UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

LEONARDO LUIZ GOMES FERREIRA

Identificação de novos inibidores da enzima aldolase de *Trypanosoma brucei*

São Carlos
2013

LEONARDO LUIZ GOMES FERREIRA

Identificação de novos inibidores da enzima aldolase de *Trypanosoma brucei*

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos, da Universidade de São Paulo, para a obtenção do título de Doutor em Ciências.

Área de concentração: Física Aplicada

Opção: Física Biomolecular

Orientador: Prof. Dr. Adriano Defini Andricopulo

Coorientador: Prof. Dr. Otavio Henrique Thiemann

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

São Carlos

2013

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação do IFSC/USP

Ferreira, Leonardo Luiz Gomes

Identificação de novos inibidores da enzima aldolase de *Trypanosoma brucei* / Leonardo Luiz Gomes Ferreira; orientador Adriano Defini Andricopulo; co-orientador Otavio Henrique Thiemann - versão corrigida -- São Carlos, 2013.
164 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2013.

1. *Trypanosoma brucei*. 2. Aldolase. 3. Triagem virtual. 4. QSAR. 5. Dinâmica molecular. I. Andricopulo, Adriano Defini, orient. II. Thiemann, Otavio Henrique, co-orient. III. Título.

AGRADECIMENTOS

A minha família, Juraci Aparecida da Rocha, Luciana Correa Vellardi e Melissa Vellardi Ferreira, pelo apoio e incentivo.

Ao Prof. Dr. Adriano Defini Andricopulo pelas oportunidades e orientação neste trabalho.

Ao Prof. Dr. Rafael V. C. Guido e a Prof. Dra. Rafaela S. Ferreira pelas contribuições.

Aos colegas do Laboratório de Química Medicinal e Computacional, Ricardo N. dos Santos, Wanessa F. Altei, Ana Isabela L. Sales, Mariana L. de Souza, Ivani Pauli, Renata K. Andricopulo, Simone M. Duarte e Fernando V. Maluf.

Aos professores e funcionários do Grupo de Cristalografia de Proteínas e Biologia Estrutural.

A FAPESP pelo apoio financeiro.

RESUMO

FERREIRA, L. L. G. **Identificação de novos inibidores da enzima aldolase de *Trypanosoma brucei***. 2013. 164 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

As doenças tropicais negligenciadas, que atingem as populações mais carentes do mundo, representam em termos humanitários e socioeconômicos uma grande preocupação global. As tripanossomíases estão entre as doenças parasitárias mais importantes, e, particularmente, a tripanossomíase africana, ou doença do sono, destaca-se como uma grave condição de saúde, causada pelo parasita unicelular *Trypanosoma brucei*. Dentre os principais alvos metabólicos considerados para o desenvolvimento de novos fármacos para o tratamento das tripanossomíases, a glicólise recebe especial atenção em função de seu papel vital no processo de produção de ATP para o parasita que vive na corrente sanguínea. Esta tese de doutorado tem como objetivo identificar novos candidatos a inibidores da enzima aldolase (EC 4.1.2.13) da via glicolítica de *T. brucei*. Considerando-se que o alvo macromolecular em questão é validado para o planejamento de fármacos, inibidores desta enzima são candidatos a novos agentes quimioterápicos. Este trabalho explora a integração de métodos experimentais e computacionais através de estratégias de planejamento de fármacos baseado na estrutura do receptor (SBDD, na sigla inglesa para *structure-based drug design*) e na estrutura do ligante (LBDD, na sigla inglesa para *ligand-based drug design*) para a identificação de inibidores da enzima alvo. Foram produzidos resultados significativos, tais como a identificação através de triagens virtuais em larga escala de novas moléculas capazes de inibir a atividade da aldolase. Adicionalmente, destaca-se a obtenção de protocolos de expressão, purificação e cristalização para a enzima alvo. Como parte da estratégia de identificação de novos inibidores da aldolase, foram desenvolvidos modelos de QSAR 2D e 3D e estudos de dinâmica molecular.

Palavras-chave: *Trypanosoma brucei*. Aldolase. Triagem virtual. QSAR. Dinâmica molecular.

ABSTRACT

FERREIRA, L. L. G. **Identification of novel inhibitors of aldolase from *Trypanosoma brucei***. 2013. 164 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Neglected tropical diseases, which affect the poorest populations across the developing world, are a major global concern. The trypanosomiasis are amongst the most serious neglected tropical diseases, and particularly, African trypanosomiasis (sleeping sickness), caused by the unicellular parasite *Trypanosoma brucei*, appears as a fatal condition. The glycolytic pathway emerges as a promising target among the metabolic pathways for the development of new drugs, due to its essential role in the ATP generating process in the bloodstream form of the parasite. The goal of this work is to identify new inhibitors for the glycolytic enzyme aldolase (EC 4.1.2.13) from *Trypanosoma brucei*. Inhibitors of this enzyme are drug candidates with high potential for clinical development, as the respective target enzyme was validated as a molecular target for the therapy of trypanosomiasis. The strategy employed in this study includes the integration of SBDD (*structure-based drug design*) and LBDD, (*ligand-based drug design*) for the identification of inhibitors of the target enzyme, through the combination of computational and experimental methodologies. Significant results were obtained, such as the identification of new small molecule inhibitors of the aldolase enzyme through high-throughput virtual screening. Additionally, it is highlighted the standardization of expression, purification and crystallization protocols for the target enzyme. As a component of the strategy for the identification of novel aldolase inhibitors, 2D and 3D QSAR models were developed, as well as molecular dynamics studies.

Keywords: *Trypanosoma brucei*. Aldolase. Virtual screening. QSAR. Molecular dynamics.

LISTA DE FIGURAS

Figura 1 –	Esquema geral de um processo típico de P&D de fármacos na indústria farmacêutica.	31
Figura 2 –	Ciclo de vida do causador da TAH, o parasita <i>Trypanosoma brucei</i> . Estágios de desenvolvimento no hospedeiro intermediário, a mosca tsé-tsé, e no hospedeiro definitivo, o ser humano.	36
Figura 3 –	Distribuição geográfica da TAH causada pela subespécie <i>T. b. gambiense</i>	37
Figura 4 –	Distribuição geográfica da TAH causada pela subespécie <i>T. b. rhodesiense</i> .	38
Figura 5 –	Estrutura química dos fármacos utilizados no tratamento da tripanossomíase africana.	39
Figura 6 –	Glicólise em <i>T. brucei</i> . A numeração corresponde às reações catalisadas pelas seguintes enzimas: (1) hexoquinase; (2) fosfoglicose isomerase; (3) fosfofrutoquinase; (4) aldolase; (5) triosefosfato isomerase; (6) gliceraldeído-3-fosfato desidrogenase; (7) fosfoglicerato quinase; (8) glicerol-3-fosfato desidrogenase; (9) glicerol quinase; (10) fosfoglicerato mutase; (11) enolase; (12) piruvato quinase.	45
Figura 7 –	Reação catalisada pela aldolase: clivagem reversível da frutose-bifosfato em diidroxiacetona-fosfato e gliceraldeído-3-fosfato.	48
Figura 8 –	Esquema simplificado da reação catalisada pela aldolase de <i>T. brucei</i> .	49
Figura 9 –	Estrutura tetramérica da aldolase de <i>T. brucei</i> .	50
Figura 10 –	Estrutura do monômero da aldolase de <i>T. brucei</i> , com destaque para o sítio ativo que está localizado na face interna barril β (amarelo).	51
Figura 11 –	Esquema do processo de planejamento baseado na estrutura do receptor (SBDD).	60
Figura 12 –	Esquema do processo de docagem molecular.	61

Figura 13 –	Esquema do processo de triagem virtual baseada na estrutura do receptor (SBVS).	64
Figura 14 –	Esquema do processo de dinâmica molecular.	67
Figura 15 –	Geração de um modelo HQSAR.	72
Figura 16 –	Geração de um modelo CoMFA.	74
Figura 17 –	Purificação de uma proteína recombinante pelo método IMAC.	79
Figura 18 –	Velocidade inicial de uma reação enzimática em função da concentração de substrato.	82
Figura 19 –	Esquema geral das reações enzimáticas.	82
Figura 20 –	Esquema do ensaio acoplado utilizado para monitorar a velocidade da reação da aldolase. FBP (frutose-bifosfato); GAP (gliceraldeído-3-fosfato); DHAP (diidroxiacetona-fosfato); G3P (glicerol-3-fosfato). A numeração corresponde às reações catalisadas pela aldolase (1); triosefosfato-isomerase (2) e glicerol-3-fosfato desidrogenase (3).	85
Figura 21–	Gel de poliacrilamida, demonstrando a progressão da purificação. A primeira coluna da direita contém a banda correspondente à aldolase.	91
Figura 22 –	Absorbância durante a gel filtração.	93
Figura 23 –	Cristais obtidos para a aldolase de <i>T. brucei</i> .	94
Figura 24 –	Padrão de difração dos cristais da aldolase de <i>T. brucei</i> .	95
Figura 25 –	Gráfico da atividade enzimática em função da concentração de enzima.	96
Figura 26 –	Gráfico da velocidade da reação em função da concentração de substrato.	97
Figura 27 –	Gráfico duplo-recíproco gerado a partir da transformação linear dos dados da velocidade em função de [S].	98

Figura 28 –	Medida cinética da atividade da aldolase realizada através da detecção da queda na concentração do cofator T-NADH, nas condições padrão do ensaio.	99
Figura 29 –	Subestrutura representativa das moléculas do conjunto de dados utilizado nas análises de QSAR.	104
Figura 30 –	Alinhamento da sequência de aminoácidos das enzimas de <i>T. brucei</i> , músculo de coelho e humana. A numeração corresponde à posição dos aminoácidos na enzima do parasita. Os aminoácidos conservados estão destacados em preto.	107
Figura 31 –	Alinhamento estrutural das enzimas de <i>T. brucei</i> (cinza), músculo de coelho (amarelo) e humana (verde).	108
Figura 32 –	Sobreposição dos resíduos de aminoácidos do sítio ativo da aldolase de coelho (amarelo) e de <i>T. brucei</i> (cinza). O inibidor manitol-bifosfato está representado em azul.	109
Figura 33 –	Conformação cristalográfica do inibidor MBP (azul) em complexo com a aldolase de coelho (amarelo). Solução de docagem para o mesmo inibidor (magenta) no sítio ativo da aldolase de <i>T. brucei</i> (cinza).	110
Figura 34 –	Distribuição da propriedade biológica nos conjuntos treinamento e teste.	111
Figura 35 –	Valores de pK_i experimentais <i>versus</i> os valores preditos e estimados para a série treinamento e série teste, obtidos para o modelo A/B/C/H/Ch/DA, 6–9 átomos (Tabela 11).	115
Figura 36 –	Mapas de contribuição para os compostos 2, 3 e 4. Os mapas mostram que a contribuição da cadeia hidrocarbônica cresce à medida que o número de átomos de carbono se aproxima de seis.	116
Figura 37 –	Mapas de contribuição para os compostos 6 (carbonila em amarelo) e 10 (carbonila em verde).	117
Figura 38 –	Conformação dos compostos 6 (A) e 10 (B) , com destaque para as interações polares entre o grupo éster (A) e o grupo amida (B) com os resíduos de lisina 156 e 239 do sítio ativo da aldolase.	118
Figura 39 –	Alinhamento molecular dos inibidores, utilizado na construção dos modelos CoMFA.	120
Figura 40 –	Gráfico dos valores experimentais <i>versus</i> os valores preditos e estimados de pK_i para a série treinamento (azul) e para a série teste (vermelho).	123
Figura 41 –	(A) Mapas de contorno CoMFA (modelo 3) para o inibidor 10 e (B) para o inibidor 23.	124

Figura 42 –	(A) Conformação do composto 4 no sítio ativo da aldolase, com destaque para as interações eletrostáticas entre os grupos fosfato do inibidor e os resíduos do sítio ativo. (B) Conformação do composto 10, que demonstra o mesmo padrão de interação, exceto pela presença de uma interação adicional entre o grupo amida do ligante e os resíduos de lisina 156 e 239.	127
Figura 43 –	(A) RMSD da aldolase na presença dos compostos 4 (linha azul) e 10 (linha laranja) durante a simulação. (B) RMSD dos compostos 4 e 10. (C) Flutuações médias de cada resíduo de aminoácido da aldolase durante toda a simulação, para o complexo com os inibidores 4 e 10, e para a apoenzima. Cada linha representa a média de cinco simulações independentes.	129
Figura 44 –	Distância das interações intermoleculares entre os ligantes e os resíduos do sítio ativo durante as simulações.	130
Figura 45 –	Sítio ativo da aldolase de <i>T. brucei</i> (cinza) sobreposto à enzima humana (verde). No centro o inibidor 6 (carbonos em azul). As diferenças estruturais estão destacadas com as siglas dos respectivos aminoácidos.	133
Figura 46 –	Estratégia de triagem virtual adotada, baseada no método <i>tiered approach</i> .	134
Figura 47 –	Diferença conformacional do resíduo de arginina 313 no sítio ativo das enzimas de <i>T. brucei</i> (cinza) e <i>L. mexicana</i> (magenta). O substrato (verde) realiza diversas interações polares com o sítio ativo da enzima de <i>L. mexicana</i> .	135
Figura 48 –	Grupos funcionais carregados negativamente encontrados em interações com resíduos de alanina, serina, arginina e lisina (grupos carboxilato, sulfonato e fosfato).	136
Figura 49 –	Estrutura, conformação e interações intermoleculares de duas moléculas selecionadas na triagem virtual.	138
Figura 50 –	Conformação de inibidores representativos da série 2,7-naftaleno, obtidas por docagem molecular e as respectivas interações intermoleculares: (A) inibidor 11; (B) 7; (C) 13 e (D) 3.	145
Figura 51 –	Conformação do inibidor 6. O anel isobenzofuranona forma uma ligação cátion- π com a arginina 313.	147

LISTA DE TABELAS

Tabela 1 –	Doenças tropicais negligenciadas e seus agentes causadores.	32
Tabela 2 –	Características dos fármacos disponíveis para o tratamento TAH.	41
Tabela 3 –	Vias metabólicas e alvos moleculares validados para o desenvolvimento de novos fármacos para a TAH.	47
Tabela 4 –	Parâmetros de distinção de fragmentos no método HQSAR.	71
Tabela 5 –	Composição do tampão de lise.	90
Tabela 6 –	Etapas da purificação e seus respectivos valores de atividade enzimática, atividade específica e fator de purificação.	92
Tabela 7 –	Condições favoráveis para o crescimento de cristais da aldolase.	93
Tabela 8 –	Condições padrão estabelecidas para os ensaios de cinética e inibição enzimática.	98
Tabela 9 –	Estrutura do conjunto de dados utilizado nos estudos de QSAR e os valores de pK_i .	105
Tabela 10 –	Resultados para os diferentes parâmetros de distinção de fragmentos, utilizando o tamanho de fragmento padrão de 4 a 7 átomos e o método de validação LOO.	112
Tabela 11 –	Resultados gerados para os diferentes tamanhos de fragmento.	113
Tabela 12 –	Resultados obtidos com o método de validação LMO para o modelo A/B/C/H/Ch/DA, tamanho de fragmento de 6 à 9 átomos.	113
Tabela 13 –	Valores experimentais e preditos de pK_i com os valores residuais para o conjunto teste	114
Tabela 14 –	Resultados estatísticos para os modelos CoMFA, obtidos em função da variação nos valores de SDC.	121
Tabela 15 –	Valores experimentais e preditos de pK_i e dos resíduos para a série teste.	122

Tabela 16 – Percentual de inibição dos compostos testados frente à enzima aldolase de *T. brucei*. 141

LISTA DE ABREVIATURAS E SIGLAS

FDA	<i>Food and Drug Administration</i>
IUPAC	<i>International Union of Pure and Applied Chemistry</i>
ADME/Tox	Absorção, distribuição, metabolismo, excreção e toxicidade
SBDD	Planejamento baseado na estrutura do receptor (<i>Structure-based drug design</i>)
VS	Triagem virtual (<i>Virtual screening</i>)
LBDD	Planejamento baseado na estrutura do ligante (<i>Ligand-based drug design</i>)
QSAR	Relações quantitativas entre estrutura e atividade (<i>Quantitative structure-activity relationships</i>)
P&D	Pesquisa e desenvolvimento
HTS	Triagem em larga escala (<i>High throughput screening</i>)
LO	Otimização de compostos-líder (<i>Lead optimization</i>)
NCE	Nova entidade química (<i>New chemical entity</i>)
ANVISA	Agência Nacional de Vigilância Sanitária
EMA	<i>European Medicines Agency</i>
IND	Candidato a novo fármaco (<i>Investigational new drug</i>)
NDA	Novo fármaco (<i>New drug application</i>)
DTN	Doenças tropicais negligenciadas
OMS	Organização Mundial de Saúde
DALY	Anos de vida produtiva perdidos por morte prematura e incapacitação (<i>Disability adjusted life year</i>)
DNDi	<i>Drugs for Neglected Diseases initiative</i>
UNICEF	<i>United Nations Children's Fund</i>
UNDP	<i>United Nations Development Programme</i>

TDR	<i>Special Programme for Research and Training in Tropical Diseases</i>
TAH	Tripanossomíase Africana Humana
<i>T. brucei</i>	<i>Trypanosoma brucei</i>
<i>T. b. rhodesiense</i>	<i>Trypanosoma brucei rhodesiense</i>
<i>T. b. gambiense</i>	<i>Trypanosoma brucei gambiense</i>
RNAi	RNA de interferência
ATP	Adenosina-trifosfato
ADP	Adenosina-difosfato
NADH	Forma reduzida da nicotinamida adenina dinucleotídeo
NAD ⁺	Forma oxidada da nicotinamida adenina dinucleotídeo
DHAP	Diidroxiacetona-fosfato
GAP	Gliceraldeído-3-fosfato
FBP	Frutose-1,6-bifosfato
EC	<i>Enzyme Comission</i>
PDB	<i>Protein Data Bank</i>
HQSAR	Holograma QSAR (<i>Hologram QSAR</i>)
CoMFA	Análise Comparativa dos Campos Moleculares (<i>Comparative molecular field analysis</i>)
UCSF	University of California San Francisco
LBVS	Triagem virtual baseada na estrutura do ligante (<i>Ligand-based virtual screening</i>)
SBVS	Triagem virtual baseada na estrutura do receptor (<i>Structure-based virtual screening</i>)
3D	Tridimensional
2D	Bidimensional
ZINC	Banco de dados <i>ZINC is not commercial</i>

DM	Dinâmica molecular
CHARMM	<i>Chemistry at Harvard Molecular Mechanics</i>
NAMD	<i>Not Another Molecular Dynamics Program</i>
MLR	Regressão linear múltipla (<i>Multiple linear regression</i>)
PLS	Mínimos quadrados parciais (<i>Partial least squares</i>)
A	Átomo
B	Ligações
C	Conectividade
H	Hidrogênios
Ch	Quiralidade
DA	Doador e receptor
IMAC	Cromatografia de afinidade por metal imobilizado (<i>Immobilized metal affinity chromatography</i>)
pET28a	Vetor de expressão pET28a
TROP	<i>Research Unit for Tropical Diseases</i>
MPD	2-metil-2,4-pentanodiol
PEG	Polietilenoglicol
E	Enzima
S	Substrato
P	Produto
$V_{\max}$	Velocidade máxima
ES	Complexo enzima-substrato
EP	Complexo enzima-produto
V_0	Velocidade inicial
K_M	Constante de Michaelis-Menten

V	Velocidade
Abs	Absorbância
ϵ	Absortividade molar
l	Caminho ótico
c	Concentração
TIM	Triose-fosfato isomerase
GDH	Glicerol-3-fosfato desidrogenase
T-NADH	Forma reduzida da tionicotinamida adenina dinocleotídeo
T-NAD ⁺	Forma oxidada da tionicotinamida adenina dinocleotídeo
P	Grupo fosfato
<i>E. coli</i>	<i>Escherichia coli</i>
LB	Luria Bertani
IPTG	Isopropil- β ,D-tiogalactopiranosídeo
TCEP	Tris(2-carboxietil)fosfina
PMSF	Fluoreto de fenilmetilsulfonila
SDS-PAGE	Eletroforese em gel de poliacrilamida e dodecil-sulfato de sódio
TRIS	2-amino-2-hidroximetil-propano-1,3-diol
BIS-TRIS	2-bis(2-hidroxietil)amino-2-(hidroximetil)-1,3-propano-diol
AM1	Método Austin 1 (<i>Austin method 1</i>)
MBP	Manitol-bifosfato
LOO	Método deixe um de fora (<i>Leave-one-out</i>)
LMO	Método deixe vários de fora (<i>Leave-many-out</i>)
q^2	Coeficiente de correlação com validação cruzada
r^2	Coeficiente de correlação sem validação cruzada

SEE	Erro padrão de estimativa (<i>Standard error of estimate</i>)
HL	Comprimento do holograma (<i>Hologram length</i>)
N	Número ótimo de componentes
SDC	Coeficiente do desvio padrão (<i>Standard deviation coefficient</i>)
SEP	Erro padrão de predição (<i>Standard error of prediction</i>)
PME	Rede de partículas de Ewald (<i>Particle Mesh Ewald</i>)
RMSD	Desvio médio da raiz quadrada (<i>Root mean square deviation</i>)
RMSF	Flutuação média da raiz quadrada (<i>Root mean square fluctuation</i>)
logP	Logarítmo do coeficiente de partição octanol/água
Arg	Arginina
Lys	Lisina
Ser	Serina
Glu	Glutamato
His	Histidina
Thr	Treonina
Phe	Fenilalanina
Ala	Alanina
Gly	Glicina
w	Água

SUMÁRIO

Capítulo 1. Introdução	25
1.1 Fármacos e Saúde Humana	27
1.2 Química Medicinal e Desenvolvimento de Substâncias Bioativas	27
1.3 Pesquisa e Desenvolvimento de Novos Fármacos	29
1.4 O Processo de Descoberta e Desenvolvimento	29
1.5 Doenças Tropicais Negligenciadas	32
1.6 Tripanossomíase Africana	35
1.7 Vias Bioquímicas e o Desenvolvimento de Fármacos para a TAH	41
1.8 Aldolase	47
Capítulo 2. Objetivos	53
Capítulo 3. Métodos Computacionais	57
3.1 Planejamento Baseado na Estrutura do Receptor	59
3.2 Docagem Molecular	60
3.3 Triagem Virtual	63
3.4 Dinâmica Molecular	66
3.5 Relações Quantitativas entre Estrutura e Atividade	68
3.6 Holograma QSAR	70
3.7 Análise Comparativa dos Campos Moleculares	72
Capítulo 4. Métodos Bioquímicos	75
4.1 Expressão, Purificação e Cristalização de Proteínas Heterólogas	77
4.2 Cinética Enzimática e Métodos de Análise	80
Capítulo 5. Padronização dos Protocolos Bioquímicos	87
5.1 Expressão, Purificação e Cristalização da Aldolase	89
5.2 Padronização dos Ensaios de Cinética Enzimática	95

Capítulo 6. Planejamento de Inibidores da Aldolase	101
6.1 Relações Quantitativas entre Estrutura e Atividade para Inibidores da Aldolase	103
6.2 O Conjunto de Dados	103
6.3 Homologia entre Aldolases e Seleção do Conjunto de Dados	107
6.4 Holograma QSAR	110
6.4.1 Mapas de Contribuição	115
6.4.2 Integração dos Mapas de Contribuição à Estrutura da Aldolase	117
6.5 Análise Comparativa dos Campos Moleculares	119
6.5.1 Alinhamento Molecular	119
6.5.2 Parâmetros dos Modelos CoMFA	120
6.5.3 Mapas de Contorno CoMFA	123
6.6 Dinâmica Molecular e as Interações Enzima-Inibidor	125
6.7 Bases Estruturais para o Planejamento de Inibidores Seletivos para a Aldolase de <i>Trypanosoma brucei</i>	132
6.8 Triagem Virtual para a Aldolase	133
6.8.1 Seleção e Preparo do Alvo Macromolecular	134
6.8.2 Triagem Virtual com o Programa DOCK	136
6.8.3 Triagem Virtual com o Programa GOLD	137
6.8.4 Seleção de Candidatos a Inibidores da Aldolase	137
Capítulo 7. Identificação de Inibidores da Aldolase	139
7.1 Avaliação Bioquímica de Candidatos a Inibidores da Aldolase	141
Capítulo 8. Conclusões e Perspectivas	149
Referências	153

Capítulo 1

Introdução

1.1 Fármacos e Saúde Humana

A saúde é um componente indispensável na manutenção do bem-estar humano, sendo um dos alicerces para a construção de uma sociedade plena, produtiva e desenvolvida. Em todo o mundo os vários setores da sociedade, incluindo os seus órgãos de governo, procuram os mais diversos meios para garantir à sua população o direito à saúde. Dentre as intervenções essenciais na saúde de nossa sociedade moderna, se encontra a farmacoterapia. É fato que os fármacos desempenham um papel central na melhoria da qualidade de vida. Por exemplo, nos últimos dez anos mais de 300 novos medicamentos foram aprovados pela agência norte-americana responsável pelo setor (FDA, na sigla inglesa para *Food and Drug Administration*), tornando-se importantes agentes indutores do aumento da expectativa de vida em diversos países¹. O papel da farmacoterapia também se destaca no auxílio à prevenção, tratamento e cura de doenças outrora consideradas fatais ou altamente debilitantes, diminuindo de forma significativa os riscos envolvidos em internações e intervenções cirúrgicas².

1.2 Química Medicinal e Desenvolvimento de Substâncias Bioativas

O processo de descoberta e desenvolvimento de novos fármacos tem se beneficiado cada vez mais de estratégias e métodos modernos de química medicinal. De acordo com a União Internacional de Química Pura e Aplicada (IUPAC, na sigla inglesa para *International Union of Pure and Applied Chemistry*), a química medicinal tem papel central na descoberta, desenvolvimento, identificação e interpretação do modo de ação molecular de compostos biologicamente ativos, bem como no planejamento de novas entidades químicas (NCEs, na sigla inglesa para *new chemical entities*). Também assume extrema importância na avaliação de propriedades farmacocinéticas e relativas à toxicidade – ADME/Tox: absorção, distribuição, metabolismo, excreção e toxicidade, além de estudos das relações entre a estrutura e atividade (SAR, na sigla inglesa para *structure-activity relationships*)³.

A química medicinal tem evoluído notavelmente nas últimas décadas em função dos grandes avanços científicos e tecnológicos das ciências moleculares. Entre estes, podem ser citados os avanços da biologia molecular e estrutural na área de espectroscopia de biomoléculas (e.g., cristalografia de raios-x, ressonância magnética nuclear) que levaram a elucidação da estrutura tridimensional de cerca de 80 mil proteínas. A exploração destes dados através de métodos de quimio- e bioinformática tem permitido a integração efetiva de ferramentas experimentais e computacionais para a compreensão dos aspectos moleculares predominantes na interação entre ligantes e macromoléculas⁴.

O emprego de dados estruturais faz parte de uma abordagem vastamente utilizada, denominada planejamento baseado na estrutura do receptor (SBDD, na sigla inglesa para *structure-based drug design*). Dentre os principais métodos de SBDD, destacam-se a triagem virtual (VS, na sigla inglesa para *virtual screening*), a docagem molecular e a dinâmica molecular⁵. Estes métodos de modelagem molecular incorporam várias ferramentas computacionais para a compreensão dos fenômenos envolvidos no processo de reconhecimento molecular. Desta maneira, novos ligantes podem ser planejados com base em simulações computacionais da energia de ligação entre moléculas pequenas e proteínas alvo⁶.

Outra abordagem amplamente explorada é o planejamento baseado na estrutura do ligante (LBDD, na sigla inglesa para *ligand-based drug design*)⁷. Nesta estratégia é utilizada a informação dos ligantes que possuem atividade biológica em sistemas bem definidos. Entre os principais métodos destacam-se a construção de modelos farmacofóricos, os estudos de SAR e de similaridade química, entre outros⁸. A análise quantitativa das relações entre a estrutura e atividade (QSAR, na sigla inglesa para *quantitative structure-activity relationships*) tem sido utilizada na identificação e quantificação das SARs. Os modelos de QSAR são utilizados para guiar a síntese de novas moléculas com propriedades superiores às aquelas que foram utilizadas na análise⁹. Os métodos de SBDD e LBDD são empregados na indústria farmacêutica e no âmbito acadêmico, com importantes contribuições no planejamento de novas moléculas de interesse terapêutico¹⁰.

Estas duas abordagens podem ainda ser empregadas de forma complementar, combinando as informações disponíveis, na geração de conhecimento¹¹. A integração dos métodos de planejamento ao trabalho químico e biológico experimental é um aspecto fundamental para a identificação de novas

moléculas bioativas com propriedades otimizadas¹². Neste contexto, onde o caráter multidisciplinar do processo de descoberta e desenvolvimento requer a exploração das várias interfaces entre as diversas ciências exatas e biológicas, a química medicinal consolida-se como campo fértil a ser explorado para o avanço da farmacoterapia.

1.3 Pesquisa e Desenvolvimento de Novos Fármacos

O setor farmacêutico, conhecido pelo seu caráter dinâmico e inovador, se tornou na economia contemporânea, um dos setores mais competitivos, caracterizado por altos níveis de investimento. O desenvolvimento de um novo fármaco, desde a sua concepção até a comercialização, requer de 12 a 15 anos de pesquisa e desenvolvimento (P&D), e investimentos da ordem de 800 milhões a 1 bilhão de dólares¹³. Estes números incluem os inúmeros fracassos que ocorrem durante o processo, pois a cada 10 mil moléculas que entram na linha de pesquisa, apenas uma chega ao mercado¹⁴.

O sucesso de um empreendimento desta natureza depende da constante introdução de fármacos pioneiros, ou no mínimo, de medicamentos de 2ª e 3ª gerações capazes de competir e se destacar no mercado. A sobrevivência destas companhias está, portanto, vinculada a um forte comprometimento com a inovação científica e tecnológica através do estabelecimento de programas robustos de P&D.

1.4 O Processo de Descoberta e Desenvolvimento

O processo de descoberta e desenvolvimento de novos fármacos é dividido em duas grandes fases: (i) pré-clínica, ou de descoberta, que leva de três a seis anos para ser concluída; e (ii) fase clínica, ou de desenvolvimento, na qual são realizados estudos para determinar a eficácia e a segurança dos candidatos a fármacos e que se estende por cerca de seis a nove anos. A fase pré-clínica inicia com estudos relacionados à fisiopatologia das doenças alvo. Incorpora ainda a

investigação do papel de vias bioquímicas e proteínas específicas nas variadas disfunções e doenças, com o objetivo de se alcançar uma melhor compreensão dos sistemas a serem explorados³. Nesta fase também ocorre a validação do alvo molecular e a análise da sua relevância no desenvolvimento de um candidato a novo fármaco. Esta é uma etapa crítica à medida que a biomacromolécula selecionada deve ser capaz de interagir com uma molécula pequena de forma que sua modulação possa produzir o efeito terapêutico desejado¹⁵.

Dentre as principais estratégias modernas de identificação e seleção de compostos capazes de interagir com uma proteína alvo, estão as triagens biológicas automatizadas em larga escala (HTS, na sigla inglesa para *high throughput screening*) e a triagem virtual (VS). As moléculas identificadas a partir destas triagens, que apresentam atividade biológica *in vitro*, são chamadas de *hits*, um termo do inglês que pode ser traduzido por ligante inicial (ou ligante bioativo)^{16,17}.

Uma vez que vários *hits* tenham sido identificados, são concentrados esforços na geração de compostos líderes, ou seja, moléculas com propriedades superiores, que apresentam características desejáveis para um fármaco, principalmente potência e seletividade, mas que ainda precisam da otimização de outras propriedades¹⁸. Para que um composto líder seja descoberto a partir de um ligante bioativo, várias estratégias de SBDD e LBDD são aplicadas em um processo denominado *hit-to-lead* (termo do inglês usado para indicar a otimização de um ligante bioativo a composto líder)¹⁹.

Os métodos de SBDD e LBDD também se consolidaram como ferramentas extremamente úteis na fase de otimização de compostos líderes (LO, termo do inglês para *lead optimization*) em relação às propriedades farmacodinâmicas e farmacocinéticas²⁰. A avaliação destas propriedades permite a priorização de compostos mais promissores^{21,22}. Os melhores compostos são submetidos a uma extensa e rigorosa avaliação da segurança para a realização de estudos em seres humanos. Os candidatos a novos fármacos, denominados de NCEs, são produzidos em torno de 3 a 6 anos, em projetos de P&D bem sucedidos²³. Após, as NCEs são submetidas às agências reguladoras para que seja aprovada a realização de testes em seres humanos (solicitação de investigação de novo fármaco, IND, na sigla inglesa para *investigational new drug*). No Brasil, é a Agência Nacional de Vigilância Sanitária (ANVISA), nos Estados Unidos é o FDA e na Europa é a *European Medicines Agency* (EMA). Esta solicitação é complexa e deve incluir todos os

resultados obtidos na fase pré-clínica, além de um planejamento detalhado para a realização dos testes clínicos¹³.

Na fase clínica I, o composto é testado em um pequeno grupo de indivíduos saudáveis, geralmente de 20 a 100 voluntários, para determinar a faixa de dosagem na qual o composto pode ser administrado com segurança e para avaliar suas propriedades farmacocinéticas¹⁴. A fase clínica II tem o propósito de avaliar, em um grupo de 100 a 500 pacientes, se o candidato a fármaco é capaz de produzir os efeitos desejados contra a doença alvo. Também são avaliados os efeitos adversos que se manifestam, além de serem realizados estudos relacionados à dose e regime de tratamento adequado¹⁵. Na fase clínica III, a NCE é testada em um número maior de pacientes, usualmente de 1000 a 5000, com o objetivo de gerar dados estatisticamente relevantes, que comprovem sua eficácia e segurança^{18,23}. Se a fase III for bem sucedida, uma nova solicitação é enviada à agência reguladora para aprovação do novo fármaco (NDA, na sigla inglesa para *new drug application*)²⁴.

Em caso de aprovação o fármaco é lançado no mercado, mas continua sendo monitorado. Isso acontece na chamada fase clínica IV (pós-lançamento), onde prosseguem os testes de segurança e eficácia em longo prazo, além do acompanhamento de efeitos colaterais inesperados. Em alguns casos, a agência pode solicitar estudos adicionais e se julgar necessário, até a retirada do fármaco do mercado²³. A sequência de etapas que caracteriza o processo de P&D de novos fármacos é esquematizada na Figura 1.

Figura 1 – Esquema geral de um processo típico de P&D de fármacos na indústria farmacêutica.

1.5 Doenças Tropicais Negligenciadas

As doenças tropicais negligenciadas (DTNs) são classificadas pela Organização Mundial de Saúde (OMS) em um único grupo de doenças apesar de constituírem condições clinicamente distintas. Têm em comum, características como forte associação com a pobreza e melhor adaptação em áreas tropicais, onde várias delas podem coexistir. Em sua maioria são doenças muito antigas, que em tempos remotos eram amplamente espalhadas pelo planeta e gradativamente se restringiram a regiões menos desenvolvidas²⁵. Embora estas doenças debilizem cerca de um bilhão de pessoas no mundo, estão relativamente escondidas em áreas rurais remotas e regiões urbanas deterioradas, onde prevalece a pobreza. A transmissão entre a população é amplamente facilitada pela precariedade das habitações, falta de água potável, saneamento básico, degradação ambiental e abundância de insetos vetores²⁶.

Atualmente 17 doenças são classificadas como DTNs pela OMS (Tabela 1). São endêmicas em cerca de 150 países, sendo que na maioria dos casos coexistem com uma ou mais doenças²⁵.

Tabela 1 – Doenças tropicais negligenciadas e seus agentes causadores.

DTN	Agente causador
Dengue	Flavivírus
Raiva	Rabdovírus
Tracoma	<i>Chlamydia trachomatis</i>
Úlcera de Buruli	<i>Mycobacterium ulcerans</i>
Sífilis	<i>Treponema pallidum</i>
Lepra	<i>Mycobacterium leprae</i>
Doença de chagas	<i>Trypanosoma cruzi</i>
Tripanossomíase africana	<i>Trypanosoma brucei</i>
Leishmanioses	<i>Leishmania spp.</i>

continua

conclusão

Cisticercose	<i>Taenia solium</i>
Dracunculíase	<i>Dracunculus medinensis</i>
Equinococose	<i>Echinococcus granulosus</i> ; <i>Echinococcus multilocularis</i>
Infecções por trematódeos	<i>Clonorchis sinensis</i> ; <i>Opisthorchis viverrini</i> ; <i>Fasciola hepática</i> ; <i>Paragonimus spp.</i>
Filariase linfática	<i>Wuchereria bancrofti</i> ; <i>Brugia malayi</i> ; <i>Brugia timori</i>
Oncocercose	<i>Onchocerca volvulus</i>
Esquistossomose	<i>Schistosoma intercalatum</i> ; <i>S. japonicum</i> ; <i>S. mansoni</i> ; <i>S. mekong</i>
Helminthíases	<i>Ancylostoma duodenale</i> ; <i>Necator americanus</i> ; <i>Ascaris lumbricoides</i> ; <i>Trichuris trichiura</i>

Alguns destes patógenos possuem ciclos de vida simples e necessitam de apenas um hospedeiro para completar todos os estágios do seu desenvolvimento. Outros possuem ciclos de vida complexos, necessitando de vários hospedeiros. Neste último caso, o hospedeiro intermediário é aquele que apresenta o parasita em sua fase larvária ou em fase de reprodução assexuada. O hospedeiro definitivo é aquele que apresenta o parasita em sua fase madura ou em fase de reprodução sexuada²⁶.

Destas 17 doenças, nove são causadas por macroparasitas e oito por microparasitas. A maioria dos microparasitas (vírus, bactérias e protozoários) se reproduzem dentro do hospedeiro e sua transmissão ocorre de quatro maneiras: (i) através de contaminação ambiental; (ii) contato íntimo, incluindo a via transplacentária; (iii) através de um vetor que pode ou não ser o hospedeiro intermediário; ou (iv) transfusão sanguínea e transplante de órgãos. Os microparasitas causam infecções agudas que podem evoluir para a morte e infecções recorrentes²⁷.

Os macroparasitas, em sua maioria, não se reproduzem no hospedeiro definitivo. Sua transmissão pode se dar através de: (i) ingestão de alimentos ou água contaminada; (ii) penetração do patógeno pela pele; ou (iii) ingestão ou contato com

um hospedeiro intermediário contaminado. As infecções causadas por macroparasitas são crônicas e apresentam menor taxa de mortalidade²⁵.

A negligência com a qual este grupo de doenças é tratado se deve a pouca força econômica e política das populações mais afetadas. Apesar da sua melhor adequação aos ambientes tropicais empobrecidos, eventualmente elas se alastram para outras regiões do planeta, o que tem resultado em uma crescente preocupação por parte de alguns países que ainda não são afetados de forma endêmica. Embora a necessidade por novos fármacos para as DTNs seja evidente, o baixo poder aquisitivo das populações dos países afetados é um fator limitante. As companhias farmacêuticas não fazem investimentos significativos, pois o retorno financeiro não é compatível com os altos investimentos em P&D²⁸. Este cenário traz consequências desastrosas para as sociedades afetadas. O impacto causado pelas DTNs na saúde coletiva da população é avaliado pelo parâmetro conhecido como DALY (na sigla inglesa para *disability adjusted life year*), que estima os anos de vida produtiva perdidos por morte prematura e pela incapacitação dos indivíduos afetados^{28,29}.

Recentemente as perspectivas em relação a este panorama têm melhorado. Existe atualmente um crescente reconhecimento do problema e um aumento dos investimentos em pesquisa principalmente através do estabelecimento de parcerias público-privadas, como a iniciativa Medicamentos para Doenças Negligenciadas (DNDi, na sigla inglesa para *Drugs for Neglected Diseases initiative*)³⁰. Deste esforço conjunto, participam governos, empresas e organizações não governamentais. Como exemplo, podemos citar o Fundo das Nações Unidas para a Infância (UNICEF, na sigla inglesa para *United Nations Children's Fund*), a OMS, o Programa de Desenvolvimento das Nações Unidas (UNDP, na sigla inglesa para, *United Nations Development Programme*) e o Banco Mundial. Em conjunto, estas quatro entidades financiam uma das mais importantes iniciativas para pesquisa em doenças negligenciadas: o Programa Especial para Treinamento e Pesquisa em Doenças Negligenciadas (TDR, na sigla inglesa para *Special Programme for Research and Training in Tropical Diseases*). Fundada em 1975, a iniciativa TDR coordena e financia esforços globais no combate às DTNs e tem alcançado resultados importantes em várias frentes de pesquisa³¹.

Estas novas modalidades de pesquisa em rede, têm contribuído substancialmente com recursos financeiros que são aplicados no desenvolvimento de novos métodos de diagnóstico, vacinas e na implementação de novas

abordagens terapêuticas, como o tratamento preventivo em larga escala. O envolvimento da indústria farmacêutica nestas iniciativas tem crescido gradativamente, principalmente através de doações de medicamentos que são utilizados em milhões de pessoas afetadas. Até o término de 2009, uma população de 887 milhões de pessoas havia recebido tratamento preventivo para algumas destas doenças, como a filaríase, a oncocercose, a esquistossomose e o tracoma³².

1.6 Tripanossomíase Africana

A tripanossomíase africana humana (TAH), ou doença do sono, é uma das mais complexas DTNs, afetando uma grande extensão do continente africano. Foi quase eliminada na década de 1960, mas ressurgiu após gradativa descontinuação dos programas de controle, colocando em risco de infecção, cerca de 60 milhões de pessoas. Na ausência de tratamento a doença é fatal³³.

O agentes causadores da TAH são parasitas da espécie *Trypanosoma brucei*, que são transmitidos aos seres humanos pela mordida da mosca do gênero *Glossina*, conhecida como tse-tsé. A mosca assume o papel de hospedeiro intermediário e vetor da doença, ao ingerir o parasita durante o repasto sanguíneo em seres humanos infectados³⁴. A Figura 2 ilustra o ciclo de vida do parasita.

Figura 2 – Ciclo de vida do causador da TAH, o parasita *Trypanosoma brucei*. Estágios de desenvolvimento no hospedeiro intermediário, a mosca tsé-tsé, e no hospedeiro definitivo, o ser humano.

A doença do sono é endêmica na África subsaariana, onde a subespécie *T. b. gambiense* afeta 24 países da região oeste e central do continente, sendo responsável por cerca de 90% dos casos (Figura 3). A subespécie *T. b. rhodesiense* é endêmica em 13 países das regiões leste e sul e causa cerca de 10% dos casos (Figura 4)²⁵. A distribuição da doença ao longo do território africano é bastante afetada pelas guerras, deslocamentos populacionais, variações climáticas e a base econômica da região. Estas correlações foram investigadas através de levantamentos epidemiológicos em regiões historicamente afetadas pela TAH e que há cerca de 10 anos não apresentam mais a doença³³.

Dados epidemiológicos obtidos entre 1999 e 2008 identificaram o surgimento de 10 mil novos casos de TAH por ano. Atualmente, estima-se que o número total de casos varie em torno de 70 mil. No entanto, devido às dificuldades de notificação e levantamento dos dados, inerentes aos países afetados, este número é considerado apenas uma fração do número real de pessoas infectadas pelo parasita²⁵.

A doença pode assumir duas formas clínicas distintas, de acordo com a subespécie causadora. A subespécie *Trypanosoma brucei gambiense* causa uma

infecção crônica, caracterizada por baixa parasitemia e evolução mais lenta. O *Trypanosoma brucei rhodesiense* causa a forma aguda, caracterizada por alta parasitemia e evolução rápida e mais agressiva dos sinais e sintomas³⁵.

As duas subespécies são exclusivamente extracelulares e a partir da sua introdução na corrente sanguínea passam a se reproduzir rapidamente, causando a fase precoce da TAH, que se caracteriza por febre, cefaléia, prostração, perda de peso, dores nas articulações, lesões de pele e vômitos. Em algumas semanas ou meses, o parasita ultrapassa a barreira hematoencefálica e invade o sistema nervoso central, passando a se reproduzir no fluido cérebro-espinhal. Nesta etapa, denominada fase tardia, são causados severos danos neurológicos e psiquiátricos, levando o indivíduo ao coma e em casos não tratados, à morte³⁶.

Figura 3 – Distribuição geográfica da TAH causada pela subespécie *T. b. gambiense*.

Figura 4 – Distribuição geográfica da TAH causada pela subespécie *T. b. rhodesiense*.

O tratamento da doença do sono baseia-se num repertório bastante limitado de fármacos, que é considerado arcaico e caracterizado pela sua baixa eficácia e alta toxicidade (provocam graves efeitos adversos). Todos são administrados por via intravenosa ou intramuscular, o que impõe dificuldades adicionais para sua manipulação e administração³⁷. Os fármacos utilizados no tratamento da TAH são apresentados na Figura 5.

Figura 5 – Estrutura química dos fármacos utilizados no tratamento da tripanossomíase africana.

A suramina foi descoberta em 1920, após observações de que certos corantes de estrutura química similar possuíam atividade tripanocida. Seu mecanismo de ação não é conhecido, e por não ultrapassar a barreira hematoencefálica, é utilizada apenas para tratar a fase precoce da doença. Uma série de efeitos colaterais é observada, como toxicidade renal e neurológica³⁸.

A pentamidina, utilizada desde 1921, é uma amidina que apresenta atividade tripanocida, para a qual não há qualquer informação acerca do mecanismo de ação. Seu uso se restringe à fase precoce da doença causada por *T. b. gambiense*, apresentando baixa atividade contra *T. b. rhodesiense*. Presume-se que sua ineficácia na fase tardia deve-se à sua baixa capacidade em cruzar a barreira hematoencefálica. Dentre os efeitos tóxicos, o mais grave é a hipoglicemia, que pode ser fatal³⁹.

O melarsoprol, um composto extremamente tóxico e que possui arsênio em sua estrutura, foi descoberto como agente tripanocida em 1949. É capaz de agir sobre as duas subespécies, mas devido a sua extrema toxicidade tem seu uso restrito apenas à fase tardia da doença. Apesar de seu mecanismo de ação não ser conhecido, sabe-se que é um pró-fármaco e que seu princípio ativo é o óxido de melarsen. Cerca de 6% dos indivíduos que são tratados com o melarsoprol morrem devido aos danos causados pelo fármaco⁴⁰.

A eflornitina, um composto inicialmente projetado para o tratamento do câncer, nunca chegou a ser utilizado para este fim. Em 1980 descobriu-se que possuía atividade tripanocida e em 1991 passou a ser utilizado para o tratamento da TAH. É o único fármaco cujo mecanismo de ação é compreendido. A eflornitina age como inibidor irreversível da ornitina descarboxilase, enzima que catalisa a primeira reação da via de produção de poliaminas, rota bioquímica essencial para a reprodução do parasita. O fármaco atua apenas na fase tardia de infecções causadas por *T. b. gambiense* e sua manipulação e administração envolvem procedimentos complexos, que limitam sua utilização em áreas remotas. Um dos aspectos negativos da eflornitina foi à ocorrência de casos fatais de choque séptico relacionados ao seu uso⁴¹. A Tabela 2 resume as principais características dos fármacos utilizados para o tratamento da TAH.

Neste contexto de sérias limitações, estudos clínicos realizados entre os anos de 2003 e 2008 identificaram que a combinação de eflornitina e nifurtimox – fármaco utilizado no tratamento da doença de Chagas – foi capaz de reduzir a duração do tratamento, aumentar a taxa de cura e por consequência, reduzir a ocorrência dos efeitos colaterais. O número de mortes relacionadas ao tratamento também foi reduzido. A partir destes dados, a OMS adotou a terapia combinada eflornitina-nifurtimox como padrão no tratamento da infecção tardia causada por *T. b. gambiense*. Os fármacos utilizados no tratamento da TAH são obtidos geralmente por meio de parcerias público-privadas entre a OMS e indústrias farmacêuticas⁴¹.

Soma-se aos aspectos já discutidos, o surgimento de cepas de parasitas resistentes. Além disso, a ausência de um fármaco que seja ativo contra as duas subespécies, nos dois estágios da doença, leva à necessidade de se identificar laboratorialmente a subespécie causadora. Frequentemente, o sistema de saúde das regiões afetadas carece de recursos humanos e de infraestrutura adequada para a realização deste tipo de diagnóstico diferencial. Neste cenário, há uma evidente necessidade de desenvolvimento de novos fármacos mais eficazes e seguros⁴².

Tabela 2 – Características dos fármacos disponíveis para o tratamento da TAH.

Fármaco	Indicação	Limitações	Mecanismo de ação	Regime de tratamento
Suramina	Fase precoce <i>T. b. rhodesiense</i>	Não ultrapassa a barreira hematoencefálica, toxicidade	Desconhecido	Injeção intravenosa, dose inicial de 200 mg e depois 1 g nos dias 1, 3, 7, 14 e 21
Pentamidina	Fase precoce <i>T. b. gambiense</i>	Não ultrapassa a barreira hematoencefálica, toxicidade	Desconhecido	Injeção intramuscular em doses diárias de 4mg/kg por 7 dias
Melarsoprol	Fase tardia <i>T. b. rhodesiense</i> e <i>T. b. gambiense</i>	Toxicidade severa	Desconhecido	3 series de 3 doses diárias via intravenosa com período de 7 dias de pausa entre elas
Eflornitina	Fase tardia <i>T. b. gambiense</i>	Regime de tratamento complexo	Inibidor irreversível da ornitina descarboxilase	Injeção intravenosa de 400mg/kg por dia divididos em 4 doses diárias por 14 dias
Nifurtimox/ Eflornitina	Fase tardia <i>T. b. gambiense</i>	Não efetivo contra <i>T. b. rhodesiense</i>	Geração de radicais livres e inibição da ornitina descarboxilase	Nifurtimox: 15mg/kg a cada 8 horas via oral, por 10 dias Eflornitina: Injeção intravenosa de 400mg/kg por dia divididos em 2 doses diárias por 7 dias

1.7 Vias Bioquímicas e o Desenvolvimento de Fármacos para a TAH

A bioquímica da família Trypanosomatidae tem sido extensivamente estudada, principalmente devido a estes organismos apresentarem várias peculiaridades em alguns de seus processos celulares. O causador da TAH é um dos representantes mais estudados desta família. Nos últimos anos várias vias metabólicas de *Trypanosoma brucei* têm sido exploradas como potenciais alvos para a descoberta de novos fármacos. A conclusão do sequenciamento genômico do parasita contribuiu de forma decisiva para a compreensão e manipulação do seu material genético e de seus produtos, resultando na investigação de várias

proteínas, como potenciais alvos moleculares⁴³. Estas vias metabólicas e os principais alvos moleculares que têm sido explorados são brevemente discutidos a seguir.

Biossíntese de poliaminas e de tripanotiona

A biossíntese de poliaminas em *Trypanosoma brucei* é essencial para o crescimento e diferenciação do parasita, como tem sido demonstrado em estudos de *knockout* gênico e RNA de interferência (RNAi). Esta rota metabólica adquire mais importância à medida que apresenta características diferentes da via humana. A TAH é a única doença para a qual a biossíntese de poliaminas é alvo para um fármaco utilizado na clínica, a eflornitina. Duas enzimas são investigadas como alvos moleculares: a ornitina descarboxilase e a adenosil-metionina descarboxilase⁴⁴.

A tripanotiona, por sua vez, é uma molécula essencial que atua no sistema redox celular, responsável pela neutralização de espécies reativas de oxigênio. Sua biossíntese está ligada à via das poliaminas através de um intermediário comum, a espermidina. Das enzimas desta via, a tripanotiona sintetase e a tripanotiona redutase são alvos para o desenvolvimento de fármacos⁴⁵.

Metabolismo de purinas e pteridinas

O metabolismo de purinas e pteridinas se torna uma via metabólica promissora a ser explorada, à medida que o parasita não possui uma *via de salvação* para estes metabólitos que são essenciais para o crescimento celular. Estes são adquiridos do hospedeiro através de uma série de proteínas transportadoras. Proteínas como a diidrofolato redutase e a pteridina redutase têm sido exploradas em outras doenças infecciosas e parasitárias, tais como a malária⁴⁶.

DNA topoisomerases

DNA topoisomerases são essenciais nos processos de replicação e transcrição do DNA. Inibidores das topoisomerases bacterianas têm demonstrado potencial tripanocida. O principal obstáculo tem sido encontrar inibidores seletivos para a enzima parasitária em relação à enzima do hospedeiro. A principal estratégia

tem sido a busca por diferenças estruturais entre a enzima de *T. brucei* e de mamíferos⁴⁴.

Biossíntese de lipídeos

Assim como ocorre na maioria dos eucariotos, a biossíntese de ácidos graxos e esfingolipídeos é essencial para o parasita, principalmente no processo de formação da membrana celular. As especificidades da via parasitária podem ser exploradas no desenvolvimento de inibidores seletivos de várias enzimas, como a elongase, a esfingolipídeo sintetase e a N-miristoil transferase. O papel essencial que estas enzimas desempenham no ciclo de vida do parasita tem sido comprovado através de estudos de RNAi⁴⁵.

Modificações proteicas

A adição de isoprenóides em determinadas proteínas é um dos principais mecanismos de transdução de sinais em tripanossomatídeos. As pesquisas acerca deste processo celular têm se concentrado na enzima farnesiltransferase, responsável por catalisar esta reação. Outra modificação proteica, também envolvida em processos de transdução de sinais, é a transferência de grupos fosfato para enzimas, principalmente aquelas localizadas na membrana celular. As enzimas que catalisam estas transferências de fosfato são as quinases (a glicogênio sintetase quinase é uma das enzimas que têm sido investigadas).

Também essencial para reprodução do parasita é o processo de clivagem de proteínas. A busca por inibidores de proteases tem sido utilizada com sucesso em doenças virais, parasitárias e câncer. Estudos de RNAi identificaram a protease cathepsina B como essencial para o desenvolvimento do parasita⁴⁷.

Via glicolítica

Dentre as vias metabólicas exploradas para o desenvolvimento de novos fármacos para a tripanossomíase africana, a via de degradação da glicose é uma das áreas mais efervescentes de pesquisa⁴⁸.

A glicólise surge como alternativa promissora à medida que é o único processo gerador de energia da forma sanguínea do parasita. Tem sido demonstrado que a glicólise em tripanossomatídeos contém várias características distintas do processo que ocorre em células de mamíferos, como por exemplo, a compartimentalização de várias enzimas envolvidas no processo, em organelas denominadas glicossomos, como ilustrado na Figura 6. Estas diferenças permitem o desenvolvimento de compostos seletivos para as enzimas do parasita⁴⁹.

No metabolismo de *T. brucei*, a glicose é a única fonte de carbono e energia. No sangue e em outros fluidos do hospedeiro, o parasita tem acesso a uma fonte ilimitada de glicose, que é internalizada e metabolizada por um processo de fermentação aeróbica, a uma taxa muito superior àquela encontrada na maioria das células eucarióticas. Enquanto na maioria dos organismos a glicólise ocorre no citosol, no parasita suas primeiras sete enzimas estão localizadas dentro do glicossomo, uma organela separada do citosol por uma membrana. Embora os glicossomos se assemelhem aos peroxissomos, eles não possuem as típicas enzimas peroxissomais, como a peroxidase e a catalase. Cerca de 90% das enzimas glicossomais são enzimas que participam do metabolismo da glicose⁴⁹.

Os organismos que possuem mitocôndria ativa utilizam a glicólise como o primeiro passo no processo de obtenção de energia, que então continua através do ciclo do ácido cítrico e da cadeia respiratória, obtendo ao final de todo o processo um rendimento de 38 moléculas de ATP⁵⁰.

Figura 6 – Glicólise em *T. brucei*. A numeração corresponde às reações catalisadas pelas seguintes enzimas: (1) hexoquinase; (2) fosfoglicose isomerase; (3) fosfofrutoquinase; (4) aldolase; (5) triosefosfato isomerase; (6) gliceraldeído-3-fosfato desidrogenase; (7) fosfogliceratoquinase; (8) glicerol-3-fosfato desidrogenase; (9) glicerol quinase; (10) fosfoglicerato mutase; (11) enolase; (12) piruvato quinase.

A única mitocôndria presente no parasita, por sua vez, não possui as enzimas do ciclo do ácido cítrico e nem da cadeia respiratória, fazendo com que o produto final da glicólise, o piruvato, seja excretado diretamente na corrente sanguínea do

hospedeiro, ao invés de ser oxidado a dióxido de carbono e água, como em outros organismos que possuem mitocôndria. O baixo rendimento da fosforilação ao nível do substrato, em termos de produção de ATP é a causa do alto consumo de glicose característico do parasita⁵¹. Dado que os tripanossomatídeos não possuem nenhuma reserva de carboidratos, como glicogênio ou creatina fosfato, a inibição do fluxo glicolítico ou a depleção do seu substrato, leva imediatamente à diminuição do nível celular de ATP, à perda de motilidade e ao desaparecimento do parasita da corrente sanguínea do hospedeiro⁵².

Atualmente, não resta dúvida de que as enzimas que participam do catabolismo da glicose constituem excelentes alvos moleculares para o desenvolvimento de fármacos. De fato, várias delas foram geneticamente validadas para este propósito, fazendo com que as discussões se deslocassem para a questão de quais enzimas constituem os melhores alvos⁴⁴.

Modelos matemáticos têm sido utilizados para prever a dinâmica do fluxo da glicose e a concentração dos metabólitos encontrados experimentalmente nos parasitas. Estes dados têm contribuído para a compreensão do grau de controle que cada enzima exerce sobre o processo. Enzimas como a gliceraldeído-3-fosfato desidrogenase, a glicerol-3-fosfato desidrogenase, a fosfogliceratoquinase, a aldolase e a fosfofrutoquinase, exercem um nível de controle considerável, especialmente em condições fisiológicas, onde a oferta de glicose não é um fator limitante. Enzimas que não possuem papel significativo na regulação da via, mas que podem ter sua atividade modulada por inibidores enzimáticos potentes, também podem constituir alvos moleculares, desde que sejam capazes de reduzir significativamente o fluxo⁵⁰. Esta abordagem foi comprovada em experimentos de *knockout* gênico, onde a supressão da atividade da triose fosfato isomerase, uma enzima considerada não essencial para a glicólise, foi capaz de diminuir a produção de ATP pela metade, o que não é capaz de suprir as necessidades metabólicas do parasita. Estes estudos indicam que um inibidor da via glicolítica, capaz de reduzir consideravelmente a produção de ATP, poderá paralisar a reprodução do patógeno⁵³. A relevância da via glicolítica pode ser constatada na Tabela 3. Dentre todas as vias bioquímicas exploradas no planejamento de fármacos, a glicólise é a que possui mais enzimas validadas⁴⁴.

Tabela 3 – Vias metabólicas e alvos moleculares validados para o desenvolvimento de novos fármacos para a TAH.

Via metabólica	Alvo molecular
Biossíntese de poliaminas e tripanotona	Ornitina descarboxilase Adenosil-metionina descarboxilase Tripanotona sintetase Tripanotona redutase
Metabolismo de purinas e pteridinas	Diidrofolato redutase Pteridina redutase
Modificações do DNA	DNA topoisomerase
Biossíntese de ácidos graxos e esfingolípídeos	Elongase Esfingolípídeo sintetase N-miristoiltransferase
Prenilação de proteínas	Farnesiltransferase
Fosforilação de proteínas	Glicogênio sintetase quinase
Clivagem de proteínas	Catepsina B
Divisão celular	Tubulina
Via glicolítica	Hexoquinase Fosfoglicose isomerase Fosfofrutoquinase Aldolase Fosfogliceratoquinase Enolase

1.8 Aldolase

As enzimas são consideradas os alvos moleculares mais atrativos para o desenvolvimento de novos fármacos, devido ao seu papel essencial em muitos processos fisiopatológicos. Este aspecto é evidenciado pelos números que demonstram que cerca de metade dos fármacos atualmente em uso clínico são inibidores enzimáticos⁵⁴⁻⁵⁷. Portanto, a identificação e caracterização de enzimas essenciais à sobrevivência de microorganismos patogênicos são fundamentais para o planejamento de novos agentes quimioterápicos.

Neste cenário, a enzima aldolase de *T. brucei* foi recentemente validada como alvo molecular. A relevância da enzima para a via glicolítica de *T. brucei* foi

demonstrada através de experimentos de RNAi, nos quais a supressão parcial da concentração intracelular da enzima foi suficiente para rapidamente causar a morte do parasita⁵⁸.

A enzima (EC 4.1.2.13), uma liase que está presente na maioria dos organismos, possui papel essencial na via de degradação da glicose. A aldolase catalisa a clivagem reversível da hexose frutose-1,6-bifosfato (FBP), resultando na formação de duas trioses: a diidroxiacetona-fosfato (DHAP) e o gliceraldeído-3-fosfato (GAP), de acordo com a Figura 7.

Figura 7 – Reação catalisada pela aldolase: clivagem reversível da frutose-bifosfato em diidroxiacetona-fosfato e gliceraldeído-3-fosfato.

Existem duas classes de aldolases. As enzimas de classe I (à qual pertence a aldolase de *T. brucei*), que estão presentes em vários grupos de organismos, desde procariotos até eucariotos, são caracterizadas pela formação de uma ligação covalente entre um resíduo de lisina e o substrato, sem necessitar de um íon metálico divalente para a catálise. O mecanismo de reação da aldolase de *T. brucei* é ilustrado na Figura 8. As enzimas de classe II, que ocorrem apenas em procariotos e eucariotos inferiores como fungos e algas, requerem um íon metálico, sendo chamadas de metaloaldolases⁵⁹.

Figura 8 – Esquema simplificado da reação catalisada pela aldolase de *T. brucei*.

A enzima é abundante em *T. brucei* e exerce um controle significativo sobre o fluxo glicolítico do parasita, o que não se confirma no caso da enzima encontrada nos eritrócitos humanos, nos quais a diminuição da atividade da aldolase não provoca efeitos adversos significativos⁶⁰.

Assim como a maioria das aldolases, a enzima de *T. brucei* é um homotetrâmero de 160 kDa, que resulta da associação de quatro cadeias polipeptídicas idênticas de 360 resíduos de aminoácidos. Cada subunidade possui um sítio ativo independente, tanto estruturalmente quanto funcionalmente, ou seja, cada sítio é formado por resíduos de apenas uma cadeia e não há fenômenos de interação entre os sítios. A enzima também não depende de cofatores ou moduladores para catalisar a reação⁵⁹. A estrutura tridimensional da aldolase de *T. brucei* é apresentada na Figura 9.

Em mamíferos, foram identificadas três isoenzimas da aldolase: a isoenzima A, presente nos músculos e nos eritrócitos; a B, presente no fígado, rins e intestino; e a C, no cérebro. A identidade da sequência primária entre as isoenzimas humanas e a parasitária é de 49%. Algumas diferenças estruturais tornam possível o planejamento de inibidores seletivos para a enzima do parasita^{59,61}.

Figura 9 – Estrutura tetramérica da aldolase de *T.brucei*.

A comparação da estrutura primária de várias isoenzimas de aldolases de classe I revelou um alto grau de identidade entre os resíduos, particularmente aqueles pertencentes ao sítio ativo. Como consequência, a estrutura da enzima é altamente conservada e um único padrão de estrutura terciária e quaternária é observado. Este enovelamento comum foi confirmado pela determinação estrutural de aldolases de várias espécies, através de cristalografia por difração de raios-x. Atualmente o banco de dados PDB (na sigla inglesa para *Protein Data Bank*) possui 85 estruturas de aldolases de classe I depositadas⁶².

A resolução da estrutura tridimensional da aldolase de *T. brucei* (código PDB: 1F2J) revelou que suas subunidades apresentam um enovelamento do tipo barril β , no qual folhas β formam um cilindro similar a um barril, que é circundado por α -hélices⁶¹. O sítio ativo localiza-se na face interna do barril que está exposta ao solvente (Figura 10), levando à principal característica do sítio, que é a presença de um grande número de resíduos carregados. Dentre estes, a lisina 239 é a responsável pela formação da ligação covalente com o substrato⁵⁹.

Dentre os aspectos que se tornaram evidentes à luz da estrutura tridimensional da enzima, destaca-se o fato de que os requerimentos necessários

para que haja complementaridade eletrostática entre os ligantes e os resíduos carregados presentes no sítio de interação, se opõem àqueles necessários para promover a passagem destes compostos através das membranas do parasita, o que torna extremamente desafiadora a busca por inibidores clinicamente úteis. Esta é uma das razões para a inexistência na literatura de registros de inibidores da aldolase de *T. brucei* com atividade contra o parasita⁴⁴.

Figura 10 – Estrutura do monômero da aldolase de *T. brucei*, com destaque para o sítio ativo que está localizado na face interna barril β (amarelo).

Capítulo 2

Objetivos

2 Objetivos

O objetivo geral desta tese de doutorado compreende a identificação de novos inibidores da enzima aldolase de *Trypanosoma brucei*, através da integração de métodos computacionais e experimentais. Os objetivos específicos incluem:

1. Desenvolvimento e padronização de protocolos de expressão, purificação e cristalização para a aldolase de *T. brucei*.
2. Desenvolvimento e padronização de ensaios de cinética enzimática para a avaliação de inibidores da enzima.
3. Desenvolvimento de modelos de relações quantitativas entre a estrutura e atividade (QSAR) para uma série de inibidores da aldolase, através dos métodos holograma QSAR (HQSAR) e análise comparativa dos campos moleculares (CoMFA).
4. Realização de simulações de dinâmica molecular para complexos formados entre a aldolase e seus inibidores, com o propósito de caracterizar os fenômenos de reconhecimento molecular.
5. Planejamento e realização de triagens virtuais em larga escala com o objetivo de identificar novos inibidores da aldolase.
6. Seleção e aquisição de moléculas identificadas nas triagens virtuais.
7. Avaliação bioquímica dos ligantes selecionados na etapa de triagem virtual frente à aldolase de *T. brucei*.

Com a finalidade de apresentar e discutir os métodos utilizados, a tese é apresentada a seguir em seções separadas para os métodos computacionais e para os métodos bioquímicos.

Capítulo 3

Métodos Computacionais

Os modelos tradicionais de descoberta e desenvolvimento de fármacos baseiam-se em abordagens como a síntese de séries de compostos ativos e a triagem sistemática (HTS) de grandes coleções de compostos em diferentes modelos experimentais. Estes modelos têm dado sinais de esgotamento por várias razões, tais como a velocidade e custo dos testes, ocorrência de falsos positivos e ausência de inovação^{4,23,63}. Neste contexto, ganhou especial importância o desenvolvimento de abordagens racionais de planejamento, como alternativa ao paradigma de “testar tudo contra tudo”. Nesta transição para modelos de pesquisa mais integrados, os métodos computacionais figuram certamente entre as ferramentas essenciais nas etapas de identificação de ligantes iniciais, conversão de ligantes iniciais em compostos líderes e otimização de compostos líderes⁶⁴. As abordagens utilizadas nesta tese são discutidas brevemente neste capítulo.

3.1 Planejamento Baseado na Estrutura do Receptor

A compreensão dos princípios pelos quais os receptores reconhecem e interagem com ligantes adquiriu grande relevância no processo de P&D de fármacos^{5,8}. O planejamento baseado na estrutura do receptor (SBDD) se refere, portanto, ao uso sistemático de informações estruturais acerca do alvo macromolecular, geralmente obtidas através de técnicas como a cristalografia de raios-x, a ressonância magnética nuclear ou a modelagem computacional por homologia. Esta abordagem busca o planejamento de pequenas moléculas com características estereoquímicas e eletrostáticas que resultem em interações de alta afinidade com o sítio de ligação. A disponibilidade deste tipo de informação possibilita a avaliação dos detalhes estruturais do sítio de ligação e o planejamento de novas moléculas que possuam os requerimentos necessários para modular significativamente a atividade do receptor⁴⁷. De forma geral, o processo caracteriza-se por sua iteratividade, ou seja, são realizados vários ciclos de planejamento *in silico*; síntese de compostos; avaliação bioquímica *in vitro*; e obtenção de estruturas 3D de complexos receptor-ligante. Ao final deste processo os dados de atividade *in vitro* são correlacionados com as informações estruturais e então se retorna ao início do ciclo, onde novas etapas de planejamento são realizadas para otimizar a

complementaridade entre a cavidade de interação e os ligantes¹¹. Este processo é esquematizado na Figura 11.

Durante a etapa de planejamento dos ligantes é importante que se considere a flexibilidade do sítio de ligação. Enzimas sofrem geralmente alterações conformacionais de forma a ajustar as distâncias e os ângulos de interação com o ligante. Esta questão pode ser abordada através do uso de técnicas como a docagem flexível e a dinâmica molecular⁶⁵.

Figura 11 – Esquema do processo de planejamento baseado na estrutura do receptor (SBDD).

3.2 Docagem molecular

A docagem molecular é um método de predição do modo de ligação de pequenas moléculas na cavidade de interação de macromoléculas^{11,15}. É uma ferramenta computacional que permite: a exploração das possíveis conformações das moléculas de interesse dentro da cavidade de interação do alvo; a avaliação das interações intermoleculares responsáveis pelo processo de reconhecimento molecular; e a quantificação da energia de ligação receptor-ligante¹⁷. Estes recursos

permitem o estabelecimento de uma classificação relativa das moléculas consideradas no processo (Figura 12).

Figura 12 – Esquema do processo de docagem molecular.

Quando a conformação predita de um ligante é comparada com os dados estruturais do complexo proteína-ligante, percebe-se que muitos algoritmos são capazes de prever com sucesso a conformação destes compostos no sítio de interação. No entanto, em muitos casos não se obtêm uma boa concordância dos dados relativos à energia de interação proteína-ligante. Obstáculos como o efeito do solvente, as perdas entrópicas devido aos graus de liberdade do ligante e da contribuição das mudanças conformacionais do receptor, são alguns desafios a serem superados pelos algoritmos de docagem^{65,66}.

De acordo com o método de busca, os programas de docagem dividem-se em duas categorias: (i) busca sistemática ou construção incremental, que tratam cada molécula como um conjunto de fragmentos, dentre os quais um determinado fragmento é inicialmente posicionado no sítio de ligação para que em seguida os outros fragmentos sejam posicionados um de cada vez, até que toda a molécula seja reconstruída^{67,68}. Exemplos de algoritmos de busca sistemática são os programas: FlexX[®] (Tripos Inc., USA) e DOCK (UCSF – Universidade da Califórnia em São Francisco, USA); (ii) métodos estocásticos, que por sua vez realizam uma varredura randômica por todo o espaço conformacional do ligante, como nos algoritmos genéticos, que inspirados nos conceitos da teoria da evolução e seleção natural, exploram os graus de liberdade torcionais dos ângulos diedrais e os graus de liberdade rotacionais e translacionais do ligante no espaço cartesiano. As soluções

para o ligante são tratadas como cromossomos que, como na biologia, podem sofrer mutações e recombinações com outros cromossomos (soluções). Um dos programas mais utilizados é o GOLD[®] – *Genetic Optimization for Ligand Docking* – (Cambridge Crystallographic Data Centre, UK)⁶⁹.

Em relação ao cálculo da energia livre de ligação entre o ligante e o receptor, há diversos tipos de funções de pontuação que visam tratar a questão. As mais utilizadas são aquelas que se baseiam: (i) em campos de força; (ii) em dados empíricos; e (iii) em conhecimento acerca das interações moleculares⁷⁰.

Os métodos baseados em campos de força utilizam tanto a energia de interação entre o ligante e o receptor, quanto a energia interna do ligante. Estas energias são obtidas através da combinação do termo de energia de van de Waals que é dado pelo potencial de Lennard-Jones e do termo de energia eletrostática que é dado pela lei de Coulomb. A função *GoldScore*, implementada no programa GOLD, é um exemplo de função baseada em campos de força.

Os métodos empíricos tentam reproduzir dados de energia de ligação de complexos receptor-ligante que foram obtidos experimentalmente. Estes dados são utilizados como coeficientes para os vários termos que compõem a função de pontuação. A função *ChemScore* é um exemplo de função baseada em dados empíricos.

As funções de pontuação baseadas em conhecimento utilizam informações relacionadas à frequência com que certas interações átomo-átomo ocorrem em complexos receptor-ligante. O algoritmo considera positivas as interações que aparecem com frequência nos bancos de dados como o PDB, e considera negativas aquelas que reconhecidamente não são encontradas na natureza.

Por fim, dentre os vários aspectos envolvidos na docagem molecular, o problema da flexibilidade do receptor é de extrema relevância. Várias alterações conformacionais podem ser observadas, desde pequenos movimentos locais, como o rearranjo de cadeias laterais no sítio de ligação, até grandes alterações globais como o deslocamento ou o rearranjo de vários elementos de estrutura secundária. O atual repertório de ferramentas de docagem que aborda este problema ainda é limitado^{65,68}. Métodos como a construção de coleções de rotâmeros, que são encontrados experimentalmente para determinados aminoácidos, e a atribuição de liberdade para os ângulos diedrais de aminoácidos envolvidos no processo de reconhecimento molecular, são empregados pelos programas de docagem. A

dinâmica molecular é um método útil para abordar a questão da flexibilidade do receptor.

3.3 Triagem Virtual

O ensaio virtual é um método computacional de triagem de compostos que é utilizado de forma alternativa ou complementar à triagem experimental (HTS). Há várias estratégias disponíveis para a sua realização, porém elas se dividem basicamente em dois tipos: (i) triagem virtual baseada na estrutura do ligante (LBVS, na sigla inglesa para *ligand-based virtual screening*); e (ii) triagem virtual baseada na estrutura do receptor (SBVS, na sigla inglesa para *structure-based virtual screening*)⁷¹.

No método de LBVS, a estrutura 3D do receptor geralmente não é conhecida. Desta forma, as propriedades moleculares dos compostos ativos são utilizadas como critérios de busca e identificação de moléculas com propriedades químicas similares. A busca pode ser baseada na subestrutura comum dos ligantes conhecidos; em modelos farmacofóricos; ou na forma tridimensional destes ligantes^{8,72}.

No método de SBVS, a estrutura 3D do receptor é conhecida e a docagem molecular de grandes coleções de compostos na cavidade de interação é a estratégia mais utilizada. Em conjunto com a predição do modo de interação, cada molécula docada é classificada através de funções de pontuação. A classificação, em conjunto com outros métodos de avaliação, é utilizada como critério para a síntese ou aquisição comercial das moléculas mais promissoras. Este grupo de compostos com características complementares às do sítio de ligação é então testado experimentalmente para a avaliação de parâmetros como potência e afinidade frente ao alvo molecular^{71,73}. O processo de SBVS é ilustrado na Figura 13.

Figura 13 – Esquema do processo de triagem virtual baseada na estrutura do receptor (SBVS).

O método de SBVS compreende uma série de etapas. Entre estas, a seleção e preparo da estrutura do receptor, seleção e preparo do banco de dados, docagem molecular e análise pós-docagem⁷⁴.

Quanto à primeira etapa, várias estruturas do receptor podem estar disponíveis em bancos de dados como o PDB. Quando há estruturas sem ligante e em complexo com ligantes, todas devem ser avaliadas em relação a critérios como a resolução, e, principalmente, alterações conformacionais decorrentes da interação com ligantes. O preparo do alvo envolve procedimentos que o tornam adequado para o processo de docagem. Envolvem basicamente a adição de átomos de hidrogênio; eliminação de moléculas de água, com exceção daquelas que foram definidas como mediadoras de importantes interações; atribuição correta dos estados de protonação e tautomerização, e atribuição das cargas parciais^{67,75}.

A seguir deve ser selecionada a base de dados de pequenas moléculas que melhor se adequa ao espaço químico definido durante o planejamento da triagem virtual. As bases de dados utilizadas na triagem virtual podem ser coleções

corporativas de compostos, bases externas de vendedores de moléculas, ou grandes bases de dados que organizam em um só sítio bases de vários fornecedores. Um exemplo de base de dados desta última categoria é a ZINC (UCSF, USA), uma interface que funciona como uma grande coleção virtual de pequenas moléculas e que permite a busca e seleção de subconjuntos de compostos de acordo com critérios estabelecidos pelo pesquisador⁷⁶. Estas ferramentas são muito úteis, à medida que reduzem o tamanho da base original que frequentemente possui milhões de compostos, muitos dos quais não apresentam características adequadas ao projeto^{72,76}.

Adicionalmente, quando o tamanho da base de dados torna-se um fator limitante, devido ao alto custo computacional, o número de compostos a ser avaliado pode ser reduzido utilizando-se inicialmente um programa de triagem mais rápido, como por exemplo, o DOCK. Desta forma, moléculas cujo tamanho não é compatível com o sítio de ligação são eliminadas e o subconjunto resultante é submetido a um programa de busca mais complexo, porém mais lento. Nesta segunda triagem os algoritmos genéticos, como o GOLD são uma boa alternativa⁷¹.

A próxima fase consiste na docagem das moléculas da base de dados na cavidade de interação do alvo molecular, na qual a complementaridade entre o receptor e os ligantes é avaliada pelas funções de pontuação, que então fornecem a ordem de classificação para as moléculas docadas^{76,77}.

A triagem virtual em larga escala geralmente envolve milhares ou milhões de compostos e gera um grande volume de dados. Portanto é indispensável que os resultados sejam avaliados após a docagem, ao invés de se proceder automaticamente à aquisição dos compostos melhor classificados. De grande utilidade na análise pós-docagem é a visualização dos complexos obtidos, quanto à complementaridade da superfície dos compostos em relação à superfície da cavidade de interação e a extensão da superfície dos ligantes que se mantém exposta ao solvente (particularmente se forem regiões hidrofóbicas). Outro aspecto que deve ser considerado é se as soluções propostas pelos programas satisfazem os requerimentos estruturais estabelecidos durante o planejamento da triagem, como por exemplo, as hipóteses farmacofóricas ou a presença de interações intermoleculares específicas^{74,75}.

3.4 Dinâmica Molecular

Os fenômenos de reconhecimento molecular envolvem alterações, de maior ou menor extensão, nas conformações do receptor e do ligante. Em alguns casos, os movimentos do receptor são pequenos e o ligante se encaixa em um sítio com pouca mobilidade. Algumas proteínas apresentam grandes alterações na cavidade de interação – da ordem de 10 Å – quando interagem com seus ligantes. Nestes casos, é preciso considerar os diferentes estados conformacionais do sítio de ligação e avaliar quais são os mais representativos⁶⁵. Tipicamente, o ligante estabiliza apenas um subconjunto das várias conformações possíveis do receptor, deslocando o equilíbrio em direção àquela na qual é obtido o ajuste de menor energia. Estas avaliações são importantes na escolha da estrutura cristalográfica mais adequada para a realização de estudos de SBDD⁷⁸. Neste contexto, as simulações por dinâmica molecular (DM) têm se consolidado como uma das ferramentas computacionais mais utilizadas para a compreensão dos fenômenos moleculares durante o processo de reconhecimento molecular⁷⁹.

Desenvolvida no final da década de 1970, a dinâmica molecular utiliza as equações de movimento de Newton, descritas na mecânica clássica, para especificar a posição e a velocidade de cada átomo no decorrer do tempo e assim fornecer a trajetória de todo o sistema molecular. É possível estudar a evolução temporal dos constituintes deste sistema a partir da sequência de posições geradas, e como consequência, determinar suas propriedades dinâmicas⁸⁰.

Inicialmente, uma determinada configuração é estabelecida para os átomos do sistema, que de acordo com potenciais de interação moleculares adquirem movimento. A partir do cálculo das forças que atuam sobre cada partícula em um dado instante, devido às interações com os outros átomos, é possível determinar as posições e velocidades em um instante posterior. Estes cálculos são realizados repetidamente até que as trajetórias moleculares sejam obtidas para um dado intervalo de tempo. O esquema geral do método é ilustrado na Figura 14.

Figura 14 – Esquema do processo de dinâmica molecular.

Os potenciais de interações moleculares, que determinam as forças que agem sobre o sistema, resultam da combinação de potenciais de interação entre átomos ligados e entre átomos não ligados. Os potenciais entre os átomos ligados referem-se às deformações dos ângulos, à rotação e à variação da distância das ligações covalentes. Os potenciais entre átomos não ligados referem-se às interações de van de Waals, modeladas pelo potencial de Lennard-Jones, e às interações eletrostáticas, obtidas pela lei de Coulomb. Com o propósito de reproduzir com fidelidade o comportamento das moléculas, as diversas propriedades – ângulos, comprimento de ligação, cargas parciais – são parametrizadas através de cálculos quânticos ou de dados experimentais. Este conjunto de parâmetros é denominado campo de força e determina a contribuição de cada tipo de interação à dinâmica molecular^{78,79}.

Há vários campos de força e programas utilizados em dinâmica molecular. Dentre estes, o campo de força CHARMM (Universidade de Harvard, USA) e o programa NAMD (Universidade de Illinois, USA) foram utilizados neste trabalho de tese para investigar a dinâmica da aldolase em complexo com inibidores^{81,82}.

Os resultados obtidos através de dinâmica molecular vêm demonstrando boa concordância com resultados experimentais e têm sido utilizados para compreender as razões moleculares para fenômenos observados experimentalmente, como a

inibição da atividade enzimática e os ajustes conformacionais induzidos pela formação dos complexos entre proteínas e ligantes⁸³. Como em qualquer outro método de planejamento molecular, a DM apresenta limitações. Entre as principais destacam-se as interações moleculares, que são aproximações clássicas de fenômenos que na realidade são quânticos. As cargas parciais dos átomos são assinaladas antes da simulação e são mantidas durante todo o processo. Porém, sabe-se que as nuvens eletrônicas em torno dos átomos estão constantemente se deslocando de acordo com o ambiente ao redor, portanto, as cargas seriam descritas de maneira mais fiel através de campos de força polarizáveis, que consideram esta flutuação⁷⁸. Outro fator limitante é o alto custo computacional demandado por simulações de sistemas da ordem de grandeza de milhares de átomos (tamanho típico dos sistemas moleculares que envolvem proteínas). Para reproduzir extensas alterações empreendidas pelos receptores durante o reconhecimento molecular, principalmente os deslocamentos de elementos de estrutura secundária e alterações de estrutura terciária, todos os possíveis estados conformacionais são explorados durante a DM. Alguns destes eventos ocorrem em escalas de tempo que excedem a capacidade computacional disponível. A despeito destas limitações, o método é capaz de dar importantes contribuições para o planejamento de fármacos, principalmente quando combinado com outros métodos de modelagem molecular⁸⁰⁻⁸².

3.5 Relações Quantitativas entre Estrutura e Atividade

Apesar dos grandes avanços realizados na área de SBDD, o processo de planejamento de compostos bioativos também pode ser realizado sem a utilização da estrutura do receptor. Nos métodos de LBDD, novas moléculas são planejadas a partir de informações extraídas de ligantes, substratos e moduladores já caracterizados. Dentre os métodos disponíveis para este fim, as relações quantitativas entre a estrutura e atividade (QSAR) estão entre os mais utilizados^{4,9}.

Os métodos de QSAR têm sido aplicados por décadas no desenvolvimento de relações entre propriedades moleculares e atividade biológica com a finalidade de se obter modelos estatísticos capazes de prever a atividade de novas entidades

químicas. Seu princípio fundamental é o conceito de que as diferenças nas propriedades moleculares são as responsáveis pelas variações na atividade biológica dos compostos de uma mesma série. É uma metodologia útil para direcionar o planejamento de novas moléculas, restringindo o espaço químico a ser explorado e aumentando a probabilidade de identificação de novos compostos ativos^{7,8}.

A atividade biológica, também chamada de propriedade alvo ou variável dependente, geralmente consiste em valores de potência ou afinidade e deve ser padronizada e validada para estudos de QSAR. As propriedades moleculares, empregadas como descritores (variáveis independentes), podem ser calculadas ou obtidas através de estudos experimentais. As correlações que descrevem a variação na propriedade alvo em função dos descritores podem ser derivadas por vários métodos estatísticos. Exemplos importantes incluem a regressão linear múltipla (MLR, na sigla inglesa para *multiple linear regression*) e a regressão por mínimos quadrados parciais (PLS, na sigla inglesa para *partial least squares*)⁸⁴.

Os estudos de QSAR dependem de vários procedimentos críticos para a obtenção de modelos robustos e preditivos: (i) escolha do conjunto de dados, que deve conter moléculas que atuem pelo mesmo mecanismo e que apresentem boa distribuição dos dados biológicos; (ii) tratamento do conjunto de dados para eliminar erros na representação das moléculas e para adequar os dados biológicos ao processo de modelagem; (iii) escolha de descritores adequados para o conjunto de dados; (iv) escolha dos conjuntos treinamento e teste de forma que todo o intervalo da propriedade biológica esteja representado em ambos; e (v) emprego de técnicas de validação interna e externa para que o modelo possa ser aplicado^{21,22}. O resultado de tal processo é a geração de modelos de QSAR, que podem ser aplicados à programas de planejamento para orientar a síntese de novas moléculas^{8,84}.

Uma abordagem que vem ganhando importância é a integração dos modelos de QSAR ao planejamento de SBDD. Modelos farmacofóricos geradas por técnicas de QSAR podem ser utilizados como filtros de bases de dados na seleção de compostos em trabalhos de triagem virtual^{12,85}.

3.6 Holograma QSAR

O método holograma QSAR (HQSAR, na sigla inglesa para *hologram quantitative structure-activity relationships*) se baseia no conceito de impressão digital molecular, no qual a estrutura molecular bidimensional é codificada em notações lineares. Estas notações consistem em sequências numéricas denominadas hologramas moleculares. Uma vez que a atividade biológica é função da estrutura molecular, os modelos de HQSAR são capazes de prever a atividade de uma molécula a partir das informações codificadas no seu holograma molecular⁸⁴⁻⁸⁶.

O holograma é gerado a partir da fragmentação da estrutura bidimensional das moléculas, na qual todos os fragmentos possíveis são considerados e codificados sob a forma de notação numérica. Os fragmentos podem ser lineares, cíclicos, ramificados e se sobrepõem uns aos outros. Cada posição do holograma contém a identidade e o número de fragmentos alocados nesta posição. Fragmentos iguais que ocorrem mais de uma vez são sempre colocados na mesma posição. Desta forma o holograma contém informações a respeito de características moleculares, tais como tipo de átomo, ligações, tamanho e quantidade dos fragmentos gerados. O holograma molecular se diferencia da impressão digital tradicional em relação ao tipo de notação, que nesta última é binária (contém apenas os caracteres 0 e 1)⁸⁷.

Durante o processo de modelagem, vários parâmetros relevantes para a obtenção dos hologramas devem ser definidos, como o comprimento do holograma, o tamanho dos fragmentos e os parâmetros utilizados para distinguir diferentes fragmentos. O comprimento do holograma define o número de posições disponíveis para a distribuição dos diferentes fragmentos moleculares. O tamanho dos fragmentos define o número mínimo e máximo de átomos que cada fragmento pode conter. Os parâmetros de distinção de fragmentos determinam quais características devem ser utilizadas para distinguir os diversos fragmentos.

Os parâmetros de distinção de fragmentos são apresentados na Tabela 4. As características do conjunto de dados devem ser analisadas para que sejam gerados modelos com a combinação adequada de parâmetros. Idealmente, várias combinações são testadas e avaliadas quanto à qualidade dos modelos produzidos.

Tabela 4 – Parâmetros de distinção de fragmentos no método HQSAR.

Parâmetro de distinção de fragmentos	Definição
Átomos (A)	Permite que os fragmentos sejam distinguidos com base nos tipos de átomos.
Ligações (B)	Permite que os fragmentos sejam distinguidos com base nos tipos de ligações: simples, duplas, triplas e aromáticas.
Conectividade (C)	Permite que os fragmentos sejam diferenciados baseando-se no estado de hibridização dos átomos contidos no fragmento.
Hidrogênios (H)	A distinção dos fragmentos é feita com base na presença e número de átomos de hidrogênio.
Quiralidade (Ch)	Permite que os fragmentos sejam distinguidos com base na presença de átomos quirais.
Doador e Receptor (DA)	Esta opção permite que os fragmentos sejam distinguidos com base na presença de átomos receptores ou doadores de interações de hidrogênio.

Após a geração dos hologramas moleculares, o método de regressão PLS correlaciona a atividade biológica com a informação contida nos hologramas, gerando um modelo estatístico de QSAR^{86,88}. Após ser validado, o modelo pode ser utilizado para prever a atividade de moléculas estruturalmente relacionadas e para planejar síntese de novas moléculas. O processo é ilustrado na Figura 15.

Figura 15 – Geração de um modelo HQSAR.

Além de prever a atividade de moléculas não testadas, o modelo HQSAR permite a identificação dos fragmentos moleculares que tem maior relação com a variação da propriedade biológica. Esta informação é disponibilizada através de um mapa de contribuição 2D, que utiliza um código de cores para indicar a contribuição de cada fragmento molecular para a atividade biológica. Estas informações são úteis para orientar a síntese de novas moléculas com propriedades otimizadas^{21,86}.

Os fragmentos destacados pelos mapas de contribuição podem ser utilizados como fragmentos chave para a busca de novos ligantes, em projetos de SBVS. Os grupos funcionais identificados pelos modelos constituem, portanto, modelos farmacofóricos bidimensionais, que podem ser empregados como critério de seleção de moléculas^{87,88}.

3.7 Análise Comparativa dos Campos Moleculares

A análise comparativa dos campos moleculares (CoMFA, na sigla inglesa para *comparative molecular field analysis*) é um dos métodos de QSAR 3D mais

utilizados. Este método se baseia no conceito de que a atividade biológica está diretamente relacionada com as interações estabelecidas durante a formação do complexo receptor-ligante⁸⁴. Estas interações possuem caráter tridimensional, portanto, o método CoMFA considera as moléculas como objetos tridimensionais e utiliza campos moleculares de interação como variáveis independentes. Os campos moleculares de interação são obtidos através do cálculo da energia de interação entre as moléculas do conjunto de dados e uma sonda virtual, que consiste em um determinado átomo com uma carga associada⁶⁶. As diferenças na atividade biológica entre um conjunto de ligantes são desta forma, descritas como função das diferenças nos campos moleculares de interação. Baseando-se no fato de que as interações no complexo ligante-receptor são predominantemente dirigidas por interações eletrostáticas e estereoquímicas, estes são os dois tipos de campos moleculares de interação utilizados pelo método CoMFA^{22,84}.

Para que os campos moleculares de interação sejam calculados de forma equivalente para todas as moléculas, o conjunto de dados precisa ser submetido a um alinhamento estrutural, para que todas as moléculas apresentem conformações semelhantes e a mesma orientação no espaço cartesiano. Esta é a principal dificuldade inerente ao método CoMFA, quando comparado aos métodos de QSAR 2D, como o HQSAR. Há vários métodos de alinhamento molecular disponíveis: os que produzem conformações moleculares de mínima energia; os que utilizam a subestrutura comum como modelo para o alinhamento de todo o conjunto de dados; e aqueles que fazem uso da estrutura do receptor, utilizando a docagem molecular para alinhar as moléculas segundo sua conformação bioativa^{3,89}.

Além da conformação e orientação, outro fator que altera significativamente o cálculo dos descritores moleculares é a atribuição das cargas atômicas parciais, à medida que são calculadas interações eletrostáticas com uma sonda molecular^{8,9}.

As moléculas alinhadas e superpostas são inseridas em uma grade tridimensional reticulada suficientemente ampla, onde em cada ponto de interseção do retículo, são calculadas as energias de interação estereoquímica e eletrostática entre a sonda virtual e os átomos próximos a este ponto. Este processo é realizado até que todos os pontos sejam mapeados pela sonda. A partir destes cálculos, é gerada uma tabela contendo centenas de colunas, onde uma das colunas corresponde ao valor da atividade biológica (variável dependente) e as outras correspondem às energias de interação (variáveis independentes). A cada ponto da

grade são atribuídas duas colunas, uma delas contendo o valor da energia de interação eletrostática e a outra o valor para a interação estereoquímica⁷. A partir desta tabela, o método de regressão PLS gera um modelo estatístico correlacionando a propriedade alvo com os descritores moleculares⁸⁷.

O modelo CoMFA é então submetido a processos de validação, principalmente externa, que é realizada por meio da predição da atividade de um conjunto teste de moléculas. O modelo validado pode ser utilizado no planejamento e predição da atividade de novas moléculas⁸⁹. O processo de construção de um modelo CoMFA é esquematizado na Figura 16.

As relações entre a estrutura e atividade podem ser investigadas adicionalmente através da interpretação dos mapas de contorno eletrostático e estereoquímico que são gerados pelo método. Estes mapas são representações gráficas que mostram regiões do espaço onde alterações nas características estereoquímicas e eletrostáticas se correlacionam de maneira mais significativa com a variação na propriedade biológica^{8,84}. Assim, como nos modelos de HQSAR, os modelos CoMFA podem ser utilizados em conjunto com metodologias baseadas na estrutura do receptor⁸⁸.

Figura 16 – Geração de um modelo CoMFA.

Capítulo 4

Métodos Bioquímicos

A era pós-genômica tem apresentado como característica uma crescente demanda pela produção de proteínas recombinantes em quantidades cada vez maiores e em escala de tempo cada vez menores. Neste contexto, os métodos em biologia molecular encontram vasta aplicação no planejamento de novas substâncias bioativas. Apesar dos numerosos avanços que têm sido realizados, a expressão de altos níveis de proteínas estáveis e funcionais é um dos gargalos em muitos projetos de pesquisa⁹⁰.

Em geral, o que se busca é a obtenção de proteínas livres de substâncias contaminantes, com o objetivo de investigar seus aspectos estruturais e/ou funcionais. Neste sentido, os métodos de obtenção e purificação de proteínas evoluíram enormemente, e uma grande variedade de proteínas alvo pode ser obtida através destes métodos⁹¹.

4.1 Expressão, Purificação e Cristalização de Proteínas Heterólogas

A obtenção de proteínas foi largamente afetada pela introdução das técnicas de recombinação do DNA, na qual proteínas de um determinado organismo podem ser produzidas por células de outro organismo. Esta abordagem, denominada expressão heteróloga ou recombinante, dispõe de diferentes sistemas para a produção de proteínas em células de bactérias, fungos, insetos e mamíferos⁹².

Na maioria dos laboratórios que se dedica à produção de proteínas, seja em pequena escala, para pesquisas de caráter exploratório, ou em grande escala, para a realização de triagens de HTS, a produção de proteínas recombinantes em bactérias é uma das técnicas mais utilizadas. Isso ocorre devido ao seu baixo custo relativo, facilidade de cultura e manipulação, rápido crescimento e maior rendimento em relação a outros sistemas de expressão^{90,93}. O maior rendimento de expressão em células bacterianas foi particularmente importante para o desenvolvimento da cristalografia de proteínas, que geralmente requer quantidades significativas de proteína em alto grau de pureza. Porém, a recuperação de proteínas de forma solúvel e com atividade, é um dos problemas do uso de bactérias, quando modificações pós-traducionais são necessárias. Adicionalmente, proteínas

expressas em grande quantidade podem precipitar e formar corpos de inclusão, tornando-se inacessíveis⁹¹.

Um sistema de expressão heteróloga em bactérias consiste basicamente em um vetor de expressão, que geralmente é um plasmídeo (uma molécula de DNA circular), que contém os elementos necessários à expressão da proteína alvo: um gene promotor; sequências gênicas reguladoras; o gene codificador da proteína de interesse; e genes que conferem resistência a determinados antibióticos.

A escolha do sistema de expressão deve levar em consideração o rendimento, os recursos laboratoriais necessários e o processo de purificação subsequente que será utilizado para isolar a proteína alvo⁹⁴.

A localização do produto dentro da célula hospedeira tem grande impacto na escolha dos métodos de purificação que serão utilizados. As bactérias podem armazenar a proteína produzida no citoplasma, podem secretá-la no meio de cultura, armazená-la no espaço periplasmático ou estocá-la de forma insolúvel em corpos de inclusão. Este aspecto afeta a quantidade de contaminantes celulares que terão que ser removidos para a obtenção da proteína pura^{92,93}.

De grande utilidade tem sido a utilização de sistemas que produzem a proteína alvo fundida com uma sequência de aminoácidos que é utilizada como marcador específico para a proteína recombinante. Um destes marcadores é a cauda de histidina, que consiste em uma sequência de quatro a dez resíduos de histidina, introduzida em uma das regiões terminais da proteína⁹⁵.

Uma vez produzida, a proteína de interesse deve ser isolada dos outros componentes celulares. Uma variedade de métodos de purificação está disponível e a sua escolha deve ser guiada pelo tipo de sistema de expressão utilizado e pela finalidade da purificação.

Quando a proteína está fusionada a um peptídeo marcador, como a cauda de histidina, as propriedades deste marcador podem ser exploradas na purificação. A cromatografia de afinidade é um método simples e adequado para a maioria dos marcadores utilizados. É baseado na ligação reversível e seletiva da proteína alvo a uma matriz sólida (fase estacionária, geralmente composta por microesferas de sefarose). A cauda de histidina tem alta afinidade por íons níquel e cobalto, que são imobilizados nas esferas de sefarose através de um quelante. Em consequência, a proteína marcada se liga ao íon e fica retida na fase estacionária, enquanto os contaminantes passam pela coluna cromatográfica. Este tipo de cromatografia de

afinidade é denominada cromatografia de afinidade por metal imobilizado (IMAC, na sigla inglesa para *immobilized metal affinity chromatography*)⁹⁶. A purificação pelo método IMAC é esquematizada na Figura 17.

Figura 17 – Purificação de uma proteína recombinante pelo método IMAC.

Quando um dos objetivos de purificação é a cristalização da proteína de interesse, geralmente é necessário que uma segunda etapa de purificação seja realizada. A cromatografia de afinidade é capaz de eliminar grande parte dos contaminantes, porém, algumas proteínas permanecem de forma residual na amostra. Portanto, a nova etapa de purificação tem por objetivo eliminar qualquer proteína contaminante que ainda se encontra na amostra. Geralmente a purificação por exclusão de tamanho (gel filtração) é utilizada nesta segunda etapa, separando as proteínas com base no seu peso molecular. Quanto maior o peso molecular da proteína menor é o seu tempo de eluição. Portanto, ao passar pela coluna de gel filtração as diferentes proteínas são separadas em diversas frações de acordo com o seu peso molecular⁹⁷.

O uso adequado destas técnicas resulta na obtenção de quantidades significativas de proteína e em grau de pureza suficiente para a realização de vários estudos de química medicinal, como os estudos de cinética enzimática, a avaliação de inibidores e os estudos estruturais. Com esta finalidade, foram desenvolvidos protocolos de expressão, purificação e cristalização para a aldolase de *T. brucei*.

O sistema de expressão foi baseado em células da bactéria *Escherichia coli* da linhagem BL21. O vetor de expressão utilizado foi o plasmídeo pET28a, no qual foi clonado o gene codificador para aldolase de *T. brucei*^{92,95}. As células foram

obtidas através de colaboração com o Prof. Paul Michels do *Research Unit for Tropical Diseases (TROP) - Université Catholique de Louvain*, Bélgica.

O método de purificação escolhido foi a combinação da purificação por afinidade com a purificação por exclusão de tamanho. Na primeira etapa foi utilizada a IMAC, explorando a presença da cauda de seis resíduos de histidina ligada à aldolase. Como fase estacionária, foram utilizadas esferas de sefarose recobertas por íons cobalto. Na segunda etapa foi utilizada a gel filtração para separar a aldolase com base no seu peso molecular. A seguir, foram realizadas triagens para definir as condições mais favoráveis para a cristalização da aldolase, com o objetivo de estabelecer um protocolo de cristalização para a enzima. O método de triagem utilizado foi o de matriz esparsa, no qual são testadas várias combinações de parâmetros que são conhecidas por favorecer a cristalização de um grande número de proteínas⁹⁸.

A triagem foi realizada através do sistema automatizado *Cartesian Dispersing Systems*[®] (Genomic Solutions), com a utilização dos kits de cristalização: Classics I[®], Classics II[®], PEG I[®], PEG II[®], MPD[®], CRYOS[®] (QIAGEN) e SaltRX I[®] (Hampton). Desta forma, foram testadas cerca de duas mil condições de cristalização. Os cristais foram obtidos através do método de difusão de vapor por gota pendente⁹⁹.

4.2 Cinética Enzimática e Métodos de Análise

Reações catalisadas por enzimas podem ser estudadas através de vários métodos com a finalidade de explorar diferentes aspectos da atividade enzimática. As análises cinéticas são a maneira mais comum para elucidar aspectos relacionados ao mecanismo de reação e à interação das enzimas com inibidores¹⁰⁰. A cinética do estado estacionário é amplamente utilizada nestes estudos, devido à conveniência para a interpretação de parâmetros importantes nas reações enzimáticas. Este tipo de análise baseia-se na análise da velocidade inicial (v_0) de uma dada reação enzimática. A velocidade da reação é dada pela variação da concentração de substrato ([S]) ou de produto ([P]) em relação ao tempo. Se apenas o tempo inicial da reação é observado, esta variação é aproximadamente linear e a

velocidade é constante, porém, conforme o substrato é consumido, a velocidade da reação diminui progressivamente¹⁰¹.

A velocidade inicial é afetada principalmente pela concentração de enzima ($[E]$) e de substrato ($[S]$). Em uma determinada $[E]$ e em baixas concentrações de $[S]$, a velocidade aumenta linearmente com o aumento de $[S]$. Com o aumento progressivo de $[S]$, a velocidade aumenta a uma taxa cada vez menor, até que finalmente é alcançado um platô onde a velocidade máxima ($V_{\max}$) é atingida, gerando uma função hiperbólica da velocidade em relação à $[S]$ (Figura 18)¹⁰². Isto ocorre devido à natureza das reações enzimáticas, na qual a enzima e o substrato formam o chamado complexo enzima-substrato (ES), que posteriormente se separa regenerando a enzima livre e formando o produto de reação (Figura 19).

De acordo com este mecanismo geral de catálise enzimática, a velocidade da reação se torna função da concentração do complexo ES. A velocidade máxima é atingida quando virtualmente todas as moléculas de enzima estão sob a forma de complexo ES. Sob estas condições a enzima está saturada pelo substrato. Todas as enzimas que seguem este padrão hiperbólico de saturação pelo substrato são estudadas sob o prisma da cinética de Michaelis-Menten¹⁰³.

Figura 18 – Velocidade inicial de uma reação enzimática em função da concentração de substrato.

Figura 19 – Esquema geral das reações enzimáticas.

Quando a enzima é colocada na presença de um grande excesso de substrato, a reação rapidamente atinge este estágio de saturação, denominado estado estacionário, onde a concentração do complexo ES se mantém constante.

A relação que descreve a dependência de v_0 em relação à $[S]$ é a equação de Michaelis-Menten (Equação 1). Uma relação importante emerge desta equação: a concentração de substrato necessária para que a velocidade inicial atinja metade da velocidade máxima é equivalente ao valor de K_M . Portanto, a constante de Michaelis-Menten (K_M) pode ser utilizada como uma medida relativa da afinidade do substrato pela enzima. Os parâmetros V_{max} e K_M podem ser obtidos experimentalmente e são importantes na avaliação de inibidores enzimáticos¹⁰⁴.

$$v = \frac{V_{max}[S]}{K_M + [S]} \quad (1)$$

A velocidade inicial das reações enzimáticas é medida através do acompanhamento de um determinado sinal que se correlaciona com o consumo de substrato ou com a formação de produto. O sinal medido geralmente se baseia em uma determinada propriedade físico-química característica do substrato ou do produto da reação. A maioria dos ensaios enzimáticos baseia-se em métodos espectroscópicos. Os tipos de detecção espectroscópica podem ser: (i) direto, no qual a própria concentração de substrato ou produto pode ser medida no decorrer do tempo; (ii) indireto, no qual os analitos não produzem um sinal que os distingue dos outros componentes presentes no ensaio, sendo necessário a adição de uma segunda reação não-enzimática que proporcione um sinal que possa ser medido; e (iii) ensaios acoplados, nos quais a reação de interesse é acoplada às reações enzimáticas subsequentes, que produzem ou consomem um determinado reagente que pode ter sua concentração monitorada no decorrer do tempo. Em um ensaio acoplado típico, o produto da reação enzimática de interesse é o substrato para a reação subsequente. As enzimas que são adicionadas ao ensaio para viabilizar um meio de detecção adequado são as enzimas acessórias^{100,101}.

Em teoria, várias reações podem ser acopladas através de intermediários comuns, porém limitações práticas surgem com o aumento do número de reações envolvidas. Cada enzima acrescentada adiciona dificuldades, à medida que as condições do meio reacional devem ser compatíveis com os requerimentos, às vezes divergentes, de cada enzima. Outro aspecto crítico para que dados corretos sejam coletados, é a necessidade de que a velocidade seja limitada pela reação de interesse e não pelas reações acessórias. Por este motivo, ensaios acoplados requerem que as concentrações das enzimas acessórias sejam muito maiores do que a concentração da enzima alvo. Também é necessário assegurar que as medidas sejam realizadas apenas no intervalo correspondente ao estado estacionário (porção linear da reação)¹⁰².

Os métodos espectrofotométricos são muito utilizados para monitorar as reações enzimáticas. Baseiam-se na capacidade do analito em absorver radiação eletromagnética em determinados comprimentos de onda. A grande utilidade destes

métodos como ferramenta analítica é o fato de que a absorção de energia radiante por uma molécula em determinado comprimento de onda pode ser correlacionado à sua concentração, de acordo com a lei de Beer (Equação 2)¹⁰⁵.

$$A = \epsilon cl \quad (2)$$

Sendo A , a absorbância da amostra em determinado comprimento de onda; c , a concentração do analito; l , o caminho ótico da amostra; e ϵ , a absorvidade molar da molécula. Utilizando a lei de Beer, a variação da absorbância em um determinado intervalo de tempo corresponde à velocidade da reação.

Sob o ângulo da cinética do estado estacionário, foi desenvolvido para a aldolase de *T. brucei* um ensaio cinético padrão para a avaliação de candidatos a inibidores selecionados por triagem virtual. Os ensaios foram conduzidos à 25°C, pH 7,6, utilizando-se o método de espectrofotometria no visível, com comprimento de onda de 405 nm. O método consiste em um ensaio enzimático acoplado, envolvendo as enzimas acessórias: triose-fosfato isomerase (TIM) e glicerol-3-fosfato desidrogenase (GDH). Um dos substratos da GDH, a tionicotinamida adenina dinucleotídeo (T-NADH) é a espécie que foi utilizada para monitorar a velocidade da reação, por apresentar uma forte absorção da radiação a 405 nm. Quando convertido a sua forma oxidada (T-NAD⁺), a molécula deixa de absorver neste comprimento de onda. O esquema do ensaio acoplado utilizado para avaliar os inibidores adquiridos é ilustrado na Figura 20.

Figura 20 – Esquema do ensaio acoplado utilizado para monitorar a velocidade da reação da aldolase. FBP (frutose-bifosfato); GAP (gliceraldeído-3-fosfato); DHAP (diidroxiacetona fosfato); G3P (glicerol-3-fosfato). A numeração corresponde às reações catalisadas pela aldolase (1); triosefosfato-isomerase (2) e glicerol-3-fosfato desidrogenase (3).

Capítulo 5

Padronização dos Protocolos Bioquímicos

5.1 Expressão, Purificação e Cristalização da Aldolase

Para que seja possível a avaliação experimental de inibidores da aldolase, se faz necessário que a enzima esteja disponível em quantidade adequada e livre de contaminantes. Com esta finalidade, foram desenvolvidos protocolos de expressão, purificação e cristalização para a aldolase de *T. brucei*.

Células da bactéria *Escherichia coli* da linhagem BL21, possuindo o plasmídeo pET28a com o gene codificador para a aldolase de *T. brucei* foram cedidas pelo Prof. Paul Michels – *Research Unit for Tropical Diseases (TROP)* - *Université Catholique de Louvain*, Bélgica.

A partir deste sistema de expressão, foi desenvolvido primeiramente um protocolo de expressão e purificação para a enzima. O protocolo final obtido é descrito a seguir.

Células de *E. coli* foram inoculadas em 40 mL de meio de cultura microbiológico Luria Bertani (LB) contendo o antibiótico canamicina a uma concentração final de 30 µg/mL, para impedir o crescimento de bactérias contaminantes. Após incubação por 14 h a uma temperatura de 37°C, 40 mL da cultura bacteriana foram inoculados em 1 L de meio LB pH 7,5, contendo 30 µg/mL de canamicina, e incubadas à 37°C, sob rotação de 200 rpm, até atingirem uma densidade ótica de 0,6 (600 nm). Nesta fase do crescimento bacteriano, foi adicionado ao meio de cultura, isopropil-β,D-tiogalactopiranosídeo (IPTG), à uma concentração final de 1 mM, para induzir a produção da aldolase. Após 4 h da adição de IPTG, a incubação foi interrompida e a cultura foi centrifugada a 12000 g por 15 min, a 4°C para separar as células e o meio de cultura. O sobrenadante foi descartado e a massa bacteriana foi ressuspensa em 30 mL de solução tampão a um pH de 8,0. A constituição deste tampão (tampão de lise) é apresentada na Tabela 5.

Tabela 5 – Composição do tampão de lise.

Trietanolamina	100 mM
NaCl	300 mM
Imidazol	4 mM
TCEP	1 mM
Glicerol	10% (v/v)
Inibidores de Protease	
PMSF	0,2 mM
Pepstatina	1 µM
Leupeptina	10 µM

TCEP, tris(2-carboxietil)fosfina, um redutor de ligações dissulfeto; PMSF, fluoreto de fenilmetilsulfonila; v/v, volume/volume.

Em seguida, as células foram lisadas pelo método de diferença de pressão, através de duas passagens da suspensão por uma prensa do tipo francesa, ajustada para atingir uma pressão de 1400 psi.

Após a lise celular foram adicionadas 1000 unidades da enzima benzonase para degradar os ácidos nucleicos presentes no lisado. O lisado foi mantido sob agitação por 30 min à temperatura ambiente.

A seguir foi adicionado sulfato de protamina ao lisado, a uma concentração final de 0,5 mg/mL, para promover a precipitação dos fragmentos de DNA e de proteínas contaminantes. Novamente o lisado foi mantido sob agitação por 30 min à temperatura ambiente.

Após o tratamento com benzonase e sulfato de protamina, o lisado foi centrifugado, a 12000 g a uma temperatura de 4°C por 30 min, para separar a aldolase dos restos celulares, ácidos nucleicos e proteínas contaminantes. O precipitado foi descartado e o sobrenadante contendo a aldolase foi mantido em gelo.

Foi adicionado a uma coluna de gravidade, 4 mL de resina TALON®. A resina foi equilibrada com 20 mL de tampão de lise. Em seguida a amostra contendo a aldolase foi aplicada à coluna, e posteriormente a coluna foi lavada com 40 mL de tampão de lise.

Em seguida foi aplicado à coluna 20 mL de tampão de lavagem, que possui a mesma constituição do tampão de lise, exceto pela concentração de imidazol que no

tampão de lavagem é de 40 mM. O tampão de lavagem tem a finalidade de eluir proteínas contaminantes que se ligam maneira mais fraca à coluna.

A eluição da aldolase foi feita com a aplicação de 15 mL de tampão de eluição à coluna. Este tampão possui a mesma composição do tampão de lise, exceto pela concentração de imidazol, que no tampão de eluição é de 200 mM.

Após a eluição, a aldolase foi concentrada para 6 mg/mL, por centrifugação, através de uma membrana com poros de 30 kDa.

Este protocolo proporciona a obtenção de aproximadamente 15 mg de proteína, com grau de pureza adequado para a realização de ensaios de cinética enzimática.

A Figura 21 consiste em um gel obtido por eletroforese em gel de poliacrilamida (SDS-PAGE), após a purificação por afinidade, onde é possível verificar a diminuição da quantidade de proteínas contaminantes com o progresso da purificação¹⁰⁶.

Figura 21 – Gel de poliacrilamida demonstrando a progressão da purificação. Numeração: (1) marcador de peso molecular; (2) lisado (3) precipitação com sulfato de protamina; (4) sobrenadante; (5) proteínas contaminantes após a passagem pela coluna; (6) aplicação de tampão de lise; (7) aplicação de tampão de lavagem; (8) eluição da aldolase. Peso molecular em kDa.

Medidas de atividade enzimática realizadas durante o processo de purificação foram realizadas com a finalidade de se verificar a eficiência do protocolo em purificar a enzima e mantê-la com atividade.

Os resultados deste estudo estão expressos na Tabela 6, onde podemos verificar que a atividade específica da enzima aumenta com o decorrer das etapas de purificação, indicando a eficiência de cada etapa. Ao final do processo, após a passagem pela coluna de afinidade, a atividade específica aumentou cerca de dez vezes.

Tabela 6 – Etapas da purificação e seus respectivos valores de atividade enzimática, atividade específica e fator de purificação.

Fração	Atividade/mL	Atividade Específica	Fator de Purificação
Lisado	8,4	3,1	
Sulfato de Protamina	12	6,2	2
IMAC	42,5	31,4	10,12

Para os estudos de cinética enzimática apenas a etapa de purificação por afinidade é suficiente, porém para os estudos cristalográficos é necessário um grau de pureza ainda maior, pois a presença de qualquer contaminante pode impedir a formação dos cristais.

Com a finalidade de estabelecer um conjunto de parâmetros ideais para a cristalização da aldolase, a etapa de purificação por cromatografia de afinidade foi seguida por uma etapa de purificação por exclusão de tamanho.

A etapa de gel filtração foi realizada com uma coluna Superdex[®] Hiloal 200 16/60 de 120 mL. A coluna foi equilibrada com solução tampão contendo 2-amino-2-hidroximetil-propano-1,3-diol (TRIS) à 100 mM e NaCl à 200 mM, pH 7,6. O tempo de purificação foi estabelecido em 3 h, com volume total de 180 mL e fluxo de 1 mL por minuto. A seguir foram aplicados à coluna 2 mL de solução a 6 mg/mL de aldolase. A eluição da enzima ocorreu na fração correspondente a 70 mL, como pode ser visualizado pelo do pico de absorbância da proteína durante a gel filtração (Figura 22).

Figura 22 – Absorbância durante a gel filtração.

Após a gel filtração, a enzima foi concentrada para 12 mg/mL para a realização de triagens para estabelecer as condições mais favoráveis de cristalização.

Foram testadas cerca de duas mil condições de cristalização. As formações compatíveis com cristais mais promissoras foram observadas nas condições descritas na Tabela 7.

Tabela 7 – Condições favoráveis para o crescimento de cristais da aldolase.

	Kit	Tampão	Precipitante
Condição 1	Classics II	100mM BIS-TRIS, pH 5,5	25% PEG 3350*
Condição 2	SaltRX I	100mM Acetato de sódio, pH 4,6	3,2 M Cloreto de sódio
Condição 3	SaltRX I	100mM Acetato de sódio, pH 4,6	3,5 M Formato de sódio

*Polietilenoglicol, porcentagem expressa em peso/volume.

Ao verificar-se o crescimento de cristais, foi realizado um processo de ajuste dos principais parâmetros (tampão, pH, agente precipitante e concentração de enzima) para o estabelecimento de um protocolo de cristalização.

Estes novos experimentos resultaram em um protocolo de cristalização para a aldolase, pelo qual são obtidos cristais da enzima a partir de três dias de incubação, nas seguintes condições: (i) tampão acetato de sódio a 100 mM; (ii) pH 4,6 a 4,8; agente precipitante PEG 3350, 25 a 27%; e (iii) concentração de proteína de 5 mg/mL. A Figura 23 mostra alguns dos cristais obtidos nestas condições.

Figura 23 – Cristais obtidos para a aldolase de *T. brucei*.

Os dados para estes cristais foram coletados em um difratômetro Rigaku Micromax-007HF, com detecção através de uma placa de imagem Rigaku R-AXIS IV++. A análise dos dados de difração coletados (Figura 24) revelou a presença de padrões de difração superpostos, provavelmente devido à presença de cristais superpostos com orientações ligeiramente diferentes, o que impossibilitou a indexação das reflexões obtidas. Não obstante, o protocolo de cristalização obtido é extremamente útil por permitir o crescimento de cristais com capacidade de difração (resolução de 2 Å), e porque não há disponível na literatura um protocolo de cristalização para a aldolase de *T. brucei*, compatível com o sistema de expressão utilizado nesta tese (pET28a).

Figura 24 – Padrão de difração dos cristais da aldolase de *T. brucei*.

5.2 Padronização dos Ensaios de Cinética Enzimática

A avaliação de compostos bioativos depende do estabelecimento de protocolos bem definidos de ensaios cinéticos e da determinação de parâmetros essenciais como a velocidade máxima e a constante de Michelis-Menten. Neste contexto, após o estabelecimento de um protocolo de obtenção da enzima, foi desenvolvido para a aldolase de *T. brucei* um ensaio cinético padrão para a avaliação de candidatos a inibidores. O método consiste em um ensaio enzimático acoplado de espectrofotometria no visível, no qual o decréscimo da concentração do T-NADH é monitorado a um comprimento de onda de 405 nm.

Na primeira etapa foi avaliado se a correlação linear entre a velocidade inicial da reação e a concentração de enzima era mantida. A Figura 25 consiste no gráfico da velocidade inicial em função da concentração de aldolase.

Figura 25 – Gráfico da atividade enzimática em função da concentração de enzima.

Os dados demonstram que a velocidade inicial da reação é uma função linear da concentração de enzima.

Os valores de K_M e V_{max} foram obtidos através de medidas da velocidade inicial da reação em diferentes concentrações de substrato. Foi empregada uma faixa de variação da concentração de substrato de aproximadamente 40 vezes. Os dados foram ajustados à equação de Michaelis-Menten (Equação 1) pelo método de regressão não-linear, implementado no programa SigmaPlot 9.1[®]. O gráfico resultante (Figura 26) revelou uma função hiperbólica da velocidade em relação à concentração de substrato.

Figura 26 – Gráfico da velocidade da reação em função da concentração de substrato.

A figura demonstra que a aldolase apresenta o fenômeno de saturação pelo substrato, com a velocidade chegando a um platô em concentrações de substrato de aproximadamente 120 μM . A transformação linear destes dados foi utilizada para gerar o gráfico de Lineweaver-Burk, ou duplo-recíproco (ilustrado na Figura 27), a partir do qual foram obtidos os valores de K_M e $V_{\max}$.

Foi determinado um valor de $K_M = 38,6 \mu\text{M}$ e de $V_{\max} = 8,6 \mu\text{M}/\text{min}$. A constante de Michaelis-Menten diverge dos valores encontrados na literatura. Para a enzima recombinante clonada no plasmídeo pET3a, o valor de K_M é de 16,8 μM e para a enzima nativa este valor é de 9,2 μM . O aumento no valor de K_M para a enzima recombinante com cauda de histidina (38,6 μM) pode ser considerado um efeito da adição destes resíduos de aminoácidos à estrutura da enzima, resultando em uma discreta diminuição da afinidade da aldolase pelo substrato.

Figura 27 – Gráfico duplo-recíproco gerado a partir da transformação linear dos dados da velocidade em função de $[S]$.

Estes parâmetros serviram como base para os experimentos que se seguiram para estabelecer o protocolo final do ensaio. A Tabela 8 apresenta as condições padrão do ensaio cinético para a aldolase.

Tabela 8 – Condições padrão estabelecidas para os ensaios de cinética e inibição enzimática.

Reagente	Concentração
Trietanolamina	25 mM
NaCl	10 mM
Frutose-bifosfato	100 μ M
T-NADH	70 μ M
Aldolase	2 nM
GDH	1 U
TIM	1,5 U

O ensaio cinético padrão para a aldolase de *T. brucei* possui boa linearidade em um intervalo de tempo suficiente para a realização das medidas. A faixa de absorbância do ensaio (em torno de 0,4) está dentro do limite no qual a lei de Beer é

válida (abaixo de 1,0). A Figura 28 apresenta o decréscimo da concentração de T-NADH, nas condições padrão do ensaio.

A partir deste método foi possível realizar a avaliação bioquímica de compostos selecionados nos ensaios virtuais, quanto à sua capacidade de inibir a atividade da aldolase de *T. brucei*.

Figura 28 – Medida cinética da atividade da aldolase realizada através da detecção da queda na concentração do cofator T-NADH, nas condições padrão do ensaio.

Capítulo 6

Planejamento de Inibidores da Aldolase

6.1 Relações Quantitativas entre Estrutura e Atividade para Inibidores da Aldolase

O processo de planejamento de inibidores da aldolase de *T. brucei* foi realizado inicialmente pela busca de inibidores da enzima que fossem adequados para o emprego de métodos de planejamento baseados na estrutura de ligantes (LBDD). Com esta finalidade, uma série compostos com atividade determinada frente à aldolase de músculo de coelho foi selecionada para o desenvolvimento de modelos de QSAR, empregando-se os métodos HQSAR e CoMFA. Posteriormente, os modelos de QSAR desenvolvidos foram integrados a métodos de planejamento baseados na estrutura do receptor (SBDD), através da docagem molecular do conjunto de inibidores na cavidade de interação da aldolase e da realização de simulações de dinâmica molecular. Os modelos de conformação bioativa propostos pelos estudos de docagem molecular foram utilizados para o alinhamento estrutural do conjunto de dados, etapa necessária para o desenvolvimento dos modelos CoMFA¹⁰⁷.

Os modelos de QSAR 2D e 3D desenvolvidos, em conjunto com as informações produzidas pelos mapas de contribuição de HQSAR, mapas de contorno CoMFA e pelos estudos de SBDD (docagem e dinâmica molecular) forneceram informações úteis acerca das bases estruturais determinantes para a inibição da atividade da aldolase de *T. brucei*.

Os modelos de HQSAR e CoMFA foram desenvolvidos e avaliados empregando-se a plataforma computacional SYBYL 8.0® (sistema operacional Linux). Os protocolos de docagem foram realizados com a utilização do programa GOLD 5.0.

6.2 O Conjunto de Dados

O conjunto de dados utilizado nos estudos de QSAR é constituído por 38 inibidores, que foram selecionados a partir de três publicações disponíveis na literatura¹⁰⁸⁻¹¹⁰. As moléculas são ésteres fosfóricos análogos do substrato da

aldolase (frutose-1,6-bifosfato). A diversidade química deste conjunto de dados é representada pela diferença no número de átomos da cadeia hidrocarbônica, pela presença de um ou dois ésteres fosfóricos nas extremidades das moléculas e pela identidade do grupo polar posicionado entre os grupos fosfato (Figura 29).

Figura 29 – Subestrutura representativa das moléculas do conjunto de dados utilizado nas análises de QSAR.

Os valores da propriedade biológica (K_i , constante de dissociação do complexo enzima-inibidor) variam de $5,0 \cdot 10^{-2}$ M a $5,5 \cdot 10^{-5}$ M (um fator de potência da ordem de 1000). Estes dados foram produzidos pelo mesmo grupo de pesquisa e sob as mesmas condições experimentais, o que satisfaz um dos requerimentos indispensáveis para a geração de modelos de QSAR¹¹¹. Com o propósito de colocar os dados na escala adequada para a construção dos modelos, os valores de K_i foram convertidos em pK_i ($-\log K_i$). As estruturas e os valores da propriedade biológica (pK_i) para o conjunto completo de inibidores são apresentados na Tabela 9.

A estrutura tridimensional dos inibidores foi gerada utilizando-se os parâmetros geométricos definidos na plataforma computacional SYBYL 8.0¹¹². Cada estrutura foi energeticamente minimizada e suas cargas atômicas foram calculadas pelo programa de cálculo quântico semi-empírico, MOPAC 6.0. Foi empregado o método AM1, com o cálculo das cargas realizado pelo método de potencial eletrostático^{113,114}.

Tabela 9 – Estrutura do conjunto de dados utilizado nos estudos de QSAR e os valores de pK_i .

Estrutura		pK_i (M)	Estrutura		pK_i (M)
1		2,44	2		2,58
3		2,85	4		3,27
5		3,14	6		3,48
7		3,28	8		3,00
9		3,48	10		4,26
11		3,60	12		3,24
13		1,67	14		2,10
15		2,14	16		1,53
17		1,31	18		1,92

continua

conclusão

19		2,26	20		2,06
21		1,64	22		1,72
23		1,88	24		2,22
25		1,82	26		1,88
27		1,79	28		1,88
29		2,30	30		2,30
31		1,96	32		2,52
33		2,18	34		1,52
35		1,30	36		1,77
37		2,30	38		2,11

6.3 Homologia entre Aldolases e Seleção do Conjunto de Dados

A sequência de aminoácidos da aldolase de *T. brucei* possui 49% de identidade em relação à enzima de músculo de coelho e à enzima humana (Figura 30). As três enzimas também se enovelam da mesma maneira, mantendo os elementos de estrutura secundária e terciária, o que lhes confere o mesmo padrão global de estrutura tridimensional (Figura 31). Como pertencem às aldolases de classe I, apresentam o mesmo mecanismo de reação⁵⁹.

Figura 30 – Alinhamento da sequência de aminoácidos das enzimas de *T. brucei*, músculo de coelho e humana. A numeração corresponde à posição dos aminoácidos da enzima do parasita. Os aminoácidos conservados estão destacados em preto.

Figura 31 – Alinhamento estrutural das enzimas de *T. brucei* (cinza), músculo de coelho (amarelo) e humana (verde).

Uma análise detalhada das enzimas de coelho e de *T. brucei* revelou uma grande conservação dos aminoácidos do sítio ativo, conforme a Figura 32. A conservação do padrão estrutural e do mecanismo de reação conduz à hipótese de que inibidores da aldolase de coelho reconheçam a cavidade de interação da enzima do parasita da mesma forma, adotando a mesma conformação bioativa. De fato, isto ocorre para o manitol-bifosfato (MBP), um inibidor das enzimas de coelho e de *T. brucei*. Este inibidor é estruturalmente relacionado ao conjunto de dados utilizado nos estudos de QSAR¹¹⁵.

Figura 32 – Sobreposição dos resíduos de aminoácidos do sítio ativo da aldolase de coelho (amarelo) e de *T. brucei* (cinza). O inibidor manitol-bifosfato está representado em azul.

O modo de ligação do MBP é ilustrado na Figura 33, no qual a conformação do inibidor em complexo com a aldolase de coelho (código PDB: 1ZAJ) obtida por cristalografia de raios-x pode ser comparada com a conformação do inibidor no sítio ativo da enzima de *T. brucei* obtida por docagem. Nota-se um modo de ligação bastante similar frente às duas enzimas. Este fato, somado aos aspectos relacionados à conservação da estrutura e ao mecanismo de ação das aldolases de coelho e de *T. brucei*, torna a seleção deste conjunto de dados, adequada para a realização dos estudos de QSAR e de SBDD desenvolvidos nesta tese.

Figura 33 – Conformação cristalográfica do inibidor MBP (azul) em complexo com a aldolase de coelho (amarelo). Solução de docagem para o mesmo inibidor (magenta) no sítio ativo da aldolase de *T. brucei* (cinza).

6.4 Holograma QSAR

O conjunto de 38 inibidores da Tabela 9 foi utilizado para a geração de modelos HQSAR. A obtenção de modelos consistentes depende de uma distribuição balanceada da diversidade estrutural e dos valores da propriedade biológica nos conjuntos treinamento e teste. Com este objetivo, análises estatísticas de agrupamento foram realizadas para selecionar os conjuntos treinamento e teste¹¹⁶. Trinta moléculas foram utilizadas como conjunto treinamento para a construção dos modelos e os oito inibidores remanescentes foram utilizados como conjunto teste para realizar a validação externa dos modelos gerados (compostos **1**, **3**, **6**, **9**, **14**, **17**, **22** e **31**). A distribuição da propriedade biológica nos dois conjuntos é representada na Figura 34.

Figura 34 – Distribuição da propriedade biológica nos conjuntos treinamento e teste.

As análises foram realizadas utilizando-se os seguintes comprimentos de holograma: 53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353 e 401. Os hologramas moleculares foram gerados através dos seguintes parâmetros para distinção de fragmentos: átomos (A), ligações (B), conexões (C), átomos de hidrogênio (H), quiralidade (Ch), e doadores e aceptores de ligações de hidrogênio (DA).

A seguir, foram utilizadas as seguintes combinações destes parâmetros, com tamanho de fragmentos de 4 a 7 átomos: A/B, A/B/C, A/B/C/H, A/B/C/H/Ch, A/B/C/H/Ch/DA, A/B/H, A/B/C/Ch, A/B/DA, A/B/C/DA, A/B/H/DA, A/B/C/Ch/DA, A/B/C/H/DA e A/B/H/Ch/DA.

Após a obtenção dos hologramas moleculares para o conjunto treinamento, o método de regressão PLS foi empregado para correlacionar os valores de pK_i aos hologramas¹¹⁷. A capacidade preditiva dos modelos foi testada pelo método de validação cruzada “deixe um de fora” (LOO, na sigla inglesa para *leave-one-out*). A Tabela 10 mostra os resultados para as diferentes combinações de parâmetros utilizadas.

Tabela 10 – Resultados para os diferentes parâmetros de distinção de fragmentos, utilizando o tamanho de fragmento padrão de 4 a 7 átomos e o método de validação LOO.

Distinção de Fragmentos	Parâmetros Estatísticos				
	q^2	r^2	SEE	HL	N
A/B	0,52	0,89	0,26	61	6
A/B/C	0,31	0,46	0,52	83	1
A/B/C/H	0,62	0,90	0,25	83	6
A/B/C/H/Ch	0,47	0,82	0,31	59	4
A/B/C/H/Ch/DA	0,77	0,96	0,15	83	6
A/B/H	0,57	0,83	0,31	59	4
A/B/C/Ch	0,30	0,51	0,50	199	1
A/B/DA	0,68	0,93	0,20	53	6
A/B/C/DA	0,62	0,77	0,34	353	2
A/B/H/DA	0,72	0,94	0,18	257	5
A/B/C/Ch/DA	0,63	0,78	0,34	352	2
A/B/C/H/DA	0,74	0,96	0,15	71	6
A/B/H/Ch/DA	0,73	0,94	0,18	257	5

q^2 , Coeficiente de correlação com validação cruzada; r^2 , Coeficiente de correlação sem validação cruzada; SEE, Erro padrão de estimativa; HL, Comprimento do holograma; N, Número ótimo de componentes.

Os melhores resultados foram obtidos com a distinção de fragmentos A/B/C/H/Ch/DA, a qual produziu um coeficiente de correlação com validação cruzada (q^2) de 0,77 e um coeficiente de correlação sem validação cruzada (r^2) de 0,96, com erro (SEE) de 0,15.

O tamanho dos fragmentos é determinante para a geração dos modelos, uma vez que através deste parâmetro é determinado o número mínimo e máximo de átomos de cada fragmento. Os efeitos da variação do tamanho dos fragmentos sobre os resultados foram investigados para o melhor modelo da Tabela 10 (A/B/C/H/Ch/DA). A Tabela 11 mostra os resultados que foram obtidos para os tamanhos de fragmento: 2–5, 3–6, 5–8, 6–9, 7–10 e 8–11.

Tabela 11 – Resultados gerados para os diferentes tamanhos de fragmento.

Tamanho de Fragmento	Parâmetros Estatísticos				
	q^2	r^2	SEE	HL	N
2–5	0,72	0,89	0,24	257	4
3–6	0,76	0,92	0,21	71	5
5–8	0,74	0,97	0,12	353	6
6–9	0,77	0,98	0,12	97	6
7–10	0,71	0,98	0,11	401	6
8–11	0,65	0,98	0,10	257	6

A variação no tamanho dos fragmentos levou a uma melhora discreta em um dos modelos da Tabela 11, quando comparado ao melhor modelo da Tabela 10 (obtido com o tamanho de fragmento 4–7). O melhor resultado estatístico ($q^2 = 0,77$; $r^2 = 0,98$; SEE = 0,12) dentre todos os modelos foi obtido com um tamanho de fragmento de 6–9 átomos. Embora o valor de q^2 não tenha sido modificado, houve um aumento discreto no valor de r^2 e uma diminuição no valor do erro (SEE).

Para melhor avaliar a estabilidade e a consistência interna do modelo mais robusto (A/B/C/H/Ch/DA, tamanho de fragmento 6-9 átomos) foi aplicado outro procedimento de validação cruzada através do método “deixe vários de fora” (LMO, na sigla inglesa para *leave-many-out*), no qual o conjunto treinamento foi dividido em cinco e dez grupos selecionados aleatoriamente. Os resultados obtidos confirmaram a consistência do modelo, à medida que os parâmetros estatísticos são comparáveis àqueles obtidos com o procedimento LOO. A Tabela 12 exibe os resultados obtidos com o método LMO.

Tabela 12 – Resultados obtidos com o método de validação LMO para o modelo A/B/C/H/Ch/DA, tamanho de fragmento de 6 à 9 átomos.

Número de Grupos de Validação Cruzada	Parâmetros Estatísticos				
	q^2	r^2	SEE	HL	N
5	0,72	0,98	0,10	257	6
10	0,76	0,97	0,12	97	6

Embora os procedimentos de validação interna sejam importantes para validar os modelos de QSAR, o teste mais importante é a capacidade do modelo de prever a propriedade biológica de moléculas não incluídas no conjunto treinamento. À medida que a atividade biológica está diretamente correlacionada aos hologramas moleculares, os modelos HQSAR devem ser capazes de prever a atividade de moléculas relacionadas a partir dos hologramas moleculares gerados para este conjunto externo de moléculas. Desta maneira, a capacidade preditiva do melhor modelo (A/B/C/H/Ch/DA, tamanho de fragmento de 6–9 átomos) foi avaliada através da predição dos valores de pK_i para um conjunto teste de oito compostos. Os resultados da validação externa são exibidos na Tabela 13, e os resultados gráficos na Figura 35.

Tabela 13 – Valores experimentais e preditos de pK_i com os valores residuais para o conjunto teste.

Inibidor	pK_i Experimental	pK_i Predito	Resíduo^a
1	2,44	2,38	0,06
3	2,85	2,95	-0,10
6	3,48	3,26	0,22
9	3,48	3,63	-0,15
14	2,10	1,88	0,22
17	1,31	1,68	-0,37
22	1,72	1,65	0,07
31	1,96	2,56	0,60

^a Diferença entre os valores experimentais e preditos de pK_i .

Figura 35 – Valores de pK_i experimentais *versus* os valores preditos e estimados para a série treinamento e série teste, obtidos para o modelo A/B/C/H/Ch/DA, 6–9 átomos (Tabela 11).

A validação externa é um dos requerimentos mais importantes no processo de qualificação de modelos de QSAR, em razão da total exclusão dos compostos da série teste da construção dos modelos. A concordância entre os valores experimentais e preditos de pK_i para o conjunto teste, em conjunto com o valor do coeficiente de predição ($r^2_{\text{pred}} = 0,86$) para o melhor modelo (A/B/C/H/Ch/DA, tamanho de fragmentos de 6–9 átomos) indica a robustez do modelo para a predição da atividade de novos inibidores dentro da diversidade estrutural considerada.

6.4.1 Mapas de Contribuição

O método de HQSAR pode ser utilizado para destacar fragmentos moleculares que contribuem positivamente ou negativamente para atividade de compostos bioativos. Através de um código de cores, a contribuição de cada fragmento individual pode ser avaliada. Nestes mapas, contribuições negativas são representadas pelas cores vermelho, vermelho-laranja e laranja. Contribuições

positivas são representadas pelas cores amarelo, verde-azul e verde. Contribuições intermediárias ou neutras são representadas pela cor branca.

Existem dois aspectos relevantes a serem considerados nos mapas de contribuição gerados para o melhor modelo de HQSAR: o número de átomos de carbono da cadeia hidrocarbônica e a presença de substituintes amida ou éster.

Três compostos do conjunto de dados que diferem apenas no número de átomos de carbono entre os dois grupos fosfato são comparados na Figura 36. O mapa de contribuição destaca a relevância da presença de seis átomos de carbono na cadeia hidrocarbônica (composto **4**) em comparação com os compostos que possuem quatro e cinco átomos (composto **2** e **3**, respectivamente), como pode ser observado pelos valores de pK_i destes inibidores. A contribuição de alguns átomos da cadeia hidrocarbônica passa de neutra para favorável à medida que o número de átomos de carbono aumenta. Isso indica que os inibidores desta classe devem possuir preferencialmente o mesmo número de átomos de carbono do substrato da aldolase (frutose-bifosfato).

Figura 36 – Mapas de contribuição para os compostos 2, 3 e 4. Os mapas mostram que a contribuição da cadeia hidrocarbônica cresce à medida que o número de átomos de carbono se aproxima de seis.

Os mapas de contribuição, gerados para os compostos mais potentes da série demonstram a importância das substituições na cadeia hidrocarbônica. Moléculas contendo grupos amida possuem maior afinidade pela enzima em relação às que contêm grupos éster. Este aspecto é destacado na Figura 37, na qual compostos que diferem apenas quanto à substituição na cadeia hidrocarbônica são comparados. É demonstrado que a carbonila do grupo amida (composto **10**) contribui mais positivamente para a atividade em comparação com a carbonila do grupo éster (composto **6**). Além disso, quando conectada a um grupo amida, a cadeia hidrocarbônica passa a contribuir positivamente para a atividade. As bases moleculares foram investigadas posteriormente através de estudos de docagem.

Figura 37 – Mapas de contribuição para os compostos **6** (carbonila em amarelo) e **10** (carbonila em verde).

6.4.2 Integração dos Mapas de Contribuição à Estrutura da Aldolase

A estrutura da aldolase de *T. brucei* (código PDB: 1F2J, resolução de 1,9 Å) propiciou a exploração das características estruturais do sítio ativo da enzima para o planejamento de novos ligantes. Neste sentido, foi utilizada neste trabalho uma abordagem que combina métodos de LBDD e SBDD, com a finalidade de explorar a importância dos fragmentos moleculares indicados pelos mapas de contribuição no processo de reconhecimento molecular entre a aldolase e os inibidores do conjunto de dados utilizado. Para tanto, foram conduzidos ensaios de docagem molecular dos compostos **6** e **10** na cavidade de ligação da aldolase de *T. brucei*.

O preparo da estrutura envolveu a adição dos átomos de hidrogênio após a remoção das moléculas de água. O programa GOLD 5.0 foi empregado para obter as conformações dos inibidores no sítio ativo da enzima. O sítio ativo foi definido

como uma esfera com raio de 10 Å em torno da cavidade de interação com o substrato. Aos resíduos de arginina e lisina presentes no sítio foi atribuída uma flexibilidade de 10 graus em torno das ligações rotacionáveis. A função de detecção de cavidades foi empregada para restringir a seleção dos átomos à superfície acessível ao solvente. A função de pontuação *GoldScore* foi utilizada para avaliar as conformações propostas. Cada molécula foi docada 20 vezes e apenas as conformações que obtiveram melhor classificação foram consideradas.

De acordo com as melhores soluções propostas pelo programa GOLD, a presença de uma amida ou um éster como substituintes da cadeia hidrocarbônica, aumenta o número de contatos polares entre os inibidores e resíduos específicos do sítio de ligação da aldolase. De fato, os resultados mostram que o átomo de oxigênio da carbonila forma interações de hidrogênio com os resíduos de lisina 156 e 239, como ilustrado na Figura 38.

Figura 38 – Conformação dos compostos 6 **(A)** e 10 **(B)**, com destaque para as interações polares entre o grupo éster **(A)** e o grupo amida **(B)** com os resíduos de lisina 156 e 239 do sítio ativo da aldolase.

Em resumo, a consistência interna e externa do melhor modelo de HQSAR (A/B/C/H/Ch/DA, tamanho de fragmento de 6–9 átomos), em conjunto com as informações obtidas a partir dos mapas de contribuição e dos estudos de docagem, permitiu a identificação de fragmentos moleculares essenciais para o reconhecimento molecular entre a aldolase e os inibidores do conjunto de dados.

6.5 Análise Comparativa dos Campos Moleculares

Para complementar as investigações de QSAR 2D, estudos de QSAR 3D foram realizados para compreender os fatores de natureza tridimensional, especificamente os campos de interação eletrostáticos e estereoquímicos, relevantes para a atividade biológica do conjunto de inibidores.

Com esta finalidade, a série de ésteres fosfóricos foi utilizada no desenvolvimento de modelos, empregando-se o método CoMFA. Os resultados foram posteriormente avaliados sob a ótica do modo de interação dos inibidores com o sítio ativo da aldolase, através de docagem e dinâmica molecular.

6.5.1 Alinhamento Molecular

O alinhamento estrutural das moléculas do conjunto de dados no espaço tridimensional é um requerimento de central importância para o método CoMFA. O alinhamento foi baseado na conformação bioativa utilizando a docagem dos inibidores no sítio ativo da aldolase de *T. brucei*. O preparo da estrutura envolveu a adição dos átomos de hidrogênio após a remoção das moléculas de água. O programa GOLD 5.0 foi empregado para obter as conformações dos inibidores no sítio ativo da enzima. O sítio ativo foi definido como uma esfera com raio de 10 Å em torno da cavidade de interação com o substrato. Aos resíduos de arginina e lisina presentes no sítio foi atribuída uma flexibilidade de 10 graus em torno das ligações rotacionáveis. A função de detecção de cavidades foi empregada para restringir a seleção dos átomos à superfície acessível ao solvente. A função de pontuação *GoldScore* foi utilizada para avaliar as conformações propostas. Cada molécula foi docada 20 vezes e apenas as conformações que obtiveram melhor classificação foram consideradas para o alinhamento molecular. O conjunto completo das moléculas alinhadas é ilustrado na Figura 39.

Figura 39 – Alinhamento molecular dos inibidores utilizado na construção dos modelos CoMFA.

6.5.2 Parâmetros dos Modelos CoMFA

Com a finalidade de gerar os modelos preditivos de QSAR, e avaliar as contribuições eletrostáticas e estereoquímicas para a afinidade dos inibidores, análises CoMFA foram realizadas a partir do alinhamento molecular ilustrado na Figura 39. Os campos de interação eletrostáticos e estereoquímicos foram calculados através dos potenciais de Lennard-Jones e Coulomb, respectivamente.

A distribuição dos inibidores dos conjuntos treinamento e teste foi a mesma adotada para o desenvolvimento dos modelos de HQSAR. Desta forma, as moléculas do conjunto treinamento foram posicionadas em uma grade tridimensional reticulada, com espaçamento de 2 Å entre os vértices do retículo e com 4 Å de extensão além da extremidade das moléculas.

Os descritores moleculares (campos eletrostáticos e estereoquímicos) foram gerados em cada ponto de interseção da grade, pelo campo de força Tripos, utilizando-se um átomo de carbono sp^3 com carga +1, como sonda molecular. Um valor de corte de 30 kcal.mol⁻¹ foi estabelecido como o valor máximo para as energias de interação eletrostáticas e estereoquímicas¹¹². O procedimento de

focalização de região foi realizado com a finalidade de refinar os modelos. Tal procedimento é realizado através do aumento do peso sobre as análises estatísticas, daqueles pontos da grade que contribuem de forma mais efetiva para o modelo. A focalização de região também melhora a resolução dos modelos e a capacidade preditiva das análises subsequentes.

Os descritores CoMFA (variáveis independentes) foram correlacionados com os valores de pK_i através do método PLS. A capacidade preditiva dos modelos foi testada através de métodos de validação cruzada e também pela predição da propriedade biológica para um conjunto teste. Além do método LOO também foi utilizado o método LMO com cinco e dez grupos selecionados aleatoriamente.

O conjunto treinamento foi submetido a várias análises CoMFA, as quais resultaram em um modelo inicial com valores de $q^2 = 0,72$ e $r^2 = 0,77$. Estes resultados foram considerados passíveis de serem melhorados, particularmente em relação ao ajuste dos dados, representado pelo parâmetro r^2 . Portanto, foi aplicado o método de focalização de região, utilizando a técnica de ajuste dos coeficientes de desvio padrão (SDC, na sigla inglesa para *standard deviation coefficient*). Foram aplicados valores de SDC iguais a 0,3; 0,6 e 0,9. Este procedimento levou a um aumento dos valores de ambos os coeficientes de correlação, além de resultar no refinamento dos mapas de contorno tridimensionais.

Os resultados estatísticos obtidos depois da aplicação do procedimento de refinamento dos modelos são apresentados na Tabela 14. Coeficientes de correlação significativos foram obtidos para os modelos 2 e 3, com valores de $q^2 = 0,80$ e $r^2 = 0,95$ para o modelo 3.

Tabela 14 – Resultados estatísticos para os modelos CoMFA, obtidos em função da variação nos valores de SDC.

Modelo	q^2	SEP	r^2	SEE	N	F	Fração	
							S	E
1	0,73	0,37	0,85	0,27	2	81,02	0,45	0,55
2	0,76	0,36	0,95	0,16	4	12,80	0,51	0,49
3	0,80	0,33	0,95	0,16	4	12,82	0,48	0,52

N, Número ótimo de componentes; F, teste F; S, contribuição estereoquímica; E, contribuição eletrostática. Modelo 1, SDC = 0,3; modelo 2, SDC = 0,6; modelo 3, SDC = 0,9.

A estabilidade do modelo 3 foi posteriormente avaliada através do método de validação cruzada LMO, o que resultou na obtenção de valores de $q^2 = 0,73$ quando cinco grupos de validação foram utilizados. Quando dez grupos foram utilizados, foi obtido um valor de $q^2 = 0,76$. Estes resultados demonstram que apesar de uma discreta diminuição, os valores de q^2 são comparáveis aos obtidos pelo método LOO, demonstrando que o modelo é capaz de manter sua capacidade preditiva.

Embora o valor dos coeficientes de correlação representem de forma adequada a consistência interna e a capacidade de predição dos modelos, o processo de validação externa é considerado o método de validação mais significativo, à medida que o modelo é utilizado para prever os valores da propriedade biológica de um conjunto externo de moléculas que não foram utilizadas na etapa de construção do modelo.

Portanto, a capacidade preditiva do modelo 3 foi avaliada através da predição da afinidade de um conjunto teste formado por oito compostos (moléculas **1**, **3**, **6**, **9**, **14**, **17**, **22** e **31**). Os valores de pK_i experimentais e preditos para os compostos da série teste são apresentados na Tabela 15. Estes resultados são representados de forma gráfica na Figura 40. A boa concordância entre os valores experimentais e preditos de pK_i para a série teste demonstra a capacidade do modelo desenvolvido em prever a afinidade de novos inibidores dentro da diversidade estrutural explorada. Esta capacidade é expressa de forma mais objetiva através do valor do coeficiente de correlação de predição ($r^2_{\text{pred}} = 0,84$).

Tabela 15 – Valores experimentais e preditos de pK_i e dos resíduos para a série teste.

Inibidor	pK_i Experimental	pK_i Predito	Resíduo
1	2,44	3,00	-0,56
3	2,85	3,13	-0,28
6	3,48	3,02	0,46
9	3,48	3,30	0,18
14	2,10	1,87	0,23
17	1,31	1,60	-0,29
22	1,72	1,77	-0,05
31	1,96	1,86	0,10

Figura 40 – Gráfico dos valores experimentais *versus* os valores preditos e estimados de pK_i para a série treinamento (azul) e para a série teste (vermelho).

6.5.3 Mapas de Contorno CoMFA

As relações existentes entre a estrutura e atividade destes inibidores foram posteriormente investigadas através da interpretação dos mapas de contorno eletrostáticos e estereoquímicos gerados pelo melhor modelo CoMFA. Os mapas de contorno determinam regiões no espaço tridimensional, representadas por poliedros, onde mudanças nas propriedades eletrostáticas e estereoquímicas dos inibidores se correlacionam com mudanças correspondentes na afinidade dos compostos pela enzima.

Os poliedros verdes no mapa estereoquímico indicam que a presença de grupos químicos volumosos em uma determinada região está correlacionada com um aumento na afinidade dos inibidores. Poliedros amarelos, por sua vez, indicam que a presença de grupos químicos volumosos é prejudicial para a afinidade dos inibidores pela enzima.

O processo de alinhamento estrutural revelou que todos os compostos do conjunto de dados que possuem apenas um grupo fosfato estão alinhados de tal

maneira que a extremidade que possui este grupo funcional está sempre orientada para a mesma região do sítio ativo.

A Figura 41 mostra os mapas de contorno para dois inibidores do conjunto de dados. O mapa de contorno estereoquímico (poliedros verdes) mostra que a presença de grupos volumosos (como o fosfato) em uma das extremidades dos inibidores é benéfica para a afinidade, à medida que o composto **23**, que não possui um grupo volumoso como o fosfato nesta posição, tem uma afinidade 1000 vezes menor do que a do composto **10**, que possui tal grupo. Em relação ao mapa eletrostático, poliedros azuis indicam que grupos eletropositivos são benéficos para a afinidade, enquanto poliedros vermelhos indicam predileção por grupos eletronegativos.

Figura 41 – (A) Mapas de contorno CoMFA (modelo 3) para o inibidor **10** e **(B)** para o inibidor **23**.

A Figura 41 mostra poliedros vermelhos em torno de um dos grupos fosfato do composto **10** e da extremidade que não possui este grupo do composto **23**. Desta forma, o mapa eletrostático correlaciona-se com os dados experimentais à medida que compostos que apresentam grupos eletronegativos nesta região tem maior afinidade pela enzima (composto **10**, $pK_i = 4,26$; composto **23**, $pK_i = 1,88$).

Adicionalmente, os poliedros vermelhos em torno dos grupos amida e éster de ambos os compostos, demonstram a importância da presença de grupos que realizam interações de hidrogênio dentro da cadeia hidrocarbônica, como indicado pelos dados experimentais.

6.6 Dinâmica Molecular e as Interações Enzima-Inibidor

A compreensão das interações que ocorrem na formação e manutenção de complexos enzima-inibidor é essencial para o planejamento de inibidores mais potentes¹¹⁸⁻¹²⁰. Neste contexto, foi empregada uma estratégia que combina os dados obtidos nos estudos de QSAR, docagem e simulações de dinâmica molecular, com a finalidade de investigar as interações essenciais que são estabelecidas entre a aldolase e o conjunto de inibidores apresentados na Tabela 9.

Simulações de dinâmica molecular foram realizadas com a estrutura da enzima em complexo com os inibidores **4** e **10** do conjunto de dados, com o objetivo de compreender os seguintes aspectos: (i) verificar a estabilidade e manutenção das interações responsáveis pela afinidade dos inibidores pela aldolase; (ii) avaliar a evolução temporal do complexo enzima-inibidor; e (iii) assinalar a ocorrência de alterações conformacionais significativas na enzima, devido à interação com os inibidores.

As simulações de dinâmica molecular foram realizadas com o programa NAMD 2.7 em um *cluster* Dell PowerEdge 1950 II¹²¹. As interações entre os átomos nestes dois complexos enzima-inibidor foram calculadas através do campo de força CHARMM 22¹²². A topologia e a parametrização para os ligantes foram obtidas através do servidor SwissParam (Instituto Suíço de Bioinformática, Suíça)¹²³.

O cálculo das cargas dos ligantes foi realizada pelo programa PC-GAMESS/Firefly 7.1. O método de cálculo foi a análise de população de Mulliken, em um nível de teoria de Hartree-Fock, com um conjunto de base 6-31G*¹²⁴⁻¹²⁵.

Em seguida as configurações iniciais dos complexos, obtidas através da docagem dos compostos **4** e **10** no sítio ativo da aldolase, foram solvatadas em uma caixa de moléculas de água com dimensões de 80x80x90 Å, apresentando uma camada de 10 Å de água a partir das extremidades da enzima. A neutralização dos

dois sistemas foi realizada através da adição de íons sódio e cloreto. Após estes procedimentos cada sistema atingiu um número de 47200 átomos.

O estado de ionização dos aminoácidos e dos inibidores foi atribuído considerando-se um pH de 7. As interações eletrostáticas de longa distância foram calculadas através do método da rede de partículas de Ewald (PME, na sigla inglesa para *Particle Mesh Ewald*), com um valor de corte de 12 Å¹²⁶⁻¹²⁸.

Os complexos foram submetidos a uma etapa preliminar de minimização de energia de 100 passos, na qual apenas as moléculas de água puderam se mover, com a finalidade de evitar conflitos estereoquímicos durante a simulação. Os dois sistemas foram posteriormente submetidos a uma etapa de minimização de energia de 20 passos, e então finalmente submetidos à dinâmica molecular em simulações de 10 ns. Foram utilizadas as condições normais de pressão e temperatura: 1 atm e 298 K.

As condições de pressão e temperatura foram mantidas pela dinâmica de Langevin e pelo controle de pressão de pistão de Langevin Nosé-Hoover¹²⁹⁻¹³¹. Todo o processo de simulação foi repetido cinco vezes para cada complexo enzima-inibidor.

Os dados de inibição frente à aldolase para estes compostos mostram que os inibidores que possuem maior afinidade pela enzima (maiores valores de pK_i , Tabela 9), possuem em sua estrutura dois ésteres fosfóricos. Além disso, a presença, dentro da cadeia hidrocarbônica, de grupos funcionais polares, é importante para o estabelecimento de ligações de hidrogênio com a enzima e, portanto, para a afinidade dos inibidores, como demonstrado pelos estudos de QSAR (Figuras 37 e 41).

Com o objetivo de racionalizar estas evidências experimentais e computacionais, os compostos **10** (o de maior afinidade do conjunto de dados, $pK_i = 4,26$; $K_i = 55 \mu\text{M}$) e **4** ($pK_i = 3,27$; $K_i = 540 \mu\text{M}$) foram selecionados para os estudos de dinâmica molecular.

A estrutura destes dois compostos difere apenas pela presença de um grupo amida no composto **10**. A conformação destes inibidores obtida por docagem molecular foi utilizada como configuração inicial para a dinâmica molecular (Figura 42). As principais interações intermoleculares para estes dois inibidores também são ilustradas.

Figura 42 – (A) Conformação do composto 4 no sítio ativo da aldolase, com destaque para as interações eletrostáticas entre os grupos fosfato do inibidor e os resíduos do sítio ativo. **(B)** Conformação do composto 10, que demonstra o mesmo padrão de interação, exceto pela presença de uma interação adicional entre o grupo amida do ligante e os resíduos de lisina 156 e 239.

A análise das conformações adotadas por estes dois inibidores revelou que os ésteres fosfóricos são necessários para o estabelecimento de interações iônicas com os resíduos de lisina 116, arginina 158 e arginina 313. Também são estabelecidas interações íon-dipolo entre os grupos fosfato e os resíduos de serina 281, alanina 312 e arginina 313. O composto **10**, que possui um grupo amida, estabelece interações adicionais entre o oxigênio do grupo amida e os resíduos de

lisina 156 e lisina 239. A estabilidade destas interações foi avaliada através das simulações de dinâmica molecular.

Os gráficos do desvio médio da raiz quadrada (RMSD, na sigla inglesa para *root mean square deviation*) obtidos durante as simulações a partir das configurações iniciais dos complexos enzima-inibidor são visualizados na Figura 43.

Na Figura 43A são apresentados o RMSD para a aldolase sob a forma de apoenzima e em complexo com os inibidores **4** e **10**. Os gráficos demonstram um padrão de evolução muito semelhante para estes três sistemas moleculares, que atingem o estado de equilíbrio em torno de 7 ns. A semelhança neste padrão sugere que o processo de reconhecimento molecular entre a aldolase e estes análogos não produz alterações conformacionais significativas na enzima. Porém a enzima em complexo com os inibidores apresenta um valor de RMSD ligeiramente menor (≈ 3 Å) quando comparado à apoenzima ($\approx 3,5$ Å), possivelmente devido à redução da liberdade conformacional dos aminoácidos da região mais próxima do sítio ativo, que é estabilizada pelos inibidores.

A análise do RMSD dos ligantes (Figura 43B) demonstra que ambos são mantidos no sítio ativo da aldolase com a mesma orientação durante toda a simulação (RMSD ≈ 1 Å). Valores de RMSD ligeiramente menores são observados para o composto **10**, em comparação ao composto **4**. Isto indica uma estabilidade maior para o composto **10**, o que corrobora os resultados experimentais que atestam que este inibidor possui afinidade dez vezes maior pela enzima do que o composto **4**; e também os resultados da docagem, que mostram duas interações de hidrogênio adicionais entre este inibidor e a enzima. Também é possível observar a partir da Figura 43B que não houve a liberação de nenhum dos inibidores do sítio ativo da enzima durante a simulação.

A Figura 43C apresenta a flutuação para cada resíduo de aminoácido, para a apoenzima e para os dois complexos enzima-inibidor (RMSF, na sigla inglesa para *root mean square fluctuation*). É possível verificar que a flutuação média dos resíduos de aminoácidos é muito semelhante para os dois complexos enzima-inibidor e ligeiramente maior para a apoenzima. Particularmente para os resíduos do sítio ativo, envolvidos na interação com os inibidores, esta flutuação é menor para os complexos enzima-inibidor do que para a apoenzima. Como esperado, a extremidade N-terminal da enzima nos três sistemas (resíduos de 1 a 20, Figura 43C) apresenta uma grande mobilidade durante toda a simulação.

Figura 43 – (A) RMSD da aldolase na presença dos compostos 4 (linha azul) e 10 (linha laranja) durante a simulação. **(B)** RMSD dos compostos 4 e 10. **(C)** Flutuações médias de cada resíduo de aminoácido da aldolase durante toda a simulação, para o complexo com os inibidores 4 e 10, e para a apoenzima. Cada linha representa a média de cinco simulações independentes.

Com objetivo de verificar a manutenção e a estabilidade do padrão de interação entre os inibidores e o sítio ativo da aldolase e para elucidar as razões moleculares responsáveis pela maior afinidade do composto **10**, a distância das interações observadas na docagem molecular para os compostos **4** e **10** foi monitorada durante as simulações de DM. Estes resultados são ilustrados

graficamente na Figura 44, onde o comprimento de cada interação pode ser observado ao longo do intervalo de simulação.

Figura 44 – Distância das interações intermoleculares entre os ligantes e os resíduos do sítio ativo durante as simulações.

Como observado na Figura 44, as interações intermoleculares foram mantidas durante todo o período de simulação. O padrão geral de todas as interações revelou-se muito similar para os compostos **4** e **10**. A análise conjunta destes dados indica que as interações eletrostáticas comuns a ambos os inibidores são a principal força estabilizadora destes ligantes no sítio ativo da enzima.

O resíduo de serina 281, através do nitrogênio da cadeia principal e da hidroxila da cadeia lateral, estabelece duas interações íon-dipolo com o oxigênio **O**

de um dos grupos fosfato (Figuras 44C e D). Também são observadas três interações eletrostáticas envolvendo o resíduo de arginina 313, demonstrando que interações com este resíduo são essenciais para a atividade destes inibidores (Figuras 44F, G e H). Estas interações consistem em duas interações íon-dipolo com o oxigênio **O4** dos ligantes, envolvendo os átomos de nitrogênio da cadeia principal e da cadeia lateral (posição E) (Figuras 44F e H). A terceira é uma interação iônica envolvendo o nitrogênio protonado da posição H e o oxigênio **O5** dos ligantes (Figura 44G).

Também é observado na Figura 44 que as interações que envolvem os resíduos de lisina 116, arginina 158 e alanina 312 também se mantêm estáveis ao longo do tempo, contribuindo para a manutenção dos inibidores no sítio ativo da aldolase (Figuras 44A, B e E).

Finalmente, a maior afinidade do composto **10** parece estar essencialmente correlacionada com as interações adicionais do tipo íon-dipolo, que o inibidor estabelece entre o átomo de oxigênio **O8** do grupo amida e os resíduos de lisina 156 e 239 (Figuras 44I e J), o que revela a importância de tal grupo polar posicionado dentro da cadeia hidrocarbônica dos inibidores. Este aspecto, corroborado pelos dados experimentais e computacionais, é compatível com o ambiente de grande polaridade encontrado no sítio ativo aldolase.

Estes estudos não apenas confirmam a estabilidade das principais interações observadas, mas também evidenciam o papel central desempenhado pelos ésteres fosfóricos no estabelecimento de uma rede de interações polares e iônicas, que se confirmaram como as principais forças responsáveis pela manutenção dos complexos entre os inibidores e a enzima. Também foi estabelecida a importância do grupo amida como receptor de interações de hidrogênio, para conferir ao composto **10** sua maior afinidade.

Dentre os resíduos de aminoácidos do sítio ativo, os polares e os carregados positivamente são os mais importantes. A serina 281 está envolvida em duas interações íon-dipolo. A arginina 313 participa de duas interações íon-dipolo e uma interação iônica. Os resíduos de lisina 116, arginina 158 e alanina 312 contribuem com duas interações iônicas e uma interação íon-dipolo respectivamente.

Os resultados gerados pelas simulações de dinâmica molecular contribuíram para a compreensão dos fenômenos moleculares experimentados pela enzima e pelos inibidores durante o processo de reconhecimento molecular. Também foram

úteis para a elucidação das razões moleculares responsáveis pelas relações existentes entre a estrutura e a atividade do conjunto de inibidores estudado.

6.7 Bases Estruturais para o Planejamento de Inibidores Seletivos para a Aldolase de *Trypanosoma brucei*

Os vários aspectos moleculares envolvidos na interação destes inibidores com a aldolase de *T. brucei*, bem como as informações produzidas pelas análises de HQSAR e CoMFA, nos levaram à investigar quais aspectos seriam mais relevantes para o planejamento de inibidores seletivos para a enzima do parasita. Para tanto, foi realizado um estudo comparativo das características estruturais da aldolase humana e da aldolase de *T. brucei*.

Dado que a estrutura das duas enzimas possui o mesmo padrão tridimensional, apresentando um alto nível de conservação dos elementos de estrutura secundária e terciária, a comparação entre as duas estruturas se concentrou na análise do sítio ativo¹³².

O exame das estruturas sobrepostas da aldolase de *T. brucei* e humana (código PDB 1QO5) revelou que apesar de haver uma grande conservação entre os aminoácidos correspondentes, existem variações pontuais que podem ser exploradas (Figura 45). Podemos observar que a glicina 302 e a treonina 38 na enzima humana são substituídas pela alanina 312 e serina 48, respectivamente, na enzima do parasita. Adicionalmente, observa-se que a fenilalanina 79, presente na aldolase humana, é substituída por um resíduo carregado, a histidina 88, na enzima de *T. brucei*. Os inibidores do conjunto de dados não ocupam a região da histidina 88, portanto a adição de grupos carregados ou polares pode ser utilizada como estratégia para ocupar esta região da cavidade de ligação, proporcionando maior seletividade para a enzima do parasita.

Figura 45 – Sítio ativo da aldolase de *T. brucei* (cinza) sobreposto à enzima humana (verde). No centro o inibidor 6 (carbonos em azul). As diferenças estruturais estão destacadas com as siglas dos respectivos aminoácidos.

6.8 Triagem Virtual para a Aldolase

A busca por novos inibidores da aldolase de *Trypanosoma brucei* foi realizada através do método de triagem virtual baseado na estrutura do receptor (SBVS). Foi utilizada uma estratégia conhecida como *tiered approach* (*abordagem em camadas*), que constitui na realização de duas etapas de triagem. Primeiramente foi utilizado o programa DOCK 3.5.54 (um algoritmo computacional mais rápido) com o propósito de excluir compostos que não são compatíveis com o sítio ativo, principalmente em relação à complementariedade estereoquímica⁷¹. Na segunda etapa, a base contendo apenas compostos complementares ao sítio de ligação foi submetida à triagem pelo programa GOLD 5.0. A análise dos resultados foi realizada com o programa de modelagem molecular Pymol 1.3 (DeLano Scientific, San Carlos, USA). A estratégia de triagem empregada é esquematizada na Figura 46.

Figura 46 – Estratégia de triagem virtual adotada, baseada no método *tiered approach*.

6.8.1 Seleção e Preparo do Alvo Macromolecular

A aldolase de *T. brucei* possui a estrutura tridimensional do monômero resolvida por cristalografia raios-x, a qual foi utilizada nas triagens virtuais (código PDB 1F2J, resolução de 1,9 Å)⁶¹. A enzima de *T. brucei* possui 81% de identidade com a enzima de *Leishmania mexicana* (parasita causador da Leishmaniose) e quando apenas o sítio ativo é comparado, a identidade é de 100%¹³³. A enzima de *L. mexicana* (código PDB 2QDG) está em complexo com o substrato – frutose-1,6-bifosfato – e quando as enzimas dos dois parasitas são sobrepostas é observado que apesar de todos resíduos do sítio ativo serem conservados, o resíduo de arginina 313 possui conformações distintas (Figura 47). Na estrutura de *L. mexicana* o resíduo se encontra voltado para o centro do sítio ativo ao passo que na estrutura de *T. brucei* está voltado para fora do sítio. Considerando que as argininas são resíduos flexíveis e que neste caso se orientam em direção ao sítio para interagir com o substrato (e com inibidores, como visto na dinâmica molecular), a arginina 313 foi modelada na estrutura de *T. brucei* para adotar a orientação encontrada na estrutura de *Leishmania*.

Triagens teste foram conduzidas com o resíduo de arginina 313 na sua conformação original encontrada na estrutura cristalográfica de *T. brucei*. Paralelamente, foram realizados ensaios com o respectivo aminoácido voltado para a cavidade de interação. A comparação dos resultados confirmou que a segunda estratégia leva a resultados mais condizentes com o modo de ligação do substrato e também com os estudos de QSAR e DM. Estes resultados foram decisivos na escolha da estratégia de se adotar a estrutura com o resíduo 313 voltado para o sítio.

Figura 47 – Diferença conformacional do resíduo de arginina 313 no sítio ativo das enzimas de *T. brucei* (cinza) e *L. mexicana* (magenta). O substrato (verde) realiza diversas interações polares com o sítio ativo da enzima de *L. mexicana*.

Após a análise de várias estruturas de aldolases de diferentes espécies depositadas no PDB, foi constatada a conservação de duas moléculas de água no sítio ativo em todas as estruturas analisadas. Estas moléculas de água foram mantidas na estrutura utilizada para a docagem para que eventuais ligações de hidrogênio pudessem ser consideradas.

O processo de preparo da estrutura selecionada para a modelagem envolveu o seguinte procedimento, realizado com o programa GOLD 5.0: (i) adição de átomos de hidrogênio; (ii) eliminação das moléculas de água (exceto águas 388 e 407); e (iii) verificação dos estados de protonação e tautomerização dos resíduos do sítio ativo.

6.8.2 Triagem Virtual com o Programa DOCK

O programa DOCK versão 3.5.54 foi utilizado na primeira etapa de triagem virtual, em colaboração com o grupo do Prof. Brian Shoichet (UCSF). O subconjunto fragmento-similar, contendo 500.000 compostos, presente no banco de dados virtual ZINC (disponível em <http://zinc.docking.org/>), foi selecionado para esta etapa de triagem⁷⁶. As moléculas presentes neste subconjunto apresentam como características: logaritmo do coeficiente de partição octanol/água ($\log P$) < ou = 3,5; peso molecular < ou = 250; número de ligações rotacionáveis < ou = 5.

Baseando-se nas interações intermoleculares observadas nas estruturas da aldolase com o substrato (*L. mexicana*, 2QDG), com o inibidor MBP (músculo de coelho, 1ZAJ) e também nos modelos de QSAR e simulações de DM, foi realizada uma busca bidimensional e incluídas no processo de triagem, moléculas apresentando grupos carregados negativamente, conhecidos por interagir com resíduos de alanina, serina, arginina e lisina (Figura 48).

Figura 48 – Grupos funcionais carregados negativamente encontrados em interações com resíduos de alanina, serina, arginina e lisina (grupos carboxilato, sulfonato e fosfato).

O sítio ativo da enzima foi definido utilizando a estrutura do substrato como referência. As coordenadas do substrato foram utilizadas para gerar as conformações iniciais das moléculas docadas de acordo com o algoritmo padrão do DOCK. A função de pontuação padrão do programa foi utilizada para avaliar e classificar as conformações propostas.

6.8.3 Triagem Virtual com o Programa GOLD

A partir dos resultados gerados na triagem inicial realizada com o DOCK, foram selecionados os 10 % da base com melhor pontuação. Com esta fração de 50.000 moléculas, foi conduzida a triagem com o programa GOLD. Nesta etapa, o sítio ativo foi definido como uma esfera com raio de 10 Å em torno da cadeia lateral da lisina 239. Com o programa GOLD é possível atribuir flexibilidade à cadeia lateral dos resíduos do sítio ativo, o que é um aspecto positivo, pois se considera um aspecto importante na formação do complexo enzima-inibidor. Aos resíduos de arginina e lisina presentes no sítio foi atribuída uma flexibilidade de 10 graus em torno de suas ligações rotacionáveis. A função de detecção de cavidades foi empregada para restringir a seleção dos átomos à superfície acessível ao solvente, e a função de pontuação padrão do programa – *GoldScore* – foi utilizada para avaliar e classificar as conformações propostas.

6.8.4 Seleção de Candidatos a Inibidores da Aldolase

Um exame detalhado dos resultados revelou a importância das interações iônicas realizadas entre moléculas com grupos carregados negativamente e o sítio ativo da aldolase. As principais interações observadas entre as moléculas melhor classificadas e o sítio ativo da enzima foram interações eletrostáticas envolvendo os seguintes resíduos: Arg313, Arg158, Lys116, Lys156, Lys239, Ser48 e Glu197. A correlação entre a carga líquida das moléculas e a sua classificação ficou evidente mediante a ordem de classificação de nucleotídeos presentes na base: os trifosfatados se classificaram melhor do que os bifosfatados, que por sua vez ficaram à frente dos que possuíam apenas um grupo fosfato.

Os 1.000 compostos melhores classificados pela função *Goldscore* foram selecionados e analisados em relação aos critérios: (i) posicionamento e conformação adotada; (ii) interações com o sítio ativo; e (iii) ausência de conflitos estereoquímicos. De acordo com esta metodologia de análise pós-docagem foram selecionados 21 compostos para a realização dos ensaios de inibição frente à

aldolase de *T. brucei*. A Figura 49 ilustra a conformação e as interações intermoleculares para dois compostos selecionados.

Figura 49 – Estrutura, conformação e interações intermoleculares de duas moléculas selecionadas na triagem virtual.

Capítulo 7

Identificação de Inibidores da Aldolase

7.1 Avaliação Bioquímica de Candidatos a Inibidores da Aldolase

A avaliação bioquímica dos 21 compostos selecionados através do método de VS descrito no capítulo 6 foi realizada através de ensaios de inibição em concentração única de inibidor. Foi empregado o ensaio cinético padrão descrito no capítulo 5, utilizando uma concentração de inibidor de 500 μM . A Tabela 16 contém a estrutura destas moléculas e a porcentagem de inibição frente à aldolase. Os ensaios foram feitos em triplicata. Como controle positivo foi utilizada a reação com todos os reagentes (exceto o inibidor), e como controle negativo foi utilizada a reação com todos os reagentes exceto o substrato.

Tabela 16 – Percentual de inibição dos compostos testados frente à enzima aldolase de *T. brucei*.

Composto	Estrutura	Inibição (%)
1		$2,0 \pm 0,7$
2		-
3		$14,0 \pm 0,1$
4		20 ± 2

continua

continuação

continua

continuação

10		13 ± 1
11		$65,0 \pm 0,9$
12		$26,8 \pm 0,8$
13		49 ± 3
14		$16,7 \pm 0,4$
15		$5,3 \pm 0,8$
16		10 ± 1

continua

conclusão

17		8 ± 1
18		9 ± 1
19		38 ± 2
20		6 ± 1
21		$25,5 \pm 0,7$

As moléculas mostradas na Tabela 16 são estruturas novas como inibidores da aldolase de *T. brucei*. Os dados bioquímicos demonstram que os compostos são capazes de abolir parcialmente a atividade da enzima, exceto o composto **2** (o-fosfo-L-serina). Quatro moléculas demonstram uma porcentagem de inibição acima de 50%. Particularmente, a série que contém um anel naftaleno entre dois grupos sulfonato nas posições 2 e 7 constitui os melhores inibidores (compostos **7**, **11** e **13**).

As interações intermoleculares, obtidas por docagem, destes compostos com o sítio ativo da aldolase, são consistentes com os dados experimentais da Tabela 16. O modo de ligação destes inibidores mostra que a presença de grupos funcionais carregados negativamente aumenta o número de contatos polares entre

os inibidores e o sítio ativo da enzima, principalmente com os resíduos de serina, lisina, arginina e glutamato. A Figura 50 mostra o modo ligação e as interações intermoleculares para compostos representativos do grupo de moléculas selecionado através da triagem virtual.

Figura 50 – Conformação de inibidores representativos da série 2,7-naftaleno, obtidas por docagem molecular e as respectivas interações intermoleculares: **(A)** inibidor 11; **(B)** 7; **(C)** 13 e **(D)** 3.

O composto **11** (ácido 3-amino-5-hidroxi-naftaleno-2,7-difulfônico), que apresenta 65% de inibição, possui dois grupos sulfonato ligados às posições 2 e 7 do anel naftaleno. Esta subestrutura demonstrou ser o fragmento molecular mais promissor. Os grupos sulfonato, como mostrado na Figura 50A, são importantes para estabelecer interações de hidrogênio com os resíduos de serina 48, arginina 158 e arginina 313. O grupo amina ligado à posição 3 forma uma interação de hidrogênio com a água 388. O grupo hidroxila ligado à posição 5 estabelece interações de hidrogênio com os resíduos de lisina 156, lisina 239 e glutamato 197.

Os outros análogos que possuem o anel naftaleno (Figuras 50B, C e D) apresentam o mesmo padrão geral de interação. A variável mais importante que define a porcentagem de inibição desta série de análogos é a posição na qual estão conectados o grupo amina e o grupo hidroxila no anel naftaleno. A ligação do grupo amina na posição 3 é crítica, uma vez que o composto **7** (Figura 50B), que difere do composto **11** apenas pela posição da amina, apresenta uma queda na capacidade de inibição da enzima (51% de inibição). De acordo com a solução de docagem, esta única diferença resulta na perda de várias interações, especificamente com a arginina 158 e com a lisina 239 (Figuras 50A e B). Esta hipótese se torna ainda mais provável, à medida que o composto **13** (Figura 50C) que não possui o grupo amina, apresenta a mesma atividade sobre a enzima, quando comparado ao composto **7**. Em relação à presença do grupo hidroxila, parece que sua posição ideal de ligação é a posição 5 do anel, à medida que sua ligação à posição 3, como no análogo **3** (Figura 50D), causa uma grande diminuição na atividade deste composto (14% de inibição), em comparação com os outros análogos, apresentados na Figura 50, que apresentam o grupo hidroxila na posição 5.

Em relação ao composto **6** que possui uma estrutura diferente dos compostos descritos acima, a conformação proposta apresenta um tipo de interação encontrada apenas para este inibidor. Uma interação cátion- π , que em termos de energia é similar a uma interação de hidrogênio, mas que ocorre entre um resíduo carregado positivamente e a nuvem eletrônica de um anel aromático, é formada entre a arginina 313 e o anel isobenzofuranona do composto **6** (Figura 51). Também, dentre os melhores inibidores, este composto é o único que alcança regiões mais profundas do sítio ativo e interage com o resíduo de serina 281. Este achado tem utilidade para o planejamento de moléculas que ocupam mais extensivamente o sítio ativo.

Figura 51 – Conformação do inibidor 6. O anel isobenzofuranona forma uma ligação cátion- π com a arginina 313.

A triagem virtual realizada nesta tese, baseada na estrutura da aldolase, identificou novos compostos capazes de inibir a atividade da aldolase de *T. brucei*. Considerando o contexto de carência de novas alternativas terapêuticas para o tratamento da doença do sono, os análogos da série 2,7-naftaleno e o sistema isobenzofuranona constituem novos esqueletos moleculares apropriados para futuras investigações em química medicinal. Adicionalmente, os estudos proporcionaram indicações acerca da natureza das principais interações intermoleculares destes compostos com o sítio ativo da enzima, acrescentando novos conhecimentos, úteis para o planejamento de inibidores mais potentes da aldolase de *Trypanosoma brucei*.

Capítulo 8

Conclusões e Perspectivas

8 Conclusões e Perspectivas

Nesta tese de doutorado foram apresentadas as metodologias e os resultados obtidos para o planejamento e identificação de novos inibidores da enzima aldolase de *Trypanosoma brucei*.

A integração dos métodos experimentais e computacionais empregados gerou novas informações a respeito dos aspectos moleculares envolvidos na interação entre a enzima e os inibidores. Como resultados desta abordagem, podemos destacar:

- O estabelecimento de protocolos de expressão e purificação para a aldolase, e de um protocolo de cristalização, capaz de propiciar a formação de cristais com capacidade de difração de raios-x.
- A padronização de um protocolo de cinética enzimática para a avaliação de inibidores da aldolase.
- A construção de modelos de QSAR que apresentam alta consistência interna e externa, capazes de prever os valores de pK_i para o conjunto teste de inibidores, com boa concordância com os valores experimentais. Estes modelos podem ser aplicados na predição da propriedade biológica de novas moléculas.
- Os mapas de contribuição 2D e de contorno 3D gerados pelos modelos de HQSAR e CoMFA forneceram evidências acerca dos fragmentos moleculares e das características eletrostáticas e estereoquímicas que podem ser utilizadas no planejamento de novos inibidores da enzima.
- Os resultados da docagem molecular e dinâmica molecular forneceram informações sobre a conformação dos inibidores do conjunto de dados no sítio ativo da aldolase e deram suporte à racionalização das principais interações intermoleculares que são responsáveis pela estabilidade do complexo entre a enzima e estes ligantes.
- As simulações de dinâmica molecular foram úteis para revelar a ausência de alterações conformacionais significativas durante o processo de

reconhecimento molecular entre a aldolase e os inibidores do conjunto de dados.

- Foram assinaladas diferenças estruturais entre a aldolase de *T. brucei* e a enzima humana, importantes para o planejamento de inibidores seletivos da enzima do parasita.
- A realização da triagem virtual resultou na identificação de novas moléculas capazes de inibir a atividade da aldolase de *T. brucei*. Os análogos da série 2,7-naftaleno e o anel isobenzofuranona se apresentam como esqueletos estruturais atrativos para trabalhos posteriores de otimização molecular. Os estudos de docagem molecular destes inibidores proporcionaram a aquisição de informações úteis para o planejamento de inibidores mais potentes da aldolase de *Trypanosoma brucei*.

REFERÊNCIAS

- 1 PHARMACEUTICAL RESEARCH AND MANUFACTURERS OF AMERICA. Medicines play a key role in improving health while reducing avoidable costs. Washington D.C., 2012. Disponível em: <<http://www.phrma.org/sites/default/files/2319/phrmamedsplaykeyroleinimprovinghealthwhilereducingavoidablecosts2012.pdf>>. Acesso em: 17 Jan. 2013.

- 2 THE AMERICAN HOSPITAL ASSOCIATION. The value of investment in health care better care, better lives. Washington D.C., 2004. Disponível em: <http://www.aha.org/content/2004/pdf/Value_Report.pdf> Acesso em: 17 Jan. 2013.

- 3 WERMUTH, C. G. **The practice of medicinal chemistry**. London: Academic Press, 2008.

- 4 OLSSON, T.; OPREA, T. I. Cheminformatics: a tool for decision-makers in drug discovery. **Current Opinion in Drug Discovery and Development**, v. 4, n. 3, p. 308-313, 2001.

- 5 VILLOUTREIX B. O.; EUDES, R.; MITEVA, M. A. Structure-based virtual ligand screening recent success stories. **Combinatorial Chemistry & High Throughput Screening**, v. 12, n. 10, p. 1000-1016, 2009.

- 6 ANDRICOPULO, A. D.; SALUM, L. B.; ABRAHAM, D. J. Structure-based drug design strategies in medicinal chemistry. **Current Topics in Medicinal Chemistry**, v. 9, n. 9, p. 771-790, 2009.

- 7 ANDRICOPULO, A. D., MONTANARI, C. A. Structure-activity relationships for the design of small-molecule inhibitors. **Mini-Reviews in Medicinal Chemistry**, v. 5, n.6, p. 585-593, 2005.

- 8 BACILIERI, M.; MORO, S. Ligand-based drug design methodologies in drug discovery process an overview. **Current Drug Discovery Technologies**, v. 3, n. 3, p. 155-165, 2006.

- 9 VERMA, J.; KHEDKAR, V. M.; COUTINHO, E. C. 3D-QSAR in drug design a review. **Current Topics in Medicinal Chemistry**, v. 10, n. 1, p. 95-115, 2010.

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

- 11 TAFT, C. A.; DA SILVA V. B.; DA SILVA, C. H. Current topics in computer-aided drug design. **Journal of Pharmaceutical Sciences**, v. 97, n. 3, p. 1089-1098, 2008.

12 JÓNSDÓTTIR, S. O.; JORGENSEN, F. S.; BRUNAK, S. Prediction methods and databases within chemoinformatics emphasis on drugs and drug candidates. **Bioinformatics**, v. 21, n. 10, p. 2145-2160, 2005.

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

14 PHARMACEUTICAL RESEARCH AND MANUFACTURERS OF AMERICA. Drug discovery and development. Washington D.C. 2007. Disponível em: <http://www.innovation.org/drug_discovery/objects/pdf/RD_Brochure.pdf>. Acesso em: 22 Jan. 2013.

15 PATRICK, G. L. **An introduction to medicinal chemistry**. New York: Oxford University Press, 2009.

16 FERREIRA, R. S.; OLIVA, G.; ANDRICOPULO, A. D. Integração das técnicas de triagem virtual e triagem biológica automatizada em alta escala oportunidades e desafios em P&D de fármacos. **Química Nova**, v. 34, n.10, p. 1770-1778, 2011.

17 TRIPATHI, A.; SAFO, M. K. In silico-screening approaches for lead generation identification of novel allosteric modulators of human-erythrocyte pyruvate kinase. **Methods in Molecular Biology**, v. 796, p. 351-367, 2012.

18 THOMAS, G. **Fundamentals of medicinal chemistry**. Chichester: Wiley, 2003.

19 LANG, M.; SEIFERT, M. H.; WOLF, K. K.; ASCHENBRENNER, A.; BAUMGARTNER, R.; WIEBER, T.; TRENTINAGLIA, V.; BLISSE, M.; TAJIMA, N.; YAMASHITA, T.; VITT, D.; NODA, H. Discovery and hit-to-lead optimization of novel allosteric glucokinase activators. **Bioorganic Medicinal Chemistry Letters**, v. 21, n. 18, p. 5417-5422, 2011.

20 DIMITRI, N. An assessment of R&D productivity in the pharmaceutical industry. **Trends in Pharmacological Sciences**, v. 32, n. 12, p. 683-685, 2011.

21 ANDRICOPULO, A. D. Structure- and ligand-based drug design advances and perspectives. **Current Topics in Medicinal Chemistry**, v. 9, n. 9, p. 754, 2009.

22 LEWIS, R. A. A general method for exploiting QSAR models in lead optimization. **Journal of Medicinal Chemistry**, v. 48, n. 5, p. 1638-1648, 2005.

23 WILLIAMS, D. A.; FOYE, W.O.; LEMKE, T. L. **Foye's principles of medicinal chemistry**. Philadelphia: Lippincott Williams and Wilkins, 2002.

24 LaMATTINA, J. L. The impact of mergers on pharmaceutical R&D. **Nature Reviews Drug Discovery**, v. 10, p. 559-60, 2011.

25 WORLD HEALTH ORGANIZATION. First WHO report on neglected tropical diseases working to overcome the global impact of neglected tropical diseases. Geneva, 2010. Disponível em: < http://whqlibdoc.who.int/publications/2010/9789241564090_eng.pdf >. Acesso em: 22 Jan. 2013.

26 FENTON, A. Worms and germs: the population dynamic consequences of microparasite-macroparasite co-infection. **Parasitology**, v. 135, n. 13, p. 1545-1560, 2008.

27 SINGH, G.; SEHGAL, R. Transfusion-transmitted parasitic infections. **Asian Journal of Transfusion Sciences**, v. 4, n. 2, p. 73-77, 2010.

28 FENWICK, A. The global burden of neglected tropical diseases **Public Health**. v. 126, n. 3, p. 233-236, 2012.

29 WORLD HEALTH ORGANIZATION. The global burden of disease 2004 update. Geneva, 2008. Disponível em: <http://www.who.int/entity/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf>. Acesso em: 22 Jan. 2013.

30 DRUGS FOR NEGLECTED DISEASES INITIATIVE. 2011 annual report towards sustainable change for neglected patients. Geneva, 2011. Disponível em: <http://www.dndi.org/images/stories/annual_report/2011/DNDi_Annual%20report%202011_low-res.pdf>. Acesso em: 25 Abr. 2013.

31 RIDLEY, R. G.; FLETCHER, E. R. Making a difference 30 years of TDR. **Nature Reviews Microbiology**, v. 6, n. 5, p. 401-407, 2008.

32 MOLYNEUX, D. H; MALECELA, M. N. Neglected tropical diseases and the millennium development goals why the other diseases matter reality versus rhetoric. **Parasites & vectors**, v. 4, n. 234, p. 1-13, 2011.

33 WORLD HEALTH ORGANIZATION. A guide for drug supply access to drugs for human african *trypanosomiasis*. Geneva, 2001. Disponível em: <http://whqlibdoc.who.int/hq/2001/WHO_CDS_CSR_EPH_2001.3.pdf>. Acesso em: 22 Jan. 2013.

34 MIGCHELSEN, S. J.; BUESCHER, P.; HOEPELMAN, A. I. M. Human african *trypanosomiasis* a review of non-endemic cases in the past 20 years. **International Journal of Infectious Diseases**. v. 15, n. 8, p. E517-E524, 2011.

- 35 BLUM, J; SCHMID, C; BURRI, C. Clinical aspects of 2541 patients with second stage human African *trypanosomiasis*. **Acta Tropica**, v. 97, n. 1, p. 55-64, 2006.
- 36 BARRETT, M. P.; CROFT, S. L. Management of trypanosomiasis and leishmaniasis. **British Medical Bulletin**. v. 104, n. 1, p. 175-196, 2012.
- 37 CHATELAIN, E.; IOSET, J. Drug discovery and development for neglected diseases: the DNDi model. **Drug Design Development and Therapy**. v. 5, p. 175-181, 2011.
- 38 BURRI, C. Chemotherapy against human African trypanosomiasis: is there a road to success? **Parasitology**, v. 137, n. 14, p. 1987-1994, 2010.
- 39 WAINWRIGHT, M. Dyes trypanosomiasis and DNA a historical and critical review **Biotechnic & Histochemistry**, v. 85, n. 6, p. 341-354, 2010.
- 40 KENNEDY, P. G. E. An alternative form of melarsoprol in sleeping sickness. **Trends in Parasitology**, v. 28, n. 8, p. 307-310, 2012.
- 41 SIMARRO, P. P.; FRANCO, J.; DIARRA, A.; Postigo, J. A.; Jannin, J. Update on field use of the available drugs for the chemotherapy of human African trypanosomiasis. **Parasitology**, v. 139, n. 7, p. 842-846, 2012.
- 42 BARRETT, M. P.; VINCENT, I. M.; BURCHMORE, R. J.; KAZIBWE, A. J.; MATOVU, E. Drug resistance in human African trypanosomiasis. **Future Microbiology**, v. 6, n. 9, p. 1037-1047, 2011.
- 43 CHOI, J.; EL-SAYED, N. M Functional genomics of trypanosomatids. **Parasite Immunology**, v. 34, n. 2-3, p. 72-79, 2012.
- 44 JACOBS, R. T.; NARE, B.; PHILLIPS, M. A. State of the art in african trypanosome drug discovery. **Current Topics in Medicinal Chemistry**. v. 11, n. 10, p. 1255-1274, 2011.
- 45 PAL, C.; BANDYOPADHYAY, U. Redox-active antiparasitic drugs. **Antioxidants & Redox Signaling**. v. 17, n. 4, p. 555-582, 2012.
- 46 SIENKIEWICZ, N.; ONG, H. B.; FAIRLAMB, A. H. *Trypanosoma brucei* pteridine reductase 1 is essential for survival *in vitro* and for virulence in mice. **Molecular Microbiology**, v. 77, n. 3, p. 658-671, 2010.
- 47 MALLARI, J. P.; SHELAT, A. A.; KOSINSKI, A.; CAFFREY, C. R.; CONNELLY, M.; ZHU, F.; MCKERROW, J. H.; GUY, R. K. Structure-guided development of-selective tbcath inhibitors. **Journal of Medicinal Chemistry**. v. 52, n. 20, p. 6489-6493, 2009.

48 MICHELS, P. A. M.; OPPERDOES, F. R. Enzymes of carbohydrate metabolism as potential drug targets. **International Journal of Parasitology**, v. 31, n. 5-6, p. 482-90, 2001.

49 OPPERDOES, F. R. Compartmentation of carbohydrate metabolism in trypanosomes. **Annual Review of Microbiology**, v. 41, p. 127-51, 1987.

50 ALBERT, M. A.; HAANSTRA, J. R.; HANNAERT, V.; VAN ROY, J.; OPPERDOES, F. R.; BAKKER, B. M.; MICHELS, P. A. Experimental and in silico analyses of glycolytic flux control in bloodstream form *Trypanosoma brucei*. **Journal of Biological Chemistry**, v. 280, n. 31, p. 28306–28315, 2005.

51 VERLINDE, C. L.; HANNAERT, V.; BLONSKI, C.; WILLSON, M.; PÉRIÉ, J. J.; FOTHERGILL-GILMORE, L. A.; OPPERDOES, F. R.; GELB, M. H.; HOL, W. G.; MICHELS, P. A. Glycolysis as a target for the design of new anti-trypanosome drugs. **Drug resistance updates: reviews and commentaries in antimicrobial and anticancer chemotherapy**. v. 4, n. 1, p. 50–65, 2001.

52 BAKKER, B. M.; MICHELS, P. A.; OPPERDOES, F. R.; WESTERHOFF, H. V. Glycolysis in bloodstream from *trypanosoma brucei* can be understood in terms of the kinetics of the glycolytic enzymes. **Journal of Biological Chemistry**, v. 272, n. 6, p. 3207–3215, 1997.

53 OLIVARES-ILLANA V.; PÉREZ-MONTFORT, R.; LÓPEZ-CALAHORRA, F.; COSTAS, M.; RODRÍGUEZ-ROMERO, A.; TUENA DE GÓMEZ-PUYOU, M.; GÓMEZ PUYOU, A. Structural differences in triosephosphate isomerase from different species and discovery of a multitrypanosomatid inhibitor. **Biochemistry**, v. 45, n. 8, p. 2556-2560, 2006.

54 COPELAND, R. A. **Evaluation of enzyme inhibitors in drug discovery**. New Jersey: Wiley, 2005.

55 PRICE, N.; STEVENS, L. **Fundamentals of enzymology**, New York: Oxford University Press, 1989.

56 SCAPIN, G. Structural biology and drug discovery. **Current Pharmaceutical Design**, v. 12, n. 17, p. 2087–2097, 2006.

57 HOLMES, C. F.; MAYNES, J. T.; PERREAULT, K. R.; DAWSON, J. F. JAMES, M. N. Molecular enzymology underlying regulation of protein phosphatase-1 by natural toxins. **Current Medicinal Chemistry**, v. 9, n. 22, p. 1981-1989, 2002.

58 JUDITH, C.; MICHELS, P. A. M.; HANNAERT, V. Genetic validation of aldolase and glyceraldehyde-3-phosphate dehydrogenase as drug targets in *Trypanosoma brucei*. **Molecular and Biochemical Parasitology**, v. 169, n. 1, p. 50-54, 2010.

59 GEFFLAUT, T.; BLONSKI, C.; PERIÉ, J.; WILLSON, M. Class I aldolases: substrate specificity mechanism inhibitors and structural aspects. **Progress in Biophysics and Molecular Biology**, v. 63, n. 3, p. 301-340, 1995.

60 GEFFLAUT, T.; BLONSKI, C.; PERIÉ, J. Slow reversible inhibitions of rabbit muscle aldolase with substrate analogues: synthesis enzymatic kinetics and UV difference spectroscopy studies. **Bioorganic & Medicinal Chemistry**, v. 4, n. 12, p. 2043-2054, 1996.

61 CHUDZIK, D. M.; MICHELS, P. A. M.; DE WALQUE, S.; HOL, W. G. J. Structures of type 2 peroxisomal targeting signals in two trypanosomatid aldolases. **Journal of Molecular Biology**, v. 300, n. 4, p. 697-707, 2000.

62 ST-JEAN, M.; LAFRANCE-VANASSE, J.; LIOTARD, B.; SYGUSCH, J. High resolution reaction intermediates of rabbit muscle fructose-1,6-bisphosphate aldolase substrate cleavage and induced fit. **Journal of Biological Chemistry**, v. 280, n. 29, p. 27262-27270, 2005.

63 HORROBIN, D. F. Innovation in the pharmaceutical industry. **Journal of the Royal Society of Medicine**, v. 93, n. 7, p. 341-345, 2000.

64 KAPETANOVIC, I. M. Computer-aided drug discovery and development (CADD) in silico-chemico-biological approach. **Chemico-biological interactions**, v. 171, n. 2, p. 165-176, 2008.

65 LIN, J. H. Accommodating protein flexibility for structure-based drug design. **Current Topics in Medicinal Chemistry**, v. 11, n. 2, p. 171-178, 2011.

66 WATERBEEMD, H. **Advanced computer assisted techniques in drug discovery**. New York: VCH, 1994.

67 SHOICHET, B. K. Virtual screening of chemical libraries. **Nature**, v. 432, n. 7019, p. 862-865, 2004.

68 KRAMER, B.; RAREY, M.; LENGAUER, T. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. **Proteins: structure, function, and bioinformatics**, v. 37, n. 2, p. 228-241, 1999.

69 JONES, G.; WILLETT, P.; GLEN, R. C.; LEACH, A. R.; TAYLOR, R. Development and validation of a genetic algorithm for flexible docking. **Journal of Molecular Biology**, v. 267, n. 3, p. 727-748, 1997.

70 KITCHEN, D. B.; DECORNEZ, H.; FURR, J. R. ; BAJORATH, J. Docking and scoring in virtual screening for drug discovery methods and applications. **Nature Reviews Drug Discovery**, v. 3, n. 11, p. 935-949, 2004.

71 LYNE, P. D. Structure-based virtual screening an overview. **Drug Discovery Today**, v. 7, n. 20, p. 649-657, 2002.

72 SIRCI, F.; ISTYASTONO, E. P.; VISCHER, H. F.; KOOISTRA, A. J.; NIJMEIJER, S.; KUIJER, M.; WIJTMANS, M.; MANNHOLD, R.; LEURS, R.; de ESCH, I. J. P.; de GRAAF, C. Virtual fragment screening discovery of histamine H(3) receptor ligands using ligand-based and protein-based molecular fingerprints. **Journal of Chemical Information and Modeling**, v. 52, n. 12, p. 3308-3324, 2012.

73 AZEVEDO, L.; SANTANA, M. F. P.; XAVIER, M. M.; OZORIO, E. P.; VILLAVICENCIO, B.; FINCK, J. A.; PROENCA, A. M.; ROCHA, K. B.; DE AZEVEDO, W. F. Recent progress of molecular docking simulations applied to development of drugs. **Current Bioinformatics**, v. 7, n. 4, p. 352-365, 2012.

74 FERREIRA, R. S.; GUIDO, R. V. C.; ANDRICOPULO, A. D.; OLIVA, G. In silico screening strategies for novel inhibitors of parasitic diseases. **Expert Opinion on Drug Discovery**, v. 6, n. 5, p. 481-489, 2011.

75 WASZKOWYCZ, B.; CLARK, D. E.; GANCIA, E. Outstanding challenges in protein-ligand docking and structure-based virtual screening. **Wiley Interdisciplinary Reviews-computational molecular science**, v. 1, n. 2, p. 229-259, 2011.

76 IRWIN, J. J.; SHOICHET, B. K. ZINC a free database of commercially available compounds for virtual screening. **Journal of Chemical Information and Modeling**, v. 45, n. 1, p. 177-182, 2005.

77 SHOICHET, B. K.; KUNTZ, I. D. Protein docking and complementarity. **Journal of Molecular Biology**, v. 221, n. 1, p. 327-346, 1991.

78 DURRANT, J. D.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. **BMC Biology**, v. 9, n. 71, p. 1-9, 2011.

79 HARVEY, M. J.; DE FABRITIIS, G. High-throughput molecular dynamics the powerful new tool for drug discovery. **Drug Discovery Today**, v. 17, n. 19-20, p. 1059-1062, 2012.

80 NICHOLS, S. E.; BARON, R.; IVETAC, A.; MCCAMMON, J. A. Predictive power of molecular dynamics receptor structures in virtual screening. **Journal of Chemical Information and Modeling**, v. 51, n. 6, p. 1439-1446, 2011.

81 PHILLIPS, J. C.; BRAUN, R.; WANG, W.; GUMBART, J.; TAJKHORSHID, E.; VILLA, E.; CHIPOT, C.; SKEEL, R. D.; KALÉ, L.; SCHULTEN, K.; Scalable molecular dynamics with NAMD. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1781-1802, 2005.

- 82 BROOKS, B. R.; BRUCCOLERI, R. E.; OLAFSON, B. D.; STATES, D. J.; SWAMINATHAN, S.; KARPLUS, M.; CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational Chemistry**, v. 4, n. 2, p. 187-217, 1983.
- 83 KIM, S. K.; LI, Y.; ABROL, R.; HEO, J.; GODDARD, W.A. 3rd. Predicted structures and dynamics for agonists and antagonists bound to serotonin 5-HT_{2B} and 5-HT_{2C} receptors. **Journal of Chemical Information and Modeling**, v. 51, n. 2, p. 420-433, 2011.
- 84 TROPSHA, A. Best practices for QSAR model development, validation, and exploitation. **Molecular Informatics**, v. 29, n. 6-7, p. 476-488, 2010.
- 85 PALMEIRA, A.; SOUSA, E.; H.; VASCONCELOS, M.; PINTO, M.; FERNANDES, M. X. Structure and ligand-based design of P-glycoprotein Inhibitors: a historical perspective. **Current Pharmaceutical Design**. v. 18, n. 27, p. 4197-4214, 2012.
- 86 FERREIRA, L. G.; ANDRICOPULO, A. D. Structure- and ligand-based structure-activity relationships for a series of inhibitors of aldolase. **Current Computer-aided drug design**, v. 8, n. 4, p. 309-316, 2012.
- 87 FREITAS, H. F.; POSTIGO, M. P.; ANDRICOPULO, A. D.; CASTILHO, M. S. Descriptor- and fragment-based QSAR models for a series of *Schistosoma mansoni* purine nucleoside inhibitors. **Journal of the Brazilian Chemical Society**. v. 22, n. 9, p. 1718-1726, 2011.
- 88 GUIDO, R. V. C.; OLIVA, G.; ANDRICOPULO, A. D. Modern drug discovery technologies opportunities and challenges in lead discovery. **Combinatorial Chemistry & High Throughput Screening**, v. 14, n. 10, p. 830-839, 2011.
- 89 FERREIRA, L. G.; LEITAO, A.; MONTANARI, C. A. ANDRICOPULO, A. D. Comparative molecular field analysis of a series of inhibitors of HIV-1 protease. **Medicinal Chemistry**, v. 7, n. 2, p. 71-79, 2011.
- 90 TERPE, K. Overview of tag protein fusions from molecular and biochemical fundamentals to commercial systems. **Applied Microbiology and Biotechnology**. v. 60, n. 5, p. 523-533, 2003.
- 91 BALDI, L.; HACKER, D. L.; ADAM, M.; WURM, F. Recombinant protein production by large-scale transient gene expression in mammalian cells state of the art and future perspectives. **Biotechnology Letters**, v. 29, n. 5, p. 677-684, 2007.
- 92 SCHMIDT, .F. R. Recombinant expression systems in the pharmaceutical industry. **Applied Microbiology and Biotechnology**, v. 65, n. 4, p. 363-372, 2004.

- 93 SØRENSEN, H. P.; MORTENSEN, K. K. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. **Journal of Biotechnology**, v. 115, n. 2, p. 113–128, 2005.
- 94 BANEY, F.; MUJACIC, M. Recombinant protein folding and misfolding in *Escherichia coli*. **Nature Biotechnology**, v. 22, n. 11, 1399 – 1408, 2004.
- 95 WOESTENENK, E. A.; HAMMARSTRÖM, M.; VAN DEN BERG, S.; HÄRD, T.; BERGLUND, H. His tag effect on solubility of human proteins produced in *Escherichia coli* a comparison between four expression vectors. **Journal of Structural and Functional Genomics**, v. 5, n. 3, p. 217–229, 2004.
- 96 LINDNER, P.; GUTH, B.; WÜLFING, C.; KREBBER, C.; STEIPE, B.; MÜLLER, F.; PLÜCKTHUN, A. Purification of native proteins from the cytoplasm and periplasm of *Escherichia coli* using IMAC and histidine tails a comparison of proteins and protocols. **Methods**, v. 4, n. 1, p. 41–56, 1992.
- 97 BATAS, B.; SCHIRALDIB, C.; CHAUDHURI, J. B. Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. **Journal of Biotechnology**, v. 68, n. 2–3, p. 149–158, 1999.
- 98 JANCARIK, J.; KIM, S. H. Sparse matrix sampling a screening method for crystallization of proteins. **Journal of Applied Crystallography**, v. 24, p. 409–411, 1991.
- 99 McPHERSON, A. Current approaches to macromolecular crystallization. **European Journal of Biochemistry/FEBS**, v. 189, n. 1, p. 1–23, 1990.
- 100 GODDARD, J-P.; REYMOND, J-L. Enzyme assays for high-throughput screening. **Current Opinion in Biotechnology**, v. 15, n. 4, p. 314–322, 2004.
- 101 CLELAND, W. W. Optimizing coupled enzyme assays. **Analytical Biochemistry**, v. 99, n. 1, p. 142–145, 1979.
- 102 EASTERBY, J. S. Coupled enzyme assays a general expression for the transient. **Biochimica et Biophysica Acta – enzymology**, v. 293, n. 2, p. 552–558, 1973.
- 103 McCLURE, W. R. Kinetic analysis of coupled enzyme assays. **Biochemistry**, v. 8, n. 7, p. 2782–2786, 1969.
- 104 COPELAND, R. A. **Enzymes a practical introduction to structure mechanism and data analysis**. New York: Wiley, 2000.

105 KRUGER, N. J. Errors and artifacts in coupled spectrophotometric assays of enzyme activity. **Phytochemistry**, v. 38, n. 5, p. 1065–1071, 1995.

106 GARFIN, D. E. Two-dimensional gel electrophoresis: an overview. **TrAC Trends in Analytical Chemistry**, v. 22, n. 5, p. 263–272, 2003.

107 BORCHHARDT, D. M.; ANDRICOPULO, A. D. CoMFA and CoMSIA 3D QSAR models for a series of cyclic imides with analgesic activity. **Medicinal Chemistry**, v. 5, n. 1, p. 66-73, 2009.

108 OGATA, H.; TAKEO, K.; KUWAHARA, A.; SUZUNO, R.; FUJIMOTO, M. SHIMIZU, J. An exploration of the binding site of aldolase using alkanediol monoglycolate bisphosphoric esters. **Biochimica et Biophysica Acta**, v. 742, n. 2, p. 384-390, 1983.

109 OGATA, H.; FUKUDA, T.; YAMAMOTO, K.; YAMASAKI, H.; FUJIMOTO, H.; FUJISAKI, S.; KAJIGAESHI, S. An exploration of the binding site of aldolase using n-(omega-hydroxyalkyl) glycolamido bisphosphoric esters. **Biochimica et Biophysica Acta**, v. 1041, n. 3, p. 254-256, 1990.

110 OGATA, H.; FUKUDA, T.; YAMAMOTO, K.; FUNAKOSHI, J.; TAKADA, K.; YASUE, N.; FUJISAKI, S.; KAJIGAESHI, S. An exploration of the binding site of aldolase using alkyl glycolamido phosphoric esters and alkyl monoglycolate phosphoric esters. **Biochimica et Biophysica Acta**, v. 1119, n. 2, p. 123-126, 1992.

111 FERREIRA, L. G.; LEITÃO, A.; MONTANARI, C. A.; ANDRICOPULO, A. D. Hologram quantitative structure-activity relationships for a class of inhibitors of HIV-1 protease. **Letters in Drug Design & Discovery**, v. 4, n. 5, p. 356-364, 2007.

112 CLARK, M.; CRAMER III, R. D.; VAN OPDENBOSCH, N. Validation of the general purpose tripos 5.2 force field. **Journal of Computational Chemistry**, v. 10, n. 8, p. 982–1012, 1989.

113 STEWART, J. J. P. MOPAC a semiempirical molecular orbital program **Journal of Computer Aided Molecular Design**, v. 4, n. 1, p. 1-105, 1990.

114 BESLER, B. H.; MERZ JR., K. M.; KOLLMAN, P. A. Atomic charges derived from semiempirical methods. **Journal of Computational Chemistry**, v. 11, n. 4, p. 431–439, 1990.

115 MABIALA-BASSILOUA, C.; ARTHUS-CARTIER, G.; HANNAERT, V.; THERISOD, H.; SYGUSCH, J.; THERISOD, M. Mannitol bis-phosphate based inhibitors of fructose 1,6-bisphosphate aldolases. **ACS Medicinal Chemistry Letters**, v. 2, n. 11, p. 804-808, 2011.

116 HAIECH, J.; KOSCIELNIAK, T.; GRASSY, G. J. Use of TSAR as a new tool to analyze the molecular dynamics trajectories of proteins. **Journal of Molecular Graphics**, v. 13, n. 1, p. 46-48, 1995.

117 BUSH, B. L.; NACHBAR, R. B. Jr. Sample-distance partial least squares PLS optimized for many variables, with application to CoMFA. **Journal of Computer Aided Molecular Design**, v. 7, n. 5, p. 587-619, 1993.

118 MAGHSOUDI, A. H.; KHODAGHOLI, F.; HADI-ALIJEANVAND, H.; ESFANDIAREI, M.; SABBAGHIAN, M.; ZAKERI, Z.; SHAERZADEH, F.; ABTAHI, S.; MAGHSOUDI, N. Homology modeling, docking molecular dynamics simulation and structural analyses of coxsakievirus B3 2A protease an enzyme involved in the pathogenesis of inflammatory myocarditis. **International Journal of Biological Macromolecules**, v. 49, n. 4, p. 487-492, 2011.

119 WLODAWER, A.; VONDRASEK, J. Inhibitors of HIV-1 protease a major success of structure-assisted drug design. **Annual Review of Biophysics and Biomolecular Structure**, v. 27, 249-284, 1998.

120 ZHANG, H.; ZAN, J.; YU, G.; JIANG, M.; LIU, P. A combination of 3D-QSAR, molecular docking and molecular dynamics simulation studies of benzimidazole-quinolinone derivatives as iNOS inhibitors. **International Journal of Molecular Sciences**, v. 13, n. 9, p. 11210–11227, 2012.

121 Nelson, M. T. NAMD: a parallel, object-oriented molecular dynamics program. **International Journal of High Performance Computing Applications**, v. 10, n. 4, p. 251–268, 1996.

122 LINDORFF-LARSEN, K.; MARAGAKIS, P.; PIANA, S.; EASTWOOD, M. P.; DROR, R. O.; SHAW, D. E. Systematic validation of protein force fields against experimental data. **PLoS ONE**, v. 7, n. 2, p. e32131, 2012.

123 ZOETE, V.; CUENDET, M. A.; GROSDIDIER, A.; MICHIELIN, O. J. SwissParam a fast force field generation tool for small organic molecules. **Computational Chemistry**, v.32, n. 11, p. 2359-68, 2011.

124 KOSZTIN, D.; IZRAILEV, S.; SCHULTEN, K. Unbinding of retinoic acid from its receptor studied by steered molecular dynamics. **Biophysical Journal**, v. 76, n. 1, p. 188-97, 1999.

125 DARIA, A.; OSOŁODKIN, D. I.; SHULGA, D. A.; OLIFERENKO, A. A.; PISAREV, S. A.; PALYULIN, V. A.; ZEFIROV, N. S. General purpose electronegativity relaxation charge models applied to comfa and comsia study of gsk-3 inhibitors. **Molecular Informatics**, v. 30, n. 2-3, p. 169–80, 2011.

126 LONG, D.; MU, Y.; YANG, D. Molecular dynamics simulation of ligand dissociation from liver fatty acid binding protein. **PLoS One**, v. 4, n. 6, p. e6081, 2009.

127 SUN, T. G.; HU, J. P.; LI, C. H.; CHEN, W. Z.; WANG, C. X. A molecular dynamics simulation study of glutamine-binding protein. **Journal of Molecular Structure THEOCHEM**, v. 725, n. 1-3, p. 9-16, 2005.

128 ESSMAN, U.; PERERA L.; BERKOWITZ, M. L.; DARDEN, T.; LEE, H.; PEDERSEN, L. G. A smooth particle mesh ewald method. **Journal of Chemical Physics**, v. 103, n. 19, p. 8577– 8593, 1995.

129 MARTÍNEZ, L.; POLIKARPOV, I.; SKAF, M. S. Only subtle protein conformational adaptations are required for ligand binding to thyroid hormone receptors simulations using a novel multipoint steered molecular dynamics approach. **Journal of Physical Chemistry B**, v. 112, n. 34, p. 10741-10751, 2008.

130 YANG, Y.; QIN, J.; LIU, H.; YAO, X. Molecular dynamics simulation free energy calculation and structure-based 3d-qsar studies of b-raf kinase inhibitors. **Journal of Chemical Information and Modeling**, v. 51, n. 3, p. 680–692, 2011.

131 FELLER, S. E.; ZHANG, Y.; PASTOR, R. W.; BROOKS, B. R. Constant pressure molecular dynamics simulation the langevin piston method. **Journal of Chemical Physics**, v. 103, n. 11, p. 4613-4621, 1995.

132 DALBY, A R.; TOLAN, D. R.; LITTLECHILD, J. A. The structure of human liver fructose-1,6-bisphosphate aldolase. **Acta Crystallographica D - biological crystallography**, v. 57, n. 11, p. 1526-1533, 2001.

133 LAFRANCE-VANASSE, J.; SYGUSCH, J. Carboxy-terminus recruitment induced by substrate binding in eukaryotic fructose bis-phosphate aldolases. **Biochemistry**, v. 46, n. 33, p. 9533-9540, 2007.