé africano infectado pelo SIV (Hahn et al., 2000).

O vírus foi isolado no final de 1983, por um grupo de pesquisadores liderados por Luc Montaigner, do Instituto Pasteur de Paris, França, que o denominou associado à linfadenopatia (*lymphadenopathy associated virus* – LAV). Simultaneamente, o grupo de Robert Gallo, trabalhando no Centers for Disease Control and Prevention (CDC), dos Estados Unidos, relatou o isolamento de um novo vírus linfotrópico de célula T humano (*human T-lymphotropic virus* – HTLV), denominando-o HTLV do tipo III, pois acreditavam que um HTLV poderia ser o agente etiológico da AIDS. Posteriormente, ficou comprovado que os vírus descritos eram os mesmos. Em 1984, ocorreu o primeiro isolamento do vírus em indivíduos assintomáticos, indicando o estado de portador, sendo este denominado de retrovírus associado à AIDS (Cunningham et al., 1996; Levy, 1996), e a partir daí foram desenvolvidos testes para detecção de anticorpos no soro de pessoas infectadas, sendo lançados no mercado em 1985 (Ferreira, 1996). Devido à confusão causada pelas várias terminologias para um mesmo agente, em 1986, o International Committee on Taxonomy of Viruses (ICTV) recomendou o uso da denominação vírus da Imunodeficiência Humana (HIV) para esse novo agente.

A pandemia causada pela Síndrome da Imunodeficiência Adquirida (SIDA ou AIDS) atinge populações em todas as regiões do globo. Desde o início da epidemia, mais de 60 milhões de pessoas já foram infectadas pelo vírus da imunodeficiência humana, o HIV (World Health Organization – WHO, 2009). Segundo o último informativo da UNAIDS em 2008, estima-se que vivem com HIV/AIDS 34.000.000 de pessoas em todo o mundo (FIGURA 1), sendo 23.600.000 somente na África Subsaariana. No ano de 2007, registraram cerca de 2.000.000 de novos casos de HIV, entre estes estima-se que 370.000 são crianças com menos de 15 anos (The Joint United Nations Programme on HIV/AIDS – UNAIDS, 2009).

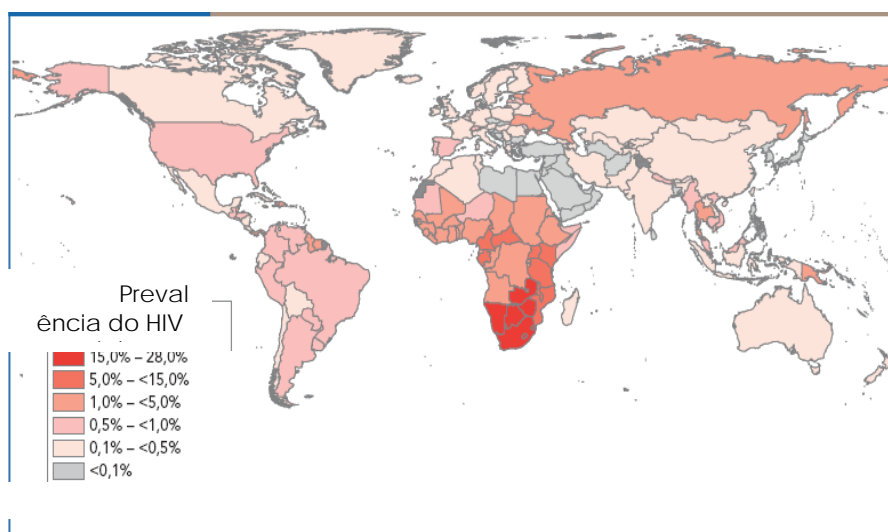


FIGURA 1 Visão global da infecção por HIV, no ano de 2007, segundo o último informe da epidemia mundial de HIV em 2008, publicado pela ONUSIDA.

Na América Latina (Figura 2 e Tabela 1) os dados mostram que 1.700.000 pessoas entre adultos e crianças vivem com HIV/AIDS, e em 2007 o número de novos casos foi de 140.000. Um total de novas infecções no mundo em 2007 foi estimado em mais 3.000.000 novos casos e que 2.000.000 de pessoas morreram devido à AIDS. E já é possível estimar que 50% dos adultos infectados são mulheres (UNAIDS, 2009).



FIGURA 2 Visão da América Latina da infecção por HIV, no ano de 2007, segundo o último informe da epidemia mundial de HIV em 2008, publicado pela ONUSIDA.

TABELA 1 Dados estimados sobre os casos de HIV no mundo e no Brasil.

Número Estimado de Pessoas que Viviam com HIV entre 2001 e 2007							
<i>Adultos e Crianças - 2007</i>		<i>Adultos e Crianças - 2001</i>			<i>Adultos Maiores de 15 anos - 2007</i>		
Pais	Estimativa	Estimativa Baixa	Estimativa	Estimativa Alta	Estimativa	Estimativa (Baixa - Alta)	
Mundial	33000000	30000000-36000000	29500000	27000000-33700000	30800000	28200000-34000000	
Brasil	730000	600000-890000	660000	520000-810000	710000	580000-870000	
Número Estimado de Pessoas que vivem com HIV							
<i>Adultos Maiores de 15 anos - 2001</i>		<i>% de Prevalência em Adultos (15-49 anos) 2001</i>		<i>% de Prevalência em Adultos (15-49 anos) 2007</i>		<i>Mulheres Maiores de 15 anos - 2007</i>	
Pais	Estimativa	Estimativa (Baixa-Alta)	Estimativa	Estimativa (Baixa-Alta)	Estimativa	Estimativa (Baixa-Alta)	Estimativa (Baixa-Alta)
Brasil	640000	510000-790000	0,6	0,5-0,8	0,6	05,-0,8	240000 200000-310000

Fonte: Conjunto de las Naciones Unidas sobre el HIV/SIDA - ONUSIDA (2009).

A AIDS é uma doença provocada por um vírus específico, que afeta o sistema de defesa imunológico. Por causa de falhas no sistema de defesa do organismo humano, microorganismos geralmente não patogênicos em condições normais podem invadir e se proliferar no sangue e nos tecidos, provocando doenças. Por este motivo, esses microorganismos são denominados de “oportunistas”. De forma similar, células cancerosas podem se multiplicar e gerar tumores (Pearl & Taylor, 1987).

O HIV já foi identificado em diversos órgãos, tecidos e fluidos humanos (e.g., sangue, saliva, lágrima, fluido cérebro- espinhal, tecido cerebral, urina, sêmen, leite materno, líquido amniótico e secreções vaginais) (Fiocruz), mas, felizmente, é um vírus extremamente lábil no meio externo, sendo facilmente inativado por agentes físicos (e.g., calor) e químicos (e.g., hipoclorito de sódio e glutaraldeído). Em condições experimentais controladas, partículas virais intracelulares podem sobreviver no meio externo por até um dia, enquanto que partículas virais livres podem sobreviver por até 15 dias em temperatura ambiente ou por até 11 dias à 37°C (Universidade Estadual Paulista - UNESP, 2009).

2.1.2 Vírus

A palavra vírus vem do latim *virus*, que se refere a venenos. Atualmente, a palavra é utilizada para descrever os vírus biológicos, além de designar, metaforicamente, qualquer coisa que se reproduza de forma parasitária. Vírus são parasitas obrigatórios do interior celular, podendo infectar qualquer organismo biológico, e isso significa que eles somente podem se reproduzir pela invasão e posseção do controle da maquinaria de auto-reprodução celular. O termo vírus refere-se, geralmente, às partículas que infectam eucariontes (organismos multicelulares e muitos organismos unicelulares), enquanto que o

termo bacteriófago ou fago é utilizado para descrever aqueles que infectam procariontes (bactérias e assemelhados) (Voet & Voet, 1995).

Os vírus apresentam uma grande variedade de forma e de tamanho. O diâmetro dos principais vírus oscila de 15-300 nm. O vírus da varíola é o maior vírus humano que se conhece (300x250x100 nm), enquanto que o da poliomielite é o menor vírus humano (20 nm de diâmetro) (Fundação Oswaldo Cruz - FIOCRUZ, 2009).

Vírus de diferentes famílias apresentam diferentes morfologias que podem ser prontamente distinguidas pelo microscópio eletrônico. Esta relação é útil para o diagnóstico de doenças virais e, especialmente para reconhecer novos vírus responsáveis por infecções. Alguns vírus têm formas parecidas, daí ser importante o uso de imunomicroscopia eletrônica. Um vírion pode se apresentar sob vários formatos: esférico (influenzavírus), de ladrilho (poxvírus), de bastão (vírus do mosaico do tabaco) e de projétil (vírus da raiva) (FIOCRUZ, 2009).

No ambiente extracelular, o vírus é denominado vírion, víron ou virião, designando uma única partícula submicroscópica viral na forma encapsulada. Esta apresenta, para cada tipo de vírus, algumas características especiais, entre elas diferentes tamanhos e formas. A estrutura de um grande número de vírus foi elucidada utilizando microscopia eletrônica (Briggs et al., 2006). Os vírus não possuem uma organização tão complexa quanto à das células, possuindo, de fato, uma estrutura simples. Eles consistem basicamente de ácido nucléico, ácido desoxiribonucleico (DNA) ou ácido ribonucleico (RNA), envolvido por uma capa protéica (capsídeo) e, em alguns casos, de uma membrana lipoprotéica (envelope). Os principais constituintes são o nucleocapsídeo, o tegumento e o envelope viral representados, esquematicamente, na Figura 3.

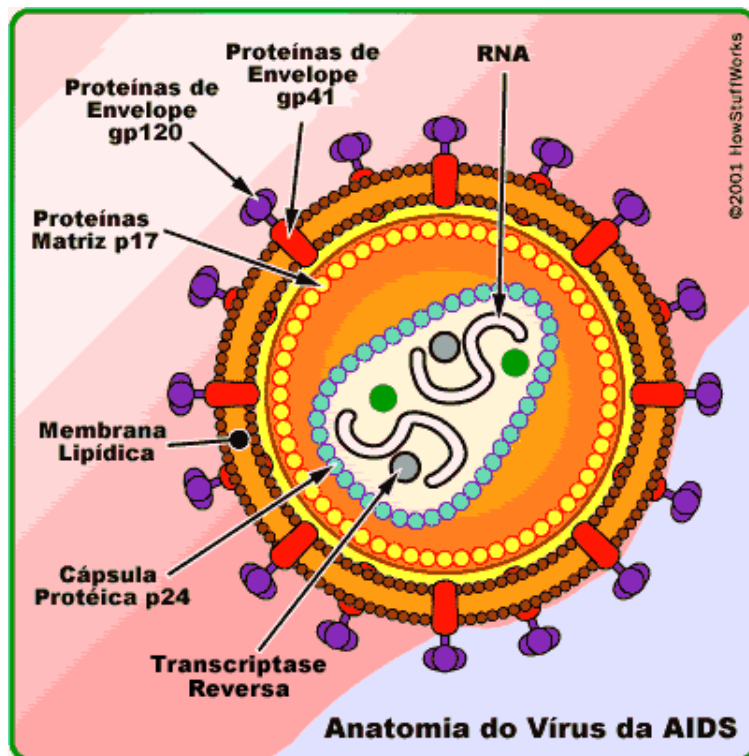


FIGURA 3 Estrutura esquemática do vírus do HIV. Fonte: <http://teusmapress.evonblogs.com.br/virus-e-suas-funcoes/> - http://www.geocities.com/mpennafort/hiv_ciclo.html

Nucleocapsídeo (Figura 3): A maior parte dos vírus tem o seu genoma protegido por uma capa protéica, chamada de capsídeo. O agrupamento das proteínas virais dá ao capsídeo uma geometria característica, normalmente icosaédrica ou helicoidal. O genoma (ácido nucléico), em conjunto com o capsídeo, constitui o nucleocapsídeo. Devido a limitações no tamanho do genoma viral, os vírus não podem codificar uma grande diversidade de proteínas. Assim, o capsídeo viral é formado de subunidades idênticas, denominadas protômeros, que se agrupam formando subunidades maiores, os

capsômeros. Assim, um capsídeo pode ser composto de centenas de capsômeros, mas ainda baseada em um simples modelo icosaédrico (Dimitrov, 2004; Cunha, 2006).

Tegumento (Figura 3): O conjunto de proteínas não glicosiladas, localizadas entre o envelope e o capsídeo, é denominado de tegumento e tem como finalidade sinalizar a montagem de alguns vírus e conferir maior rigidez à estrutura (Dimitrov, 2004; Cunha, 2006).

Envelope viral: Muitos vírus possuem estruturas de membrana complexas envolvendo o nucleocapsídeo. O envelope viral consiste em uma bicamada lipídica com proteínas, em geral glicoproteínas, embebidas nesta. A membrana lipídica provém da célula hospedeira, muito embora as proteínas sejam codificadas exclusivamente pelo vírus (Dimitrov, 2004; Cunha, 2006).

Vírus que possuem DNA como material genético podem empregar diretamente a maquinaria celular para transcrição de seus genes, sua replicação e reparo de seu DNA. Isso permite a alguns vírus ter um genoma grande como os herpesvírus, que evoluíram de forma a produzir alguns genes próprios (e.g., para síntese de nucleotídeos e polimerases próprias), ficando mais independentes do metabolismo celular. Por outro lado, os vírus de RNA devem dispor de enzimas próprias para serem processados, como por exemplo, enzimas de replicação tais como RNA replicases e transcriptases reversas.

2.1.3 O Vírus HIV

O HIV tem como material genético o RNA, e é classificado como retrovírus, não podendo ser inserido nos cromossomos da célula do hospedeiro constituídos por DNA. Para isso, este vírus possui uma enzima específica, a transcriptase reversa (TR), que permite a transcrição reversa do RNA viral em DNA viral. Esta enzima, também conhecida como "DNA-polimerase-RNA-

dirigida" foi descoberta por Howard Temin e David Baltimore em 1970 (Weatherall, 2001; Cunha, 2006).

O HIV pertence à família *Retroviridae*, ao gênero *Lentivirinae*, que caracteriza os vírus de sintomatologia lenta no hospedeiro (Hu et al., 1996; Soares, 2003). Os lentivírus são considerados retrovírus complexos, uma vez que além dos genes *gag*, *pol* e *env* (presentes em todos os retrovírus), possuem uma série de genes que codificam proteínas acessórias ou regulatórias (Soares, 2003). As principais funções exercidas por estes genes incluem o controle da transcrição, o processamento do RNA, a montagem do vírus e a regulação da expressão gênica viral, entre outras. O gene *gag* codifica proteínas do nucleocapsídeo (p17, p24, p7 e p6); o gene *pol* codifica as enzimas protease (p10), a RT (p160 e p51) e integrase (p32); e o gene *env* codifica as proteínas do envelope (gp120 e gp41) (White & Fenner, 1994).

Na Figura 4 são apresentadas as proteínas virais associadas à maturação do vírus do HIV. Todos os lentivírus são envelopados por uma bicamada lipídica (amarela, Figura 5) que é derivada da membrana da célula do hospedeiro (Summers & Turner, 1999).

Genes adicionais (*tat*, *ver*, *vif*, *vpu*, *nef*, *vpr*) codificam proteínas acessórias e regulatórias do vírus (Tabela 2). As proteínas acessórias são necessárias em diferentes passos da maturação da partícula viral, afetando a infecciosidade, a replicação e a montagem do vírus. As proteínas regulatórias são responsáveis por regular diferentes estágios da replicação viral (Vaishnav & Wong-Staal, 1991).

O esquema da estrutura "clássica" de um genoma retroviral é 5'-LTR-gag-pol-env-LTR-3'. As regiões LTR ("terminal repetido longo") representam as duas extremidades do genoma viral que estão ligadas ao DNA celular da célula hospedeira após integração e que não codificam nenhuma proteína viral.

TABELA 2 Resumo das funções das proteínas codificadas pelos nove genes do HIV.

Gene	Proteínas Codificadas e/ou Função das Proteínas
<i>Gag</i> (Kjems et al., 2006)	Proteína Estrutural que forma a camada ao redor do RNA viral (nucleocapsídio).
<i>Env</i> (Kjems et al., 2006)	Glicoproteínas da camada externa (envelope) do vírion.
<i>Pol</i> (Kjems et al., 2006)	Enzimas: transcriptase reversa, protease, integrase
<i>Tat</i> (Wei et al., 1998)	Estimula a transcrição do DNA proviral do HIV em RNA, promove a elongação do RNA, estimula o transporte do RNA do HIV do núcleo para o citoplasma e é essencial para a tradução.
<i>Ver</i> (Wei et al., 1998)	Possui as funções do gene Tat.e é também um fator nuclear de exportação importante na troca da expressão precoce das proteínas reguladoras pelas proteínas estruturais que são sintetizadas mais tardiamente.
<i>Vpr</i> (Miller et al., 1997)	Parece ser essencial para a replicação viral em células como os macrófagos
<i>Nef</i> (Aiken et al., 1994; Collins et al., 1998; Peter, 1998)	Pode induzir a sub-regulação dos CD4 e das moléculas HLA de classe I e II a partir da superfície das células infectadas com VIH-1, também interferir com a ativação das células T por ligação a várias proteínas que estão envolvidas nas vias intracelulares de transdução de sinal
<i>Vif</i> (Mariani et al., 2003)	Importante para os mecanismos intracelulares de transporte dos componentes virais
<i>Vpu</i> (Bour et al., 1995)	Importante para o processo de formação de vesículas, e é também envolvida quando os complexos CD4-gp160 são degradados no retículo endoplasmático

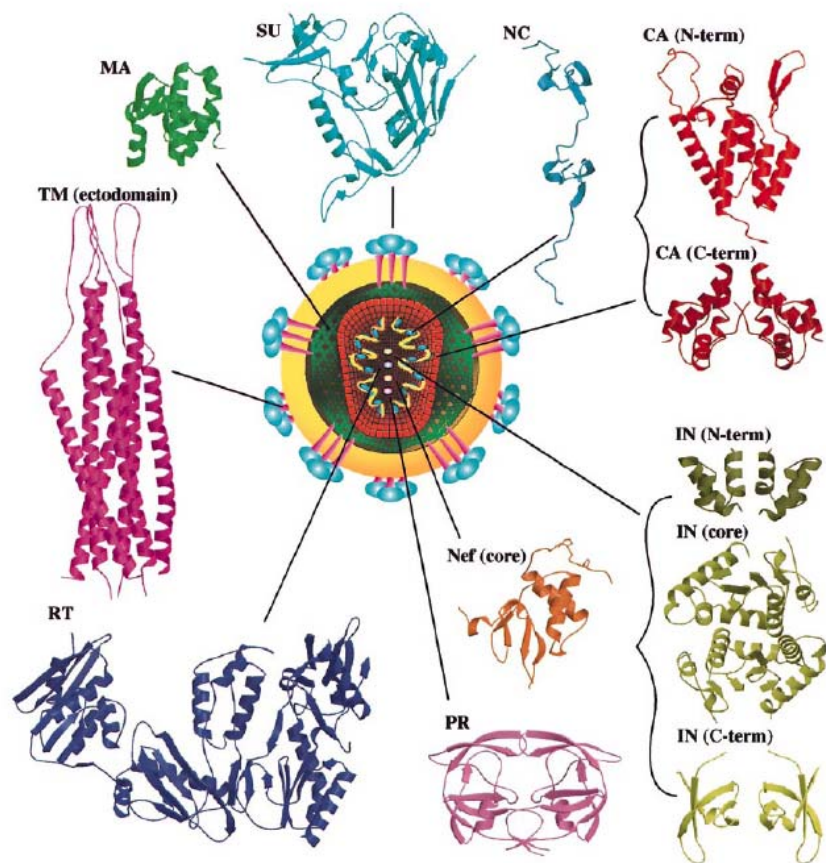


FIGURA 4 Desenho do vírion do HIV maduro circundado pela representação em fitas das proteínas virais e dos fragmentos de proteína caracterizados estruturalmente. As estruturas das proteínas foram representadas na mesma escala. Onde, SU= glicoproteínas de superfície, TM=proteínas da transmembrana, MA=matriz protéica, CA= proteína do capsídeo, NC= proteína do nucleocapsídeo, PR=protease, RT=transcriptase reversa, IN=integrase, Nef=proteínas acessórias. Adaptado de Summers & Turner (1999).

Até o momento são conhecidos dois tipos antigênicos (variáveis) do HIV, o tipo 1 (HIV-1) e o tipo 2 (HIV-2), que apresentam de 40 a 45% de homologia entre suas estruturas primárias, onde 39 dos 99 aminoácidos são idênticos ao serem alinhados (Gustchina & Weber, 1991; Priestle et al., 1995). Ambos os tipos estão presentes em amostras de pacientes com AIDS, sendo que o tipo 1, em comparação com o tipo 2, é mais virulento, mais disseminado pelo mundo, apresenta maior taxa de mutação e maior variedade genética (subtipos).

2.1.4 Ciclo viral

Antes do processo de replicação propriamente dito ocorrer, qualquer vírus deve ser transportado do meio extracelular (e.g., o sangue ou a linfa do organismo) para o interior de uma célula hospedeira. Cada tipo de vírus apresenta uma maior ou menor afinidade com um determinado tipo de célula, e, para atingir este alvo, *i.e.*, a célula hospedeira, o vírus deve inicialmente escapar dos mecanismos de defesa do sistema imune do organismo sendo capaz de se aproveitar qualquer falha temporária que o organismo possa apresentar.

O HIV tem a particularidade de ser capaz de atacar os próprios mecanismos de defesa do organismo, *i.e.*, as células do sistema imunológico que, normalmente, são encarregadas de eliminá-los. O HIV infecta, particularmente, células de defesa contendo em sua superfície o receptor CD4 (Douek et al., 2002), uma macromolécula encontrada principalmente nas células-T auxiliares (linfócitos), responsáveis pela secreção de citocinas, e também nos macrófagos, células de defesa responsáveis pela fagocitose de microorganismos patogênicos.

O processo de replicação do HIV, cujo material genético é composto por RNA, é um pouco mais complexo do que o de vírus composto por DNA. Assim, o ciclo de replicação do HIV pode ser dividido, resumidamente, em seis estágios

(fusão, transcrição reversa, integração, transcrição, tradução e formação viral), conforme é ilustrado, esquematicamente, na Figura 5 (Sierra et al., 2005).

Fusão. O HIV, ao ser reconhecido pela célula-T do organismo, via interação entre o complexo gp120-gp41 e o receptor CD4 da célula hospedeira, ativa outros receptores da célula hospedeira (e.g., CCR5 e CXCR4) e funde-se à célula-T, após a dissociação da gp-120 e gp41 ter ocorrido, permitindo a entrada do capsídeo na célula (Peçanha et al., 2002).

Transcrição reversa. Após a fusão, o capsídeo é desencapado liberando o RNA e enzimas virais dentro da célula hospedeira. A transcriptase reversa viral traduz o material genético de RNA para DNA de fita dupla e, em seguida, o complexo núcleo protéico, incluindo o DNA recém formado e as proteínas p31 (integrase) e p17, é transportado para o núcleo da célula hospedeira em um processo mediado pela proteína vpr (Peçanha et al., 2002).

Integração. No núcleo da célula hospedeira, a integrase viral catalisa a integração do DNA viral com o DNA da célula hospedeira, processo que permite ao HIV “reprogramar” a célula hospedeira para “fabricar” novas partículas virais (Peçanha et al., 2002).

Transcrição. O DNA transcreve-se, formando o RNA viral, os quais são, algumas vezes, chamados de RNA mensageiro. A proteína *Tat* auxilia regulando o processo de transcrição do RNA (Aboul-Ela et al., 1995).

Tradução. O conjunto de RNAs transcritos é transportado para o citoplasma, e as poliproteínas virais, constituídas dentro da célula humana, produzirão a nova partícula do HIV. Estas poliproteínas são criadas através da

tradução das informações fornecidas pelo mensageiro do RNA, regulado pela proteína rev (Peçanha et al., 2002).

Formação Viral. As poliproteínas virais são cortadas em partes menores pela protease viral, formando uma nova partícula do HIV, que inclui as enzimas e proteínas necessárias para a repetição do processo replicativo. Uma vez formada, a nova partícula viral sai da célula hospedeira e passa para a corrente sanguínea, podendo, assim, infectar outras células. Aproximadamente 10,3 bilhões de novos vírions são formados diariamente em pessoas que não usam medicamento algum (Swanstrom & Erona, 2000).

Após a ligação e fusão para a superfície da célula hospedeira, o HIV-1 entra no núcleo da célula hospedeira. A transcriptase reversa sintetiza duplas-fitas provirais de DNA, que são transportadas para o núcleo. A integração ocorre e o subsequente provirus exprime o RNA viral que irá sintetizar proteínas virais usando os ribossomos do hospedeiro. Novos vírions são criados pela montagem e brotação através da membrana celular infectada e posterior maturação devido a ações de proteases (Pomerantz & Horn, 2003).

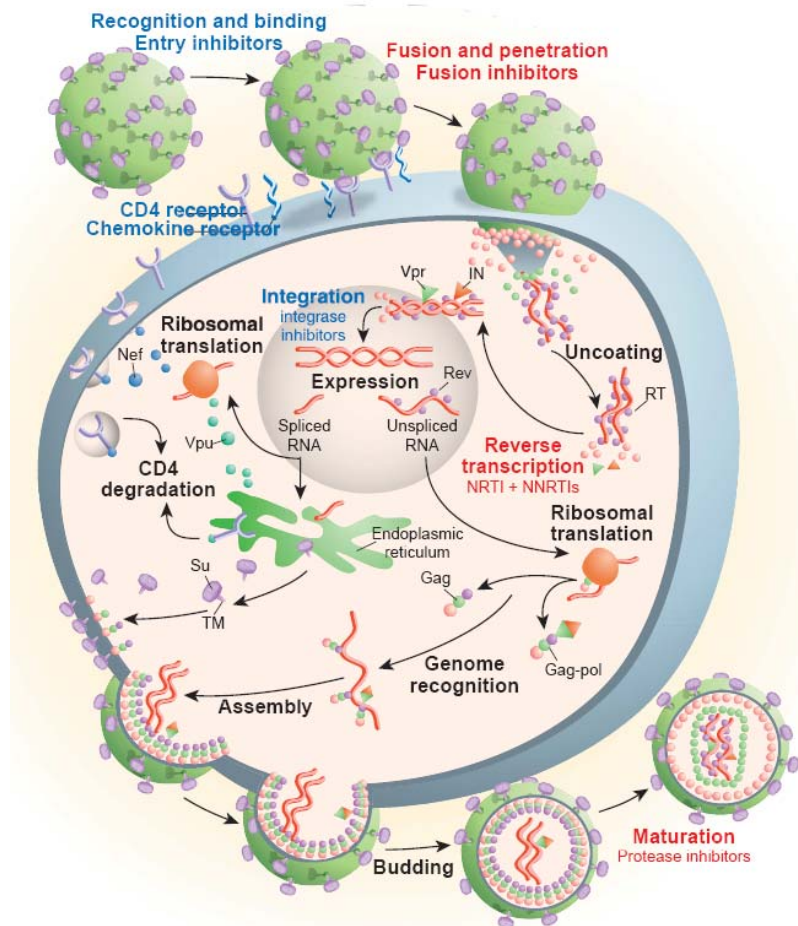
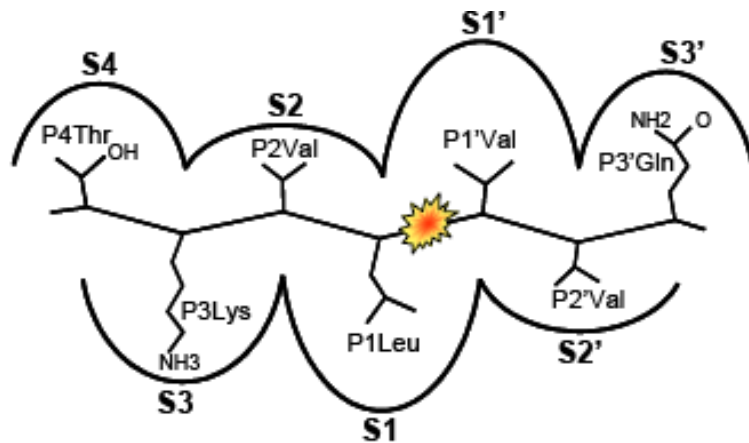


FIGURA 5 Ciclo de vida do HIV. Os textos em vermelho se referem aos alvos inibidos por agentes existentes aprovados pela FDA. Os textos em azul referem-se aos alvos que são inibidos por agentes que poderão vir a estar disponíveis no futuro.

2.1.5 Estrutura da HIV-protease

As primeiras estruturas da protease nativa foram quase simultaneamente resolvidas por três grupos, duas utilizando proteína produzida por expressão heteróloga em *E. coli*, pelos Merck Research Laboratories (Navia et al., 1989) e Birbeck College (Lapatto et al., 1989); e uma outra utilizando uma proteína produzida por síntese total, pela NCI-Frederick (Wlodawer & Vondrasek, 1998). No entanto, somente as coordenadas dos C- α da estrutura resolvida pelos pesquisadores da Merck foi publicada, e ainda assim, notou-se que essa estrutura possuía erros grosseiros no traçado da cadeia (Lapatto et al., 1989; Wlodawer & Vondrasek, 1998; Kent et al., 2000). Esta enzima é indispensável para a maturação do vírus da AIDS impedindo a formação de novos vírions, fazendo desta proteína um dos alvos primários para terapia com fármacos inibidores (Batista, 2005).

De forma geral, proteases de retrovírus possuem um sítio ativo composto de sete subsítios, divididos no mínimo, quatro à direita e três à esquerda do ponto de clivagem do substrato. O subsítio de clivagem é denominado subsítio S1/S1' e é formado por dois resíduos de aspartato. A protease reconhece o local de clivagem na estrutura do substrato protéico pela presença de uma seqüência característica de cinco a seis resíduos hidrofóbicos (Figura 6) (Tozser et al., 2000).



S4 – Asp29, Asp30, Met46, Ile47, Gly48, Leu76

S3 – Arg8', Leu23', Asp29, Gly48, Gly49, Ile50, Thr80', Pro81', Val82'

S2 – Ala28, Asp29, Asp30, Val32, Ile47, Gly49, Ile50, Leu76, Ile84

S1 – Arg8', Leu23', Asp25', Asp25, Gly27, Gly49, Ile50, Thr80', Val82', Ile84'

S1' - Arg8', Leu23', Asp25', Asp25, Gly27, Gly49, Ile50, Thr80', Val82', Ile84'

S2' - Ala28', Asp29', Asp30', Val32', Ile47', Gly49', Ile50', Leu76', Ile81'

S3' - Arg8, Leu23, Asp29', Gly48', Gly49', Ile50', Thr80, Pro81, Val82

FIGURA 6 Representação esquemática do sítio catalítico da HIV-Pr associado com um substrato polipeptídico. S e S' identificam os subsítios (resíduos de aminoácidos) da enzima e P e P' identificam os subsítios (resíduos de aminoácidos) do substrato polipeptídico. O polipeptídeo representado é Pro-Ala-Val-Ser-Leu-Met-Thr. O destaque em amarelo e vermelho representa o local de clivagem. Adaptado da Tozser et al. (2000)

A região entre os resíduos 43-58, chamada de alça (*flap*), contém as fitas- β e se fecha sobre o sítio ativo da protease complexada com inibidores ou substratos, formando um teto (Figura 7). Do ponto de vista mecânico, as duas grandes alças (ida: do resíduo 43 aos 50; volta: do resíduo 51 aos 58), localizadas na parte superior da protease abrir-se-iam, permitindo o encaixe da

poliproteína (certamente em forma distendida) em seu sítio ativo (Collins et al., 1995a,b; Batista, 2005). Esses *flaps* encontram-se abertos na estrutura não complexada e esse fechamento sobre o sítio ativo constitui a maior diferença entre as estruturas complexada e não-complexada da protease (Miller et al., 1989; Fitzgerald et al., 1990; Wlodawer & Vondrasek, 1998). No entanto, as estruturas das proteases são bastante conservadas entre os diferentes complexos, com RMSD entre os C- α raramente excedendo 0,6 Å (Batista, 2005).

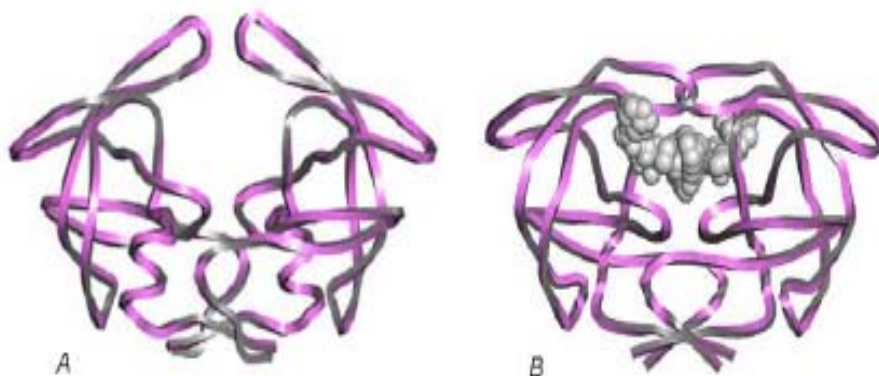


FIGURA 7 Mudança conformacional provocada pela presença do inibidor no sítio ativo. (A) HIV-1 Protease em sua forma livre e em (B) o complexo com o inibidor. A ligação do inibidor induz uma mudança conformacional na qual os *flaps* se fecham sobre o sítio ativo.

A HIV-1 protease é então, um homodímero de simetria do tipo C₂, contendo um único sítio catalítico, caracterizado pela presença da tríade Asp-Thr-Gli (Figura 8) característica das aspartil-proteases. No dímero da HIV-Pr, os monômeros (cadeias A e A') não estão ligados covalentemente e cada cadeia é formada por 99 resíduos de aminoácidos, contendo uma α -hélice (normalmente do resíduo 87 ao 93) e nove folhas- β (Figura 9). O substrato fica alojado numa

cavidade formada por estruturas em forma de laço (loop) e a entrada do substrato nesta cavidade parece estar condicionada a presença de estruturas flexíveis que funcionariam com tampas (flaps) que, presumivelmente, estariam associadas a afinidade da enzima pelo substrato. Dois resíduos de Asp catalíticos, Asp25 e Asp25', identificados no sítio ativo da HIV-Pr, ligam-se a uma molécula de água responsável pela hidrólise das poliproteínas (substratos).

Uma vez que o núcleo da célula infectada começa a produzir longas cadeias polipeptídicas não funcionais, nomeadas como poliproteínas gag-pol precursoras, na seqüência de eventos, tem início à formação de vírions imaturos próximos à superfície celular. A protease desempenha seu papel de “tesoura molecular” ao clivar diversas poliproteínas em proteínas menores e funcionais, de forma a gerar proteínas maduras. Esta proteólise específica pode ocorrer durante a formação do vírion até após sua liberação no plasma (Henderson et al., 1988). O processamento incompleto das poliproteínas resulta na formação de partículas virais imaturas, não infecciosas (Crawford & Goff, 1985; Batista, 2005).

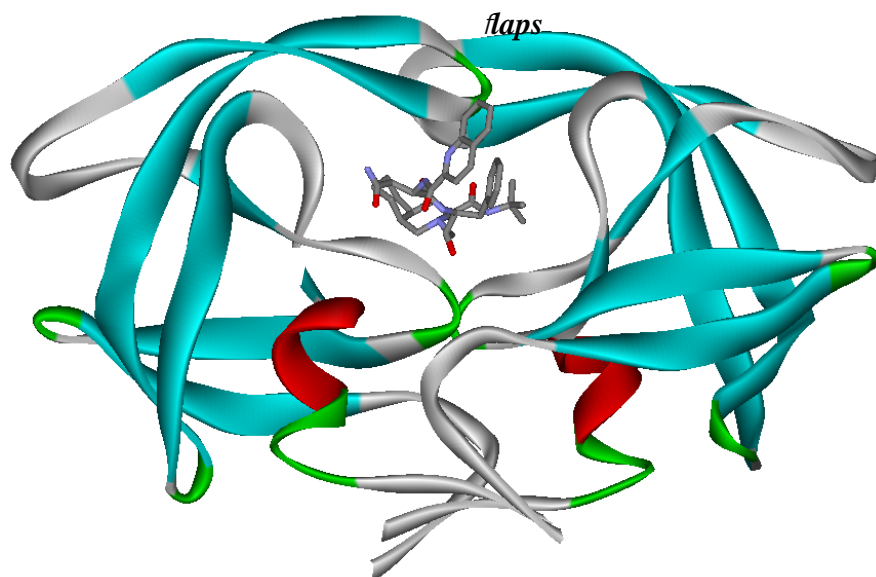


FIGURA 8 Ilustração em 3D da HIV-1protease complexada com o inibidor Saquinavir (código 1HXB no PDB). A cor vermelha identifica as hélices α , ciano as folhas β e cinza as voltas sem estrutura secundária fixa. O Saquinavir segue a representação: vermelho para átomos de oxigênio, cinza para carbono e azul para nitrogênio (átomos de hidrogênio foram omitidos).

2.1.6 Inibidores da HIV-1 protease

O primeiro inibidor de protease de HIV-1 descoberto, foi a pepstatina, um inibidor de pepsina, que também é uma protease aspártica (Martins, 2007). Os inibidores de protease (IPs) pertencem a uma classe de compostos peptidomiméticos que se ligam ao sítio ativo da enzima, e bloqueiam a produção de vírus infecciosos de células infectadas através da inibição da clivagem das poliproteínas precursoras necessárias para produzir vírions. As partículas produzidas por células tratadas com IPs contem precursores não processados e são não- infecciosas (Peng et al., 1989). Há dez IPs aprovados pela FDA (*Food and Drug Administration*) e em uso atualmente: saquinavir (SQV), indinavir

(IDV), ritonavir (RTV), nelfinavir (NFV), lopinavir (LPV), amprenavir (APV), atazanavir (ATV), fosamprenavir (FOS), tipranavir (TPV) e darunavir. Destes, os mais comumente utilizados no Brasil são o nelfinavir, indinavir, saquinavir e ritonavir. LPV, APV e ATV são menos freqüentes nos regimes de tratamento e TPV ainda não é utilizado no Brasil. RTV é freqüentemente usado combinado a outros IPs porque, aparentemente, tem a capacidade de diminuir a metabolização destas outras drogas, aumentando, assim, o nível basal de cada uma delas, o que resulta numa maior eficiência (Oliveros, 2005).

A busca de novos agentes mais eficazes contra o vírus HIV visa não somente reduzir os efeitos adversos e o número de doses e/ou tempo de tratamento, pois ainda não se conhece uma cura para a AIDS, mas principalmente atender aos pacientes com elevada resistência aos medicamentos.

2.2 Inibidores peptídicos com simetria C-2

2.2.1 Grupo diol vicinal

O padrão estrutural que combina um grupo diol vicinal (1,2-dihidroxietano ou 1,2- etanodiol ou etileno glicol) central em compostos peptídicos ou peptidomiméticos com simetria C-2 se tornou uma estratégia modelo no planejamento de uma série de potentes inibidores da HIV-1 protease contendo este padrão estrutural. Estes compostos têm como principal finalidade explorar a simetria C2 (moléculas com dois eixos/planos de simetria) inerente a HIV-1 protease. O argumento central desta estratégia é a suposição de que, ao ligar-se no sítio ativo da enzima, o inibidor adota uma orientação em que os eixos de simetria de ambos, inibidor e enzima, coincidem. Entretanto, estudos de cristalografia de raios-X dos inibidores A-77003, A-76889, A- 76928 e A-78791 (Tabela 3) com simetria C2 e grupamento diol vicinal central, demonstram que ele interage com a HIV-1 protease numa conformação simétrica ou assimétrica (Hosur & Bhat, 1994; Cunha, 2006).

TABELA 3 Estrutura química dos análogos diol baseado na simetria C-2 .

Compostos	X	Y
A-77003	<i>R</i> -OH	<i>S</i> -OH
A-76889	<i>R</i> -OH	<i>R</i> -OH
A-76928	<i>S</i> -OH	<i>S</i> -OH
A-78791	<i>S</i> -OH	H

As interações por ligação hidrogênio entre os resíduos aspartato catalíticos da HIV- 1 Pr e o grupo (R)-OH dos diastereoisômeros (R,S) e (R,R), evidentemente, induzem um modo de ligação assimétrico. A sobreposição destes isômeros demonstra que eles apresentam conformações similares, contudo, o (R,S)-diol é 10 vezes mais potente do que (R,R)-diol. Os autores deste estudo (Yaakov & Caflish, 2003) sugerem que, provavelmente, o segundo grupo (R)-OH do isômero (R,R) deve fazer contato de van der Waals repulsivo com o resíduo de Gly27' (distância de ~2,5Å entre os átomos de oxigênio do grupo hidroxila e da Gly27').

Alguns dados da literatura sugerem que apesar da “aparente” simetria C₂, a HIV-1 protease não deve ser simétrica quando está interagindo com um ligante (Appelt, 1993), visto que, apenas um grupo aspartato encontra-se ionizado (Baldwin et al., 1989; Lapatto et al., 1989; Okimoto et al., 2000) e a flexibilidade de cada monômero é distinta (Prabu-Jeyabalan et al., 2000).

Desta forma, não é tão relevante basear-se exclusivamente na simetria C₂ para o planejamento de novos inibidores para HIV-1 Pr. Outros fatores como, por exemplo, a presença do grupo diol vicinal central podem ser mais determinantes (Buth et al., 1995).

A partir de 2001, Samuelsson e colaboradores (Samuelsson et al., 2001a,b,c, 2003, 2005, 2006) desenvolveram uma série de inibidores peptídicos

da HIV-1 protease contendo um grupo diol vicinal como unidade central e apresentando simetria C-2. Estruturalmente, a série apresenta dois peptídeos idênticos unidos por grupo diol vicinal (R1-R2-diol-R2-R1), como pode ser exemplificado pelo composto (2*R*,3*R*,4*R*,5*R*)-2,5-bis((2,4-difluorobenzil)oxi)-3,4-diidroxi-*N,N*-bis((1*R*,2*S*)-2-hidroxi-2,3-diidro-1*H*-inden-1-il)hexanodiamino (Figura 9), que apresenta elevada potência em ensaios de culturas de células infectadas com HIV-1 protease.

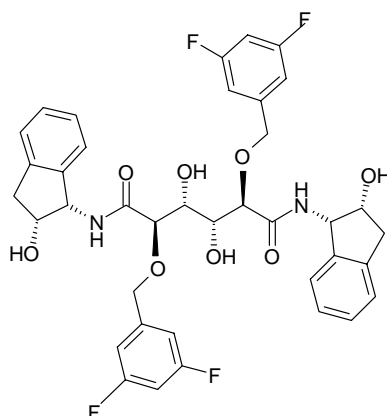


FIGURA 9 Estrutura dos dois peptídeos idênticos, unidos por um grupo diol vicinal (R1-R2-diol-R2-R1), como pode ser exemplificado pelo composto derivados de (1*S*,2*R*)-(-)*cis*-1-amino-2-indanol.

Em 1998, um curto e estereo-controlado método para a síntese de inibidores C2-simétrico para a HIV-1 protease foi reportada. Nestes inibidores, uma unidade de dihidroxietileno serve como um estado de transição mimetizando o fragmento e grupos benzilóxi como P1/P1' nas cadeias laterais. O núcleo dessas estruturas de carboidratos base é subsequentemente usado para explorar o impacto da elongação das cadeias laterais de P1/P1' pela catálise de paládio em

reações acopladas. Esta operação permite diversos inibidores (Figura 10) da HIV-1 protease com altas potências em enzimas e culturas de células (Samuelsson et al., 2005).

Samuelsson e colaboradores observaram em estudos de correlação estrutura-atividade, que inibidores fluorados se mostraram potentes inibidores quando monofluorossustituídos, e em alguns casos podem ser comparados a inibidores como o ritonavir; fato que pode ser em parte explicado pelo aperfeiçoamento da permeabilidade celular atribuído a alta lipofilicidade de compostos fluoro substituídos (Samuelsson et al., 2001b).

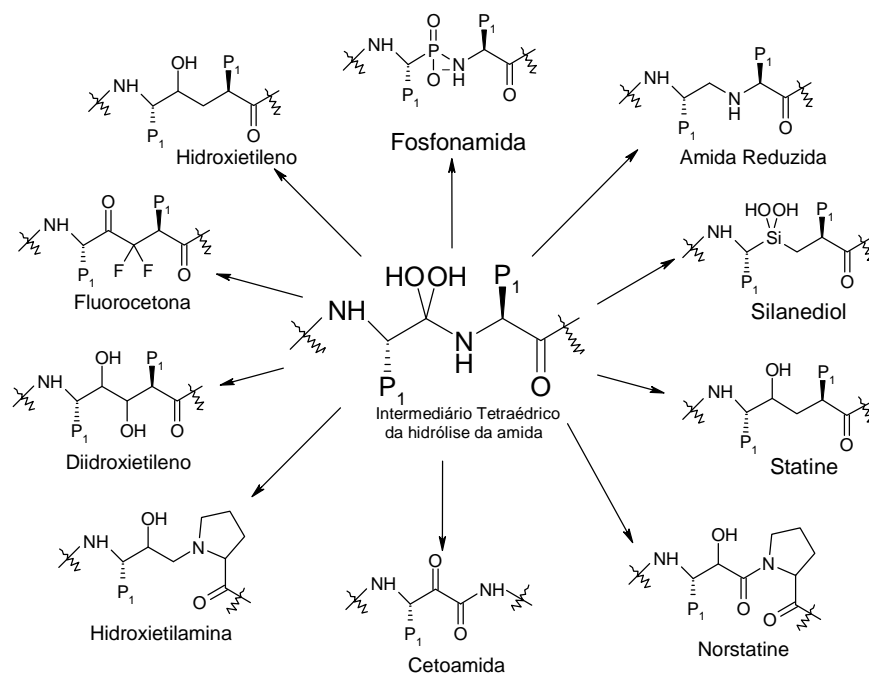


FIGURA 10 Estado de transição e compostos gerados posteriormente. Adaptado de Wong & Brik (2003).

Adicionalmente, Torsten e colaboradores, utilizaram a cristalografia de raios-X para decifrar a relação estrutura-atividade para uma série de inibidores fluorados. Em geral, a substituição com flúor resulta na utilização eficiente de volumes acessíveis em subsítios, associados com um aumento do número de interações de van der Waals e área da superfície na qual se encaixam os inibidores. Isto reflete em uma significativa inibição da protease, embora menos eficiente do que seu análogo não-substituído. A redução geral na eficácia de ligação associada com a substituição com flúor é contraditório com respeito ao aumento do número de interações de van der Waals. Estruturalmente e bioquimicamente, dados sugerem que diflúor substituintes na posição *orto-meta* de grupos simétricos na cadeia do diol vicinal de inibidores de protease, preservam a eficácia e significância da potência antiviral (Torsten et al., 2004).

Estudos também mostram a ordem de preferência da PR por diferentes isósteros: hidroxietileno (CH(OH)-CH₂) > difluorostatona (CO-CF₂-CO) > estatina (CH(OH)-CH₂-CO) > fosfinato (P(O₂)H-CH₂) > amida reduzida (CH₂-NH). Normalmente, grupos hidrofóbicos e aminoácidos não naturais são incorporados a estes isósteros para aumentar a potência do inibidor. Por exemplo, o isóster hidroxietilamina facilmente acomoda um grupo aminoácido, mimetizando assim a ligação Phe-Pro ou Tyr-Pro, encontrada em vários sítios de clivagem das poliproteínas do HIV (Batista, 2005). E ainda que, embora o comprimento mínimo do substrato seja de sete aminoácidos, um mimético tri ou tetrapeptídico pode ser um inibidor potente da PR. Em todos os casos, exceto para a amida reduzida, o grupo hidroxila da junção não hidrolisável é posicionado entre os ácidos aspárticos da protease, com distância para uma interação de ligação de hidrogênio com pelo menos um oxigênio do carboxilato dos ácidos aspárticos. As ligações de hidrogênio se formam, principalmente entre os átomos da cadeia principal da enzima e do substrato/inibidor e seguem

um padrão bastante similar entre os complexos (Figura 11) (Wlodawer & Vondrasek, 1998; Batista, 2005).

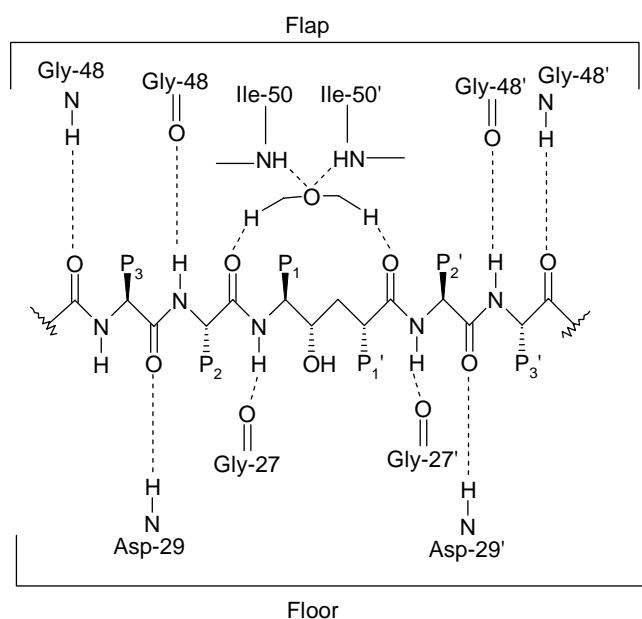


FIGURA 11 Esquema geral de ligações de hidrogênio entre HIV-1 PR e um substrato representativo. Adaptado de Garg et al. (1999).

2.3 QSAR

2.3.1 História do desenvolvimento do QSAR

Tem quase 40 anos desde que a relação quantitativa de estrutura-atividade (QSAR) teve seu primeiro paradigma encontrado na prática da agroquímica, química farmacêutica, toxicologia, e eventualmente mais facetas da química. Isto pode ser atribuída à força do postulado inicial de que a atividade é uma função da estrutura como descrito pelas atribuições eletrônicas, hidrofobicidade, e propriedades estéricas bem como o rápido e extensivo

desenvolvimento em metodologias e técnicas computacionais que têm resultado em delinear e refinar as muitas variáveis e aproximações que definem o paradigma (Selassie, 2003).

Se utilizados corretamente, métodos computacionais para análise da correlação estrutura-atividade (QSAR) podem fornecer um menor custo e um menor tempo no desenvolvimento de uma nova droga quando comparados a procedimentos empíricos (Patrick, 1995). Os mais recentes métodos desenvolvidos nesta linha são os métodos de QSAR-4D (Hopfinger et al., 1997), QSAR 5D e 6D (Vedani et al., 2005) e MIA-QSAR (Freitas et al., 2005).

2.3.2 Relações quantitativas entre estrutura e atividade/propriedade (QSAR/QSPR)

As relações quantitativas entre estrutura e atividade/propriedade iniciaram-se na década de 1960 com os trabalhos de Hansh & Fujita (1964) que tiveram como objetivo a construção de modelos matemáticos correlacionando propriedades moleculares e estruturais com a atividade biológica (QSAR) ou qualquer tipo de propriedade física ou química (QSPR). Assim, os métodos de QSPR requerem que propriedades moleculares de uma série de compostos, que na maioria das vezes são análogos, sejam conhecidas e estatisticamente correlacionadas com propriedades de caráter diverso associadas a características estruturais e/ou com propriedades moleculares (Ramos, 2005). Dentre os descritores estruturais é comum o emprego de descritores de contribuição por grupos, topológicos, peso molecular, número de doadores e aceptores de ligação hidrogênio, número de ligações que sofrem rotação, etc. Os descritores de caráter quântico são também muito úteis e utilizados no estudo de inibição de corrosão, sendo eles: momentos de dipolo (Dip), energia do HOMO (EH), energia do LUMO (EL), diferenças HOMO-LUMO (Dif), polarizabilidade (Pol), cargas em sítios específicos, etc (Lipkowitz & Boyd, 1990). Assim, a receita para o sucesso em correlações estrutura-propriedade depende fundamentalmente

do equilíbrio entre os diversos descritores utilizados, sendo essencial o uso de descritores variados representando de maneira equilibrada o caráter local e coletivo com que se manifesta a particular propriedade em estudo.

2.4 Métodos e índices estatísticos

2.4.1 Coeficiente de correlação linear (r)

O estudo da correlação tem por objetivo medir e avaliar o grau de relação existente entre duas variáveis aleatórias (Toledo et al., 1985). O coeficiente de correlação linear (r) é o instrumento empregado para medir a correlação linear, ou seja, é a medida do grau de relação linear entre duas variáveis, usualmente chamadas de X e Y, e pode assumir qualquer valor variando de -1 a +1, onde +1 corresponde a uma correlação direta máxima e -1, a uma correlação inversa máxima. Um valor de coeficiente de correlação linear igual a zero indica que não há qualquer relação linear entre as variáveis. Contudo, a conclusão de que não há correlação linear significativa não quer dizer que X e Y não estejam relacionados de alguma forma, e.g., correlação não linear (Kubinyi, 1997).

Para avaliar o grau de correlação entre duas variáveis, ou seja, medir o grau de ajuste dos valores em torno de uma reta, deve-se utilizar o coeficiente de correlação de Pearson, que é dado, segundo Martins & Donaire (1993), por:

$$r = \frac{n \sum_{i=1}^n x_i \cdot y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right]} \sqrt{\left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \quad (1)$$

Onde n é o número de observações.

2.4.2 Coeficiente de correlação linear quadrático (r^2 ou R^2)

O coeficiente de correlação linear quadrático, r^2 , também conhecido como coeficiente de determinação, é a medida do grau de ajuste da equação de regressão múltipla aos dados amostrais. Ele representa a proporção de discrepância de uma variável que é compartilhada com a outra variável, ou seja, a proporção da discrepância de uma variável que pode ser predita da outra variável. A proporção de discrepância que não pode ser predita ou considerada pela outra variável, $1 - R^2$, é conhecida como coeficiente de alienação (Gaudio & Zandonade, 2001).

2.4.3 Análise de regressão pelo método dos mínimos quadrados parciais (PLS)

PLS é um método de regressão onde uma matriz de dados (X) é correlacionada com o vetor da função resposta (y) numa combinação linear e por meio das componentes principais. As componentes principais são otimizadas para melhor descrever a relação entre os blocos x e y simultaneamente (e por isto, chamadas de variáveis latentes). Variáveis latentes são combinações lineares das variáveis originais, portanto levam em conta todas as variáveis dando peso (loadings) a cada um dos scores (descritores projetados). Os coeficientes do vetor de regressão indicam quais descritores são importantes na modelagem da propriedade (Gaudio et al., 2002). A dimensionalidade dos dados originais é reduzida para o número de variáveis latentes do modelo que apresentar o menor resíduo da variância de y . O modelo obtido por PLS não requer mais compostos que descritores, como na regressão linear múltipla (RLM), além disso, nem todas as variáveis são importantes, já na RLM todas são (Gaudio et al., 2002).

2.4.4 Validação cruzada

A validação cruzada é uma metodologia utilizada para a escolha do número de componentes principais baseada na avaliação da magnitude dos erros de previsão de um dado modelo de calibração. Esta avaliação é feita pela comparação das previsões das concentrações previamente conhecidas (c_i $i = 1:n$), e em resumo consiste do seguinte:

1 - Remove-se uma ou mais amostras i do conjunto de calibração e constrói-se o modelo como anteriormente.

2 - Usa-se o novo modelo para prever os dados removidos c'_i .

3 - Calcula-se o erro de previsão ($c_i - c'_i$).

4 - Calcula-se a soma dos quadrados dos erros de previsão: PRESS $\sum_i (c_i - c'_i)^2$ ou a raiz quadrada RMSEP, que é na realidade um desvio padrão $\sqrt{\frac{\sum_i (c_i - c'_i)^2}{n}}$ onde n é o número de amostras do conjunto de calibração (Ferreira et al., 1999).

O coeficiente de correlação linear quadrático da validação cruzada, q^2 , é definido segundo a equação:

$$r_{CV}^2 = q^2 = 1 - \frac{PRESS}{\sum_y (Y_{Exp} - Y_{médio})^2} \quad (2)$$

Onde, PRESS (Predictive Sum of Squares) = $\sum_y (Y_{Pred} - Y_{Exp})^2$; Y_{Exp} é o valor experimental da atividade biológica; Y_{Pred} é o valor predito pelo modelo; e $Y_{médio}$ é a média dos valores da atividade (ou $(\sum y)^2/n$) (Kubinyi, 1993).

2.4.5 Índice de falta de ajuste

A técnica de PLS inclui um índice denominado de “índice de falta de ajuste” (em inglês, lack-of-fit, LOF), implementado por Friedman (Karki & Kulkarni, 2001), e definido pela Equação 1.

$$LOF = \frac{LSE}{1 - \frac{(c+dp)^2}{M}} \quad (3)$$

onde LSE (*least squares error*) é o erro dos mínimos quadrados calculado pela diferença entre os valores experimentais e preditos da atividade; d (smoothing factor) é um parâmetro ajustável que controla o número de variáveis no modelo; c é o número de variáveis independentes; p é o número total de características contidas em todas as variáveis independentes; e M é o número de amostras no conjunto de treinamento (Cunha et al., 2004).

2.4.6 Validação externa dos modelos

Segundo Tropsha & Golbraikh (2002) a validação externa é um requisito absoluto para o desenvolvimento de um verdadeiro modelo preditivo QSAR.

O melhor modelo de QSAR-5D e 6D é submetido à validação externa usando o conjunto de teste contendo doze compostos que não foram incluídos no desenvolvimento dos modelos de QSAR. Este processo verificará a capacidade do modelo de prever a atividade biológica destes doze compostos.

2.4.7 Teste de *Scramble*

O *scramble test* deve ser um teste obrigatório para qualquer modelo de QSAR. A filosofia subjacente é a seguinte: a sensibilidade para com os dados biológicos utilizados para estabelecer um QSAR é testada pela randomização ao acaso das atividades experimentais dos ligantes do treinamento conjunto

(ΔG_{exp}) e então a simulação é repetida de modo idêntico. Se $R^2_{\text{Treinamento}}$ é comparável ao valor obtido utilizando a randomização dos dados de ΔG , o modelo é inútil já que não é sensível aos dados experimentais utilizados para estabelecer a relação no QSAR. Como uma randomização selecionada, deve-se acidentalmente ser fechado os verdadeiros valores de ΔG , assim, Vedani e colaboradores sugerem que se execute uma série de 2-5 diferentes testes de scramble (Quasar User's Guide).

3 OBJETIVOS

Um alvo no combate a AIDS é a enzima HIV-1 protease, que cliva diversas poliproteínas em proteínas menores e funcionais, de forma a gerar novos vírus. HIV-1 protease é um homodímero, onde cada monômero é constituído por 99 aminoácidos com um resíduo aspartato catalítico localizado no centro do sítio ativo. Neste trabalho, estudos quantitativos de estrutura-atividade em 5D e 6D de 46 inibidores peptídicos com grupamento diol vicinal central, serão realizados usando os programas Symposar, Raptor 2.0 (*Receptor as Poly Tier Object Representation*) e Quasar 6.0 (*The quase-atomistic receptor modeling*) para propor novos candidatos a fármacos anti-HIV. Doze compostos, que não farão parte dos cálculos dos modelos, serão usados para validar os métodos 5D-QSAR e 6D-QSAR. Para planejar a estratégia a ser empregada para o desenvolvimento de potenciais inibidores anti-HIV, foram estabelecidos os seguintes objetivos neste trabalho:

i. Desenvolver uma correlação quantitativa entre a estrutura e a atividade, de um conjunto de treinamento de 34 compostos, sintetizados e avaliados farmacologicamente, usando os métodos de QSAR.

ii. Comparar os valores de q^2 , $R^2_{\text{Treinamento}}$ e R^2_{Teste} dos diferentes modelos de QSAR e visualizar os modelos através de mapas de contorno.

iii. Propor modificações dos substituintes nas posições R1, R1', R2 e R2' (Figura 12) a fim de se propor um novo potencial composto com atividade inibitória.

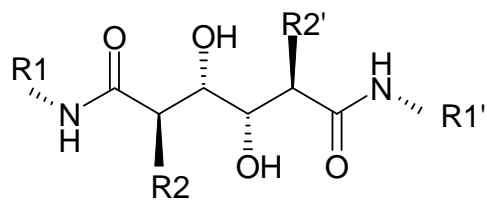


FIGURA 12 Estrutura do esqueleto principal dos inibidores peptídicos da HIV-1 protease com grupamento diol vicinal central.

4 MATERIAL E MÉTODOS

4.1 Dados biológicos

Na Tabela 4 está relacionada à série de inibidores análogos ao (2*R*,3*R*,4*R*,5*R*)-2,5-*bis*((2,4-difluorobenzil)oxi)-3,4-diidroxi-*N,N'*-*bis*((1*R*,2*S*)-2-hidroxi-2,3-diidro-1*H*-inden-1-il)hexanodiamino e seus respectivos valores de atividade inibitória da HIV-1 protease.

Os modelos de QSAR-5D/6D foram desenvolvidos usando um conjunto de 34 compostos (1, 3-5, 7-9, 11, 12, 15-22, 24, 26-27, 29, 31-33, 35-37, 39-41, 43-46) denominado conjunto de treinamento e validados externamente usando um conjunto de 12 compostos (2, 6, 10, 13-14, 23, 25, 28, 30, 34, 38, 42) (Tabela 4), denominado conjunto de teste.

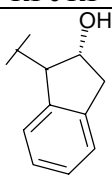
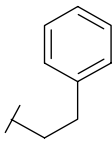
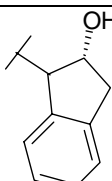
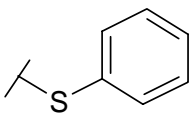
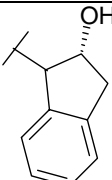
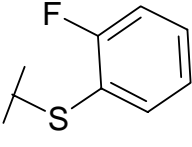
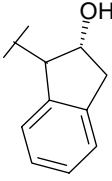
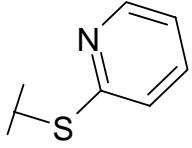
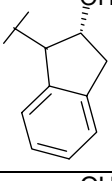
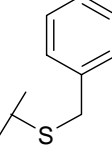
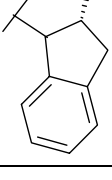
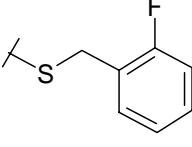
Esses compostos foram selecionados a partir de uma série de compostos a partir de uma série de inibidores peptídicos com diol vicinal da HIV-1 protease desenvolvidos por Samuelsson (Samuelsson et al., 2001a,b,c, 2003, 2005, 2006), onde a atividade biológica foi avaliada como a constante de atividade inibitória (*K_i*) da HIV-1 protease purificada a partir de *Escherichia coli* (Samuelsson et al., 2001a,b,c, 2003, 2005, 2006). Os valores de *K_i* expressos em nM foram convertidos para M e então transformados para $-\log K_i$ (*pK_i*). A faixa de *pK_i* dos compostos (conjunto de teste e treinamento) abrange quatro ordens de magnitude (5,69-9,24) em unidade logarítmica, e os valores de atividade biológica apresentam uma distribuição regular dentro dessa faixa. O gráfico da Figura 13 corresponde à distribuição dos valores de *pK_i* dos 46 compostos listados na Tabela 4.

O composto (2*R*,3*R*,4*R*,5*R*)-2,5-*bis*((2,4-difluorobenzil)oxi)-3,4-diidroxi-*N,N'*-*bis*((1*R*,2*S*)-2-hidroxi-2,3-diidro-1*H*-inden-1-il) hexanodiamino,

corresponde ao composto **14**, e foi usado como referência neste trabalho por ser conhecida sua estrutura cristalográfica complexada a HIV-1 protease.

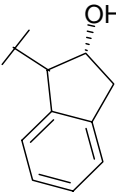
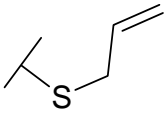
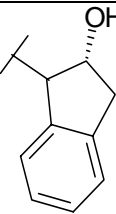
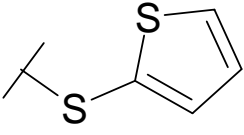
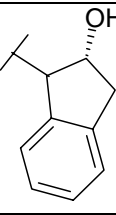
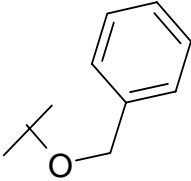
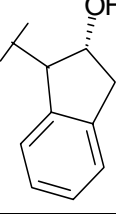
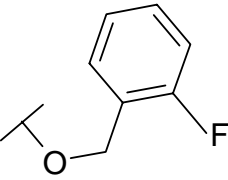
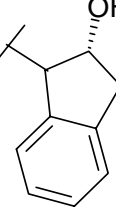
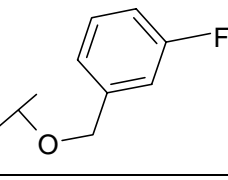
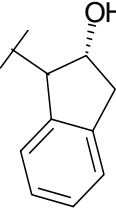
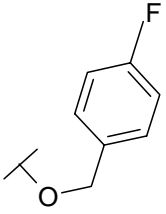
A divisão entre os compostos do conjunto de treinamento e do conjunto de teste foi feita separando-se os 46 compostos, estruturalmente relacionados, em quatro subgrupos de acordo com as faixas de valores de atividade biológica, e assim retirados dois compostos da primeira faixa, um da segunda, três da terceira e seis da quarta faixa, formando um conjunto de teste com doze compostos, o que corresponde a ~30% dos compostos totais. O Grupo I compreende compostos com valores de pK_i de 5,5 a 6,5; o Grupo II os compostos com valores de pK_i de 6,5 a 7,5; o Grupo III os compostos com valores de pK_i de 7,5 a 8,5; e o Grupo IV, os compostos com valores de pK_i de 8,5 a 9,5. O composto **14**, incluído no Grupo IV, não foi considerado na seleção randômica dos compostos do conjunto teste porque ele é o composto de referência do conjunto de treinamento.

TABELA 4 Estrutura dos 46 inibidores peptídicos com carbono vicinal central da HIV-1 protease com os respectivos valores de K_i (nM).

Inibidor	R1 e R1'	R2 e R2'	K_i (nM)
1			0,60
2			1,80
3			2,30
4			285
5			2,8
6			4,3

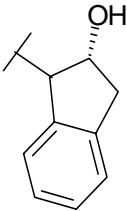
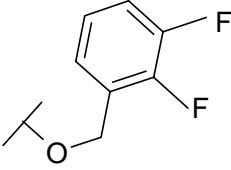
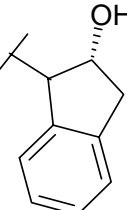
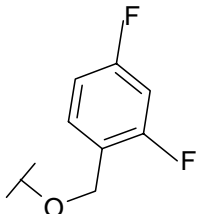
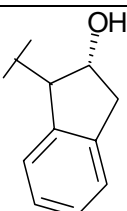
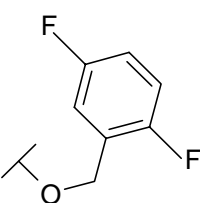
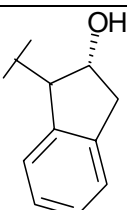
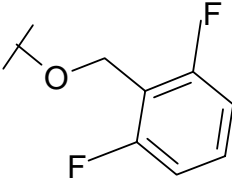
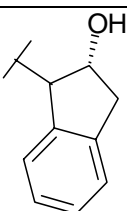
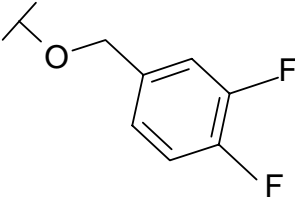
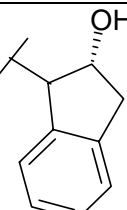
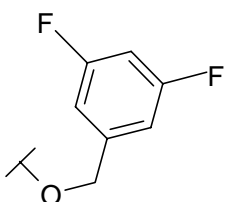
...continua...

TABELA 4, Cont.

7			16
8			0,5
9			1,22
10			3,26
11			7,13
12			5,39

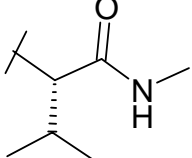
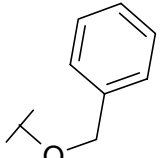
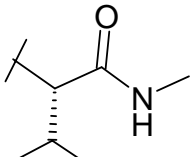
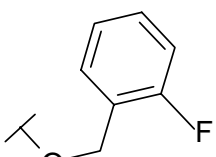
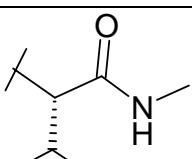
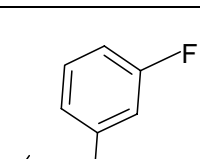
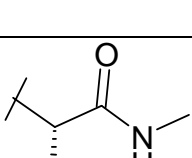
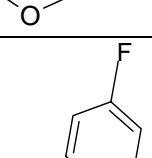
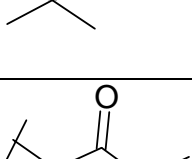
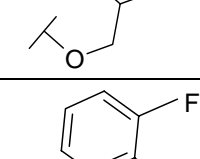
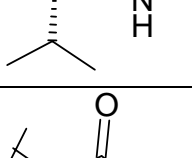
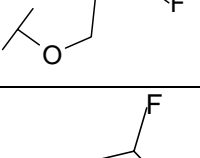
...continua...

TABELA 4, Cont.

13			4,00
14			1,65
15			3,29
16			1,64
17			3,90
18			5,00

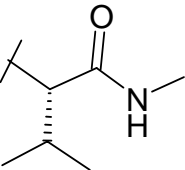
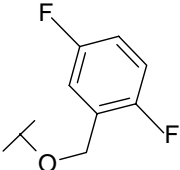
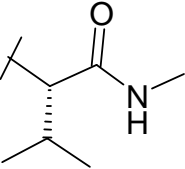
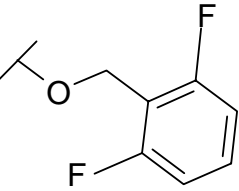
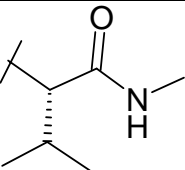
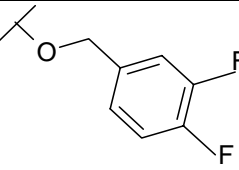
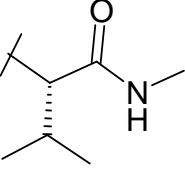
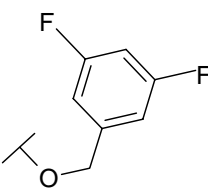
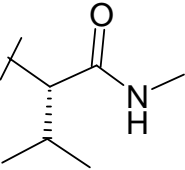
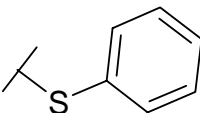
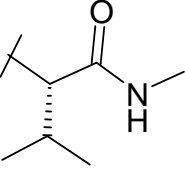
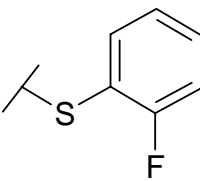
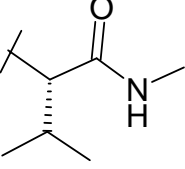
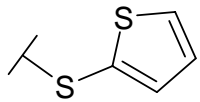
...continua...

TABELA 4, Cont.

19			0,80
20			1,92
21			1,83
22			1,79
23			1,28
24			1,12

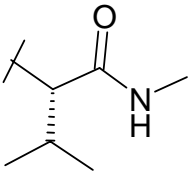
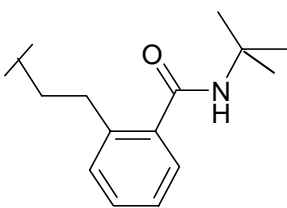
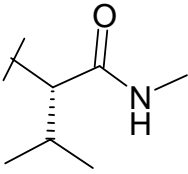
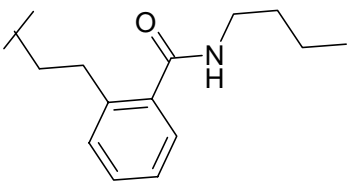
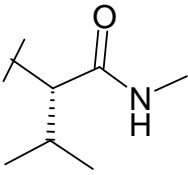
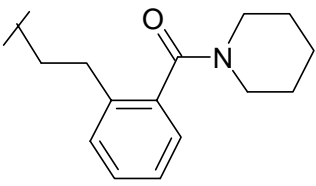
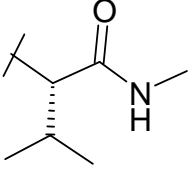
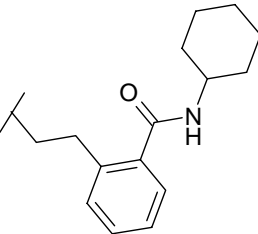
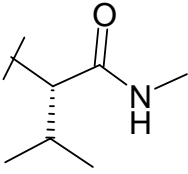
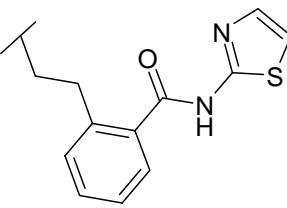
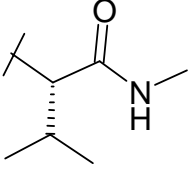
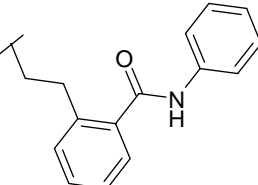
...continua...

TABELA 4, Cont.

25			1,98
26			1,73
27			1,98
28			1,98
29			1,80
30			0,57
31			1,00

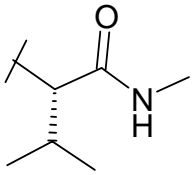
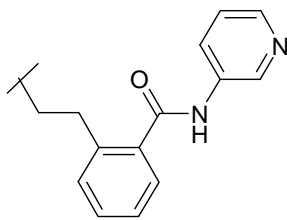
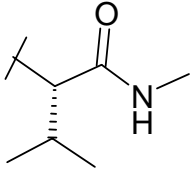
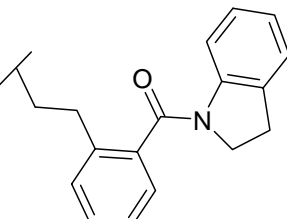
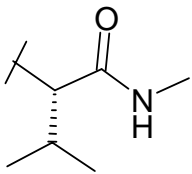
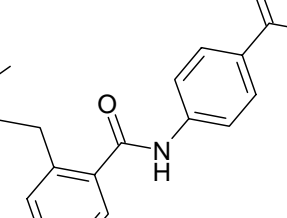
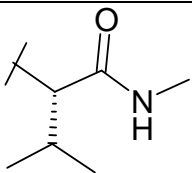
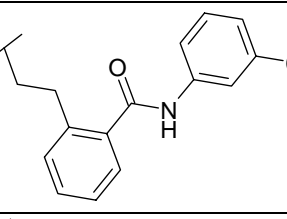
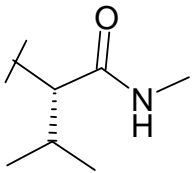
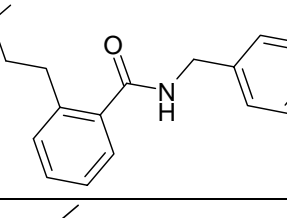
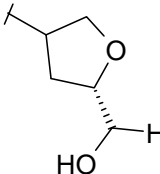
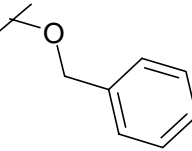
...continua...

TABELA 4, Cont.

32			700
33			600
34			400
35			200
36			700
37			7,00

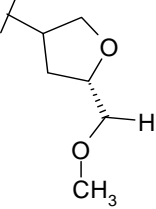
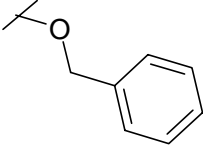
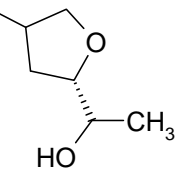
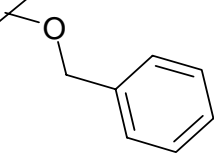
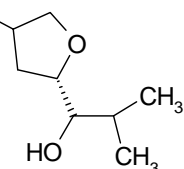
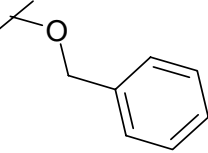
...continua...

TABELA 4, Cont.

38			70
39			800
40			200
41			6,0
42			2000
43			4,80

...continua...

TABELA 4, Cont.

44			103
45			208
46			1590

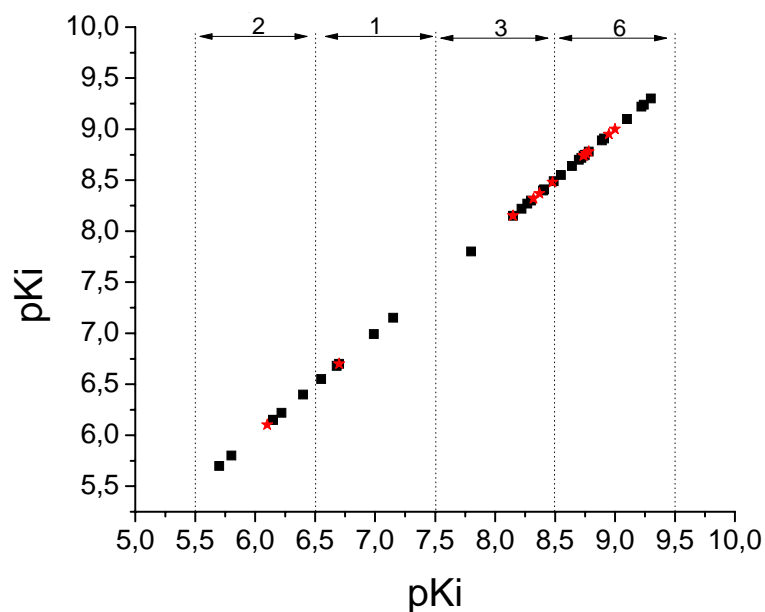


FIGURA 13 Gráfico da distribuição dos valores de constante de inibição (pK_i) dos 46 compostos em estudo. Os compostos de treinamento estão representados por (■) e os do conjunto de teste por estrela vermelha.

4.2 QSAR-4D

4.2.1 Módulo *Symposar*

O módulo *Symposar* é um software implementado no pacote de programas *Biograf^x*, que tem como alvo criar um conjunto de dados consistente com uma alta similaridade de propriedades físico-químicas no espaço 3D entre os ligantes, mas que permita configurações alternativas posteriores (conjunto de dados 4D) (Lill, 2006).

As moléculas dos ligantes são superpostas em uma ou várias moléculas modelo, primeiro com base na similaridade 2D da subestrutura, e subsequentemente, no espaço 3D, relacionado com a similaridade dos campos

físico-químicos. Estes dois processos combinam a velocidade da busca de similaridade 2D com a acurácia e autenticidade da proteína – interações de ligantes no espaço 3D. As moléculas são tratadas como flexíveis e são inteiramente relaxadas no fim do processo de alinhamento (Lill, 2006).

O esquema da Figura 14 mostra o procedimento geral do processo de alinhamento combinando buscas 2D e 3D de similaridades entre as estruturas. A molécula a ser alinhada deve ser representada pelos diferentes estereoisômeros e estados de protonação, é primeiramente alinhada no espaço 2D de uma ou mais moléculas modelo. Cada modelo pode ser representado pelos diferentes estados de protonação, estereoisômeros, conformações e orientações que representam o modo de ligação dos compostos modelos. Depois de alinhada a molécula no espaço 3D com base na estrutura modelo, a subestrutura é minimizada (Lill, 2006).

Com o intuito de refinar as soluções obtidas, uma pesquisa com o método de Monte Carlo subsequente a minimização é realizada, a fim de se identificar combinações ótimas entre as características farmacofóricas 3D e as moléculas modelos. Opcionalmente, a melhor solução 3D (máximo de similaridade com o melhor modelo de correspondência) pode ser utilizada como modelo para o posterior alinhamento de outras moléculas. Se esta opção for selecionada, as moléculas serão alinhadas na ordem de suas afinidades experimentais vinculadas à proteína alvo (Lill, 2006).

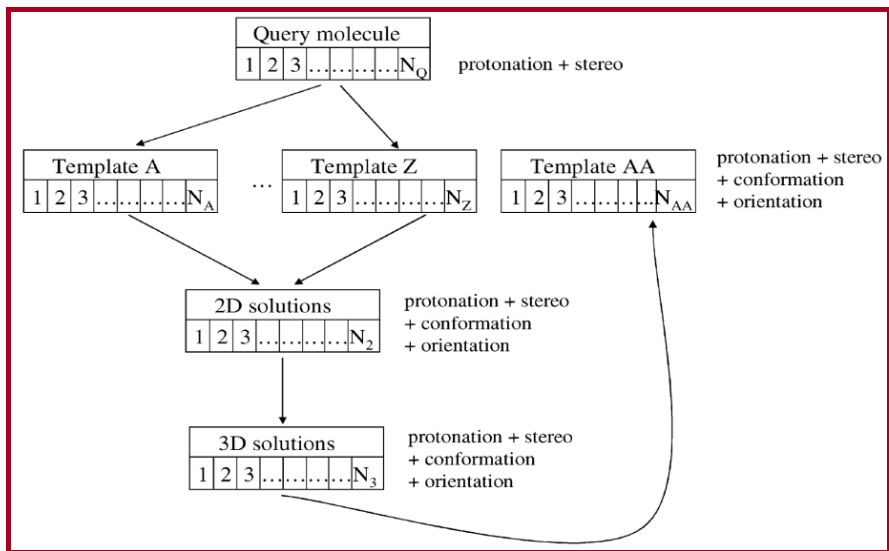


FIGURA 14 Procedimento geral do processo de alinhamento combinando buscas 2D e 3D de similaridades entre as estruturas no *Symposar*.

O campo de força direcional (Figura 15) que está incorporado no pacote de programas Biograf^x, que inclui os softwares *Symposar*, *Raptor* e *Quasar* - sendo de extrema importância para quantificar ligações hidrogênio e interações metal-ligante de forma adequada.

$$\begin{aligned}
E_{\text{total}} = & \sum_{\text{bonds}} K_r (r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_{\text{eq}})^2 + \\
& \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{\text{nb pairs}} \frac{q_i \cdot q_j}{4\pi\epsilon_0 D(r) r_{ij}} + \sum_{\text{nb pairs}} \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^6} \\
& + \sum_{\text{H-bonds}} \left(\frac{C}{r_{\text{H-Acc}}^{12}} - \frac{D}{r_{\text{H-Acc}}^{10}} \right) \cdot \cos^2(\theta_{\text{Don-H-Acc}}) \cdot \cos^n(\omega_{\text{H-Lig-LP}}) \\
& + \sum_{\text{metal-ligand pairs}} \frac{q_i^{\text{CT}} \cdot q_j^{\text{CT}}}{4\pi\epsilon_0 D(r) r_{ij}} + \sum_{\text{metal-ligand pairs}} \left(\frac{E}{r_{\text{M-Lig}}^{12}} - \frac{F}{r_{\text{M-Lig}}^{10}} \right) \\
& + (E_{\text{MC}} + E_{\text{LFS}}) \cdot \prod_{\text{indep. angles}} \cos^2(\psi_{\text{Lig-M-Lig}} - \psi_{\text{eq}}) \cdot \frac{1}{n} \sum_{\text{1st shell ligands}} \cos^n(\omega_{\text{M-Lig-LP}})
\end{aligned}$$

FIGURA 15 Campo de Força Direcional implementado no pacote de programas Biograf^x.

Adaptado de Vedani & Hutha (1990).

4.2.2 QSAR-5D

4.2.3 QSAR-5D – adição da quinta dimensão

A quinta dimensão refere-se à representação do receptor. A simulação do encaixe induzido é uma maneira explícita a simulação de uma adaptação topológica do modelo da superfície do sítio de ligação para moléculas individuais do ligante. Inclusive, a superfície do modelo do sítio de ligação pode ser ligeiramente encolhida ou expandida, dependendo do tamanho e da topologia da ligação do ligante. Como a identificação da magnitude correta e mecanismo do encaixe induzido não é possível na ausência de uma estrutura verdadeira da proteína alvo, diferentes protocolos (magnitude dependente de efeitos estéricos, eletrostáticos, ligação hidrogênio ou potencial lipofílico) de encaixe induzido

são presentes como cenários alternativos (QSAR-5D) no QSAR. Uma penalidade de energia é incluída proporcionalmente à magnitude do encaixe induzido. Além disso, atributos de ligação hidrogênio (doadoras ou aceptoras) no modelo de receptor são permitidas para girar, dependendo das propriedades da ligação hidrogênio da molécula do ligante (Lill, 2007).

Recentemente, foi desenvolvido um método que é capaz de simular anisotropicamente a simulação do encaixe induzido. Duas camadas espacialmente separadas permitem ambas, a simulação de adaptação da proteína local assim como grande mudanças conformacionais, tal como as envolvidas no agonismo e antagonismo em receptores nucleares. Variações na distribuição das propriedades entre a camada interna e externa representam as diferentes naturezas físico-químicas de diferentes sub-bolsas resultado do encaixe induzido do movimento de proteínas (Lill, 2007).

A simulação do encaixe induzido de inibidores peptídicos no sítio ativo poderá ser simulado adaptando uma superfície de van der Waals à topologia de cada molécula do ligante do conjunto de treinamento e de teste. Este procedimento mimetiza um encaixe local induzido, o que pode ser representado de forma isotrópica (linearmente), anisotrópica ou através da minimização de energia (Vedani & Dobler, 2002). Esta parte do trabalho será realizada com o programa RAPTOR 2.0 (Raptor 2.0 User's Guide).

4.2.4 O Conceito raptor

Raptor é uma abordagem de modelagem-receptor baseada na relação quantitativa estrutura-atividade multidimensional (QSARs). Destina-se a interpretar intuitivamente modelos de sítios de ligação de proteínas e predizer com acurácia energias livres relativas do sítio de ligação. Para alcançarmos este ultimo objetivo, é de extrema importância a simulação do encaixe induzido. O *Raptor* explicita e permite anisotropicamente a representação da dupla camada

do receptor. O encaixe induzido não está limitado aos aspectos estéricos, mas inclui a variação físico-química dos campos ao mesmo. A *scoring function* usada para avaliar interações ligante-receptor inclui um termo direcional para ligação hidrogênio, um termo para interações hidrofóbicas e assim, trata dos efeitos de solvatação implícita. Isto faz com que haja independência da abordagem do modelo de carga parcial. Como consequência, permite-se uma boa modelagem das moléculas do ligante no sítio de ligação do receptor com diferentes cargas líquidas (Raptor 2.0 User's Guide).

4.2.5 Score function

A função no programa Raptor compreende hidrofobicidade (ΔG_{HO}) e a ligação hidrogênio (ΔG_{HB}), bem como condições para o custo da adaptação topológica (ΔG_{IF}) e as mudanças na entropia ($T\Delta S$) sobre o ligante:

$$\Delta G = \Delta G_{const} + f_{HO}\Delta G_{HO} + f_{HB}\Delta G_{HB} + f_{IF}\Delta G_{IF} + f_{T\Delta S}\Delta G_{T\Delta S} \quad (4)$$

ΔG_{const} é uma contribuição de energia para a energia de ligação racionalizada como uma perda global da entropia translacional e rotacional do ligante ou como um ganho global de entropia devido à dessolvatação do sítio de ligação. f_{HO} , f_{HB} , f_{IF} e $f_{T\Delta S}$ são fatores escalares que são inerentes a um determinado modelo de receptor, eles são otimizados durante a simulação para cada fármaco-alvo específico e, normalmente, limitado a intervalos específicos (ex. $f_{HO} = 0,75-1,25$).

No Raptor, a topologia da molécula do ligante é representada pela sua superfície acessível ao solvente ($r_{probe}=1,4\text{\AA}$). Os campos (ligação hidrogênio e hidrofobicidade) gerados pelos átomos do ligante são projetados sobre esta superfície, mais precisamente, em pontos discretos localizados em posições equidistantes.

A adaptação de ambos, campo e topologia do receptor para cada ligante, é alcançado por uma combinação de ajuste estérico com a topologia do próprio ligante (Figura 16) e uma componente de atração ou repulsão entre ligante e receptor. Este último é obtido pela correlação de suas propriedades físico-químicas (hidrofobicidade e propensão a ligação hidrogênio) no espaço 3D.

Durante o processo de adaptação estérica, os campos gerados pelo sítio de ligação da proteína para o ligante da superfície acessível ao solvente (SAS) são calculados por interpolação linear entre a camada interna e externa da dupla camada, se o ligante SAS tem porções localizadas no local (setas tracejadas na Figura 16). Para pontos da superfície localizados dentro da camada interna (ligantes muito pequenos), este último pode adaptar apenas em parte, para a topologia do ligante, onde a magnitude do encaixe induzido é proporcional à distância e orientação dos vetores normais destas superfícies (setas sólidas na Figura 16) (Raptor User's Guide).

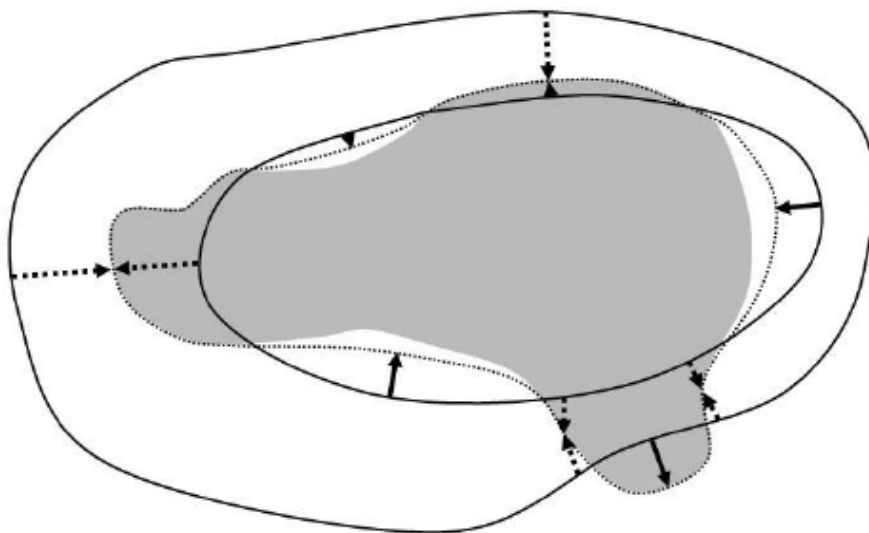


FIGURA 16 Esboço da molécula do ligante (representado seu SAS; superfície cinza) na representação da dupla camada do receptor (linhas sólidas). Durante o processo de adaptação estérica, os campos gerados pelo sítio de ligação da proteína para o ligante SAS são calculados por interpolação linear entre interior e exterior da camada, se o ligante SAS ficar entre as duas camadas (setas tracejadas). Para pontos localizados na superfície dentro da camada interna, este último pode adaptar apenas em parte para a topologia do ligante (flechas sólidas; linha pontilhada). Adaptado de Lill et al. (2004).

4.2.6 Algoritmo de otimização

Para obter a melhor relação entre os valores experimentais e calculados para os ligantes do conjunto de treinamento, o algoritmo subjacente no *Raptor* otimiza a natureza e distribuição das unidades mapeadas na dupla camada representando o receptor. Essas propriedades incluem a hidrofobicidade e propensão à ligação hidrogênio, as variáveis que descrevem o processo de adaptação, bem como o *scaling factor* para associar a energia com o encaixe

induzido. Como critério de qualidade no processo de otimização, a seguinte expressão é minimizada:

$$Q = \sum_{\text{Ligantes}} (\Delta G_{\text{calc}} - \Delta G_{\text{exp}})^4 \quad (5)$$

A quarta potência foi deliberadamente escolhida, por ela ter gerado pequenos desvios (que normalmente se aproximam de incerteza experimental).

Em detalhe, os campos hidrofóbicos e de ligação hidrogênio são projetados para a dupla camada com um peso ponderado pelo ΔG do correspondente ligante. Os centros de domínio são atribuídos à posição de manifestações inequívocas da hidrofobicidade e propensão à ligação hidrogênio. A atribuição de cada ponto na camada para o domínio é decidida através da comparação de seu campo com aquele experimentado no centro do domínio.

4.3 QSAR-6D

4.3.1 O conceito Quasar

Um possível sítio de ligação quase-atomístico refere-se a um elevado nível de abstração do modelo. As informações essenciais sobre o hipotético receptor é fornecida por meio de um envelope 3D, que rodeia as moléculas do ligante com uma superfície de Van der Waals, mapeando-a. A topologia desta superfície 3D do sítio de ligação inclui a representação das propriedades mapeadas de interesse, como a hidrofobicidade, cargas parciais, potencial eletrostático, e propensão à ligação hidrogênio.

O software *Quasar*, receptor quase-atomístico, é uma ferramenta de modelagem QSAR-6D: a quarta dimensão refere-se à possibilidade de representar cada molécula por um conjunto de conformações, orientações, estados de protonação e tautômeros - reduzindo assim os preconceitos associados com a escolha de conformações bioativas; a quinta dimensão refere-

se à possibilidade de considerar um conjunto de diferentes modelos de encaixe induzido; a sexta dimensão permite a avaliação simultânea de diferentes modelos de solvatação. Além disso, o *Quasar* permite a simulação de encaixe induzido local, ligação hidrogênio flip-flop, e vários efeitos de solvatação. Interações ligante-receptor são estimadas com base em um campo de força direcional. Uma família de modelos “quasi-atomística” de receptor é então gerada usando um algoritmo genético combinado com a ponderação da validação cruzada (Quasar User’s Guide).

4.3.2 Geração de uma família inicial de estruturas parentais

As propriedades quase-atomísticas são mapeadas para distribuir igualmente pontos individuais na superfície (Tabela 5).

TABELA 5 Propriedades Atomísticas Mapeadas para a Superfície do Receptor no *Quasar*.

Partícula	Potencial tipo não-ligado ¹	Carga Elétrica	Função não ligada de profundidade
Hidrofóbica, neutra	6/12	-	-0.024 kcal/mol ²
Hidrofóbica, positiva	6/12 + eletrostático	+0.10	-0.024 kcal/mol ²
Hidrofóbica, negativa	6/12 + eletrostático	-0.10	-0.024 kcal/mol ²
Ligação H doadora	10/12	-	-2.5/-2.0/-1.2 kcal/mol ³
Ligação H aceptora	10/12	-	-2.5/-2.0/-1.2 kcal/mol ³

...continua...

TABELA 5, Cont.

Ponte Salina, positiva	10/12 + eletrostático	+0.25	-2.5/-2.0/-1.2 kcal/mol ³
Ponte Salina, negativa	10/12 + eletrostático	-0.25	-2.5/-2.0/-1.2 kcal/mol ³
Ligação H flip- flop ⁴	10/12	-	-2.5/-2.0/-1.2 kcal/mol ³
Superfície Solvente	simétrico 10/12 ⁵	-	-0.97/-0.80/-0.46 kcal/mol ^{3,6}
Vácuo	-	-	-

1. Os valores de i, j referem-se aos coeficientes de atração e repulsão das funções potenciais não ligadas usadas para a interação ligante-receptor. A forma geral deste potencial é: $E(r)=A/r^i-C/r^j$.

2. A função adaptada tem forma geral $E(r)=A/r^{12}-C/r^6$. Os coeficientes A e C são calculados de acordo com, $A=-\varepsilon(r_i+r_j)^{12}$ e $C = -2\varepsilon(r_i+r_j)^6$, respectivamente, e $\varepsilon=(\varepsilon_i \cdot \varepsilon_j)^{1/2}$. Os números dados representam ε_j, r_j e r_i correspondente a energia de van der Waals do raio de dois átomos interagindo.

3. Valores para interações $-O-H\cdots Y$, $>N-H\cdots Y$, and $-S-H\cdots Y$ ligação hidrogênio, respectivamente, onde “Y” denota a ligação hidrogênio aceptora. Valores idênticos são usados para $X\cdots O$, $X\cdots N$, e arranjo $X\cdots S$, onde “X” denota a ligação hidrogênio doadora.

4. Partículas com ligação hidrogênio flip-flop podem adaptar suas propriedades (ligação hidrogênio doadora ou aceptora) para cada molécula ligante dentro dos limites farmacofóricos, dependendo do seu grupo de interação funcional.

5. Para evitar forças repulsivas entre a superfície do solvente e qualquer molécula do ligante, um potencial simétrico 10/12 (espelhado em $r = r_0$) é usado. Isto representa uma possível aproximação a um solvente móvel.

6. Como as partículas virtuais são diferentes em raio superior a uma molécula de água, a energia associada deve ser corrigida para diferentes volumes: $E =$

$(2r_{vp}/2,75)^3 E_0$; por exemplo, para $r_{vp}=0,8 \text{ \AA}$, $E=0,197E_0$. O $2,75\text{\AA}$ correspondem à média da distância de O-H...O na ligação hidrogênio.

Se houver indícios experimentais para uma cavidade do receptor acessível ao solvente, partes da superfície do receptor podem ser permitidas ao representar o solvente. Alternativamente, as regiões podem evoluir dinamicamente a se tornar "v\u00e1cuo", isto \u00e9, n\u00e3o povoadas por qualquer propriedade. Os melhores resultados s\u00e3o normalmente obtidos com uma popula\u00e7\u00e3o inicial de 200-500 modelos gerados randomicamente.

4.3.3 Estimativa da energia livre relativa do s\u00edtio de liga\u00e7\u00e3o

O conceito de Vedani et al. (1995, 1998) e Zbinden et al. (1998), engloba a abordagem combinada de Blaney et al. (1982) com o m\u00e9todo de Still et al. (1990) para estimar a energia de solvata\u00e7\u00e3o do ligante, e um termo para corrigir a perda de entropia acerca da liga\u00e7\u00e3o do receptor seguindo Searle & Williams (1992):

$$E_{bdg} \approx E_{lig-rec} - E_{solv.,lig} - T\Delta S_{bdg} + \Delta E_{int.,lig} + E_{env.adapt.,lig} \quad (6)$$

$E_{lig-rec} < 0$; $T\Delta S_{bdg} < 0$; $E_{solv.,lig} > 0$ (ligantes hidrof\u00f3bicos), < 0 (ligantes hidrof\u00edlicos); $\Delta E_{int.,lig} > 0$; $E_{env.adapt.,lig} > 0$

O termo $E_{lig-rec}$ corresponde \u00e0 contribui\u00e7\u00e3o ent\u00e1lpica para a intera\u00e7\u00e3o ligante-receptor e \u00e9 determinado por meio de um campo de for\u00e7a direcional (Vedani & Huhta, 1990; Vedani et al., 1995, 1998). O termo $E_{solv.,lig}$ corresponde \u00e0 energia necess\u00e1ria para tirar o solvente do ligante quando s\u00e3o necess\u00e1rias liga\u00e7\u00f5es de um ambiente aquoso para a cavidade hidrof\u00f3bica do receptor. Este termo \u00e9 calculado usando um algoritmo desenvolvido e validado por Styll et al. (1990). A aproxima\u00e7\u00e3o de Blaney \u00e9 baseada na suposi\u00e7\u00e3o de que todos os

ligantes são igualmente ancorados dentro do receptor e, portanto, diferenças na energia de solvatação do complexo ligante-receptor são insignificantes. Para sistemas onde os ligantes expõem uma fração diferente de sua superfície para a acessibilidade do solvente em parte do sítio de ligação, é possível definir a estática ou a dinâmica da acessibilidade do solvente à região e, assim, corrigir esta situação. $T\Delta S_{bdg}$ é estimado pela designação de uma quantia de 0,7 kcal/mol para toda ligação livremente rotacional, excluindo grupos terminais $-CH_3$. O termo $\Delta E_{int.,lig}$ conta para um aumento da energia potencial interna do ligante, enquanto vinculados ao receptor (relativo a uma conformação de referência com tensão-livre em solução aquosa). Esta correção parece necessária, já que a energia interna de uma molécula do ligante pode aumentar, enquanto sua interação com o receptor é maximizada.

$E_{env.adapt.,lig}$ está associado com a absorção da energia mediante modificação da envelope médio do receptor para o envelope individual do receptor. Quando a representação da múltipla-conformação (ou múltipla orientação) é utilizada, a contribuição de uma entidade individual para o total de energia é calculado por um fator Boltzmann normalizado:

$$E_{bdg,tot} = \sum E_{bdg,ind} \cdot \exp(-w_i \cdot E_{bdg,ind,menor}) \quad (7)$$

onde, $w_i = (E_{bdg,ind}/E_{bdg,ind,menor})^{-1}$ é o fator de normalização

Energia livre do sítio de ligação, ΔG_{pred} , são então preditas por meio de uma regressão linear entre ΔG_{exp} e E_{bdg} usando as moléculas de ligante do conjunto de treinamento:

$$\Delta G_{pred} = |a| \cdot E_{bdg} + b \quad (8)$$

O declive e o intercepto da equação acima são inerentes a um determinado modelo de receptor, sendo posteriormente aplicados para prever a energia relativa da ligação das diferentes moléculas do ligante estabelecidas no conjunto de treinamento. Em contraste com outros métodos, Vedani e colaboradores calibraram cada sistema de receptor com um conjunto de treinamento (Vedani et al., 1995, 1998; Zbinden et al., 1998), em vez de aplicar uma função universal para os vários sistemas de proteína-ligante.

4.3.4 Análise da família modelo

Um critério obrigatório para validar uma família de receptores modelos é a sua habilidade para prever sua energia livre relativa do sítio de ligação para um conjunto de teste externo de moléculas ligantes, não utilizados durante o modelo de construção. Um desafio mais sério é a chamada corrida teste (do inglês, *scramble test*) (Rogers & Hopfinger, 1994). Aqui, os dados de ΔG do conjunto de treinamento são aleatoriamente misturados com relação aos verdadeiros valores biológicos e, a simulação é repetida em outras condições idênticas. Se os ligantes do conjunto de teste são preditos similarmente e corretamente comparados com a verdadeira simulação utilizando decodificadores dos dados de ΔG , o modelo é inútil, já que não é sensível aos dados biológicos como deveria ser para um QSAR.

Outros critérios incluem o valor de q^2 - validação cruzada, a falta de ajuste (do inglês, *lack-of-fit*) para os ligantes do conjunto de treinamento, a variação de ΔG_{pred} ao longo de todos os membros do modelo da família, bem como a uniformidade da distribuição das unidades mapeadas para o envelope do receptor. O teste final, porém, é a predição e verificação experimental da atividade dos novos compostos.

5 RESULTADOS E DISCUSSÃO

5.1 QSAR-4D – módulo symposar

O módulo Symposar é um software implementado no pacote de programas Biograf^x, que tem como alvo criar um conjunto de dados consistente com uma alta similaridade de propriedades físico-químicas no espaço 3D entre os ligantes, mas que permita configurações alternativas posteriores (conjunto de dados 4D). O módulo symposar é um precursor do QSAR 5 e 6D, dessa forma o resultado obtido nesta etapa está apresentado na Figura 17, onde o conjunto de conformações e orientações dos compostos estão alinhados na melhor conformação obtida gerada pelo programa. Os alinhamentos realizados geraram um máximo de duas conformações por ligante, com um total de 84 conformações geradas.

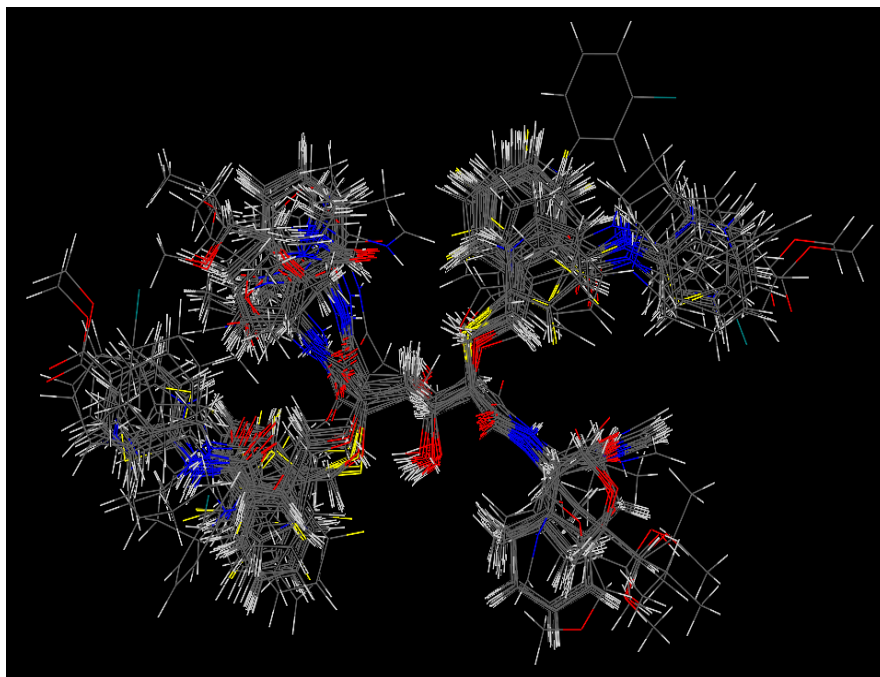


FIGURA 17 Estereorepresentação do alinhamento obtido para a série de compostos em estudo no *Symposar*.

5.2 QSAR-5D

5.2.1 Análise dos modelos de QSAR-5D

O conceito Raptor foi aplicado para a predição de afinidades de ligações associadas aos inibidores peptídicos com diol vicinal central (34 compostos no conjunto de treinamento e 12 compostos no conjunto de teste).

O intervalo de valores de pK_i dos 46 compostos abrangem quatro ordens de magnitude (5,69 a 9,24) em escala logarítmica. O programa *Symposar* gerou um conjunto de conformações e orientações para cada ligante e possíveis alinhamentos espaciais em 3D. Com base nesses alinhamentos foram realizados os cálculos de QSAR-5D usando o programa *Raptor*.

Os valores de pK_i predito e experimental, bem como *factor off* ($K_{i_{exp}}/K_{i_{calc}}$ para $K_{i_{exp}}/K_{i_{calc}} > 1.0$ e $K_{i_{calc}}/K_{i_{exp}}$, caso contrário) estão apresentados na Tabela 6 para o conjunto de treinamento e o conjunto teste. Observa-se que para o conjunto de treinamento o máximo de desvio de qualquer ligante a partir do valor experimental é um *factor off* de 3,084, e para o conjunto de teste, o desvio máximo de qualquer ligante a partir do valor experimental é um *factor off* de 16,205.

TABELA 6 Valores de pK_i experimental e predito para o conjunto de treinamento e conjunto de teste obtidos no QSAR-5D.

Ligantes	Experimental	Predito	Fator off
<i>Conjunto de Treinamento</i>			
1	9,22	8,96	1,831
3	8,64	8,76	1,331
4	6,55	7,01	2,983
5	8,55	8,61	1,129
7	7,80	7,49	2,045
8	9,30	8,99	2,053
9	8,91	8,56	1,130
11	8,15	8,21	1,161
12	8,27	8,34	1,176
15	8,48	8,50	1,052
16	8,78	9,03	1,770
17	8,41	8,46	1,135
18	8,30	8,44	1,367
19	9,10	8,82	1,885
20	8,72	8,77	1,125
21	8,74	8,86	1,344
22	8,75	8,76	1,032
24	8,95	9,09	1,375
26	8,76	8,96	1,567
27	8,70	9,01	2,024
29	8,74	8,68	1,152
31	9,00	9,29	1,959

...continua...

TABELA 6, Cont.

32	6,15	6,52	2,324
33	6,22	5,74	3,059
35	6,70	7,01	2,051
36	6,15	6,63	2,994
37	8,15	8,09	1,146
39	6,10	6,51	2,617
40	6,70	6,95	1,783
41	8,22	8,33	1,276
43	8,32	8,04	1,992
44	6,99	6,94	1,109
45	6,68	6,80	1,329
46	5,80	6,29	3,084
<i>Conjunto de Teste</i>			
2	8,74	8,74	1,016
6	8,37	8,54	1,498
10	8,49	8,13	2,251
13	8,40	8,08	2,080
14	8,78	8,71	1,175
23	8,89	8,77	1,326
25	8,70	9,01	2,027
28	8,70	9,18	2,979
30	9,24	8,50	5,505
34	6,40	6,50	1,276
38	7,15	6,75	2,526
42	5,70	6,91	16,205

A Figura 18 apresenta o gráfico dos valores de pK_i experimental e predito para os dois conjuntos, treinamento e teste. Uma família de 15 modelos de receptores foi criada resultando em um valor de q^2 (coeficiente de correlação da validação cruzada – *leave one out*) de 0,937 e $R^2_{\text{Treinamento}}$ de 0,944 para o conjunto de treinamento e um valor de R^2_{Teste} de 0,793 para o conjunto de teste. Um valor de q^2 , inferior 0,5 demonstra que o modelo “quase-atomístico” gerado possa não predizer uma boa identificação das interações hidrofóbicas/ligação de hidrogênio existentes, pelo menos não em sua totalidade.

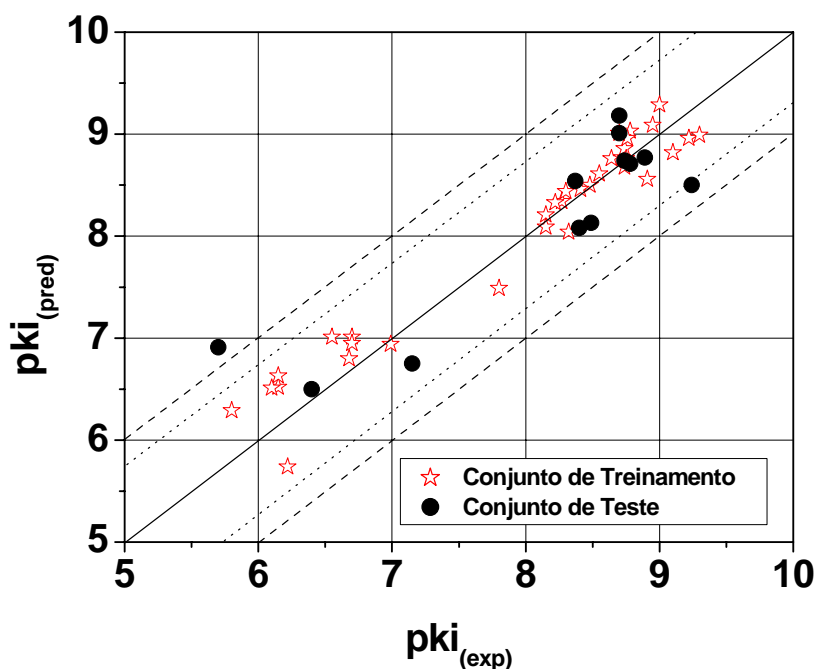


FIGURA 18 Gráfico dos valores de pK_i experimental e predito para os valores obtidos no QSAR-5D.

A estereorepresentação da dupla camada obtida no *Raptor* (QSAR-5D) que melhor se adequa a série de ligantes peptídicos é descrita na Figura 19. A figura mostra todos os inibidores com suas conformações dentro da dupla camada do receptor, que é representada por uma superfície colorida que corresponde à hidrofobicidade e/ou pretensão à ligação hidrogênio. A escala de hidrofobicidade da superfície pode variar de marrom (muito hidrofóbico) para amarelo (hidrofóbico) e cinza (neutro ou hidrofóbico). As regiões hidrofílicas estão representadas em vermelho (aceptor de ligação de hidrogênio) e azul (doador de ligação de hidrogênio).

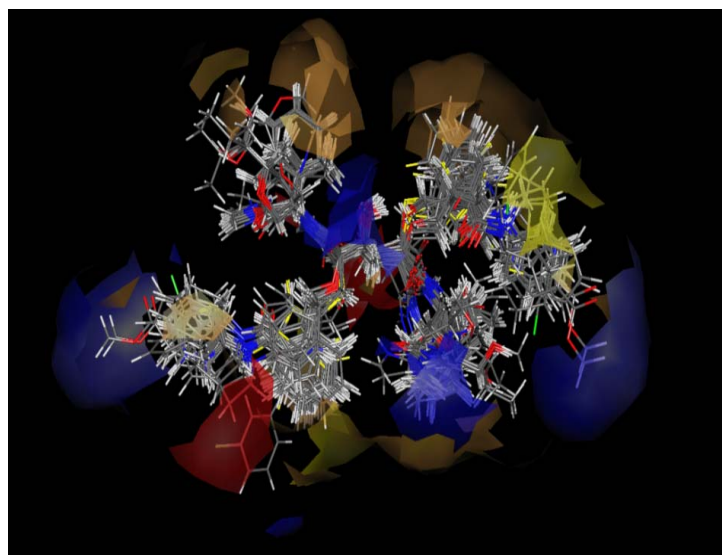


FIGURA 19 Estereorepresentação dos modelos do receptor obtidos no encaixe induzido pela representação da dupla camada mapeada de acordo com as propriedades físico-químicas. As cores significam: castanho = muito hidrofóbico; amarelo = hidrofóbico; cinza = neutro ou hidrofóbico; vermelho = acceptor de ligação de hidrogênio; azul = doadora de ligação de hidrogênio.

Com o intuito de melhor entendermos o modelo obtido, destacamos os inibidores que possuem maior (8) e menor (43) valor de K_i , representados na Figura 20, pela dupla camada, interna e externa. Esta escolha é feita, para melhor observarmos que a superfície hipotética gerada é a mesma para todos os compostos, e assim, podemos comparar e melhor visualizar os substituintes que favorecem a atividade inibitória como o composto 8, ou então, que desfavorecem a atividade inibitória como o composto 43.

Observa-se na camada interna, que as regiões do modelo do receptor próximas aos substituintes em R2/R2' (ver Tabela 4) são predominantemente hidrofóbicas. Próximo ao diol vicinal central encontra-se uma região acceptora de

ligação hidrogênio, presente também em R1', indicando que nesta região devemos ter um substituinte doador de ligação hidrogênio, e em R1 temos representada uma região hidrofóbica. Encontra-se também no corpo central da estrutura uma região vermelha, relativamente próxima aos oxigênios presentes no diol, que indica uma região acceptora de ligação hidrogênio. Na camada externa, representada por pontos, observa-se que regiões vermelhas em cada substituinte R1/R1' e R2/R2', indicando que estes dependendo do substituinte, podem ser regiões acceptoras de ligações hidrogênio. Destaca-se também, pontos marrons no corpo central, indicando regiões hidrofóbicas.

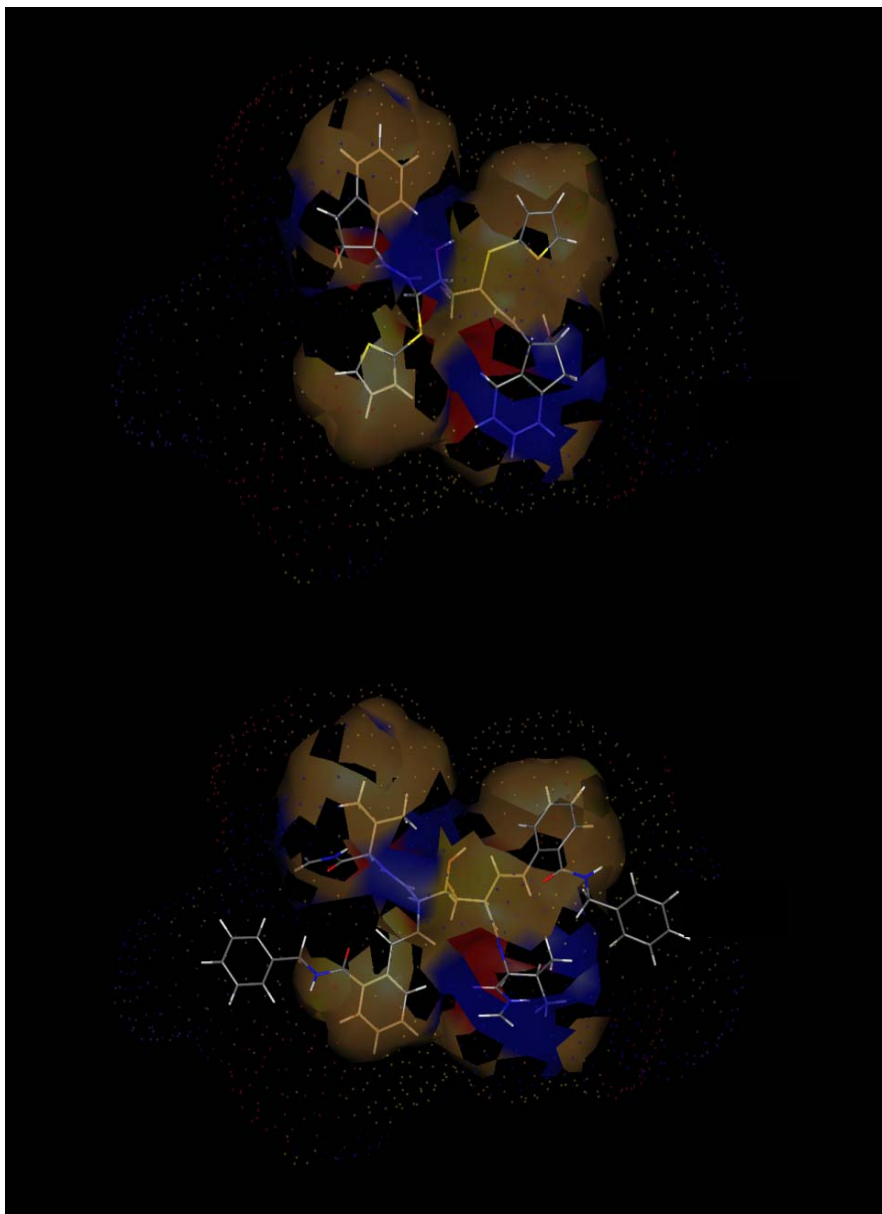


FIGURA 20 Estereorepresentação da dupla camada para os inibidores 8 (A) e 43 (B). A camada interna representada em *surface* e a externa em pontos.

Para identificarmos potenciais sítios e funcionalidades permitidas a um aumento da afinidade de ligação, a funcionalidade de grupos individuais de ambos, conjunto de treinamento e teste, são analisados pela sua contribuição à energia livre ao sítio de ligação, ΔG . A Tabela 7 lista todos os detalhes entálpicos (eletrostático, van der Waals e ligação hidrogênio), entrópicos, solvatação, e contribuições do encaixe-induzido para o cálculo da afinidade do sítio de ligação. Podemos observar na tabela, que o valor mais favorável obtido para a energia de ligação hidrogênio foi para o composto 24, que é um composto difluoro-substituído em R2/R2' e possui também em R1/R1' átomos de N e O com pares de elétrons livres para a realização da ligação hidrogênio. O valor menos favorável obtido para a série de compostos foi para o composto 4, que apresenta em R2/R2' um átomo de enxofre ligado a um anel piridínico e em R1/R1' um 2,3-diidro-1*H*-inden-2-ol ligado a uma hidroxila, sendo menos propenso a ligação de hidrogênio, como também pode ser observado pelos seus valores experimentais.

TABELA 7 Detalhes entálpicos, entrópicos, solvatação, e contribuições do encaixe-induzido para o cálculo da afinidade do sítio de ligação.

Ligante	$\Delta G(\text{exp})$	$\Delta G(\text{pred})$	$\Delta \Delta G$	$\Delta G(\text{IF})^*$	$T\Delta S^*$	Hidrofobicidade	Lig, H	$st\Delta v$
1	-12,362	-12,010	-0,35	0,815	3,415	-8,067	-9,845	0,039
2	-11,722	-11,713	-0,01	0,293	2,775	-7,286	-9,166	0,375
3	-11,579	-11,745	0,17	0,418	2,775	-7,589	-9,022	0,182
4	-8,774	-9,393	0,62	0,259	2,775	-6,514	-7,585	0,065
5	-11,465	-11,536	0,07	1,276	3,628	-7,337	-10,776	0,235
6	-11,215	-11,450	0,24	1,238	3,628	-7,487	-10,502	0,444
7	-10,450	-10,034	-0,42	1,075	3,628	-5,660	-10,750	0,075
8	-12,468	-12,049	-0,42	0,027	2,775	-7,446	-9,078	0,341
9	-11,948	-11,877	-0,07	0,922	3,628	-7,820	-10,280	0,264
10	-11,376	-10,904	-0,47	1,060	3,628	-7,143	-10,121	0,276
11	-10,921	-11,008	0,09	1,085	3,628	-7,273	-10,120	0,193
12	-11,084	-11,178	0,09	1,104	3,628	-7,181	-10,402	0,293
13	-11,257	-10,831	-0,43	1,043	3,628	-7,206	-9,968	0,206
14	-11,773	-11,679	-0,09	1,024	3,628	-7,739	-10,265	0,378
15	-11,371	-11,400	0,03	1,119	3,628	-7,592	-10,228	0,243
16	-11,776	-12,109	0,33	0,863	3,628	-7,898	-10,373	0,048
17	-11,272	-11,346	0,07	1,149	3,628	-7,487	-10,308	0,185
18	-11,127	-11,309	0,18	1,161	3,628	-7,512	-10,259	0,158
19	-12,194	-11,825	-0,37	2,065	4,055	-7,040	-12,577	0,034
20	-11,684	-11,753	0,07	2,092	4,055	-7,156	-12,416	0,109
21	-11,712	-11,884	0,17	2,148	4,055	-7,362	-12,397	0,092
22	-11,725	-11,744	0,02	2,160	4,055	-7,182	-12,449	0,100

..continua...

TABELA 7, Cont.

23	-11,921	-11,757	-0,16	2,116	4,055	-7,424	-12,176	0,176
24	-11,998	-12,184	0,19	2,172	4,055	-7,505	-12,579	0,143
25	-11,667	-12,078	0,41	2,197	4,055	-7,662	-12,341	0,128
26	-11,745	-12,007	0,26	2,132	4,055	-7,402	-12,464	0,112
27	-11,667	-12,078	0,41	2,231	4,055	-7,632	-12,405	0,076
28	-11,667	-12,302	0,64	2,304	4,055	-7,918	-12,415	0,144
29	-11,722	-11,640	-0,08	1,541	3,202	-6,426	-11,629	0,216
30	-12,391	-11,398	-0,99	1,688	3,202	-6,680	-11,281	0,431
31	-12,064	-12,455	0,39	1,337	3,202	-6,525	-12,141	0,220
32	-8,250	-8,741	0,49	4,585	3,842	-8,405	-10,435	0,215
33	-8,340	-7,689	-0,65	5,414	5,123	-8,475	-11,422	0,074
34	-8,576	-8,718	0,14	5,477	4,055	-8,545	-11,378	0,732
35	-8,980	-9,398	0,42	5,166	4,269	-8,171	-12,334	0,165
36	-8,250	-8,888	0,64	4,663	4,055	-7,963	-11,316	0,281
37	-10,931	-10,852	-0,08	4,603	3,628	-8,328	-12,427	0,227
38	-9,591	-9,052	-0,54	4,881	4,055	-7,824	-11,836	0,610
39	-8,173	-8,733	0,56	5,378	3,628	-8,180	-11,232	0,252
40	-8,980	-9,317	0,34	4,455	3,415	-7,232	-11,627	0,149
41	-11,021	-11,163	0,14	2,823	3,628	-8,388	-10,898	0,198
42	-7,639	-9,261	1,62	5,080	4,055	-8,649	-11,419	0,742
43	-11,151	-10,771	-0,38	2,225	4,055	-4,714	-14,009	0,173
44	-9,366	-9,306	-0,06	1,967	4,055	-5,307	-11,694	0,219
45	-8,957	-9,123	0,17	2,780	4,055	-5,186	-12,445	0,373
46	-7,773	-8,429	0,66	2,503	4,842	-5,381	-11,706	0,109

* $\Delta G_{(HO)}$ = hidrofobicidade, $\Delta G_{(HB)}$ = ligação hidrogênio, $\Delta G_{(TF)}$ = custo da adaptação topológica e TAS = mudança na entropia.

O QSAR-5D também gera como resultados todos os coeficientes necessários para calcularmos os valores de ΔG dos modelos gerados (Tabela 8). Podemos observar na tabela que, dentre os modelos gerados, o modelo que apresentou o maior valor do coeficiente de validação da correlação cruzada foi o modelo 6 com um q^2 igual a 0,937. Assim, se necessário podemos calcular os valores de energia para cada um dos inibidores de cada modelo gerado.

TABELA 8 Coeficientes das equações geradas para os modelos de QSAR-5D pelo software Raptor.

$$\Delta G = \Delta G_{const} + f_{HO}\Delta G_{HO} + f_{HB}\Delta G_{HB} + f_{IF}\Delta G_{IF} + f_{TAS}\Delta G_{TAS}$$

Modelo	ΔG_{const}	f_{HO}	f_{HB}	f_{IF}	f_{TAS}	q^2	R2	r2 preditivo	Máx. Desvio
1	1.386	0.823	1.245	1.145	0.303	0.925	0.934	0.683	2.628
2	1.589	0.768	1.249	1.032	0.302	0.916	0.927	0.749	4.677
3	1.987	0.792	1.147	1.058	0.300	0.907	0.936	0.760	6.301
4	1.861	0.788	1.248	1.338	0.304	0.934	0.946	0.808	4.088
5	1.993	0.750	1.245	1.225	0.304	0.903	0.925	0.648	5.487
6	1.958	0.780	1.188	1.208	0.301	0.937	0.942	0.324	2.141
7	1.663	0.757	1.161	1.194	0.300	0.927	0.922	0.747	2.508
8	1.514	0.754	0.859	1.188	0.301	0.931	0.927	0.548	2.747
9	1.919	0.751	1.101	1.154	0.301	0.928	0.935	0.879	3.061
10	1.972	0.762	1.219	1.175	0.301	0.935	0.936	0.860	2.700
11	0.327	0.773	1.035	1.290	0.301	0.933	0.928	0.660	2.525
12	1.995	0.751	1.124	1.087	0.352	0.920	0.923	0.749	2.884
13	1.399	0.761	1.241	1.259	0.301	0.921	0.921	0.658	2.476
14	1.870	0.778	1.250	1.183	0.302	0.933	0.934	0.659	2.499
15	1.649	0.833	1.175	1.062	0.301	0.932	0.927	0.717	2.472

O modelo do receptor obtido foi validado através do *scramble test* (Figura 21) tendo como resultado os valores de $R^2_{\text{Treinamento}}$ de 0,652 para o conjunto de treinamento. *Scramble test* consiste em randomizar os valores de atividade dos inibidores em relação à estrutura e repetir os cálculos para obtenção do modelo do receptor. Se o valor de q^2 for inferior que o encontrado sem o *scramble test*, o modelo é aceito; por outro lado, se o valor de q^2 for comparável ao obtido na simulação sem o *scramble test*, o modelo do receptor obtido não é sensível aos dados experimentais. Dessa forma, o modelo obtido para o QSAR, foi validado e podemos classificá-lo como um bom modelo de predição.

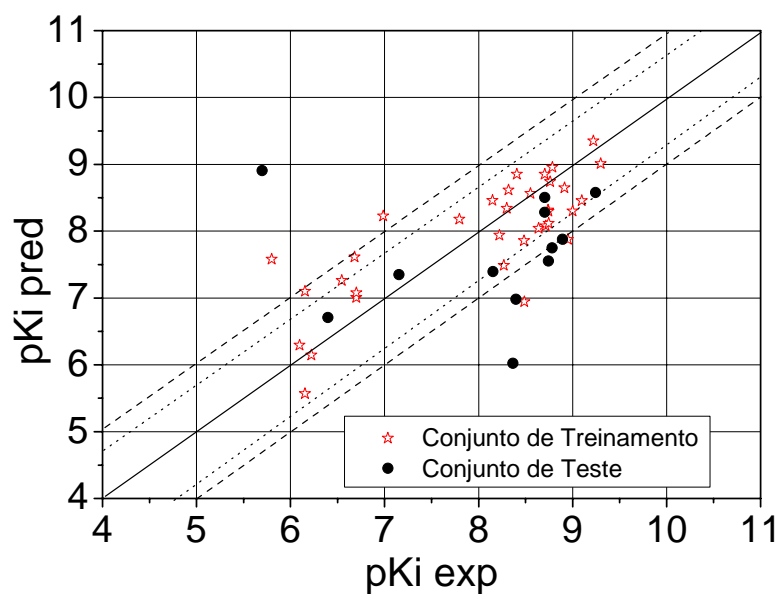


FIGURA 21 Gráfico obtido no *Scramble Test* dos valores de pKi experimental e predito para os valores obtidos no QSAR-5D.

Por fim, de acordo com os resultados obtidos para o QSAR-5D, temos que dentre a série estudada diversos compostos que se enquadram na superfície hipotética gerada, e dentre eles podemos destacar como exemplo os inibidores 1, 8, 9 e 19 como sendo os que apresentam em R2/R2' substituintes hidrofóbicos e em R1/R1' substituintes que apresentam características hidrofóbicas e hidrofílicas, ou seja, que possuem em sua estrutura grupos de caráter hidrofóbicos como cadeias carbônicas, mas também grupos hidrofílicos que possam vir a realizar ligações hidrogênio, como por exemplo a presença de uma hidroxila. Tudo isso de acordo com a superfície hipotética gerada pelo QSAR-5D e também pelos valores experimentais encontrados na literatura. Como exemplo, temos na Figura 22 o inibidor 1, que possui em R2/R2' apenas cadeias carbônicas e em R1/R1' um substituinte que tem caráter tanto hidrofóbico quanto hidrofílico (Figura 24).

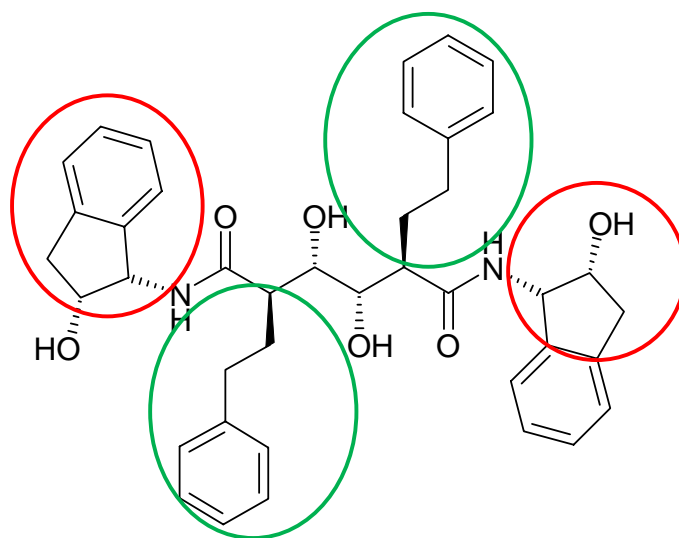


FIGURA 22 Estrutura do inibidor que mais se encaixa à superfície hipotética do receptor gerado no QSAR-5D. Em vermelho, o substituinte R1/R1' e em verde, o grupo hidrofóbico substituinte em R2/R2'.

5.3 QSAR-6D

5.3.1 Análise dos modelos de QSAR-6D

O conceito de modelagem de receptor através do software *Quasar*, desenvolvido por Vedani e colaboradores (Vedani et al., 2005), é baseado no QSAR-6D e explicitamente permite a simulação do encaixe induzido. O software *Quasar* gera uma família de receptores “quase-atômico” que são otimizados através de um algoritmo genético. O receptor hipotético é caracterizado por uma superfície de van der Waals tridimensional que circunda as moléculas do ligante e que é preenchida com propriedades atômicas mapeadas para os mesmos. A topologia desta superfície imita a forma 3D do sítio de ligação; as propriedades mapeadas representam outras informações de interesse, como a hidrofobicidade, potencial eletrostático, e propensão a ligação hidrogênio.

O conceito *Quasar* (6D-QSAR), permite a consideração simultânea de diferentes modelos de solvatação. Isto pode ser alcançado pelo mapeamento explícito de partes da superfície com propriedades do solvente (posição e tamanho são otimizados pelo algoritmo genético) ou implicitamente.

Das 46 moléculas utilizadas em nosso estudo, definiram-se 34 compostos como conjunto de treinamento e 12 compostos como conjunto de teste. No *Quasar*, foi formada uma família composta por 200 modelos que evoluíram ao longo dos 5000 *crossovers*, correspondendo a 200 gerações. A simulação atingiu uma validação cruzada q^2 igual a 0,707, um valor $R^2_{\text{Treinamento}}$ de 0,717 e um valor de R^2_{Teste} de 0,719. A superfície do receptor é mostrada na Figura 23.

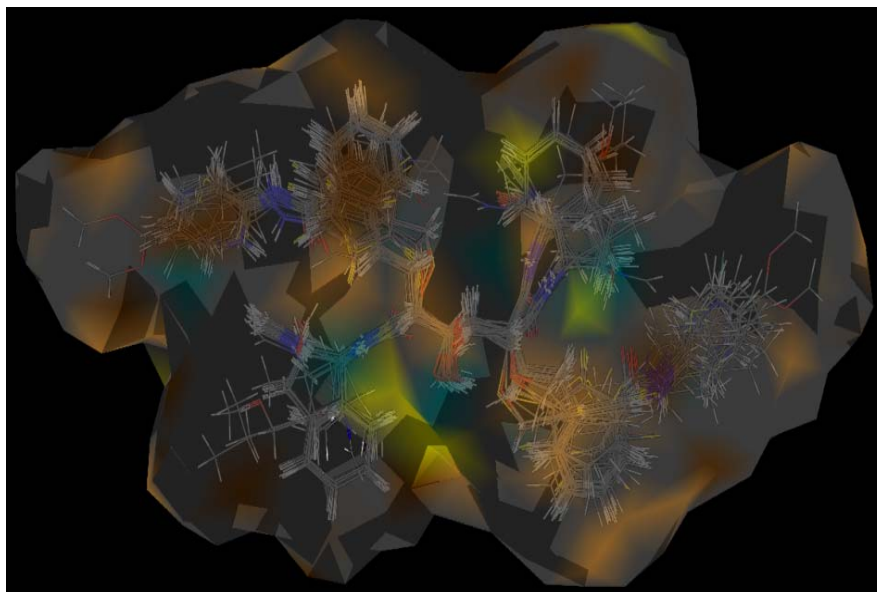


FIGURA 23 Superfície mapeada pelas propriedades “quase atomísticas” dos inibidores peptídicos obtida no QSAR-6D. As cores significam: verde=ligação hidrogênio doadora, amarelo=ligação hidrogênio aceptora, marrom claro= região hidrofóbica carregadas positivamente, marrom escuro= região hidrofóbica carregadas negativamente e cinza= região hidrofóbicas neutra.

Podemos observar na Figura 23 a representação do modelo “quase-atomístico” do sítio de ligação que melhor se adapta aos inibidores peptídicos em estudo. O modelo do receptor obtido pode ser descrito como predominantemente hidrofóbico, com três regiões caracterizadas por ligações hidrogênio que acomodam substituintes hidrofílicos em sua região central. As cores mostram que a região amarela representa ligação hidrogênio aceptora, ou seja, o inibidor deverá possuir em R1 um substituinte doador de ligação hidrogênio. A região em verde representa uma superfície doadora de ligações hidrogênio, onde o ligante deverá apresentar uma região aceptora de ligação

hidrogênio, o que está de acordo com a presença do diol nesta região. As regiões em marrom-claro e marrom-escuro representam regiões hidrofóbicas carregadas positivamente e negativamente, respectivamente; dessa forma o ligante deve apresentar regiões hidrofóbicas com cargas opostas às obtidas na superfície do receptor.

As estereorepresentações mostradas na Figura 24 mostram, respectivamente, as estruturas dos ligantes com menor valor de K_i (Inibidor 8) e a de maior valor de K_i (Inibidor 42) descritos na literatura dentro da superfície do receptor. Podemos dizer que, um inibidor que preenche os requisitos necessários à superfície do receptor hipotético gerado, possui em R1/R1' a presença de um substituinte hidrofílico e em R2/R2' substituinte predominantemente hidrofóbico. Desta forma, este composto será um bom inibidor, ou seja, um inibidor mais potente; enquanto que o inibidor de menor potência, possivelmente não atende a tais requisitos, como por exemplo, uma região hidrofóbica com carga oposta à da superfície.

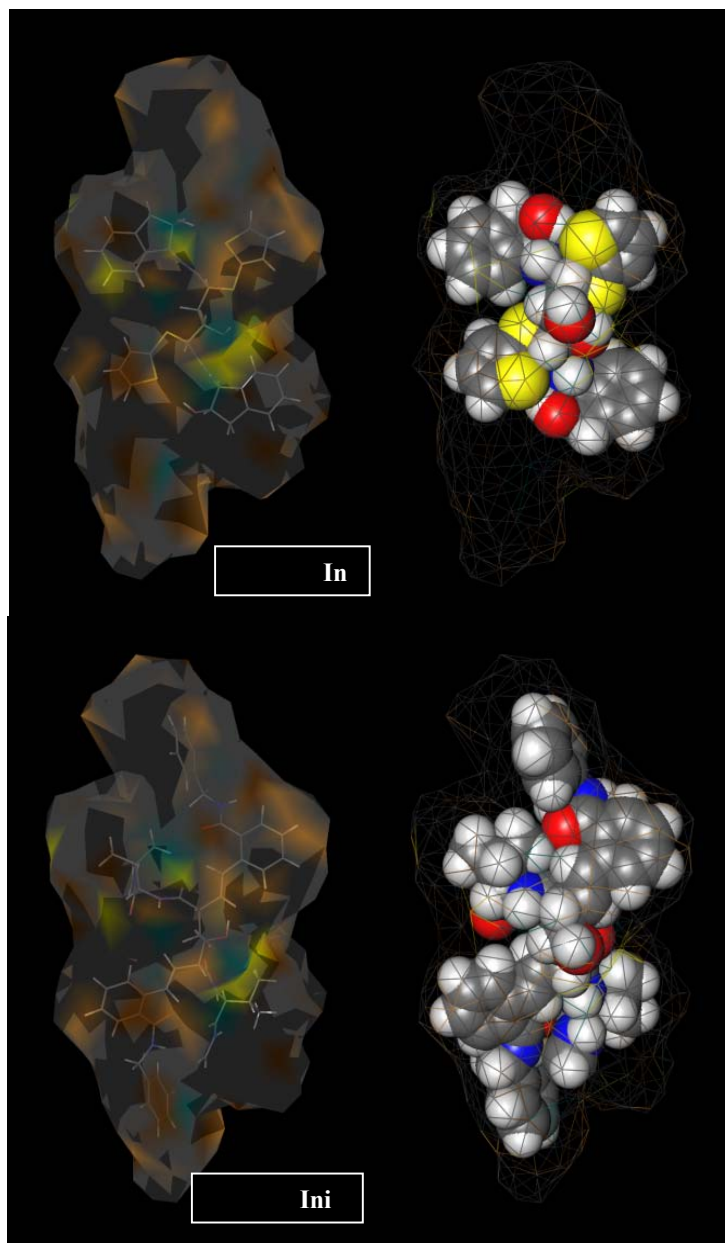


FIGURA 24 Superfície mapeada pelas propriedades “quase atômicas” dos inibidores peptídicos 8 e 42, que apresentam o maior e o menor valor de inibição experimental (valor de K_i), obtida no QSAR-6D.

A estereorepresentação da superfície do receptor foi apresentada na Figura 25; a correlação entre os valores de $pK_{i_{pred}}$ e $pK_{i_{exp}}$ obtidos no *Quasar* (QSAR-6D), são descritos na Figura 25 em forma de gráfico de $pK_{i_{exp}}$ versus $pK_{i_{pred}}$ para os conjuntos de treinamento e teste; e comparados na Tabela 9. O maior desvio observado para o conjunto de treinamento foi de 1,49 para o inibidor 46, e o menor de 0,01 para o inibidor 35; para o conjunto de teste, temos 1,48 para o inibidor 42 e 0,02 para o inibidor 13.

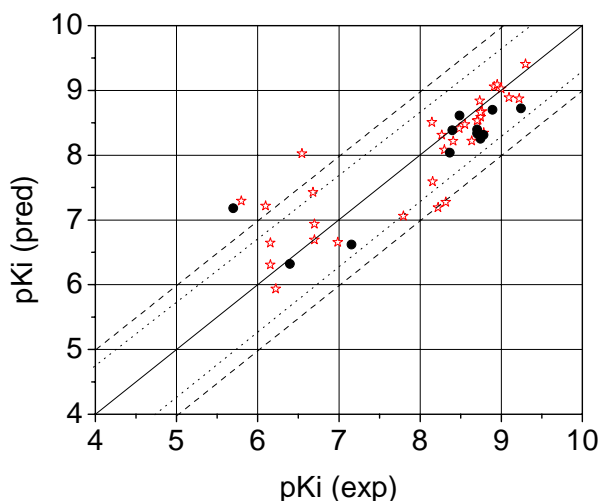


FIGURA 25 Gráfico dos valores de pK_i experimental e predito para os valores obtidos no QSAR-6D.

Os valores de ΔG predito e experimental, assim como seu desvio ($\Delta\Delta G$), também são apresentados na Tabela 9. Observa-se que o maior valor de $\Delta\Delta G$ obtido para o conjunto de treinamento foi de 1,400 para o inibidor **43** e o menor foi de 0,09 para o inibidor **35**; já para o conjunto de teste, temos o maior valor de 1,985 para o inibidor **42** e o menor de 0,017 para o inibidor **13**. Temos ainda,

que os inibidores **13**, **35** e **42** apresentaram as mesmas características quanto aos maiores e menores valores em relação ao pKi predito obtido no 6D.

TABELA 9 Valores de pKi experimental e predito para o conjunto de treinamento obtidos no QSAR-6D.

Inibidor	pki Experimental	pki Predito	Desvio	ΔG (exp)	ΔG (pred)	$\Delta\Delta G$
<i>Conjunto de Treinamento</i>						
1	9,22	8,87	0,35	-12.362	-11.892 ± 0.153	0.470
3	8,64	8,22	0,41	-11.579	-11.022 ± 0.151	0.557
4	6,55	8,03	-1,48	-8.774	-10.755 ± 0.183	-1.981
5	8,55	8,47	0,08	-11.465	-11.361 ± 0.133	0.104
7	7,80	7,06	0,74	-10.450	-9.462 ± 0.144	0.988
8	9,30	9,41	-0,11	-12.468	-12.610 ± 0.179	-0.142
9	8,91	9,06	-0,15	-11.948	-12.144 ± 0.117	-0.196
11	8,15	8,51	-0,36	-10.921	-11.405 ± 0.137	-0.484
12	8,27	8,31	-0,04	-11.084	-11.142 ± 0.138	-0.058
15	8,48	8,42	0,07	-11.371	-11.280 ± 0.132	0.091
16	8,78	8,34	0,44	-11.776	-11.185 ± 0.148	0.591
17	8,41	8,22	0,19	-11.272	-11.013 ± 0.153	0.259
18	8,30	8,08	0,22	-11.127	-10.832 ± 0.146	0.295
19	9,10	8,89	0,20	-12.194	-11.919 ± 0.101	0.275
20	8,72	8,54	0,17	-11.684	-11.449 ± 0.138	0.235
21	8,74	8,84	-0,10	-11.712	-11.850 ± 0.131	-0.138
22	8,75	8,59	0,16	-11.725	-11.513 ± 0.127	0.212
24	8,95	9,09	-0,14	-11.998	-12.187 ± 0.181	-0.189
26	8,76	8,67	0,09	-11.745	-11.629 ± 0.162	0.116
27	8,70	8,53	0,17	-11.667	-11.437 ± 0.167	0.230
29	8,74	8,66	0,08	-11.722	-11.608 ± 0.114	0.114
31	9,00	9,02	-0,02	-12.064	-12.097 ± 0.170	-0.033
32	6,15	6,31	-0,15	-8.250	-8.456 ± 0.125	-0.206
33	6,22	5,93	0,29	-8.340	-7.955 ± 0.138	0.385
35	6,70	6,69	0,01	-8.980	-8.971 ± 0.132	0.009
36	6,15	6,64	-0,49	-8.250	-8.908 ± 0.136	-0.658
37	8,15	7,59	0,56	-10.931	-10.177 ± 0.194	0.754
39	6,10	7,21	-1,12	-8.173	-9.671 ± 0.127	-1.498
40	6,70	6,94	-0,24	-8.980	-9.299 ± 0.201	-0.319
41	8,22	7,19	1,03	-11.021	-9.640 ± 0.140	1.381
43	8,32	7,27	1,04	-11.151	-9.751 ± 0.170	1.400
44	6,99	6,66	0,33	-9.366	-8.923 ± 0.175	0.443
45	6,68	7,43	-0,75	-8.957	-9.958 ± 0.159	-1.001
46	5,80	7,29	-1,49	-7.773	-9.778 ± 0.150	-2.005
<i>Conjunto de Teste</i>						
2	8,74	8,25	0,49	-11.722	-11.064 ± 0.147	0.658
6	8,37	8,04	0,33	-11.215	-10.775 ± 0.154	0.440
10	8,49	8,61	-0,12	-11.376	-11.540 ± 0.148	-0.164
13	8,40	8,38	0,02	-11.257	-11.240 ± 0.188	0.017
14	8,78	8,31	0,47	-11.773	-11.142 ± 0.140	0.631

...continua...

TABELA 9, Cont.

23	8,89	8,70	0,19	-11.921	-11.665 ± 0.174	0.256
25	8,70	8,40	0,30	-11.667	-11.255 ± 0.136	0.412
28	8,70	8,33	0,37	-11.667	-11.173 ± 0.159	0.494
30	9,24	8,72	0,52	-12.391	-11.694 ± 0.162	0.697
34	6,40	6,32	0,08	-8.576	-8.475 ± 0.146	0.101
38	7,15	6,62	0,53	-9.591	-8.875 ± 0.139	0.716
42	5,70	7,18	-1,48	-7.639	-9.624 ± 0.154	-1.985

Para identificarmos potenciais sítios e funcionalidades permitidas a um aumento da afinidade de ligação, a funcionalidade de grupos individuais de ambos - conjunto de treinamento e teste - é analisada pela sua contribuição à energia livre ao sítio de ligação, ΔG . Isso pode ser calculado através dos detalhes entálpicos (eletrostático, van der Waals e ligação hidrogênio), ligações e ângulos de torção, e contribuições do encaixe-induzido para o cálculo da afinidade do sítio de ligação.

O modelo do receptor obtido foi validado através do *scramble test* (Figura 26) tendo como resultado os valores de q^2 (coeficiente de correlação linear quadrático de validação cruzada) de -0,053 e um valor de $R^2_{\text{Treinamento}}$ de 0,001, sendo então validado como um bom modelo de predição.

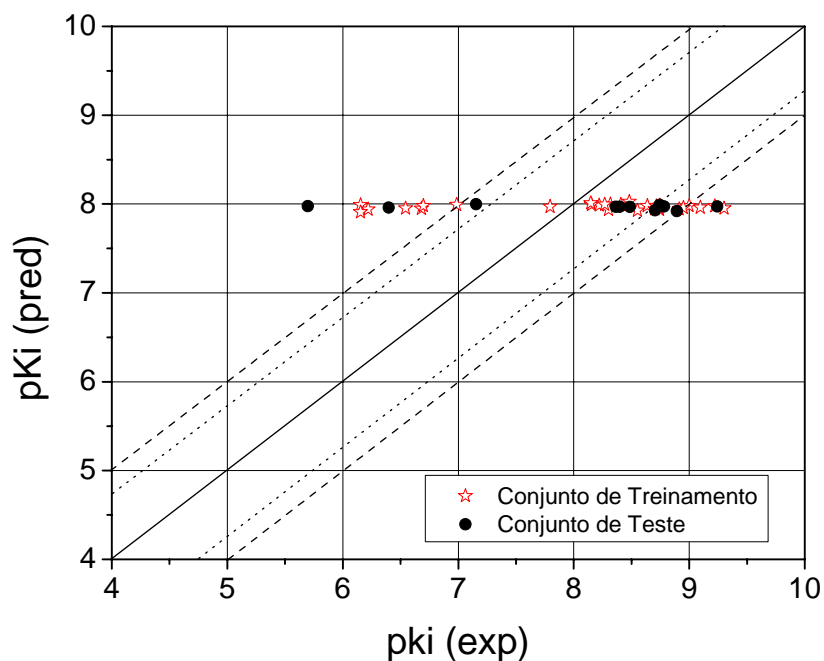


FIGURA 26 Gráfico obtido no *Scramble Test* dos valores de pKi experimental e predito para os valores obtidos no QSAR-6D.

Dessa forma, de acordo com os resultados obtidos para o QSAR-6D, temos que dentre a série estudada diversos compostos que se enquadram na superfície hipotética gerada, e dentre eles podemos destacar novamente os inibidores 1, 8, 9 e 19 como sendo aqueles que melhor se adaptam a superfície hipotética do receptor gerada através dos diferentes modelos de solvatação simultaneamente, por apresentarem em R1/R1' substituintes que possuem caráter hidrofílico e hidrofóbico de acordo com a superfície obtida e em R2/R2', um substituinte predominantemente hidrofóbico (Figura 19 e 23). Portanto, os resultados obtidos no QSAR-6D, corroboram com os resultados obtidos no QSAR-5D, fornecendo-nos informações cruciais para o entendimento da série de compostos estudada.

Na literatura, encontramos que os valores mínimos de q^2 e $r^2_{\text{preditivo}}$ aceitáveis em trabalhos de QSAR são respectivamente, 0,5 e 0,7 (Vedani & Dobler, 2002, 2008). Dessa forma, os nossos resultados sugerem que o trabalho desenvolvido apresentou resultados aceitáveis e confiáveis, de acordo com os valores encontrados na literatura para outros conjuntos de dados.

Os estudos de QSAR-5D e 6D, nos forneceram informações características dos substituintes R1 e R2, para um melhor encaixe no sítio do receptor. Temos então, em R2/R2' uma preferência por substituintes hidrofóbicos, e em R1/R1' substituintes que possuem características tanto hidrofóbicas quanto hidrofílicas, em que dentro da nossa série de compostos em estudo, destaca-se os compostos 1, 8, 9 e 19 como sendo os que mais se aproximaram aos resultados obtidos.

Portanto, a partir dos resultados obtidos, podemos então propor a estrutura de um composto que melhor preencha os requisitos do receptor hipotético gerado, propor modificações significativas em R1/R1' e R2/R2', e calcular sua energia (ΔG) através do QSAR-5D e então prever sua potência inibitória pela relação entre energia e K_i .

Nossa proposta de uma estrutura hipotética está representada na Figura 27, onde em R1 temos um substituinte hidrofóbico, R1' um substituinte hidrofílico, podendo vir a realizar ligações hidrogênio com o receptor, e em R2/R2' um substituinte predominantemente hidrofóbico, mantendo em seu esqueleto principal um peptídeo com diol vicinal central.

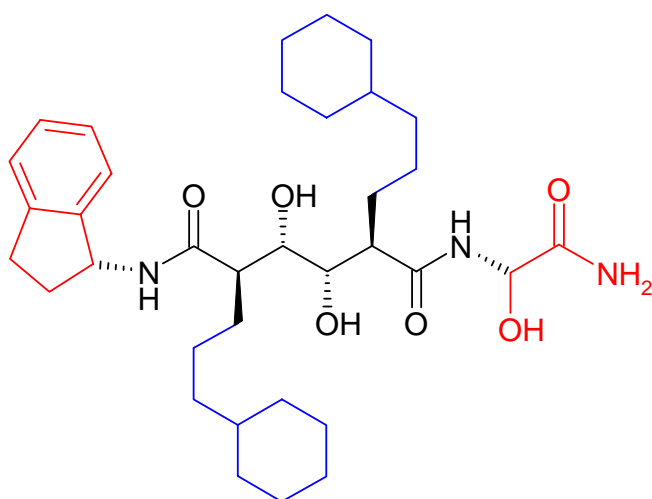


FIGURA 27 Estrutura hipotética proposta para estudos posteriores baseados nos resultados obtidos no estudo de QSAR de inibidores peptídicos com diol vicinal central.

6 CONCLUSÕES

Neste estudo, construímos e avaliamos modelos obtidos dos métodos de QSAR-5D e 6D para inibidores da HIV-1 protease, baseado numa série de inibidores peptídicos com grupamento diol vicinal central. Os melhores modelos no estudo de QSAR compreendem importantes características que podem ser aplicadas no planejamento de novos inibidores a partir das propriedades físico-químicas desse receptor obtido.

Pelo modelo de QSAR-5D obtido, um bom inibidor, análogo aos inibidores peptídicos com diol vicinal central será aquele que tiver em R1/R1' grupos parte hidrofóbicos e parte hidrofílicos como cadeias carbônicas com um heteroátomo com pares de elétrons livres ou ainda presença de hidroxilas, e em R2/R2' grupos predominantemente hidrofóbicos. No estudo do QSAR-6D, que leva em conta diversos modelos de solvatação, apresentou resultados que corroboraram com os resultados obtidos no 5D. Portanto, dentro da série de inibidores estudada, destaca-se os compostos 1, 8, 9 e 19 como sendo os que mais se encaixam à superfície hipotética gerada, e a partir deles propor novas estruturas .

Então concluímos que o estudo multimensional de QSAR nos fornece informações crucias no estudo de novos potenciais inibidores da HIV-1 protease, auxiliando-nos no entendimento da influência dos grupos substituintes, assim como o encaixe induzido desses compostos e a presença de solvente. Podemos, nesse contexto, sugerir compostos inéditos como potenciais inibidores da protease do HIV.

Perspectivas

Temos como perspectivas, dar continuidade no estudo sobre a estrutura hipotética proposta, fazendo as simulações de QSAR-5D e 6D, a fim de avaliar sua capacidade inibitória, e posteriormente sugerir novos compostos inéditos que poderão vir a ser utilizados no combate ao HIV.

Além disso, como as metodologias de QSAR-5D e 6D são recentes e ainda pouco utilizadas no Brasil, temos como proposta fazer a simulação dessa série de compostos utilizando a metodologia do QSAR-4D desenvolvido por Hopfinger et al. (1997), a fim de se estudar o comportamento dessa série de compostos em função do tempo de simulação de dinâmica molecular, e assim comparar os resultados obtidos e as metodologias estudadas.

REFERÊNCIAS BIBLIOGRÁFICAS

ABOUL-ELA, F.; KARN, J.; VARANI, J. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. **Journal of Molecular Biology**, Dordrecht, v. 253, n. 2, p. 313-332, Oct. 1995.

AIKEN, C.; KONNER, J.; LANDAU, N. R.; LENBURG, M. E.; TRONO, D. Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. **Cell**, Cambridge, v. 76, n. 5, p. 853-864, Mar. 1994.

APPELT, K. Crystal structures of HIV-1 protease-inhibitor complexes. **Perspectives in Drug Discovery and Design**, Leiden, v. 1, n. 1, p. 23-48, Aug. 1993.

BALDWIN, E.; WEBER, I. T.; SELK, L. M.; CLAWSON, L.; SCHNEIDER, J.; KENT, S. B. H. Conserved folding in retroviral proteases-crystal structure of synthetic HIV-1. **Science**, New York, v. 245, n. 4918, p. 616-621, Aug. 1989.

BATISTA, P. R. **Estudos computacionais da protease do HIV-1: abertura das alças e diferenças entre subtipos**. 2005. 159 p. Dissertação (Mestrado em Biofísica) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

BLANEY, J. M.; WEINER, P. K.; DEARING, A.; KOLLMAN, P. A.; JORGENSEN, E. C.; OATLEY, S. J.; BURRIDGE, J. M.; BLAKE, J. F. Molecular mechanics simulation of protein-ligand interactions: binding of thyroid analogues to prealbumin. **Journal of the American Chemical Society**, Easton, v. 104, n. 23, p. 6424-6434, Nov. 1982.

BOUR, S.; SCHUBERT, U.; STREBEL, K. The human immunodeficiency virus type 1 vpu protein specifically binds to the cytoplasmic domain of CD4: implications for the mechanism of degradation. **Journal of Virology**, Washington, v. 69, n. 3, p. 1510-1520, Mar. 1995.

BRIGGS, J. A. G.; MARC, C. J.; SIMON, M. N.; FULLER, S. D.; VOGT, V. M. Cryo-electron microscopy reveals conserved and divergent features of gag packing in immature particles of rous sarcoma virus and human immunodeficiency virus. **Journal of Molecular Biology**, Dordrecht, v. 355, n. 1, p. 157-168, Jan. 2006.

BUTH, K. H.; PEYMAN, A.; HANSEN, J.; KNOLLE, J.; MEICHSNER, C. HIV protease inhibitor HOE/BAY 793, structure-activity relationships in a series of C2- symmetric diols. **Bioorganic & Medicinal Chemistry**, New York, v. 3, n. 5, p. 559-571, May 1995.

COLLINS, J. R.; BURT, S. K.; ERICKSON, J. W. Activated dynamics of flap opening in HIV-1 protease. **Advances in Experimental Medicine and Biology**, New York, v. 362, p. 455-460, 1995a.

COLLINS, J. R.; BURT, S. K.; ERICKSON, J. W. Flap opening in HIV-1 protease simulated by 'activated' molecular dynamics. **Nature Structural Biology**, New York, v. 2, n. 4, p. 334-338, Apr. 1995b.

COLLINS, K. L.; CHEN, B. K.; WALKER, B. D.; BALTIMORE, D. HIV-1 nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. **Nature**, London, v. 391, p. 397-401, Jan. 1998.

CONJUNTO DE LAS NACIONES UNIDAS SOBRE EL HIV/SIDA. **Informe sobre la epidemia mundial de sida**. Disponível em: <http://www.unaids.org/es/KnowledgeCentre/HIVData/GlobalReport/2008/2008_Global_report.asp>. Acesso em: 12 abr. 2009.

CRAWFORD, S.; GOFF, S. P. A deletion mutation in the 5' part of the POL gene of moloney murine leukemia virus blocks proteolytic processing of the GAG and POL polyproteins. **Journal of Virology**, Washington, v. 53, n. 3, p. 899-907, Mar. 1985.

CUNHA, E. F. F. da. **Estudos quantitativos de correlação estrutural-atividade em 3D e 4D de inibidores peptídicos da HIV-1 protease**. 2006. 194 p. Tese (Doutorado em Química Orgânica) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

CUNHA, E. F. F. da; MARTINS, R. C. A.; ALBUQUERQUE, M. G.; ALENCASTRO, R. B. LIV-3D-QSAR model for estrogen receptor ligands. **Journal of Molecular Modeling**, Washington, v. 10, n. 4, p. 297-304, Aug. 2004.

CUNNINGHAM, A. L.; DWYER, D. E.; MILLS, J.; MONTAGNIER, L. Structure and function of HIV. **Medical Journal of Australia**, Victoria, v. 164, n. 3, p. 161-165, Feb. 1996.

DIMITROV, S. D. Virus entry: molecular mechanism and biomedical applications. **Nature Review Microbiology**, London, v. 2, n. 2, p. 109-122, Feb. 2004.

DOUEK, D. C.; BRECHLEY, J. M.; BETTS, M. R.; AMBROZAK, D. R.; HILL, B. J.; OKAMOTO, Y.; CASAZZA, J. P.; KURUPPU, J.; KUNSTMAN, K.; WOLINSKY, S.; GROSSMAN, Z.; DYBUL, M.; OXENIUS, A.; PRICE, D. A.; CONNORS, M.; KOUPECKI, R. A. HIV preferentially infects HIV-specific CD4⁺ T cells. **Nature**, London, v. 417, p. 95-98, May 2002.

FERREIRA, J. **AIDS e infecção pelo HIV**. Porto Alegre: Artmed, 1996. 178 p.

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Quimiometria I: calibração multivariada, um tutorial. **Química Nova**, São Paulo, v. 22, n. 5, p. 724-731, set./out. 1999.

FITZGERALD, P. M.; MCKEEVER, B. M.; MIDDLESWORTH, J. F. van; SPRINGER, J. P.; HEIMBACH, J. C.; LEU, C. T.; HERBER, W. K.; DIXON, R. A.; DARKE, P. L. Crystallographic analysis of a complex between human immunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0-Å resolution. **Journal of Biological Chemistry**, Bethesda, v. 265, n. 24, p. 14209-14219, Aug. 1990.

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

FUNDAÇÃO OSWALDO CRUZ. **Vírus**. Disponível em:
<<http://www.fiocruz.br>>. Acesso em: 10 mar. 2009.

GARG, R.; GUPTA, S. P.; GAO, H.; BABU, M. S.; DEBNATH, A. K.; HANSCH, C. Comparative quantitative structure-activity relationship studies on anti-HIV drugs. **Chemical Reviews**, Washington, v. 99, n. 12, p. 3525-3601, Nov. 1999.

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

GAUDIO, A. C.; ZANDONADE, E. Proposição, validação e análise dos modelos que correlacionam estrutura química e atividade biológica. **Química Nova**, São Paulo, v. 24, n. 5, p. 658-671, set./out. 2001.

GUSTCHINA, A.; WEBER, I. T. Comparative analysis of the sequences and structures of HIV-1 and HIV-2 proteases. **Proteins: Structure, Function and Bioinformatics**, New York, v. 10, n. 4, p. 325-339, Apr. 1991.

HAHN, B. H.; SHAW, G. M.; COCK, K. M. de; SHARP, P. M. AIDS as a zoonosis: scientific and public health implications. **Science**, New York, v. 287, n. 5453, p. 607-614, Jan. 2000.

HANSH, C.; FUJITA, T. Pi-Sigma-Rho analysis: a method for the correlation of biological activity and chemical structure. **Journal of the American Chemistry Society**, Easton, v. 86, n. 8, p. 1616-1621, Apr. 1964.

HENDERSON, L. E.; BENVENISTE, R. E.; SOWDER, R.; COPELAND, T. D.; SCHULTZ, A. M.; OROSZLAN, S. Molecular characterization of gag proteins from simian immunodeficiency virus (SIV_{Mne}). **Journal of Virology**, Washington, v. 62, n. 8, p. 2587-2595, Aug. 1988.

HOPFINGER, A. J.; WANG, S.; TOKARSKI, J. S.; JIN, B.; ALBUQUERQUE, M.; MADHAV, P. J.; DURAISWAMI, C. J. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. **Journal of the American Chemical Society**, Easton, v. 119, n. 5, p. 10509-10524, Oct. 1997.

HOSUR, M. V.; BHAT, T. N. Influence of stereochemistry on activity and binding modes for C2 symmetry-bases diol inhibitors of HIV-1 protease. **Journal of the American Chemical Society**, Easton, v. 116, n. 3, p. 847-855, Feb. 1994.

HU, D. J.; DONDERO, T. J.; RAYFIELD, M. A.; GEORGE, J. R.; SCHOCHETMAN, G.; JAFFE, H. W.; LUO, C. C.; KALISH, M. L.; WENIGER, B. G.; PAU, C. P.; SCHABLE, C. A.; CURRAN, J. W. The emerging genetic diversity of HIV: the importance of global surveillance for diagnostics, research, and prevention. **Journal of the American Medical Association**, Chicago, v. 275, n. 3, p. 210-216, Jan. 1996.

KARKI, R. G.; KULKARNI, V. M. Three-dimensional quantitative structure-activity relationship (3D-QSAR) of 3-aryloxazolidin-2-one antibacterials. **Bioorganic & Medicinal Chemistry**, New York, v. 9, n. 12, p. 3153-3160, Dec. 2001.

KENT, S.; MARSHALL, G. R.; WLODAWER, A. Determining the 3D structure of HIV-1 protease. **Science**, New York, v. 288, n. 5471, p. 1590-1593, June 2000.

KJEMS, J.; BERKHOUT, B.; DAS, A. T.; REINERT, L. S.; LUTZELBERGER, M. A Novel splice donor site in the *gag-pol* gene is required for HIV-1 RNA stability. **The Journal of Biological Chemistry**, Baltimore, v. 281, n. 27, p. 18644-18651, July 2006.

KUBINYI, H. QSAR: hanch analysis and related approaches. In: MANNHOLD, R.; KROGSGAARD-LARSEN, P.; TIMMERMAN, H. (Ed.). **Methods and principles in medicinal chemistry**. Weinheim: VHC, 1993. v. 1, p. 63-97, 240 p.

KUBINYI, H. QSAR and 3D QSAR in drug design: part 1: methodology. **Drug Discovery Today**, Kidlington, v. 2, n. 11, p. 457-467, Nov. 1997.

LAPATTO, R.; BLUNDELL, T.; HEMMINGS, A.; OVERINGTON, J.; WILDERSPIN, A.; WOOD, S.; MERSON, J. R.; WHITTLE, P. J.; DANLEY, D. E.; GEOGHEGAN, K. S. B. H. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. **Science**, New York, v. 245, n. 4918, p. 616-621, Aug. 1989.

LEVY, J. A. Heterogeneity in transmission and pathogenesis. In: MANN, J. M.; TARANTOLA, D. J. M. (Ed.). **AIDS in the world II**. New York: Oxford University, 1996. p. 177-185.

LILL, M. A. **RAPTOR 2.0**: user's manual biographics laboratory 3R. Basel: Friedensgasse, 2006. 54 p.

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

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

LIPKOWITZ, K. B.; BOYD, D. B. **Reviews in computational chemistry**. New York: VHC, 1990. v. 1, 384 p.

MARIANI, R.; CHEN, D.; SCHRÖFELBAUER, B. Species-specific exclusion of APOBEC3G from HIV-1 virions by vif. **Cell**, Cambridge, v. 114, n. 1, p. 21-31, July 2003.

MARTINS, G. A.; DONAIRE, D. **Princípios de estatística**. São Paulo: Atlas, 1993. 56 p.

MARTINS, N. H. **Ensaio enzimático de proteases de HIV-1 de subtipos brasileiros**. 2007. 119 p. Dissertação (Mestrado em Física Biomolecular) - Universidade de São Carlos, São Carlos.

MILLER, M. D.; FARNET, C. M.; BUSHMAN, F. D. Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. **Journal of Virology**, Washington, v. 71, n. 7, p. 5382-5390, July 1997.

MILLER, M.; JASKÓLSKI, M.; RAO, J. K. M.; LEIS, J.; WLODAWER, A. Crystal structure of a retroviral protease proves relationship to aspartic protease family. **Nature**, London, v. 337, p. 576-579, Feb. 1989.

NAVIA, M. A.; FITZGERALD, P. M.; MCKEEVER, B. M.; LEU, C. T.; HEIMBACH, J. C.; HERBER, W. K.; SIGAL, I. S.; DARKE, P. L.; SPRINGER, J. P. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. **Nature**, London, v. 337, n. 6208, p. 615-620, Feb. 1989.

OKIMOTO, N.; TSUKUI, T.; KITAYAMA, K. Molecular dynamics study of HIV-1 protease-substrate complex: roles of the water molecules at the loop structures of the active site. **Journal of the American Chemical Society**, Easton, v. 122, n. 23, p. 5613-5622, May 2000.

OLIVEROS, M. P. R. **Prevalência e covariação de mutações relacionadas a resistência aos inibidores de protease do subtipo F do HIV-1**. 2005. 118 p. Dissertação (Mestrado em Fisiopatologia Experimental) - Faculdade de Medicina, Universidade de São Paulo, São Paulo.

PATRICK, G. L. **An introduction to medicinal chemistry**. Oxford: Oxford University, 1995. 646 p.

PEARL, L. H.; TAYLOR, W. R. A structural model for the retroviral proteases. **Nature**, London, v. 329, n. 6137, p. 351-354, Sept. 1987.

PEÇANHA, E. P.; ANTUNES, O. A. C.; TANURI, A. Estratégias farmacológicas para a terapia anti-AIDS. **Química Nova**, São Paulo, v. 25, n. 6, p. 1108-1116, nov./dez. 2002.

PENG, C.; HO, B. K.; CHANG, T. W.; CHANG, N. T. Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity. **Journal of Virology**, Washington, v. 63, n. 6, p. 2550-2556, June 1989.

PETER, F. HIV nef: the mother of all evil? **Immunity**, Washington, v. 9, n. 4, p. 433-437, Oct. 1998.

POMERANTZ, R. J.; HORN, D. L. Twenty years of therapy for HIV-1 infection. **Nature Medicine**, London, v. 9, n. 1, p. 867-873, Jan. 2003.

PRABU-JEYABALAN, M.; NALIVAICA, E.; SCHIFFER, C. A. How does a symmetric dimer recognize an asymmetric substrate?: a substrate complex of HIV-1 protease. **Journal of Molecular Biology**, Dordrecht, v. 301, n. 5, p. 1207-1220, Sept. 2000.

PRIESTLE, J. P.; FASSLER, A.; ROSEL, J.; TINTELNOT-BLOMLEY, M.; STROP, P.; GRUTTER, M. G. Comparative analysis of the X-ray structures of HIV-1 and HIV-2 proteases in complex with CGP 53820, a novel pseudosymmetric inhibitor. **Structure**, Philadelphia, v. 3, n. 4, p. 381-389, Apr. 1995.

RAMOS, A. L. Q. B. **Busca de descritores para modelagem quimiométrica de inibidores de corrosão**. 2005. 135 p. Dissertação (Mestrado em Físico-Química) - Universidade Federal Fluminense, Niterói.

ROGERS, D.; HOPFINGER, A. J. Application of genetic function approximation to quantitative structure-activity relationship and quantitative structure-property relationship. **Journal of Chemical Information and Computer Sciences**, Washington, v. 34, n. 4, p. 854-866, July 1994.

SAMUELSSON, B.; CLASSON, B.; HALLBERG, A. Synthesis of novel, potent, diol-based HIV-1 protease inhibitors via intermolecular pinacol homocoupling of (2S)-2-Benzoyloxymethyl-4-phenylbutanal. **Journal of Medicinal Chemistry**, New Jersey, v. 44, n. 21, p. 3407-3416, Sept. 2001a.

SAMUELSSON, B.; HALLBERG, A.; WANNBERG, J.; KAISER, N. K.; VRANG, L.; LARHED, M.; HALLBERG, A. High-speed synthesis of potent C2-symmetric HIV-1 protease inhibitors by in-situ aminocarbonylations. **Journal of Combinatorial Chemistry**, Washington, v. 7, n. 4, p. 611-617, May 2005.

SAMUELSSON, B.; MUHLMAN, C. B.; HALLBERG, A. Synthesis of potent C2-symmetric, diol-based HIV-1 protease inhibitors: investigation of thioalkyl and thioaryl P1/P1' substituents. **Journal of Medicinal Chemistry**, New Jersey, v. 44, n. 19, p. 3402-3406, Aug. 2001b.

SAMUELSSON, B.; OSCARSSON, K.; LAHMANN, M.; LINDBERG, J.; KANGASMETSÄ, J.; UNGE, T.; OSCARSON, S.; HALLBERG, A. Design and synthesis of HIV-1 protease inhibitors novel tetrahydrofuran P2/P2'-groups interacting with Asp29/30 of the HIV-1 protease: determination of binding from X-ray crystal structure of inhibitor protease complex. **Bioorganic & Medicinal Chemistry**, New York, v. 11, n. 6, p. 1107-1115, Mar. 2003.

SAMUELSSON, B.; PYRING, D.; LINDBERG, J.; ROSENQUIST, A.; ZUCCARELLO, G.; KVARNSTRO, M. I.; ZHANG, H.; VRANG, L.; UNGE, T.; HALLBERG, A. Design and synthesis of potent C2-symmetric diol-based HIV-1 protease inhibitors: effects of fluoro substitution. **Journal of Medicinal Chemistry**, New Jersey, v. 44, n. 19, p. 3083-3091, Aug. 2001c.

SAMUELSSON, B.; WANNBERG, J.; SABNIS, Y. A.; VRANG, L.; KARLE, A. A new structural theme in C2-symmetric HIV-1 protease inhibitors: ortho-substituted P1/P10 side chains. **Bioorganic & Medicinal Chemistry**, New York, v. 14, n. 5, p. 5303-5315, Aug. 2006.

SEARLE, M. S.; WILLIAMS, D. H. The cost of conformational order: entropy changes in molecular associations. **Journal of the American Chemical Society**, Easton, v. 114, n. 27, p. 10690-10697, Dec. 1992.

SELASSIE, C. D. History of quantitative structure-activity relationship. In: ABRAHAM, D. J. **Burger's medicinal chemistry and drug discovery**. 6. ed. New York: J.Wiley, 2003. v. 1, cap. 1, p. 1-48.

SIERRA, S.; KUPFER, B.; KAISER, R. Basics of the virology of HIV-1 and its replication. **Journal of Clinical Virology**, Amsterdam, v. 34, n. 4, p. 233-244, Dec. 2005.

SOARES, E. A. J. M. **Caracterização epidemiológica e molecular do vírus da imunodeficiência humana do tipo 1 no sul do Brasil**. 2003. 160 p. Dissertação (Mestrado em Genética) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

STILL, W. C.; TEMPCZYK, A.; HAWLEY, R. C.; HENDRICKSON, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. **Journal of the American Chemical Society**, Easton, v. 112, n. 16, p. 6127-6129, Aug. 1990.

SUMMERS, M. F.; TURNER, B. G. Structural biology of HIV. **Journal of Molecular Biology**, Dordrecht, v. 285, n. 1, p. 1-32, Jan. 1999.

SWANSTROM, R.; ERONA, J. Human immunodeficiency virus type-1 protease inhibitors: therapeutic successes and failures, suppression and resistance. **Pharmacology & Therapeutics**, Saint Louis, v. 86, n. 2, p. 145-170, May 2000.

THE JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS. **Programme HIV/AIDS**. Disponível em: <<http://www.unaids.org/en/default.asp>>. Acesso em: 12 abr. 2009.

TOLEDO, G. S.; FONSECA, J. S.; MARTINS, G. A. **Estatística aplicada**. 2. ed. São Paulo: Atlas, 1985. 396 p.

TORSTEN, U.; SAMUELSSON, B.; HALLBERG, A.; CLASSON, B.; VRANG, L.; ZHANG, H.; KVARNSTROM, I.; ZUCCARELLO, G.; ROSENQUIST, A.; LOWGREN, S.; PYRING, D.; LINDEBERG, J. Symmetric fluoro-substituted diol-based HIV protease inhibitors: ortho-fluorinated and meta-fluorinated P1/P1' benzyloxy side groups significantly improve the antiviral activity and preserve binding efficacy. **European Journal Biochemistry**, Berlin, v. 271, n. 22, p. 4594-4602, Nov. 2004.

TOZSER, J.; ZAHUCZKY, G.; BAGOSSI, P.; LOUIS, J. M.; COPELAND, T. D.; OROSZLAN, S.; HARRISSON, R. W.; WEBER, I. T. Comparison of the substrate specificity of the human T-cell leukemia virus and human immunodeficiency virus proteinases. **European Journal Biochemistry**, Berlin, v. 267, n. 20, p. 6287-6295, Oct. 2000.

TROPSHA, A.; GOLBRAIKH, A. Beware of q^2 . **Journal of Molecular Graphics Modelling**, Washington, v. 20, n. 4, p. 269-276, Jan. 2002.

UNIVERSIDADE ESTADUAL PAULISTA. Faculdade de Arquitetura, Artes e Comunicação. **Considerações preliminares:** o que é o HIV? Disponível em: <<http://www.faac.unesp>>. Acesso em: 10 mar. 2009.

VAISHNAV, Y. N.; WONG-STAAAL, F. The biochemistry of AIDS. **Annual Reviews Biochemistry**, Palo Alto, v. 60, p. 577- 630, 1991.

VEDANI, A.; DOBLER, M. 5D-QSAR: the key for simulating induced fit? **Journal of Medicinal Chemistry**, New York, v. 45, n. 11, p. 2139-2149, Apr. 2002.

VEDANI, A.; DOBLER, M. **QUASAR 5.2:** user's manual biographics laboratory 3R, Friedensgasse 35, CH-4056. Basel: Laboratory 3R, 2008. 55 p.

VEDANI, A.; DOBLER, M.; LILL, M. A. Combining protein modeling and 6D-QSAR: simulating the binding of structurally diverse ligands to the estrogen receptor. **Journal of Medicinal Chemistry**, New York, v. 48, n. 11, p. 3700-3703, May 2005.

VEDANI, A.; DOBLER, M.; ZBINDEN, P. Quasi-atomistic receptor surface models: a bridge between 3D-QSAR and receptor modeling. **Journal of the American Chemical Society**, Easton, v. 120, n. 18, p. 4471-4477, Apr. 1998.

VEDANI, A.; HUHTA, D. W. A new force field for modeling metalloproteins. **Journal of the American Chemical Society**, Easton, v. 112, n. 12, p. 4759-4767, June 1990.

VEDANI, A.; ZBINDEN, P.; SNYDER, J. P.; GREENIDGE, P. A. Pseudoreceptor modeling: the construction of three-dimensional receptor surrogates. **Journal of the American Chemical Society**, Easton, v. 117, n. 17, p. 4987-4994, May 1995.

VOET, D.; VOET, J. **Biochemistry**. New York: J.Wiley, 1995. 1392 p.

WEATHERALL, D. J. Historical review towards molecular medicine: reminiscences of the haemoglobin field, 1960-2000. **British Journal of Haematology**, Oxford, v. 115, n. 4, p. 729-738, Dec. 2001.

WEI, P.; GARBER, M. E.; FANG, S. M.; FISCHER, W. H.; JONES, K. A. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. **Cell**, Cambridge, v. 92, n. 4, p. 451-462, Feb. 1998.

WHITE, D. O.; FENNER, F. **Medical virology**. San Diego: Academic, 1994. 603 p.

WLODAWER, A.; VONDRASEK, J. Inhibitors of HIV-1 protease: a major success of structure: assisted drug design. **Annual Review Biophysical Biomolecular Structure**, New York, v. 27, p. 249-284, June 1998.

WONG, C.; BRIK, A. HIV-1 protease: mechanism and drug discovery. **Organic Biological Chemistry**, New York, v. 1, n. 4, p. 5-14, Dec. 2003.

WORLD HEALTH ORGANIZATION. **HIV/AIDS**. Disponível em: <<http://www.who.int/en/>>. Acesso em: 15 jan. 2009.

YAAKOV, L.; CAFLISCH, A. Flexibility of monomeric and dimeric HIV-1 protease. **The Journal of Physical Chemistry B**, Washington, v. 107, n. 17, p. 3068-3079, May 2003.

ZBINDEN, P.; DOBLER, M.; FOLKERS, G.; VEDANI, A. *PrGen*: pseudoreceptor modeling using receptor-mediated ligand alignment and pharmacophore equilibration. **Quantitative Structure-Activity Relationships**, Weinheim, v. 17, n. 2, p. 122-130, June 1998.

Artigos Publicados

Recent Patents on Inflammation & Allergy Drug Discovery 2007, 1, 000-000

1

Targeting Inhibition of COX-2: A Review of Patents, 2002 - 2006

Elaine F. F. da Cunha*, Teodorico C. Ramalho, Daniela Josa, Melissa S. Caetano and Thais C. S. de Souza

Departamento de Química, Universidade Federal de Lavras (UFLA), 37200-000, Caixa Postal 3037, Lavras, Minas Gerais, Brazil

Received: March 19, 2007; Accepted: April 16, 2007; Revised: April 30, 2007

Abstract: The main COX inhibitors are the non-steroidal anti-inflammatory drugs (NSAIDs). NSAIDs exert anti-inflammatory and analgesic effects through the inhibition of prostaglandin synthesis by blocking COX activity. Currently two COX isoenzymes are known, COX-1 and COX-2. Prostaglandins influenced by COX-1 maintain the integrity of the gastric mucosa. On the other hand, prostaglandins influenced by COX-2 mediate the inflammatory process. The common anti-inflammatory drugs (like aspirin, ibuprofen, and naproxen) all act by blocking the action of both the COX-1 and COX-2 enzymes. The COX-2 inhibitors represent a new class of drugs that do not affect COX-1, but selectively block COX-2. This selective action provides the benefits of reducing inflammation without irritating the stomach. This review will focus on the most recent developments published in the field, paying particular attention to promising COX-2 inhibitors, their chemistry and biological evaluation, and to new chemical and pharmaceutical processes. Moreover, we will discuss recent patents of structural analogs of the COX-2 inhibitors celecoxib and valdecoxib, and novel potential pyridazine, triazole, indole, thione derivatives as a future target for the treatment of inflammation, pain and other diseases.

Keywords: Patents, COX-2, anti-inflammatory drugs.

1. INTRODUCTION

Inflammation is critical for our survival against a hostile world. It is a local reaction of the vascular and supporting elements of a tissue to injury resulting in the formation of protein-rich exudates [1]. Classical inflammation is usually associated with pain and can result from continuing wear of an overactive immune system that is constantly turned on so that the body attacks itself. Physical agents, chemical agents, immunological reactions, and infection by pathogenic organism can cause acute or chronic inflammations [2]. Silent inflammation happens when the immune system continues to attack at the cellular level (heart, immune system and brain) without the perception of pain. These continued insults often result in heart disease, cancer and Alzheimer. Inflammatory responses are thought to be mediated in part by the prostaglandins derived from arachidonic acid (AA) by the action of the enzyme cyclooxygenase. There are drugs, which block the synthesis of prostaglandin through the inhibition of the enzyme cyclooxygenase (COX). Through this mechanism, the drugs belonging to the class of non-steroidal anti-inflammatory drugs (NSAID) show, along with the anti-inflammatory action, negative side effects on the gastrointestinal apparatus.

1.1. Enzyme Cyclooxygenase

The enzyme cyclooxygenase (COX, EC 1.14.99.1) catalyzes the first step of the transformation of arachidonic acid (AA) into different lipid mediators prostaglandins and thromboxanes and other eicosanoids, each of which is involved in some aspect of the inflammatory response [3]. In general, prostaglandins occur at very low levels in tissues, but they have profound biological activities. In structure, they are best considered as derivatives of a C20 saturated fatty acid (prostanic acid), which does not itself occur in nature. Thromboxane A (TXA) contains an unstable bicyclic oxygenated ring structure, while thromboxane B (TXB) has a stable oxane ring [4]. The first step in the synthesis of prostanoids is the release of the substrate fatty acid, such as arachidonic acid, from the cellular phospholipids, by the action of the enzyme phospholipase A2. Next, the free acids are acted upon by one of two isoenzymes, cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2). COX-1 constitutively expressed in most tissues of the human body and provides prostaglandins homeostatic functions, and maintaining the

integrity of the gastric mucosa [3]. On the other hand, prostaglandins influenced by COX-2 are expressed in normal tissues at low levels and it is highly induced by pro-inflammatory mediators in the setting of inflammation, injury, and pain. Fig. 1 illustrates the first step in their synthesis which is the release of a substrate fatty acid, from the cellular phospholipids, by the action of the enzyme phospholipase A2, then the free acids are acted upon by COX-1 and COX-2 (alternatively termed prostaglandin endoperoxide H synthases-1 and -2 denominated PGHS-1 and PGHS-2, respectively). Both enzymes catalyze the same two reactions at different sites. Thus, PGH2 is converted to PGE2 by a prostaglandin E synthase. PGE2 in particular has potent pro-inflammatory effects, including fever induction and pain enhancement.

Cyclooxygenase enzymes require one catalytic turnover from their peroxidase active site to initiate the cyclooxygenase reaction, and the residue of amino acid His 207 (residue numbering based on sheep seminal vesicle COX-1) is critical for peroxidase activity [5]. The COX active site is located at the end of a hydrophobic channel in the enzyme that extends from the membrane-binding region towards the haem region. Inhibition of the COXs is the main target for Non-Steroidal Anti-Inflammatory Drugs (NSAIDs).

1.2. Non-Steroidal Anti-Inflammatory Drugs: Action Mechanisms

There are three classes of cyclooxygenase inhibitors. The differences between them are based on the selectivity regarding the different isozymes, COX-1 and COX-2 [6]. Those classes are: 1) aspirin, is shown to trigger a covalent acetylation in the enzyme, irreversibly blocking its activity. 2) indomethacin, naproxen, ibuprofen, nabumetone, and other traditional Non-Steroidal Anti-Inflammatory Drugs (NSAIDs), and 3) COX-2 specific inhibitors (COXIB) [7]. The common anti-inflammatory drugs (like aspirin, ibuprofen, and naproxen) all act by blocking the action of both the COX-1 and COX-2 enzymes. The COX-2 inhibitors represent a new class of drugs that do not affect COX-1, but selectively block COX-2. COXIBs (celecoxib, parecoxib, rofecoxib, valdecoxib, etoricoxib, lumiracoxib and deracoxib. Fig. 2) were developed to provide a selective agent for COX-2, which at therapeutic doses, demonstrated comparable therapeutic benefit to conventional non-selective NSAIDs, without the attendant COX-1-mediated toxicities.

Celecoxib (1), the first COXIB developed [8] and rofecoxib (2) [9] have been licensed by the Food and Drug Administration (FDA) for the management of inflammatory chronic pain of osteoarthritis,

*Address correspondence to this author at the Departamento de Química, Universidade Federal de Lavras (UFLA), 37200-000, Caixa Postal 3037, Lavras, Minas Gerais, Brazil. Tel: +55-35-38291891; Fax: +55-35-38291271; E-mail: elaine_cunha@ufla.br; effcid@dacafe.com

Analysis of Wild-Type and Gly96Ala Mutant EPSP Synthase Structures via *in Silico* Docking with Inhibitors and Molecular Dynamics Simulation

Melissa S. Caetano, Teodorico C. Ramalho*, Elaine F. F. da Cunha, Daniela Josa and Thais C. S. Souza

Departamento de Química, Universidade Federal de Lavras, Campus Universitário, 37200-000, Lavras-MG, Brazil

Abstract: The high frequency of contamination by herbicides suggests the need for more active and selective herbicides. Glyphosate is the active component of one of the top-selling herbicides, which is also a potent EPSP synthase inhibitor. That is a key enzyme in the shikimic acid pathway, which is found only in plants and some microorganisms. Thus, EPSP synthase is regarded as a prime target for herbicides. In this line, molecular modeling studies using molecular dynamics simulations and molecular docking techniques were performed to understand the interaction of glyphosate and its analogs with the wild type enzyme and Gly96Ala mutant EPSP synthase. Our findings indicate some key points for the designing new selective glyphosate derivatives.

Keywords: EPSP synthase, glyphosate, molecular dynamics, molecular docking.

1. INTRODUCTION

The availability of computer chemistry programs and the internet databases are, currently, fundamental instrument for the discovery and the design of drugs. This information makes possible a quick analysis of biological activity *versus* physical-chemical proprieties of a great quantity of

synthase or other enzymes in this important pathway [3]. The enzyme EPSP synthase catalyzes the transfer of the enolpyruvyl moiety of phosphoenolpyruvate (PEP) to the 5-hydroxyl of shikimate-3-phosphate (S3P) to produce enolpyruvyl shikimate-3-phosphate and inorganic phosphate [4, 11-12]. The EPSP is an important branch-point

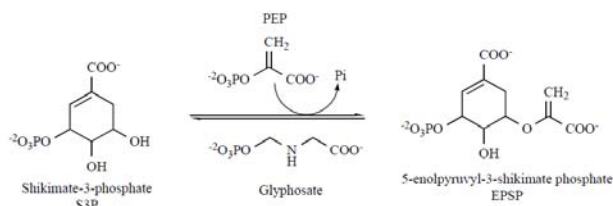


Fig. (1). Reaction pathway of Shikimate-3-phosphate.

interesting molecules [1,2]. This account summarizes recent efforts to more clearly define the molecular mode of action of glyphosate [3], [*N*-(phosphonomethyl)glycine] the world's most successful commercial broad-spectrum herbicide, which the active ingredient of the weed control agent Roundup [4-6]. That herbicide belongs to chemical group derived from glycine, classified as non-selective [7]. It is well-known that the glyphosate inhibits EPSP synthase (5-enolpyruvylshikimate 3-phosphate synthase) in a reversible reaction, which is competitive versus PEP and uncompetitive versus S3P [4, 8-10]. No other classes of commercial herbicides are known which inhibit EPSP

intermediate for the synthesis of aromatic amino acids (phenylalanine, tyrosine, and tryptophan). Glyphosate appears to occupy the binding site of the second substrate of EPSP synthase, phosphoenol pyruvate- PEP, mimicking an intermediate state of the ternary enzyme- substrates complex¹⁴.

While some modification of the PEP structure is tolerated, even minor structural changes in the glyphosate skeleton lead to a significant loss in inhibitor potency and reduced herbicidal activity. Only two closely related analogs, *N*-hydroxyglyphosate (1) and *N*-aminoglyphosate (2), exhibit inhibition properties nearly comparable to glyphosate (3) [3, 15].

One of the first glyphosate-insensitive enzymes reported was a Gly96Ala mutant of EPSP synthase from *Klebsiella pneumoniae*. The mutant enzyme is insensitive to glyphosate with unaltered affinity for its first substrate, shikimate-3-

*Address for correspondence to this author at the Departamento de Química, Universidade Federal de Lavras, Campus Universitário, 37200-000, Lavras-MG, Brazil; Tel: +55-35-38291891; Fax: +55-35-38291271; E-mail: teo@ufla.br

Hypothesis

Homology Modeling of Wild-type, D516V, and H526L *Mycobacterium Tuberculosis* RNA Polymerase and Their Molecular Docking Study with Inhibitors

<http://www.jbsdonline.com>

Daniela Josa
Elaine F. F. da Cunha*
Teodorico C. Ramalho
Thais C. S. Souza
Melissa S. Caetano

Abstract

Rifamycins (Rifs) are antibiotic widely used for the treatment of tuberculosis (TB); nevertheless, their efficacy has been limited by a high percentage of mutations, principally in the *rpoB* gene. In this work, the first three-dimensional molecular model of the hypothetical structures for the wild-type and D516V and H526L mutants of *Mycobacterium tuberculosis* (*m*RNA Polymerase) were elucidated by a homology modeling method. In addition, the orientations and binding affinities of some Rifs with those new structures were investigated. Our findings could be helpful for the design of new more potent rifamycin analogs.

Key words: Homology modeling; Rifamycins; *Mycobacterium tuberculosis* RNA polymerase; and Docking.

Introduction

Rifamycins (Rifs) are antibiotics extremely effective against tuberculosis and exert their activity by specific inhibition of bacterial DNA-dependent RNA polymerase (DDRP) (1, 2). Introduced in 1968 rifampin (RIF) has promoted a meaningful advance in tuberculosis (TB) chemotherapy (3). However its efficiency has been limited by mutations principally in *rpoB* gene (4-9). The mutations in the position 516, 526, and 531 codon, are responsible for approximately 86% of the strains reported (10). Rif-resist TB has contributed to become a global emergency with about 8 million new TB cases and 2 million deaths worldwide annually (9). Whereas several Rif analogs have been synthesized (1, 11-13), surprisingly little attention has been given to the molecular base of their action. To worsen this picture, the 3D-structure of *Mycobacterium Tuberculosis* RNA polymerase is unknown so far.

The present paper describes the first three-dimensional molecular model of the hypothetical structures for the wild-type and D516V and H526L mutants of *m*RNA Polymerase obtained through the use of the homology modeling technique. In addition, we investigated the orientations and binding affinities of some Rifs with those new structures.

Methodology

The amino acid primary sequence of *m*RNA Polymerase was retrieved from the Swiss-Model Data Bank (14-17). The search for sequences similar to *m*RNA Polymerase within the Protein Data Bank (PDB) (18) was performed with the Basic Local Alignment Search Tool program (BLAST) (19). The search for the best template for modeling was carried out by choosing structures possessing a high degree of sequence similarity with *m*RNA Polymerase. The crystal structural coordinates of RNA polymerase of *Thermus thermophilus* (*t*RNA Polymerase) at 2.5 Å resolution (PDB code: 2A69) (20) was used as template structure to build a three-dimensional model of *m*RNA Polymerase. Sequence align-

Departamento de Química
Universidade Federal de Lavras
Campus Universitário
Lavras, 3037, 37200-000, MG, Brazil

*Phone: +55-35-38291891
Fax: +55-35-38291271
Email: elaine_cunha@ufla.br
effcid@dcafe.com

373

RESEARCH ARTICLE

Molecular modelling of *Mycobacterium tuberculosis* acetolactate synthase catalytic subunit and its molecular docking study with inhibitors

Thais C.S. Souza, Daniela Josa, Teodorico C. Ramalho, Melissa Soares Caetano and Elaine F.F. da Cunha*

Departamento de Química, Universidade Federal de Lavras, Lavras, MG, Brazil

(Received 16 February 2008; final version received 13 April 2008)

Mycobacterium tuberculosis is a leading cause of infectious disease in the world today. This outlook is aggravated by a growing number of *M. tuberculosis* infections in individuals who are immunocompromised as a result of HIV infections. Thus, new and more potent anti-TB agents are necessary. Therefore, acetolactate synthase (mtALS) was selected as a target enzyme to combat *M. tuberculosis*. In this work, the three-dimensional molecular model of the hypothetical structure for the ALS catalytic subunit of *M. tuberculosis* was elucidated by homology modelling. In addition, the orientations and binding affinities of sulphonylurea inhibitors with the new structure was investigated. Our findings could be helpful for the design of new, more potent mtAHAS inhibitors.

Keywords: homology modelling; molecular docking; sulphonylureas; *Mycobacterium tuberculosis* acetolactate synthase

Introduction

Acetolactate synthase (ALS; also known as aceto-hydroxyacid synthase; EC 2.2.1.6) is the first common enzyme in the biosynthetic pathway of the branched-chain amino acids in bacteria, fungi and higher plants [1]. The enzyme catalyses two parallel reactions: the condensation of two pyruvate molecules to give rise to 2-acetolactate in the first step of the valine and leucine synthetic pathway, and the condensation of pyruvate and 2-oxobutyrate to yield 2-aceto-2-hydroxybutyrate in the second step of isoleucine biosynthesis [2]. Thiamin diphosphate (ThDP) is the cofactor of this enzyme and coordinates to a divalent metal ion such as Mg^{2+} . In addition, ALS also binds a molecule of FAD, although this cofactor is unusual because the reactions catalysed by ALS involve no net reduction or oxidation [3]. Bacterial ALSs are composed of large and small subunits, in an $\alpha_2\beta_2$ heterotetramer. It has been generally accepted that the large (60–70 kDa) subunits are catalytic, where pairs of catalytic subunits form an intimate dimer containing two active sites. The small ones are regulatory (9.5–54 kDa) [4]. The molecular size of the plant and fungal enzymes is less certain and the enzymes, as isolated, frequently contain the catalytic subunits only [5].

ALS constitutes a target for five structurally diverse inhibitor classes: sulphonylureas, imidazolinones, triazolopyrimidines, pyrimidinylthio (or oxy)-benzoates and sulphonylamino-carbonyl triazolinones [1]. The sulphonylurea herbicides were discovered by Levitt et al. and the most active compounds show an aromatic ring attached to the

sulphur atom and a heterocyclic ring (pyrimidine or triazine) attached to the distal nitrogen atom. The most known are: pyrazosulfuron ethyl (PSE), primisulfuron methyl (PSM), sulfometuron methyl (SM), metsulfuron methyl (MM), chlorimuron ethyl (CE), chlorsulfuron (CS), thifensulfuron methyl (TM), nicosulfuron (NS) and triasulfuron (TS). The sensitivity of ALS to the different sulphonylureas varies substantially. For example, Duggleby et al. have purified the wild-type and mutant yeast ALS and determined their sensitivity to six sulphonylureas [6]. Yeast ALS showed apparent K_i values of 3.25 (CE), 5.08 (SM), 9.40 (MM), 5.25 (TM) and 127.0 nM (CS). They observed that yeast ALS has an extended N-terminal sequence that is not present in bacterial ALS. Additionally, Chang et al. expressed, purified and characterised the *Arabidopsis thaliana* ALS and observed that it is strongly inhibited by the sulphonylurea herbicides [7,8]. *A. thaliana* ALS showed apparent K_i values of 10.8 (CE), 25.5 (SM), 36.2 (MM), 72.2 (TM) and 54.6 nM (CS). Choi et al. searched the inhibitory activities of various sulphonylurea herbicides against the catalytic subunit of *Mycobacterium tuberculosis* acetolactate synthase (mtALS). They suggested that ALS might serve as a target protein for the development of novel anti-tuberculosis therapeutics [9,10].

M. tuberculosis is a leading cause of morbidity and mortality in the world today [11–13]. This fact is aggravated by a growing number of Mtb infections in individuals who are immunocompromised as a result of HIV infections [14]. In addition, the ongoing selection of

*Corresponding author. Email: elaine_cunha@ufla.br; effcid@dacafe.com