



MARIA CAROLINA GUIMARÃES

**PLANEJAMENTO COMPUTACIONAL DE
COMPOSTOS ANTI-HIV-1 COM MÚLTIPLOS
ALVOS**

LAVRAS – MG

2014

MARIA CAROLINA GUIMARÃES

**PLANEJAMENTO COMPUTACIONAL DE COMPOSTOS ANTI-HIV-1
COM MÚLTIPLOS ALVOS**

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

Orientador

Dr. Matheus Puggina de Freitas

LAVRAS – MG

2014

**Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e
Serviços da Biblioteca Universitária da UFLA**

Guimarães, Maria Carolina.

Planejamento computacional de compostos anti-HIV-1 com múltiplos alvos / Maria Carolina Guimarães. – Lavras : UFLA, 2014.

119 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2014.

Orientador: Matheus Puggina de Freitas.

Bibliografia.

1. AIDS. 2. Medicamentos. 3. Antirretrovirais. 4. Química computacional. I. Universidade Federal de Lavras. II. Título.

CDD – 616.9792

MARIA CAROLINA GUIMARÃES

**PLANEJAMENTO COMPUTACIONAL DE COMPOSTOS ANTI-HIV-1
COM MÚLTIPLOS ALVOS**

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

APROVADA em 26 de fevereiro de 2014.

Dra. Elaine Fontes Ferreira da Cunha UFLA

Dra. Melissa Soares Caetano UFOP

Dr. Matheus Puggina de Freitas
Orientador

**LAVRAS – MG
2014**

AGRADECIMENTOS

Agradeço, em primeiro lugar, a Deus, por ser a força que me impulsiona a prosseguir e a luz que ilumina o meu caminho.

A toda a minha família, em especial aos meus amados pais, Regina e Márcio; ao meu irmão, Lucas e ao meu noivo, Cristiano. Sem a presença e o apoio de vocês esta conquista não seria possível.

Obrigada ao meu orientador, Matheus, por todos os ensinamentos e aos membros da banca, pela disposição e considerações que, certamente, em muito contribuirão para este trabalho.

Aos meus amigos de toda a vida e aos professores e colegas do Laboratório de Modelagem Molecular, por todo o auxílio e colaboração nos momentos de dificuldades.

À Universidade Federal de Lavras, pela oportunidade de realizar este trabalho, bem como ao CNPq, pelo financiamento dos recursos necessários ao desenvolvimento do mesmo.

A todos o meu muito obrigada!

RESUMO

Compostos anti-HIV-1 compreendem inibidores de alguns alvos biológicos diferentes, tais como as enzimas HIV-protease, integrase e transcriptase reversa, assim como proteínas de fusão ou entrada, visando reduzir ou bloquear a replicação do vírus HIV no organismo infectado. Atualmente, a terapia medicamentosa se dá pela combinação de fármacos administrados como coquetéis inibidores das diferentes etapas da replicação viral. Com o objetivo de propor compostos anti-HIV-1 com múltiplos alvos como alternativa para um tratamento menos exaustivo, foram desenvolvidos modelos MIA-QSAR para compostos com atividade denominada anti-HIV-1 frente às células da linhagem MT4 (modelo 1) e inibidores da enzima transcriptase reversa (modelo 2) que apresentaram resultados estatísticos satisfatórios obtidos das análises, sendo para o modelo 1 r^2_{cal} de 0,934, q^2 de 0,677 e r^2_{teste} de 0,639 e, para o modelo 2 r^2_{cal} de 0,910, q^2 de 0,663 e r^2_{teste} de 0,667. Em razão da similaridade estrutural entre os compostos destas duas classes, novos compostos (**A-F**) foram propostos, os quais tiveram suas bioatividades preditas pelos modelos construídos, além de passarem por estudos de ancoramento molecular (*docking*) para validar tais predições. Um novo composto proposto, **B**, apresentou-se como um promissor agente com múltiplos alvos, com elevado valor predito de pIC_{50} (8,69) e *docking score* (-129,7 kcal mol⁻¹) maior ou semelhante aos demais compostos, além de ligações de hidrogênio adicionais com os resíduos de aminoácidos Lys-101, Lys-103 e Tyr-318 da enzima transcriptase reversa, bem como interação π -*stacking* com Tyr-188. Análises teóricas preliminares dos parâmetros farmacocinéticos evidenciaram os compostos **A** e **B** como comparáveis aos melhores compostos das 3 séries (compostos **112** e **113**), podendo ser bem absorvidos, distribuídos, metabolizados e excretados.

Palavras-chave: Antirretrovirais. Transcriptase reversa. Múltiplos alvos. QSAR, MIA-QSAR.

ABSTRACT

HIV-1 inhibitors include compounds of some different biological targets, such as the HIV protease, reverse transcriptase and integrase enzymes, as well as fusion or entrance proteins, to reduce or block the replication of the HIV virus in the infected organism. Currently, drug therapy is given by the combination of drugs administered as cocktails of inhibitors of the different stages of viral replication. With the aim of proposing anti-HIV-1 compounds with multiple target as an alternative to improve the treatment, MIA-QSAR models were developed for compounds with anti-HIV-1 activity towards the MT4 cell line (model 1) and inhibitors of the reverse transcriptase enzyme (model 2), giving satisfactory statistical results, that is, Model 1 with r^2_{cal} of 0.934, q^2 of 0.677 and r^2_{test} of 0.639, while for Model 2 r^2_{cal} of 0.910, $q^2 = 0.663$ and r^2_{test} 0.667. Given the structural similarity between the compounds of these two classes, new compounds (**A-F**) were proposed and their bioactivities were estimated by the QSAR models, and additional docking studies were performed to validate these predictions. The proposed compound **B** was calculated to be a promising agent toward multiple targets due to its high predicted pIC_{50} value (8.69) and docking score (-129.7 kcal mol⁻¹), which are larger than or similar to other compounds, possibly because of hydrogen bonding with the amino acid residues Lys-101, Lys -103 and Tyr-318 of the reverse transcriptase enzyme and π -stacking interactions with Tyr-188. Preliminary theoretical analysis of pharmacokinetic parameters showed that compounds **A** and **B** are comparable to the best compounds of each class (compounds **112** and **113**) and, therefore, they are expected to be well absorbed, distributed, metabolized and excreted.

Keywords: Antiretroviral drugs. Reverse transcriptase. Multiple targets. QSAR. MIA-QSAR.

LISTA DE FIGURAS

Figura 1	Estrutura de um vírus e morfologia do retrovírus causador da AIDS.....	21
Figura 2	Ciclo de vida resumido da replicação viral do HIV	24
Figura 3	Fórmula estrutural da zidovudina e da lamivudina	25
Figura 4	Estruturas químicas dos inibidores análogos de nucleosídeos da transcriptase reversa (INTR) disponíveis no mercado para uso clínico	27
Figura 5	Representantes dos inibidores análogos não nucleosídeos da TR para o tratamento da AIDS	28
Figura 6	Estruturas químicas dos medicamentos disponíveis no mercado, capazes de inibir a protease viral.....	29
Figura 7	Sequência de aminoácidos do fármaco enfuvirtina	32
Figura 8	Estrutura química do fármaco maraviroc.....	33
Figura 9	Estrutura química do fármaco raltegravir	33
Figura 10	Estrutura química do fármaco etravirina	34
Figura 11	Adsorção e fusão do vírus HIV-1 à membrana plasmática da célula infectada	36
Figura 12	Enzima HIV-1 TR	37
Figura 13	Enzima HIV-1 IN e seus domínios.....	39
Figura 14	HIV-1 PR, com destaque (círculo vermelho) para as alças que se fecham sobre o sítio ativo da enzima, quando complexada.....	40
Figura 15	Estruturas químicas dos análogos de nucleotídeos da TR adefovir e tenofovir. Os compostos 1 e 2 são seus pró-fármacos desenvolvidos para contornar o problema da baixa solubilidade em água e biodisponibilidade, quando administrados via oral.....	42
Figura 16	Etapas do processo de descoberta e desenvolvimento de fármacos.....	45
Figura 17	Conversão das estruturas químicas em uma matriz de dados binários, em que os <i>pixels</i> pretos da imagem da estrutura correspondem ao dígito 0 e <i>pixels</i> brancos (espaços em branco) correspondem ao dígito 765, de acordo com o sistema de cores RGB.....	55
Figura 18	Área de trabalho do aplicativo <i>Paint</i> na dimensão 500x500 <i>pixels</i> . Escolha do <i>pixel</i> no aumento de 800%, para alinhamento e fixação na coordenada 250x190 <i>pixels</i>	68
Figura 19	Sobreposição das estruturas químicas e formação do arranjo tridimensional. Em negrito está a porção similar das estruturas sobrepostas e as porções variáveis são os substituintes que	

	respondem pela variação na atividade biológica de cada composto.....	69
Figura 20	Interface do programa Matlab, utilizado na geração da matriz X de descritores. Abaixo são ilustrados a construção do arranjo tridimensional e o desdobramento da matriz	72
Figura 21	Calibração do modelo MIA-QSAR por meio da regressão por mínimos quadrados parciais (PLS).....	76
Figura 22	Validação externa ou <i>test set</i>	77
Figura 23	Randomização do bloco Y e sucessiva regressão com a matriz X intacta via PLS. O processo é repetido algumas vezes, para atestar a correlação do modelo obtido pela calibração	78
Figura 24	Compostos biologicamente mais ativos das séries 1 e 3, de acordo com os valores de atividade experimental, que foram combinados (parte em vermelho de uma molécula com a parte preta de outra) para a proposta de novos inibidores com atividade anti-HIV-1.....	91
Figura 25	Compostos biologicamente mais ativos, experimentalmente, das séries 2 e 3, que foram combinados (parte em vermelho de uma molécula com a parte preta de outra) para a proposta de novos inibidores com atividade inibitória da TR, conforme modelo MIA-QSAR 2.....	92
Figura 26	Novos compostos com atividade anti-HIV-1 (A-C), inibidor da TR (D) e compostos propostos com dupla ou mais atividades (E e F). Os compostos A-D tiveram suas atividades biológicas preditas pelos modelos construídos de acordo com suas respectivas atividades exercidas. Contudo, os valores de <i>docking score</i> (kcal mol ⁻¹) são referentes apenas à enzima TR.....	93
Figura 27	Estruturas químicas do AZT e dos compostos 112 e 113 que, ao serem combinadas, deram origem aos compostos E e F com suposta múltipla atividade antirretroviral	95
Figura 28	Composto B ancorado no sítio ativo da enzima TR.....	99

LISTA DE GRÁFICOS

Gráfico 1	Gráfico de <i>scores</i> da componente principal 2 (CP2) eixo Y <i>versus</i> componente principal 1 (CP1) eixo X	85
Gráfico 2	Determinação do número ótimo de variáveis latentes PLS para o modelo1	86
Gráfico 3	pIC ₅₀ experimentais <i>versus</i> ajustados/preditos da série de compostos com atividade anti-HIV-1 (modelo 1).....	87
Gráfico 4	Número ótimo de variáveis latentes (PLS) para o modelo 2	88
Gráfico 5	pIC ₅₀ experimentais <i>versus</i> ajustados/preditos da série de compostos com atividade inibitória de TR (modelo 2).....	90
Gráfico 6	Correlação entre valores de pIC ₅₀ experimentais dos inibidores da TR selecionados e os valores de <i>docking scores</i>	98

LISTA DE TABELAS

Tabela 1	Histórico da metodologia QSAR, em ordem cronológica.....	49
Tabela 2	Séries de compostos usadas na construção dos modelos MIA-QSAR.....	63
Tabela 3	Resultados estatísticos obtidos para o modelo por meio da utilização de 9 variáveis latentes PLS.....	86
Tabela 4	Parâmetros estatísticos obtidos para o modelo MIA-QSAR 2	89
Tabela 5	Cálculo dos parâmetros da regra dos cinco de Lipinski e TPSA para avaliação de alguns dos compostos com valores de atividades experimentais e para os compostos propostos (A-F)	101
Tabela 6	Parâmetros farmacocinéticos computacionais para os compostos com atividades experimentais e para os compostos propostos (A - F).....	102

LISTA DE ABREVIATURAS

3TC	Lamivudina
AZT	3-azido-2,3-didesoxitimidina
FTC	Entricitabina
ddl	Didanosina
ABC	Abacavir
bmp	Bitmaps
Br	Bromo
CH ₃	Metil
gp	Glicoproteína
Log	Logaritmo
Lys	Lisina
Tyr	Tirosina
U	Inibidores e fusão ou entrada

LISTA DE SIGLAS

3D	Terceira dimensão
4D	Quarta dimensão
5D	Quinta dimensão
6D	Sexta dimensão
7D	Sétima dimensão
ADME	Absorção, distribuição, metabolismo e excreção
ADMET	Absorção, distribuição, metabolismo, excreção e toxicidade
AIDS	<i>Acquired Immunodeficiency Syndrome</i>
ANVISA	Agência Nacional de Vigilância Sanitária
BBB	<i>Blood brain barrier penetration</i>
CP	Componente principal
CP1	Componente principal 1
CP2	Componente principal 2
DNA	<i>Deoxyribonucleic acid</i>
DQI	Departamento de Química
DST	Doença sexualmente transmissível
ED	<i>Effective dose</i>
F	<i>Oral bioavailability</i>
FDA	<i>Food and drug administration</i>
HAART	<i>Highly active antiretroviral therapy</i>
HCA	<i>Hierarchical clusters methods</i>

HIA	<i>Human intestinal absorption</i>
HIV	<i>Human Immunodeficiency Virus</i>
HQSAR	Holograma QSAR
IC	<i>Inhibitory molar concentration</i>
ICTV	<i>International Committee on Taxonomy of Viruses</i>
IN	Integrase
INTR	Inibidor nucleosídeo da Transcriptase Reversa
INNTR	Inibidor não nucleosídeo da Transcriptase Reversa
LOOCV	<i>Leave one out cross validation</i>
MIA-QSAR	<i>Multivariate image analysis-QSAR</i>
MLR	Regressão linear múltipla
MVD	Molegro virtual docker
MW	<i>Molecular weight</i>
NDA	<i>New drug application</i>
OMS	Organização Mundial da Saúde
PCA	<i>Principal component analysis</i>
PCR	<i>Principal Components Regression</i>
PDB	<i>Protein data bank</i>
PLS	<i>Partial least square</i>
PR	Protease
QSAR	<i>Quantitative structure activity relationship</i>
RGB	<i>Red green blue</i>
RML	Regressão linear múltipla
RMSE	<i>Root mean square error</i>
RMSECV	<i>Root mean square error cross validation</i>
RNA	<i>ribonucleic acid</i>
SUS	Sistema Único de Saúde
TPSA	<i>Topological polar surface area</i>
TR	Transcriptase reversa
UFLA	Universidade Federal de Lavras
VL	Variável latente

LISTA DE SÍMBOLOS

Log P	Coeficiente de partição octanol/água
r^2	Coeficiente de correlação ao quadrado
q^2	Coeficiente de correlação com validação cruzada
pIC_{50}	Log de IC_{50}
IC_{50}	50% de concentração inibitória em mol L ⁻¹
r^2_{Cal}	Coeficiente de correlação da calibração
pIC_{50Exp}	Log de IC_{50} experimental
$pIC_{50LOOCV}$	Log de IC_{50} LOOCV
pIC_{50Pred}	Log de IC_{50} predito
r^2_{Teste}	Coeficiente de correlação da validação externa
r^2_m	Coeficiente de correlação modificado
R^2_{pe}	Diferença entre coeficiente de correlação randomizado/calibrado
E_{Score}	Função <i>score</i>
E_{Inter}	Energia de interação ligante-proteína
n_{ON}	n° de aceptores de ligação de hidrogênio
n_{OHNH}	n° de doadores de ligação de hidrogênio
y_i	Valores de pIC_{50}
\hat{y}_i	Valores estimados/preditos de pIC_{50}
$\bar{\square}$	Valores médios de pIC_{50}
r^2_0	Coeficiente de correlação ao quadrado sem intercepto
R^2_{rand}	Coeficiente de correlação do $Y_{randomization}$
Kcal mol ⁻¹	Quilocaloria por mol
n_{viol}	N de violações à regra de Lipinski
n_{rotb}	Valores de flexibilidade molecular

SUMÁRIO

1	INTRODUÇÃO	15
2	REFERENCIAL TEÓRICO	18
2.1	AIDS.....	18
2.2	HIV.....	19
2.3	Fármacos antirretrovirais	24
2.3.1	Terapia antirretroviral altamente ativa.....	30
2.4	Atividade anti-HIV-1	35
2.5	Inibidores da transcriptase reversa (anti-TR-1).....	41
2.6	A química medicinal e o planejamento de fármacos.....	44
2.7	QSAR como ferramenta para a elaboração de novos candidatos a fármacos	46
2.7.1	Critérios e regras em metodologias QSAR	51
2.7.2	Análise multivariada de imagens aplicada ao QSAR	53
2.8	<i>Docking</i> e análises de parâmetros farmacocinéticos	56
3	OBJETIVO GERAL.....	59
3.1	Objetivos específicos	59
4	METODOLOGIA.....	59
4.1	A construção dos modelos MIA-QSAR.....	61
4.2	Identificação de <i>outliers</i>	74
4.3	Validações e calibração dos modelos MIA-QSAR	75
4.4	<i>Docking</i>	79
4.5	Parâmetros farmacocinéticos e biodisponibilidade	80
5	RESULTADOS E DISCUSSÃO.....	82
5.1	Modelo 1: anti-HIV-1	83
5.2	Modelo 2: anti-TR-1	87
5.3	Proposta de novos possíveis inibidores.....	90
5.4	Estudos de <i>docking</i> ou ancoramento molecular	95
5.5	Análises dos parâmetros farmacocinéticos	99
6	CONCLUSÃO	104
	REFERÊNCIAS	105

1 INTRODUÇÃO

O vírus da imunodeficiência humana (HIV, do inglês *human immunodeficiency virus*) é o retrovírus patogênico causador da AIDS, doença caracterizada pela perda progressiva de linfócitos T auxiliares CD4⁺ do sistema imunológico humano, responsável pela defesa do organismo contra infecções e doenças (CHEN et al., 2008). A depressão de tais células, provocada pela proliferação do vírus HIV no organismo infectado, favorece o desenvolvimento de doenças oportunistas, complicações neurológicas e neoplasias que são, na maioria dos casos, responsáveis pelo alto índice de óbitos em portadores da doença (HUTCHINSON, 2001).

Essa epidemia continua a ser um dos grandes desafios para a saúde global, uma vez que, segundo a Organização Mundial da Saúde (OMS), cerca de 35,3 milhões de pessoas vivem com HIV em todo o mundo. Somente em 2012, 2,3 milhões de pessoas foram infectadas e 1,6 milhão morreram em decorrência da doença, principalmente devido ao acesso inadequado aos serviços de tratamento e atenção (AIDS EPIDEMIC UPDATE, 2013).

O estudo do ciclo replicativo do HIV possibilitou a identificação de alguns alvos macromoleculares suscetíveis à intervenção terapêutica (BRITO, 2011). Atualmente, o tratamento consiste no uso combinado de diferentes antirretrovirais, conhecidos como coquetéis, que atuam como inibidores das etapas chave do ciclo de reprodução do vírus: fusão, transcrição reversa, integração, transcrição e tradução com formação das proteínas virais (BRITO et al., 2008; CUNICO; GOMES; VELLASCO JUNIOR, 2008).

Uma vez que não existe cura para a doença, o uso dos medicamentos anti-HIV tem como objetivo reduzir ao máximo a carga viral no menor intervalo de tempo possível, evitando o enfraquecimento do sistema imunológico do

organismo infectado, proporcionando melhor qualidade e expectativa de vida ao indivíduo portador da doença.

Apesar dos inúmeros avanços na terapêutica contra o vírus HIV, a eficácia do tratamento é limitada pelo surgimento de cepas resistentes aos antirretrovirais disponíveis, à toxicidade desses fármacos e pela baixa resposta terapêutica de alguns indivíduos. Como consequência, a descoberta de novos alvos terapêuticos, bem como o desenvolvimento de potentes inibidores para os alvos já conhecidos, continua sendo um desafio para a comunidade científica (CUNICO; GOMES; VELLASCO JUNIOR, 2008).

Nesse sentido, técnicas computacionais, como a relação quantitativa entre estrutura química e atividade (QSAR, do inglês *Quantitative Structure Activity Relationship*), têm se destacado como uma ferramenta eficaz no desenvolvimento de novos protótipos bioativos, ao se combinar subestruturas químicas (FREITAS, 2007; GOODARZI; FREITAS; FERREIRA, 2009), evitando a síntese aleatória e testes biológicos onerosos nas etapas que envolvem a síntese e produção dos fármacos (FREITAS; BROWN; MARTINS, 2005). Além disso, a aplicação de estudos de ancoramento molecular ou *docking* são úteis na validação e auxiliam na interpretação dos dados bioquímicos, associados a estudos preliminares de parâmetros farmacocinéticos ADME (absorção, distribuição, metabolismo e excreção).

Visando colaborar para um tratamento menos exaustivo para os pacientes que fazem uso dos coquetéis anti-HIV, o presente trabalho foi realizado com o objetivo de elaborar novos protótipos de antirretrovirais que apresentem atividade protetora das células da linhagem MT4, ou seja, impedem a infecção dos linfócitos T CD4⁺ pelo vírus HIV-1 e à transcrição reversa do HIV simultaneamente, pela aplicação da metodologia MIA-QSAR e pela posterior combinação de subestruturas dos compostos com maiores atividades biológicas de cada série de dados utilizada no estudo. Também foram realizados

estudos de *docking* e análise do perfil farmacocinético das estruturas propostas, a fim de validar o caráter preditivo do modelo construído e colaborar para um melhor entendimento do perfil farmacológico dos novos inibidores propostos.

2 REFERENCIAL TEÓRICO

2.1 AIDS

A Síndrome da Imunodeficiência Adquirida (AIDS, do inglês *Acquired Immunodeficiency Syndrome*) é a doença infecciosa sexualmente transmissível que mais mata em todo o mundo, fato que a mantém como um importante problema de saúde pública mundial. É uma doença que não tem cura e a maior parte da população infectada vive em países em desenvolvimento, sobretudo na África. Contudo, no Brasil, já foram notificados cerca de 600 mil casos (CUNICO; GOMES; VELLASCO JUNIOR, 2008).

Acredita-se que o vírus da imunodeficiência símia (SIV, do inglês *simian immunodeficiency virus*), que acomete o sistema imunológico de macacos, tenha chegado à espécie humana após o contato de caçadores humanos com primatas através do sangue, da carne ou da mordida de uma espécie de chimpanzé infectado (HAHN et al., 2000). Embora os primeiros sintomas de uma doença até então desconhecida tenha surgido por volta de 1978, em homossexuais dos Estados Unidos e da Suécia (HARMS; FELDMER, 2002), a doença foi descoberta em 1981, nos Estados Unidos, tendo o vírus HIV sido isolado em 1983, a partir de um paciente com linfadenopatia generalizada (CHEN et al., 2008).

Em 1984, foi realizado o primeiro isolamento do vírus em indivíduos assintomáticos associado à AIDS, sendo denominado estado de portador do vírus e, a partir daí, por fazer parte da família *Retroviridae*, foi classificado como um retrovírus (CUNNINGHAM et al., 1996; LEVY, 1996), ou seja, vírus que apresenta genoma constituído por duas fitas simples de RNA, cuja principal característica é a transcrição reversa do material genético por meio da enzima transcriptase reversa viral.

Após atingir e adaptar-se à espécie humana, o retrovírus HIV tornou-se transmissível entre humanos, principalmente através de sangue, sêmen, secreção vaginal e leite materno de pessoas infectadas, fase facilitada pelo aumento na densidade populacional e pelo comportamento da população (HILL; ROSENBLOOM; NOWAK, 2012).

É importante ressaltar que a infecção pelo HIV não determina a AIDS imediatamente e os aspectos de como e quando os pacientes infectados irão progredir para a expressão da doença permanecem controversos (SALDANHA; ANDRADE; BECK, 2009).

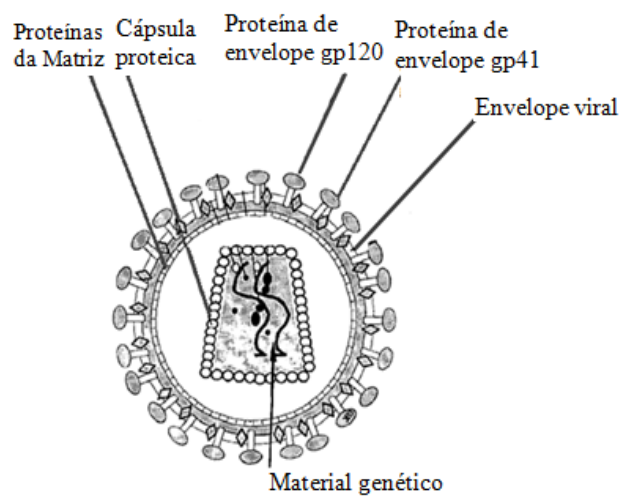
Segundo o Departamento de DST (doenças sexualmente transmissíveis), AIDS e Hepatites Virais do Ministério da Saúde, a AIDS corresponde ao estágio mais avançado da infecção pelo HIV, quando o sistema imunológico já não consegue exercer sua função de defesa devido à redução considerável do número de linfócitos T CD4⁺, o que torna o organismo vulnerável a diversas doenças, desde um simples resfriado a infecções mais graves, como tuberculose ou câncer que, muitas vezes, levam o paciente a óbito (BRITO et al., 2000).

2.2 HIV

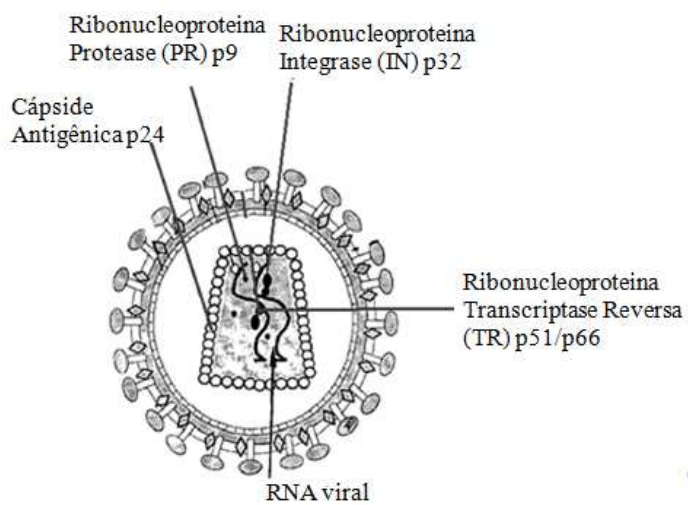
O HIV é um dos retrovírus da família *Retroviridae* e gênero *Lentivirus* (do latim *lentus*) que, como tal, compartilham determinadas propriedades, como longo período de incubação antes da manifestação dos primeiros sintomas da doença, infecção das células do sangue e do sistema nervoso e supressão do sistema imune (VAISHNAV; WONSTAAL, 1991).

Ele se constitui de um germe microscópico com 80 a 100 nm de diâmetro, com envelope lipoproteico ligeiramente esférico, capsídeo cilíndrico e genoma formado por duas fitas molde (*sense*) simples de RNA linear (FERREIRA; RIFFEL; SANT'ANA, 2010).

Por apresentar o RNA como material genético, o HIV deve dispor de enzimas próprias para que seja processado, existindo sob a forma de um complexo de ribonucleoproteínas que, como mostrado na Figura 1, é constituído pela transcriptase reversa TR (p51/p66), a protease PR (p9) e a integrase IN (p32), sendo o núcleo viral composto por uma grande cápside antigênica (p24) (ROY; LEONARD, 2005). Além disso, devido à sua incapacidade de autorreplicação, precisa infectar uma célula que servirá como hospedeira para a produção de novos vírus (VAISHNAV; WONSTAAL, 1991).



(a) Estrutura viral



(b) Retrovirus HIV

Figura 1 Estrutura de um vírus e morfologia do retrovírus causador da AIDS
 Fonte adaptado de Google Imagens (2013)

Existem dois tipos de HIV, o HIV-1 e o HIV-2, com similaridade genômica de apenas 40% (FERREIRA; RIFFEL; SANT'ANA, 2010). O HIV-1 corresponde à estirpe viral responsável pela maior epidemia mundial e é mais virulento e mutável que o HIV-2, que causa a AIDS mais lentamente e está restrito, principalmente, à África Ocidental (HILL; ROSENBLOOM; NOWAK, 2012). Como consequência, os indivíduos infectados apenas com o HIV-2 têm, em geral, sobrevida mais longa (BRITO, 2011).

Os alvos da infecção pelo HIV em humanos são células que apresentam receptores CD4 em suas superfícies, como macrófagos (células que percorrem o organismo combatendo microorganismos) e, principalmente, os linfócitos T-auxiliares do sistema imunológico que se encontram em circulação na corrente sanguínea (BRITO, 2011; RAFFANTI; HAAS, 2006). Tais receptores são glicoproteínas monoméricas da superfamília das imunoglobulinas que apresentam quatro domínios e são utilizados pelo HIV-1 para adentrarem nas células alvo.

As duas glicoproteínas virais do envelope, gp120 de superfície e gp41 transmembrana, interagem com os receptores celulares CD4, favorecendo a fusão vírus-célula, que permite que o material genético do vírus HIV seja lançado no citoplasma da célula hospedeira, dando início ao processo de replicação viral.

Ainda no citoplasma da célula infectada, o RNA genômico é transcrito em uma fita dupla de DNA pró-viral pela TR viral e, após a transcrição reversa, o DNA é integrado às proteínas virais e celulares. Este complexo nucleoproteico é, então, transportado para o núcleo celular e a dupla fita de DNA é inserida no cromossomo do hospedeiro por ação da IN viral, o que permite a ocorrência das primeiras transições do DNA pró-viral pela RNA polimerase II celular com a produção de RNAs genômico e mensageiro virais. No citoplasma, as fitas de RNA mensageiro virais são traduzidas produzindo as poliproteínas que darão

origem às proteínas virais (CIUFFI; BUSHMAN, 2006; FERREIRA; RIFFEL; SANT'ANA, 2010; SIERRA; KUPFER; KAISER, 2005).

Por fim, as proteínas do envelope viral migram e se inserem na membrana plasmática juntamente com as poliproteínas Gag e Gag-Pol, na qual começam a montagem do vírion (partícula viral completa) direcionada pela poliproteína Gag. Dessa forma, as enzimas virais, o RNA genômico e os demais compostos celulares se associam no nucleocapsídeo imaturo que, mais tarde, irá brotar da membrana plasmática ainda como um vírion imaturo (FERREIRA; RIFFEL; SANT'ANA, 2010; SIERRA; KUPFER; KAISER, 2005) como mostrado, de maneira simplificada, na Figura 2.

A maturidade irá ocorrer com a ativação da PR através do brotamento que, autocataliticamente, cliva as poliproteínas Gag e Gag-Pol, liberando as proteínas estruturais e as enzimas. As proteínas individuais sofrerão futuras interações, formando o nucleocapsídeo cônico e a proteína MA é associada ao envelope viral. É o processamento das proteínas virais pela protease que, além de amadurecer os vírions formados e liberados, os tornam partículas virais infecciosas (CIUFFI; BUSHMAN, 2006; FERREIRA; RIFFEL; SANT'ANA, 2010; SIERRA; KUPFER; KAISER, 2005). Assim, os vírions liberados da replicação estão prontos para infectar novas células e proliferar por todo o organismo.

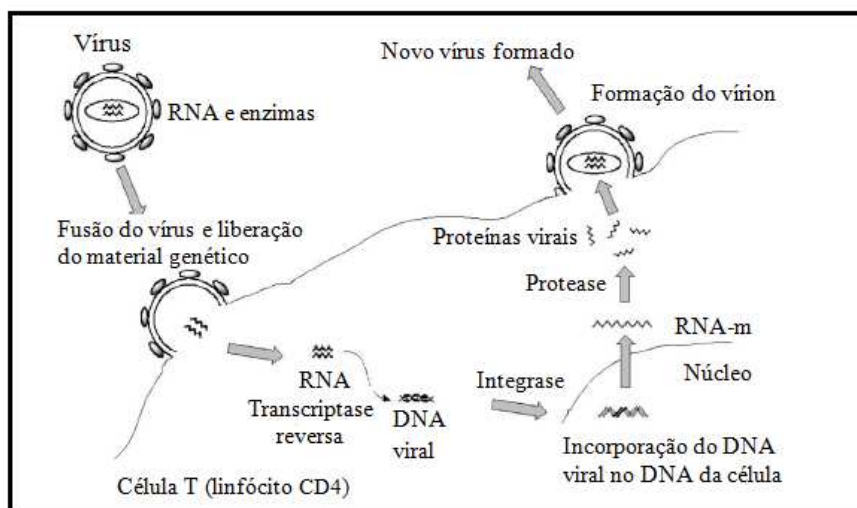


Figura 2 Ciclo de vida resumido da replicação viral do HIV
 Fonte Cunico, Gomes e Vellasco Junior (2008)

À medida em que o vírus vai se replicando, o número de linfócitos T $CD4^+$ é reduzido e o sistema imunológico torna-se fragilizado. Perante o potencial catastrófico do vírus no organismo e partindo de conhecimentos acerca do mecanismo de replicação viral, fármacos antirretrovirais foram propostos para o tratamento da AIDS.

2.3 Fármacos antirretrovirais

A descoberta do mecanismo biossintético do DNA, na década de 1950, trouxe consigo a esperança de que um nucleosídeo, ou seja, um composto que apresenta uma base nitrogenada ligada a uma pentose (ribose ou desoxirribose) quimicamente modificada, pudesse bloqueá-lo. A partir desta hipótese, vários análogos de nucleosídeos foram sintetizados na tentativa de, inicialmente, ser utilizados como inibidores da replicação de células tumorais (SOUZA, 2005).

Diante disso, Horwitz et al. (1964) sintetizaram o composto denominado AZT (3-azido-2,3-didesoxitimidina), conhecido como azidotimidina, zidovudina ou retrovir, com inicial atividade antitumoral (HORWITZ et al., 1964; SOUZA, 2005).

Somente em 1985 o AZT foi reconhecido como um inibidor da transcriptase reversa do vírus HIV, após sua aprovação pelo órgão norte-americano de controle sobre produtos farmacêuticos, o *Food and Drug Administration* (FDA), comercializado pelos laboratórios Burroughs-Wellcome, a partir de 1987 (SOUZA, 2005).

Com a descoberta do AZT, inúmeras substâncias foram sintetizadas para o combate da replicação do HIV e, em 1995, um segundo inibidor da transcrição reversa, conhecido como lamivudina (3TC), também teve seu uso como antirretroviral aprovado pelo FDA (SOUZA; STORPIRTIS, 2004; STYRT; PIAZZA-HEPP; CHIKAMI, 1996), sendo liberado para uso clínico. As estruturas químicas de ambos os inibidores da TR (AZT e 3TC) encontram-se representadas na Figura 3.

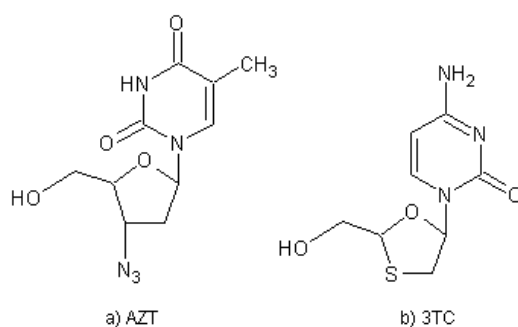


Figura 3 Fórmula estrutural da zidovudina e da lamivudina
Fonte adaptado de Souza e Storpirtis (2004)

Apesar do emprego de inibidores da transcriptase reversa no tratamento da AIDS desde 1987, avanços mais significativos com a aplicação da terapia antirretroviral só foram possíveis graças aos estudos que esclareceram a imunopatogênese desta infecção (DOURADO et al., 2006; FAUCI, 1988; HAMMER; YENI, 1998; RIO et al., 1991).

Para inibir a replicação viral, três classes de fármacos foram, então, desenvolvidas. São elas: os inibidores análogos de nucleosídeos da transcriptase reversa (INTR), os inibidores não análogos de nucleosídeos da transcriptase reversa (INNTR) e os inibidores da protease (PR) (SOUZA; STORPIRTIS, 2004).

Os INTR (Figura 4) são caracterizados por sua capacidade de se incorporar à cadeia de DNA que o vírus cria, tornando-a defeituosa e, assim, impede a reprodução do HIV no organismo infectado (AYMARD et al., 2000; BRASIL, 2009; SOUZA; STORPIRTIS, 2004).

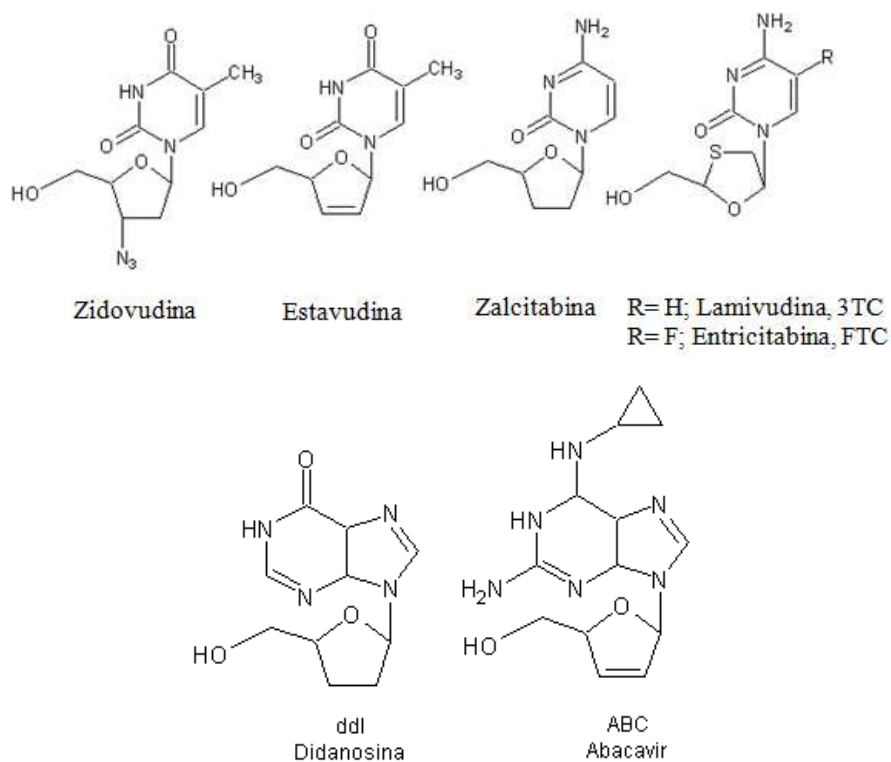


Figura 4 Estruturas químicas dos inibidores análogos de nucleosídeos da transcriptase reversa (INTR) disponíveis no mercado para uso clínico

Fonte adaptado de Souza (2005)

Os inibidores não análogos de nucleosídeos da transcriptase reversa (INNTR), representados na Figura 5, são os bloqueadores diretos da ação da enzima transcriptase reversa, ou seja, impedem que a enzima realize a transcrição do RNA viral em DNA pró-viral e dê sequência aos eventos responsáveis pela formação de novos vírus (AYMARD et al., 2000; BRASIL, 2009; SOUZA; STORPIRTIS, 2004).

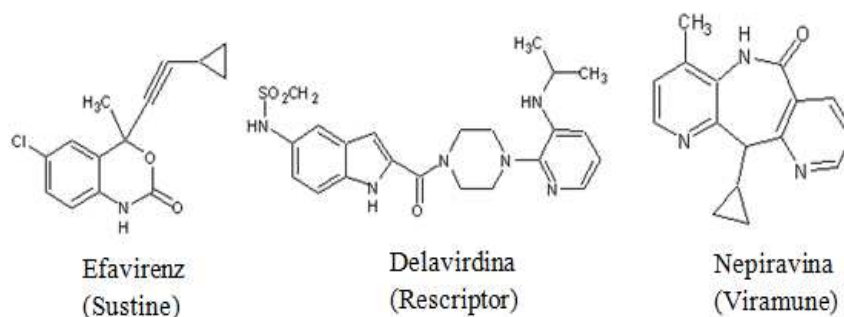


Figura 5 Representantes dos inibidores análogos não nucleosídeos da TR para o tratamento da AIDS
 Fonte adaptado de Souza (2005)

Já os inibidores da PR (Figura 6) atuam bloqueando a ação da enzima protease, impedindo a produção de proteínas virais para as novas cópias de células infectadas pelo HIV (AYMARD et al., 2000; BRASIL, 2009; SOUZA; STORPIRTIS, 2004).

Estudos desenvolvidos na década de 1990 mostraram que as interrupções na replicação do HIV, provocadas pelo uso dos antirretrovirais, atreladas à rapidez na replicação viral e meia vida plasmática (48 horas), facilitavam o desenvolvimento de resistência viral aos fármacos até então disponíveis. Com isso, o tratamento monoterápico passou a ser substituído por associações entre tais medicamentos, para melhorar a efetividade da terapia antirretroviral (DHAMI et al., 2009; GREENE et al., 2008; MIRANDA et al., 2010).

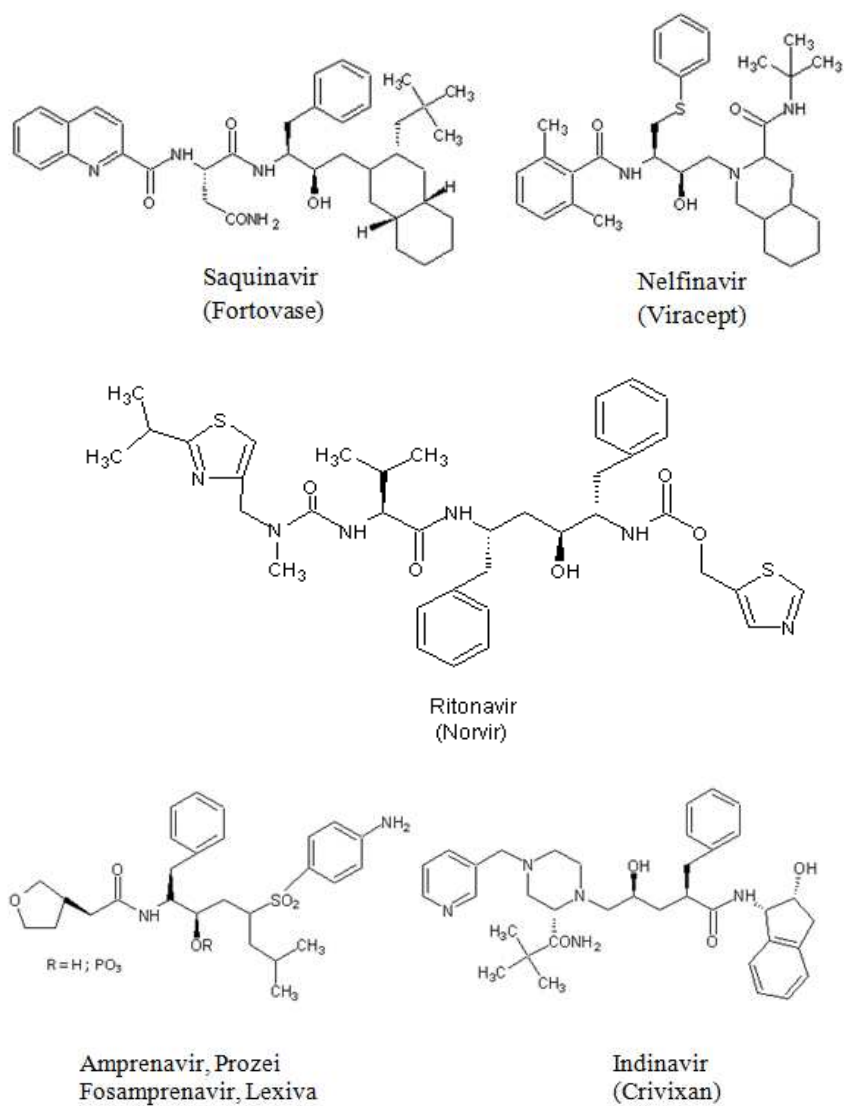
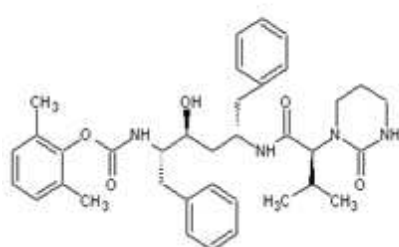
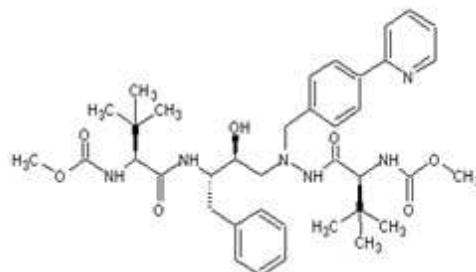


Figura 6 Estruturas químicas dos medicamentos disponíveis no mercado, capazes de inibir a protease viral
Fonte adaptado de Souza (2005)

(...Continua...)



Lopinavir

Atazanavir
(BMS-232632)

No dia 13 de novembro de 1996 foi sancionada, no Brasil, a Lei nº 9.313, que garante aos portadores do HIV/AIDS o direito de receber gratuitamente, pelo Sistema Único de Saúde (SUS), todos os medicamentos necessários ao tratamento correto da doença (SALDANHA; ANDRADE; BECK, 2009). A política para a assistência aos indivíduos infectados pelo HIV e/ou com AIDS, estabelecida nesta lei, inclui também outras modalidades assistenciais quimioproláticas e que visam à redução das internações hospitalares, tais como assistência ambulatorial especializada, hospital-dia e assistência domiciliar terapêutica (DOURADO et al., 2006).

2.3.1 Terapia antirretroviral altamente ativa

Instituída em 1996, a terapia antirretroviral altamente ativa (HAART do inglês *Highly Active Antiretroviral Therapy*) se dá pela combinação dos inibidores da protease e da transcriptase reversa, de forma a ser efetiva na redução da carga viral plasmática de HIV-1 (COLOMBRINI; LOPES; FIGUEIREDO, 2006; JOHNSTON et al., 1998; PATERSON et al., 2000). Esta combinação de fármacos com diferentes ações, estratégia popularmente

conhecida como coquetel, visa melhorar a eficácia terapêutica e retardar o surgimento de resistência viral (DHAMI et al., 2009; GREENE et al., 2008; MIRANDA et al., 2010) com o conseqüente declínio da morbidade e da mortalidade dos portadores do HIV/AIDS.

A resistência viral desenvolvida nesse tipo de tratamento pode estar relacionada a fatores distintos e variar entre as classes de antirretrovirais, podendo ocorrer, inclusive, resistência cruzada, dependendo da eficácia da posologia e da adesão do paciente ao tratamento (GARDNER et al., 2009; MIRANDA et al., 2010; ZUCKERMAN; PILLAY, 2007).

A estratégia de se combinar fármacos antirretrovirais com diferentes alvos de ação provou ser uma estratégia mais eficiente que os regimes monoterápicos (BRITO, 2011; DE CLERCQ, 2004, 2010; FATTORUSSO et al., 2005). As combinações mais comuns incluem dois ou três inibidores da TR, sendo dois INTR associados a um INNTR e um inibidor da PR (BRITO, 2011; JANSSEN et al., 2005), podendo ser manejada, para cada paciente, de acordo com o aparecimento de resistências, toxicidade e comorbidades (BRASIL, 2009; FERREIRA; RIFFEL; SANT'ANA, 2010).

Nos últimos anos, entretanto, a necessidade de se buscar novas classes de antirretrovirais ficou mais evidente e moléculas capazes de atuar em novos alvos relacionados ao agente etiológico e à fisiopatologia da AIDS, e, portanto, não susceptíveis à resistência cruzada com os fármacos até então existentes, passaram a ser amplamente estudadas (MIRANDA et al., 2010).

Além dos já conhecidos inibidores da transcrição reversa e da protease, outros dois alvos passaram a ser objeto de estudo pela comunidade científica: os inibidores de fusão ou entrada (U), que impedem a fusão e a conseqüente entrada do vírus HIV na célula hospedeira e os inibidores da integrase (IN), que bloqueiam a atividade da enzima integrase responsável pela integração entre o

DNA do HIV e o DNA da célula humana (VILAR; SANTANA; URIARTE, 2006).

Desde 2003, quatro novos fármacos foram aprovados pelo FDA. São eles a enfuvirtina, o maraviroc, o raltegravir e a etravirina (BRITO, 2011; DE CLERCQ, 2010). A enfuvirtina, cujo nome comercial é FuzeonTM (Roche), também conhecida como T-20, foi o primeiro fármaco de uma nova classe a ser aprovado desde a introdução dos inibidores de protease (BRITO, 2011; DE CLERCQ, 2004, 2010; LABONTE; LEBBOS; KIRKPATRICK, 2003). Trata-se de um polipeptídeo linear com 36 aminoácidos, conforme Figura 7, que atua prevenindo a fusão e a entrada viral na célula humana (BRITO, 2011; KILBY et al., 1998; LALEZARI, 2003).

CH₃CO-Tyr-Thr-Ser-Leu-Ile-Glu-Glu-Ser-Gln-Asn-Gln-Gln-Glu-Lys-Asn-
Glu-Gln-Glu-Leu-Leu-Glu-Leu-Asp-Lys-Trp-Ala-Ser-Leu-Trp-Asn-Trp-Phe-NH₂

Figura 7 Sequência de aminoácidos do fármaco enfuvirtina
Fonte: Brito (2011)

O maraviroc, ou SelzentryTM (Pfizer), é um inovador inibidor da entrada no HIV-1 em células humanas (BRITO, 2011; DE CLERCQ, 2010), representado estruturalmente na Figura 8, aprovado pelo FDA, em agosto de 2007, como o primeiro fármaco para o tratamento do HIV-1 com especificidade para o receptor de quimiocina CCR5. É um antagonista seletivo dos receptores de quimiocinas que previne a interação da gp120 com o receptor CCR5, necessário para a entrada do vírus na célula, atuando em combinação com outros antirretrovirais (BRITO, 2011; KURITZKES; KAR; KIRKPATRICK, 2008; TSIBRIS; KURITZKES, 2007).

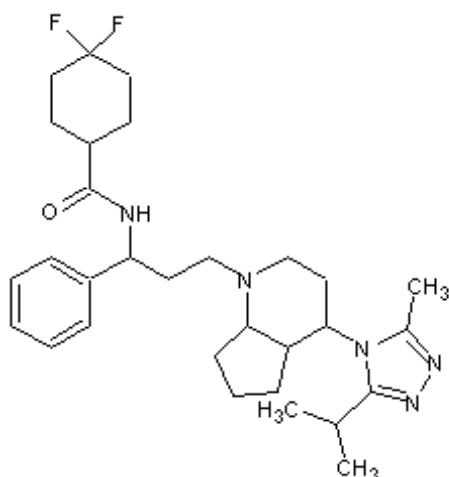


Figura 8 Estrutura química do fármaco maraviroc
Fonte adaptado de Brito (2011)

O raltegravir (Figura 9) foi o primeiro fármaco inibidor da integrase do HIV-1 aprovado pelo FDA, em outubro de 2007 (BRITO, 2011; DE CLERCQ, 2010; DEEKS, 2008). Atua inibindo a integração do DNA viral ao genoma da célula hospedeira (BRITO, 2011; BRITO et al., 2008; POMMIER; JOHNSON; MARCHAND, 2005) e corresponde a um alvo terapêutico atrativo pelo papel chave que exerce, assim como a transcriptase reversa e a protease (ALBUQUERQUE et al., 2007; BRITO, 2011).

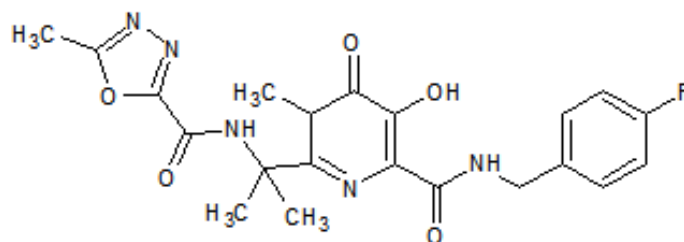


Figura 9 Estrutura química do fármaco raltegravir
Fonte adaptado de Brito (2011)

A etravirina, comercialmente conhecida como Intelence™ ou Tibotec, é um inibidor não nucleosídeo da TR aprovado pelo FDA, em janeiro de 2008, representado quimicamente na Figura 10. É indicado para o tratamento do HIV-1, em combinação com outros antirretrovirais para pacientes que já fizeram algum tratamento prévio e que tenham variantes de HIV-1 resistentes a outros fármacos dessa classe (BRITO, 2011; DE CLERCQ, 2010; HAUBRICH et al., 2008). Assim como os demais INNTR, atua ligando-se ao sítio alostérico da TR, próximo ao sítio catalítico, impedindo a replicação viral (BRITO, 2011; BRITO et al., 2000).

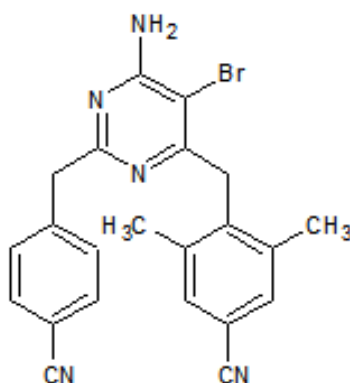


Figura 10 Estrutura química do fármaco etravirina
Fonte adaptado de Brito (2011)

Embora muitos avanços continuem ocorrendo na terapia antirretroviral, ainda não há cura para a AIDS e o coquetel antirretroviral deve ser administrado de modo contínuo e por tempo prolongado (por toda a vida do paciente). A adequação do estilo de vida deve ser abordada antes do início do tratamento e regularmente durante o acompanhamento da doença, assim como devem ser esclarecidos o risco de efeitos adversos e a importância do manejo terapêutico para reduzir tais efeitos.

Por se tratar de uma combinação de fármacos, o uso do coquetel anti-HIV torna o tratamento exaustivo e a não adesão é uma das principais causas (somada aos efeitos adversos e resistências) da falha terapêutica. Nesse sentido, medicamentos que apresentem múltiplas atividades tornam-se uma alternativa na busca por melhorias na terapia antirretroviral ao colaborar na redução da quantidade de drogas a serem administradas, atreladas à possibilidade de que apresentem melhor perfil farmacocinético e potência frente ao vírus HIV.

2.4 Atividade anti-HIV-1

A atividade anti-HIV-1 está relacionada à interrupção seletiva de uma ou mais etapas específicas e necessárias para a ocorrência do ciclo de reprodução do vírus HIV-1 no organismo infectado.

As enzimas virais responsáveis pela replicação do HIV-1 são a transcriptase reversa (TR), a integrase (IN) e a protease (PR), além das glicoproteínas gp41 e gp120 para a fusão vírus-célula. Uma vez que o HIV tenha alcançado o interior das células humanas, o mesmo pode ter sua replicação atenuada ou destruída por meio de medicamentos que atuem na inibição dessas proteínas virais essenciais. Como foi dito anteriormente, tais alvos, incluindo a gp41 e gp120, têm sido aplicados ao desenvolvimento dos fármacos antirretrovirais usados na terapia atual.

O ponto primário de contato entre o envelope viral e a célula hospedeira (linfócitos T e macrófagos) se dá entre a gp120 do envelope viral e o receptor celular CD4, o que causa mudanças conformacionais na gp120, favorecendo sua ligação ao correceptor quimiocínico CCR5 ou CXCR4 (receptores acoplados à proteína G) essenciais ao reconhecimento da célula alvo (CUNICO; GOMES; VELLASCO JUNIOR, 2008; MELO, 2011; ROCHE et al., 2013).

A interação da gp120 com os receptores celulares também provoca alterações conformacionais na gp41, que resultam na inserção e na exposição do peptídeo de fusão na membrana celular, desencadeando o processo de fusão das membranas e formação de um poro, conforme Figura 11, por onde o capsídeo viral penetra no citoplasma da célula hospedeira (MELO, 2011).

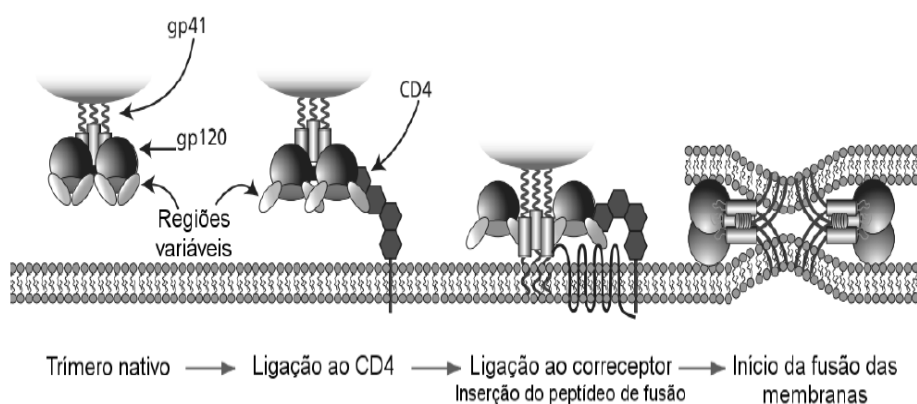


Figura 11 Adsorção e fusão do vírus HIV-1 à membrana plasmática da célula infectada

Fonte Doms (2004) e Melo (2011)

A inibição seletiva dos correceptores quimiocínicos CCR5 e CXCR4 impede a interação com a gp120, bloqueando a entrada do vírus na célula. Um exemplo desse tipo de inibição é a provocada pelo medicamento antirretroviral Maraviroc, citado anteriormente. A fusão vírus-célula também pode ser impedida por meio da inibição do rearranjo estrutural das glicoproteínas do envelope viral (gp41 e gp120), assim como ocorre pelo uso do anti-HIV-1 enfuvirtina que atua na gp41, impedindo a entrada do vírus na célula.

Uma vez ocorrida a fusão, a enzima viral transcriptase reversa (TR) recodifica o material genético do HIV, convertendo-o de RNA para DNA, pois só assim ele consegue entrar no núcleo da célula e assumir o controle dos

mecanismos de replicação da célula hospedeira (BARBOSA, 2009; SOUZA, 2005).

No início da transcrição viral ocorre a ligação do iniciador tRNA^{lys} ao sítio PBS (do inglês, *primer binding site*) na extremidade 5' do RNA viral, para que a TR inicie a síntese da fita de DNA complementar. O domínio de Rnase H da TR degrada o RNA usado como molde, permitindo que o segmento de DNA complementar recém-sintetizado associe-se à extremidade 3' do RNA viral para completar a síntese da fita de DNA (MELO, 2011).

Isolada do HIV-1 (Figura 12) a TR apresenta-se sob a forma de heterodímeros, consistindo em quantidades equimolares de dois polipeptídeos de pesos moleculares de 66kDa e 51kDa, derivados de clivagens proteolíticas da subunidade 66kDa, catalisada por uma protease codificada pelo vírus. Os sítios de atividade polimerásicas estão associados aos domínios de ambos os polipeptídeos, enquanto o sítio para a atividade ribonucleásica H está associado ao domínio terminal carboxílico do polipeptídeo de 66kDa (PEREIRA et al., 1996).

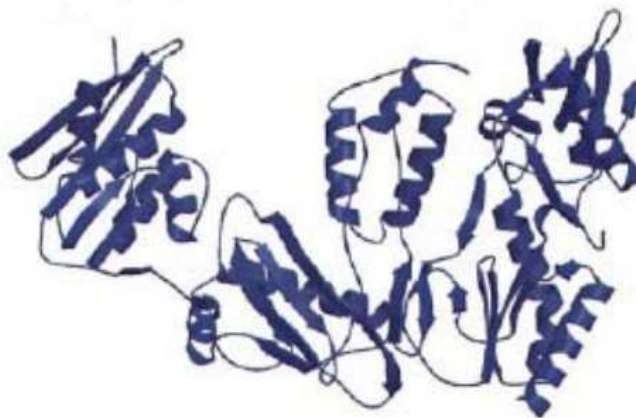


Figura 12 Enzima HIV-1 TR
Fonte adaptado de Souza (2009)

A TR representa um alvo de relevante importância para a elaboração de drogas capazes de bloquear infecções por retrovírus, podendo ser inibida de forma direta ou competitiva.

Após a transcrição reversa, dupla fita de DNA é transportada do citoplasma para o núcleo da célula, na qual, em associação com a enzima integrase (IN) ligada a aos terminos do DNA viral, se integra ao genoma da célula hospedeira (PEREIRA et al., 1996). O processo de integração ocorre com mais frequência em regiões transcricionalmente ativas e é modulada pela interação da proteína LEDGF/p75 com a cromatina celular (MELO, 2011).

A enzima IN do HIV-1 apresenta 32kDa e 288 aminoácidos que formam uma cadeia polipeptídica simples que se dobra em três domínios funcionais (MELO, 2006), conforme Figura 13. O domínio N-terminal (resíduos de aminoácidos 1-50), no qual se encontram dois resíduos de histidina e dois de cisteína que promovem a ligação com o Zn^{2+} ; o domínio central catalítico (resíduos 50-212), com aminoácidos ácidos altamente conservados Asp64 (denominado D64), Asp116 (D116) e Glu52 (E152), comumente designados “DDE”, onde se ligam dois íons de Mg^{2+} e, por fim, os resíduos 213 a 288 formam o domínio de ligação C-terminal do DNA, que contém resíduos de aminoácidos básicos e aparenta ter a função de ligar e orientar o DNA proviral ao genoma do hospedeiro (CARVALHO, 2011; MELO, 2006).

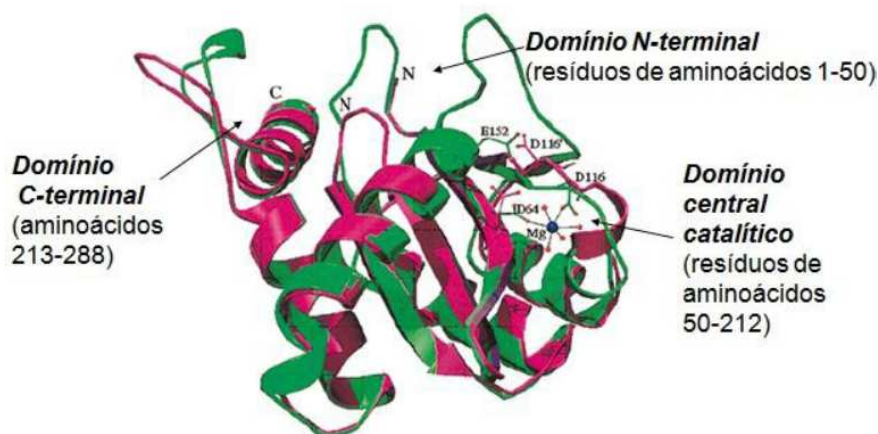


Figura 13 Enzima HIV-1 IN e seus domínios
Fonte Carvalho (2011)

O domínio central DDE (D64, D116 e E152) é crucial para a atividade da enzima integrase que atua por meio de duas etapas no ciclo de replicação viral que podem ser inibidas com a finalidade de bloquear a multiplicação viral.

Em primeiro lugar, a IN catalisa a clivagem endonucleotídica do 3'-terminal do DNA viral clivado (c-DNA) e, posteriormente, permanece ligada no complexo pré-integração catalisando a inserção do c-DNA viral no cromossomo hospedeiro, processo denominado Etapa de Transferência de Fita, no qual o átomo de oxigênio 3'-terminal do DNA viral clivado ataca a ligação fosfodiéster do DNA hospedeiro através dos dois íons de Mg^{2+} (CARVALHO, 2011).

Após a integração, o ciclo replicativo do HIV-1 segue através da transcrição do DNA em RNA viral (RNA mensageiro, RNAm), com o auxílio da proteína Tat. O conjunto de RNAs transcritos é transportado para o citoplasma da célula, no qual são produzidas poliproteínas virais, por meio da traduções das informações fornecidas pelo RNAm (SOUZA, 2009).

Após a montagem e o brotamento, a partícula viral formada sofrerá maturação por ação da enzima HIV-1 protease (PR), que irá clivar as poliproteínas em proteínas menores e funcionais, formando uma nova partícula

infecciosa que inclui as enzimas e as proteínas necessárias para que o processo replicativo se repita (MELO, 2011; SOUZA, 2009).

A HIV-1 PR é um homodímero que contém um único sítio catalítico. A região entre os resíduos de aminoácidos 43-85 denominada alça (*flip*) contém as fitas- β , como mostrado na Figura 14, que se fecha sobre o sítio ativo da PR complexada com inibidores ou substratos. Os monômeros não estão ligados covalentemente e cada cadeia é formada por 99 resíduos de aminoácidos, contendo uma α -hélice (aminoácidos 87-93) e nove folhas- β (SOUZA, 2009).

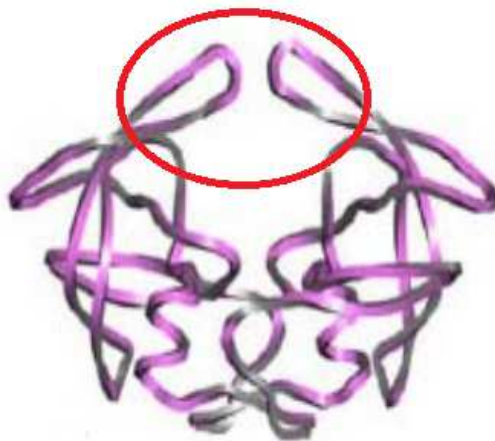


Figura 14 HIV-1 PR, com destaque (círculo vermelho) para as alças que se fecham sobre o sítio ativo da enzima, quando complexada
Fonte adaptado de Souza (2009)

O processamento incompleto das poliproteínas resulta na formação de partículas virais imaturas não infecciosas, o que mantém a PR como um alvo atrativo no desenvolvimento de antirretrovirais, facilitado pela determinação da estrutura tridimensional da proteína isolada e pelo conhecimento dos sítios de clivagem e do mecanismo de catálise dessa classe de enzimas.

2.5 Inibidores da transcriptase reversa (anti-TR-1)

Mais de duas décadas após a descoberta do vírus HIV, ainda nos deparamos com inúmeros desafios para a compreensão dos vários aspectos que essa epidemia apresenta (POLEJACK; SEIDL, 2010). A prevalência da infecção e as taxas de incidência variam consideravelmente e países subdesenvolvidos, como a África Subsaariana, são os mais afetados (NASH; EIUL, 2006; POLEJACK; SEIDL, 2010).

Inibidores da transcrição reversa do HIV foram os primeiros antirretrovirais estudados e lançados no mercado para o tratamento clínico da AIDS. A transcriptase reversa viral catalisa a converção do RNA viral em DNA pró-viral, um estágio considerado chave na replicação do vírus HIV (RAVICHANDRAN; AGRAWAL, 2007). Seu papel central a mantém como o principal alvo (WEITMAN et al., 2011), tanto em termos da terapia antirretroviral a ser adotada para cada paciente de acordo com a evolução da doença, como para o desenvolvimento de novos e eficientes fármacos para o combate da AIDS.

A transcrição do RNA genômico viral é considerada RNA-dependente, ocorrendo a síntese da fita complementar de DNA a partir da fita molde de RNA. Ao mesmo tempo, esta fita dupla de DNA e RNA é hidrolisada pela ribonuclease H (RNase H) da TR, seguida pela atividade da DNA polimerase que sintetiza a segunda cadeia de DNA, partindo da fita de DNA já sintetizada como modelo. Essas três atividades de TR interligadas geram o DNA de fita dupla que, subsequentemente, é conectado ao DNA genômico celular por uma segunda enzima retroviral, a integrase do HIV (WEITMAN et al., 2011).

O bloqueio de cada uma das atividades da TR tem provado ser útil na proteção de células-alvo, ao impedir a entrada no núcleo celular e a continuação dos mecanismos de replicação da célula hospedeira (SOUZA, 2005).

Como foi dito anteriormente, os inibidores da TR incluem os análogos nucleosídeos (INTR) e os não nucleosídeos (INNTR), contudo, atualmente, estão disponíveis no mercado, para uso clínico, os análogos de nucleotídeos tenofovir e adefovir, cujas estruturas químicas estão disponíveis na Figura 15.

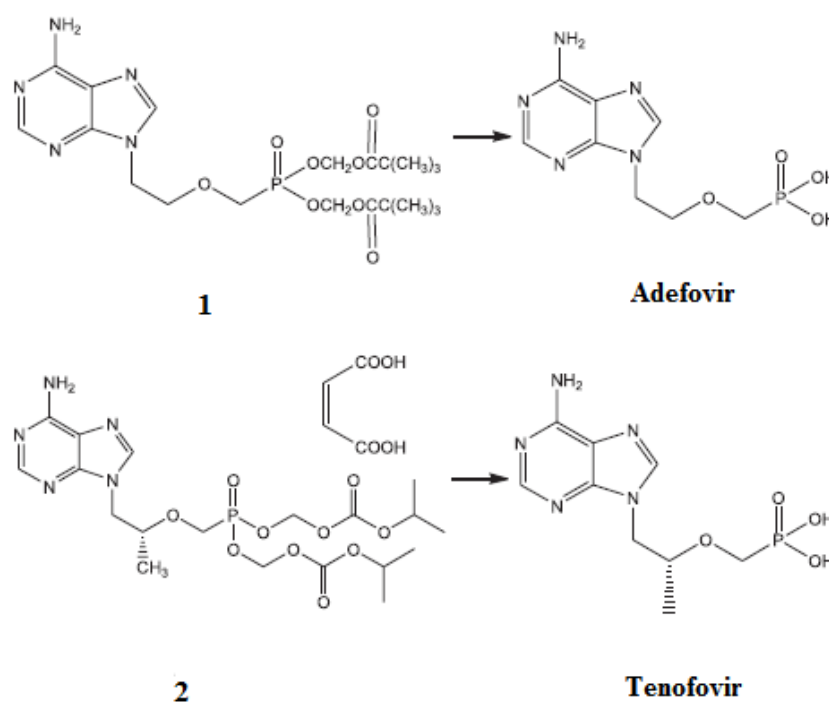


Figura 15 Estruturas químicas dos análogos de nucleotídeos da TR adefovir e tenofovir. Os compostos 1 e 2 são seus pró-fármacos desenvolvidos para contornar o problema da baixa solubilidade em água e biodisponibilidade, quando administrados via oral

Fonte adaptado de Cunico, Gomes e Vellasco Junior (2008)

Os inibidores de nucleosídeos diferem dos inibidores de nucleotídeos pela necessidade de serem transformados em seus metabólitos, ou seja, eles atuam como pró-fármacos na inibição da transcriptase reversa do HIV e, conseqüentemente, impedem a formação do DNA viral (CUNICO; GOMES;

VELLASCO JUNIOR, 2008). Esta transformação ocorre por ação de enzimas quinases do hospedeiro que trifosforilam intracelularmente estes análogos (SOUZA; STORPIRTIS, 2004). Em contrapartida, os inibidores nucleotídeos já apresentam um radical fosfato em sua estrutura (Figura 15), estando prontos ou ativos para a inibição direta da atividade da transcriptase reversa após duas fosforilações (CUNICO; GOMES; VELLASCO JUNIOR, 2008).

A enzima AMPquinase é a responsável pelas duas fosforilações do adenovir e do tenofovir, para que eles possam impedir diretamente a transcrição do RNA em DNA (CUNICO; GOMES; VELLASCO JUNIOR, 2008).

Os INNTR são inibidores não competitivos que se ligam a um sítio alostérico da enzima e, como resultado desta interação, o sítio ativo responsável pela formação da dupla hélice do DNA tem sua mobilidade e flexibilidade restritas, reduzindo a eficiência enzimática (CUNICO; GOMES; VELLASCO JUNIOR, 2008; SHEN et al., 2003). Fármacos desta classe (nevirapina, delavirdina e efavirenz) apresentam grande potencial, sendo amplamente utilizados como componentes da HAART, em especial a nevirapina, que é um dos poucos fármacos utilizados na prevenção da transmissão do HIV de mãe para filho (transmissão vertical) (CUNICO; GOMES; VELLASCO JUNIOR, 2008).

Embora mais de 30 compostos estruturalmente diferentes já tenham sido identificados como inibidores da TR do HIV, grande parte dessas estruturas tem suas fases clínicas suspensas por motivos como toxicidade, propriedade farmacocinética ruim, baixa eficácia *in vivo* e potência insuficiente contra vírus mutantes (CUNICO; GOMES; VELLASCO JUNIOR, 2008; DE CLERCQ, 2004).

Contudo, a utilização dos inibidores da TR em terapia de combinação continua apresentando resultados animadores (JI et al., 2007), pois favorecem a diminuição da resistência aos fármacos, além de serem indicados para prevenir a

transmissão vertical da doença (SOUZA; STORPIRTIS, 2004). Tais características mantêm este alvo de ação como objeto de estudos na busca por inibidores mais potentes, com melhores perfis farmacocinéticos, menores efeitos colaterais e melhor espectro de atividade a diferentes cepas resistentes.

2.6 A química medicinal e o planejamento de fármacos

As condições adequadas para o surgimento do processo de descoberta racional de um novo fármaco foram criadas a partir do início do século XX, atribuição da Química Medicinal, fruto da multiplicidade de fatores que envolvem o planejamento molecular de novas estruturas capazes de apresentar os efeitos farmacológicos desejados, com biodisponibilidade adequada ao seu uso terapêutico (BARREIRO, 2002; MONTANARI, 2011).

Como uma disciplina multidisciplinar envolvendo a tríade Química Orgânica–Farmacologia–Química Computacional, a Química Medicinal se dedica ao estudo das razões moleculares da ação dos fármacos, a relação entre estrutura química e atividade biológica/farmacológica, além do planejamento racional de novos agentes terapêuticos (BARREIRO, 2002; MONTANARI, 2011; WERMUTH, 1996).

Estudos estimam que o custo total para o desenvolvimento de um novo fármaco pode ultrapassar um bilhão de dólares, além de envolver um tempo médio de doze a quinze anos, até que todas as etapas, desde o seu planejamento até o lançamento no mercado para uso clínico, sejam cumpridas (MONTANARI, 2011). Este dispendioso sistema é dividido em duas grandes fases: (i) descoberta, fase pré-clínica ou pesquisa básica e (ii) desenvolvimento ou fase clínica, como ilustrado na Figura 16 (BARREIRO, 2002; GUIDO; ANDRICOPULO; OLIVA, 2010; LOMBARDINO; LOWE, 2004).

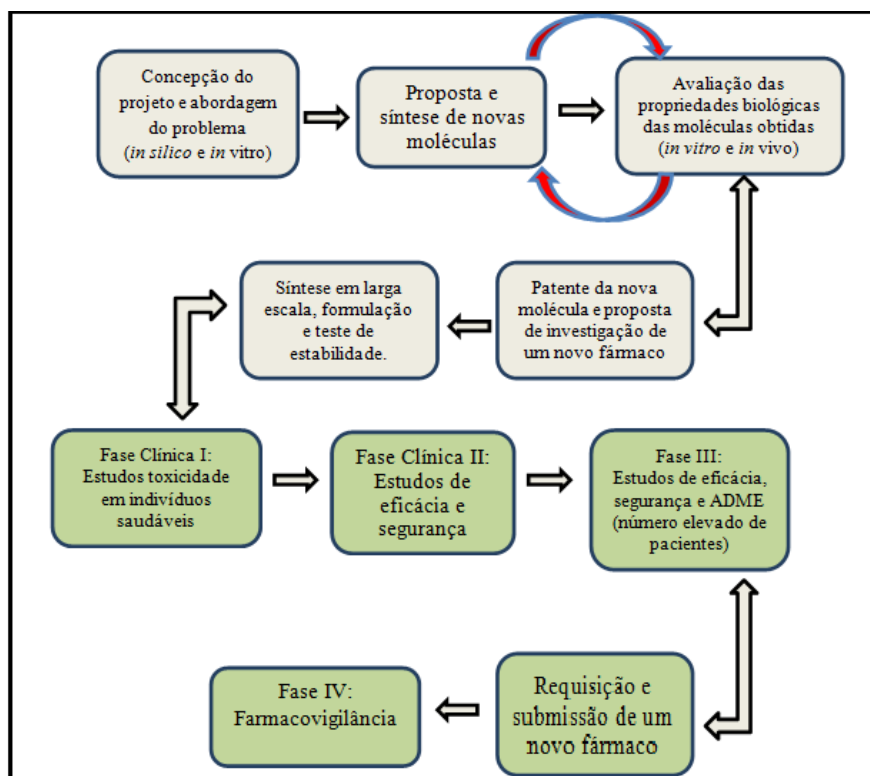


Figura 16 Etapas do processo de descoberta e desenvolvimento de fármacos
Fonte adaptado de Guido, Andricopulo e Oliva (2010)

Com a finalidade de reduzir o tempo e os altos custos envolvidos no processo, modelos *in silico* ou assistidos por computadores estão sendo aplicados na seleção de um número menor de moléculas a serem avaliadas em ensaios *in vitro* e *in vivo*, maximizando as chances de identificação de novas entidades químicas (NCEs) qualificadas (MONTANARI, 2011).

Por vezes, embora o novo medicamento chegue ao seu destino final (comercialização), o sucesso farmacoterapêutico não acontece devido a problemas posteriores quanto à sua farmacocinética, à perda de sua eficácia terapêutica, ao desenvolvimento de efeitos adversos consideráveis e, até mesmo, por razões comerciais (SILVA, 2013; WATERBEEMD; GIFFORD, 2003).

Como a ação dos fármacos é resultante de interações destes com sistemas biológicos, há dependência entre fatores relacionados à sua estrutura química e suas propriedades físico-químicas que irão influenciar a interação do medicamento com a biofase e sua consequente distribuição nos diversos compartimentos que compõem o sistema biológico (TAVARES, 2004).

Nos últimos 40 anos, foram desenvolvidas várias técnicas para auxiliar e simplificar a busca por novos alvos terapêuticos, o que fez do processo, antes intuitivo, agora racional (SILVA, 2013). A modelagem molecular representa um dos mais importantes avanços no planejamento e na descoberta de um fármaco, sendo também uma importante ferramenta na otimização de um protótipo já existente ou, até mesmo, obtido pelo próprio estudo de modelagem molecular.

Seu desenvolvimento se deu, em grande parte, devido aos avanços computacionais que permitiram a obtenção de propriedades específicas de determinada molécula que poderão influenciar a interação com o receptor. Por meio de programas de computador é possível desenhar a estrutura da molécula e realizar cálculos de otimização geométrica, bem como estudos de análise conformacional, facilitando o planejamento racional de novas moléculas que satisfaçam às propriedades eletrônicas e estruturais para um melhor encaixe no sítio receptor (RODRIGUES, 2001) e, conseqüentemente, um possível melhor efeito farmacológico.

2.7 QSAR como ferramenta para a elaboração de novos candidatos a fármacos

Diante dos avanços científico-tecnológicos observados em diversas áreas, como a biologia estrutural, molecular e, especialmente, a química computacional, a busca racional por novos compostos tornou-se uma realidade,

sendo esta última uma das áreas mais promissoras deste novo século (BARREIRO, 2002; RAUPP; SERRANO; MARTINS, 2008).

Inseridos neste contexto encontram-se os estudos QSAR (do inglês *Quantitative Structure-Activity Relationships*), que tratam da capacidade de se correlacionar quantitativamente a atividade e/ou propriedade química ou física com a estrutura molecular de um determinado composto, com a finalidade de colaborar para o planejamento de novas substâncias com perfil terapêutico específico e que atenda às necessidades (NEVES et al., 1998; TAVARES, 2004).

A aplicação desta metodologia, além de tentar contornar o problema gerado pelos altos custos da produção e desenvolvimento de um novo fármaco por meio de métodos convencionais, uma vez que dispensa as sínteses experimentais exploratórias, também apresenta o potencial de reduzir o tempo e os esforços necessários para a criação ou a melhoria dos fármacos atuais em termos de eficácia (AFANTITIS et al., 2006).

As técnicas utilizadas em QSAR tiveram início em 1863, conforme Tabela 1, quando Cross observou que a toxicidade de álcoois em mamíferos aumentava quando sua solubilidade em água diminuía (MARTINS, 2010). Contudo, a sistematização das análises neste tipo de metodologia deve ser associada ao trabalho de Hansch e Fujita, de 1964, no qual foram utilizados como descritores o coeficiente de partição octanol/água (LogP), a constante de Hammett σ atuando como um descritor eletrônico e o parâmetro de lipofilicidade π (equação 1), definido em analogia ao descritor eletrônico (MARTINS, 2010).

Assim, o modelo de Hansch e Fujita (1964) foi definido pela relação dos efeitos dos substituintes na mesma estrutura básica com a atividade biológica de ácidos benzóicos, conforme equação 2.

$$\sigma = \text{Log } K_x - \text{Log } K_0 \quad (1)$$

$$\pi = \text{Log } P_x - \text{Log } P_H \quad (2)$$

em que P_H representa o coeficiente de partição do composto de origem não substituído (com H) e P_x é o termo correspondente para o derivado substituído por x.

Tabela 1 Histórico da metodologia QSAR, em ordem cronológica

Cross - 1863	Observou que a toxicidade de alcoóis em mamíferos aumentava quando suas solubilidades em água diminuíam.
Crum Brown e Fraser - 1868	Postularam que existia uma relação entre as atividades biológicas e as estruturas químicas.
Richet - 1893	Propôs que a toxicidade de alguns alcoóis e éteres eram inversamente proporcionais a suas solubilidades em água.
Meyer e Overton - 1900	Descreveram um parâmetro que pode ser considerado como um precursor do atual LogP, o coeficiente de partição octanol/água.
Hammett 1937	Dissociação de ácidos benzoicos meta ou para-substituídos em H ₂ O a 25°C, sendo denominada constante de Hammett (σ).
Ferguson 1939	Estudou o comportamento de diversas propriedades em relação à atividade tóxica de diferentes séries homólogas de compostos.
Free e Wilson 1964	Postularam que para uma série de compostos similares, a contribuição do substituinte é aditiva e dependente de seu tipo e sua posição.
Hansch e Fujita 1964	Sistematização das análises em QSAR.

Fonte: Hammett (1937), Hansch et al. (2002), Martins (2010) e Silva (2013)

Assim, já no final da década de 1960, os modelos QSAR começaram a ser utilizados como uma ferramenta prática na identificação das propriedades relacionadas com a atividade biológica exercida (MONTANARI, 2011).

Em 1988, as técnicas QSAR sofreram uma grande transformação a partir da inserção de parâmetros moleculares tridimensionais, ficando conhecido como QSAR 3D. A partir daí, foi considerada a influência de diferentes confôrmeros, estereoisômeros ou enantiômeros e esse tipo de modelo implica no alinhamento das estruturas moleculares de acordo com um farmacóforo comum, derivado do conhecimento da interação fármaco-receptor (MARTINS, 2010). Introduzida por Cramer, a análise comparativa do campo molecular CoMFA, do inglês *comparative molecular field analysis*, é baseada na representação das moléculas em sua forma tridimensional, alinhadas no interior de uma caixa 3D virtual ou grade (grid) (CRAMER; PATERSON; BUNCE, 1988; MARTINS, 2010).

Associadas aos conceitos de similaridade entre diferentes aspectos moleculares, outras abordagens QSAR foram propostas representando a evolução desse tipo de metodologia. Assim, Hopfinger et al. (1997) propuseram o chamado QSAR 4D, que incorpora liberdade conformacional ao desenvolvimento de modelos do tipo 3D, fazendo com que a mudança do estado molecular constitua a quarta dimensão. A geometria do receptor também foi incorporada ao método QSAR 4D, para tornar a representação do sistema mais fiel à realidade. Esse método foi denominado QSAR 4D dependente do receptor (RD-4D-QSAR) (PAN; TSENG; HOPFINGER, 2003).

A representação do receptor, com a simulação do encaixe induzido e a adaptação topológica do modelo da superfície do sítio de ligação para moléculas individuais do ligante, é característica do QSAR 5-D (VEDANI; DOBLER, 2002), enquanto a possibilidade da simulação *in silico* dos efeitos adversos provocados por compostos químicos tornou-se realidade por meio da avaliação simultânea de diferentes modelos de solvatação, dando origem ao QSAR 6D (VEDANI; DOBLER; LILL, 2005).

Diante da necessidade da inclusão de modelos baseados na dimensão do receptor ou na interação virtual entre receptor-ligante, obtidos por modelagem

por homologia, foi desenvolvido o QSAR 7D. Tal classificação leva em consideração a dimensionalidade ou a complexidade de estruturas proteicas (POLANSKI, 2009). Apesar de não denominá-lo dessa forma, Martin (2009) sugere que a sétima dimensão em QSAR deveria considerar efeitos de tautomerização.

Diversas outras abordagens vêm sendo utilizadas em estudos de QSAR, entre as quais se destacam CoMSIA (KLEBE; ABRAHAM; MEITZNER, 1994), HQSAR (LOWIS, 1997), MIA-QSAR (FREITAS; BROWN; MARTINS, 2005) e LQTA-QSAR (MARTINS et al., 2009), cada uma com suas limitações, vantagens e desvantagens, compartilhando o objetivo comum da obtenção de modelos preditivos e aplicáveis o mais próximo possível da realidade.

2.7.1 Critérios e regras em metodologias QSAR

A construção de modelos matemáticos que correlacionem a estrutura química e a atividade de uma série de compostos análogos (FERREIRA; MONTANARI; GAUDIO, 2002) envolve uma série de procedimentos e regras que acompanharam a evolução que esse tipo de metodologia vem sofrendo ao longo dos anos.

Inicialmente, é preciso realizar uma busca na literatura pela definição do problema que se deseja investigar, obtendo-se conhecimento sobre a doença de interesse e dos mecanismos de ação dos compostos químicos disponíveis para combatê-la (MONTANARI, 2011).

A partir daí é selecionada uma série congênere de, no mínimo, 20 a 30 compostos com similaridade estrutural e que apresentem atividades e/ou propriedades químicas experimentalmente conhecidas e medidas conforme o mesmo protocolo de ensaio, expressas como a dose efetiva (ED) ou a

concentração molar inibitória (IC) necessária para produzir determinado efeito biológico (MARTINS, 2010; MONTANARI, 2011).

Essa busca na literatura não garante que todas as moléculas selecionadas deverão, obrigatoriamente, participar da construção do modelo QSAR. Moléculas identificadas como *outliers*, ou seja, aquelas que apresentam um comportamento não análogo ao das demais, devem ser criteriosamente analisadas, para que a molécula seja mantida ou não no modelo. Trata-se de um processo laborioso, pois, muitas vezes, a retirada de um composto do modelo pode prejudicar sua capacidade preditiva e sua não remoção pode influenciar a estabilidade do mesmo (MARTINS, 2010; SILVA, 2013).

Com isso, os descritores podem ser obtidos e, após seleção prévia, são selecionados aqueles que se mostrarem úteis para o estudo, ou seja, aqueles que, devido a mudanças na estrutura molecular, causam uma variação semelhante na resposta biológica que produzem (MONTANARI, 2011).

O próximo passo é constituído pela construção e a validação dos modelos construídos quanto à capacidade preditiva, à robustez e à significância (MONTANARI, 2011). A confiabilidade do modelo não pode ser garantida apenas por valores ótimos de r^2 ($>0,8$) ou q^2 ($>0,5$), mas por um conjunto de testes estatísticos, entre os quais se destaca a validação externa (MARTINS, 2010).

Segundo Golbraikh e Tropsha (2002), a validação externa consiste na escolha de um conjunto de amostras ou *test set*, que não fará parte da construção do modelo, mas que terá suas atividades biológicas calculadas pelo modelo construído. Este tipo de validação é a única maneira de se alcançar um modelo QSAR confiável, sendo, atualmente, uma validação obrigatória em metodologias QSAR (GOLBRAIKH; TROPSHA, 2002; MARTINS, 2010).

Como as estruturas químicas não contêm a informação relacionada à atividade biológica de forma explícita, as informações referentes às atividades

biológicas devem ser extraídas das estruturas na forma de descritores moleculares, o que exige cuidados mais complexos (MARTINS, 2010). Tais descritores acentuam diferentes propriedades químicas implícitas na molécula, que vão desde parâmetros físico-químicos, eletrônicos ou mecânico-quânticos, até características geométricas ou topológicas das moléculas, podendo, então, ser correlacionadas diretamente com a atividade (MARTINS, 2010; WEBER, 2008).

Além disso, deve-se atentar para os erros estruturais ou experimentais presentes no conjunto de amostras, pois, enquanto os primeiros podem ser facilmente identificados e contornados, equívocos de caráter experimentais não podem ser sanados durante a modelagem, uma vez que ainda não se dispõe de métodos capazes de reconhecê-los. A remoção de compostos que se encontram fora do limite de correlação do método empregado, como, por exemplo, compostos inorgânicos, organometálicos, sais e misturas, ajuda a minimizar os erros envolvidos no processo de modelagem molecular (SILVA, 2013).

2.7.2 Análise multivariada de imagens aplicada ao QSAR

Uma variedade de descritores multidimensionais tem sido extensivamente aplicada na geração de modelos (PINHEIRO et al., 2008) que requerem a triagem conformacional e o alinhamento 3D dos ligantes (GOODARZI; FREITAS, 2009), exigindo um maior custo computacional. Contudo, descritores moleculares 2D, usualmente descritores físico-químicos referidos em análises QSAR clássicas, não têm se mostrado inferiores aos descritores 3D, sendo extremamente potentes quanto à conveniência e à simplicidade dos cálculos (CORMANICH; NUNES; FREITAS, 2012; TIAN; ZHOU; LI, 2007).

A análise multivariada de imagens aplicada em QSAR (MIA-QSAR, do inglês *Multivariate Image Analysis applied to QSAR*) (FREITAS; BROWN; MARTINS, 2005) tem se mostrado igualmente preditiva, porém, mais rápida, barata e simples de operar, uma vez que utiliza imagens 2D para a geração de descritores (FREITAS, 2007), em comparação aos demais métodos disponíveis.

Em MIA-QSAR, as imagens desenhadas por meio de programas para desenho de estruturas químicas em duas dimensões, de livre acesso, têm seus *pixels* tratados numericamente como binários, que irão compor a matriz de descritores essencial para a construção do modelo, como mostrado na Figura 17. Baseado no sistema de cores RGB, ou *red-green-blue*, o pixel preto corresponde ao dígito 0 e pixels brancos serão representados pelo número 765, obtido pela soma do valor correspondente a cada cor do sistema RGB, ou seja, 255 (FREITAS; BROWN; MARTINS, 2005).

As modificações estruturais, ou mudança na posição dos substituintes em uma série congênere de moléculas, correspondem a alterações nas coordenadas dos *pixels* da imagem, e essas alterações explicam a variância no bloco **Y**, o bloco correspondente às variáveis dependentes, tais como as atividades biológicas (CORMANICH; NUNES; FREITAS, 2012).

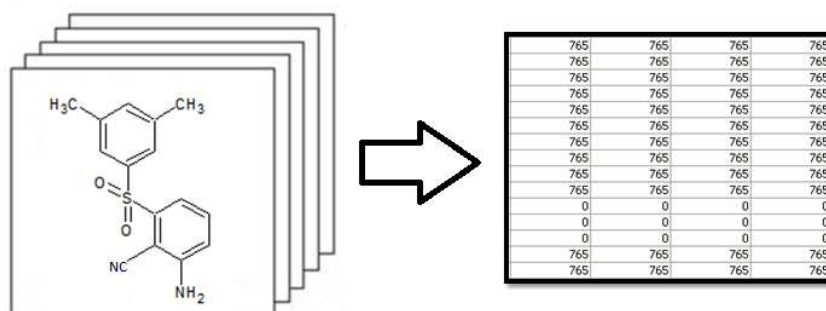


Figura 17 Conversão das estruturas químicas em uma matriz de dados binários, em que os *pixels* pretos da imagem da estrutura correspondem ao dígito 0 e *pixels* brancos (espaços em branco) correspondem ao dígito 765, de acordo com o sistema de cores RGB

É possível perceber a importância de se desenhar corretamente as estruturas para a construção do modelo MIA-QSAR, uma vez que a série de descritores altera-se conforme a representação 2D da estrutura química, obtendo-se a correlação (SILVA, 2013). Por isso, é importante manter a mesma disposição dos átomos, dos grupos substituintes e a direção da cadeia carbônica para todas as estruturas da série, embora seja válido ressaltar que diferenças, como a representação do grupo substituinte metila, seja como -Me ou como -CH₃, não irá afetar os resultados, desde que todos os compostos sejam desenhados da mesma forma, respeitando esse critério (GOODARZI; FREITAS; JENSEN, 2009; SILVA, 2013).

Para comprovar que descritores MIA podem codificar propriedades químicas e que a correlação não é obra do acaso, recentemente foi divulgado o trabalho desenvolvido por Cormanich, Nunes e Freitas (2012), no qual foi construído um modelo MIA-QSAR, a partir de uma série congênere de 70 derivados benzimidazólicos com atividade tricomonocida experimentalmente

conhecida. Este modelo foi comparado a um segundo modelo MIA-QSAR, cujas imagens correspondiam às letras do alfabeto e foram correlacionadas aos respectivos números do ordenamento do alfabeto, ou seja, a letra A corresponde ao número 1, à letra B ao número 2 e assim sucessivamente. Como era de se esperar, o primeiro modelo apresentou boa correlação, enquanto o segundo modelo não resultou em correlação alguma, demonstrando a utilidade deste método na predição da atividade biológica de compostos congêneres de uma determinada classe de substâncias bioativas (CORMANICH; NUNES; FREITAS, 2012).

Ao agrupar cálculos matemáticos com propriedades físico-químicas, a gama de possibilidades de estudos via MIA-QSAR é ampliada (SILVA, 2013), destacando-se a relevância deste método como um auxiliar na descoberta e no desenvolvimento de novos fármacos. Nesse sentido, o método proporciona uma análise muito adequada, uma vez que a interpretação visual desempenha um importante papel em todas as formas de análises de dados, sendo facilmente acessível para ambos, recém-chegados e usuários avançados (FREITAS; BROWN; MARTINS, 2005).

2.8 Docking e análises de parâmetros farmacocinéticos

O *docking*, ou ancoramento molecular, é uma das mais importantes técnicas de investigação de compostos capazes de interagir com um dado sítio macromolecular que inclui enzimas, receptores proteicos, ácidos nucleicos ou canais iônicos. É um procedimento essencialmente computacional que gera um grande número de estruturas de possíveis ligantes, analisa a viabilidade do modo de ligação e classifica os candidatos segundo alguns critérios (MONTANARI, 2011).

Atrelado às metodologias QSAR, o *docking* pode colaborar na validação e na interpretação de dados bioquímicos, ao avaliar a potencialidade ligante de cada estrutura por meio de cálculos da energia potencial de interação ou encaixe rígido, da minimização da energia potencial do sistema (encaixe parcialmente flexível) ou por meio da simulação da dinâmica molecular (encaixe flexível), que pode incluir o cálculo de propriedades termodinâmicas como parâmetros de avaliação (MONTANARI, 2011).

A aplicação de uma função de pontuação pode facilitar a escolha do ligante pela classificação dos candidatos, de acordo com o grau de interação com o sítio de ligação, baseada na soma de contribuições específicas, como interações hidrofóbicas, de van der Waals, ligações de hidrogênio e interações eletrostáticas (MONTANARI, 2011).

Pelo uso do programa Molegro Virtual Docker® (MVD) (THOMSEN; CHRISTENSEN, 2006), por exemplo, é possível calcular a interação ligante-alvo por meio do uso de suas estruturas tridimensionais que podem ser obtidas por modelos de comparação ou homologia e por experimentos de cristalografia de raios X ou de espectroscopia de ressonância magnética nuclear (RMN) (THOMSEN; CHRISTENSEN, 2006).

A análise energética permite prever se processos moleculares podem ou não ocorrer sob determinadas condições, em que energia equivale à soma da energia potencial calculada a partir de forças intra e intermoleculares atrativas ou repulsivas e a energia cinética decorrente da movimentação molecular. Como as técnicas de química computacional têm, por definição, que a menor energia do sistema corresponde à estabilidade, encontrar a conformação mais estável de uma molécula é encontrar sua forma de menor energia (CAFFARENA, 2008).

Colaborando para o conhecimento acerca do comportamento de candidatos a fármaco, estão disponíveis ferramentas computacionais que

auxiliam na compreensão da biodisponibilidade oral dos compostos propostos e na seleção de candidatos considerados promissores.

A fase farmacocinética determina o tempo e o curso do princípio ativo pelo corpo humano e refere-se à absorção (A), à distribuição (D), ao metabolismo (M) e à excreção (E) que garantem os níveis de concentração do fármaco na faixa terapêutica desejada por um período prolongado (MONTANARI, 2011; WILLIAMS; LEMKE, 2002).

Para que uma molécula seja considerada uma boa candidata a fármaco, além dos parâmetros ADME, alguns requisitos devem ser atendidos. Os critérios avaliados incluem a “Regra dos Cinco” de Lipinski, a qual considera os valores de $\log P \leq 5$, peso molecular ≤ 500 , número de aceptores de ligação de hidrogênio (nON) ≤ 10 e número de doadores de ligação de hidrogênio (nOHNH) ≤ 5 (LIPINSKI et al., 1997), relevantes para o estudo de um protótipo.

Propriedades moleculares, como solubilidade, lipofilia, ligação de hidrogênio e área de superfície polar (TPSA), afetam diretamente o perfil farmacocinético de fármacos e têm contribuído para a otimização de propriedades de ADME (MONTANARI, 2011).

3 OBJETIVO GERAL

O presente trabalho foi desenvolvido com o objetivo geral de elaborar e propor novas estruturas químicas com múltipla atividade frente ao vírus HIV-1, ou seja, compostos caracterizados por apresentar atividade anti-HIV-1 em células da linhagem MT4 e capazes de inibir a enzima transcriptase reversa viral (anti-TR-1), simultaneamente.

3.1 Objetivos específicos

Os objetivos específicos foram:

- a) modelar as atividades biológicas de duas séries de antirretrovirais com atividades diferentes frente ao vírus HIV-1, utilizando o método MIA-QSAR e, pela combinação das subestruturas mais ativas de cada série, propor estruturas com potencial múltipla atividade, empregando-se os modelos MIA-QSAR construídos;
- b) realizar estudos de *docking* utilizando o programa Molegro Virtual Docker®, para avaliação bioquímica das interações enzima-substrato das estruturas propostas;
- c) calcular parâmetros ADME das moléculas propostas para avaliação farmacocinética preliminar dos possíveis candidatos a antirretrovirais.

4 METODOLOGIA

Este trabalho foi desenvolvido no Laboratório de Modelagem Molecular do Departamento de Química da Universidade Federal de Lavras (UFLA), por meio do uso de uma estação de trabalho Samsung WriteMaster™, e dividido em três etapas:

- a) determinação do conjunto de dados;
- b) construção dos modelos MIA-QSAR (1 e 2) e proposta de novos inibidores com suposta múltipla atividade;
- c) *docking* e avaliação de parâmetros farmacocinéticos.

Inicialmente, foi realizada uma busca, na literatura, por um conjunto de moléculas pertencentes a uma série congênere, uma vez que, para a aplicação da metodologia MIA-QSAR, é exigida similaridade estrutural para que seja efetuado o alinhamento bidimensional, além da necessidade dos compostos apresentarem atividades biológicas experimentalmente conhecidas.

Foram selecionadas para o estudo três séries de compostos que apresentam atividade antirretroviral frente às células da linhagem MT4 (linfócitos T CD4⁺), aqui denominada como anti-HIV-1 e atividade inibitória da enzima transcriptase reversa (TR) do vírus HIV-1 ou anti-TR-1. Assim, o método MIA-QSAR foi empregado para a construção de dois modelos, de acordo com a atividade antirretroviral de cada série determinada para o estudo.

Em seguida, foram propostas estruturas com ambas as atividades descritas, por meio da combinação das subestruturas mais ativas do conjunto de dados envolvido no estudo.

Para avaliar a afinidade ligante-receptor, foi aplicada a técnica de ancoramento molecular, ou *docking*. Também foram efetuadas avaliações preliminares acerca de alguns parâmetros farmacocinéticos dos inibidores

propostos, por meio da utilização dos programas Molinspiration, PK/DB e da Plataforma ACD/Percepta, disponíveis nos sites: <http://molinspiration.com>, <http://miro.ifsc.usp.br/pkdb/> e <http://acdlabs.com/products/percepta/>, respectivamente.

4.1 A construção dos modelos MIA-QSAR

Foram obtidas, da literatura (FREEMAN et al., 1995; KUMAR; SINGH; TIWARI, 2011; LEONARD; ROY, 2004), três séries de compostos com certa similaridade estrutural, conforme Figura 19 (subestrutura em negrito), e valores experimentais de atividade antirretroviral (anti-HIV-1 e anti-TR-1), totalizando 120 estruturas químicas, como mostrado na Tabela 2. Dois modelos MIA-QSAR, modelo 1 ou anti-HIV-1 e modelo 2 ou anti-TR-1, foram construídos, conforme o tipo de atividade biológica das estruturas químicas do conjunto de dados.

Uma das séries utilizadas (série 3) apresentou certas estruturas químicas com valores experimentais para ambos os tipos de atividade biológica (anti-HIV-1 e anti-TR-1) e, por esta razão, tais estruturas foram incluídas nos dois modelos construídos, conforme a atividade apresentada.

As estruturas químicas foram sistematicamente desenhadas, independentemente do tipo de atividade biológica desempenhada, utilizando-se o programa para desenho em 2D, ChemSketch (ADVANCED CHEMISTRY DEVELOPMENT - ACD, 2009).

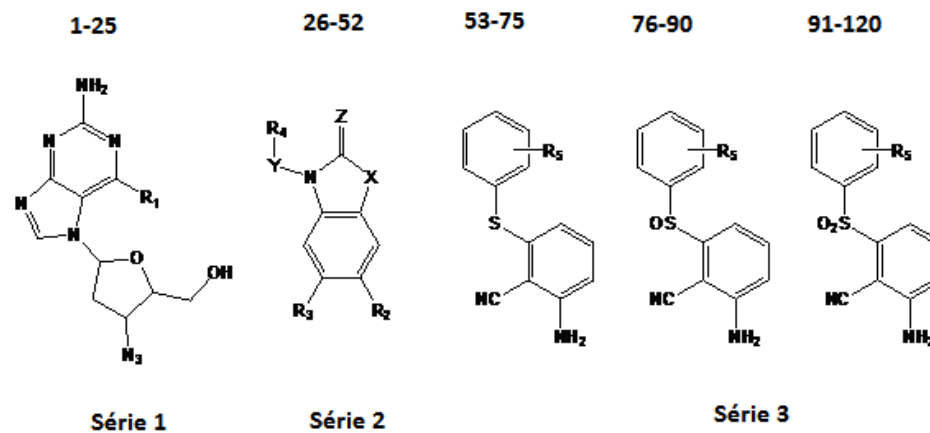
Como ilustrado na Figura 18, cada imagem foi cuidadosamente transferida para uma área de trabalho do aplicativo *Paint* do Microsoft Windows, com dimensões previamente definidas em 500x500 *pixels*. A definição do tamanho da área de trabalho no *Paint* é obrigatória, pois todas as

matrizes deverão apresentar a mesma dimensão para o posterior alinhamento das estruturas.

Em seguida, um *pixel* localizado na porção superior do anel benzênico comum a todas as estruturas químicas (em destaque na Figura 18) foi selecionado e, uma a uma, as estruturas foram movidas, de tal forma que este *pixel* fosse ajustado na coordenada 250x190 *pixels* da área de trabalho.

Este procedimento corresponde ao alinhamento 2D e é feito manualmente com o auxílio de um *mouse*. A ferramenta ZOOM, do aplicativo *Paint*, permite que a imagem seja aumentada de 100% para 800%, por exemplo, o que facilita a escolha e a seleção do *pixel* e o arraste da estrutura química até a coordenada a ser fixada.

Tabela 2 Séries de compostos usadas na construção dos modelos MIA-QSAR



Composto	R ₁	R ₂	R ₃	R ₄	X	Y	Z	R ₅	Anti-HIV-1	TR
1	AZT								8,00	
2	OH								5,30	
3	OCH ₃								5,05	
4	OCH ₂ CH ₃								4,64	
5*	OCH ₂ CH ₂ CH ₃								5,05	
6	OCH(CH ₃) ₂								4,66	
7	O(CH ₂) ₃ CH ₃								4,72	
8	O-ciclobutil								5,30	
9	OC ₆ H ₅								4,55	
10*	OCH ₂ C ₆ H ₅								5,10	
11	NH ₂								5,70	
12	HNCH ₃								4,28	
13*	HNCH ₂ CH ₃								4,77	

Tabela 2, continuação

Composto	R ₁	R ₂	R ₃	R ₄	X	Y	Z	R ₅	Anti-HIV-1	TR
14	HN(CH ₂) ₂ CH ₃								4,92	
15	HN-ciclopropil								4,74	
16	HN(CH ₂) ₃ CH ₃								4,54	
17	HN-ciclobutil								4,85	
18*	HN(CH ₂) ₂ .fenol								3,95	
19	CH ₃ NCH ₃								4,68	
20	CH ₃ NCH ₂ CH ₃								4,92	
21*	CH ₃ N(CH ₂) ₂ CH ₃								4,96	
22	CH ₃ N-ciclopropil								5,40	
23	N(CH ₂) ₃								5,05	
24*	N(CH ₂) ₄								5,10	
25	Cl								5,22	
26		H	Cl	2,6-difluorofenol	NH	CH ₂	O			6,62
27		Cl	H	2,6-difluorofenol	NH	CH ₂	O			4,53
28		H	Cl	2,6-difluorofenol	O	CH ₂	O			5,43
29		H	Cl	2,6-dichlorofenol	NH	CH ₂	O			5,42
30*		H	Cl	4-i-propilfenol	NH	CH ₂	O			5,55
31		H	Cl	α-naftol	NH	CH ₂	O			6,06
32		H	Cl	β-naftol	NH	CH ₂	O			5,20
33		H	Cl	3,5-dimetilfenol	NH	CH ₂	O			7,49
34		H	Cl	3,5-dimetilfenol	NH	SO ₂	O			8,70
35*		H	Cl	2,6-difluorofenol	NH	SO ₂	O			5,81
36		H	Cl	2,6-difluorofenol	NH	CH ₂	S			6,13
37		H	Cl	2,6-difluorofenol	NH	CH ₂	O			6,27
38		H	Cl	3,5-dimetilfenol	NH	CH ₂	S			7,12
39		H	Cl	3,5-dimetilfenol	NCOCH ₃	CH ₂	O			7,93
40*		H	Cl	3,5-difluorofenol	NH	CH ₂	O			6,54
41		H	Cl	3,5-difluorofenol	NH	SO ₂	O			7,16
42		H	Cl	3,5-difluorofenol	NCOCH ₃	CH ₂	O			5,78
43		H	CH ₃	2,6-difluorofenol	NH	CH ₂	O			4,73
44		H	CH ₃	2,6-difluorofenol	NH	SO ₂	O			6,71
45*		H	CH ₃	3,5-dimetilfenol	NH	CH ₂	O			7,12
46		H	CH ₃	3,5-dimetilfenol	NH	SO ₂	O			8,20

Tabela 2, continuação

Composto	R ₁	R ₂	R ₃	R ₄	X	Y	Z	R ₅	Anti-HIV-1	TR
47		H	CH ₃	3,5-difluorofenol	NH	SO ₂	O			7,19
48		F	Cl	3,5-dimetilfenol	NH	CH ₂	O			6,03
49		F	Cl	3,5-dimetilfenol	NH	SO ₂	O			6,35
50*		F	Cl	2,6-difluorofenol	NH	CH ₂	O			4,54
51		H	CF ₃	3,5-dimetilfenol	NH	CH ₂	O			6,37
52		H	CF ₃	2,6-difluorofenol	NH	CH ₂	O			5,26
53								H	4,84	5,06
54								2-OCH ₃	5,37	5,57
55*								3-OCH ₃	5,22	5,82
56								2-CH ₃	4,80	-
57								3-CH ₃	5,21	6,02
58								4-CH ₃	3,94	5,24
59								2-Cl	5,39	5,14
60*								3-Cl	5,13	4,80
61								4-Cl	-	4,92
62								2-Br	4,52	-
63								3-Br	5,29	4,82
64								3-F	5,01	4,92
65*								2-CN	-	5,04
66								3-CN	5,76	5,96
67								4-CN	4,36	-
68								3-CF ₃	4,89	5,15
69								3-NH ₂	4,50	-
70*								2,5-Cl ₂	-	5,46
71								3,5-(CH ₃) ₂	6,37	5,96
72								3,5-Cl ₂	-	6,92
73								3-Cl, 5-CH ₃	5,75	5,77
74								3-OCH ₃ , 5-CH ₃	5,70	6,85
75*								3-OCH ₃ , 5-CF ₃	5,29	4,89
76								2-OCH ₃	5,32	4,92
77								3-OCH ₃	4,80	4,72
78								2-CH ₃	4,03	-
79								3-CH ₃	4,53	5,00

Tabela 2, continuação

Composto	R ₁	R ₂	R ₃	R ₄	X	Y	Z	R ₅	Anti-HIV-1	TR
80*								4-CH ₃	4,31	-
81								2-Br	4,41	-
82								3-Br	7,10	5,32
83								4-Br	4,69	-
84								2-CN	5,41	5,00
85*								3-CN	4,85	-
86								3-CF ₃	4,40	-
87								3,5-(CH ₃) ₂	6,47	6,30
88								2,5-Cl ₂	5,01	5,20
89								3-Cl, 5-CH ₃	6,49	6,28
90*								3-OCH ₃ , 5-CF ₃	5,68	6,04
91								H	5,70	5,16
92								2-OCH ₃	6,22	5,85
93								3-OCH ₃	6,04	6,22
94								4-OCH ₃	4,60	4,88
95*								2-CH ₃	5,64	5,35
96								3-CH ₃	6,40	6,70
97								4-CH ₃	5,02	5,14
98								2-Cl	5,39	5,23
99								3-Cl	6,23	6,40
100*								4-Cl	5,52	-
101								2-Br	5,30	4,92
102								3-Br	6,27	6,70
103								4-Br	4,70	-
104								2-F	5,52	5,30
105*								3-F	5,52	-
106								2-CN	5,27	5,22
107								3-CN	5,62	5,74
108								4-CN	4,10	-
109								3-CF ₃	4,46	5,28
110*								2,5-Cl ₂	6,52	6,52
111								3,5-Cl ₂	7,15	7,52
112								3,5-(CH ₃) ₂	8,00	8,15

Tabela 2, conclusão

Composto	R ₁	R ₂	R ₃	R ₄	X	Y	Z	R ₅	Anti-HIV-1	TR
113								3-Br, 5-CH ₃	7,70	8,52
114								3-Cl, 5-CH ₃	7,52	8,30
115*								3-OCH ₃ , 5-CH ₃	7,30	8,00
116								3-OCH ₃ , 5-CF ₃	7,05	7,40
117								3-OH, 5-CH ₃	6,37	-
118								3-OCH ₂ CH ₃ , 5-CH ₃	7,22	-
119								3-O(CH ₂) ₂ CH ₃ , 5-CH ₃	7,22	-
120*								3-O(CH ₂) ₃ CH ₃ , 5-CH ₃	6,22	6,40

Valores de atividade anti-HIV-1 e TR expressos em pIC₅₀

* Compostos destinados à validação externa ou *test set* dos modelos construídos

Fonte: Freeman et al. (1995), Kumar, Singh e Tiwari (2011) e Leonard e Roy (2004)

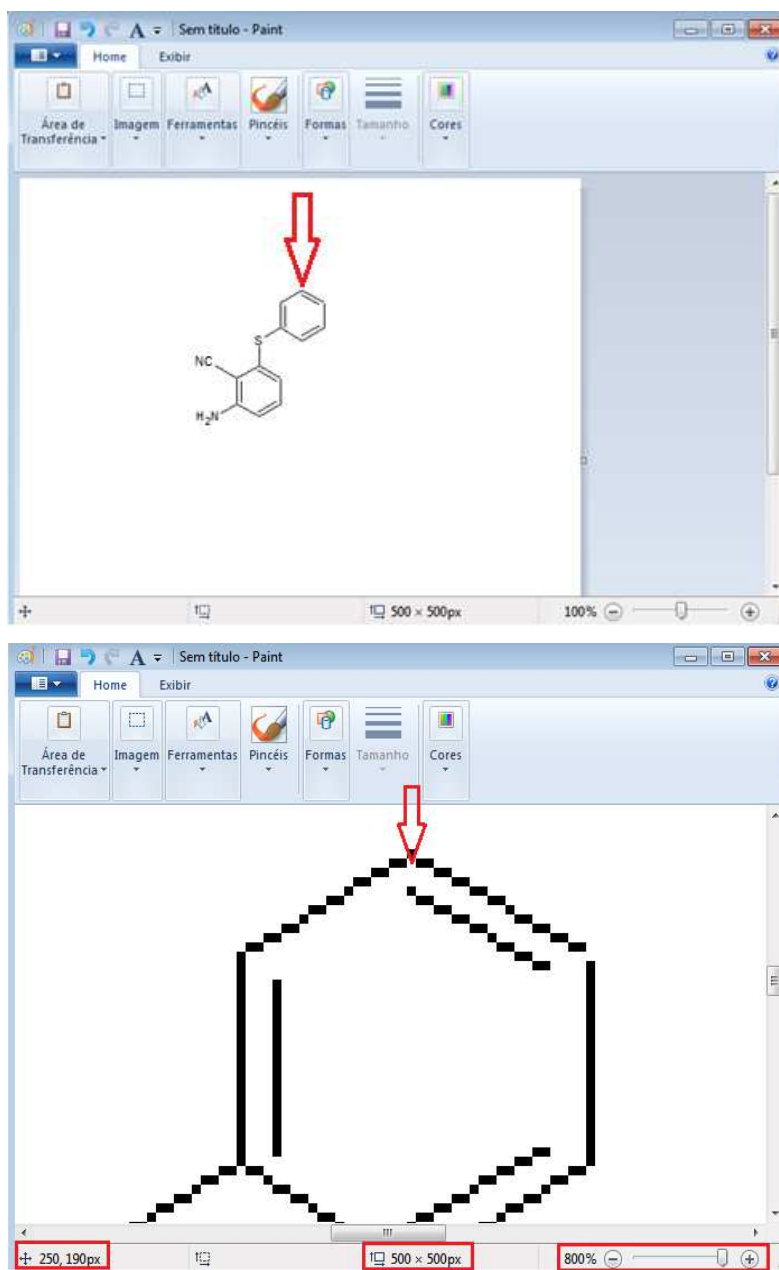


Figura 18 Área de trabalho do aplicativo *Paint* na dimensão 500x500 *pixels*. Escolha do *pixel* no aumento de 800%, para alinhamento e fixação na coordenada 250x190 *pixels*

Após conferir se a estrutura química encontrava-se fixada na coordenada correta (250x190 *pixels*), cada imagem foi salva como *bitmaps* (.bmp). Esta etapa é considerada fundamental na análise, pois permite que cada imagem em um plano bidimensional seja sobreposta à outra, dando origem a um arranjo tridimensional em que as partes comuns entre as estruturas da série congênere estejam congruentes, conforme Figura 19. Dessa forma, os substituintes e as porções das moléculas que variam entre si explicam a variância dos dados biológicos (diferentes valores de atividade biológica).

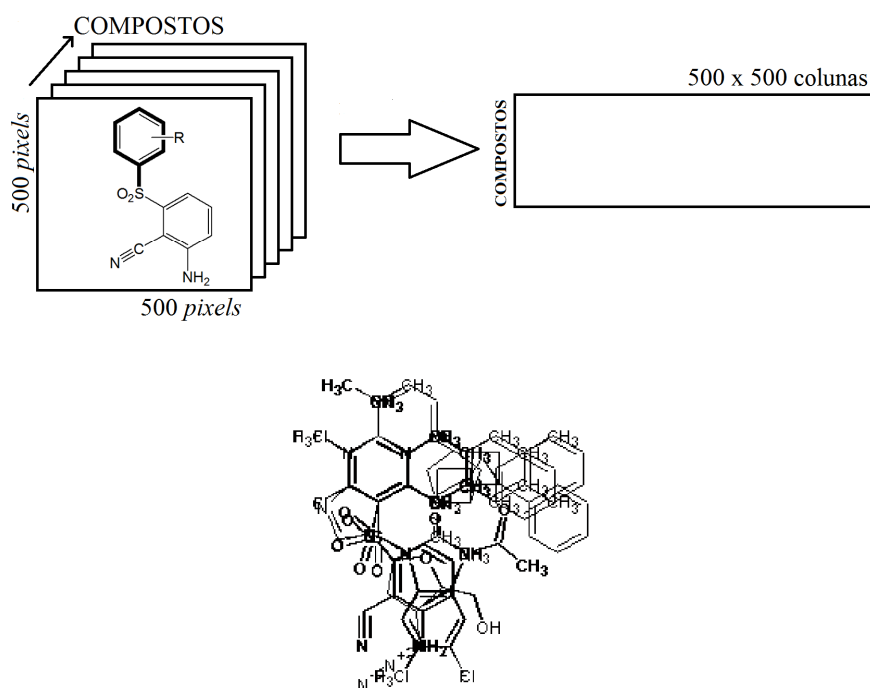


Figura 19 Sobreposição das estruturas químicas e formação do arranjo tridimensional. Em negrito está a porção similar das estruturas sobrepostas e as porções variáveis são os substituintes que respondem pela variação na atividade biológica de cada composto

Fonte: Guimarães et al. (2014)

Em MIA-QSAR, as imagens correspondem ao desenho das estruturas químicas em 2D e o método se fundamenta no uso dos *pixels* das imagens como descritores, uma vez que os *pixels* podem ser tratados numericamente como binários, em que a cor branca equivale ao dígito 765 e os *pixels* pretos correspondem ao dígito 0, de acordo com o sistema de cores RGB. O dígito 765 provém da soma das cores vermelha, verde e azul (*red, green e blue* do RGB), cada uma equivalente a 255 e a ausência das três cores do sistema responde pela cor preta ou dígito 0.

As modificações estruturais, ou mudanças na posição dos substituintes em uma série congênere de estruturas químicas, respondem por alterações nas coordenadas dos *pixels* da imagem, refletindo na variância do bloco Y das atividades biológicas.

A formação do arranjo tridimensional pelo alinhamento ou a sobreposição das estruturas é importante pela possibilidade de ser desdobrado em um arranjo bidimensional de descritores (matriz X), o que permite sua regressão com o bloco Y das atividades biológicas, por meio de mínimos quadrados parciais (PLS bilinear, do inglês *partial least squares*), método de calibração multivariada mais indicado quando o número de descritores excede o número total de amostras do estudo. O procedimento para o tratamento e a análise das imagens, ilustrado na Figura 20, e os comandos a seguir demonstram como as imagens podem ser carregadas e convertidas em binários por meio do programa Matlab.

Cada *pixel* da área de trabalho definida em 500x500 corresponde a um descritor, o que resultou, após o desdobramento, em uma matriz X composta por 250.000 descritores ($500 \times 500 = 250.000$). Cada linha da matriz X corresponde a uma estrutura química pertencente ao conjunto de dados, sendo descrita por uma sequência de descritores 0 ou 765, de acordo com a cor do *pixel* da imagem.

O conjunto de dados composto por um total de 120 moléculas foi dividido de acordo com os dois tipos de atividades biológicas desempenhadas pelas estruturas químicas selecionadas para o estudo (anti-HIV-1 e anti-TR-1). Dessa forma, o conjunto referente à atividade anti-HIV-1 foi composto por 90 estruturas, enquanto o conjunto com atividade inibitória frente à enzima transcriptase reversa (TR) era formado por 78 compostos. Assim, foram obtidas duas matrizes X, uma para cada modelo a ser construído, de acordo com a atividade biológica experimental apresentada.

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

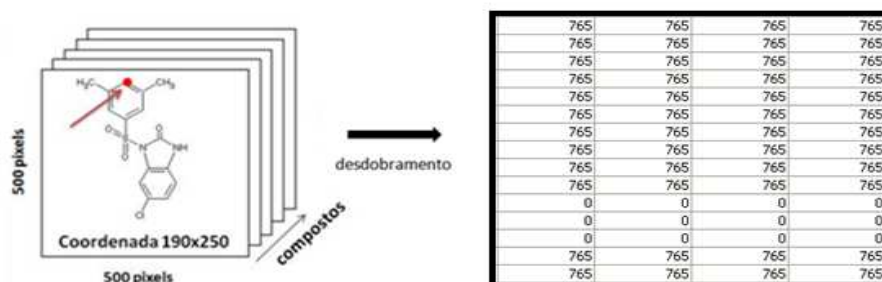


Figura 20 Interface do programa Matlab, utilizado na geração da matriz X de descritores. Abaixo são ilustrados a construção do arranjo tridimensional e o desdobramento da matriz

A matriz X da série com atividade anti-HIV-1 apresenta dimensão de 90 linhas (compostos) por 250.000 colunas (descritores), enquanto a matriz X da série anti-TR-1 foi composta por 78 linhas por 250.000 colunas, pois as etapas de desenho das estruturas químicas e o alinhamento no *Paint* foram realizadas simultaneamente, sendo definida a mesma dimensão da área de trabalho

(500x500 *pixels*) e mesma coordenada (250x190 *pixels*) para o total de 120 moléculas participantes do estudo.

Isto explica o fato de ambas as matrizes apresentarem o mesmo número de colunas ($500 \times 500 = 250.000$). A conversão de estruturas químicas em sequências numéricas, ferramenta do método MIA-QSAR, é uma analogia ao funcionamento dos computadores que, por meio de códigos binários (0 e 1), formam sequências que são traduzidas em letras, comandos e figuras, dentre outros.

Com a finalidade de agilizar os cálculos, as colunas com variância zero, como os espaços em branco comuns a todas as imagens ou as partes congruentes das estruturas químicas, foram removidas pelo programa Matlab 7.5 (MATHWORKS, 2007), ocasionando a redução dos descritores da matriz X e a consequente redução da demanda computacional. Tal procedimento reduziu a matriz X das atividades anti-HIV-1 para a dimensão de 90x4015, ou seja, o número de descritores desta matriz passou de 250.000 para 4015, enquanto a matriz X anti-TR-1 passou a apresentar dimensões de 78x2977.

Para a construção dos modelos MIA-QSAR, os conjuntos de dados (90 amostras do modelo 1 e 78 do modelo 2) devem ser divididos em conjunto treinamento e teste. O conjunto teste, ou *test set*, destinado à validação externa obrigatória nesse tipo de metodologia é composto por 20% do total das amostras destinadas ao estudo.

Antes de se efetuar a divisão dos conjuntos de dados em treinamento e teste, foi realizada a análise da presença de *outliers* para ambas as séries, conforme descrito a seguir.

4.2 Identificação de *outliers*

A detecção de amostras anômalas, conhecidas na literatura como *outliers*, é aplicada para garantir que as amostras que irão compor o conjunto treinamento a ser utilizado na construção do modelo de regressão formem um conjunto homogêneo (MARTINS; FERREIRA, 2013).

Compostos estruturalmente diferentes dos demais ou com valores de atividade biológica atípicos podem influenciar de maneira negativa os resultados obtidos pelo modelo construído e sua remoção do conjunto de dados pode melhorar a qualidade estatística do modelo, embora deva ser reconsiderada a deleção, no caso de conjuntos de amostras de tamanho reduzido (MARTINS, 2010).

A identificação de estruturas anômalas em um conjunto de dados pode ser realizada por um procedimento considerado comum em quimiometria. A análise de componentes principais (PCA, do inglês *Principal Components Regression*) é utilizada devido à sua capacidade de compreensão dos dados em função da existência de correlação entre diversas variáveis medidas (SABIN; FERRÃO; FURTADO, 2004).

Por meio de PCA, dados complexos são apresentados sob a forma de gráficos que evidenciam informações relevantes do sistema. PCA (PEARSON, 1901) cria p variáveis latentes (Y) como combinações lineares dos p variáveis originais (X), de tal maneira que os novos eixos ortogonais (componentes principais, PCs) são construídos para explicar a variação máxima possível em apenas algumas dimensões

$$Y_i = e_i^T X = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p \quad (3)$$

em que o vetor e_i , desconhecido estabelece a combinação linear, para $i = 1, \dots, p$.

Uma propriedade importante das PCs é que elas são completamente não correlacionadas e ortogonais entre si, ou seja, a informação contida em uma delas não está presente em outra. Além disso, outro fato importante está na quantidade de informação dos dados originais que cada uma dessas novas variáveis é capaz de descrever (MONTANARI, 2011).

Desse modo, foi desenvolvida uma análise de componentes principais (PCA) para agrupar as amostras com propriedades semelhantes, por meio do uso do programa Chemoface (NUNES et al., 2012), com a intenção de identificar possíveis *outliers* no conjunto de dados.

É importante deixar claro que a remoção de amostras deve ser realizada de forma cautelosa, pois uma explicação química ou biológica deve ser dada para cada amostra classificada como atípica (MARTINS; FERREIRA, 2013).

4.3 Validações e calibração dos modelos MIA-QSAR

Para a construção dos modelos MIA-QSAR, os dois conjuntos de amostras (89 anti-HIV-1, pois uma amostra foi removida e 78 inibitórias da TR) foram divididos entre si, em grupos treinamento (cerca de 80% dos compostos de cada série) e teste (20% dos conjuntos de dados). As amostras pertencentes ao grupo teste foram escolhidas de forma aleatória, evitando-se os valores extremos de pIC_{50} , mantendo-se uma distribuição regular de tais valores ao longo do conjunto formado.

As validações estatísticas são necessárias para que o modelo construído possa ser considerado robusto e preditivo. Além da validação externa (conjunto teste), também é aplicada a validação interna ou cruzada *leave-one-out* (LOOCV, deixe uma de fora), para se determinar o número ótimo de variáveis latentes ou componentes PLS para a construção do modelo MIA-QSAR.

Na validação cruzada, o conjunto treinamento é dividido em grupos e vários modelos com o mesmo número de variáveis latentes são construídos, sempre deixando uma molécula de fora do modelo. A variável dependente (atividade biológica) é, então, prevista pelo modelo construído para a molécula que ficou de fora e este processo é repetido até que todas as moléculas do conjunto treinamento tenham sido deixadas de fora uma vez, caracterizando o *leave-one-out*. Dessa forma, é possível ter uma ideia inicial do poder de previsibilidade e robustez, e bons modelos devem apresentar valores de $q^2 > 0,5$ (r^2_{Loocv}) e $r^2_{\text{cal}} > 0,8$ (calibração), segundo a literatura especializada.

Após a definição do número ótimo de variáveis latentes pela validação cruzada, os modelos MIA-QSAR foram obtidos por meio de uma regressão por mínimos quadrados parciais (PLS) entre a matriz de descritores e o bloco Y referente às atividades biológicas das estruturas químicas do conjunto treinamento, como ilustrado na Figura 21. O método PLS, geralmente, produz modelos mais parcimoniosos, com um número menor de fatores e mantendo um bom ajuste, sendo um dos mais utilizados em metodologias QSAR (MARTINS; FERREIRA, 2013).

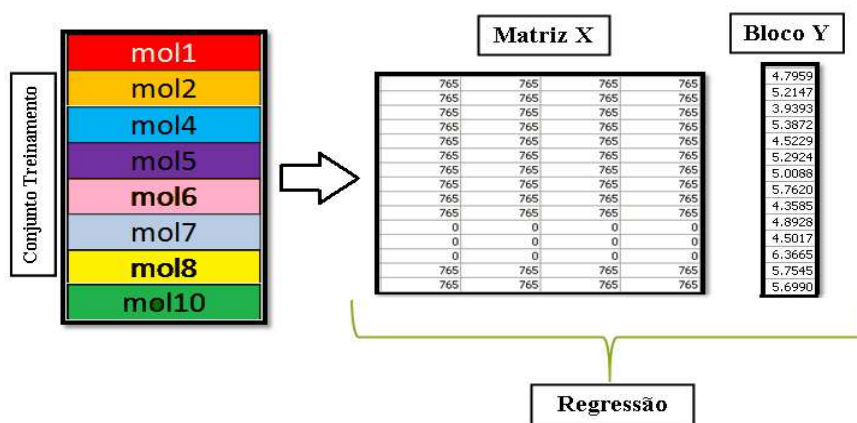


Figura 21 Calibração do modelo MIA-QSAR por meio da regressão por mínimos quadrados parciais (PLS)

Segundo Golbraikh e Tropsha (2002), a forma mais segura de se atestar o caráter preditivo de modelos QSAR (que inclui o método MIA-QSAR) é por meio da aplicação da validação externa para um grupo teste (*test set*). Este tipo de validação, ilustrada na Figura 22, utiliza o modelo construído para prever a atividade biológica (pIC_{50Pred}) das moléculas pertencentes ao conjunto teste, ou seja, moléculas que não participaram da construção do modelo. Como os valores experimentais das atividades biológicas de tais moléculas são conhecidos, é possível comparar os valores experimentais e preditos por meio de parâmetros estatísticos similares aos da validação cruzada. Bons modelos devem apresentar valores de $r^2_{teste} \geq 0,5$.

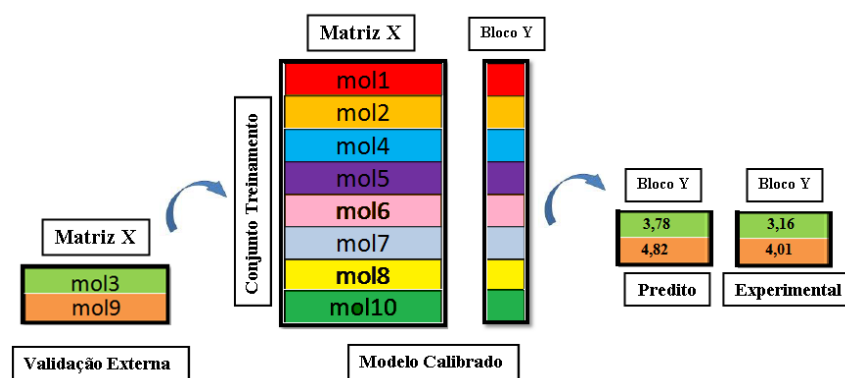


Figura 22 Validação externa ou *test set*

Uma limitação da validação cruzada é que ela não fornece sozinha a significância estatística do poder de predição estimado para o modelo e, para isso, é utilizada uma validação adicional denominada *Y-randomization*. Nesse tipo de validação, esquematizada na Figura 23, o bloco Y, referente aos valores das atividades biológicas experimentais (pIC_{50Exp}), é randomizado, de forma que cada pIC_{50Exp} não mais corresponda à sua verdadeira amostra (estrutura química que é responsável pela atividade exercida), sendo este processo realizado com a

finalidade de atestar que o modelo obtido não é obra do acaso. Após a regressão da matriz X com o bloco Y randomizado, esperam-se valores de r^2_{rand} consideravelmente menores que os resultados obtidos para o r^2_{Cal} , a fim de provar a relação existente entre uma estrutura química e determinada atividade biológica.

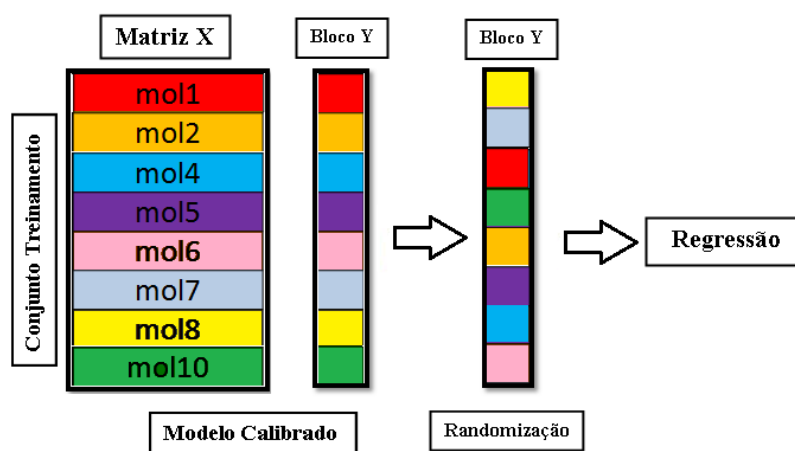


Figura 23 Randomização do bloco Y e sucessiva regressão com a matriz X intacta via PLS. O processo é repetido algumas vezes, para atestar a correlação do modelo obtido pela calibração

Outras validações adicionais podem ser aplicadas nesse tipo de metodologia, sendo sugeridos dois cálculos simples e úteis para esta finalidade. O primeiro, proposto por Roy et al. (2010), refere-se à determinação do r^2_m ou coeficiente de correlação ao quadrado modificado, obtido conforme a Equação 4, que comprova que os resultados da validação externa não somente exibem boa correlação entre pIC_{50} experimentais e preditos, mas que ambos apresentam valores similares (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2010, 2012),

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

em que r^2 e r_0^2 correspondem aos valores do coeficiente de correlação ao quadrado entre os valores de pIC_{50} observados e preditos para o conjunto teste com e sem intercepto, respectivamente.

Já o segundo, denominado ${}^c r_p^2$, proposto por Mitra, Saha e Roy (2010), é aplicado com a finalidade de validar o *Y-randomization*. Este parâmetro determina a diferença estatística entre os valores de r_{Cal}^2 (r^2 da calibração) e r_{rand}^2 (r^2 da randomização) e é calculado de acordo com a Equação 4. São aceitáveis valores de r_m^2 e ${}^c r_p^2$ iguais ou superiores a 0,5 (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2010, 2012).

$${}^c r_p^2 = r (r^2 - r_{rand}^2)^{1/2} \quad (5)$$

Após a construção dos dois modelos, foram propostos novos candidatos a antirretrovirais com ambas as atividades simultaneamente, por meio da estratégia de combinar as subestruturas mais ativas de cada série.

Tais compostos com suposta múltipla atividade frente ao vírus HIV-1 passaram por estudos de *docking* e análise dos parâmetros farmacocinéticos a seguir, com a finalidade de atestar o desenvolvimento das atividades biológicas sugeridas.

4.4 Docking

Após a busca pela estrutura cristalográfica da TR e do complexo glicoproteico gp120 e gp41 (associadas à atividade anti-HIV) no Banco de Dados de Proteínas (PDB), representadas pelos códigos 3C6T, 1G9N e 2OT5,

respectivamente, foram realizados os estudos de ancoramento molecular, ou *docking*, para as estruturas propostas. O programa Molegro Virtual Docker® (MVD) (THOMSEN; CHRISTENSEN, 2006) foi utilizado para prever a conformação mais provável de cada ligante no interior das macromoléculas selecionadas no PDB. A função *score* MolDock (E_{score}) utilizada na MVD® é definida por

$$E_{score} = E_{Inter} + E_{Intra} \quad (6)$$

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

4.5 Parâmetros farmacocinéticos e biodisponibilidade

A “regra dos cinco” de Lipinski, assim conhecida pelo fato de os valores de fronteira serem 5, 500, 2*5 e 5, é aplicada a candidatos a fármacos com a finalidade de avaliar o perfil farmacocinético de tais compostos.

De acordo com esta regra, que estabelece peso molecular ≤ 500 , $\log P \leq 5$, número de aceptores de ligação hidrogênio (n_{ON}) ≤ 10 e número de doadores de ligação hidrogênio (n_{OHNH}) ≤ 5 , moléculas que violam mais de um desses valores de referência podem apresentar problemas relacionados à sua biodisponibilidade por influenciar no comportamento da molécula no organismo vivo (LIPINSKI et al., 1997).

Por meio do programa Molinspiration (MOLINSPIRATION CHEMINFORMATICS, 1986), foram calculados e avaliados parâmetros que incluem a regra de Lipinski, juntamente com área superficial polar da molécula, denominado TPSA.

A análise computacional do perfil físico-químico, farmacocinético e de segurança (ADMET), que envolve a absorção, a distribuição, o metabolismo e a excreção, foi realizada para as moléculas com múltipla atividade propostas no estudo por meio do módulo *Percepta* do programa ACDLabs (PLATAFORMA..., 2013).

Introspecções similares foram obtidas utilizando-se o programa Hologram QSAR Technique, desenvolvido por Moda et al. (2008), de acesso por meio da página www.pkdb.ifsc.usp.br. Esta base permite a avaliação computacional da absorção intestinal humana, a biodisponibilidade oral, a penetração à barreira hematoencefálica e a solubilidade de novos candidatos a fármacos.

5 RESULTADOS E DISCUSSÃO

A estratégia utilizada neste estudo foi a construção de dois modelos MIA-QSAR, a partir da combinação de três séries de estruturas químicas com dados experimentais em pIC_{50} (IC_{50} em mol/L^{-1}) para atividades anti-HIV-1 e de inibição da transcriptase reversa do HIV-1 (anti-TR-1).

Vale ressaltar que se entende por atividade anti-HIV-1 a inibição da atividade de uma ou mais enzimas (TR, IN ou PR), proteínas do envelope viral gp41 e gp120 ou correceptores (CCR5 e CXCR4) envolvidos no processo de infecção da célula alvo e replicação do HIV-1 no organismo infectado.

Diante das três séries de compostos disponíveis, o modelo 1 foi obtido pela combinação das séries 1 e 3, num total de 90 compostos com atividade anti-HIV-1 e o modelo 2 se deu pela combinação dos dados das séries 2 e 3, totalizando 78 compostos com atividade inibitória da TR (anti-TR-1).

Contudo, o composto 1 (substituinte AZT) foi identificado como um *outlier*, provavelmente devido à complexidade da estrutura de seu substituinte R e, por este motivo, foi removido do conjunto de dados. Assim, 89 compostos com atividade anti-HIV-1 foram selecionados para participar do estudo e divididos em conjunto treinamento (71 compostos - 80% do total de compostos do grupo) destinado à calibração e à validação cruzada, e conjunto teste, formado por 18 estruturas, para o desenvolvimento da validação externa obrigatória para esse tipo de metodologia.

O conjunto de dados do modelo 2 não apresentou *outliers* e seus 78 compostos foram, então, divididos, dos quais 63 foram destinados ao conjunto treinamento e 15 compostos, ao conjunto teste.

Após a construção dos modelos 1 e 2, novos inibidores foram propostos pela combinação de subestruturas dos compostos mais ativos de cada uma das séries 1 e 3, e 2 e 3, caracterizados como multialvos por apresentarem porções

anti-HIV-1 e anti-RT-1 que, de acordo com o estudo, são atividades que não se correlacionam, sendo, portanto, distintas.

5.1 Modelo 1: anti-HIV-1

A atividade anti-HIV-1, aqui descrita, foi testada experimentalmente em âmbito celular (frente às células MT4), tendo sido observada estreita correlação entre a regulação dos receptores de superfície CD4 e a potência da atividade anti-HIV-1, para uma série de compostos macrociclos (VERMEIRE et al., 2003).

Contudo, não foi possível definir o alvo enzimático (CD4, gp41 e gp120, TR, IN ou PR), visto que há a possibilidade de inibição de um ou mais alvos biológicos, simultaneamente.

Uma vez que não houve correlação linear entre a atividade anti-HIV-1 e anti-TR-1 aplicada neste estudo, os compostos propostos pela combinação de subestruturas mais ativas de ambas as séries podem apresentar um alvo diferente da TR e, já que as atividades anti-TR-1 são conhecidas, pode-se dizer que tais compostos têm, pelo menos, dois alvos distintos.

Como foi dito anteriormente, o método utilizado na construção dos modelos foi o MIA-QSAR (FREITAS; BROWN; MARTINS, 2005), em que os descritores binários (0 e 765), derivados dos *pixels* das imagens das estruturas químicas bidimensionais, foram correlacionados aos dados biológicos correspondentes por meio de regressão por mínimos quadrados parciais (PLS).

Nesse tipo de metodologia, as imagens das estruturas químicas do conjunto de dados são sobrepostas a partir da determinação de um *pixel* comum a todas as estruturas da série e a porção variável ou os substituintes respondem pela variância no bloco Y referente às bioatividades, enquanto as partes

congruentes da série são utilizadas no alinhamento 2D, necessário para o desdobramento da matriz X que sofrerá a regressão.

Valores de r^2 e q^2 foram utilizados na avaliação dos modelos, sendo definidos como

$$1 - \frac{[\sum(y_i - \hat{y}_i)^2 / \sum(y_i - \bar{y})^2]}{\quad} \quad (7)$$

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

Embora o modelo 1 compreenda os compostos da série 1 (atividade anti-HIV-1 para determinadas células da linhagem MT4, como os linfócitos T CD4⁺) e da série 3 (estruturas com atividade inibitória da TR e atividade anti-HIV-1), não existe correlação entre a potência anti-HIV-1 e a inibição da TR para compostos da série 3, ou seja, o fato de determinado composto ser altamente ativo (elevado pIC_{50}) contra um dos alvos biológicos não significa, necessariamente, que apresente eficácia similar frente ao segundo ou mais alvos.

Numa análise preliminar utilizando os pixels das imagens para a aplicação a técnica PCA, o composto 1, o mais ativo de sua série, representado no gráfico como molécula 69, foi identificado como um *outlier*, por ter apresentado um comportamento anômalo quando comparado aos demais, provavelmente devido à complexidade da estrutura de seu substituinte R (AZT), não podendo fazer parte do modelo de calibração.

Foi construído um gráfico (Gráfico 1) de *scores* da componente principal 2 (CP2) eixo Y *versus* componente principal 1 (CP1) eixo X. A CP1 é definida pela direção que descreve a máxima variância dos dados originais e, neste caso, explica 90, 95% do sistema. A segunda componente principal (CP2)

tem a direção de máxima variância dos dados no subespaço ortogonal à CP1, aqui explicando 3,57% do sistema.

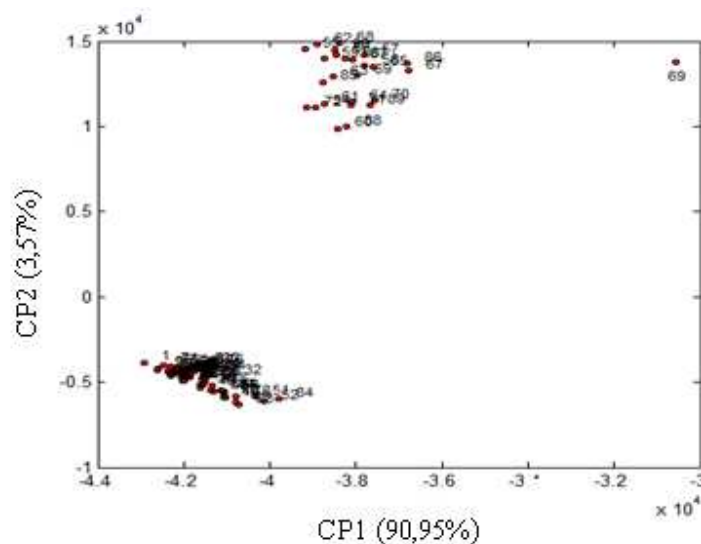


Gráfico 1 Gráfico de *scores* da componente principal 2 (CP2) eixo Y versus componente principal 1 (CP1) eixo X

Após sua remoção do conjunto de dados, o modelo 1 foi construído a partir das 89 estruturas restantes nesta série (71 no conjunto treinamento e 18 no conjunto teste).

Como resultado, o modelo 1 apresentou um r^2_{Cal} de 0,934 (erro quadrado médio, RMSE, de 0,239), utilizando-se 9 variáveis latentes PLS, conforme o Gráfico 2. A validação cruzada *leave-one-out* resultou em um q^2 de 0,678 (RMSE de 0,535), enquanto a validação externa, atualmente obrigatória para atestar a confiabilidade de modelos QSAR, apresentou r^2_{teste} igual a 0,632 (RMSE de 0,489).

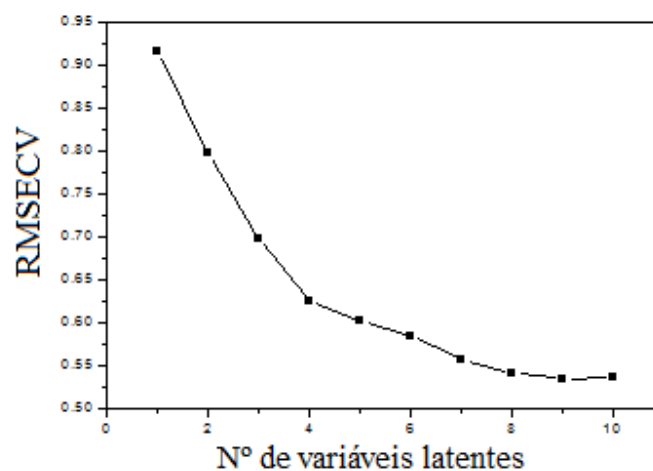


Gráfico 2 Determinação do número ótimo de variáveis latentes PLS para o modelo 1

Além disso, foi desenvolvida uma validação adicional, denominada *Y-randomization*, baseada na randomização do bloco Y (atividades biológicas) e sua posterior regressão com a matriz X de descritores (mantida intacta). Como resultado para uma média de 10 repetições, foi obtido r^2_{rand} de 0,646, consideravelmente menor que o resultado obtido para o r^2_{Cal} , conforme Tabela 3, o que prova que a relação existente entre uma estrutura química e determinada atividade biológica não é obra do acaso.

Tabela 3 Resultados estatísticos obtidos para o modelo por meio da utilização de 9 variáveis latentes PLS

Parâmetro	Calibração	LOOCV	Conjunto teste	<i>Y-randomization</i> ^a
r^2	0,934	0,677	0,632	0,646±0,05
RMSE	0,239	0,535	0,489	0,552±0,04

^a média de 10 repetições

Para garantir maior confiabilidade ao modelo MIA-QSAR construído, foram realizadas mais duas validações adicionais: o r_m^2 ou r^2 modificado e o r_p^2 . O r_m^2 obtido para o modelo 1 foi de 0,500, enquanto o r_p^2 foi de 0,519, considerados aceitáveis, conforme os valores de referência (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2010, 2012).

Os bons resultados do modelo 1 (incluindo as validações) confirmam sua satisfatória capacidade preditiva, representada pelo Gráfico 3.

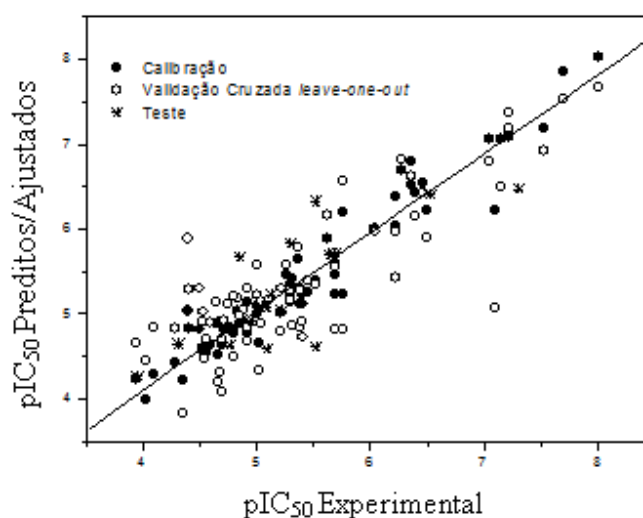


Gráfico 3 pIC₅₀ experimentais *versus* ajustados/preditos da série de compostos com atividade anti-HIV-1 (modelo 1).

5.2 Modelo 2: anti-TR-1

Abordagem semelhante à utilizada no desenvolvimento do modelo 1 foi aplicada às séries 2 e 3, na construção do modelo 2. Dados de atividade inibitória frente à enzima TR estão disponíveis para ambas as séries (2 e 3), embora a série 3 apresente também dados sobre a atividade anti-HIV-1.

Assim, os novos candidatos a inibidores da TR com adicional atividade anti-HIV-1 podem ser propostos pela combinação de subestruturas dessas duas diferentes classes. Uma vez que dados de variância de ambas as séries tenham sido incluídos em um único modelo QSAR (modelo 2), estruturas mistas podem ter suas atividades inibitórias da TR estimadas pelo modelo, bem como podem ser validadas por estudos de *docking*.

Partindo dos 78 compostos com bioatividade inibitória da TR (63 no conjunto treinamento e 15 no conjunto teste), o modelo 2 foi construído por meio de regressão por mínimos quadrados parciais (PLS) e, utilizando 8 variáveis latentes (conforme Gráfico 4), foi obtido um satisfatório modelo de calibração ($r^2 = 0,910$ e $RMSE = 0,318$).

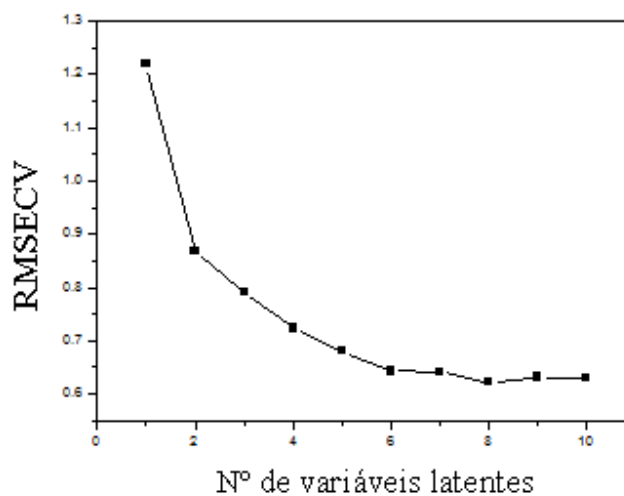


Gráfico 4 Número ótimo de variáveis latentes (PLS) para o modelo 2

A validação cruzada *leave-one-out* resultou em um q^2 de 0,663 ($RMSE = 0,621$) e a validação externa apresentou r^2_{teste} igual a 0,667 ($RMSE = 0,636$). Com a randomização do bloco Y e sua regressão com a matriz X intacta, foi

obtido, para uma média de 10 repetições, um $r^2 = 0,609$ com RMSE= 0,663. Os dados estatísticos das análises realizadas estão mostrados na Tabela 6.

Tabela 4 Parâmetros estatísticos obtidos para o modelo MIA-QSAR 2

Parâmetro	Calibração	LOOCV	Conjunto teste	Y-randomization ^a
r^2	0,910	0,663	0,667	0,609±0,02
RMSE	0,318	0,621	0,636	0,663±0,03

^a média de 10 repetições

Para o modelo 2 também foram realizadas as validações adicionais para o r^2_{teste} e r^2_{rand} . O r^2_m , que avalia o resultado obtido para o r^2_{teste} , foi satisfatório (0,570), igualmente ao r^2_p , usado para validar o r^2_{rand} , que também se manteve dentro do recomendado pela literatura (0,523) (MITRA; SAHA; ROY, 2010; OJHA et al., 2011; ROY et al., 2010, 2012).

O modelo 2, assim como o modelo 1, mostrou-se útil na predição de novos compostos, tendo confirmada a sua satisfatória capacidade preditiva pelo Gráfico 5.

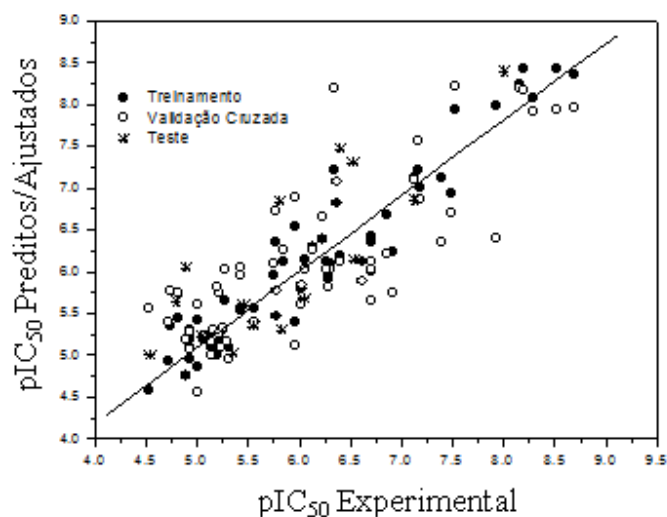


Gráfico 5 pIC_{50} experimentais *versus* ajustados/preditos da série de compostos com atividade inibitória de TR (modelo 2)

5.3 Proposta de novos possíveis inibidores

É importante deixar claro que o fato de determinado composto da série 3 (que apresenta valores de atividade biológica anti-HIV-1 e anti-TR-1) ser altamente ativo contra um dos alvos biológicos (TR ou anti-HIV-1) não significa, necessariamente, que o mesmo apresente potência similar frente ao segundo ou mais alvos, ou múltipla eficácia.

Após a construção dos modelos MIA-QSAR 1 e 2, os compostos mais ativos (maiores valores de bioatividade) de cada série tiveram suas subestruturas combinadas, de acordo com cada tipo de atividade biológica apresentada (anti-HIV-1 e inibidores da TR) e, em seguida, seus pIC_{50} foram preditos por seus respectivos modelos.

Dessa forma, em relação ao modelo 1 (atividade anti-HIV-1), as moléculas mais ativas das séries 1 e 3 (11, 112 e 113, Figura 24) foram

combinadas por meio de suas subestruturas (parte em vermelho e preto), dando origem a diferentes compostos com suposta atividade anti-HIV-1.

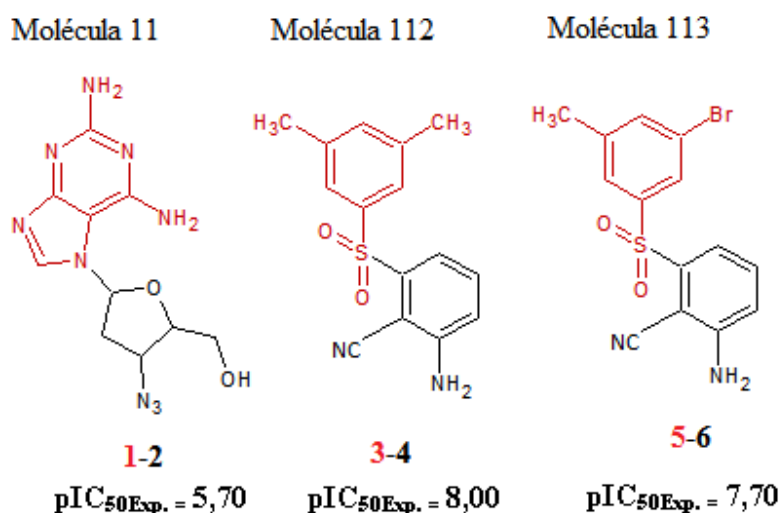


Figura 24 Compostos biologicamente mais ativos das séries 1 e 3, de acordo com os valores de atividade experimental, que foram combinados (parte em vermelho de uma molécula com a parte preta de outra) para a proposta de novos inibidores com atividade anti-HIV-1

As combinações relevantes para este trabalho foram as realizadas pela parte em vermelho da molécula 112, denominada como **3**, com a parte preta da molécula 11, referente ao número 2 e, dessa forma, as seguintes combinações foram possíveis: **3-2** (molécula A), **5-2** (molécula B) e **1-4** (molécula C).

Já em relação à proposta de novos inibidores com atividade inibitória da TR (referente ao modelo 2), foram utilizadas as moléculas mais ativas experimentalmente das séries 2 e 3 (34 e 113), conforme Figura 25, utilizando-se o método de combinação entre suas subestruturas para propor novos inibidores.

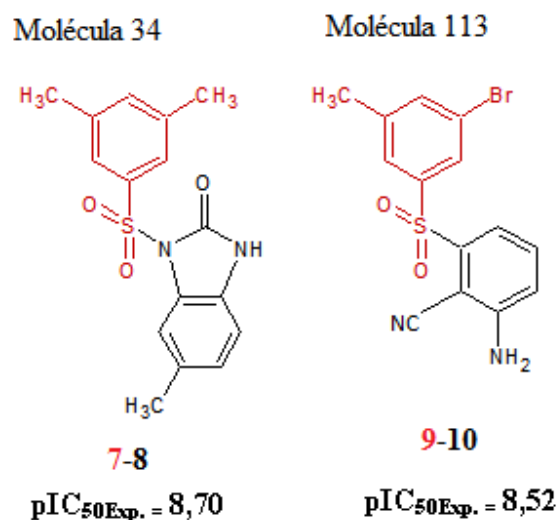


Figura 25 Compostos biologicamente mais ativos, experimentalmente, das séries 2 e 3, que foram combinados (parte em vermelho de uma molécula com a parte preta de outra) para a proposta de novos inibidores com atividade inibitória da TR, conforme modelo MIA-QSAR 2

A combinação das subestruturas 9-8 deu origem ao composto D (Figura 26). Porém, ao combinar as subestruturas 7-10, foi obtida uma estrutura presente no conjunto de dados (molécula 112) e, por esta razão, tal possibilidade foi descartada.

Ao utilizar a técnica de combinar subestruturas de compostos altamente ativos é importante que se mantenha o esqueleto básico comum a todas as estruturas. Assim, semelhante ao que foi descrito em trabalhos anteriores (ANTUNES; FREITAS; RITTNER, 2008; PINHEIRO et al., 2008; SILVA, 2013), foi observado sinergismo por meio do uso de tais combinações, ou seja, o composto B originado a partir deste tipo de metodologia apresentou bioatividade, calculada pelo modelo MIA-QSAR 1 superior ao de seus compostos de origem.

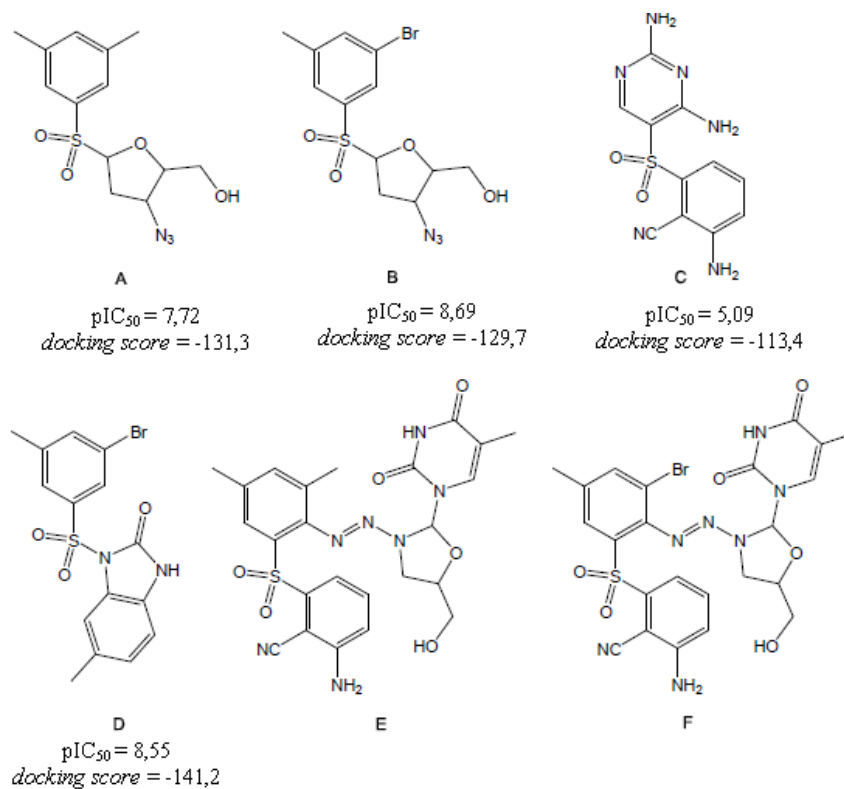


Figura 26 Novos compostos com atividade anti-HIV-1 (A-C), inibidor da TR (D) e compostos propostos com dupla ou mais atividades (E e F). Os compostos A-D tiveram suas atividades biológicas preditas pelos modelos construídos de acordo com suas respectivas atividades exercidas. Contudo, os valores de *docking score* ($kcal\ mol^{-1}$) são referentes apenas à enzima TR

Notavelmente, compostos que apresentam o grupo sulfóxido são os mais ativos (maiores valores de atividade biológica) da série 3, revelando a influência da ligação polar S=O sobre a atividade anti-HIV-1. Em associação, grupos não polares e volumosos no anel benzênico induziram positivamente o efeito das interações hidrofóbicas nas posições 3 e 5, refletindo nos valores de atividades biológicas de tais compostos. Além disso, o substituinte amino (NH_2) pode atuar tanto como receptor de ligações de hidrogênio quanto como doador de prótons,

indicando ser o responsável pela maior atividade do composto 11 da série 1, exceto quando comparado ao composto 1, cuja bioatividade não pode ser assegurada pelo modelo 1, por não ter participado de sua construção (*outlier* excluído do modelo).

O composto A também apresentou elevado valor de pIC_{50} predito (7,72), enquanto C não alcançou atividade considerável ($pIC_{50} = 5,09$). Isto revela que o anel benzênico substituído desempenha importante papel em termos de bioatividade, ao contrário do fragmento de purina, nas séries congêneres de atividade anti-HIV-1.

Apesar de não ter sido observado efeito sinérgico na estrutura D, vale ressaltar a importância de se manter o grupo substituinte 3,5-(CH₃)₂ em vez do átomo de bromo, apesar de sua preferência por derivados 2-aminobenzonitrila. Embora haja semelhança entre os substituintes metila e bromo, por exemplo, em relação à hidrofobicidade de ambos, eles diferem significativamente em outros aspectos que podem afetar a energia de ligação da estrutura (enzima) que os contém.

Além dos compostos A-D, foram propostas as estruturas E e F que, por serem obtidas pela combinação da porção AZT (atividade anti-HIV-1 disponível no mercado para uso clínico) do composto 1 com as estruturas dos inibidores mais ativos da TR (112 e 113), representados na Figura 27, esperam-se, como resultado, novos compostos com múltipla atividade antirretroviral.

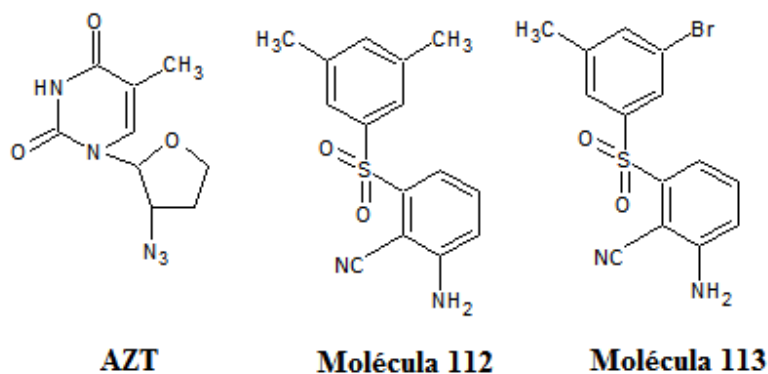


Figura 27 Estruturas químicas do AZT e dos compostos 112 e 113 que, ao serem combinadas, deram origem aos compostos E e F com suposta múltipla atividade antirretroviral

O grupo azido (N₃) do composto AZT foi ligado ao anel benzênico (carbono 2) da molécula 112, dando origem ao composto E. Já a ligação do AZT (pelo mesmo grupo citado anteriormente) à molécula 113 gerou o composto F, representado na Figura 26. Contudo, como o composto 1 (que contém a porção AZT) não participou da calibração do modelo MIA-QSAR 1 por ter sido identificado como um *outlier*, não foi possível calcular as atividades biológicas das estruturas propostas E e F pelos modelos construídos.

O desempenho dos compostos propostos (A-F) em relação às atividades biológicas apresentadas foi avaliado por meio da aplicação de estudos de ancoramento molecular (*docking*), com a finalidade de se identificar a presença de dupla ou mais atividades em uma mesma estrutura química.

5.4 Estudos de *docking* ou ancoramento molecular

Foram obtidas, por meio do PDB, as estruturas cristalinas da TR e das glicoproteínas gp120 e gp41, supostamente relacionadas à atividade anti-HIV-1

(frente às células MT4), representadas pelos respectivos códigos 3C6T, 1G9N e 2OT5.

A partir daí foram realizados estudos de ancoramento molecular para avaliar o encaixe e o alinhamento das estruturas propostas A-F no interior dos sítios ativos da TR, gp120 e gp41, por meio do programa Molegro Virtual Docker® (THOMSEN; CHRISTENSEN, 2006), sendo as sobreposições baseadas no composto AZT disponível comercialmente como um inibidor da HIV-TR 1.

Teoricamente, a porção AZT do composto 1 (que não participou da construção dos modelos MIA-QSAR) pode ser combinada às estruturas dos inibidores mais ativos da TR (112 e 113), como descrito anteriormente, dando origem a potentes compostos (E e F) que deveriam incluir os tipos de atividade biológica que lhes deram origem (anti-HIV-1 e inibitória da TR) e, com isso, as energias correspondentes à ligação desses compostos (E e F) no interior da enzima TR e das glicoproteínas gp120 e gp41 (referentes à atividade anti-HIV-1) poderiam ser calculadas utilizando-se abordagens baseadas na interação ligante-receptor.

Assim, o AZT foi combinado aos compostos 112 e 113 (inibidores da TR mais ativos da série). Porém, ao se realizar os estudos de *docking* para os compostos obtidos por esta combinação (E e F), devido aos superficiais sítios de ligação da gp120 e gp41, não foram possíveis a previsão do encaixe e o alinhamento dos ligantes devido aos altos graus de liberdade das estruturas no interior dos sítios ativos, dificultando a precisão dos resultados e comprometendo sua confiabilidade.

O ancoramento de alguns compostos com dados de atividade anti-HIV-1 experimentais (série 1 e 3) no interior das gp120 e gp41 resultou em energias de ligação (E_{Score}) que não se correlacionam com os valores de suas respectivas bioatividades experimentais, devendo ser descartada esta análise do estudo.

Diferente do que foi relatado em estudos anteriores (CORMANICH; NUNES; FREITAS, 2012), neste caso a referida atividade anti-HIV-1 deve ser descrita por mecanismos mais complexos envolvendo, provavelmente, a inibição de mais de um alvo enzimático e a presença de cofatores.

Diante da impossibilidade do ancoramento das moléculas propostas no interior das gp120 e gp41 para avaliar a atividade anti-HIV-1, os dados de *docking score* apresentados na Figura 26 para os compostos A-D referem-se à interação de tais compostos apenas com a enzima TR, pois todos eles apresentam, em sua estrutura química, porções de inibidores desta enzima.

Os dados obtidos pelo *docking* correlacionam-se razoavelmente bem ($R^2 = 0,684$, Gráfico 6) com os valores de atividade experimental dos compostos selecionados para comparação (34, 72, 87, 112 e 113), apesar dos valores residuais relativamente elevados para 87 e 112. Como o *docking* pode ser utilizado para avaliar as predições feitas por meio de metodologias QSAR e valores de coeficiente de correlação ao quadrado (r^2) acima de 0,5 serem geralmente considerados aceitáveis em validações deste tipo de metodologia, tais abordagens podem ser utilizadas para gerar compreensões confiáveis a respeito da inibição da TR gerada pelos compostos propostos (A-D).

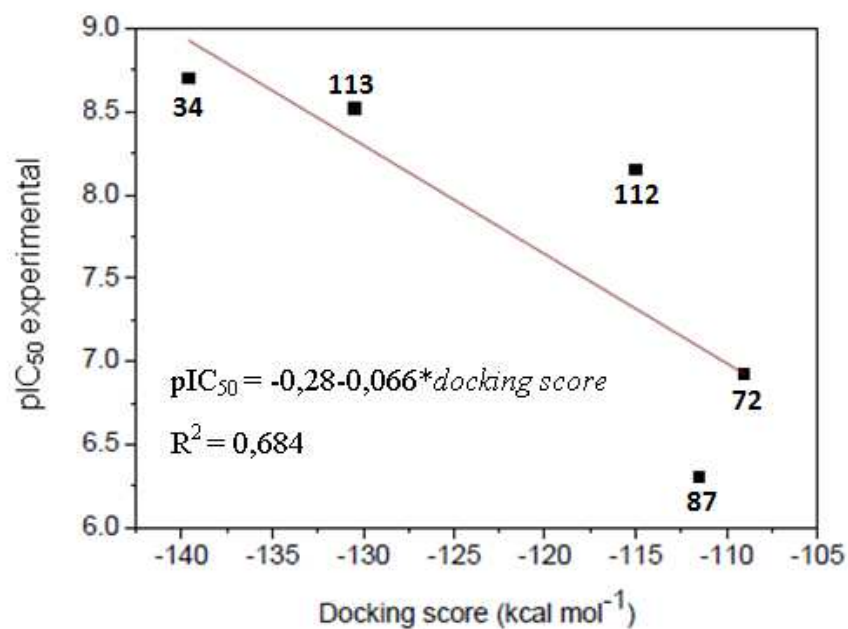


Gráfico 6 Correlação entre valores de pIC_{50} experimentais dos inibidores da TR selecionados e os valores de *docking scores*

O resultado do *docking* correspondente ao composto D foi similar ao do melhor inibidor experimental da TR (composto 34, Tabela 2). Surpreendentemente, o composto B, que se mostrou um promissor candidato anti-HIV-1 (de acordo com o modelo 1), também se mostrou um promissor inibidor da TR com *docking score* de $-129,7\text{ kcal mol}^{-1}$.

Além disso, como mostrado na Figura 28, o composto B apresentou ligação de hidrogênio com os resíduos de aminoácidos Lys-101, Lys-103 e Tyr-318 e interagiu com Tyr-188 através de ligação do tipo π -stacking.

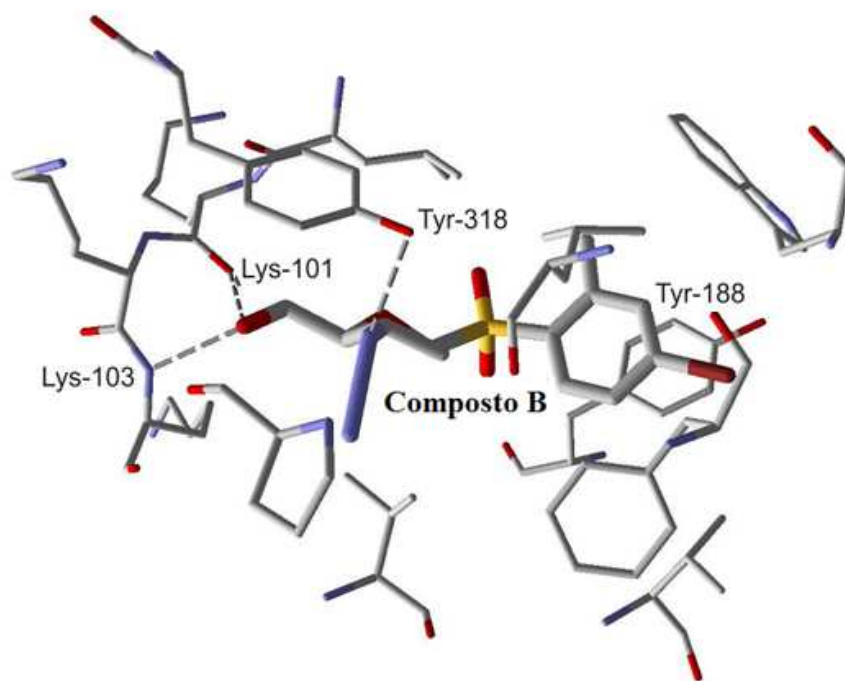


Figura 28 Composto B ancorado no sítio ativo da enzima TR

O composto A, embora tenha apresentado afinidade com a TR ($J = -131,13 \text{ kcal mol}^{-1}$) ao estabelecer interações de hidrogênio com o resíduo de aminoácido Lys-101 e π -stacking com Tyr-188, mostrou-se com potencial de atividade anti-HIV-1 inferior aos seus compostos de origem, ao contrário do apresentado pelo composto B.

5.5 Análises dos parâmetros farmacocinéticos

Os compostos propostos A e B (Figura 26) se mostraram promissores agentes antirretrovirais e, embora pertençam ao modelo 1 (atividade anti-HIV-1), apresentam em suas estruturas químicas porções com atividade experimental

inibitória da TR (derivadas dos compostos 112 e 113), podendo, portanto, atuar em mais de um alvo biológico.

Para ser considerado um bom fármaco são necessários, além de considerável potência, boa absorção e perfil farmacocinético adequado. Os compostos mais ativos 1, 34, 112 e 113, além dos propostos A-F, tiveram suas capacidades de absorção avaliadas por meio do uso do programa Molinspiration (MOLINSPIRATION CHEMINFORMATICS, 1986), que se baseia na aplicação da regra dos cinco de Lipinski e dados de TPSA (provável habilidade de absorção) para avaliar e prever a biodisponibilidade de novos candidatos a fármacos.

Os resultados estão descritos na Tabela 5, na qual é possível perceber que os compostos propostos A-C apresentaram propriedades melhoradas, quando comparados ao composto 1 (mais ativo da série anti-HIV-1). Os compostos A e B não violaram nenhuma das cinco propriedades envolvidas na regra de Lipinski, sendo considerados candidatos promissores, quando comparados aos compostos com melhores atividades experimentais (112 e 113).

Tabela 5 Cálculo dos parâmetros da regra dos cinco de Lipinski e TPSA para avaliação de alguns dos compostos com valores de atividades experimentais e para os compostos propostos (A-F)

Composto	LogP	TPSA	MW	non	noHNH	Violações	Volume
1	-5,223	270,5	544,5	19	7	3	448,2
34	3,529	71,9	336,8	5	1	0	266,8
112	2,747	83,9	286,4	4	2	0	248,1
113	3,107	83,9	351,2	4	2	0	249,5
A	1,767	113,3	311,4	7	1	0	263,7
B	2,128	113,3	376,2	7	1	0	265,0
C	-0,447	161,8	290,3	8	6	1	229,3
D	3,66	71,9	381,2	5	1	0	271,2
E	2,053	196,2	539,6	13	4	2	449,8
F	2,414	196,2	604,4	13	4	2	451,1

Além disso, tais compostos (1, 34, 112, 113, A- F) foram submetidos à avaliação computacional ADME (absorção, distribuição, metabolismo e excreção), que utiliza como parâmetro a regra dos cinco de Lipinski (LIPINSKI et al., 1997) em conjunto com propriedades farmacocinéticas obtidas por meio do programa Hologram QSAR Technique (MODA et al., 2005). De acordo com os dados da Tabela 6, os compostos propostos A-C se mostraram melhores na maioria das propriedades analisadas, tendendo a ser mais bem absorvidos.

Tabela 6 Parâmetros farmacocinéticos computacionais para os compostos com atividades experimentais e para os compostos propostos (A-F)

Composto	HIA (%)	F (%)	PPB (%)	BBB (logBB)	Solubilidade (logS) ^b
1	83,33	-	47,45	-2,66	-3,21
34	93,35	-	94,50	-0,77	-5,73
112	94,30	85,98	86,97	-0,44	-3,85
113	95,27	85,93	83,43	-0,27	-3,92
A	94,68	-	65,66	-1,15	-2,54
B	95,66	-	62,12	-0,98	-2,60
C	100,00	97,88	77,66	-0,82	-2,58
D	100,00	80,36	83,67	-0,78	-5,35
E	78,17	66,41	100,00	-1,78	-5,47
F	79,36	67,81	100,00	-1,67	-5,53

É esperado que D comporte-se de forma muito semelhante aos inibidores da TR experimentalmente disponíveis (34, 112 e 113), uma vez que seus dados farmacocinéticos e da regra dos cinco de Lipinski foram bastante consistentes. Contudo, é importante ressaltar que D não apresentou efeito sinérgico em relação à atividade predita anti-TR-1, ou seja, não apresentou atividade biológica superior à de seus compostos de origem (34 e 113).

Em relação ao perfil físico-químico, farmacocinético e de segurança dos compostos propostos, foi realizada análise computacional utilizando-se a *Plataforma ACD/Percepta* (PLATAFORMA..., 2013). Dentre os compostos analisados (A-F), B se destacou ao apresentar bons resultados físico-químicos, mostrando ser um composto com boa solubilidade e que não viola nenhum dos parâmetros referentes à regra dos cinco de Lipinski, assim como os compostos A e D.

Sobre os parâmetros farmacocinéticos de absorção, distribuição, metabolismo e excreção por transporte passivo, o composto B mostrou-se altamente permeável à membrana das células intestinais Caco-2 e altamente absorvido pelo intestino humano (100% para HIA). A permeabilidade intestinal de candidatos a fármacos, juntamente com a solubilidade, afeta sua biodisponibilidade oral. Compostos com HIA > 70%, como é o caso do composto B, são considerados bem absorvidos quando administrados via oral.

Com relação ao perfil de segurança, B não se comportou como substrato da glicoproteína P (Pgp), que é a responsável pelo transporte de compostos a uma grande variedade de células, sendo este dado clinicamente relevante. Candidatos a fármacos que são substratos ou inibidores da Pgp são considerados indesejáveis, devendo ser excluídos das fases iniciais da concepção de novos medicamentos.

Além disso, B não é inibidor de quatro das cinco enzimas envolvidas no metabolismo de fármacos CYP1A2, CYP2C9, CYP2D6 e CYP3A4, sendo indefinido para CYP2C19. Tais análises colaboram para a detecção precoce de compostos que poderão apresentar problemas metabólicos ou interações medicamentosas relacionadas ao metabolismo dos mesmos. Também se mostrou indefinido para o teste de genotoxicidade Ames, utilizado para avaliar o potencial mutagênico de candidatos a fármacos e não inibidor do canal de potássio hERG, que atesta a segurança do mesmo em relação a uma possível cardiotoxicidade.

6 CONCLUSÃO

A combinação de subestruturas de inibidores da transcriptase reversa (anti-TR-1) e compostos com atividade anti-HIV-1 (sem qualquer relação com os dados inibitórios da TR) proporcionou, como resultado, compostos com altas atividades preditas e bons parâmetros farmacocinéticos calculados, podendo atuar como múltiplos alvos frente ao vírus HIV-1.

A presença do grupo sulfonila (SO₂) e do anel benzênico dissubstituído (3-Br e 5-CH₃) desempenhou importante papel em relação à bioatividade dos compostos propostos.

Dentre eles, o composto B mostrou-se promissor para o exercício da múltipla atividade proposta neste estudo, com destaque para o efeito sinérgico observado para a atividade anti-HIV-1 e bons resultados frente à transcriptase reversa viral, confirmados por estudos de *docking* que identificaram a ocorrência de ligação de hidrogênio e π -*stacking* com resíduos específicos de aminoácidos.

Além disso, ao realizar análises referentes ao perfil físico-químico e farmacocinético dos compostos propostos (A-F), B se destacou em relação à sua biodisponibilidade oral e segurança, apresentando bons resultados em relação a possíveis cardio e genotoxicidade.

Este estudo serve de suporte para o desenvolvimento de novos compostos derivados de B, úteis e, até mesmo, mais potentes como antirretrovirais com duplo ou mais alvos.

REFERÊNCIAS

ADVANCED CHEMISTRY DEVELOPMENT. **ACD/ChemSketch**. Version 12.01. Toronto, 2009. Disponível em: <http://alainsheratan.com/chemsk_t12.pdf>. Acesso em: 10 abr. 2013.

AFANTITIS, A. et al. A novel simple QSAR model for the prediction of anti-HIV activity using multiple linear regression analysis. **Molecular Diversity**, New York, v. 10, n. 3, p. 405-414, Aug. 2006.

AIDS EPIDEMIC UPDATE. **Epidemiological status**. Disponível em: <<http://www.unaids.org/en/dataanalysis/>>. Acesso em: 10 nov. 2013.

ALBUQUERQUE, M. G. et al. Multidimensional-QSAR: beyond the third dimension in drug design. In: TAFT, C. A.; SILVA, C. H. T. P. (Ed.). **Current methods in medicinal chemistry and biological physics**. Trivandrum: Research Signpost, 2007. p. 1-10.

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

AYMARD, G. et al. Determination of twelve antiretroviral agents in human plasma sample using reversed-phase high-performance liquid chromatography. **Journal of Chromatography B**, Amsterdam, v. 744, n. 21, p. 227-240, July 2000.

BARBOSA, J. E. F. **Estudo da atividade citotóxica e antiviral (anti-HSV-1 e anti-HIV-1) dos derivados N-amino-triazóis**. 2009. 108 p. Dissertação (Mestrado em Neuroimunologia) - Universidade Federal Fluminense, Niterói, 2009.

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

BRASIL. Ministério da Saúde. **Portal da saúde: AIDS Brasil**. Brasília, 2009. Disponível em: <<http://www.portal.saude.gov.br>>. Acesso em: 10 abr. 2013.

BRITO, A. M. Fármacos recentes usados para o tratamento da infecção pelo HIV-1: enfuvirtina, maraviroc, raltegravir e etravirina. **Revista de Ciências Farmacêuticas Básica e Avançada**, Araraquara, v. 32, n. 2, p. 159-168, 2011.

BRITO, A. M. et al. 3D-QSAR CoMFA of a series of DABO derivatives as HIV-1 reverse transcriptase inhibitors. **Journal of Chemical Information and Modeling**, Washington, v. 39, n. 8, p. 456-62, Aug. 2008.

BRITO, A. M. et al. AIDS e infecção pelo HIV no Brasil: uma epidemia multifacetada. **Química Nova**, São Paulo, v. 34, n. 2, p. 207-217, mar./abr. 2000.

CAFFARENA, E. **Modelagem molecular: avanços em engenharia de proteínas e síntese de peptídeos**. Rio de Janeiro: FIOCRUZ, 2008. Disponível em: <http://www.ioc.fiocruz.br/peptideos2008/pdf/modelagem_ernesto.pdf>. Acesso em: 16 ago. 2013.

CARVALHO, L. L. de. **Modelagem molecular de uma série de compostos inibidores da enzima integrase do vírus HIV-1**. 2011. 112 p. Tese (Doutorado em Química) Universidade de São Paulo, São Carlos, 2011.

CHEN, K. et al. Quantitative structure-activity relationship studies on 1-aryl-tetrahydroisoquinoline analogs as active anti-HIV agents. **Bioorganic & Medicinal Chemistry Letters**, New York, v. 18, n. 20, p. 5381-5386, Oct. 2008.

CIUFFI, A.; BUSHMAN, F. D. Retroviral DNA integration: HIV and the role of LEDGF/p75. **Trends in Genetics**, London, v. 22, n. 7, p. 388-395, 2006.

COLOMBRINI, M. R. C.; LOPES, M. H. B. M. de; FIGUEIREDO, R. M. de. Adesão à terapia antiretroviral para HIV/AIDS. **Revista da Escola de Enfermagem da USP**, São Paulo, v. 40, n. 4, p. 576-581, 2006.

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

CRAMER, R. D.; PATERSON, D. E.; BUNCE, J. D. Comparative molecular field analysis (CoMFA): I., effect of shape on binding of steroids to carrier proteins. **Journal of the American Chemical Society**, Easton, v. 110, p. 5959-5967, 1988.

CUNICO, W.; GOMES, C. R. B.; VELLASCO JUNIOR, W. T. HIV: recentes avanços na pesquisa de fármacos. **Química Nova**, São Paulo, v. 31, n. 8, p. 2111-2117, 2008.

CUNNINGHAM, A. L. et al. Structure and function of HIV. **Medical Journal of Australia**, Sydney, v. 164, n. 3, p. 161-171, Feb. 1996.

DE CLERCQ, E. Antiretroviral drugs. **Current Opinion in Pharmacology**, London, v. 10, n. 5, p. 507-515, Oct. 2010.

DE CLERCQ, E. Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs): past, present, and future. **Chemistry & Biodiversity**, Weinheim, v. 1, p. 44-64, Jan. 2004.

DEEKS, S. G. Raltegravir. **Nature Reviews Drug Discovery**, London, v. 7, p. 117-119, Feb. 2008.

DHAMI, H. et al. The chemokine system and CCR5 antagonists: potential in HIV treatment and other novel therapies. **Journal of Clinical Pharmacy & Therapeutics**, Oxford, v. 34, n. 2, p. 147-160, Apr. 2009.

DOMS, R. W. Unwelcome guests with master keys: how HIV enters cells and how it can be stopped. **Topics in HIV Medicine**, San Francisco, v. 12, n. 4, p. 100-103, 2004.

DOURADO, I. et al. Tendências da epidemia de AIDS no Brasil após a terapia anti-retroviral. **Revista de Saúde Pública**, Rio de Janeiro, v. 40, p. 9-17, abr. 2006. Suplemento.

FATTORUSSO, C. et al. Specific targeting highly conserved residues in the HIV-1 reverse transcriptase primer grip region: design, synthesis, and biological evaluation of novel, potent, and broad spectrum NNRTIs with antiviral activity. **Journal of Medicinal Chemistry**, Washington, v. 48, n. 23, p. 7153-7165, Nov. 2005.

FAUCI, A. S. The human immunodeficiency virus: infectivity and mechanisms of pathogenesis. **Science**, New York, v. 239, n. 12, p. 617-22, 1988.

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

FERREIRA, R. C. S.; RIFFEL, A.; SANT'ANA, A. E. G. HIV: mecanismo de replicação, alvos farmacológicos e inibição por produtos derivados de plantas. **Química Nova**, São Paulo, v. 33, n. 8, p. 1743-1755, 2010.

FREEMAN, G. A. et al. 2-amino-9-(3-azido-2,3-dideoxy- β -D-erythro-pentofuranosyl)-6-substituted-9H-purines: synthesis and anti-HIV activity. **Bioorganic & Medicinal Chemistry**, Oxford, v. 3, n. 4, p. 447-458, 1995.

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

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, 2005.

GARDNER, M. E. et al. Antiretroviral medication adherence and the development of class-specific antiretroviral resistance. **AIDS**, London, v. 23, n. 9, p. 1035-1046, June 2009.

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

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

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

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

GOOGLE IMAGENS. Disponível em: <<https://www.google.com.br>>. Acesso em: 20 nov. 2013.

GREENE, W. C. et al. Novel targets for HIV therapy. **Antiviral Research**, Amsterdam, v. 80, n. 3, p. 251-265, 2008.

GUIDO, R. V. C.; ANDRICOPULO, A. D.; OLIVA, G. Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas. **Estudos Avançados**, São Paulo, v. 24, n. 70, p. 81-98, 2010.

GUIMARÃES, M. C. et al. Computer-assisted design of dual-target anti-HIV-1 compounds. **Medicinal Chemistry Research**, Cambridge, v. 23, n. 14, p. 1548-1558, 2014.

HAHN, B. H. et al. AIDS as a zoonosis: scientific and public health implications. **Science**, New York, v. 287, n. 5453, p. 607-614, Jan. 2000.

HAMMER, S. M.; YENI, P. Antiretroviral therapy: where are we? **AIDS**, London, v. 12, p. 181-188, 1998. Supplement.

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

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

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

HARMS, G.; FELDMIEIER, H. HIV infection and tropical parasitic diseases: deleterious interactions in both directions? **Tropical Medicine & International Health**, Oxford, v. 7, n. 6, p. 479-488, 2002.

HAUBRICH, R. et al. Etravirine. **Nature Reviews Drug Discovery**, London, v. 7, p. 287-296, 2008.

HILL, A. L.; ROSENBLOOM, D. I. S.; NOWAK, M. A. Evolutionary dynamics of HIV at multiple spatial and temporal scales. **Journal of Molecular Medicine**, Amsterdam, v. 90, n. 5, p. 543-561, May 2012.

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

HORWITZ, J. P. et al. Nucleosides. **Organic Chemistry**, Washington, v. 29, p. 2076-2084, 1964.

HUTCHINSON, J. F. The biology and evolution of HIV. **Annual Review of Anthropology**, Palo Alto, v. 30, p. 85-108, 2001.

JANSSEN, P. A. J. et al. Search of a novel anti-HIV drug: multidisciplinary coordination in the discovery of 4-[[4-[[4-[(1E)-2-Cyanoethenyl]-2,6-dimethylphenyl] amino]pyrimidinyl]amino]benzotrile (R278474, Rilpivirine). **Journal of Medicinal Chemistry**, Washington, v. 48, n. 6, p. 1901-1909, Nov. 2005.

JI, L. et al. Synthesis and anti-HIV activity evaluation of 1-[(alkenyl or alkynyl or alkyloxy)methyl]-5-alkyl-6-(1-naphthoyl)-2,4 pyrimidinediones as novel non-nucleoside HIV-1 reverse transcriptase inhibitors. **European Journal of Medicinal Chemistry**, Paris, v. 42, n. 2, p. 198-204, Feb. 2007.

JOHNSTON, B. E. et al. Adherence to highly active anti-retroviral therapy among HIV: infected patients of the inner city. In: WORLD AIDS CONFERENCE, 12., 1998, Geneva. **Abstracts...** Geneva: WAC, 1998. v. 1, p. 599.

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

KUMAR, S.; SINGH, V.; TIWARI, M. QSAR modeling of the inhibition of reverse transcriptase reverse enzyme with benzimidazolone analogs. **Medicinal Chemistry Research**, Cambridge, v. 20, n. 9, p. 1530-1541, Dec. 2011.

KURITZKES, D.; KAR, S.; KIRKPATRICK, P. Maraviroc. **Nature Reviews Drug Discovery**, London, v. 7, n. 1, p. 15-17, Jan. 2008.

LABONTE, J.; LEBBOS, J.; KIRKPATRICK, P. Enfuvirtide. **Nature Reviews Drug Discovery**, London, v. 2, n. 5, p. 345-347, May 2003.

LALEZARI, J. P. Enfuvirtide, an HIV-1 fusion inhibitor for drug-resistant HIV infection in North and South America. **New England Journal of Medicine**, Boston, v. 348, n. 22, p. 2175-2185, May 2003.

LEONARD, J. T.; ROY, K. Classical QSAR modeling of HIV-1 reverse transcriptase inhibitor 2-amino-6-arylsulfonylbenzotrioles and congeners. **QSAR & Combinatorial Science**, Weinheim, v. 23, n. 1, p. 23-35, Feb. 2004.

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.

LIPINSKI, C. A. et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews**, London, v. 23, n. 1/3, p. 3-25, 1997.

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

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

MARTIN, Y. C. Let's not forget tautomers. **Journal of Computer-Aided Molecular Design**, Dordrecht, v. 23, n. 10, p. 693-704, Oct. 2009.

MARTINS, J. P. A. **Química computacional aplicada a QSAR**. São Paulo: CENAPAD, 2010. Disponível em: <http://www.cenapad.unicamp.br/servicos/treinamentos/apostilas/apostila_QuimComp.pdf>. Acesso em: 10 abr. 2013.

MARTINS, J. P. A.; FERREIRA, M. M. C. *QSAR Modelling*: um novo pacote computacional *open source* para gerar e validar modelos QSAR. **Química Nova**, São Paulo, v. 36, n. 4, p. 554-560, mar. 2013.

MATHWORKS. **MatLab**. Version 7.5. Natick, 2007. Disponível em: <<http://www.mathworks.com/products/matlab/whatsnew.html>>. Acesso em: 10 abr. 2013.

MELO, E. B. de. Inibidores da HIV-Integrase: potencial abordagem farmacológica para tratamento da AIDS. **Química Nova**, São Paulo, v. 29, n. 3, p. 555-562, mar. 2006.

MELO, F. L. de. **Caracterização biológica e molecular de recombinantes naturais de HIV-1**. 2011. 104 p. Tese (Doutorado em Microbiologia) - Universidade de São Paulo, São Paulo, 2011.

MIRANDA, A. S. de et al. Maraviroque: uma inovação terapêutica para o tratamento da AIDS. **Revista Virtual de Química**, Niterói, v. 2, n. 2, p. 130-139, 2010.

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

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

MOLINSPIRATION CHEMINFORMATICS. **Molinspiration**. New Ulica, 1986. Disponível em: <<http://www.molinspiration.com/>>. Acesso em: 20 out. 2013.

MONTANARI, C. A. **Química medicinal: métodos e fundamentos em planejamento de fármacos**. São Paulo: EDUSP, 2011. 710 p.

NASH, D.; EIUL, B. The impact of HIV on cities in the era of treatment: a global perspective. **Technology in Society**, Philadelphia, v. 28, n. 1/2, p. 195-223, Jan./Apr. 2006.

NEVES, J. P. et al. TOP: um programa de cálculo de descritores topológicos para uso em correlações entre estrutura e atividade. **Química Nova**, São Paulo, v. 21, n. 6, p. 109-113, nov./dez. 1998.

NUNES, C. A. et al. Chemoface: a novel free user-friendly interface for chemometrics. **Journal of the Brazilian Chemical Society**, São Paulo, v. 23, n. 11, p. 2003-2010, 2012.

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

PAN, D.; TSENG, Y.; HOPFINGER, A. J. Quantitative structure-based design: formalism and application of receptor-dependent RD-4D-QSAR analysis to a set of glucose analogue inhibitors of glycogen phosphorylase. **Journal of Chemical Information and Computer Sciences**, Washington, v. 43, n. 5, p. 1591-1607, 2003.

PATERSON, D. L. et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. **Annals of Internal Medicine**, Philadelphia, v. 133, n. 1, p. 21-30, 2000.

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

PEREIRA, H. S. et al. Inibição seletiva das atividades das enzimas transcriptase reversa do vírus HIV-1 e DNA polimerases humanas por derivados dipirazolo-piridina. **Jornal Brasileiro de Doenças Sexualmente Transmissíveis**, Rio de Janeiro, v. 8, n. 4, p. 14-18, 1996.

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

PLATAFORMA ACD/Percepta. Disponível em: <<http://www.acdlabs.com>>. Acesso em: 6 out. 2013.

POLANSKI, J. Receptor dependent multidimensional QSAR for modeling drug-receptor interactions. **Current Medicinal Chemistry**, Wageningen, v. 16, n. 25, p. 3243-3257, Sept. 2009.

POLEJACK, L.; SEIDL, E. M. F. Monitoramento e avaliação da adesão ao tratamento antirretroviral para HIV/AIDS: desafios e possibilidades. **Ciência e Saúde Coletiva**, Santa Maria, v. 15, n. 10, p. 1201-1208, 2010.

POMMIER, Y.; JOHNSON, A. A.; MARCHAND, C. Integrase inhibitors to treat HIV/AIDS. **Nature Reviews Drug Discovery**, London, v. 4, p. 236-248, Mar. 2005.

RAFFANTI, S.; HAAS, D. W. **Goodman & Gilman as bases farmacológicas da terapêutica**. 11. ed. Rio de Janeiro: McGraw-Hill, 2006. 1671 p.

RAUPP, D.; SERRANO, A.; MARTINS, T. L. C. A evolução da química computacional e sua contribuição para a educação em química. **Revista Liberato**, Nova Hamburgo, v. 9, n. 12, p. 13-22, 2008.

RAVICHANDRAN, V.; AGRAWAL, R. K. Predicting anti-HIV activity of PETT derivatives: CoMFA approach. **Bioorganic & Medicinal Chemistry Letters**, New York, v. 17, n. 8, p. 2197-2202, Apr. 2007.

RIO, C. del; SOTO-RAMIREZ, E. L. E. Bridging the gap. **AIDS Clinical Care**, Waltham, v. 4, p. 1-5, 1991.

ROCHE, M. et al. A common mechanism of clinical HIV-1 resistance of the CCR5 antagonist maraviroc despite divergent resistance levels and lack of common gp120 resistance mutations. **Retrovirology**, London, v. 10, n. 43, p. 1-20, 2013.

RODRIGUES, C. R. Processos modernos no desenvolvimento de fármacos: modelagem molecular. **Cadernos Temáticos de Química Nova na Escola**, São Paulo, v. 3, n. 3, p. 43-50, maio 2001.

ROY, K. et al. Comparative studies on some metrics for external validation of QSPR models. **Journal of Chemical Information and Modeling**, Washington, v. 52, n. 2, p. 346-408, Feb. 2012.

ROY, K.; LEONARD, J. T. QSAR by LFER model of cytotoxicity data of anti-HIV 5-phenyl-1-phenylamino-1H-imidazolone derivatives using principal component factor analysis and genetic function approximation. **Bioorganic & Medicinal Chemistry**, Oxford, v. 13, n. 7, p. 2967-2973, Apr. 2005.

ROY, P. P. et al. On two novel parameters for validations of predictive QSAR models. **Molecules**, Basel, v. 15, p. 604-605, Jan. 2010.

SABIN, J. G.; FERRÃO, M. F.; FURTADO, J. C. Análise multivariada aplicada na identificação de fármacos antidepressivos parte II: análise por componentes principais (PCA) e o método de classificação SIMCA. **Revista Brasileira de Ciências Farmacêuticas**, São Paulo, v. 40, n. 3, p. 387-396, jul./set. 2004.

SALDANHA, J. S.; ANDRADE, C. S.; BECK, S. T. Grau de adesão ao tratamento com anti-retrovirais entre indivíduos HIV positivos atendidos no hospital universitário de Santa Maria. **Saúde**, Santa Maria, v. 35, n. 1, p. 4-9, 2009.

SHEN, L. et al. Steered molecular dynamics simulation on the binding of nrti to HIV-1 RT. **Biophysical Journal**, New York, v. 84, n. 6, p. 3547-3563, June 2003.

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. 33-244, 2005.

SILVA, D. G. **Modelagem racional de novos dipeptídeos modificados como inibidores da enzima conversora de angiotensina**. 2013. 121 p. Dissertação (Mestrado em Agroquímica) - Universidade Federal de Lavras, Lavras, 2013.

SOUZA, J. de; STORPIRTIS, S. Atividade anti-retroviral e propriedades farmacocinéticas da associação entre lamivudina e zidovudina. **Revista Brasileira de Ciências Farmacêuticas**, São Paulo, v. 40, n. 1, p. 9-19, jan./mar. 2004.

SOUZA, M. V. D. de. Fármacos inibidores de fusão: uma nova estratégia no combate à replicação do vírus HIV. **Acta Farmacéutica Bonaerense**, Buenos Aires, v. 24, n. 2, p. 291-300, 2005.

SOUZA, T. C. S. 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, 2009.

STYRT, B. A.; PIAZZA-HEPP, T. D.; CHIKAMI, G. K. Clinical toxicity of antiretroviral nucleoside analogs. **Antiviral Research**, Amsterdam, v. 31, n. 3, p. 121-135, July 1996.

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

THOMSEN, R.; CHISTENSEN, M. H. MolDock: a new technique for high-accuracy molecular docking. **Journal of Medicinal Chemistry**, Washington, v. 49, n. 11, p. 3315-3321, Apr. 2006.

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

TSIBRIS, A. M.; KURITZKES, D. R. Chemokine antagonists as therapeutics: focus on HIV-1. **Annual Review of Medicine**, Palo Alto, v. 58, p. 445-459, 2007.

VAISHNAV, Y. N.; WONGSTAAL, F. The biochemistry of AIDS. **Annual Review of 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**, Washington, v. 45, n. 11, p. 2139-2149, May 2002.

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

VERMEIRE, K. et al. The anti-HIV potency of cyclotriazadisulfonamide analogs is directly correlated with their ability to down-modulate CD4 receptor. **Molecular Pharmacology**, New York, v. 63, n. 1, p. 203-210, 2003.

VILAR, S.; SANTANA, L.; URIARTE, E. Probabilistic neural network model for the in silico evaluation of anti-HIV activity and mechanism of action. **Journal of Medicinal Chemistry**, Washington, v. 49, n. 3, p. 1118-1124, Jan. 2006.

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

WEBER, K. C. **Modelagem molecular de compostos arilpiperazínicos e suas interações com o receptor 5-HT_{1a}**. 2008. 138 p. Tese (Doutorado em Química) - Universidade de São Paulo, São Carlos, 2008.

WEITMAN, M. et al. Structure-activity relationship studies of 1-(4-chloro-2,5-dimethoxyphenyl)-3-(3-propoxypropyl)thiourea, a non-nucleoside reverse transcriptase inhibitor of human immunodeficiency virus type-1. **European Journal of Medicinal Chemistry**, Paris, v. 46, n. 2, p. 447-467, Feb. 2011.

WERMUTH, C. G. **The practice of medicinal chemistry**. New York: Academic, 1996. 968 p.

WILLIAMS, D. A.; LEMKE, T. L. **Foye's principles of medicinal chemistry**. 5th ed. Philadelphia: Lippincott Williams & Wilkins, 2002. 1114 p.

ZUCKERMAN, M.; PILLAY, D. HIV diagnosis and antiretroviral therapy. **Medicine**, London, v. 36, n. 3, p. 172-174, Mar. 2007.