Geraldo Rodrigues Sartori

**Planejamento de Inibidores Baseado em Fragmentos Moleculares para a
Enzima Gliceraldeído-3-fosfato Desidrogenase de *Trypanosoma cruzi***

Dissertação apresentada ao Programa de Pós-Graduação do Instituto de Química de São Carlos, da Universidade de São Paulo, para a obtenção do título de mestre em Ciências.

Área de concentração: Química Orgânica e Biológica

Orientador: Prof. Dr. Carlos Alberto Montanari

São Carlos, 2012

Exemplar revisado.

**O exemplar original encontra-se disponível no
Serviço de Pós-Graduação do IQSC-USP.**

Agradecimentos

Aos meus pais Geraldo e Nadir pela educação, pelo amor, carinho, incentivo e conversas, até mesmo de âmbito científico;

Ao meu irmão Gustavo pela amizade, pela confiança, por ser irmão;

Ao Prof. Dr. Carlos A. Montanari, pela orientação, desafios, pelo apoio e também pela amizade. À sua esposa, Dr. Malu Montanari, pelo carinho, pelas conversas;

Aos meus amigos e colegas do NEQUIMED, Igor Prokopczyk, Dr. Josmar Rocha, Jean Ribeiro, Leandro Avelar, Ricardo Goulart, Alexander Patiño, Cristian Reyes, Ms. Fabiana Rosini, Ms. Emanuella Barreto, Ms. Fabyana Soares, Prof. Dr. Andrei Leitão e demais integrantes do grupo pela convivência, conversas, apoio, churrascos, comemorações e discussões científicas bastante construtivas;

Aos meus amigos Lennon David, Murilo Moreira, Rafael Moraes, Felipe Cardoso, Thay Tavares, Ivan Silva, Ju Cheleski, Helton Wiggers, Sabrina Matos, Vinicius Sciutti, Rodrigo Berté, Atilio Tomazini, e todos os demais, pelos momentos de descontração, de alegria.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pela bolsa concedida durante os dois anos de Mestrado.

Às demais agências de fomento pelo apoio financeiro concedido ao NEQUIMED e essencial na execução desse trabalho.

Epígrafe

“A VIDA É UMA PEÇA DE TEATRO QUE NÃO PERMITE ENSAIOS. POR ISSO CANTE, CHORE, RIA E VIVA INTENSAMENTE ANTES QUE A CORTINA SE FECHE E A PEÇA TERMINE SEM APLAUSOS.”

Charlie Chaplin

Resumo

A Doença de Chagas, endêmica na América Latina, é causada pelo parasito tripanossomatídeo *Trypanosoma cruzi* e atualmente já se espalha para o restante do mundo devido à migração humana. Os dois medicamentos disponíveis para o tratamento dessa doença, o Nifurtimox, (banido do Brasil), e o Benzonidazol, são eficazes somente na etapa aguda da doença e possuem efeitos colaterais severos. Recentemente, três novas substâncias para o tratamento chegaram à fase clínica de testes contra essa doença, mas ainda é necessária a pesquisa de novas moléculas contra esse parasito. A enzima Gliceraldeído-3-fosfato Desidrogenase (GAPDH) foi selecionada como alvo para busca de moléculas potencialmente tripanossomicidas. De forma a inibir essa enzima, a busca de moléculas baseou-se na abordagem de fragmentos moleculares para encontrar substâncias com elevada eficiência de ligante. A partir de um banco de dados comercial de 500 mil moléculas, filtros moleculares de solubilidade e da Regra dos Três foram aplicados para montar uma biblioteca focalizada de moléculas. Essa biblioteca foi submetida a estudos integrados baseados na estrutura do alvo macromolecular (docagem) e do substrato da enzima (similaridade química e eletrostática) e a inspeção visual dos fragmentos bem classificados em ambas técnicas foi realizada de modo a selecionar cinco compostos de classes químicas diversas. Essas substâncias foram submetidas a ensaios enzimáticos *in vitro* por meio de espectroscopia de fluorescência, sendo encontrado então um fragmento com K_i de $(425 \pm 53) \mu\text{M}$ e eficiência de ligante de 0,33, valor bastante promissor para abordagem baseada em fragmentos moleculares. Estudos de simulação por Dinâmica Molecular (MD) foram feitos para as cinco moléculas adquiridas, com energia de interação ligante-enzima calculada usando o método MM-GB/SA. A classificação das moléculas por essa energia foi idêntica à obtida experimentalmente. Além disso, a MD possibilitou a predição de modo de interação dos fragmentos no sítio ativo da TcGAPDH e a identificação de uma nova cavidade passível de modulação de sua atividade. Uma segunda série de fragmentos foi selecionada baseada no fragmento ativo de modo a construir uma relação estrutura-atividade (SAR) teórica por MD. A SAR sugere que a presença de um átomo nitrogênio capaz de doar ligações de hidrogênio é importante para a interação, com o resíduo de aminoácido Asp210. Este resíduo de aminoácido desponta como um possível ponto de seletividade para a enzima humana, que possui uma leucina nessa posição. Além disso, a posição na substituição do anel central também está diretamente relacionada à interação da molécula com a enzima, com

uma substituição 2,3 em um anel de cinco membros a mais favorável. Este trabalho identificou pela primeira vez um fragmento molecular com alta eficiência de ligante para a enzima TcGAPDH, com o auxílio do uso conjunto de técnicas baseadas na estrutura do ligante e do alvo, para seleção de moléculas, e espectroscopia de fluorescência para identificação de atividade inibitória frente à enzima. Método de simulação por MD conseguiu reproduzir os resultados experimentais e prover informações teóricas de SAR para o composto ativo.

Abstract

Chagas disease is a parasitic illness endemic in Latin America caused by the trypanosomatid parasite *Trypanosoma cruzi* that spreads around the world due to people migration. Nowadays, Benznidazole and Nifurtimox (banned in Brazil), are used for the treatment of this disease but causes severe side effects to patients. Recently, three new molecules have reached clinical trials phase in the development of drugs against Chagas disease but it is still necessary to develop new drugs. In this studies, the enzyme Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH) was used as a target for the search of new antitrypanosomatid molecules. It belongs to the glycolytic pathway, the major one for parasite's energy generation. With the aim searching a new molecule that inhibits this enzyme, the fragment-based approach guided the search of molecules with high ligand efficiency. A focused compound library was assembled from a database of 500,000 molecules using molecular and solubility filters and the Rule of Three. The integrated use of ligand (chemical and electrostatic similarity) and target (molecular docking) based drug design was carried out to rank the molecular fragments by a consensual score. Through visual inspection of the top 500 molecules five diverse fragments were selected for the *in vitro* enzymatic assays using fluorescence spectroscopy. One of these molecules shows a K_i equals to $(425 \pm 53) \mu\text{M}$ and ligand efficiency equals to 0,33, a promising value for the fragment-based approach. Additionally, Molecular Dynamics simulations (MD) were carried out with these fragments and the predicted energy of interaction for fragment-enzyme complex was able to rank the molecules as using the experimental results. Furthermore, the MD was useful to predict the mode of interaction of the fragments in the active site of enzyme and to reveal a new cavity close to the substrate binding site. A second generation of compounds was selected based on the structure of the active fragment to construct theoretical structure-activity relationship (SAR) using MD. SAR indicates that the presence of a nitrogen with hydrogen-bond donor property is important to the interaction, making hydrogen-bonding with the amino acid residue Asp210. In addition, MD shows the influence of different substituent position in the central ring in the energy of the interaction, with a 2,3 substitution at five-atom ring the most favorable. This study identifies the first molecular fragment with high ligand efficiency for the enzyme TcGAPDH, with the combined use of ligand and target-based tools and fluorescence spectroscopy, for selection and identification of active compounds against the enzyme. MD was able to reproduce experimental

results and generate theoretical information of SAR to the active molecular fragment

Lista de Figuras

Figura 1.1 - Ciclo de vida do <i>Trypanosoma cruzi</i> , mostrando as fases de vida, com as modificações morfológicas em diferentes tecidos, tanto em mamíferos quanto em outros triatomídeos.....	24
Figura 1.2 - Evolução do número de parasitas e número de anticorpos na corrente sanguínea, em função do tempo decorrido após a picada	25
Figura 1.3 - Expansão da doença de Chagas pelo mundo a partir de países endêmicos.....	28
Figura 1.4-Estrutura química das moléculas já em fase terapêutica e moléculas em fase clínica para a Doença de Chagas	32
Figura 1.5 - O caminho do fármaco no organismo, desde sua administração até a eliminação.....	34
Figura 1.6 - Evolução da molécula durante a gênese de fármacos	37
Figura 1.7 - Distribuição dos alvos biomacromoleculares para os fármacos atualmente na terapêutica	38
Figura 1.8 - Fases pré-clínica e clínica necessária para a aprovação de uma molécula como fármaco	41
Figura 1.9 - Via glicolítica e sua distribuição em insetos Triatomíneos	43
Figura 1.10 - Representações da enzima GAPDH, código PDB 1QXS. A) Cada cor representando um monômero da TcGAPDH. B) Em azul, domínio de Rossmann, em verde, domínio C-terminal e em vermelho, sítio catalítico.	45
Figura 1.11 - Esquema da reação catalisada pela enzima GAPDH	46
Figura 1.12 - Esquematização do mecanismo catalítico de reação da enzima GAPDH.....	47
Figura 1.13 - Representação esquemática do espaço químico (cubo) e do espaço químico biologicamente relevante (elipsóides).....	49
Figura 1.14 - Esquema de diferentes estratégias de abordagem para o planejamento <i>in silico</i> de moléculas bioativas	52
Figura 1.15 - Influência da complexidade química na interação da pequena molécula com o sítio ativo da biomacromolécula	54
Figura 1.16 - Limite de detecção de afinidade para diferentes técnicas e comparação com valores médios de afinidade para cada tamanho de molécula	56
Figura 1.17 - Representação de diferentes métodos de crescimento de fragmento. A) Ligação de fragmentos. B) Auto ligação de fragmentos. C) Crescimento a partir de uma âncora. D) Otimização de fragmento.	57
Figura 1.18 - Comparação de fragmentos moleculares com moléculas de maior massa molecular e delimitação da área de alta eficiência de ligante.	59
Figura 3.1 - Esquema representando as etapas da docagem molecular: usando a estrutura de um receptor e um conjunto de moléculas, os programas de docagem molecular avaliam a complementaridade do ligante em relação ao sítio ativo do receptor. Aquelas mais bem pontuadas são avaliadas numa próxima etapa através de um ensaio experimental.	64
Figura 3.2 - Tempo de ocorrência de fenômenos em um sistema proteico	69
Figura 3.3 - Procedimento para o calculo de trajetória de átomos utilizando simulação por Dinâmica Molecular	70
Figura 3.4 - Fluxograma esquematizado da etapa de seleção de moléculas para testes <i>in vitro</i> e de simulação por Dinâmica Molecular	75
Figura 3.5 - Rampa de aquecimento utilizada para o processo de <i>Simulated Annealing</i>	84
Figura 4.1 - Distribuição do número de moléculas do banco de dados original, em vermelho e do banco de dados de solubilidade, em verde, em função do a) número de aceitadores de ligação de hidrogênio, b) número de doadores de ligação de hidrogênio, c) massa molecular, d) ligações de rotação livre, e) coeficiente de partição, cLogP, f) área superficial polar.	88
Figura 4.2 - Esquematização da interação dos ligantes co-cristalizados para as estruturas de código PDB a) 3CIF, b)1QXS, c)1K3T.....	91
Figura 4.3 - a) Representação do sítio da estrutura 1QXS preparada para docagem molecular utilizando o FlexX. Aminoácidos mostrados são a SER247 e THR167 b) Átomos em bastão mostrando os resíduos de aminoácidos Ser247 (abaixo) e Thr167. Em branco, a estrutura da GAPDH humana (1U8F) e em rosa, a estrutura da TcGAPDH (1QXS) c) Alinhamento unidimensional entre as sequencias de aminoácidos da enzima GAPDH humana com a enzima GAPDH de <i>Trypanosoma cruzi</i>	94
Figura 4.4 - Distribuição do número de moléculas do banco de solubilidade, em vermelho e das moléculas que	

passaram pelo filtro do FlexX, em verde, em função do a) número de aceitadores de ligação de hidrogênio, b) número de doadores de ligação de hidrogênio, c) massa molecular, d) ligações de rotação livre, e) coeficiente de partição, cLogP, f) área superficial polar.....	96
Figura 4.5 - Distribuição de valores de energia de docagem molecular predito pelo programa Glide-SP.....	99
Figura 4.6 - Distribuição de valores de energia de docagem molecular predita pelo programa <i>Glide-XP</i>	100
Figura 4.7 - Estrutura bidimensional das moléculas selecionadas na primeira geração para o sítio do G3P..	102
Figura 4.8 - Gráfico de Michaelis-Menten para a molécula NEQUIMED 293.....	104
Figura 4.9 - Gráfico de RMSD em função do tempo para a simulação por MD para o sistema contendo somente a enzima TcGAPDH.	106
Figura 4.10 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator, levando em conta a posição dos átomos a) da proteína b) do cofator NAD ⁺	107
Figura 4.11 - Conformação e posicionamento do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.....	108
Figura 4.12 - Conformação e posicionamento da molécula NEQUIMED 293 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.....	110
Figura 4.13 - Conformação e posicionamento da molécula NEQUIMED 292-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.....	111
Figura 4.14 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 295-3, levando em conta a posição dos átomos do ligante.....	112
Figura 4.15 - Conformação e posicionamento da molécula NEQUIMED 295-3 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.....	112
Figura 4.16 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 358T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho	113
Figura 4.17 - Conformação e posicionamento da molécula NEQUIMED 294-2 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.....	114
Figura 4.18 - Conformação e posicionamento da molécula NEQUIMED 291 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.....	115
Figura 4.19 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 358T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 358T-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em verde) e Ser224 (em ciano).	119
Figura 4.20 - Conformação e posicionamento da molécula NEQUIMED 358T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.	120
Figura 4.21 - Conformação e posicionamento da molécula NEQUIMED 357T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação	121
Figura 4.22 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 357T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho. b) Gráfico de energia de interação da molécula NEQUIMED 357T-1 calculado por MMGBSA para cada passo da simulação, após a etapa de arrefecimento simulado.	121
Figura 4.23 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 354T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho. b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 354T-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em vermelho)	122
Figura 4.24 - Conformação e posicionamento da molécula NEQUIMED 354T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.	123
Figura 4.25 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 359T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho. b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 359T-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em azul), Ser224 (em vermelho) e Thr167 (em ciano).....	124
Figura 4.26 - Conformação e posicionamento da molécula NEQUIMED 359T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação	125
Figura 4.27 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 350-1, levando em conta a posição dos átomos do ligante, em preto, e do	

NAD ⁺ , em vermelho.b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 350-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em ciano) e Ser224 (em azul).....	126
Figura 4.28 - Conformação e posicionamento da molécula NEQUIMED 350-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.....	127
Figura 4.29 - Conformação e posicionamento da molécula NEQUIMED 360T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.	127
Figura 4.30 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 360T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho.b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 360T-1 com os resíduos de aminoácidos Ser247 (em preto) e Thr167 (em vermelho).....	128
Figura 4.31 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 356T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 356T-1 com o resíduo de aminoácido Ser247.	129
Figura 4.32 - Conformação e posicionamento da molécula NEQUIMED 356T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.	130
Figura 4.33 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 361T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 361T-1 com os resíduos de aminoácidos Ser247 (em preto), Asp210 (em Pink) e Arg249 (em ciano).....	130
Figura 4.34 - Conformação e posicionamento da molécula NEQUIMED 361T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.	131
Figura 4.35 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 355T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD ⁺ , em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 355T-1 com o resíduo de aminoácido Asp210.	132
Figura 4.36 - Conformação e posicionamento da molécula NEQUIMED 355T-1 e do cofator NAD ⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.	132

Lista de Tabelas

Tabela 1.1 - Inibição enzimática necessária para levar à diminuição de 50% na produção de glicose em <i>Trypanosoma brucei</i> . ALD, Frutose-1,6-difosfato aldolase; GDH, Glicerol-3-fosfato desidrogenase; GAPDH, Gliceraldeído-3-fosfato desidrogenase; PGK, 3-fosfoglicerato cinase; HK, Hexocinase; PFK, 6-fosfofruto-cinase; PYK, Piruvato cinase	44
Tabela 3.1 - Principais parâmetros definidos como entrada para o programa FILTER	76
Tabela 3.2 - Parâmetros modificados ao se utilizar o programa OMEGA	80
Tabela 3.3 - Concentração de D,L-G3P utilizada em cada coluna da placa de 96 poços para ensaio de inibição enzimática utilizando espectroscopia de fluorescência	82
Tabela 4.1 - Informações cristalográficas das estruturas da TcGAPDH depositadas no <i>PDB</i>	90
Tabela 4.2 - Estado de protonação e tautomerização dos resíduos de aminoácidos de Histidina para a enzima TcGAPDH.....	93
Tabela 4.3 - Dados estatísticos da distribuição de valores de energia, em kcal/mol, de docagem molecular preditos pelo programa <i>Glide-SP</i>	98
Tabela 4.4 - Dados estatísticos da distribuição de valores de energia, em kcal/mol de docagem molecular predita pelo programa <i>Glide-XP</i>	100
Tabela 4.5 - Dados de constante de inibição obtidos para as moléculas de primeira geração em ensaio enzimático contra a TcGAPDH utilizando espectroscopia de fluorescência.....	104
Tabela 4.6 - Energia de interação calculada computacionalmente por MMPBSA para as moléculas de primeira geração.	109
Tabela 4.7 - Estrutura bidimensional das moléculas selecionadas na segunda geração para o sítio do G3P...116	
Tabela 4.8 – Valor médio de energia de interação calculada computacionalmente por MMPBSA para as moléculas de segunda geração	117

Lista de Abreviaturas

μM	micromolar, 10^{-6} M, unidade de medida de concentração
1,3 BPG	1,3-bisfosfoglicerato
Å	angstrom, 10^{-10} metros, unidade de medida de distância
ADME/Tox	Administração, Distribuição, Metabolismo e Toxicidade
ADN	ácido desoxirribonucleico
ALD	Frutose-1,6-difosfato aldolase
ANVISA	Agência Nacional de Vigilância Sanitária
Arg	resíduo de aminoácido Arginina
ARN	ácido ribonucleico
Asp	resíduo de aminoácido aspartato
ATP	adenosina trifosfato
CAS	<i>Chemical Abstract Society</i>
Cys	resíduo de aminoácido cisteína
Da	Dalton, unidade de massa atômica
MD	Dinâmica Molecular
DMSO	dimetilsulfóxido
E.C.	<i>Enzyme Commission</i> - Classificação das enzimas baseado nas reações catalisadas pelas mesmas
E1224	derivado monolisina do Ravuconazol
EDTA	ácido etilenodiamino tetra-acético
FDA	<i>Food and Drug Administration</i> , órgão estadunidense de regulamentação de fármacos
fs	femtosssegundo, 10^{-15} segundo, unidade de medida de tempo
G3P	Gliceraldeído-3-fosfato
GAPDH	Gliceraldeído-3-fosfato Desidrogenase
GDH	Glicerol-3-fosfato desidrogenase
HEPES	ácido 2-[4-(2-hidroxietil)1-piperazinil]-etanosulfônico
His	resíduo de aminoácido histidina
HK	Hexocinase
HTS	<i>High Throughput Screening</i>

IC ₅₀	Concetração necessária para inibir 50% da atividade
IgG	imunoglobulinas séricas
IgM	imunoglobulinas M
IND	Sigla do inglês para Investigação para Novo Fármaco
iOWH	<i>One World Health</i>
IQSC	Instituto de Química de São Carlos
ITC	sigla em inglês para Calorimetria de Titulação Isotérmica
IUPAC	sigla em inglês para União Internacional de Química Pura e Aplicada
K	Kelvin, unidade de medida de temperatura
K-11777	molécula em fase clínica I contra Doença de Chagas
Kcal.mol ⁻¹	quilocaloria por mol, unidade de medida de energia
K _d	Constante de dissociação
kg	quilograma
K _i	Constante de inibição
l	comprimento de ligação
LBDD	sigla em inglês para Planejamento baseado na estrutura do ligante
LE	sigla em inglês para Eficiência do Ligante
Leu	resíduo de aminoácido Leucina
mg	miligrama
mM	milimolar, 10 ⁻³ M, unidade de medida de concentração
MM	Massa molecular
NAD ⁺	Nicotinamida Adenina Dinucleotídeo
NADH	Nicotinamida Adenina Dinucleotídeo reduzida
NADP ⁺	Nicotinamida Adenina Dinucleotídeo Fosfato
NCE	sigla em inglês para novas substâncias química bioativas
NDA	Sigla do inglês para Submissão de Novo Fármaco
NEQUIMED	Grupo de estudos em Química Medicinal
NIAID	sigla em inglês para Instituto Nacional de Alergia e Doenças Infecciosas, órgão estadunidense
nM	nanomolar, 10 ⁻⁹ M, unidade de medida de concentração
ns	nanossegundo, 10 ⁻⁹ segundos, unidade de medida de tempo
OMS	Organização Mundial da Saúde

PDB	<i>Protein Data Bank</i>
PFK	6-fosfofruto-cinase
PGK	3-fosfoglicerato cinase
pH	medida de acidez e basicidade de uma solução aquosa
P _i	Fosfato inorgânico
pK _a	escala logarítmica da constante de acidez
PMF	Potencial de força média
ps	picossegundos, 10 ⁻¹² segundos, unidade de medida de tempo
PYK	Piruvato cinase
q	carga parcial atômica
QSAR	Relação quantitativa estrutura-atividade
R	Constante universal dos gases
r _{ij}	distância interatômica
RMN	Ressonância magnética nuclear
RMSD	Raiz Quadrada do Desvio Quadrático Médio
SA	Simulated Annealing, ou arrefecimento simulado
SAR	sigla em inglês para relação estrutura-atividade
SBDD	sigla em inglês para Planejamento baseado na estrutura do alvo
Ser	resíduo de aminoácido serina
SPR	sigla em inglês para relação estrutura-propriedade
SPR	sigla em inglês para Ressonância Plasmônica de Superfície
T	temperatura
T. cruzi	Trypanosoma cruzi
TcOYE	<i>Old Yellow Enzyme</i> de <i>Trypanosoma cruzi</i>
Thr	resíduo de aminoácido treonina
UDP	Uridina Difosfato
USP	Universidade de São Paulo
v/v	volume/volume
θ	ângulos de ligação
n_{ϕ}	parâmetro de deslocamento de fase

Conteúdo

1	INTRODUÇÃO	22
1.1	DOENÇA DE CHAGAS	22
1.1.1	<i>Formas de Transmissão.....</i>	22
1.1.2	<i>Ciclo de vida do parasito.....</i>	23
1.1.3	<i>Sintomatologia.....</i>	25
1.1.4	<i>Panorama atual.....</i>	27
1.1.5	<i>Tratamento.....</i>	29
1.2	QUÍMICA MEDICINAL E O DESENVOLVIMENTO DE FÁRMACOS.....	33
1.2.1	<i>O fármaco no organismo – fases de atividade.....</i>	33
1.2.1.1	Fase Farmacêutica.....	33
1.2.1.2	Fase Farmacocinética.....	33
1.2.1.3	Fase Farmacodinâmica	35
1.2.2	<i>A Química Medicinal no planejamento de fármacos.....</i>	36
1.2.2.1	Descoberta e Desenvolvimento de fármacos	37
1.2.2.1.1	A descoberta.....	37
1.2.2.1.2	O desenvolvimento.....	39
1.3	O ALVO BIOMACROMOLECULAR	41
1.3.1	<i>A via glicolítica.....</i>	41
1.3.2	<i>A enzima Gliceraldeído-3-Fosfato desidrogenase.....</i>	44
1.3.2.1	Informações estruturais.....	45
1.3.2.2	Reação catalisada pela GAPDH.....	45
1.3.2.3	Inibidores conhecidos	47
1.4	O ESPAÇO QUÍMICO E FERRAMENTAS PARA INVESTIGÁ-LO.....	48
1.4.1	<i>O espaço químico</i>	49
1.4.2	<i>O espaço químico biologicamente relevante</i>	49
1.4.3	<i>A quiminformática</i>	50
1.4.4	<i>Abordagens computacionais de modelagem baseada em campo de força molecular</i>	51
1.5	FRAGMENTOS MOLECULARES COMO FERRAMENTA PARA A DESCOBERTA DE MOLÉCULAS BIOATIVAS	52
1.5.1	<i>Histórico e fundamentação teórica.....</i>	53
1.5.2	<i>Fragmentos moleculares e fármaco-similares.....</i>	53
1.5.3	<i>Montagem de um banco de dados de fragmentos: Regra dos Três.....</i>	54
1.5.4	<i>Detecção de fragmentos moleculares como ligantes.....</i>	55
1.5.5	<i>Crescimento e otimização de um fragmento com atividade biológica.....</i>	56
1.5.6	<i>Eficiência do Ligante.....</i>	58
2	OBJETIVOS	60
3	MATERIAIS E MÉTODOS.....	62
3.1	MATERIAIS	62
3.2	MÉTODOS	63
3.2.1	<i>Ensaio virtual baseado na estrutura do alvo: Docagem molecular</i>	63
3.2.1.1	Representação do Alvo	64
3.2.1.2	Busca no espaço conformacional do ligante	65
3.2.1.2.1	Algoritmo estocástico	65
3.2.1.2.2	Algoritmo sistemático ou construção incremental.....	66
3.2.1.3	Algoritmos de pontuação.....	66
3.2.1.3.1	Algoritmo de pontuação baseado em campo de força.....	67
3.2.1.3.2	Funções de pontuação empíricas	67
3.2.1.3.3	Funções baseadas em conhecimento	68
3.2.2	<i>Simulação por Dinâmica Molecular.....</i>	68
3.2.3	<i>Similaridade química.....</i>	72

3.2.4	<i>Espectroscopia de Fluorescência e detecção de reação enzimática</i>	73
3.3	METODOLOGIA UTILIZADA NO TRABALHO	75
3.3.1	<i>Seleção de moléculas</i>	75
3.3.1.1	Montagem do banco de dados de fragmentos moleculares.....	75
3.3.1.2	Escolha e preparação da estrutura alvo macromolecular.....	76
3.3.1.3	Preparação da coleção de ligantes para a etapa de docagem molecular.....	77
3.3.1.4	Preparação do substrato natural da enzima, o G3P.....	77
3.3.1.5	Docagem molecular utilizando o programa FlexX.....	78
3.3.1.6	Preparação do sistema para docagem molecular pelo programa Glide e posterior docagem molecular no modo standard precision (SP).....	78
3.3.1.7	Estudo consensual.....	79
3.3.1.7.1	Docagem molecular pelo Glide no modo Extra-Precision.....	79
3.3.1.7.2	Similaridade eletrostática e por cor – ROCS e EON.....	79
3.3.1.7.3	Análise consensual e seleção de moléculas para a compra.....	80
3.3.2	<i>Ensaios bioquímicos por espectroscopia de fluorescência</i>	81
3.3.2.1	Análise dos dados.....	82
3.3.3	<i>Simulação por Dinâmica Molecular</i>	82
3.3.3.1	Simulação por Dinâmica molecular da TcGAPDH na forma apo.....	82
3.3.3.2	Dinâmica molecular das moléculas selecionadas.....	85
3.3.3.3	Cálculo da energia de interação por meio de MMPBSA.....	86
3.3.4	<i>Seleção de compostos análogos às moléculas de boa eficiência do ligante</i>	86
4	RESULTADOS E DISCUSSÃO	88
4.1	SELEÇÃO, ENSAIO ENZIMÁTICO E DINÂMICA MOLECULAR PARA O SÍTIO DO G3P DA TcGAPDH	88
4.1.1	<i>Seleção de compostos in silico</i>	88
4.1.1.1	Montagem do Banco de Dados.....	88
4.1.1.2	Escolha da estrutura cristalográfica de TcGAPDH a ser utilizada para os processos de docagem molecular e simulação por dinâmica molecular.....	90
4.1.1.3	Preparação da estrutura cristalográfica para docagem molecular.....	92
4.1.1.4	Preparação do G3P para os processos de docagem molecular e similaridade química.....	93
4.1.1.5	Docagem molecular do banco de dados.....	93
4.1.1.5.1	Docagem molecular utilizando o programa FlexX.....	93
4.1.1.5.2	Glide-SP.....	98
4.1.1.6	Estudo consensual.....	99
4.1.1.6.1	Abordagem baseada na estrutura do Alvo.....	99
4.1.1.6.2	Similaridade de forma e eletrostática.....	101
4.1.1.6.3	Consensus e seleção de moléculas para compra.....	101
4.1.2	<i>Teste de atividade por Espectroscopia de Fluorescência</i>	103
4.1.3	<i>Dinâmica molecular para confirmar posição de fragmento no sítio</i>	105
4.1.3.1	Controle – Apo.....	106
4.1.3.2	Controle – Holo.....	107
4.1.3.3	Dinâmica molecular das moléculas selecionadas para o sítio do G3P – Primeira geração.....	108
4.1.3.3.1	Molécula NEQUIMED 293.....	109
4.1.3.3.2	Molécula NEQUIMED 292-1.....	110
4.1.3.3.3	Molécula NEQUIMED 295-3.....	111
4.1.3.3.4	Molécula NEQUIMED 294.....	113
4.1.3.3.5	Molécula NEQUIMED 291.....	114
4.1.4	<i>Crescimento da molécula ativa para sítios adjacentes</i>	115
4.1.5	<i>Seleção de moléculas similares à molécula mais ativa – Segunda geração</i>	115
4.1.5.1	Simulação por Dinâmica molecular da segunda geração de compostos.....	116
4.1.5.1.1	Molécula NEQUIMED 358T-1.....	118
4.1.5.1.2	Molécula NEQUIMED 357T-1.....	120
4.1.5.1.3	Molécula NEQUIMED 354T-1.....	122
4.1.5.1.4	Molécula NEQUIMED 359T-1.....	123
4.1.5.1.5	Molécula NEQUIMED 350T,-1.....	125
4.1.5.1.6	Molécula NEQUIMED 360T-1.....	127
4.1.5.1.7	Molécula NEQUIMED 356T.....	128

4.1.5.1.8	Molécula NEQUIMED 361T	130
4.1.5.1.9	Molécula NEQUIMED 355T-1	131
4.1.6	<i>Análise final</i>	132
5	CONCLUSÕES	136
6	PERSPECTIVAS	138
	REFERÊNCIAS.....	140
	APÊNDICE 1 – PARÂMETROS DE FILTRO DAS MOLÉCULAS	151
	APÊNDICE 2 - PARÂMETROS UTILIZADOS PARA SIMULAÇÃO POR DINÂMICA MOLECULAR ..	154
	APÊNDICE 3 – GRÁFICOS DE SIMULAÇÃO POR DINÂMICA MOLECULAR	157

1 Introdução

1.1 Doença de Chagas

O parasito *Trypanosoma cruzi* é o agente etiológico da Doença de Chagas, também conhecida como esquizotripanose ou tripanossomíase americana. Este parasito foi descrito pelo médico sanitário brasileiro Carlos Ribeiro Justiniano das Chagas em 1909, ao chefiar trabalhos de tratamento de malária em Minas Gerais. Ao analisar o sangue dos chamados “barbeiros”, insetos triatomíneos comuns na região, identificou a presença de um tripanossoma até então desconhecido pela comunidade científica. Sabendo que o inseto se alimentava de sangue humano, o médico coletou o sangue de várias pessoas da região que tinham contato direto com o vetor, de modo a verificar se o tripanossoma encontrado infectava o ser humano. A primeira amostra de sangue infectado foi obtida de uma menina de dois anos de idade que havia sido picada pelo vetor poucas semanas antes da coleta. A partir dos relatos da mãe da criança foram obtidos dados de sintomatologia da doença, os quais eram semelhantes aos sintomas apresentados por ratos contaminados com amostras do mesmo parasito obtidos diretamente do inseto vetor.¹

1.1.1 Formas de Transmissão

A principal forma de transmissão da doença de Chagas é por meio do contato com fezes do inseto vetor, o triatomídeo, contaminadas com o protozoário *Trypanosoma cruzi* (*T. cruzi*). Este inseto, durante o repasto sanguíneo, defeca próximo à ferida causada pela picada. O mamífero, ao coçar a ferida, espalha as fezes pela pele e pela ferida, fazendo com que o parasito entre na corrente sistêmica, levando à infecção.

Passados 103 anos da descoberta e descrição da doença, os países Brasil,²⁻³ Chile⁴ e Uruguai⁵ são considerados pela Organização Pan-americana de Saúde livres da transmissão de Doença de Chagas causada pelo inseto triatomídeo, o principal inseto transmissor da doença. A eliminação do inseto vetor é a forma mais eficiente existente de se evitar o contágio da doença por meio do ciclo natural do parasito, já que ainda não se possui qualquer vacina contra essa enfermidade.⁶

Todavia, somente a extinção da transmissão do parasito pelo inseto vetor não é suficiente para evitar a propagação da doença, já que existem modos alternativos de exposição ao parasito e,

consequentemente, infecção pelo mesmo. São conhecidos casos de infecção por transfusão de sangue, transplante de órgãos, infecção congênita pela transmissão do *T. cruzi* da mãe para o filho durante a gravidez (a chamada transmissão vertical), ingestão de alimentos contaminados e exposição acidental em laboratório.

A transfusão de sangue contaminado pode ser considerada o segundo meio de transmissão da doença, ficando atrás somente da transmissão direta. Foram verificados casos de transmissão da doença por transfusão sanguínea no Chile, Argentina e Brasil, na Bolívia e somente sete casos nos Estados Unidos e Canadá nas últimas duas décadas. Nos EUA e Canadá, o parasito foi detectado em cerca de 800 amostras de sangue desde 2007, quando foi implantada a triagem de sangue contaminado com *T. cruzi*.²⁻³

Uma segunda forma de contaminação pelo protozoário entre seres humanos, sem intermediários, é o transplante de órgãos, para o qual já foram relatados casos na América Latina tanto de novas infecções quanto de reativação da doença em pacientes em geral imunossuprimidos que receberam coração, rim e medula óssea, além de outros órgãos.

Outra forma de transmissão entre seres humanos é a transmissão de mãe para filho durante a gravidez, ou também nomeado como transmissão vertical, uma importante forma de contaminação em áreas endêmicas e não endêmicas. A presença do parasito durante a gravidez pode acarretar, entre outros, aborto, nascimento prematuro, morte infantil e sequelas graves no recém-nascido.⁴

No caso da contaminação por ingestão de alimentos, a infecção está relacionada à ingestão de fezes do inseto contaminadas com *T. cruzi* presentes no alimento. Recentemente foram descritos casos da contaminação pela ingestão de caldo de cana,¹⁰ polpa de açaí¹¹ no Brasil e suco de goiaba contaminado na Venezuela, esse último levando à contaminação de 103 pessoas, das cerca de 1000 que se acreditam terem ingerido o suco.¹² No caso da ingestão do caldo de cana de açúcar contaminado, 23 pessoas apresentaram infecção por *T. cruzi*, sendo que três delas morreram.

1.1.2 Ciclo de vida do parasito

Considerando o mecanismo natural de transmissão pelo inseto triatomíneo, o ciclo de vida do parasito no mamífero hospedeiro definitivo (Figura 1.1) inicia-se pela entrada do *Trypanosoma cruzi* na forma tripomastigota metacíclico, presente na urina ou fezes do

“barbeiro”, na ferida aberta na pele do mamífero durante a picada do inseto vetor, iniciando-se a parasitemia. Neste mesmo local, o macrófago internaliza o parasito por endocitose. Dentro do macrófago, ocorre a diferenciação do parasito para a forma amastigota. Ainda no interior do macrófago essa forma se reproduz por divisão binária simples e posteriormente volta a se converter em forma tripomastigota. A célula hospedeira então libera o parasito para o interstício. O *T. cruzi* na forma tripomastigota logo migra para a corrente sanguínea, podendo atingir outras células e ser destruído pelo sistema imunológico, ou ser ingerido pelo inseto vetor ao picar o homem.

Figura 1.1 - Ciclo de vida do *Trypanosoma cruzi*, mostrando as fases de vida, com as modificações morfológicas em diferentes tecidos, tanto em mamíferos quanto em outros triatomídeos

Fonte: da Rocha, J. R. Planejamento de Inibidores das Enzimas Gliceraldeído-3-fosfato Desidrogenase e Diidroorotato Desidrogenase de *Trypanosoma cruzi*. (2010). **Tese (Doutorado)** - Doutorado em Ciências - área de concentração: Físico-Química - Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2010.

Durante a fase inicial da doença (fase aguda), o número de parasitos no sangue na forma tripomastigota é bastante elevado, o que pode causar a morte do mamífero hospedeiro,

principalmente crianças da espécie humana. Nesta fase, juntamente com o aumento da quantidade de parasito no corpo, existe o aumento na produção de imunoglobinas M (IgM), responsáveis pela resposta inicial ao antígeno e também imunoglobulinas séricas (IgG), responsáveis pela resposta imune secundária. Com o passar do tempo, o *T. cruzi* passa a ser encontrado preferencialmente no interior das células, na forma amastigota, ainda com uma grande quantidade de IgG presente, caracterizando a fase crônica da doença (Figura 1.2).

Figura 1.2 - Evolução do número de parasitas e número de anticorpos na corrente sanguínea, em função do tempo decorrido após a picada

Fonte: Coura, J.R., Borges-Pereira, J. Chagas disease: 100 years after its discovery. A systemic review. *Acta Tropica*, v. 115, n. 1-2, p. 5-13, 2010.⁵

Ao picar um hospedeiro, os triatomídeos são infectados ao ingerir a forma tripomastigota do parasito presente no sangue do mamífero. No estômago do invertebrado, o tripanossoma se transforma em forma epimastigota. Esta última se divide por fissão binária no intestino médio, mantendo a infecção no inseto. No reto do triatomídeo, a forma epimastigota se diferencia à forma tripomastigota novamente, sendo eliminada por fezes ou urina, reestabelecendo o ciclo (Figura 1.1).

1.1.3 Sintomatologia

A sintomatologia da doença varia de acordo com a fase em que esta se encontra, aguda ou crônica, sendo que a fase aguda ocorre de forma sintomática ou assintomática, de 4 a 10 dias

depois da picada do triatomíneo, regredindo em um ou dois meses (Figura 1.2). Os sintomas se iniciam através de manifestações locais, quando o parasito penetra na conjuntiva ou na pele, causando o sinal de Romaña e o chagoma de inoculação, respectivamente. O sinal de Romaña se caracteriza por edema bupalperal unilateral, aumento de volume, palpável, de linfonodos pré-auriculares, submandibulares e outros, celulite do tecido gorduroso periorbitário e palperal, com grande quantidade de parasito presente no interior e exterior das células. Já os sintomas cutâneos são caracterizados pelo aparecimento em qualquer parte do corpo de chagoma primário, uma inflamação aguda local na derme e hipoderme, no local de inoculação do parasito, sendo que a lesão lembra, microscopicamente, um furúnculo, que não chega à formação de pus e regride lentamente, seguido de descamação, além de linfadenite-satélite.

Manifestações clínicas mais gerais da fase aguda são representadas pelo aparecimento de febre, edema localizado e generalizado, hepatomegalia (aumento no tamanho do fígado) e esplenomegalia (aumento do tamanho do baço). Casos menos comuns são insuficiência cardíaca e perturbações neurológicas. Essa última é rara, acometendo geralmente crianças muito jovens ou pacientes imunossuprimidos, como consequência de meningoencefalite.

A fase crônica da doença pode ser subdividida na fase crônica assintomática ou indeterminada e na fase crônica sintomática. A primeira delas ocorre logo após o fim da fase aguda e nessa o paciente passa por um longo período (10 a 30 anos) sem qualquer sintoma relacionado à doença. Nessa fase, os exames sorológicos e/ou parasitológicos apresentam resultado positivo para a presença do parasito no organismo, mas com tamanho normal do coração, do esôfago, e do cólon, órgãos geralmente afetados por aumento de tamanho durante a fase crônica sintomática. Além disso, resultado de eletrocardiograma durante essa fase da doença costuma se apresentar normal. Porém, do ponto de vista imunológico, a doença apresenta sinais de atividade, com a presença de anticorpos líticos no sangue do paciente, usados para detectar a presença de parasito vivo no organismo.

A fase crônica sintomática é caracterizada pelo aparecimento de sintomas da doença, principalmente no sistema cardiocirculatório e/ou digestivo, modificando inteiramente a fisionomia anatômica do miocárdio e do tubo digestivo. Esses sintomas estão relacionados com uma reativação intensa de processo inflamatório, afetando tais órgãos.

A cardiopatia na fase crônica da doença não guarda relação direta com a presença e a quantidade de parasitos nesse tecido. Essa inflamação tende a ser predominantemente linfocitária

e distribuída por todo o miocárdio, acompanhada de intensa fibrose. O fato de serem observados linfócitos agredindo fibras cardíacas saudáveis indica existirem outros fatores, que não o parasitário, na manutenção do caso clínico de inflamação. Várias teorias foram construídas para explicar esse fenômeno e uma delas em especial teve impacto negativo por mais de duas décadas no estudo de medicamentos tripanossomicidas e vacinas. Isso ocorreu pois a hipótese da participação central do reconhecimento autoimune na patogênese da cardiopatia, proposta na década de 70, afirmava que os processos autoimunes seriam provocados por antígenos do *T.cruzi* no início da infecção, na fase aguda da doença. Assim, uma influência externa no sistema imunológico poderia agravar o quadro do paciente. Essa teoria foi revista por Kierszenbaum em 2005, rebatendo esta hipótese.⁶

Na forma digestiva da fase crônica da doença, os sintomas são caracterizados pelo megaesôfago, podendo surgir em qualquer idade, principalmente entre os 20 e 40 anos e também pelo megacólon, que em grande parte das vezes acontece posteriormente ao aumento do tamanho do esôfago. A detecção do megacólon é mais difícil, pois compartilha sintomas com outras doenças digestivas. Alguns dos sintomas mais graves são obstrução intestinal e perfuração, podendo levar a um quadro de peritonite. Já o megaesôfago é caracterizado e detectado devido ao aparecimento de sintomas como disfagia, odinofagia, dor retroesternal, regurgitação, pirose, soluço, tosse e sialose.

1.1.4 Panorama atual

Atualmente estima-se que mais de 10 milhões de pessoas estejam contaminadas com a Doença de Chagas no mundo todo, principalmente na América Latina. No ano de 2008, foi estimado que a doença vitimou mais de dez mil pessoas em todo o globo. Na área de risco de contaminação, definida como a região onde existe a presença natural do inseto vetor, vivem atualmente cerca de 25 milhões de pessoas.⁷

Essa doença é classificada pela Organização Mundial de Saúde (OMS) como negligenciada, por estar ligada à pobreza da população mais vulnerável à enfermidade, como indígenas e habitantes de áreas rurais. Porém a doença tem um grande impacto econômico, devido à perda de produtividade comercial da pessoa enferma e aos gastos com saúde. Esse impacto econômico está estimado em cerca de 800 milhões de dólares por país onde a doença é endêmica, em cada ano. Em toda a América Latina, esse número é estimado em 18 bilhões de

dólares perdidos por ano devido aos efeitos da Doença de Chagas. Além da América Latina, onde é endêmica, a Doença de Chagas começa a se espalhar por todo o planeta. (Figura 1.3).

Figura 1.3 - Expansão da doença de Chagas pelo mundo a partir de países endêmicos

Fonte: Coura, J. R.; Vinas, P. A. Chagas disease: a new worldwide challenge. *Nature*, v. 465, n. n7301_suppl, p. S6-S7, 2010.⁸

A expansão dessa doença para países não endêmicos, como os da América do Norte, Europa, Ásia e Oceania, está relacionada à migração de habitantes da América Latina contaminados com o *T. cruzi* para essas regiões. A propagação da doença nos países fora da América Latina acontece somente devido às formas de transmissão entre seres humanos, sem intermediários, seja por meio de transfusão de sangue, doação de órgãos ou transmissão vertical, já que em geral não existe nessas regiões um inseto vetor capaz de completar o ciclo de vida do parasito fora do mamífero. Como não se trata de uma doença muito conhecida fora das regiões endêmicas, medidas preventivas para evitar a propagação do parasito por meio de sangue e órgãos contaminados eram inexistentes até recentemente. Nos Estados Unidos da América (EUA), por exemplo, o controle de sangue contaminado se iniciou no ano de 2007.³

Apesar de os EUA se localizarem fora da região endêmica da doença, o parasito foi encontrado em cães domésticos, mamíferos selvagens e em triatomídeos autóctones, nos estados mais próximos ao México. Esse fato possibilita a transmissão da doença por meio do inseto, aumentando consideravelmente a chance de contágio, quando comparado a uma transmissão somente por meios não-vetoriais. Além disso, foram relatados casos de doadores de sangue

contaminados em 42 dos estados estadunidenses, o que deixa suas autoridades em alerta contra a doença.

1.1.5 Tratamento

Apesar de a Doença de Chagas afetar um grande número de pessoas, principalmente na América Latina e se encontrar em expansão para países de todo o mundo devido à migração, é considerada pela OMS uma doença supernegligenciada, sem qualquer fármaco eficaz e seguro para o tratamento em qualquer fase da moléstia.

O termo doença negligenciada, ou supernegligenciada, está embasado no fato de as doenças assim classificadas não possuírem uma forma eficaz de tratamento e também não atraírem a atenção de empresas farmacêuticas na busca e desenvolvimento de novos fármacos para as enfermidades nas últimas três décadas, estando as ações de pesquisa restritas ao meio acadêmico.⁹⁻¹⁰ A justificativa para o não investimento de verbas por parte das grandes empresas farmacêuticas nessas doenças reside no fato que o mercado comercial para estes fármacos é bastante restrito e voltado para populações de baixa renda, sendo necessário o subsídio governamental para disponibilizar os mesmos aos enfermos. Desta forma, acaba se tornando uma doença com baixo retorno financeiro para as grandes indústrias farmacêuticas.

No caso da Doença de Chagas, existem dois medicamentos para o seu tratamento, o Nifurtimox® e o Benzonidazol® (Fig 1.4), ambos descobertos de forma empírica na década de 70, com severos efeitos colaterais e ineficazes na fase crônica da doença. Por essa razão, o Nifurtimox® foi retirado do mercado no Brasil e sua venda está restrita a países da América Central.¹¹

O Nifurtimox®, ou 5-nitrofurano (3-methyl-4-(5-nitrofurfurylideneamine) tetrahydro-4H-1,4-tiazine-1,1-dioxide), e o Benzonidazol®, N-benzyl-2-nitroimidazole acetamida, possuem elevada taxa de cura na fase aguda, chegando a 80% de cura (número varia de acordo com a localização geográfica da realização do teste, provavelmente devido a diferenças entre as cepas do *T. cruzi* em regiões diversas).¹²⁻¹³ Em alguns estudos nos últimos 15 anos, o Benzonidazol® mostrou uma taxa de cura de até 70% para casos de doentes no início da fase crônica, número que não foi atingido em estudos mais recentes, utilizando tanto Nifurtimox® quanto Benzonidazol®.¹⁴

Um dos pontos fracos destes dois medicamentos é que são incapazes de exibir cura parasitológica em pacientes na fase crônica da doença. A razão para a diferença entre o poder de cura na fase aguda e na fase crônica ainda não é bem entendida, mas suspeita-se que seja devido às suas propriedades farmacocinéticas inadequadas, como tempo de meia vida relativamente curto e baixa permeabilidade a tecidos, impossibilitando a sua chegada ao parasito na forma amastigota, que se encontra no interior da célula.

Adicionalmente à ineficiência no tratamento de pacientes na fase crônica e ao longo tempo de administração necessário (em média 60 dias para o tratamento da fase aguda, a 8-10 mg/kg/dia para o Nifurtimox® e 5 mg/kg/dia para o Benzonidazol®, posologia essa inadequada para administração em crianças),¹⁵ ambos os medicamentos apresentam diversos efeitos colaterais. Os efeitos de uso do Nifurtimox® são anorexia, náusea, vômitos, causando perda acentuada de massa corporal, insônia, irritabilidade e, em menor escala, polineuropatia periférica. Já para o Benzonidazol®, os efeitos colaterais descritos são dermatopatia alérgica, síndromes gastrointestinais e menos frequentemente, depressão na medula óssea, purpura trombocitopênica, polineuropatia e polineurite de nervos periféricos.¹⁴

Além desses efeitos, ambos os medicamentos são genotóxicos, ou seja, têm capacidades de interagir com ADN (ácido desoxirribonucleico), o que pode acarretar problemas genéticos, característica descoberta somente quando foi descrito o mecanismo de ação de ambos os fármacos no organismo, mais de dez anos depois de sua liberação para comercialização.

Ambos os fármacos agem pela formação de radicais livres e/ou formação de metabólitos eletrófilos, com o grupo nitro sendo reduzido à amina pela ação de nitrorredutases. Esses radicais livres formados ligam-se a lipídeos, proteínas e ADN, danificando-os.¹⁶ Especificamente, o mecanismo de ação do Nifurtimox® está relacionado à produção de H_2O_2 e O_2^- a partir do radical nitro. Já no caso do Benzonidazol®, não existe a formação destes íons óxidos. O efeito é exercido pelos metabólitos reduzidos, os quais se ligam às macromoléculas de forma covalente, inativando-as. Apesar de se conhecer esse mecanismo desde a década de 80, detalhes moleculares deste mecanismo de ação foram elucidados recentemente. Esses fármacos são ativados pela família das Flavoproteínas nitrorredutases do tipo I, como a *Old Yellow Enzyme* (TcOYE). Essa ativação produz os intermediários reativos e os metabólitos reduzidos dos fármacos, que são responsáveis pelos efeitos adversos.¹⁷

Dadas todas as limitações do atual tratamento para a Doença de Chagas, é necessária a busca de novos medicamentos. Novos fármacos, para serem considerados promissores, precisam ser ativos contra as formas infectivas em células de mamíferos, tripomastigota e amastigota, principalmente na segunda, que está presente na forma crônica da doença, com desejável estudo de atividade em modelos animais. É desejável também que o novo fármaco tenha atividade contra diferentes cepas de *T. cruzi*, de modo que seja útil para diferentes regiões afetadas pela doença e seja pouco susceptível à resistência do parasita. Baixa toxidez e baixo risco de interação com citocromos P450 são essenciais, essa última para evitar interação medicamentosa, principalmente pelo fato de vários pacientes também ingerirem outros fármacos para controlar os sintomas da doença. Devido ao fato de a doença atingir principalmente pessoas pobres e que geralmente moram em lugares afastados, é imprescindível que o medicamento tenha grande biodisponibilidade quando administrado por via oral. Sua posologia deve ser também adequada para administração com pouco ou nenhum efeito colateral em qualquer faixa etária.¹⁸

A busca por fármacos contra a Doença de Chagas ganhou força na última década, devido a esforços conjuntos da OMS, DNDi (Drug for Neglected Disease Initiative), Wellcome, iOWH e NIAID¹⁹ no desenvolvimento de medicamentos para doenças negligenciadas e a confirmação da expansão da doença para países de elevada renda (EUA e países da Europa), atraindo um maior interesse das indústrias farmacêuticas.

Um dos alvos mais estudados para o tratamento da doença é a via de biossíntese de ergosterol, pois após esforços nas últimas duas décadas, foi demonstrado que o *T. cruzi* é dependente da mesma, assim como os fungos. Experimentos iniciais mostraram que os inibidores da biossíntese de ergosterol comerciais para fungos (Itraconazol, Cetoconazol, Terbinafina) atuam como supressores do crescimento do parasito, mas não na sua cura. Desenvolvimento de novos triazóis baseados naqueles já conhecidos levou à descoberta do Posaconazol, molécula estruturalmente análoga ao Itraconazol, que leva à cura de modelos de murina tanto para fase aguda quanto para a fase crônica. Essa molécula foi também capaz de matar cepas resistentes do *T. cruzi*, erradicar o parasito na forma amastigota intracelular em cultura de cardiomiócitos e obter melhores resultados que o Benzonidazol® na prevenção de danos às células cardíacas. Além do Posaconazol, foi desenvolvido também o Ravuconazol, para o qual foi desenvolvido um pro-fármaco solúvel em água, E1224 (derivado de mono-lisina), que

também possui grande poder de cura de animais infectados com o parasito. Tanto o E1224 quanto o Posaconazol encontram-se em fase clínica atualmente. (

Figura 1.4-).¹⁴

Uma terceira molécula, que se encontra em fase clínica I para o tratamento da Doença de Chagas foi desenvolvida a partir de compostos que tinham como alvo cisteíno-proteases para o tratamento de osteoporose e psoríase, por meio de sua inibição. Os resultados não foram o esperado para esse fim, mas serviram como base para o desenvolvimento de inibidores da cruzaina, cisteíno-protease presente em *Trypanosoma cruzi* e um alvo validado para o tratamento da Doença de Chagas. Depois de uma década de estudos, a molécula K-11777 se mostrou eficaz na cura de células ou animais infectados com o parasito em doses não tóxicas,²⁰ e então foi aprovada para entrar em fases clínicas de teste.²¹

Figura 1.4-Estrutura química das moléculas já em fase terapêutica e moléculas em fase clínica para a Doença de Chagas

Apesar de existirem três moléculas em fase clínica de testes para o tratamento contra a doença de Chagas, a busca de outras moléculas para o tratamento e cura da doença ainda se faz necessária, no caso dessas moléculas falharem em se tornarem fármacos ou tiverem baixa eficiência. Algumas enzimas têm se mostrado promissoras para serem utilizadas como alvo para busca de moléculas contra essa enfermidade. Por exemplo, pode-se listar a cruzaina, diidroorotato desidrogenase, a via de biossíntese de ergosterol, enzimas da via glicolítica, como por exemplo, a GAPDH, entre outras.

1.2 Química medicinal e o desenvolvimento de fármacos

O desenvolvimento de um novo fármaco é um processo bastante caro e demorado, sendo estimado em 2003 o gasto de mais de 1 bilhão de dólares em até 25 anos de desenvolvimento²². Mesmo diante dessas cifras e do tempo para desenvolvimento de um novo agente terapêutico, não é assegurado que ao final de todo o processo se terá um novo medicamento no mercado.

1.2.1 O fármaco no organismo – fases de atividade

O fármaco, desde a sua administração até a ação farmacológica no corpo, passa por três fases: farmacêutica, farmacocinética e farmacodinâmica. Essas fases devem ser bem entendidas de modo a minimizar o risco de efeitos colaterais e maximizar a sua interação com o alvo biológico, e assim, a resposta farmacológica desejada.

1.2.1.1 Fase Farmacêutica

A fase farmacêutica está relacionada com o modo de administração do medicamento no organismo. Os métodos de administração hoje existente são divididos em via local, via enteral, que consiste em administração oral, bucal/sublingual e retal, via parenteral, por meio de injeção intramuscular, subcutânea ou intravenosa, via intranasal e inalação, com cada um deles tendo suas vantagens e desvantagens. O modo de administração mais desejado no meio farmacêutico é por via oral, por ser a rota mais fácil e segura. Em geral, os medicamentos são administrados por outras vias somente em casos em que o paciente não apresente condições de engolir, que seja necessária uma absorção imediata ou que o fármaco tenha problemas em ser absorvido pelo trato gastrointestinal de forma eficaz. Para cada um dos modos de administração e características do fármaco, a formulação de administração deve ser estudada de modo a otimizar a sua absorção pelo organismo, podendo-se optar por administração oral por meio de cápsulas, comprimidos ou na forma líquida.²³

1.2.1.2 Fase Farmacocinética

A fase farmacocinética corresponde ao trajeto do fármaco desde a sua administração até a sua chegada ao alvo macromolecular. Nessa fase, é de grande importância que o fármaco tenha

características de Absorção, Distribuição, Metabolismo e Excreção, além de baixa toxidez, ou em outras palavras, um bom perfil ADME/Tox. É esta fase a principal responsável pelo alto índice de atrito observado para os fármacos quando estão em fase clínica. A Absorção é o fenômeno da passagem do fármaco do seu meio de administração para a corrente sistêmica, o que é resumido na Figura 1.5²³.

Figura 1.5 - O caminho do fármaco no organismo, desde sua administração até a eliminação

Fonte: Adaptado de: Wermuth, C. G. *The Practice of Medicinal Chemistry*, 3 ed. Londres: Academic Press, 2008

Para a administração oral, a absorção acontecerá principalmente no estômago, intestino grosso e intestino delgado. Neste último, principalmente no duodeno, acontece grande parte da absorção de nutrientes pelo organismo, e, também, dos fármacos. Os principais fatores que guiam a absorção do fármaco são a formulação de administração, mobilidade pelo trato gastrointestinal e perfusão do trato gastrointestinal. Uma vez absorvido, o fármaco deve ser distribuído pelo organismo. A distribuição para o local de ação pode ser limitada tanto pelo fluxo sanguíneo para o tecido/órgão de destino (distribuição limitada pela perfusão), pela incapacidade do fármaco em atravessar a barreira sangue-tecido devido às propriedades físico-químicas do composto

(distribuição limitada pela permeabilidade), pela alta afinidade do fármaco com certos tecidos, o que diminui a sua disponibilidade no sangue, causando acumulação do fármaco, ou ao contrário, a sua alta concentração no sangue, devido à ligação com proteínas plasmáticas, diminuindo também a sua biodisponibilidade.

Durante todo esse processo, o fármaco está sujeito ao metabolismo e à excreção do organismo. O metabolismo, também denominado de biotransformação, acontece principalmente no fígado, mas pode também acontecer na mucosa intestinal, rins, pulmão, pele, placenta e plasma. No caso da biotransformação hepática, que acontece nos hepatócitos, em geral o fármaco é metabolizado a moléculas mais solúveis. Esse metabolismo, quando consiste na adição de grupos funcionais reativos por meio de oxidação, redução e hidrólise é descrito como de Fase I, sendo as reações realizadas em grande parte pelas enzimas da família Citocromo P450. O metabolismo por meio de conjugação de grupos polares como ácidos glicurônicos, sulfonatos, glutathione e aminoácidos, é classificado como sendo de Fase II. Nessa fase, as reações são catalisadas por metiltransferases, sulfotransferases, N-acetiltransferases, UDP-glucoroniltransferases e glutathione-S-transferases.²⁴ Essas duas fases de metabolismo tornam a molécula mais solúvel e possibilitam a sua excreção pelos rins ou pela vesícula biliar. A biotransformação pela Fase II, em geral, inativa a molécula, enquanto a biotransformação pela fase I pode levar à forma ativa do fármaco, o denominado pró-fármaco. A transformação do fármaco em seu metabólito, às vezes reativo, pode levar ao surgimento de severos efeitos colaterais e também causar toxidez.

Na etapa de excreção, ou seja, eliminação do fármaco e seus metabólitos do corpo, os rins exercem um papel fundamental. São eles os responsáveis pela limpeza de produtos metabólitos e toxinas do sangue, para posterior eliminação por meio da urina. Além dos rins, a vesícula biliar é outro órgão que atua diretamente na excreção de fármacos. Para que seja excretado pela bile, o composto, com massa molecular entre 400 e 500 Da, deve conter um forte grupo polar. Grande fração dos fármacos assim excretados possui um grupo glucurônico conjugado, fruto do metabolismo de Fase II. A bile encaminha o líquido biliar para o intestino, para que seja eliminado por meio das fezes.²³

1.2.1.3 Fase Farmacodinâmica

Por fim, a fase farmacodinâmica corresponde à interação específica do fármaco com o

alvo macromolecular de interesse. Essa interação acontece por meio de complementaridade estéreo-eletrônica do composto com o alvo, de modo a modular a sua atividade biológica. Desta forma, a fase farmacodinâmica é de extrema importância para que o fármaco administrado exerça a sua ação farmacológica. Assim, grandes esforços e verbas estão despendidos no planejamento e otimização do composto para se tornar fármaco principalmente no que diz respeito à potência e afinidade do mesmo com o alvo macromolecular escolhido, mas sem se descuidar das propriedades farmacocinéticas. É importante também assegurar, ou ao menos maximizar, a sua seletividade para com este alvo, de forma a minimizar as reações adversas indesejadas.²⁵

1.2.2 A Química Medicinal no planejamento de fármacos

Com vários fatores diretamente envolvidos, o lançamento de um novo fármaco no mercado é um campo de pesquisa de alto risco e demanda profissionais com conhecimento em diversas áreas, desde a fisiopatologia da doença, seleção de alvo terapêutico até os testes *in vitro* e *in vivo* e compreensão de relação estrutura-propriedade e estrutura-atividade, além de propriedades ADME/Tox, passando pela seleção, otimização e síntese de moléculas. O desenvolvimento de fármacos envolve a integração de várias frentes da ciência, a qual está reunida na sub-área da química denominada Química Medicinal.²⁶

Segundo a União Internacional de Química Pura e Aplicada (IUPAC, do inglês):

A Química Medicinal envolve a descoberta, o desenvolvimento e a identificação e interpretação do modo de ação de compostos biologicamente ativos ao nível molecular. A ênfase é colocada em fármacos, mas o interesse da Química Medicinal também está voltado para o estudo, identificação e síntese de seus produtos metabólicos bem como dos compostos relacionados."

Esta é uma disciplina baseada na química que engloba inovação, descoberta e desenvolvimento de novas substâncias químicas bioativas (NCE), por meio de métodos baseados na estrutura do ligante ou do receptor, síntese ou modificação molecular, descrição das moléculas desde a sua constituição atômica (passando por relações estrutura-propriedade, SPR) até suas características estruturais quando da(s) interação(ões) com alvos biológicos de interesse terapêutico;²⁷ compreensão, em nível molecular, de processos bioquímico-farmacológicos, toxicológicos e farmacocinéticos (ADME/Tox) e a geração de relações estrutura química-atividade farmacológica (SAR/QSAR). Ela também está implicitamente relacionada com a proposição e validação de modelos matemáticos através dos estudos de relações entre a estrutura

química e a atividade farmacológica e/ou toxicológica e/ou farmacocinética. A Química Medicinal é, portanto, híbrida e opera conjuntamente com outras especialidades como biofísica, biologia molecular, bioquímica, clínica médica, físico-química, fisiologia, patologia, química computacional, química biológica, química farmacêutica, química orgânica e química quântica, entre outras.²⁸

1.2.2.1 Descoberta e Desenvolvimento de fármacos

O desenvolvimento de novos fármacos pode ser dividido em duas grandes fases: descoberta e desenvolvimento.

1.2.2.1.1 A descoberta

A primeira etapa da gênese de fármacos, a descoberta, visa encontrar uma molécula que possua características básicas de um fármaco, tanto na potência/afinidade com o alvo quanto nas características ADME/Tox e nos parâmetros físico-químicos. Esta molécula é denominada composto-líder (*lead*, do inglês) e está pronta para avançar na gênese de fármacos. Um panorama geral desta etapa é apresentado na Figura 1.6²⁹.

Figura 1.6 - Evolução da molécula durante a gênese de fármacos

Fonte : Adaptado de Bleicher, K.H., Böhm, H.-J., Müller, K., Alanine, A.I. Hit and lead generation: Beyond high-throughput screening. *Nature Rev. Drug Disc.*, v. 2, n. 5, p. 369-378, 2003

O primeiro procedimento para identificar moléculas que venham a se tornar um candidato a fármaco para o tratamento de uma determinada patologia é a escolha do alvo biológico adequado. Este alvo deve ser passível de modulação pelo futuro fármaco e essa modificação na função deve acarretar mudanças fisiológicas que caracterizem a cura da doença. É desejável que o alvo escolhido já seja validado, ou seja, tenha sido demonstrada a importância direta do mesmo na manifestação da doença.³⁰ Ainda que o alvo não esteja rigorosamente validado, o estudo de moléculas que modulem a sua ação pode servir como um futuro meio de validação do alvo macromolecular. A Figura 1.7 mostra em quais classes de alvos biológicos os fármacos aprovados para comercialização pelo órgão de regulamentação de fármacos estadunidense, FDA (sigla do inglês Food and Drug Administration)³¹:

Figura 1.7 - Distribuição dos alvos biomacromoleculares para os fármacos atualmente na terapêutica

Fonte: Adaptado de Hopkins, A.L., Groom, C.R.. The druggable genome. **Nature Rev. Drug Disc.**, v. 1, n. 9, p. 727-730, 2002.

Escolhido o alvo para estudo, parte-se então para a identificação de ligantes. Essa identificação pode ser feita de inúmeras maneiras, como o ensaio enzimático *in vitro* de moléculas em larga escala (HTS, do inglês *High Throughput Screening*), teste de bibliotecas de moléculas com característica de fragmentos (*Fragment-like*), ensaios *in silico* e ensaios por cristalografia de raios-X, ressonância magnética nuclear e/ou espectrometria de massas. Essa primeira etapa é definida como ensaio primário³² e o ideal é que faça uso de técnicas complementares para confirmação do resultado.

As moléculas que apresentem, nesses ensaios, atividade biológica contra o alvo

biomacromolecular são consideradas como ligantes. Testes de validação são feitos então para confirmar que se trata de fato de um ligante e não de um falso positivo, por conta das condições do ensaio primário.

A molécula confirmada como ligante, passa para a etapa de otimização a composto-líder. Nesse estágio, o ganho na atividade contra o alvo fica em segundo plano, quando se trata de compostos provenientes de HTS. Essas modificações são feitas de modo a chegar a uma molécula com características básicas de um fármaco, com dados de atividade *in vivo* e propriedades ADME/Tox adequadas. As características que em geral são levadas em conta durante o processo de otimização são a seletividade para o alvo macromolecular em estudo, mecanismo de ligação, propriedades farmacocinéticas, tratabilidade química e liberdade de patenteamento. Já quando o ligante é fragmento-similar, os parâmetros em geral avaliados para escolher a molécula que será otimizada, e para sua posterior otimização, são o conhecimento do modo de interação do ligante com o alvo e a facilidade química em se produzir análogos.

Algumas técnicas são utilizadas para se construir, a partir do ligante, um composto-líder como, por exemplo, substituição bioisostérica, fragmentação do ligante e evolução do ligante. As moléculas similares a fragmentos são otimizadas a ligantes maiores por métodos que serão abordados posteriormente, na seção 1.5.5.

Ao se encontrar um composto líder, é feito seu refinamento molecular para melhoria nas propriedades farmacocinéticas e farmacodinâmicas, diminuição na possível toxidez e otimização de parâmetros químicos em modelos animais, *in vivo* e *in vitro*, com descrição de propriedades toxicológicas, tempo de meia vida e dosagem de administração. Estudo reprodutível de toxicidade em animais machos e fêmeas, além de estudo de longo prazo de carcinogênese é necessário para o composto ser considerado candidato a fármaco.³³

1.2.2.1.2 O desenvolvimento

Concluída a etapa de busca pelo candidato a fármaco, todos os dados obtidos até o momento são enviados à agência reguladora do país onde se quer registrar o composto como fármaco (ANVISA no Brasil e FDA nos Estados Unidos) na forma de um documento, o Comunicado Especial, no caso da ANVISA ou equivalentemente para o FDA, a solicitação de registro para Investigação de Novo Fármaco (IND, do inglês Investigational New Drug).

Com a aprovação do órgão regulamentador, ensaios clínicos em humanos são realizados

com o composto candidato a fármaco. A Fase I consiste na administração do composto em voluntários sadios, de modo a avaliar a segurança, a tolerância, os parâmetros farmacodinâmicos e efeitos farmacocinéticos do composto no organismo humano. Nesta fase, o gasto fica por volta de 500 mil dólares, podendo chegar até a um milhão e meio de dólares, caso sejam necessários estudos especiais.³³

Na Fase II, o candidato a fármaco é administrado em pessoas que possuem a moléstia para o qual o fármaco foi desenvolvido, na intenção de avaliar a sua eficácia e segurança. Essa fase pode ser dividida em duas etapas. A primeira etapa é a prova de conceito, onde o candidato é em geral administrado em dose única na dosagem máxima tolerada em um pequeno número de pacientes (12 -100), enquanto a segunda consiste na administração em diferentes dosagens na população alvo, de modo a determinar a dosagem mínima efetiva, a dosagem ineficaz e também a dosagem ótima para o composto. Esta fase pode custar vários milhões de dólares e envolve centenas de pacientes. Mesmo após a aprovação nesta fase, às vezes as indústrias farmacêuticas desistem de continuar o processo devido a resultados inferiores ao esperado inicialmente ou à ausência de mercado que justifique o investimento dos elevados custos da Fase III.³³

A Fase III, a última antes de o candidato a fármaco se tornar um fármaco, pode envolver milhares de pacientes, para criar um banco de dados adequado para assegurar a eficácia e o perfil de segurança do composto. Esta fase é realizada em diversos países de modo a verificar a eficácia e segurança para diferentes populações. Nesta fase, são confirmadas as doses clínicas, frequências e período de administração. Após o fim dessa fase, é feito um pedido de Registro de Novo Fármaco na Anvisa ou Registro de Aplicação como novo Fármaco, NDA (do inglês New Drug Application), no FDA, o qual pode ficar em revisão por até 15 meses para aprovação.

Quando o fármaco é lançado no mercado, seu uso continua sendo monitorado pelas agências regulamentadoras, com a indústria farmacêutica tendo de oferecer periodicamente relatórios sobre a segurança do medicamento, no que é descrito como período de farmacovigilância na Europa, ou Fase IV. Caso sejam descobertos efeitos colaterais graves nessa fase, o fármaco pode ser retirado do mercado, como foi o caso do rofecoxib. Todo esse processo é representado na Figura 1.8²⁶.

Figura 1.8 - Fases pré-clínica e clínica necessária para a aprovação de uma molécula como fármaco

Fonte: Adaptado de Lombardino, J.G., Lowe III, J.A.. The role of the medicinal chemist in drug discovery - Then and now. *Nature Rev. Drug Disc.*, v. 3, n. 10, p. 853-862, 2004.

Os fármacos órfãos, aqueles voltados para o tratamento de doenças que afetem menos de 200 mil pessoas no país, segundo a definição do FDA, possuem uma série de incentivos econômicos e regulatórios, como uma maior rapidez no julgamento dos resultados e período de exclusividade de comercialização por sete anos. Na Europa, essa classe inclui doenças negligenciadas, como é o caso da Doença de Chagas.

1.3 O alvo biomacromolecular

1.3.1 A via glicolítica

A via glicolítica em tripanossomatídeos é sugerida como uma interessante via a ser modulada por compostos químicos para o tratamento de doenças negligenciadas, incluindo a doença de Chagas.³⁴⁻³⁵ Este interesse reside no fato de que essa via é essencial na produção de ATP para tripanossomatídeos, sendo que nesses organismos esta via está organizada de forma bastante peculiar, compartimentalizada no interior de glicossomos. Além disso, várias das enzimas dessa via possuem características estruturais e cinéticas diferentes das encontradas nos mamíferos.

Esta via, como demonstrado em extensivos estudos reportados na literatura, é a única

fonte de geração de energia para a forma sanguínea do *Trypanosoma brucei* (*T. brucei*), causador da tripanossomíase africana, também conhecida como a doença do Sono, já que esse parasito não possui cadeia respiratória ou ciclo de Krebs associado à glicólise.³⁶ Para *Trypanosoma cruzi* e *Leishmania spp*, apesar do papel da via glicolítica não ser tão pronunciado, também foi evidenciado que estes utilizam a glicose como a principal fonte de energia.

Na forma amastigota, forma do parasito encontrada no interior das células do hospedeiro, acredita-se que a glicólise tenha um papel essencial na geração de energia,³⁷ apesar de Hellemond ter mostrado em 2009 que processos mitocondriais também têm uma grande contribuição, a partir de aminoácidos abundantes no interior do organismo, como prolina e glutamina.³⁸⁻³⁹ O trabalho de Silber e colaboradores em 2009 demonstrou que, na fase amastigota, a concentração de prolina dentro do parasita é consideravelmente maior do que em qualquer outra forma.⁴⁰

A via glicolítica é constituída de dez enzimas, as quais são responsáveis por converter glicose em piruvato, com a produção de duas moléculas de ATP e duas moléculas de NADH, como mostrado na Figura 1.9.

Figura 1.9 - Via glicolítica e sua distribuição em insetos Triatomíneos

Na figura, nota-se que sete das enzimas que participam da glicólise estão confinadas no interior do glicossomo, ao contrário do que acontece em outros organismos, onde todas as enzimas estão presentes no citoplasma celular.

Diante de todas as enzimas da via glicolítica, qual seria a ideal de se buscar compostos inibidores de sua atividade para que o parasito morra? Na busca pela resposta a esta pergunta, Bakker e colaboradores, em 1999, fizeram um estudo teórico utilizando dados experimentais de cinética para obter qual enzima controla a glicólise na fase sanguínea do *T. brucei*, sob condições fisiológicas, dados resumidos na **Erro! Fonte de referência não encontrada.**⁴¹

Tabela 1.1 - Inibição enzimática necessária para levar à diminuição de 50% na produção de glicose em *Trypanosoma brucei*. ALD, Frutose-1,6-difosfato aldolase; GDH, Glicerol-3-fosfato desidrogenase; GAPDH, Gliceraldeído-3-fosfato desidrogenase; PGK, 3-fosfoglicerato cinase; HK, Hexocinase; PFK, 6-fosfofruto-cinase; PYK, Piruvato cinase

Enzima	Inibição necessária
	Redução do fluxo em 50%
	%
Transporte de glicose	51
HK	93
PFK	93
ALD	76
GAPDH	84
PGK	85
PYK	97
GDH	83

Fonte: Adaptado de Bakker, B.M., Michels, P.A.M., Oppendoes, F.R., Westerhoff, H.V.. What controls glycolysis in bloodstream form *Trypanosoma brucei*? **J. Biol. Chem.**, v. 274, n. 21, p. 14551-14559, 1999

Este estudo mostra que o transporte de glicose para o glicossomo seria a etapa mais recomendada de se inibir para causar a morte do parasito, seguido por inibição da ALD, GDH, GAPDH e PGK, seguindo a ordem das enzimas que exercem maior controle na via glicolítica, necessitando de uma inibição de 76, 83, 84 e 85 % das mesmas, respectivamente, para reduzir o fluxo glicolítico em 50 %.

1.3.2 A enzima Gliceraldeído-3-Fosfato desidrogenase

A Gliceraldeído-6-fosfato desidrogenase (GAPDH) tem sido estudada no meio científico como um promissor alvo contra a doença de Chagas.⁴² Frutos desse esforço, existem hoje cinco estruturas depositadas no Banco de Dados de Proteínas (PDB, do inglês Protein Data Bank),⁴³ na forma holoenzima e também com inibidores, reversíveis e irreversíveis, co-cristalizados.

1.3.2.1 Informações estruturais

A GAPDH de *T. cruzi* é um homotetrâmero, com cada unidade monomérica contendo 359 aminoácidos e dois domínios. O domínio N-terminal, por meio de duas unidades do enovelamento de Rossmann, acomoda o cofator NAD⁺ enquanto o domínio C-terminal acomoda o substrato gliceraldeído-3-fosfato (G3P) e o fosfato inorgânico (P_i). A cavidade entre os dois domínios representa o sítio catalítico da enzima, contendo a cisteína catalítica 166 (Cys166) Figura 1.10. Quando comparada com a enzima humana, apresenta uma identidade de até 48% da sequência de aminoácidos com a GAPDH de espermatozóides, com diferenças no sítio ativo, principalmente na cavidade de interação do grupo adenosina do cofator. A identidade de aminoácidos entre a enzima de humanos e de *T. cruzi* varia devido ao fato de existir pelo menos três isoformas da GAPDH em humanos, presentes em diferentes tecidos.

Figura 1.10 - Representações da enzima GAPDH, código PDB 1QXS. A) Cada cor representando um monômero da TcGAPDH. B) Em azul, domínio de Rossmann, em verde, domínio C-terminal e em vermelho, sítio catalítico.

1.3.2.2 Reação catalisada pela GAPDH

A GAPDH é a sexta enzima da via glicolítica, com número de classificação enzimática (E.C. number, do inglês Enzyme classification number) 1.2.1.12 ou seja, é uma oxidoredutase (E.C. 1) que atua em aldeídos ou grupos oxo da molécula doadora (E.C. 1.2) e tem como molécula aceitadora o NAD⁺ ou NADP⁺ (E.C. 1.2.1). Essa enzima é responsável por catalisar a

transformação de G3P a 1,3-bisfosfoglicerato (1,3 BPG), utilizando um fosfato inorgânico e tendo como cofator o NAD^+ , o qual é reduzido a NADH. (Figura 1.11)

Figura 1.11 - Esquema da reação catalisada pela enzima GAPDH

Para que a reação de formação de G3P a 1,3-BPG aconteça, inicialmente a His194 abstrai o próton da Cys166, levando à formação da cisteína ativada. Esta então realiza um ataque nucleofílico ao grupo carbonílico do G3P, formando um intermediário tetraédrico hemitioacetal. Adjacente a esse intermediário, o NAD^+ abstrai o hidreto do que era o G3P, formando NADH e saindo do sítio. Ao mesmo tempo, o hemitioacetal é oxidado a tioéster, utilizando uma molécula de água do meio. Uma nova molécula de NAD^+ substitui o NADH e essa carga positiva atua como estabilizadora da carga negativa do oxigênio carbonílico do tioéster do intermediário. Por fim, o fosfato inorgânico ataca o tioéster, formando um novo intermediário tetraédrico, que rapidamente se colapsa, liberando 1,3-BPG e o grupo tiol da CYS166 catalítica. (Figura 1.12)

Figura 1.12 - Esquematisação do mecanismo catalítico de reação da enzima GAPDH

Fonte: da Rocha, J. R. Planejamento de Inibidores das Enzimas Gliceraldeído-3-fosfato Desidrogenase e Diidroorotato Desidrogenase de *Trypanosoma cruzi*. (2010). **Tese (Doutorado)** - Doutorado em Ciências - área de concentração: Físico-Química - Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2010

1.3.2.3 Inibidores conhecidos

Por se tratar de uma enzima bastante estudada, várias moléculas com característica de ligantes já foram descritas. As moléculas que inibem reversivelmente são provenientes de produtos naturais, análogos de substrato, produto e/ou cofator e de base de dados comerciais.

As classes de produtos naturais que mostraram atividade de inibição da GAPDH são os flavonoides, com IC₅₀ entre 46 e 86 µM,⁴⁴ as coumarinas, com 13 moléculas testadas e IC₅₀ entre 64 e 617 µM,⁴⁵ dentre elas a chalepina, co-cristalizada com a TcGAPDH (estrutura de código PDB 1K3T), além de derivados de ácido anacárdico, com três moléculas ensaiadas e IC₅₀ entre 38 e 161 µM.⁴⁶ e as biantronas, com constante de afinidade, K_i, igual a 27,3 µM.⁴⁷

Em relação aos inibidores análogos aos substratos ou produtos, Ladame e colaboradores testaram uma série de oito compostos análogos ao produto 1,3-BPG, encontrando ligantes com potência entre 700 e 2000 µM.⁴⁸ Estes inibidores, apesar de serem bastante semelhantes ao produto da reação catalisada pela enzima, mostraram-se bastante fracos, principalmente quando comparados aos resultados já obtidos por produtos naturais.

Por outro lado, inibidores similares ao cofator NAD⁺ foram identificados como bastante

potentes e seletivos para a TcGAPDH. A molécula com maior atividade contra a enzima apresentou um IC_{50} de 0,1 μM para a GAPDH de *T. cruzi*, sendo inativa contra a GAPDH humana a uma concentração de até 500 μM .⁴⁹ Porém, essa molécula possui uma massa molecular maior do que 800 Da e propriedades farmacocinéticas inadequadas para o seu prosseguimento na gênese de fármacos. Em nosso grupo, NEQUIMED, recentemente, a Ms. Fabyana Soares mostrou que um grupo de seis moléculas similares ao anel nicotinamida do NAD⁺ possui K_i em valores de baixo micromolar em relação a este cofator contra a enzima do parasito, sendo que uma delas apresentou seletividade em relação à enzima humana.⁵⁰

Adicionalmente, estão descritos na literatura inibidores análogos simultaneamente ao substrato G3P e ao cofator NAD⁺. Pesquisadores do grupo de Wilson encontraram em 2005 inibidores com IC_{50} entre 4 e 2000 μM , com K_i em relação ao substrato entre 1050 e 4 μM , sendo todos definidos como inibidores competitivos. Além disso, o K_i medido para essas moléculas em relação ao cofator varia entre 950 e 10 μM , tendo seus modos de ação descritos como não-competitivos ou mistos.⁵¹

O Dr. Renato Ferreira de Freitas, durante seu doutorado no NEQUIMED, encontrou inibidores para a TcGAPDH realizando uma busca por bancos de dados de moléculas *in silico*. Essas moléculas apresentaram K_i na faixa de baixo micromolar para o sítio do G3P, sendo ativas contra o parasito. É o primeiro caso relatado na literatura de inibidores não covalentes da GAPDH encontrados a partir de uma fonte sintética e baseado prioritariamente na estrutura do alvo, encontrando esqueletos inéditos inibidores desta enzima.⁵²

Apesar de todos os inibidores já descritos na literatura, os mesmos ainda não são compostos-líderes, devido, principalmente, à falta de atividade em testes contra o parasito e problemas quanto às características químicas e propriedades farmacocinéticas, além de dificuldade de se obter em grande escala, no caso de produtos naturais.

1.4 O espaço químico e ferramentas para investigá-lo

No ano de 2010, foram lançados no mercado 16 novos fármacos, número que vem caindo nos últimos anos, apesar do aumento na quantia investida pela indústria farmacêutica na pesquisa de novos medicamentos.⁵³

1.4.1 O espaço químico

Um dos motivos para o descompasso do aumento entre a quantia investida e a diminuição ou estagnação de novos fármacos lançados no mercado é a dificuldade intrínseca em explorar a diversidade existente no espaço químico, que, por definição, é o espaço total de descritores que abrange todas as pequenas moléculas constituídas de carbono que podem, em princípio, ser criadas.⁵⁴ Em uma conta rápida, seria possível sintetizar 10^{29} moléculas diferentes derivados de *n*-hexano a partir de uma biblioteca de 150 substituintes, enumerando todas as possibilidades desde hexano substituído em uma posição até substituído em quatorze posições. Estima-se que possam existir mais 10^{60} moléculas contendo em sua constituição C, N, O, S e um total de até 30 átomos,⁵⁵ o que já supera o número de partículas elementares no universo cosmológico.⁵⁶

1.4.2 O espaço químico biologicamente relevante

O verdadeiro desafio do químico medicinal é mapear, dentro de todo o espaço químico, regiões onde seja mais provável se encontrar moléculas com atividade biológica adequada. Essa região é denominada de espaço químico biologicamente relevante (Figura 1.13).⁵⁷

Figura 1.13 - Representação esquemática do espaço químico (cubo) e do espaço químico biologicamente relevante (elipsóides)

Fonte: Adaptado de Lipinski, C., Hopkins, A.. Navigating chemical space for biology and medicine. **Nature**, v. 432, n. 7019, p. 855-861, 2004.

Desde o advento da química combinatória, em 1992, o número de substâncias químicas conhecidas vem aumentando exponencialmente, sendo que este número está próximo de 65 milhões de moléculas atualmente, segundo o Chemical Abstract Society (CAS).⁵⁸ As indústrias farmacêuticas, não raro, fazem ensaio enzimático em larga escala (HTS) de cerca de um milhão de compostos, muitas vezes escolhidos de forma aleatória contra determinados alvos biológicos. Considerando o gasto de 40 centavos de dólar por molécula testada, já incluso o preço da síntese e realizando o ensaio contra 25 diferentes alvos, a quantia investida é de 10 milhões de dólares.⁵⁷ Mesmo assim, a varredura do espaço químico com essas moléculas é bem baixa, principalmente pelo fato de que parte desses bancos é composto por moléculas que foram desenvolvidas para antigos alvos de interesse da mesma. Por isso, várias das grandes farmacêuticas estão investindo dinheiro na diversificação molecular de seu banco de dados *in house*.⁵⁹

1.4.3 A quiminformática

Neste panorama, encaixa-se a quiminformática, cujo principal papel é encontrar uma forma racional de descrever todo o espaço químico e, assim, navegar por ele. O desenvolvimento de ferramentas capazes de varrer todo o espaço químico de maneira lógica é de grande importância para o desenvolvimento de moléculas biologicamente ativas e fármacos.

A quiminformática, que antes era definida como “a aplicação de métodos de informática para resolução de problemas químicos”, por Gasteiger (2003) ou como “um termo genérico que compreende o planejamento, criação, organização, gerenciamento, recuperação, análise, disseminação, visualização e uso de informações químicas” por Brown (2006), agora recebe uma definição mais objetiva de Varnek (2011, p. 23), como “um campo baseado na representação de moléculas como objetos (gráficos ou vetores) no espaço químico”.⁵⁶

Essa disciplina trata as moléculas como vetores numéricos, contendo informações sobre sua estrutura e propriedade para representação no espaço químico. Esse tipo de tratamento matemático é bastante eficaz para processamento de grandes bases de dados, construção de modelos de relação estrutura-atividade, sejam qualitativos (SAR) ou quantitativos (QSAR), e estrutura-propriedade (QSPR, do inglês Quantitative Structure-Property Relationship), busca por similaridade química e agrupamento de banco de dados por similaridade de descritores.⁵⁶ Atualmente, a quiminformática já é empregada em toda a gênese de descoberta de fármacos, principalmente devido à redução de custos e pelo aproveitamento de conhecimento químico de

ligantes, substratos e produtos de enzimas e receptores para guiar a busca de moléculas com características similares. Essa abordagem também pode ser denominada como planejamento molecular baseado na estrutura do ligante.

1.4.4 Abordagens computacionais de modelagem baseada em campo de força molecular

Além da quiminformática, abordagens computacionais de modelagem baseada em campo de força molecular têm sido bastante empregadas na triagem virtual do banco de moléculas contra a estrutura tridimensional do alvo biomacromolecular conhecida por cristalografia de Raios-X, RMN ou por modelagem *in silico*. Desta forma, esse tipo de enfoque é denominado planejamento baseado na estrutura do alvo biomacromolecular.

Ao passo que a quiminformática descreve o sistema por meio de descritores moleculares, abordagens baseadas em campo de força molecular descrevem o modelo levando em conta a posição de átomos e ligações do sistema. Seu foco é voltado para o estudo de interações intermoleculares entre o alvo e o ligante, por meio de campos moleculares baseados, em geral, em Mecânica Clássica. Dentre as ferramentas utilizadas nessa área, destacam-se a docagem molecular e simulação por dinâmica molecular. A primeira está focada na predição de modo de interação ligante-alvo e cálculo de energia de ligação entre a macromolécula e o ligante. Já a segunda, trata-se de uma simulação física de todos os átomos do sistema com o interesse em estudar a evolução temporal da dinâmica molecular do complexo, utilizando assim como a docagem, campos de força molecular para o cálculo de energia do sistema. O uso de cada uma dessas abordagens depende do tipo de informação que se tem disponível sobre o sistema biológico de interesse. (Figura 1.14)

Figura 1.14 - Esquema de diferentes estratégias de abordagem para o planejamento *in silico* de moléculas bioativas

A integração do planejamento baseado no ligante e no alvo já foi mostrada como bastante eficaz, com dados na literatura mostrando maior poder de recuperação de moléculas biologicamente ativas do que quando cada abordagem é utilizada separadamente.⁶⁰

1.5 Fragmentos moleculares como ferramenta para a descoberta de moléculas bioativas

Além do problema da navegabilidade do espaço químico, existe também a dificuldade em montar um banco de dados que seja representativo quanto à sua diversidade molecular. Uma proposta surgida e consolidada nos últimos 15 anos é iniciar a busca a partir de um banco de dados de moléculas pequenas, de até 250 Da. Esta estratégia, conhecida como planejamento baseado em fragmentos, baseia-se em um banco de dados menor, porém mais representativo em relação ao número de moléculas.⁶¹

Desde que o uso de fragmentos para a busca de ligantes de alta afinidade com proteínas foi relatado com sucesso, em publicação por Shuker e colaboradores, dos Laboratórios Abbott, em 1996,⁶² o planejamento de ligantes/compostos-líderes baseados em fragmentos tem sido amplamente utilizado pelas indústrias farmacêuticas, laboratórios e academia.⁶³ Passados 15 anos da primeira descrição de sucesso do uso desta técnica, pelo menos treze candidatos a fármaco que foram desenvolvidos utilizando essa estratégia estão agora em fase clínica.⁶⁴

1.5.1 Histórico e fundamentação teórica

A fundamentação teórica e os princípios básicos da metodologia começaram com Jencks e colaboradores, em 1981, com a introdução da hipótese que moléculas maiores poderiam ser formadas por fragmentos moleculares que possuísem características otimizadas para se ligar ao alvo. Em seguida, o grupo de Andrews, em 1984, elaborou um modelo teórico para estimar qual seria o valor máximo no ganho de potência na molécula otimizada devido à adição de um grupo.⁶⁵ Esse trabalho foi complementado por Kuntz e seus colaboradores que, em 1999, mostraram que a potência de uma molécula poderia aumentar em até 1,5 kcal.mol⁻¹ por átomo adicionado, até que a potência atingisse 12 kcal.mol⁻¹ (equivalente a um inibidor com K_d aproximadamente a 1 nanomolar).⁶⁶ Já em 2006, Hadjuk e colaboradores mostraram que uma otimização de sucesso de ligante para composto-líder tinha um acréscimo de em média 0,3 kcal.mol⁻¹ por átomo adicionado à molécula, novamente até o limite de 12 kcal.mol⁻¹.⁶¹ Esse valor de 0,3 kcal.mol⁻¹ já havia sido sugerido por Hopkins e colaboradores, em 2004, como o limite mínimo de energia por átomo na molécula para que o fragmento fosse selecionado como um possível candidato a composto-líder.⁶⁷⁻⁶⁴

1.5.2 Fragmentos moleculares e fármaco-similares

Começar a busca de moléculas com característica de fármacos (ou seja, com uma potência elevada contra o alvo) a partir de pequenas moléculas com baixa afinidade, com K_i da ordem de alto micromolar a milimolar, parece ser contraditório, ainda mais com a existência do HTS, o qual já fornece ligantes com potência próxima à desejada. Porém, por ter uma massa molecular menor, o fragmento possui um espaço químico mais restrito do que quando comparado com as moléculas utilizadas em HTS, com massa molecular perto de 500 Da. São estimados que existam 10⁸ compostos com até 250 Da, um número bem menor do que os 10⁶⁰ compostos preditos como existentes para moléculas de até 500 Da. Dentre os compostos possíveis com massa até 250 Da, 10⁶ são preditos como sinteticamente viáveis, número da mesma ordem de grandeza do número de moléculas testadas por HTS.⁶⁴

Além disso, a taxa de acerto em ensaios baseados em fragmentos moleculares é consideravelmente maior do que para HTS. Isso é explicado por provavelmente possuírem uma boa complementaridade com o alvo, devido ao fato de serem moléculas de menor complexidade

(Figura 1.15).⁶⁸

Figura 1.15 - Influência da complexidade química na interação da pequena molécula com o sítio ativo da biomacromolécula

Fonte : Adaptado de Carr, R.A.E., Congreve, M., Murray, C.W., Rees, D.C.. Fragment-based lead discovery: Leads by design. **Drug Disc. Today**, v. 10, n. 14, p. 987-992, 2005

Por terem uma maior complementaridade com o sítio, sua interação com o alvo é de alta qualidade, o que caracteriza uma elevada eficiência do ligante, que representa o quanto cada átomo contribui para a interação com a proteína, em kcal/mol. Isso faz com que a otimização futura seja mais fácil, com a adição de grupos que igualmente contribuem com a interação. Assim, o processo resulta em candidatos a fármacos menores, com todos os átomos contribuindo efetivamente para a interação e com mais liberdade para melhorar propriedades farmacocinéticas sem a perda de potência (Figura 1.15).⁶⁹

Quando a abordagem estava ainda no início, existia um temor na comunidade científica de que essas pequenas moléculas não apresentassem um complexo único e estável com a proteína para que fossem utilizadas no planejamento de fármacos. Esse receio em grande parte se desfez com a confirmação experimental de que os fragmentos, além de apresentarem um modo de interação estável com a macromolécula, apresentavam também relação estrutura-atividade (SAR) significativos.⁶¹

1.5.3 Montagem de um banco de dados de fragmentos: Regra dos Três

Para definir quais compostos são considerados fragmentos, Congreve e colaboradores descreveram, a partir de análise de diversos fragmentos da literatura que haviam mostrado interação contra vários alvos diferentes, a chamada ‘Regra dos Três’. Esse nome tem correlação

direta com a Regra dos 5 de Lipinski, regra empírica para identificar moléculas que tenham boa possibilidade de serem absorvidas via oral.⁷⁰

Esses pesquisadores encontraram que, em média, esses fragmentos possuíam uma massa molecular menor do que 300 Da, menos do que 3 átomos capazes de doar ligações de hidrogênio, menos do que 3 átomos capazes de aceitar ligações de hidrogênio e coeficiente de partição calculado menor do que três. Além disso, foi notado também que os fragmentos estudados em média possuem três ou menos ligações livres e uma área superficial polar menor ou igual a 60 Å².⁷¹ Apesar de essa não ser a única regra para a descrição de fragmentos, as outras em geral são pequenas variações desta, possibilitando construir um banco de dados de fragmentos mais focado para o sistema de interesse.

Além dessas características, é fundamental selecionar para a montagem do banco de dados de fragmentos somente aquelas moléculas com uma elevada solubilidade em água, com um corte de no mínimo 1 mM. É importante também retirar do banco de dados aquelas que têm tendência a formar agregados e que tenham atividade tóxica predita.⁷²

1.5.4 Detecção de fragmentos moleculares como ligantes

Para a aplicação do banco de dados em ensaios *in vitro* contra o alvo, faz-se necessário o uso de técnicas bastante sensíveis para a detecção da afinidade e interação do fragmento perante um alvo biológico, devido ao fato de serem inibidores de baixa afinidade. A primeira técnica utilizada com esse propósito foi o RMN, seguido do uso da cristalografia de raios-X, primeiramente empregada pelos Laboratórios Abbot em 2000. A partir daí, técnicas biofísicas vêm sendo utilizadas para mensurar a afinidade do fragmento com o alvo como, por exemplo, Calorimetria de Titulação Isotérmica (ITC), Ressonância Plasmônica de Superfície (SPR, do inglês Surface Plasmon Resonance), *tethering*, termoforese de microescala, *Time-resolved Fluorescence Resonance Energy Transfer* (TR-FRET), eletroforese capilar, TINS e espectroscopia de fluorescência diferencial. Algumas dessas técnicas, assim como a sua sensibilidade, estão mostradas na Figura 1.16⁷³.

Figura 1.16 - Limite de detecção de afinidade para diferentes técnicas e comparação com valores médios de afinidade para cada tamanho de molécula

Fonte: Adaptado de Hubbard, R.E., Murray, J.B.. Experiences in fragment-based lead discovery. **Methods in Enzymology**, v. 493, p. 509-531, 2011.

Uma das técnicas mais utilizadas atualmente, depois de RMN e cristalografia por Raios-X é a SPR, principalmente devido à sua alta sensibilidade, sendo capaz de detectar moléculas com afinidade maior do que 1 mM.

1.5.5 Crescimento e otimização de um fragmento com atividade biológica

Uma vez encontrado um fragmento com atividade moduladora sobre o alvo, faz-se necessário o seu crescimento e otimização para compostos-líderes. As técnicas utilizadas hoje são ligação, crescimento, autoligação e otimização de fragmentos, de modo que o ligante encontrado na primeira etapa tenha sua afinidade e propriedades farmacocinéticas melhoradas para chegar a um composto-líder.

A abordagem de ligação de fragmentos é utilizada quando existem duas ou mais cavidades próximas entre si que sejam passíveis de serem exploradas. Neste caso, são encontrados fragmentos que se ligam a cada uma das cavidades de interesse e é construída uma ponte, geralmente flexível, para conectar as duas moléculas (Figura 1.17a).

A abordagem de autoligação é similar à de ligação de fragmentos. Neste caso, a ligação entre os dois é feita por meio de reação química de grupos complementares, quando ambos já estão interagindo com o alvo. Uma esquematização está representada na Figura 1.17b. O terceiro tipo de tratamento é o crescimento do ligante. Neste caso, a partir de um fragmento identificado para uma cavidade, grupos químicos são adicionados ao mesmo para que ocupe outras cavidades vizinhas, aumentando assim a sua afinidade com o alvo. Figura 1.17c

Dentre esses três métodos, o mais utilizado atualmente é o crescimento do ligante. A razão é simples: existe uma dificuldade maior em identificar dois ligantes em sítios adjacentes e, além disso, o ato de unir os dois fragmentos requer muitas vezes estratégias múltiplas de síntese para que o composto desejado seja construído, sendo o composto final bastante sensível à ponte de ligação entre os fragmentos. Por outro lado, o crescimento do ligante é uma abordagem mais iterativa, onde o fragmento inicial deve manter a sua posição geométrica no sítio, a despeito da adição de grupos químicos.⁷⁴

Quando já se tem um protótipo de composto-líder, a estratégia de otimização ocorre com a substituição de regiões da molécula por outros fragmentos. Nessa etapa, além do incremento na afinidade, existe um interesse em melhorar as propriedades farmacocinéticas do composto, sendo essa a principal razão pela alta taxa de atrito na fase clínica⁷³. Figura 1.17d

Figura 1.17 - Representação de diferentes métodos de crescimento de fragmento. A) Ligação de fragmentos. B) Auto ligação de fragmentos. C) Crescimento a partir de uma âncora. D) Otimização de fragmento.

Fonte : Adaptado de Hubbard, R.E., Murray, J.B.. Experiences in fragment-based lead discovery. **Methods in Enzymology**, v. 493, p. 509-531, 2011.

O processo de otimização de qualquer molécula em geral resulta no aumento da lipofilia e da massa molecular do composto.⁷⁵ Como o fragmento possui uma baixa massa molecular, esse aumento não faz com que o mesmo deixe de obedecer à Regra dos Cinco de Lipinski. Além disso, um estudo mostrou que, ao contrário dos compostos-líderes gerados a partir de HTS, os

gerados pela estratégia baseada em fragmentos têm sua otimização predita com grande acurácia.⁷⁶ Para acompanhar a evolução do fragmento quanto à sua afinidade durante a otimização, utiliza-se o parâmetro de Eficiência do Ligante (LE, do inglês Ligand Efficiency).

1.5.6 Eficiência do Ligante

Este parâmetro foi formalizado por Hopkins e colaboradores como uma métrica para a seleção de compostos-líderes, com o objetivo de auxiliar na seleção de quais fragmentos vão seguir no processo de otimização e também guiar a evolução da molécula desde fragmento-similares até compostos-líderes. A definição de LE é dada a seguir:

$$LE = \frac{RT \ln K_i}{N_{\text{Átomos pesados}}} = \frac{\Delta G}{N_{\text{Átomos pesados}}} \quad (1)$$

onde R é a constante universal dos gases, T a temperatura, K_i a constante de inibição da proteína pela molécula e $N_{\text{átomos pesados}}$ é o número de átomos diferente de hidrogênio presentes na molécula.

O valor máximo possível de eficiência do ligante, como mostrado por Kuntz e colaboradores, é de 1,5 kcal.mol⁻¹ e o valor mínimo para ser uma molécula atrativa para ser desenvolvida a um fármaco é 0,3 kcal.mol⁻¹, como mostrado por Hadjuk e colaboradores. Nesse limite inferior, uma molécula com 500 Da, aproximadamente 38 átomos pesados, teria uma constante de ligação menor do que 10 nM.⁷⁷ (Figura 1.18).

Figura 1.18 - Comparação de fragmentos moleculares com moléculas de maior massa molecular e delimitação da área de alta eficiência de ligante.

Fonte: Adaptado de Siegal, G., Eiso, A.B., Schultz, J.. Integration of fragment screening and library design. **Drug Disc. Today**, v. 12, n. 23/24 p. 1032-1039, 2007.

Vários trabalhos na literatura têm mostrado que mesmo iniciando a otimização do ligante com LE igual a 0,3, é possível chegar a uma molécula final com boa afinidade pelo alvo, mantendo a mesma LE durante todo o processo pelo menos nos casos em que a abordagem baseada na estrutura do alvo foi utilizada durante todo o processo.⁷⁸

Apesar de vários casos de moléculas que alcançaram a fase clínica sendo planejados pela abordagem de fragmentos serem reportados na Literatura, as indústrias farmacêuticas estão deixando um pouco de lado essa e todas as outras abordagens cujo produto final é uma pequena molécula, devido à alta taxa de atrito ainda existente na fase clínica, por conta principalmente das propriedades farmacocinéticas inadequadas para administração do fármaco em seres humanos. O foco tem se voltado para os chamados fármacos biológicos, como por exemplo, anticorpos, que historicamente têm uma menor taxa de atrito quando comparado às pequenas moléculas. Porém, certas áreas, como antiparasitários, antibactericidas e antivirais, têm seu tratamento dependente de pequenas moléculas, necessitando ainda da estratégia baseada em fragmentos entre outras.⁶³

Neste trabalho, a abordagem de planejamento baseado em fragmentos foi unida com planejamento *in silico* de moléculas, por meio da estratégia baseada na estrutura do ligante e do alvo biomacromolecular, para busca de moléculas de alta eficiência do ligante contra um importante alvo terapêutico, a GAPDH de *T. cruzi*, que no futuro venham a chegar ao nível de compostos-líderes. Como se trata de um alvo parasitário, a abordagem de pequenas moléculas é bastante válida para a busca de novos fármacos.

2 Objetivos

- Encontrar fragmentos ativos contra a enzima TcGAPDH, propor o modo de interação da mesma no sítio ativo da enzima e prever a contribuição de cada região da molécula para a interação.

Para isso, foram propostas as seguintes metas:

- i. Construir uma coleção de compostos disponíveis comercialmente com características de fragmentos (MM 120-250 Da);
- ii. Identificar a presença, a conformação e a energia de interação no sítio de ligação do G3P da TcGAPDH por meio de diferentes programas de docagem molecular;
- iii. Estudar a similaridade química dos fragmentos com o substrato natural (G3P);
- iv. Selecionar moléculas para compra baseado nos itens (ii) e (iii) e em inspeção visual;
- v. Determinar os valores K_i para os fragmentos selecionados frente à TcGAPDH utilizando espectroscopia de fluorescência;
- vi. Prever o modo de interação do fragmento com o sítio ativo da enzima computacionalmente;
- vii. Construir banco de segunda geração de moléculas comerciais com mesmo esqueleto que a molécula ativa;
- viii. Prever computacionalmente a relação estrutura-atividade do fragmento ativo.

3 Materiais e métodos

3.1 Materiais

Este trabalho foi realizado no Grupo de Estudos em Química Medicinal – NEQUIMED, do Instituto de Química de São Carlos, Universidade de São Paulo (IQSC-USP). A etapa de planejamento molecular *in silico* das moléculas foi realizada no laboratório de Quiminformática e Modelagem Molecular, enquanto a parte experimental de teste *in vitro* de moléculas selecionadas foi executada no Laboratório de Biotermodinâmica de Fármacos, ambos deste grupo.

Na parte computacional, a base de dados de moléculas disponíveis comercialmente, ZINC,⁷⁹ foi utilizada como fonte de moléculas para montagem de um banco de dados. Os programas FILTER (Openeye, EUA)⁸⁰ e VolSurf (Molecular Discovery, Itália)⁸¹ foram os escolhidos para selecionar as moléculas para o banco de dados focado de fragmentos para o trabalho. O ensaio virtual de moléculas foi feito utilizando os programas para docagem molecular FlexX⁸² (BioSolveIt, Alemanha), e Glide (Schrodinger, EUA),⁸³ MakeFragLib, OMEGA,⁸⁴ Rocs⁸⁵ e Eon⁸⁶ (Openeye,USA) além da plataforma MAESTRO⁸⁷ (Schrodinger, EUA). Para visualização, foram utilizados os programas VIDA,⁸⁸ (Openeye, EUA) e CHIMERA (UCSF, EUA).⁸⁹ O computador utilizado para essa primeira parte possui a seguinte configuração: Core i7 870 2,91GHz, 4GB DDR3, com o sistema operacional Fedora 13.

A etapa de dinâmica molecular foi executada por meio do programa AMBER11 (UCSF, EUA)⁹⁰, implementado para ser executado em placas de processamento gráfico (tecnologia CUDA), e de seus programas auxiliares, implementados no AmberTools1.4. Esse cálculo foi feito em uma estação de trabalho SUPERMICRO, com dois processadores Xeon 5600, 24 Gb DDR3 ECC, duas placas Nvidia TESLA C2070. O sistema operacional instalado na máquina é o Fedora 13, complementado com o driver CUDA-Toolkit 3.2.16 e pacote MPICH2 para paralelização de placas de processamento gráfico ou e processadores.

Para a parte experimental, foram utilizados os reagentes *D,L*-gliceraldeído-3-fosfato (G3P), ácido 2-[4-(2-hidroxietil)1-piperazinil]-etanosulfônico (HEPES), ácido etilenodiamino tetra-acético (EDTA), 2-mercaptoetanol, fosfato de sódio hepta hidratado (NaH₂PO₄·7H₂O), nicotinamida adenina dinucleotídeo (NAD⁺) e o dimetilsulfóxido DMSO), todos adquiridos da empresa Sigma-Aldrich. A enzima TcGAPDH foi expressa e purificada *in house*, de acordo com

o protocolo estabelecido pela Dra. Juliana Cheleski.⁹¹ Os experimentos foram realizados no fluorímetro modelo Synergy HTMicroplate Reader (Biotek, EUA). Os compostos selecionados para o ensaio foram adquiridos das empresas Enamine, Princeton Research Biomolecular, ChemDiv, Asinex.

3.2 Métodos

3.2.1 Ensaio virtual baseado na estrutura do alvo: Docagem molecular

A docagem molecular é um método da química computacional que consiste em prever a conformação e orientação (pose) de um ligante de acordo com as restrições do sítio de interação do alvo e a partir disso, estimar a sua energia de interação e classificar os ligantes de acordo com a sua afinidade com o alvo macromolecular (proteína, enzima, RNA, DNA, membranas, entre outros).

Essa técnica é aplicada em diferentes estágios no processo de desenvolvimento de fármacos com o objetivo de cumprir três propostas principais: prever a pose de interação de uma molécula conhecida no sítio ativo da biomacromolécula, identificar novos ligantes usando o ensaio virtual e prever a afinidade de ligação de compostos relacionados a partir de uma série de ligantes ativos.⁹² Extensas coleções de compostos, geralmente disponíveis comercialmente, são docadas na estrutura do alvo em estudo por um programa computacional de docagem molecular.⁹³ Cada composto é amostrado em milhares de possíveis conformações e pontuado com base na sua complementaridade eletrostática hidrofóbica e estereoquímica com o alvo. Os ligantes mais bem pontuados por essa etapa são selecionados para uma próxima etapa, que pode ser o ensaio experimental (Figura 3.1).⁹⁴

Figura 3.1 - Esquema representando as etapas da docagem molecular: usando a estrutura de um receptor e um conjunto de moléculas, os programas de docagem molecular avaliam a complementaridade do ligante em relação ao sítio ativo do receptor. Aquelas mais bem pontuadas são avaliadas numa próxima etapa através de um ensaio experimental.

Fonte : Adaptado de: Shoichet, B. K., Virtual screening of chemical libraries. *Nature*, V. 432, n. 7019, p. 862-865, 2004.

Para que essa técnica seja viável, é preciso que se tenha a estrutura tridimensional do alvo macromolecular escolhido, com parâmetros cristalográficos adequados. Em geral, a estrutura é obtida por meio de técnicas experimentais como a cristalografia de raios-x e RMN. Porém, com o avanço da proteômica, os alvos farmacêuticos vêm sendo identificados em uma velocidade muito maior do que a resolução de sua estrutura tridimensional, de modo que se observa o aumento no número de estudos de docagem molecular que se utilizam estruturas previstas computacionalmente por modelagem molecular, principalmente por modelagem comparativa.⁹⁵ O banco de dados de ligantes que passará pelo processo de triagem virtual frente ao alvo escolhido é obtido por coleções *in house* de moléculas, lista de moléculas sinteticamente viáveis para o grupo ou moléculas disponíveis comercialmente.

Três são os processos chave em uma docagem molecular: representação do alvo, busca no espaço conformacional do ligante e proteína, e classificação das possíveis soluções obtidas.

3.2.1.1 Representação do Alvo

Na maioria dos programas de docagem molecular disponíveis atualmente, o alvo é tratado como rígido, como no caso da teoria chave-fechadura para interação de pequena molécula – alvo, enquanto o ligante é tratado como flexível. Essa teoria, apesar de já ter sido substituída pelo modelo de ajuste induzido por grande parte dos pesquisadores nos últimos anos, ainda é bastante utilizada pelos programas de docagem molecular, devido principalmente ao alto custo

computacional relacionado ao cálculo de mudanças conformacionais de uma biomacromolécula, necessário para a implementação do modo de interação ligante-macromolécula.

A representação do sistema na docagem molecular é feita de forma atômica, por meio de grades ou por superfície. A representação atômica é adotada em geral em conjunto com as funções de pontuação e já durante a fase final de classificação das moléculas, devido à elevada complexidade computacional de tratar o sistema desta forma. A representação por meio de superfícies é usada quase que exclusivamente para docagem molecular macromolécula-macromolécula. Para docagem molecular entre pequenas moléculas e macromoléculas, o método mais comumente utilizado é a representação por meio de grades. Basicamente, essa representação armazena informações sobre as contribuições eletrostáticas e de van der Waals em cada ponto de uma grade tridimensional que representa o sistema macromolecular. Essa é a única informação do alvo que vai ser utilizada durante os processos posteriores de docagem molecular.

3.2.1.2 Busca no espaço conformacional do ligante

A busca pelo espaço conformacional do ligante para encontrar a pose da molécula no sítio é um dos grandes problemas enfrentados pelo método de docagem molecular. O número de conformações possíveis de uma molécula aumenta exponencialmente com o número de ligações de rotação livre. Por exemplo, considerando uma molécula com 10 ligações de rotação livre, são possíveis 59.409 conformações se considerada apenas três conformações para cada ligação livre. Caso sejam consideradas seis conformações possíveis para cada ligação de rotação livre, o número de conformações possíveis chega a $3,48 \times 10^9$ para uma única molécula. Várias estratégias são adotadas atualmente para contornar esse problema, dentre elas, o método estocástico e o sistemático.

3.2.1.2.1 Algoritmo estocástico

O método estocástico faz a varredura por todo espaço conformacional da molécula de forma aleatória. Tal método pode usar dois algoritmos para o estabelecimento das conformações: Monte Carlo e algoritmos genéticos.

O método de Monte Carlo gera de forma aleatória uma nova conformação para a molécula a partir de uma conformação inicial. Essa nova conformação tem sua energia de interação com o

alvo calculada e caso seja menor do que a conformação inicial, a molécula é aceita. Caso contrário, é sujeita ao critério de Metrópolis, o qual consiste na geração de um número aleatório entre zero e um e a comparação desse número ao fator de Boltzmann ($e^{-\frac{\Delta E}{k_B T}}$). Caso o número seja maior, a conformação é rejeitada e caso contrário, é aceita. Esse procedimento é repetido até que o número de conformações desejado seja atingido. Esse método é implementado nos programas de docagem molecular Autodock-Vina, MOE-Dock, SLIDE, entre outros.

O método de busca por algoritmo genético se baseia na teoria da evolução e seleção natural proposta por Darwin. No caso das conformações da molécula, suas características de interação com o alvo são descritas, analogamente, por meio de objetos conhecidos como “cromossomos”. O cruzamento desses “cromossomos” gera uma nova série de conformações. As que possuem características “genéticas” favoráveis à interação são mantidas, enquanto as que possuem características desfavoráveis são excluídas. Esse processo é repetido por um número pré-definido de vezes, determinado pelo usuário. Dentre os programas que utilizam dessa metodologia para varrer o espaço conformacional do ligante destacam-se o DOCK, o GOLD e também o AutoDock.

3.2.1.2.2 Algoritmo sistemático ou construção incremental

O método por construção incremental baseia-se na fragmentação da molécula e posterior reconstrução da mesma diretamente no sítio de interação do alvo. Essa metodologia pode ser feita pela docagem molecular dos diferentes fragmentos no sítio de interação e posterior ligação entre os fragmentos (também conhecido como método *de novo*) ou pela divisão dos fragmentos gerados em núcleos (porção rígida) e cadeias laterais (porção flexível). Os núcleos são então docados no sítio de interação do alvo e posteriormente as cadeias laterais são adicionadas de forma incremental para a reconstrução da molécula. Os principais programas que utilizam essa metodologia são o DOCK, o FlexX, e o Glide, entre outros.

3.2.1.3 Algoritmos de pontuação

A classificação das poses dos ligantes do banco de dados de forma correta é de fundamental importância para triagem virtual de moléculas baseada na estrutura do alvo. Uma falha nessa etapa torna o usuário incapaz de distinguir uma molécula ativa de um falso-positivo. Técnicas de simulação de energia livre, uma forma eficaz de realizar a predição da energia de

interação da pequena molécula com o alvo, estão sendo desenvolvidas e já apresentam resultados satisfatórios de predição,⁹⁶ porém ainda se trata de um algoritmo inviável de ser aplicado em larga escala, devido ao seu elevado gasto computacional. Os algoritmos utilizados nessa etapa da docagem molecular são obtidos através de simplificações e aproximações e não levam em conta de forma completa todos os parâmetros que determinam o reconhecimento molecular, como a perda de entropia. Dentre eles, os mais utilizados são os baseados em campo de força, os empíricos e funções de pontuação baseadas em conhecimento.

3.2.1.3.1 Algoritmo de pontuação baseado em campo de força

O método baseado em campo de força calcula a energia de interação alvo-ligante e a energia interna do ligante (sujeito a restrições moleculares devido à interação com o alvo) por meio de campos de força moleculares, como Amber⁹⁷ e Tripos.⁹⁸ Em geral, a interação entre o ligante e o alvo é descrita por meio dos termos eletrostático (formulação de Coulomb dielétrica dependente da distância que descreve contribuição carga-carga) e potencial de Lennard-Jones. Entre principais desvantagens desse método está no fato de não incluírem termos de solvatação e entropia para a mensuração da energia de interação. Programas mais recentes, como as novas versões do AutoDock, do Gold e do Glide passaram a incluir termos que levam em conta tanto a entropia quanto fenômenos de solvatação para o cálculo de energia.

3.2.1.3.2 Funções de pontuação empíricas

Essa classe de funções de pontuação é baseada no princípio de que a energia de interação pode ser aproximada como a soma ponderada de termos não correlacionados. O coeficiente de ponderação desses vários termos é obtido através de uma regressão linear utilizando dados experimentais de energia de interação e dados estruturais, baseados em estruturas obtidas por cristalografia de Raios-X. Em outras palavras, esse tipo de função de pontuação é construído de modo a reproduzir resultados experimentais, por exemplo, de energia de interação, como a soma de diversas funções parametrizadas.

As funções de pontuação podem incluir contribuições não-entálpicas, como o chamado termo de rotação, uma aproximação para perda de entropia devido à ligação com o alvo e dependente do número de ligações de rotação livre presentes no ligante. Funções mais robustas

também levam em conta efeitos de solvatação e dessolvatação do alvo e do ligante. Entre os mais utilizados, se destacam o ChemScore e o F-Score.

3.2.1.3.3 Funções baseadas em conhecimento

Essas funções, ao contrário das outras, têm como foco reproduzir estruturas tridimensionais de complexos obtidas experimentalmente, ao invés de calcular a energia de interação do complexo. O sistema alvo-ligante é descrito por meio de potenciais de interação entre pares atômicos relativamente simples, permitindo uma varredura de grandes bancos de compostos. Por outro lado, possui como desvantagem o fato de ser baseado em um número limitado de complexos alvo-ligante. Entre os algoritmos que se destacam nessa categoria são o PMF (do inglês *Potential of mean force*) e o DrugScore.⁹⁹

3.2.2 Simulação por Dinâmica Molecular

A simulação por dinâmica molecular (MD) é um método de modelagem computacional que simula o comportamento de uma molécula ou um sistema molecular ao longo do tempo. Essa abordagem possui a capacidade de simular as condições *in vitro* e também proporcionar informações detalhadas sobre o comportamento conformacional de um complexo molecular. O crescente interesse dos cientistas em conhecer o mundo em nível atômico é saciado em parte pela cristalografia de raios-X, capaz de descrever o sistema molecular com resoluções bastante elevadas, entre 4 e 0,5 Å. Essa mesma técnica, através imagens obtidas a partir de vários cristais, pode também prover informações sobre o comportamento dinâmico do sistema.

O grande interesse no estudo da dinâmica das moléculas em nível atômico está no entendimento de fenômenos como movimento de domínios, influência da presença de uma pequena molécula na conformação da enzima, enovelamento proteico, rotação de cadeias laterais, dissociação de complexos, entre outros. Esses fenômenos acontecem em escala de tempo diferentes, desde femtossegundos até segundos, minutos, até horas (Figura 3.2).¹⁰⁰

Figura 3.2 - Tempo de ocorrência de fenômenos em um sistema proteico

Fonte : Adaptado de Skjærven, L., Reuter, N., Martinez, A., Dynamics, flexibility and ligand-induced conformational changes in biological macromolecules: a computational approach, **Future Medicinal Chemistry**, v. 3, n. 16, p. 2079-2100, 2011

Devido à grande dificuldade e o alto custo de se estudar a dinâmica de sistemas macromoleculares por meio de técnicas experimentais, a MD surge como uma ferramenta computacional robusta para se obter as mesmas informações possíveis de serem obtidas experimentalmente, como movimento de cadeias proteicas, transições alostéricas, enovelamento e desenovelamento de proteínas devido a presença de agentes desnaturantes ou devido a efeito de temperatura, estabilidade de complexos pequena molécula-macromolécula ou macromolécula-macromolécula, entre vários outros. Porém, devido ao número de processos descritos por mecânica quântica presente em movimentos atômicos de um sistema molecular ser considerável, o cálculo dos mesmos por essa mecânica para um sistema complexo é inviável computacionalmente. Assim, o método de MD clássica é construído baseado em aproximações baseadas na mecânica clássica newtoniana para descrever o comportamento do sistema, reduzindo enormemente a complexidade computacional. O processo de cálculo da trajetória do sistema por MD pode ser resumido na Figura 3.3.¹⁰¹

Figura 3.3 - Procedimento para o cálculo de trajetória de átomos utilizando simulação por Dinâmica Molecular

Fonte : Adaptado de Durrant, J.D., McCammon, J.A., Molecular dynamics simulations and drug Discovery, **BMC Biology**, v. 9, n. 71, 2011.

O modelo inicial do sistema para se iniciar a dinâmica molecular pode ser obtido de qualquer fonte que dê a informação da posição de cada um dos átomos, podendo ser cristalografia por Raios-X, RMN ou modelagem molecular comparativa, *ab-initio*.

O cálculo das forças atuantes em cada átomo do sistema é feito por meio dos campos de força moleculares, conjunto de funções de energia potencial dependentes da distância entre os átomos e de parâmetros que descrevem as características das interações do sistema, considerando os átomos como uma massa pontual e também a aproximação de Born-Oppheneimer.¹⁰² Para biomoléculas, os campos de força comumente utilizados são o CHARMM, GROMOS, AMBER, entre vários outros. Esses campos de força utilizam o princípio de aditividade para calcular a energia do sistema por meio da soma de potenciais e a transferabilidade para utilizar parâmetros em comum para grupos químicos similares.

A descrição da energia do sistema é feita como a soma da energia de interação de átomos ligados com a energia de interação entre átomos não-ligados:

$$V(r) = E_{\text{ligado}} + E_{\text{não-ligado}} \quad (2)$$

A energia para átomos ligados está relacionada com rotações torcionais, movimento harmônico das ligações, tanto em relação ao comprimento quanto ao ângulo. Matematicamente, a energia para átomos ligados pode ser descrita em geral como:

$$E_{ligado} = \frac{1}{2} \sum_{ligação} k_l (l - l_0)^2 + \frac{1}{2} \sum_{ligação} k_\theta (\theta - \theta_0)^2 + \frac{1}{2} \sum_{diedro} k_\phi (1 + \cos[n_\phi \phi - n_\phi]) \quad (3)$$

onde l representa o comprimento de ligação, k representa as constante de forças, θ representa os ângulos de ligação, Φ representa o ângulo torcional e n_ϕ e n_ϕ representam respectivamente parametros de periodicidade e deslocamento de fase para cada átomo. Já a energia para átomos não ligados é descrita como função da repulsão e atração eletrostática e de van der Waals.

$$E_{n\grave{a}o-ligado} = \frac{1}{2} \sum_{ij} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right) + \frac{1}{2} \sum_{ij} \frac{q_i q_j}{r_{ij}} \quad (4)$$

onde o primeiro termo representa interação entre dois átomos, modelado pelo potencial 6-12 de Lennard-Jones, onde A_{ij} e B_{ij} são coeficientes obtidos experimentalmente, e o segundo termo representa o potencial de Coulomb como função de cargas parciais q_i e q_j e da distância entre as mesmas r_{ij} . Todos esses termos são parametrizados nos campos de força de modo a reproduzirem os resultados obtidos por meio de cálculos e simulações quânticas.

O campo de força é responsável então por determinar a força atuante por todo o sistema em um único átomo e a partir de da segunda lei de Newton (via integração numérica) a posição e velocidade dos átomos são calculadas para um Δt de um a dois femtossegundos. Esse procedimento é executado milhares de vezes, até atingir o tempo de simulação desejado. No caso de uma simulação de 1 ns, é necessario que se rode 10^6 passos, evidenciando o alto custo computacional de uma MD. O passo de poucos femtossegundos é necessario para que se obtenha uma trajetória estável e para evitar instabilidade no movimento da molécula. Devido a isso, a MD está restrita em geral a um espaço conformacional restrito. A restrição quanto ao espaço conformacional está relacionada também à limitação energética, sendo possível o complexo simulado por MD ficar preso em mínimos locais de energia durante toda a simulação, principalmente se o tempo de simulação for curto. Alguns artifícios são utilizados para possibilitar o sistema a ultrapassar barreiras energéticas, tais como a Dinâmica Molecular

acelerada (aMD, do inglês accelerated Molecular Dynamics), que diminui a barreira de energia potencial do sistema, o arrefecimento simulado (SA, do inglês Simulated Annealing), que fornece energia ao sistema por meio do aumento de temperatura por breve período de tempo, a *Replica Exchange*, a qual faz a simulação paralelamente a diferentes temperaturas, correlacionando-as de modo que o sistema consiga acessar novas conformações só disponíveis em temperaturas mais elevadas, entre outros.

O desenvolvimento de *hardwares* tem possibilitado o avanço no tempo de simulação de MD. Vale destacar o desenvolvimento da linguagem CUDA para o processamento de cálculos matemáticos em unidades de processamento gráfico, usualmente utilizadas como placas de vídeo para melhorar o processamento de imagem. A adaptação do algoritmo da MD para possibilitar a sua execução nessa plataforma levou a um ganho de 10x na velocidade de cálculo. Como exemplo, o *benchmark* do Amber mostra que um sistema de aproximadamente 91 mil átomos em 12 núcleos de processador x5670 tem um desempenho de cálculo de 3,97 ns/dia, enquanto o mesmo sistema em uma placa de processamento gráfico Tesla C2070 tem um desempenho de 11,08 ns/dia, uma melhora de quase três vezes no tempo de cálculo.¹⁰³ Apesar dessas limitações um grande número de trabalhos tem mostrado uma boa concordância entre os resultados obtidos experimentalmente com os resultados obtidos por simulação de MD. Os programas mais utilizados atualmente são Amber, Gromacs, NAMD.

3.2.3 Similaridade química

A similaridade química é uma ferramenta da química computacional que busca recuperar as moléculas mais similares a uma determinada molécula referência. Por essa razão, é comumente utilizada para busca de moléculas bioativas quando se tem pouco ou nenhum conhecimento do alvo biomacromolecular e se conhece pelo menos uma molécula que se ligue a esse alvo, podendo ser um substrato, cofator ou inibidor.

Entre os principais métodos de medida de similaridade química se destacam a similaridade por forma, por cor e similaridade eletrostática, similaridade de densidade eletrônica, de campo de força molecular, todas elas tridimensionais, além da similaridade bidimensional, também denominada similaridade topológica.¹⁰⁴

A similaridade por forma e cor se baseia na sobreposição de volume molecular do conjunto de compostos teste em relação à molécula de referência. Enquanto a similaridade por

forma considera todos os átomos igualmente, só importando o volume do mesmo, a similaridade por cor, além da sobreposição de volume, pontua melhor as moléculas que tenham uma maior sobreposição de heteroátomos em relação ao composto de referência.

Por se tratar de uma sobreposição tridimensional, diferentes conformações tanto da molécula de referência quanto do banco de moléculas teste retornarão resultados consideravelmente diferentes. De modo a contornar essa limitação, em geral é construído antes de calcular a sobreposição e a similaridade, um banco multiconformacional do banco inicial de moléculas. Em geral, os programas como MakeFragLib e CHOMP, para gerar esses bancos, fragmentam a molécula original em vários fragmentos menores, respeitando restrições de quebra de ligação, as quais são particulares de cada programa. Então, as moléculas têm suas ligações químicas reconstruídas, utilizando programas como o OMEGA2, de forma quase que aleatória, respeitando somente critérios de torção de ligação química e contatos estereoquímicos.

Uma maneira de se buscar o melhor alinhamento e sobreposição para o cálculo de similaridade é por meio da descrição do volume da molécula em termos de funções gaussianas. A descrição matemática do volume permite uma minimização para a sobreposição máxima de modo rápido e eficiente. O programa pioneiro nessa forma de sobreposição é o ROCS.

Uma vez as moléculas alinhadas com a estrutura de referência, são pontuadas de acordo com diferentes métricas, das quais se destacam o coeficiente de Tanimoto e de Tversky. Para similaridade de forma ou por cor, o volume molecular é utilizado para a pontuação e classificação das moléculas, enquanto para similaridade eletrostática, o potencial eletrostático da molécula de referência e de teste é calculado e utilizado para comparação e classificação. Para similaridade por forma e eletrostática, os programas largamente utilizados são respectivamente o ROCS e o EON.

3.2.4 Espectroscopia de Fluorescência e detecção de reação enzimática

A fluorescência de uma molécula acontece devido à emissão de fótons devido à volta para o orbital de menor energia do elétron pareado. Substâncias com significativa fluorescência em grande parte contêm elétrons delocalizados formalmente presentes em duplas ligações conjugadas. Essa é uma transição quanticamente permitida, com frequência de emissão da ordem de 10^{-8} s^{-1} . Desta forma, o tempo médio durante o qual o fluoróforo permanece no estado excitado, ou tempo de meia vida, é próximo a 10^{-8} s .

A razão de moléculas em cada estado (excitado e fundamental) em condições sem excitação específica é dada por

$$R = e^{-\Delta E/kT} \quad (5)$$

Onde ΔE é a diferença de energia entre os níveis excitado e fundamental, k a constante de Boltzmann, T a temperatura em K. Considerando a temperatura ambiente a 300 K, a razão fica em torno de 0.01, o que implica que grande parte das moléculas estará no estado fundamental. O aumento de energia devido à vibração térmica não é suficiente para os fluoróforos popularem o primeiro estado excitado S1, em grande parte dos casos. Apesar de esse número ser baixo, é possível detectar a população do S1 causada por indução térmica por espectro diferencial de absorção realizado a diferentes temperaturas.

Durante uma absorção de luz pela molécula, o elétron pode ser excitado do nível fundamental para os níveis eletrônicos excitados S1 ou S2. Porém a relaxação para o nível excitado mais baixo acontece em cerca de 10^{-12} s, ou seja, quatro ordens de grandeza mais rápido que uma emissão de fluorescência. Deste modo, a fluorescência em geral resulta de um estado excitado termicamente equilibrado. Fenômeno semelhante acontece durante a emissão dos fótons, onde o elétron pode ir para estados vibracionais excitados do estado eletrônico fundamental. Esses fenômenos são observados nos espectros de absorção e emissão, respectivamente. Geralmente, excitação eletrônica não modifica o espaçamento dos níveis de energia vibracionais, de modo que os espectros de absorção e emissão são bastante semelhantes.¹⁰⁵

Essa técnica é usada para ensaios enzimáticos *in vitro* em sistemas onde o substrato ou o produto da reação fluoresce e então a sua concentração é acompanhada pela amplitude do pico de fluorescência no comprimento de onda específico do composto. A enzima TcGAPDH, utilizada nesse trabalho, possui como cofator o NAD^+ , o qual é convertido a NADH pela reação enzimática. Tanto o substrato quanto o produto absorvem luz em comprimento ultravioleta a 259 nm, devido à presença da adenina, e também a 339 nm para o NADH. Porém, somente o NADH fluoresce, a um comprimento de onda de 460 nm, com tempo de meia vida de 0,4 ns.¹⁰⁵ Desta forma, é possível monitorar o surgimento do produto NADH em função do tempo de reação e a partir daí, obter parâmetros cinéticos da reação enzimática, na presença e na ausência de inibidores.

3.3 Metodologia utilizada no trabalho

3.3.1 Seleção de moléculas

Uma esquematização simplificada do que foi feito durante a etapa de seleção de moléculas para ensaios *in vitro* contra a enzima TcGAPDH está esquematizada na Figura 3.4

Figura 3.4 - Fluxograma esquematizado da etapa de seleção de moléculas para testes *in vitro* e de simulação por Dinâmica Molecular

3.3.1.1 Montagem do banco de dados de fragmentos moleculares

A primeira etapa do trabalho consistiu na montagem do banco de dados somente com moléculas com características de fragmentos moleculares, comercialmente disponíveis e com solubilidade elevada, pelo menos por predição *in silico*, para posterior seleção de compostos de interesse.

Para essa montagem, as moléculas comercialmente disponíveis das empresas Asinex, Chembridge, Chemdiv, ChemicalBlock, Enamine, Life Chemicals, Maybridge, Otava, Princeton, Specs, Timtec, Vitas-M foram baixadas a partir do banco virtual de domínio público de moléculas comercialmente disponíveis ZINC. Essas moléculas foram então submetidas ao programa FILTER, programa que atua como um filtro molecular de compostos, removendo do banco as moléculas com características indesejadas a partir dos parâmetros de entrada dados pelo usuário.

No nosso caso, os principais parâmetros utilizados estão resumidos na Tabela 3.1. A lista completa de parâmetros pode ser encontrada no Apêndice 1 – Parâmetros de filtro das moléculas

Tabela 3.1 - Principais parâmetros definidos como entrada para o programa FILTER

Massa molecular	Até 250 Da	Área superficial polar	Até 60 Å ²
Coeficiente de partição	Entre -3 e 3	Número de ligações livres	Até 3
Aceitadores de Ligação de hidrogênio	Até 3	Átomos permitidos	H, C, N, O, S, P, Cl, F, Br
Doadores de ligação de hidrogênio	Até 3		

Os principais filtros moleculares utilizados foram exatamente os que definem a Regra dos Três, como descrito por Congreve. Além desses mostrados, é importante ressaltar também que foram incluídos no filtro parâmetros para remover moléculas que são preditas como propensas à formação de agregados em solução e que possuem grupamentos reativos. Um filtro para remoção de moléculas que são preditas como de baixa solubilidade em água também foi utilizado.

As moléculas que passaram por esse primeiro filtro tiveram a sua solubilidade em água predita quantitativamente pelo VolSurf. Como já dito, um filtro preliminar de solubilidade foi inserido na parametrização do Filter, porém o mesmo retorna um resultado qualitativo, insuficiente para os propósitos desse trabalho.

Foram selecionadas apenas as moléculas com solubilidade predita pelo VolSurf maior do que 1 mM. Esse valor foi determinado por ser da ordem de grandeza da afinidade de fragmentos com alvos macromoleculares. Como foram realizados experimentos *in vitro*, considerou-se como importante que o fragmento molecular fosse solúvel nessas concentrações para uma detecção mais acurada de sua afinidade, por espectroscopia de fluorescência.

3.3.1.2 Escolha e preparação da estrutura alvo macromolecular.

No início do trabalho, a enzima TcGAPDH possuía, cinco estruturas cristalográficas

depositadas no banco de dados de proteínas PDB (código PDB 1QXS, 1K3T, 1ML3, 3DMT, 3IDS). Uma análise nas informações estruturais de cada uma delas foi feita, observando parâmetros como resolução, R e R_{free} , presença de ligante no sítio do substrato.

Uma vez escolhida a estrutura cristalográfica de código PDB 1QXS, essa foi preparada utilizando o ambiente de modelagem molecular Maestro, através do *script ProteinPreparation*. Por este *script*, foi feito o ajuste nas ordens de ligação de toda a proteína, corrigidas cadeias laterais incompletas, adicionados átomos de hidrogênio e retiradas águas cristalográficas.

Ainda com o *script ProteinPreparation*, foi feita a predição do estado de protonação e tautomerização dos aminoácidos histidina, aspartato e glutamato da proteína. Os átomos de hidrogênio adicionados foram minimizados de forma a aumentar a capacidade de doação de ligações de hidrogênio entre os próprios aminoácidos da proteína e com o cofator e ligante cocrystalizados. Por fim, uma leve minimização (500 passos) foi realizada de modo a relaxar a estrutura cristalográfica. Depois, as coordenadas finais foram salvas em formato PDB para os estudos por docagem molecular.

3.3.1.3 *Preparação da coleção de ligantes para a etapa de docagem molecular*

Todos os fragmentos do banco de dados foram tratados por meio do *LigPrep*, implementado no Maestro. O *LigPrep* gera várias conformações e diferentes isômeros da molécula inicial além de gerar os possíveis estados de protonação em um determinado pH, estabelecido por usuário. Por opção, foi solicitado ao programa que exportasse no resultado as 10 conformações de menor energia da molécula e o estado de protonação mais provável a um pH de $(8,0 \pm 1,0)$, que é o pH no qual será feito o experimento *in vitro*.

3.3.1.4 *Preparação do substrato natural da enzima, o G3P*

O G3P foi retirado da estrutura cristalográfica de código PDB 3CIF, a estrutura de menor resolução contendo G3P co-cristalizado com a GAPDH no banco de dados PDB. Essa estrutura é da GAPDH de *Cryptosporidium parvum* contendo o substrato co-cristalizado, a 2 Å de resolução. Não existe qualquer estrutura tridimensional de GAPDH de *Homo sapiens* ou *Trypanosoma cruzi* cocrystalizada com essa molécula. No ambiente do Maestro, ajustou-se a ordem de ligação do ligante, adicionaram-se os átomos hidrogênio e todo o resto da estrutura cristalográfica foi

apagado. Usou-se o *LigPrep* para calcular o estado de protonação do G3P para o pH ($8,0 \pm 1,0$). A estrutura com a protonação mais provável para essa faixa de pH foi retornada pelo programa e salva em formato mol2 e sdf, sem modificar a conformação da molécula..

3.3.1.5 Docagem molecular utilizando o programa FlexX

Para a docagem molecular utilizando o programa FlexX, implementado no pacote SYBYL foi utilizada a proteína preparada na seção 4.1.1.3. O sítio de docagem molecular para esse programa foi definido a partir da posição do cofator NAD^+ e do substrato co-cristalizados na estrutura cristalográfica escolhida: uma região de 15 Å em torno dessas duas moléculas selecionadas. O cofator foi desconsiderado como parte do receptor durante a docagem molecular, sendo removido após a definição do sítio.

Além disso, foi imposta uma restrição de que, para serem consideradas, as moléculas docadas necessariamente teriam que realizar ao menos uma ligação de hidrogênio com os resíduos de aminoácidos Treonina 167 (Thr167) ou Serina 247 (Ser247), ambos mostrados por Freitas e colaboradores como de especial importância para o reconhecimento de moléculas ativas planejadas contra o sítio do substrato G3P.⁵²

Todos os outros parâmetros do programa foram mantidos como na configuração padrão, com a única modificação de que somente uma única pose da molécula docada deveria ser retornada. Uma vez preparado o sistema para a docagem molecular, todos os ligantes do banco foram submetidos a esse processo.

3.3.1.6 Preparação do sistema para docagem molecular pelo programa Glide e posterior docagem molecular no modo standard precision (SP)

Ao passo do que foi feito para o programa FlexX, a estrutura preparada na etapa 4.1.1.3 foi utilizada para a construção do Grid. Para isso, uma caixa de 15 Å centrada no centróide do ligante cocrystalizado foi definida. Adicionalmente, a restrição farmacofórica nos resíduos de aminoácido Ser247 e Thr167 também foi definida, assim como para o caso do FlexX.

Com o *Grid* da proteína calculado, foi definido o modo de docagem molecular SP (*Standard Precision*), uma função de pontuação um pouco mais flexível e mais rápida, indicada para um primeiro filtro baseado na estrutura do receptor para bancos de tamanho intermediário. Além disso, definiu-se o sistema de modo que o composto docado somente teria a sua energia de

interação com o alvo calculada se fizesse ao menos uma ligação de hidrogênio com pelo menos um dos resíduos de aminoácidos definidos como parte da restrição farmacofórica na seção 4.1.1.5.1.

Desta vez, somente os compostos do banco que passaram pelo filtro da docagem molecular realizada pelo FlexX foram utilizados aqui. As moléculas que passaram por essa etapa ficaram agrupadas em um novo banco, que foi utilizado nas próximas três etapas, que resultaram na seleção de moléculas para compra, baseado na escolha consensual.

3.3.1.7 *Estudo consensual*

3.3.1.7.1 Docagem molecular pelo Glide no modo Extra-Precision

Para essa etapa, o novo banco foi docado utilizando o Glide no modo XP (*Extra Precision*). Esse modo penaliza mais rigorosamente moléculas e conformações que realizam interações físico-químicas inapropriadas. Também reduz o peso de interações de grupos carregados da molécula com a proteína.

A docagem molecular foi realizada utilizando o mesmo Grid gerado no passo 4.1.1.5.2 assim como a obrigatoriedade de obedecer ao menos uma das restrições farmacofóricas

3.3.1.7.2 Similaridade eletrostática e por cor – ROCS e EON

Até esta etapa, o planejamento molecular foi inteiramente baseado na estrutura do alvo molecular (no caso, a TcGAPDH). Adicionalmente, foi realizado também um estudo baseado na estrutura do ligante. Isso foi feito pela análise da similaridade, eletrostática e por cor, de cada uma das moléculas que passaram por todos os passos anteriores com o substrato natural G3P da enzima em estudo.

Na etapa de preparação do novo banco de dados para o cálculo de similaridade por cor, o primeiro passo foi gerar múltiplas conformações dos fragmentos no banco de dados. Foi então utilizado o programa MakeFragLib (MFL) (configurações *default*) para fragmentar os compostos do banco de dados e o programa OMEGA para gerar múltiplas conformações, a partir dos fragmentos gerados pelo MFL. No OMEGA, alguns parâmetros foram modificados em relação ao default, como mostra a Tabela 3.2. Todos os parâmetros restantes foram mantidos como o original.

Tabela 3.2 - Parâmetros modificados ao se utilizar o programa OMEGA

Parâmetro	Valor default	Valor modificado	Significado do parâmetro
<i>FromCT</i>	True	False	Modificar ou não o comprimento de ligação e os ângulos da molécula
<i>Ewindow</i>	10	25	Diferença mínima de energia em relação à molécula original para aceitar a nova conformação
<i>maxconfs</i>	500	400	Número máximo de conformações geradas

Com o banco de dados multiconformacional preparado e com o G3P preparado como foi descrito na seção 4.1.1.4, foi feito então o cálculo de similaridade por cor pelo programa ROCS. A molécula de referência foi definida como o G3P e foi calculada a similaridade de todas as moléculas do banco de dados multiconformacional em relação à molécula de referência. Todos os valores foram mantidos como indicados pelo programa, com a única modificação no parâmetro *eon_input*, o qual foi ativado para o ROCS deixar seu arquivo de saída preparado em um formato que possa ser lido pelo EON.

O programa EON foi utilizado então na etapa subsequente para calcular a similaridade eletrostática. Neste programa, todos os parâmetros foram mantidos como originalmente propostos.

3.3.1.7.3 Análise consensual e seleção de moléculas para a compra

Com os resultados da energia de interação predita pela docagem molecular pelo Glide-XP, a similaridade eletrostática e por cor, um reescalonamento dos valores foi feito. Para a energia de interação os valores esperados são negativos (valores positivos de energia de interação significam interação desfavorável), porém existe a possibilidade de surgirem valores positivos, enquanto para os valores de similaridade química, os limites possíveis são zero (nenhuma similaridade observada) e um (composto teste e composto de referência são idênticos). Por conta disso, a equação utilizada reescalonar os valores de cada uma das técnicas entre zero e um é diferente. Para os valores de energia de interação, a equação é:

$$X_{reesc} = (x - x_{m\acute{a}x}) / (x_{m\acute{i}n} - x_{m\acute{a}x}) \quad (6)$$

enquanto a equação para o reescalonamento dos valores de similaridade é:

$$X_{reesc} = (x - x_{min}) / (x_{máx} - x_{min}) \quad (7)$$

onde X_{reesc} é o valor re-escalado para a molécula em questão, x , o valor original, x_{min} o menor valor para aquele parâmetro em todo o banco e $x_{máx}$ o maior valor para aquele parâmetro em todo o banco.

Agora, com os valores escalonados, foi possível somar os diferentes parâmetros sem que nenhum deles tivesse uma contribuição maior para o resultado final simplesmente devido à amplitude intrínseca de seu valor. Então, o valor consensual para cada um dos fragmentos do banco foi calculado como sendo a média aritmética de seus três valores escalonados, como mostrado na equação abaixo

$$x_{cons} = \frac{x_{reesc}^{G-XP} + x_{reesc}^{EON} + x_{reesc}^{ROCS}}{3} \quad (8)$$

onde x_{cons} é o valor consensual para aquela molécula, x_{reesc}^{G-XP} o valor escalonado obtido pelo Glide-XP, x_{reesc}^{EON} o valor escalonado obtido pelo EON e x_{reesc}^{ROCS} o valor escalonado obtido pelo ROCS.

Os fragmentos foram ordenados de forma decrescente por valor e as 100 melhores que ainda estavam disponíveis comercialmente nas empresas selecionadas inicialmente foram submetidas a uma inspeção visual utilizando o programa VIDA. Ao final, cinco fragmentos foram selecionados para compra e para ensaio *in vitro* bioquímico com a TcGAPDH.

3.3.2 Ensaios bioquímicos por espectroscopia de fluorescência

Os ensaios bioquímicos utilizando o fluorímetro Synergy HTMicroplate Reader da Biotek foram feitos utilizando uma placa de 96 poços. Os filtros de banda de excitação e emissão utilizados foram de 340/40 nm e 460/40 nm respectivamente. O volume final do poço para cada experimento foi de 200 μ L. Os experimentos para cada poço foram feitos em triplicata.

Os poços que sem a presença de inibidor continham Nad^+ 1 mM, enzima TcGAPDH 2 nM, Triton 0,01% v/v, tampão HEPES 100 mM e pH 8,0, MDSO 5 % v/v. A concentração do G3P variou nas colunas como segue a Tabela 3.3

Tabela 3.3 - Concentração de D,L-G3P utilizada em cada coluna da placa de 96 poços para ensaio de inibição enzimática utilizando espectroscopia de fluorescência

# coluna	1	2	3	4	5	6	7	8	9	10	11	12
[G3P] μM	3000	2100	1470	1029	720	504	252	126	63	32	16	8

A reação foi disparada momentos antes de se iniciar a leitura, com a placa já posicionada na gaveta do fluorímetro, pela injeção de íon fosfato de modo que sua concentração final no poço fosse de 20 μM

Os poços com presença de inibidor foram preenchidos da mesma maneira como descrito para os poços sem inibidor, com a diferença que a solução de MDSO 5% v/v foi substituída por uma solução contendo o inibidor diluído em MDSO, de modo que estivessem a uma concentração final no poço de 500 μM e 5% v/v.

3.3.2.1 Análise dos dados

As constantes K_M , k_{cat} e V_{max} foram determinadas a partir dos dados obtidos experimentalmente, através do ajuste não-linear dos mínimos quadrados à equação de Michaelis-Menten, usando o programa *SigmaPlot* 10.0.

3.3.3 Simulação por Dinâmica Molecular

3.3.3.1 Simulação por Dinâmica molecular da TcGAPDH na forma apo

No arquivo PDB obtido durante a preparação da proteína para docagem molecular, sem qualquer cofator ou substrato, todos os átomos de hidrogênio foram retirados utilizando a plataforma *Maestro*. A forma de protonação dos resíduos de aminoácidos, calculada para o pH 8.0 por meio do servidor H++,¹⁰⁶ servidor online para cálculo de pKa dos grupos ionizáveis da proteína para um determinado pH, foi inserida manualmente no arquivo PDB.

A preparação do arquivo de parâmetros para o Amber teve como arquivos de entrada a proteína preparada na seção 4.1.1.3 sem os ligantes cocomplexados. O campo de força escolhido para a parametrização dos resíduos de aminoácidos foi o campo MMFF99SB (parametrização já

incluída na biblioteca do *Leap*). Além da parametrização de todos os átomos no campo de força determinado, foram adicionados os átomos de hidrogênio obedecendo estado de protonação previamente calculado e uma caixa de água octaédrica truncada de 10 Å do tipo TIP3P¹⁰⁷. Íons de cloro (Cl⁻) foram adicionados de forma a neutralizar a carga total de todo o sistema.

Nessa etapa, foram salvos os arquivos de formato *inpcrd* e *prmtop*, respectivamente arquivo inicial de coordenadas e arquivo de parâmetros do sistema, para o complexo com e sem as moléculas de água.

Os dois primeiros processos da simulação por dinâmica molecular são de minimização do sistema. Primeiro, a minimização das moléculas de água e íon: 2000 ciclos de minimização com gradiente conjugado a cada 1000 passos, com os átomos da proteína fixos com constante de restrição de 500.0 kcal mol⁻¹ Å⁻². O segundo processo de minimização consistiu em 3500 ciclos de minimização com conjugação de gradiente a cada 200 passos, sem qualquer restrição.

Em seguida, na etapa de aquecimento, o sistema teve sua temperatura elevada de 0 a 300 K, com os átomos de proteína sujeitos a uma leve restrição, com constante de força de 10,0 kcal mol⁻¹ Å⁻², de modo a evitar que os mesmos adquirissem velocidade muito elevadas devido a possíveis repulsões eletrostáticas. O cálculo foi realizado a volume constante (ntb = 1), com o banho térmico do tipo Langevin, com frequência de colisão de moléculas de 2,0 ps⁻¹. Nessa etapa, foram realizados 25.000 passos espaçados por 2 fs., um total de 50 ps. de simulação.

O passo subsequente foi a homogeneização de densidade do sistema. Nessa etapa, foi considerado novamente o sistema a volume constante (*ensemble* canônico, NVT), banho térmico do tipo Langevin, com frequência de colisão de 4 ps⁻¹, durante 100.000 passos, espaçados por 2 fs., um total 200 ps. Não houve qualquer restrição no movimento dos átomos.

Por fim a última etapa antes da simulação propriamente dita foi o equilíbrio do sistema. Nesse passo os parâmetros são os mesmos utilizados durante a dinâmica molecular: sistema considerando pressão constante, *ensemble* isotérmico/isobárico, NPT (ntb=2), banho térmico do tipo Langevin, frequência de colisão de moléculas de 4,0 ps⁻¹. Esse processo foi feito em 100.000 passos, com espaçamento de 2 fs., totalizando 200 ps. Essa etapa é importante para reequilibrar o sistema à condição de pressão constante e retirar todas as restrições que por ventura ainda estivessem no sistema.

A simulação por dinâmica molecular foi feita utilizando os mesmos parâmetros empregados na etapa de equilíbrio, com a diferença no tempo total: a simulação foi feita por 10

ns., sendo 2 ns antes da etapa de arrefecimento simulado (SA, do inglês *Simulated Annealing*) e 8 ns após a etapa de SA.

A etapa de SA foi feita após o sistema se equilibrar, determinado pelo acompanhamento da variação do RMSD dos carbonos alfa do sistema proteico, e com o intuito de possibilitar que o sistema supere barreiras energéticas mais elevadas, o que faz com que saia de um mínimo local de energia e possa ir para outra região do espaço conformacional e energético. Com isso, o ligante pode sair do mínimo local estabelecido pela docagem molecular e acessar outras conformações que não seriam possíveis na temperatura da simulação. Nessa etapa, foi feita restrição de ângulos diedro, obtidas do próprio sistema, para evitar que os aminoácidos adotassem conformações biologicamente inaceitáveis quando o sistema fosse aquecido a temperaturas maiores, como, por exemplo, formação de resíduos de aminoácidos na conformação *cis*. O único parâmetro modificado foi o tempo entre cada passo, reduzido para 0,5 fs. A rampa de aquecimento e resfriamento realizada nesse processo consistiu em um aumento de temperatura de 300 K a 500 K em 20 ps, estabilização do sistema a essa temperatura por 50 ps, seguido de um resfriamento gradual à temperatura de 200 K durante 30 ps e uma nova estabilização de 50 ps a 200 K. Por fim, o sistema foi aquecido gradualmente por 10 ps à temperatura de simulação, 300 K. Essa etapa está esquematizada na Figura 3.5. Todos os processos de simulação foram realizados com restrição de vibração e de energia de ligações contendo átomos de hidrogênio (ntf = ntc = 2)

Figura 3.5 - Rampa de aquecimento utilizada para o processo de *Simulated Annealing*

Após os 10 ns de simulação, o resultado da dinâmica foi analisado por meio dos programas VMD e Chimera.

3.3.3.2 Dinâmica molecular das moléculas selecionadas

As moléculas selecionadas para a compra foram submetidas a uma simulação por dinâmica molecular. Inicialmente, construiu-se o complexo enzima-ligante, utilizando como base o sistema gerado a partir da docagem molecular pelo *Glide-Extra Precision*. A docagem molecular foi repetida da mesma forma da descrita na seção 3.3.1.7.1, com a possibilidade de se obter até 10 complexos enzima-fragmento para cada molécula.

Separadamente, o ligante teve sua carga calculada por meio do Gaussian03,¹⁰⁸ com método Hartree-Fock e função de base 6-31G*. Então os arquivos de extensão *frcmod* e *lib* foram parametrizados por meio das ferramentas implementadas no pacote AmberTools 1.4: *antechamber*, *atomtype*, *prepgen* *parmchk* e *Leap*, nesse último utilizando o campo de força GAFF. Os arquivos do tipo *frcmod* e *lib* arquivos contêm informações sobre a conectividade dos átomos e suas cargas parciais, sendo importantes para a ferramenta *Leap* reconhecer de forma adequada o ligante, quando complexado com a proteína.

A preparação da biblioteca de parâmetros deste sistema para o Amber teve como arquivos de entrada o complexo obtido pela docagem molecular e o ligante parametrizado, com extensão *frcmod* e *lib* do ligante. O campo de força escolhido para a parametrização dos aminoácidos foi o campo MMFF99SB (parametrização já incluída na biblioteca do *Leap*). Além da parametrização de todos os átomos no campo de força determinado, foram adicionados os átomos de hidrogênio obedecendo estado de protonação previamente calculado e uma caixa de água octaédrica truncada de 10 Å do tipo TIP3. Íons cloro (Cl⁻) foram adicionados de forma a neutralizar a carga total de todo o sistema. Este número é por volta de 20, variável devido à presença de carga também na pequena molécula.

Nessa etapa, foram salvos os arquivos de formato *inpcrd* e *prmtop*, respectivamente arquivo inicial de coordenadas e arquivo de parâmetros do sistema, para o complexo com e sem as moléculas de água. Esse segundo foi utilizado na etapa de cálculo de energia de interação do ligante com a enzima por meio do método de MMGB/SA. Já pensando no cálculo de MMGB/SA que foi realizado em etapas posteriores, geraram-se os arquivos *prmtop* e *inpcrd* para um sistema contendo somente a enzima com o cofator, que será representado como o receptor, da mesma forma como foi descrito no parágrafo anterior. Os arquivos de entrada de configuração para as etapas pré-MD e MD são os mesmos utilizados na seção 3.3.3.1

Após os 10 ns de simulação, o resultado da dinâmica foi analisado por meio dos

programas VMD e *Chimera*. Todo esse procedimento foi feito para todos os complexos gerados para cada uma das cinco moléculas selecionadas para compra.

3.3.3.3 Cálculo da energia de interação por meio de MMPBSA

Para o cálculo de energia de interação entre o ligante e a macromolécula, foi utilizado o programa MMPBSA, implementado no pacote do AmberTools1.4. Os arquivos de entrada foram as coordenadas de todas as configurações geradas durante os oito últimos nanossegundos de simulação, e o arquivo de parâmetros, extensão *prmtop*, para o complexo solvatado, o complexo sem água, o ligante e o receptor.

3.3.4 Seleção de compostos análogos às moléculas de boa eficiência do ligante

Uma segunda série de compostos planejados para o sítio do G3P foi selecionada para testes *in vitro*. Essa seleção foi baseada na estrutura do composto mais ativo contra a TcGAPDH e foi feita com o intuito de validar a interação do mesmo contra a enzima. Para isso, foi feita a busca de compostos comerciais similares, com a modificação um grupo funcional por vez, por outros com características químicas diferentes ao presente na molécula original, já na intenção de fazer um estudo de relação estrutura-atividade. Foram adquiridos 10 compostos para o ligante mais ativo, com os quais foram feitas simulações por dinâmica molecular, da mesma forma como foi descrito anteriormente.

4 Resultados e discussão

4.1 Seleção, ensaio enzimático e dinâmica molecular para o sítio do G3P da TcGAPDH

4.1.1 Seleção de compostos in silico

4.1.1.1 Montagem do Banco de Dados

O banco de dados bruto, obtido a partir do banco de dados ZINC possuía um total de 444.149 moléculas, das quais 109.866 preencheram todos os requisitos do filtro, descritos na seção 3.3.1.1. O próximo passo do projeto foi manter no banco de dados de trabalho somente as moléculas que possuíam solubilidade em água maior do que 1 mM, restando 55934 moléculas, denominadas como banco de dados de solubilidade, que foram submetidas ao processo de docagem molecular.

A distribuição das características das moléculas do banco de dados inicial e após os dois processos de filtragem estão representadas abaixo, na Figura 4.1.

Figura 4.1 - Distribuição do número de moléculas do banco de dados original, em vermelho e do banco de dados de solubilidade, em verde, em função do a) número de aceitadores de ligação de hidrogênio, b) número de doadores de ligação de hidrogênio, c) massa molecular, d) ligações de rotação livre, e) coeficiente de partição, cLogP, f) área superficial polar.

Da Figura 4.1, nota-se que apesar do banco inicial já ser composto por moléculas pequenas, de até 350 Da de massa molecular, um grande número delas viola ao menos um item da Regra dos Três para fragmentos moleculares. Outro fator que provavelmente foi responsável por eliminar um grande número de moléculas aparentemente foi o filtro para retirada de moléculas preditas como toxicofóricas ou como capazes de formar agregados. A Figura 4.1 mostra a predominância de moléculas flexíveis (alto número de ligações rotacionáveis), com um número maior de moléculas com capacidade de aceitar uma ou duas ligações de hidrogênio, e área superficial polar próxima ao limite de corte da Regra dos Três. Apesar desses fatores, a maior parte das moléculas possui uma massa molecular de até 200 Da.

As c.a 56000 moléculas que passaram pelos filtros tiveram o seu estado de protonação calculado para pH ($8,0 \pm 1,0$) pelo programa *LigPrep*, do ambiente de trabalho *Maestro*. Essas moléculas então estão prontas para serem utilizadas na etapa de docagem molecular.

4.1.1.2 Escolha da estrutura cristalográfica de TcGAPDH a ser utilizada para os processos de docagem molecular e simulação por dinâmica molecular

A escolha da estrutura cristalográfica a ser utilizada foi feita por meio de uma análise dos parâmetros cristalográficos, da presença e/ou ausência do cofator NAD⁺ e de outros ligantes cocrystalizados. Além disso, observou-se a posição relativa dos resíduos de aminoácidos no sítio catalítico da enzima para as diferentes estruturas. As informações cristalográficas estão resumidas na Tabela 4.1.

Tabela 4.1 - Informações cristalográficas das estruturas da TcGAPDH depositadas no PDB

Código PDB	Resolução	R_{factor}	R_{free}	cofator	Outros ligantes
1ML3	2,5	0,199	0,264	sim	covalente
1K3T	1,95	0,201	0,28	não	não covalente
1QXS	2,75	0,202	0,276	sim	não covalente
3IDS	1,80	0,202	0,241	sim	covalente
3MDT	2,30	0,190	0,252	sim	covalente

A estrutura 3IDS não foi publicada e então foi descartada. Uma seleção observando somente a resolução da estrutura levaria à escolha da estrutura cristalográfica de código PDB 1K3T. Porém uma análise da estrutura mostra que o ligante cocrystalizado realiza somente uma interação de hidrogênio com o resíduo de aminoácido His194, por meio do grupo cetona, enquanto o restante da molécula aparentemente não interage por ligações de hidrogênio com o restante do sítio. Ao contrário deste ligante, os outros cocrystalizados possuem uma maior interação com os resíduos de aminoácidos Ser247, Ser224, Arg149, Thr167 e Thr226, os quais estão diretamente envolvidos no reconhecimento do substrato G3P, como mostrado na Figura 4.2.

Figura 4.2 - Esquemática da interação dos ligantes co-cristalizados para as estruturas de código PDB a) 3CIF, b) 1QXS, c) 1K3T.

Quando se observa o posicionamento do ligante cocrystalizado na estrutura 1K3T, nota-se que este se encontra atravessado entre o sítio de interação do cofator e do substrato e, por esse motivo, o sítio de interação com o substrato se deformou. Deste modo, essa estrutura cristalográfica foi rejeitada. Pelo fato do ligante complexado em sua estrutura não ocupar o sítio de interação do substrato, a estrutura cristalográfica de código PDB 3MDT também foi descartada.

Uma comparação entre as outras estruturas cristalográficas mostra um sítio bem conservado, com todos os átomos na mesma posição relativa, inclusive com a sobreposição do fosfato do ligante S70 da estrutura 1QXS com o fosfato do inibidor covalente CYX da estrutura 1ML3. Fez-se um alinhamento tridimensional entre as estruturas, com RMSD de 0,409 Å entre ambas, para a cadeia principal dos aminoácidos. Desta forma, ambas as estruturas estão aptas para o planejamento baseado no sítio do substrato. A escolha da estrutura 1QXS se deve ao fato desta possuir um inibidor não covalente complexado, o que facilita na preparação da estrutura da proteína para os experimentos *in silico*, já que todos os resíduos de aminoácidos do sistema já estão descritos nos *softwares* utilizados. No caso de um inibidor covalente, seria necessário

modificar a estrutura cristalográfica de modo a restituir a cisteína catalítica Cys166 à sua forma encontrada na enzima ativa.

Além disso, os parâmetros cristalográficos da estrutura 1QXS e da 1ML3 são bastante similares, tanto na resolução quanto nos parâmetros R. Deste modo, faria pouca diferença entre a escolha de uma estrutura ou de outra em questão de qualidade dos dados cristalográficos.

4.1.1.3 Preparação da estrutura cristalográfica para docagem molecular

A etapa inicial da preparação da estrutura cristalográfica do alvo, de retirada de águas de cristalização, foi feita pelo fato de não se conhecer quais são exatamente as águas cristalográficas, que estão relacionadas ao processo de reconhecimento molecular ou de catálise enzimática.

A etapa de correção ou modelagem de cadeias laterais é importante principalmente em casos de estruturas cristalográficas de resolução não tão alta, como é o caso da proteína em estudo. O fato de se possuir resolução maior do que 2 Å pode levar à dificuldade na determinação da posição de cadeias laterais, o que leva alguns cristalógrafos a deixarem a estrutura da proteína incompleta ou com aminoácidos com cadeias laterais em múltiplas posições. No caso da estrutura escolhida, ao menos no sítio ativo todos os resíduos de aminoácido estavam completos e tinham cadeia lateral com conformação única. A estrutura 1K3T, por exemplo, possui a cadeia lateral da Cys166 repetida e o usuário teria de tratar a mesma antes de começar os cálculos.

O programa *Protein Preparation Wizard* adiciona átomos de hidrogênio considerando o estado de protonação dos resíduos de aminoácido do arquivo de entrada. Porém, esses resíduos de aminoácidos estão sujeitos à mudança de pKa devido ao microambiente químico no qual se localizam. Baseado nessas informações, o *LigPrep* define a posição do hidrogênio nos nitrogênios ϵ e δ , ou ambos (protonado, carga +1), da histidina. Baseado na informação do microambiente químico e na capacidade de realizar ligações de hidrogênio, define as posições dos átomos de hidrogênios de grupamentos tiol e hidroxila e define os rotâmeros mais prováveis para os resíduos de aminoácidos glutamina, asparagina e histidina. A protonação das histidinas resultante é mostrada na Tabela 4.2.

Tabela 4.2 - Estado de protonação e tautomerização dos resíduos de aminoácidos de Histidina para a enzima TcGAPDH

Histidina com 'H' no nitrogênio delta	194
Histidina com 'H' no nitrogênio épsilon	123, 150, 151, 158, 159, 176, 349
Histidina protonada	341

A última etapa da preparação foi uma minimização de energia para ajustar as coordenadas atômicas da proteína para uma conformação mais relaxada. Essa minimização de energia é capaz de fazer somente modificações locais na estrutura, sem alterações bruscas na mesma. Após essa etapa, a estrutura preparada foi salva para os processos de docagem molecular.

4.1.1.4 Preparação do G3P para os processos de docagem molecular e similaridade química

A estrutura do G3P foi obtida por meio da estrutura cristalográfica de código PDB 3CIF. A obtenção da molécula a partir da estrutura cristalográfica de um complexo proteico foi para ter o composto em sua conformação de interação com a enzima, importante para as etapas de similaridade por cor e eletrostática. A preparação pelo *LigPrep* foi útil para determinar o estado de protonação mais provável da molécula e ajustar a ordem de ligação da mesma, sem qualquer modificação na conformação da molécula. Essa foi utilizada também como molécula de entrada para docagem molecular.

4.1.1.5 Docagem molecular do banco de dados

Para a seleção dos programas a serem utilizados na etapa de docagem molecular, foi utilizado o critério de que o mesmo devia aceitar a definição de restrições por ligação de hidrogênio na docagem molecular. Dentre os disponíveis no nosso laboratório, o Glide e o FlexX eram os programas capazes de fazer esse tipo de restrição. Ambos os programas foram utilizados, por possuírem método de busca conformacional e de pontuação diferentes.

4.1.1.5.1 Docagem molecular utilizando o programa FlexX

A parte inicial do processo de docagem molecular foi realizada utilizando o programa *FlexX*. Foram feitas duas restrições de docagem molecular, forçando que as moléculas

interagissem por ligações de hidrogênio com os resíduos de aminoácidos Ser247 ou com a Thr167. A escolha dessas restrições foi feita baseada nas estruturas cristalográficas, onde tanto o ligante cocrystalizado S70 quando o ligante ligado covalentemente à cisteína catalítica, representado por CYX, interagem de forma mais pronunciada com a região os resíduos de aminoácidos citados se localizam. Esses dois resíduos de aminoácidos estão a 6,5 Å de distância e delimitam grande parte da cavidade onde se liga o fosfato inorgânico. A Figura 4.3.a mostra o sítio da enzima com os átomos dos dois resíduos de aminoácidos citados em destaque.

Figura 4.3 - a) Representação do sítio da estrutura 1QXS preparada para docagem molecular utilizando o FlexX. Aminoácidos mostrados são a SER247 e THR167 b) Átomos em bastão mostrando os resíduos de aminoácidos Ser247 (abaixo) e Thr167. Em branco, a estrutura da GAPDH humana (1U8F) e em rosa, a estrutura da TcGAPDH (1QXS) c) Alinhamento unidimensional entre as sequências de aminoácidos da enzima GAPDH humana com a enzima GAPDH de *Trypanosoma cruzi*

Na Figura 4.3.b, são mostrados os resíduos de aminoácidos que são diferentes na estrutura da enzima humana em relação à TcGAPDH. Nota-se que um dos resíduos de aminoácidos diferentes é exatamente a Ser247, a qual possui uma alanina na mesma posição da enzima humana, o que indica que forçar uma interação da molécula com esse resíduo de aminoácido pode levar a uma seletividade dessa molécula com a TcGAPDH em detrimento à HsGAPDH.

O programa de docagem molecular busca a melhor interação das moléculas do banco com a estrutura cristalográfica tratada, sempre tentando respeitar as restrições informadas no início do processo. As moléculas que não tinham energia favorável quando eram forçadas a fazer tal interação já eram descartadas do processo e não apareciam no arquivo de saída. O arquivo de saída guarda a informação da posição de cada átomo e suas energias de interação com a enzima. De todas as moléculas do banco de dados que entraram, 6546 moléculas passaram pelo filtro, correspondendo a 12% das que foram docadas.

Os parâmetros da Regra dos Três para fragmentos foram mensurados novamente para esse banco de dados restante, de modo a tentar delinear o tipo de moléculas que melhor interage com o sítio da GAPDH de acordo com o sistema de pontuação do FlexX (Figura 4.4)

Figura 4.4 - Distribuição do número de moléculas do banco de solubilidade, em vermelho e das moléculas que passaram pelo filtro do FlexX, em verde, em função do a) número de aceptadores de ligação de hidrogênio, b) número de doadores de ligação de hidrogênio, c) massa molecular, d) ligações de rotação livre, e) coeficiente de partição, cLogP, f) área superficial polar.

Observando a distribuição de propriedades dos fragmentos que tiveram sua pose predita pelo programa FlexX, nota-se que houve uma seleção de algumas características de modo a interagir com o sítio. Percebeu-se que as moléculas que passaram pelo filtro do método de docagem molecular com restrição farmacofórica possuíam em geral um valor de cLogP perto de

zero e área superficial polar perto do limite de 60 Å² da regra dos três. Além disso, foram selecionadas as moléculas que possuíam maior capacidade de aceitar ligações de hidrogênio e pouca ou nenhuma capacidade de atuar como doadora de ligação de hidrogênio. As duas últimas propriedades refletem bem a região do sítio na qual foi feita a busca: trata-se de uma cavidade com grande quantidade de resíduos de serina e treonina, os quais possuem uma hidroxila capaz de doar ligação de hidrogênio. Além disso, a presença de Arg249 nessa região contribui para o alto poder de doação de ligação de hidrogênio dessa cavidade. Essas características, juntamente com as duas primeiras, mostram que se trata de uma cavidade com grande preferência por moléculas hidrofílicas.

As outras duas propriedades estudadas, massa molecular e número de ligações com rotação livre, não estão diretamente relacionadas com hidrofília de moléculas e não houve uma seleção explícita pelo programa baseado nessas propriedades.

Foi feito um segundo experimento, rotacionando os átomos de hidrogênio em 180°. Isso foi feito com o intuito de selecionar moléculas que agissem como doadores de ligações de hidrogênio, e não mais como aceptadores, como era no primeiro caso.

O resultado dessa mudança foi que um número bem menor de moléculas foi pontuada pelo FlexX, um total de 1689, e quando se compara com as moléculas obtidas pela docagem molecular com os átomos de hidrogênio posicionados para o sítio, nota-se que nenhum composto está presente nos dois bancos, ou seja, nenhum composto foi capaz de realizar ligações de hidrogênio como aceptador e doador de átomo de hidrogênio, com os resíduos de aminoácido selecionados para a restrição.

Comparando a energia das moléculas que passaram em cada um dos experimentos anteriores, observou-se que o experimento onde os resíduos de aminoácidos do sítio se comportavam como doadores de ligações de hidrogênio possibilitavam uma interação de menor energia do que quando os resíduos se comportavam como aceptadores de ligações de hidrogênio. Isso indica que o sítio tem uma maior afinidade por moléculas que se comportam como aceptadores de ligações de hidrogênio. Por esse motivo, desta etapa em diante, foram utilizadas somente as moléculas obtidas quando os átomos de hidrogênio estavam voltados para a cavidade de interação do substrato.

4.1.1.5.2 Glide-SP

O primeiro passo para a docagem molecular utilizando o *Glide* é delimitar uma região onde o programa irá tentar encaixar as moléculas e calcular a energia de interação. Para isso, utilizou-se o centro de massa do ligante cocrystalizado S70 como referência para o centro do cubo de 10 Å de lado. Além disso, foi feita a mesma restrição do *FlexX*, de exigir que as moléculas interagissem com os resíduos de aminoácidos Thr167 e Ser247.

Neste procedimento, foram docadas todas as moléculas que foram classificadas pelo *FlexX*, utilizando o modo *Standard Precision (SP)* do programa *Glide*. A escolha do modo *Standard Precision* na primeira etapa foi por se tratar de uma primeira etapa de seleção, não necessitando de uma robustez na função de pontuação para o banco.

Só foram passadas para as etapas posteriores de seleção as moléculas que foram selecionadas para ambos os programas, possuindo então um valor consensual, caso fosse calculado, de no mínimo 0,5. A análise consensual não foi empregada da maneira tradicional porque existem vários trabalhos na literatura que mostram que o *Glide* possui um maior poder de predição para proteínas com sítios de características diversas e consegue uma boa predição do modo nativo de ligação da pequena molécula com o alvo, com a classificação dessas conformações nas primeiras posições.⁹³ Desta forma o planejamento foi feito de modo a priorizar a função de pontuação do *Glide*.

O resultado desta docagem molecular foi que 240 moléculas foram barradas, sobrando no arquivo de saída 6305 moléculas. Os dados estatísticos da pontuação das moléculas pelo *Glide*, modo *Standard Precision* estão representados na Tabela 4.3. Dessas, as 3000 primeiras moléculas foram selecionadas para dar continuidade ao trabalho. Essa seleção se baseou no valor da média de energia obtida.

Tabela 4.3 - Dados estatísticos da distribuição de valores de energia, em kcal/mol, de docagem molecular preditos pelo programa *Glide-SP*

Média	Mediana	Máximo	Mínimo
-23.86	-23.85	-0.17	-39.55

Figura 4.5 - Distribuição de valores de energia de docagem molecular predito pelo programa Glide-SP

4.1.1.6 Estudo consensual

O estudo consensual de técnicas baseadas na estrutura do ligante e do alvo vem ganhado espaço no planejamento de moléculas com capacidade de ter alta afinidade pelo alvo. Vários grupos mostraram aplicações de sucesso dessa metodologia para a descoberta de inibidores ativos em concentrações de baixo micromolar a nanomolar, para diferentes alvos. Em nosso laboratório, o Dr. Helton Wiggers mostrou que o uso combinado das duas abordagens leva a uma melhor recuperação de moléculas ativas para o alvo cruzaina, quando comparado com o uso isolado de cada uma das abordagens.⁶⁰ Deste modo, como se conhece tanto o substrato natural quanto a estrutura do alvo, essa abordagem foi utilizada na segunda etapa para a escolha de fragmentos a serem testados contra o alvo *in vitro*.

4.1.1.6.1 Abordagem baseada na estrutura do Alvo

A escolha do programa *Glide*, modo XP (*Glide-XP*) foi pelo fato de possuir um algoritmo de pontuação bastante robusto, o qual bonifica a presença de grupos hidrofóbicos e hidrofílicos em cavidades hidrofóbicas e hidrofílicas, respectivamente. Além do algoritmo de pontuação mais preciso e amplo, esse modo está amparado por um sistema de busca de conformação mais abrangente, acoplado ao algoritmo de pontuação. Esse método de busca é baseado no posicionamento de uma âncora, em geral um anel, e uma reconstrução da molécula, com vários

ciclos de otimização para minimizar todos os contatos negativos de acordo com os termos de pontuação do *Glide-XP*.

O *Glide* apresenta a contribuição de termos como as interações de van der Waals favoráveis e desfavoráveis da molécula com a enzima, interações eletrostáticas e por ligação de hidrogênio, o que pode ser bastante útil na etapa de otimização e crescimento dos fragmentos e vai ser utilizado em etapas posteriores do trabalho.

O resultado final da docagem molecular foi que 1824 moléculas das 3000 docadas foram barradas, ou seja, não foram capazes de adotar uma conformação de modo a satisfazer as restrições de ligação de hidrogênio e o rigor da função de pontuação. Os dados estatísticos e a distribuição de energias obtida pela docagem molecular estão representados na Tabela 4.4 e na Figura 4.6 .

Figura 4.6 - Distribuição de valores de energia de docagem molecular predita pelo programa *Glide-XP*

Tabela 4.4 - Dados estatísticos da distribuição de valores de energia, em kcal/mol de docagem molecular predita pelo programa *Glide-XP*

Média	Mediana	Máximo	Mínimo
-27.13	-23.04	-17.36	-39.32

Apesar da molécula com energia de interação menos negativa com a enzima ter um ΔH calculado de $-23,85 \text{ kcal.mol}^{-1}$ pelo *Glide-SP*, nota-se que ao passar o conjunto de dados pelo *Glide-XP*, a molécula com energia de interação menos negativa possui um ΔH calculado de -

17,36 kcal.mol⁻¹. Essa diferença confirma que a função de pontuação do *Glide-XP* é mais rigorosa quando se comparada com o *Glide-SP*, punindo, por exemplo, posicionamento de grupos hidrofílicos em cavidades hidrofóbicas do sítio ativo da enzima.

Observando as interações de van der Waals mostradas pelo programa nas moléculas que passaram pelo teste, percebe-se claramente que a maioria delas possui grande parte das interações como favoráveis e algumas poucas muito desfavoráveis.

Durante essa etapa, a molécula de G3P preparada anteriormente também foi docada. A sua energia predita pelo programa foi de -46,365 kcal.mol⁻¹.

4.1.1.6.2 Similaridade de forma e eletrostática

Utilizando o programa *OMEGA*, inicialmente foram gerados confômeros para as moléculas presentes na atual etapa do projeto. Esses confômeros então foram utilizados como arquivos de entrada para o programa *ROCS*. Os valores de similaridade gerados por esse programa foram armazenados e as conformações foram utilizadas como entrada para o programa *EON*. Este programa retornou os valores de similaridade eletrostática do banco em relação à conformação ativa do G3P.

4.1.1.6.3 Consensus e seleção de moléculas para compra

Os valores obtidos de similaridade eletrostática e estrutural entre o banco de dados e a molécula de referência assim como os valores de energia obtidos pela docagem molecular utilizando o *Glide-XP* foram reescalados, de modo que os valores máximos e mínimos de cada um deles seja 1 e 0, respectivamente.

As 500 moléculas mais bem ranqueadas foram selecionadas e submetidas a uma inspeção visual, separando as moléculas com melhor interação e complementaridade com o sítio e a um estudo de dissimilaridade entre elas (para evitar comprar moléculas que tivessem um mesmo esqueleto químico). Na inspeção visual, foram levados em conta fatores como capacidade de toda a molécula realizar ligações de hidrogênio com a enzima (evitando aquelas que somente uma pequena região dela faz interações com a enzima), complementaridade eletrostática e de forma com a enzima, variação de grupamento químico na região da molécula que interage especificamente com os resíduos Ser247 e Thr167 utilizados como restrição no processo de

docagem molecular.

Foram separadas 19 moléculas possíveis de serem adquiridas, as quais foram buscadas no banco de dados Emolecules¹⁰⁹, para verificar se estavam realmente disponíveis para compra. Destas, foram encontradas oito moléculas, sendo que as cinco melhores classificadas pelo estudo consensual foram adquiridas. A estrutura bidimensional das moléculas selecionadas, definidas como moléculas de primeira geração, são mostradas na Figura 4.7

Figura 4.7 - Estrutura bidimensional das moléculas selecionadas na primeira geração para o sítio do G3P

<p>NEQUIMED 293</p>		<p>NEQUIMED 294</p>
<p>NEQUIMED 292</p>		<p>NEQUIMED 291</p>
<p>NEQUIMED 295</p>

Nota-se pela estrutura química que as moléculas NEQUIMED 292 e NEQUIMED 294 são bastante similares, com diferença na presença de um grupamento éter a mais na estrutura NEQUIMED 294 e a troca de um grupo carboxila por um grupo amino da estrutura NEQUIMED 292 para a estrutura NEQUIMED 294. Essas pequenas variações na estrutura química dos compostos foi suficiente para que fossem posicionadas pelo *Glide-XP* em diferentes cavidades do sítio. Em comum, somente a interação das duas moléculas com o sítio de ligação do fosfato inorgânico, devido às restrições de docagem molecular, já que inclusive o anel em comum foi posicionado em diferentes regiões. Apesar de atuarem em sítios diferentes em ambos casos,

apresentaram uma boa complementaridade, o que pode se traduzir em uma boa eficiência do ligante nos testes *in vitro*.

A molécula NEQUIMED 291 foi selecionada visualmente pela presença do grupo sulfona, que pelo resultado da docagem molecular está fazendo o papel do fosfato inorgânico. A molécula NEQUIMED 293 possui nessa cavidade outro grupo que pode mimetizar as características do fosfato em relação à capacidade de realizar ligações de hidrogênio, o grupo carboxílico. Outra característica que se nota comum nesses dois compostos é um átomo de nitrogênio capaz de doar ligação de hidrogênio para resíduos de aminoácido na enzima.

Durante o processo de seleção visual foi observado que grande parte das moléculas mais bem classificadas nos processos de docagem molecular possuía um grupamento carboxílico nas proximidades dos resíduos de aminoácido selecionados como restrição. Esse fato provavelmente está relacionado à presença de inúmeros resíduos de aminoácidos polares capazes de doar ligações de hidrogênio nessa região e o fato do grupo carboxílico ser um bom aceitador de ligação de hidrogênio e ser um grupo largamente encontrado em moléculas orgânicas. O mesmo foi observado para as moléculas mais bem ranqueadas pelo processo de seleção baseada na estrutura do ligante. Neste caso, a recuperação de varias moléculas com esse grupo possivelmente foi devida à elevada similaridade eletrostática com o grupo fosfato presente no G3P e também à similaridade por cor elevada com a função cetona também presente no substrato natural.

A escolha de somente um composto contendo o grupo carboxílico foi baseada no conhecimento prévio, devido a trabalhos anteriores do grupo, de busca de compostos bioativos planejados para o sítio do G3P, de que moléculas com esse grupamento interagindo com o sítio do fosfato inorgânico quando testadas contra a TcGAPDH não apresentam atividade inibitória contra a enzima.

4.1.2 Teste de atividade por Espectroscopia de Fluorescência

Os ensaios para a determinação do K_i das moléculas de primeira geração contra a TcGAPDH por espectroscopia de fluorescência mostraram somente uma das moléculas apresentou inibição maior ou igual a 50% da enzima quanto feito o ensaio único a 500 μ M de inibidor. Das outras quatro moléculas, as moléculas NEQUIMED 292 e NEQUIMED 291 apresentaram atividade inibitória, com um K_i determinado acima de 1 mM, o que também é esperado para fragmentos moleculares. Porém, como a concentração de molécula utilizada para o

ensaio foi bem menor do que o K_i da mesma, o erro associado à medida foi bastante da mesma ordem de grandeza que a própria medida, o que impediu de aproveitá-los. Tabela 4.5

Tabela 4.5 - Dados de constante de inibição obtidos para as moléculas de primeira geração em ensaio enzimático contra a TcGAPDH utilizando espectroscopia de fluorescência

Molécula	$K_i \pm \text{erro} (\mu\text{M})$	R^2
NEQUIMED 293	425 ± 53	0,974
NEQUIMED 292	3116 ± 1962	0,959
NEQUIMED 294	-	-
NEQUIMED 291	3225 ± 1653	0,971
NEQUIMED 295	5525 ± 6433	0,936

Os experimentos não foram repetidos em concentrações mais elevadas pela insuficiência de fragmentos para tal operação. A Figura 4.8 mostra o gráfico de Michaelis-Menten para a molécula NEQUIMED 293, a qual apresentou maior atividade.

Figura 4.8 - Gráfico de Michaelis-Menten para a molécula NEQUIMED 293

O K_i obtido para a molécula NEQUIMED 293 por meio dos dados experimentais expressos no gráfico foi de $425 \pm 53 \mu\text{M}$. Para essa molécula, a eficiência de ligante é de 0,33, um valor elevado, já acima do limiar de 0,3 para ser considerado um fragmento promissor para desenvolvimento a moléculas com características de fármacos.

É importante ressaltar que a TcGAPDH possui um sítio bastante extenso, o que dificulta na localização de inibidores com boa afinidade com a mesma. Se tratando de fragmentos moleculares, moléculas com uma massa molecular abaixo de 300 Da, a dificuldade é ainda maior.

4.1.3 Dinâmica molecular para confirmar posição de fragmento no sítio

Paralelamente ao estudo de inibição enzimática, foi feito um estudo *in silico* do comportamento da enzima e do ligante em água, por meio de MD. Por esse método, é possível simular a dinâmica do complexo enzima-fragmento em solução, e verificar se as interações observadas pelo método de docagem molecular são conservadas e fortes o suficiente para manter a pequena molécula complexada na macromolécula em uma pose estável no sistema. Além disso, permite observar possíveis modificações conformacionais da enzima em presença do fragmento complexado.

Para obter-se um padrão do comportamento da enzima durante uma MD, foi feita inicialmente uma simulação da TcGAPDH na presença do cofator e na sua forma apo. Para a simulação utilizando a enzima complexada somente com o cofator, foi utilizada a estrutura 1QXS para a preparação do arquivo de entrada da MD, após a retirada do ligante cocomplexado. O mesmo aconteceu com a forma apo da proteína: a estrutura 1QXS teve todos os ligantes cocomplexados retirados, inclusive o NAD^+ , para posterior preparação.

Após os 10 ns de simulação, acompanhou-se a evolução de cada um dos componentes do complexo em função do tempo de simulação. Uma das maneiras de se acompanhar essa evolução é por meio do cálculo do RMSD de cada configuração de cada um dos elementos em relação a uma configuração predeterminada. Escolheu-se, portanto, analisar o processo em relação à configuração inicial da simulação e em relação à configuração final da mesma.

Além da evolução estrutural, é importante também analisar a estabilidade das ligações de hidrogênio estabelecidas entre os ligantes, quando presente, e a proteína. Nesse processo, verifica-se por quanto tempo uma determinada ligação se mantém ao longo da simulação. No caso do nosso alvo de estudo, devido à presença de vários resíduos de aminoácidos com as mesmas características para ligações de hidrogênio, é possível que a pequena molécula perca a interação com um resíduo de aminoácido, mas forme outra interação com resíduos similares próximos.

4.1.3.1 Controle – Apo

Para a enzima apo, observou-se o comportamento dos resíduos de aminoácidos da cadeia principal mostrado na Figura 4.9

Figura 4.9 - Gráfico de RMSD em função do tempo para a simulação por MD para o sistema contendo somente a enzima TcGAPDH.

Observa-se um aumento gradual do RMSD da proteína, até uma estabilização perto de 2,5 Å. Desta forma, mostra-se que a MD como foi configurada não leva a movimentos inesperados e bruscos do sistema proteico, com variação de 2,5 Å em 2,3 ns sendo possível prosseguir para as próximas etapas. O comportamento do gráfico pode ser explicado como uma relaxação da estrutura cristalográfica da biomacromolécula, já que em geral a formação de cristal leva a uma restrição de movimentos dos átomos da mesma. Nota-se que a estabilização do sistema acontece momentos antes de ocorrer o arrefecimento simulado (SA) (2,5 ns), indicando que a escolha do momento de oferecer energia ao sistema foi acertada.

A MD, no tempo de simulação realizado, não foi capaz de identificar modificação estrutural do sistema devido à ausência do cofator no sítio ativo da enzima. Esse resultado já era esperado, pois modificações alostéricas do sistema acontecem em escala de tempo superior (da ordem de microssegundo a segundo)¹⁰⁰ ao que foi realizado. Uma inspeção visual do sítio ativo da enzima após a MD mostra que a cavidade de acomodação do NAD^+ se manteve intacta, sem maiores modificações, enquanto a cavidade de interação do G3P sofreu modificações de cerca de 2 Å na cadeia lateral dos resíduos de aminoácidos, mas suficiente para deformar o sítio de maneira a dificultar o planejamento de inibidores para o sítio do substrato.

4.1.3.2 Controle – Holo

O comportamento dos resíduos de aminoácidos do sistema proteico e do cofator NAD^+ está representado na forma de RMSD em função do tempo de simulação na Figura 4.10

Figura 4.10 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator, levando em conta a posição dos átomos a) da proteína b) do cofator NAD^+ .

O comportamento dos resíduos de aminoácidos se assemelha bastante ao observado para a MD na ausência do NAD^+ , o que se sugere que a relaxação dos resíduos de aminoácidos se trata de um fenômeno mais pronunciado do que qualquer modificação conformacional causada pelo cofator, no tempo de simulação realizado.

O comportamento do cofator durante a simulação pode ser dividido em dois patamares. O primeiro é basicamente a posição original cristalográfica do mesmo no sítio ativo da enzima. O segundo, que ocorre a 5 ns de simulação, trata-se de uma pequena modificação de torção das ligações fosfodiéster. Essa modificação altera ligeiramente a conformação do cofator no sítio, porém os dois grupos de maior importância para interação com a proteína, a nicotinamida e a adenina, mantêm-se na mesma cavidade, sem a perda de ligações de hidrogênio, como mostrado na Figura 4.11:

Figura 4.11 - Conformação e posicionamento do cofator NAD^+ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.3.3 Dinâmica molecular das moléculas selecionadas para o sítio do G3P – Primeira geração

As simulações de MD foram realizadas utilizando os mesmos parâmetros que foram utilizados para as simulações controle. Desta forma, foi possível perceber que o padrão de relaxação da estrutura proteica durante a simulação se mantém idêntico ao observado para a simulação contendo somente a enzima TcGAPDH, ou seja, a presença ou ausência de ligantes no sítio ativo da enzima não afeta a estrutura do complexo proteico de forma mensurável no tempo de simulação realizado. Além disso, observou-se que a presença do ligante pode influenciar na estabilidade do NAD^+ no sítio ativo da enzima, inclusive com a sua saída, ao menos parcial, do sítio, como foi o caso da simulação para o ligante o ligante NEQUIMED 294. A tabela 4.6 mostra a energia de interação calculada para cada uma das simulações por MD realizadas por meio da técnica de MMGB/SA.

Tabela 4.6 - Energia de interação calculada computacionalmente por MMPBSA para as moléculas de primeira geração.

Molécula	Modelo	Energia (kcal/mol)	Desvio médio (kcal/mol)
NEQUIMED 293	1	-18,43	4,18
NEQUIMED 292	1	-16,74	4,78
NEQUIMED 295	3	-12,01	3,93
	1	-9,00	2,41
	2	-6,15	5,01
NEQUIMED 294	2	-10,72	4,66
	1	-7,63	3,74
NEQUIMED 291	1	-4,26	3,09

4.1.3.3.1 Molécula NEQUIMED 293

Os resultados de simulação por MD para essa molécula mostram que a mesma é estável em uma posição do sítio, alcançada logo após a etapa de SA. Além de estabilidade conformacional, observada pelo gráfico de RMSD, nota-se também que existe uma grande estabilidade energética, com a energia em média sempre por volta de -20 kJ.mol^{-1} , valor bem próximo do obtido para toda a simulação. Adicionalmente, a tabela 4.6 mostra que esse valor é significativamente maior do que para as outras moléculas. Esse resultado está de acordo com o obtido experimentalmente, onde essa foi a única molécula que teve um K_i mensurável quando feitos testes de inibição da enzima com uma concentração de $500 \mu\text{M}$ de inibidor. Nota-se que a molécula se afastou 4 angstrom da região onde interage o G3P, porém também ocasionou em uma mudança conformacional do sistema de modo que o substrato tenha dificuldades de interagir com a enzima (Apêndice 3 – Gráficos de simulação por Dinâmica Molecular). Essa modificação sugere que a interação do grupo carboxila com os resíduos de aminoácidos selecionados como restrição não é forte o suficiente para manter a molécula naquela posição.

Nota-se que esse grupo passa a interagir com a Arg249, porém sem perder interação do grupo amino da pirrolidina com o resíduo Asp210, o que leva à torção da molécula observada na Figura 4.12. A interação do grupo amino da pirrolidina com o Asp210 induz à abertura de uma nova cavidade na enzima. Essa cavidade permanecia fechada acima de 85% do tempo na simulação com a enzima na forma apo e na forma holo. A abertura dessa cavidade provavelmente

acontece devido ao fato de o Asp210 perder interação de hidrogênio que fazia com a Ser224 para interagir com o grupo amino. Além disso, comparando essa região com a enzima humana, percebe-se que o resíduo de aminoácido que ocupa a mesma posição ocupada pelo resíduo de aminoácido Asp210 na TcGAPDH é uma leucina, resíduo de aminoácido com características hidrofóbicas e incapaz de doar ou aceitar ligações de hidrogênio por meio de sua cadeia lateral. Estes resíduos possuem praticamente o mesmo tamanho, diferindo pela troca de um grupo carboxil do aspartato por um grupo isopropil da leucina. Em trabalhos posteriores, esse pode ser um ponto para busca de seletividade da molécula em relação à enzima humana

É importante ressaltar também que o cofator NAD^+ se manteve na mesma posição durante toda a simulação, informação obtida ao se observar o RMSD dessa molécula durante a MD, um indicativo da modificação da pose da molécula no sítio da proteína. Um RMSD elevado, acima de dois angstroms, mostra um deslocamento translacional da molécula no sítio, enquanto flutuações menores de RMSD dão a informação de modificação conformacional da molécula no sítio.

Figura 4.12 - Conformação e posicionamento da molécula NEQUIMED 293 e do cofator NAD^+ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.3.3.2 Molécula NEQUIMED 292-1

A MD para esse fragmento mostrou que a pose indicada pela docagem molecular, com o grupo benzodioxol fazendo interações de hidrogênio com os resíduos de aminoácidos Ser247 e Thr167, provavelmente não é a correta. Como se pode observar na Figura 4.13, os 10 ns de simulação mais o SA foram suficientes para rotacionar a molécula em 180° , agora com o grupo

carboxílico voltado para os resíduos de aminoácidos de restrição.

Uma análise da energia de interação por passo de simulação mostra que por 1 ns de simulação, o composto possuía uma energia de interação de cerca de -31 kJ.mol^{-1} , enquanto durante todo o restante da simulação após a etapa de SA, possuía uma energia média de interação de -14 kJ.mol^{-1} (Apêndice 3 – Gráficos de simulação por Dinâmica Molecular) Essa pose com melhor energia de interação possivelmente não era estável quando se considera todo o sistema macromolecular envolvido, o que fez com que o fragmento saísse rapidamente dessa conformação. Desta forma, convém considerar a energia média de interação por volta de -14 kJ.mol^{-1} . Esse valor de energia poderia indicar uma atividade da molécula em ensaios *in vitro*, o que não foi observado, para as condições de ensaio, onde o erro ficou na mesma escala que a medida

Figura 4.13 - Conformação e posicionamento da molécula NEQUIMED 292-1 e do cofator NAD^+ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.3.3.3 Molécula NEQUIMED 295-3

Essa molécula durante a simulação por MD sofreu modificações conformacionais consideráveis, como é possível se perceber pela Figura 4.14. Nota-se também que essa modificação ocorreu após o AS, mas mesmo após ter alcançado essa conformação, a molécula não ficou estável, o que sugere que a interação da molécula com a enzima não é específica. (Figura 4.14)

Figura 4.14 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 295-3, levando em conta a posição dos átomos do ligante

Quando se observa o comportamento da molécula no sítio ativo dessa enzima, mostrado na Figura 4.15, percebe-se que a mesma faz interação com a enzima em uma localidade longe do sítio de interação do G3P, sem causar deformação na estrutura da enzima.

Figura 4.15 - Conformação e posicionamento da molécula NEQUIMED 295-3 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

As outras duas simulações realizadas com essa molécula, com diferentes poses de início, também indicam que a molécula tem uma grande tendência de sair completamente do sítio da proteína, alcançando valores positivos de energia, o que indica que a interação não é favorável. Deste modo, apesar da energia de interação ter retornado um valor atrativo, era de se esperar que a molécula fosse inativa em ensaios enzimáticos *in vitro*, o que realmente foi observado.

4.1.3.3.4 Molécula NEQUIMED 294

Nota-se para essa molécula que nas duas simulações por MD, o ligante terminou em uma conformação bem similar, com um RMSD entre a posição final da molécula nas duas MD menor que 2 Å. É importante ressaltar também a mudança de conformação do NAD⁺, com o grupo amida da nicotinamida perdendo a interação com os resíduos de aminoácidos Glu336 e Asn335, o que ocorre em ambos os modelos. Provavelmente isso é devido à mudança conformacional do complexo proteico para estabilizar a pose final do ligante, o que faz com que a interação do cofator com os resíduos de aminoácidos Glu e Asn sejam perdidas. Percebe-se que na simulação 2, o grupo amino do ligante está mais próximo do cofator, o que pode ter desestabilizado a interação do grupo amida com a proteína. Com a mudança de posição do ligante, talvez o NAD⁺ não tenha conseguido reestabelecer a interação, o que levaria a uma mudança na sua conformação. Essa hipótese ganha mais força quando se observa o gráfico de RMSD do ligante e do NAD⁺. Neste gráfico, percebe-se que o ligante muda a sua conformação depois de 3,5 ns de simulação, enquanto o NAD⁺ sai do sítio depois de 5 ns de simulação. Figura 4.16

Figura 4.16 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 358T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD⁺, em vermelho

Pelo gráfico de energia por *frame* (cada 50 *frames* correspondem a 1 ns de simulação), nota-se que a energia de interação ligante-enzima é praticamente constante durante todo o tempo, sendo um pouco maior quando o ligante interage na cavidade superior, como mostra na Figura 4.17. No entanto, essa conformação aparenta ser desfavorável quando se observa toda a energia do sistema, já que logo a molécula sai dessa conformação

Figura 4.17 - Conformação e posicionamento da molécula NEQUIMED 294-2 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação

4.1.3.3.5 Molécula NEQUIMED 291

Nota-se pelo gráfico de RMSD do ligante em função do tempo que a sua conformação muda bastante, chegando a um RMSD de 14 Å. Isso mostra que a conformação (a única) obtida por docagem molecular para essa molécula não representa o modo de interação da mesma no sítio da proteína. Inicialmente, a molécula estava interagindo na cavidade de reconhecimento do fosfato inorgânico e do G3P, e no final da MD, estava localizada na superfície da proteína, em uma região logo acima do NAD⁺. Por grande parte da simulação, nota-se que a energia de interação do ligante com a enzima é maior do que -5 kcal/mol, um valor já considerado alto, indicando uma interação desfavorável entre o ligante e o enzima. Desta forma, unindo as informações de posicionamento da molécula no sítio e a energia de interação, é de se esperar que essa molécula não apresente atividade no ensaio *in vitro*, como foi observado.

Figura 4.18 - Conformação e posicionamento da molécula NEQUIMED 291 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.4 Crescimento da molécula ativa para sítios adjacentes

Em seguida, foram buscadas, em um banco de dados de moléculas comerciais, estruturas químicas que possuíam o mesmo esqueleto do fragmento ativo, com crescimento para a nova cavidade exposta pela dinâmica molecular. O requisito básico para que a molécula fizesse parte do banco de dados foi a manutenção do esqueleto químico da molécula ativa, com grupos funcionais diferentes nas extremidades do mesmo. Todas as moléculas foram docadas utilizando somente restrição de ligação de hidrogênio do fragmento 'crescido' com a Thr167 ou Ser247. Foi observado que nenhuma molécula entrou na cavidade esperada, mas esse resultado pode ser explicado pelo fato de a mesma ser de relativo difícil acesso, já que a cavidade se localiza acima de um degrau e perpendicular à posição do fragmento ativo.

4.1.5 Seleção de moléculas similares à molécula mais ativa – Segunda geração

No total, foram selecionadas para aquisição 12 moléculas baseadas na molécula mais ativa. As moléculas foram selecionadas de tal modo a trazer informações sobre relação estrutura-atividade e esclarecer o modo de interação da molécula no sítio da enzima.

Percebe-se que uma característica por vez foi modificada. Foram modificados separadamente o anel furano, o ácido carboxílico e o anel pirrolidina. Além disso, o comprimento da cadeia carbônica entre o furano e a pirrolidina, assim como a posição na substituição no anel

furano foram modificadas. A estrutura química das moléculas selecionadas é mostrada na Tabela 4.7:

Tabela 4.7 - Estrutura bidimensional das moléculas selecionadas na segunda geração para o sítio do G3P

NEQUIMED 350T		NEQUIMED 356T
NEQUIMED 351T		NEQUIMED 357T
NEQUIMED 352T		NEQUIMED 358T
NEQUIMED 353T		NEQUIMED 359T
NEQUIMED 354T		NEQUIMED 360T
NEQUIMED 355T		NEQUIMED 361T

4.1.5.1 Simulação por Dinâmica molecular da segunda geração de compostos

Foi feita a simulação por MD de nove dos compostos selecionados para a compra, também selecionando mais de uma pose predita na docagem molecular para a simulação. As

poses escolhidas foram aquelas de energia de interação próxima e conformação diferente do ligante no sítio ativo. Três das moléculas selecionadas para segunda geração (NEQUIMED 351T, 302T e 303T) não tiveram sua trajetória calculada por MD, devido ao fato de não terem sido produzidas poses para as mesmas na etapa de docagem molecular. Isso aconteceu provavelmente porque a molécula não conseguiu se acomodar no sítio de modo a realizar interações favoráveis e ainda cumprir as restrições farmacofóricas inseridas no modelo.

Ao passo do que foi observado para a os resultados de MD para os compostos de primeira geração, o padrão de RMSD para os átomos da proteína se manteve, independente do ligante complexado na sua estrutura. Para o caso do NAD^+ , notou-se que o mesmo apresentou um aumento de RMSD quando o ligante NEQUIMED 358T estava complexado no sítio. Novamente esse aumento de RMSD aconteceu logo após a perda da interação do grupo nicotinamida com os resíduos de aminoácidos Glu335 e Asn336.

Os resultados de energia média de interação ligante – enzima para cada modelo estão mostrados na Tabela 4.8.

Tabela 4.8 – Valor médio de energia de interação calculada computacionalmente por MMPBSA para as moléculas de segunda geração

Molécula	Modelo	Energia (kcal/mol)	Desvio padrão (kcal/mol)
NEQUIMED 358T	1	-29,74	3,78
NEQUIMED 357T	1	-19,04	6,49
	2	-17,95	3,23
NEQUIMED 354T	1	-17,56	3,26
	2	-14,79	3,70
NEQUIMED 359T	1	-16,52	3,12
NEQUIMED 350T	2	-15,90	2,98
	1	-12,69	3,49
NEQUIMED 360T	1	-15,88	5,73
NEQUIMED 356T	1	-12,26	2,53
NEQUIMED 361T	1	-10,17	3,52
NEQUIMED 355T	1	-1,40	3,23

Nota-se que a molécula mais bem classificada possui um valor de energia significativamente mais negativo, quando comparado com as outras moléculas. Isso sugere que a complementaridade da mesma com o sítio bastante superior às outras moléculas testadas *in silico*. As três moléculas subsequentes têm energia bastante similar à energia obtida para a molécula original. Mesmo as moléculas com energia mais alta possuem valores mais favoráveis às outras moléculas selecionadas na primeira geração. A exceção recai na molécula pior classificada, a qual possui valor de energia próximo de zero. Essa molécula perde toda a interação com o sítio da proteína, sugerindo que a modificação realizada na molécula levou à desestabilização da interação que estabilizava a molécula no interior do sítio da proteína.

4.1.5.1.1 Molécula NEQUIMED 358T-1

A molécula NEQUIMED 358T foi escolhida para avaliar o efeito da posição de substituição no anel furano do fragmento selecionado na primeira série. Inicialmente, a substituição estava nas posições 2 e 5 do anel, sendo modificada para a posição 2 e 3 do anel furano, para o NEQUIMED 358T. A presença do grupo metila na posição 5 foi simplesmente por questões de disponibilidade comercial.

Essa modificação levou a uma diminuição bastante pronunciada na energia de interação ligante-enzima, calculado por MMGBSA. Nota-se que a MD levou a uma modificação considerável na conformação da proteína e na pose tanto do NAD^+ quanto do ligante. O ligante se afastou da região definida com restrição para docagem molecular e se posicionou bem próximo ao resíduo de aminoácido Asp210. Essa modificação ocorreu durante o SA, quando o RMSD do ligante deu um salto de 2 a 6 Å. Figura 4.19.a Neste mesmo momento, o fragmento perde interação com a Ser247 e ganha interações com o Asp210, Ser224 e Arg249. Figura 4.19.b

Figura 4.19 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 358T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD⁺, em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 358T-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em verde) e Ser224 (em ciano).

Como já notado para o fragmento de primeira geração, o nitrogênio com hidrogênio disponível para ligação de hidrogênio mantém a molécula interagindo com o Asp210 e com a Ser224, o que faz com que a interação de hidrogênio intermolecular entre esses dois resíduos de aminoácidos se rompa e uma nova cavidade seja aberta. Esse fenômeno provavelmente ocorreu com a interação do nitrogênio do grupo pirrolidina com a Ser224, já no início da simulação. No caso dessa molécula, além desse efeito, foi observada também uma modificação na posição da Arg249. Essa modificação levou a uma diminuição de volume disponível para interação com a pequena molécula na região descrita como o sítio de ligação do G3P e proporcionou uma grande estabilidade do ligante, já que o mesmo faz pelo menos duas ligações de hidrogênio, por meio do grupo carboxílico, com este resíduo de aminoácido.

Considerando toda a molécula, percebe-se que todos os átomos capazes de realizar ligações de hidrogênio estão fazendo pelo menos uma com resíduos de aminoácidos da proteína. Além disso, observa-se uma boa complementaridade estrutural do ligante com a proteína. Todos esses fatores ajudam a explicar o valor de energia de interação teórica obtida.

A modificação no posicionamento do NAD⁺ está relacionada à modificação do sítio ativo causada pela presença do ligante. Nota-se que o grupo nicotinamida do NAD⁺ perde a sua interação com os resíduos de aminoácidos Glu336 e Asn335 logo após a mudança brusca de posicionamento do fragmento. Uma vez que houve essa perda de interação, o grupo nicotinamida passou a se afastar cada vez mais de seu sítio de interação, como visto na Figura 4.20. Provavelmente isso se deve à modificação na Arg249, ocorrida para estabilizar a posição do

ligante.

Figura 4.20 - Conformação e posicionamento da molécula NEQUIMED 358T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.5.1.2 Molécula NEQUIMED 357T-1

Esse fragmento apresentou uma interação mais forte com a enzima quando comparada com o fragmento selecionado na primeira geração de moléculas. Porém esse aumento pode estar simplesmente relacionado ao aumento na massa molecular da molécula. A principal modificação nessa molécula foi a inserção de um anel após a pirrolidina, com o intuito de verificar o comportamento do sistema quando se insere um grupo volumoso naquela posição. Além disso, por motivos de disponibilidade comercial, trocou-se o grupo carboxílico por um metil-éster. O método de docagem molecular foi incapaz de reproduzir a pose relativa dos grupos obtida para o ligante mais ativo, o que já era esperado pela falta de espaço no sítio da proteína para acomodar um grupo volumoso naquela posição. Percebe-se na simulação por MD que o ligante alarga a cavidade acima da cisteína catalítica e então se acomoda inteiramente nessa região, como mostrado na Figura 4.21.

Figura 4.21 - Conformação e posicionamento da molécula NEQUIMED 357T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação

Apesar do método de docagem molecular ter posicionado o grupo metil-éster nas proximidades dos resíduos de aminoácido selecionados como restrição, a MD foi capaz de retirá-lo desta posição sugerindo que este grupamento não é eficiente em manter a interação com a cavidade de interação do G3P. Mesmo assim, o ligante não aparenta estar estável no sítio da enzima, como se vê na Figura 4.22.

Figura 4.22 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 357T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD⁺, em vermelho. b) Gráfico de energia de interação da molécula NEQUIMED 357T-1 calculado por MMGBSA para cada passo da simulação, após a etapa de arrefecimento simulado.

Também se observa que a energia de interação do ligante com a enzima não é estável durante a dinâmica molecular, inclusive atingindo valores bastante elevados, próximos a zero. Desta forma, com esse resultado, é mais provável que o ligante seja inativo em experimentos enzimáticos *in vitro*.

4.1.5.1.3 Molécula NEQUIMED 354T-1

Essa molécula apresentou uma energia de interação calculada por MMPBSA ligeiramente inferior à molécula mais ativa. Como se trata de uma molécula com quatro átomos a menos (substituição do anel pirrolidina por uma amina), espera-se que essa molécula tenha uma eficiência de ligante superior à molécula original. Essa modificação nos sugere que o anel não tem grande importância para a interação do ligante com o alvo e o que realmente faz o reconhecimento molecular naquela região é o átomo de nitrogênio com um hidrogênio ligado.

A molécula permaneceu praticamente na mesma posição durante toda a MD (Figura 4.24). Foi observado pelo gráfico de RMSD do ligante em função do tempo que o fragmento saiu da posição próxima à inicial por volta de 0,5 ns após o início da simulação, chegando a um RMSD de 8 Å. (Figura 4.23)

Figura 4.23 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 354T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD⁺, em vermelho. b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 354T-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em vermelho)

Durante esse tempo, a molécula perdeu todas as ligações de hidrogênio que fazia com a Ser247 e a Arg249. Porém, durante a etapa de resfriamento do arrefecimento simulado, a molécula voltou para a uma conformação bastante próxima da original, com um RMSD de 2 Å. Logo quando o fragmento iniciava a sua volta para a posição original, percebeu-se a formação de ligação de hidrogênio com o resíduo de aminoácido Asp210, o que pode ter contribuído para a abertura da cavidade. Após isso, o fragmento se fixou na posição obtida pelo método de docagem molecular e reestabeleceu as ligações de hidrogênio com a Ser247 e Arg249. Esse pequeno RMSD é explicado por uma pequena modificação conformacional da molécula de modo a entrar

na cavidade entre os resíduos de aminoácidos Ser224 e Asp210.

Figura 4.24 - Conformação e posicionamento da molécula NEQUIMED 354T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.5.1.4 Molécula NEQUIMED 359T-1

Com a mudança do grupo pirrolidina por um grupo fenil, a docagem molecular foi incapaz de posicionar o anel na mesma cavidade aonde se encontra o anel pirrolidina da molécula de primeira geração, provavelmente devido a efeito de impedimento estereoquímico. Porém a MD, com a flexibilização dos átomos da enzima, relaxou o sítio de forma a possibilitar a acomodação do anel aromático nas proximidades do resíduo de aminoácido Asp210, acomodação essa que ocorreu de forma suave. Figura 4.25

Figura 4.25 - Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 359T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD^+ , em vermelho. b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 359T-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em azul), Ser224 (em vermelho) e Thr167 (em ciano)

Enquanto a molécula permaneceu próxima à pose obtida por docagem molecular, a carboxila fez interações de hidrogênio com Ser247, Ser224 e Arg249, de forma alternada. Após a troca de posição, o ácido carboxílico se posicionou um pouco mais para cima da posição original, perdendo todas as interações anteriores e formando ligações de hidrogênio somente com a Thr167.

Observa-se que a modificação crucial para possibilitar a mudança de conformação do ligante foi a modificação na posição da Arg249, representada como uma saliência azul à esquerda da cisteína catalítica, representada em amarelo na Figura 4.26

Figura 4.26 - Conformação e posicionamento da molécula NEQUIMED 359T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação

Nota-se, nessa simulação, que devido à ausência de nitrogênio protonado nas proximidades dos resíduos de aminoácidos Asp210 e Ser224, a cavidade se manteve fechada por toda a MD. Apesar de toda a modificação do ligante, o NAD⁺ permanece na mesma posição, o que indica que a presença do ligante e a modificação conformacional da proteína para acomodá-lo não influenciaram na estabilidade do cofator.

4.1.5.1.5 Molécula NEQUIMED 350T,-1

A modificação realizada nessa molécula foi a mudança no tamanho do anel e a adição de um átomo de nitrogênio no mesmo. Como esperado, a docagem molecular conseguiu posicionar o ligante na mesma posição do ligante original. A simulação mostra a tendência do ligante em abrir uma cavidade lateral, comportamento já observado em uma simulação anterior para o composto NEQUIMED 294, da primeira geração. Pelo visto, a adição do nitrogênio na extremidade da molécula ajudou a abertura da cavidade, mas isso aconteceu sem qualquer melhora na energia.

O fragmento sofreu a mudança de pose, entrando na cavidade entre o Asp210 e Ser224 durante a etapa de SA, como mostra a Figura 4.27.

Figura 4.27 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 350-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD^+ , em vermelho. b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 350-1 com os resíduos de aminoácidos Ser247 (em preto), Arg249 (em ciano) e Ser224 (em azul).

Até esse instante, a molécula fazia interação de hidrogênio somente com a Ser247, por meio da carboxila, e com a Arg249, por meio do anel furano. Após isso, a molécula passou a fazer quatro interações de hidrogênio, uma com o Asp210, uma com a Ser224, ambas por meio do anel piperazina e duas com a Arg249 por meio do ácido carboxílico (Figura 4.28). Essas ligações contribuíram para estabilizar o complexo e tornar a energia de interação mais favorável, estável a -16 kJ.mol^{-1} . Ao se verificar o tempo de formação das ligações de hidrogênio com esses resíduos de aminoácidos, percebe-se que a molécula começou a fazer ligação de hidrogênio com a Ser224 ao mesmo tempo que se formava a segunda ligação com a Arg249. Somente depois que essas ligações estavam estabelecidas é que o Asp210 começou a fazer interação com a molécula. Esse fato pode ser explicado pela manutenção da molécula, por meio das ligações de hidrogênio com a Ser224 e Arg249, em uma posição que favoreça a interação da mesma com o Asp210. Além disso, já foi observado anteriormente que na ausência de qualquer molécula na região entre o Asp210 e a Ser224, estes dois resíduos de aminoácido estão interagindo entre si por meio de ligação de hidrogênio. A interação do fragmento com a Ser224 rompe essa interação entre os resíduos de aminoácido e possibilita que o Asp210, localizado um pouco mais longe do ligante, também esteja disponível para essa interação, por meio de mudança conformacional.

Nos ensaios experimentais, é esperado Ki da mesma ordem de grandeza da molécula original, talvez um pouco maior, devido à semelhança no valor de energia obtido por MMPBSA. Novamente o cofator NAD^+ se manteve na mesma posição durante toda a dinâmica.

Figura 4.28 - Conformação e posicionamento da molécula NEQUIMED 350-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.5.1.6 Molécula NEQUIMED 360T-1

A troca do grupo pirrolidina por um grupo triazol levou a uma inversão na pose predita pelo método de docagem molecular. Esse método posicionou o anel na posição onde o ácido carboxílico do fragmento de primeira geração se encontrava. Esse novo posicionamento já era esperado, já que os triazóis são conhecidos como bioisósteros de carboxilas. Já o anel furano foi posicionado em frente à cisteína catalítica, como pode ser visto na Figura 4.29.

Figura 4.29 - Conformação e posicionamento da molécula NEQUIMED 360T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

Durante toda a MD, o anel triazol se manteve na mesma posição, com leve rotação para otimizar as interações de hidrogênio com a cavidade altamente hidrofílica onde se localiza, com a Thr167 e principalmente com a Ser247, com a qual fez ligações de hidrogênio durante os 10 ns de simulação. A força de interação do anel triazol para essa pose possivelmente diminuiu graus de

liberdade da molécula, o que levou à manutenção do anel furano próximo à cisteína catalítica, com poucas mudanças conformacionais durante todo o tempo de MD. Figura 4.30

Figura 4.30 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 360T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD^+ , em vermelho. b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 360T-1 com os resíduos de aminoácidos Ser247 (em preto) e Thr167 (em vermelho).

O grupo ácido carboxílico tenta se aproximar dos resíduos de aminoácidos Ser224 e Asp210, porém está limitado pela conformação adotada pelo anel triazol. Durante grande parte da simulação, os átomos de oxigênio do furano e da carboxila mantinham ligação de hidrogênio com a Arg249. Esta ligação de hidrogênio deixa a molécula tensionada e pode levar a desestabilização da molécula no sítio da enzima.

O fato de a energia de interação predita pelo MMGBSA ser ligeiramente menor do que o predito para o fragmento ativo possivelmente se deve à falta de interação do ligante com os resíduos de aminoácidos Ser224 e Asp210, como acontece para o fragmento de primeira geração por meio do anel pirrolidina. Essa ausência pode explicar a instabilidade na energia de interação observada para esse ligante.

4.1.5.1.7 Molécula NEQUIMED 356T

A modificação realizada nessa molécula foi a troca do grupo pirrolidina por um grupo aminofenil. Além disso, por motivos de disponibilidade comercial, o grupo carboxílico foi substituído por um grupo metil-éster. A docagem molecular mais uma vez não conseguiu posicionar a molécula da mesma forma que o fragmento da primeira geração está posicionado. O ligante permaneceu estável nessa conformação até o SA, quando o grupo aminofenil saiu da cavidade acima da cisteína catalítica. Apesar de todo o movimento do aminofenil, o metiléster

manteve durante toda a simulação a ligação de hidrogênio com a Ser247. Figura 4.31

Figura 4.31 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 356T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD^+ , em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 356T-1 com o resíduo de aminoácido Ser247.

Observa-se também uma modificação na posição do NAD^+ , porém bem suave. O mesmo mantém a interação com os resíduos de aminoácidos Glu335 e Asn336 e talvez por essa razão não sofreu grandes modificações conformacionais. O que se nota é que o anel de adenina modificou a sua posição no sítio, mas aparentemente essa alteração não é capaz de desestabilizar a interação do cofator com o sítio da enzima. Figura 4.32

A diferença entre a energia predita para o NEQUIMED 356T e para o fragmento NEQUIMED 293 pode ser devido à troca dos grupos ou simplesmente ao erro de cálculo de energia pelo método de MMPBSA, que costuma ser da ordem de 4 kcal/mol. De qualquer forma, é esperado que este ligante tenha um K_i experimental da mesma ordem de grandeza que a molécula original.

Figura 4.32 - Conformação e posicionamento da molécula NEQUIMED 356T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.5.1.8 Molécula NEQUIMED 361T

Nesta molécula, houve a troca do grupo furano por um fenil. Com essa modificação na estrutura, o método de docagem molecular foi capaz de posicionar a molécula na mesma posição que a molécula mais ativa da primeira geração. Nota-se que a simulação por MD novamente levou à abertura da cavidade lateral, provavelmente devido à presença do nitrogênio protonado. A molécula não é muito estável no sítio da enzima, tendo alcançado durante a MD quatro patamares de estabilidade, como mostrado na Figura 4.33

Figura 4.33 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 361T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD⁺, em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 361T-1 com os resíduos de aminoácidos Ser247 (em preto), Asp210 (em Pink) e Arg249 (em ciano)

Durante toda a simulação, o fragmento mantém a interação de hidrogênio com o Asp210 e

Arg249 (Figura 4.33). Já a sua interação com a Ser247 se mantém até a etapa de arrefecimento simulado, deixando de existir após isso. É durante essa etapa que a molécula começa a entrar na cavidade entre os resíduos de aminoácidos Ser224 e Asp210. Pelo fato de o anel aromático apresentar substituições na posição *para*, a movimentação do fragmento para o interior da cavidade aberta necessariamente desloca o ácido carboxílico para longe da cavidade de ligação do G3P. Apesar desse deslocamento, somente por 1 ns durante a simulação, o mesmo perde interação com a Arg249. Após a última mudança de patamar, a molécula perde interação com todos os resíduos de aminoácidos acima citados, atingindo a pose mostrada na Figura 4.34.

Energeticamente, essa molécula é consideravelmente inferior que a molécula de origem, o que sugere que a retirada do grupo furano ou a modificação da posição relativa entre os grupos carboxílico e pirrolidina, têm uma influência negativa na interação do fragmento com a enzima.

Figura 4.34 - Conformação e posicionamento da molécula NEQUIMED 361T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.5.1.9 Molécula NEQUIMED 355T-1

Esse ligante foi selecionado para verificar a importância do grupo carboxílico na molécula. Para isso, esse grupamento foi substituído por um grupo metil-éster. A docagem molecular conseguiu posicionar a molécula em uma pose bastante similar à obtida para o fragmento de primeira geração. Porém a MD indica que o NEQUIMED -305 não é energeticamente estável no sítio da enzima. Durante a simulação a molécula saiu do sítio, realizando interações com a superfície da biomacromolécula Figura 4.36.

Figura 4.35 - a) Gráfico de RMSD para simulação por MD da enzima TcGAPDH complexada com seu cofator e com a molécula NEQUIMED 355T-1, levando em conta a posição dos átomos do ligante, em preto, e do NAD⁺, em vermelho b) Gráfico de número de ligações de hidrogênio realizadas pelo NEQUIMED 355T-1 com o resíduo de aminoácido Asp210.

Essa saída pode ter sido ocasionada pela perda de interação do metil-éster com a enzima (não foi detectada nenhuma interação do ligante com a Ser247 ou Thr167) (Figura 4.35) e logo, o aumento da energia cinética do fragmento, o que ocasionou na perda de interação do anel pirrolidina com os resíduos de aminoácido Asp210, o que aconteceu logo após o SA.

Figura 4.36 - Conformação e posicionamento da molécula NEQUIMED 355T-1 e do cofator NAD⁺ no sítio ativo da TcGAPDH a) ao início da simulação por MD b) ao final da simulação.

4.1.6 Análise final

Pelos estudos de seleção de moléculas *in silico* e ensaios enzimáticos *in vitro* por espectroscopia de fluorescência e simulação por MD para as moléculas de primeira geração,

percebe-se que:

- O sítio ativo da TcGAPDH possui, *in silico*, uma preferência por ligantes com cLogP próximo a zero, com grande número de átomos capazes de aceitar ligações de hidrogênio e com elevada área superficial polar, nos limites da Regra dos 3 para fragmentos-similares;
- A técnica de espectroscopia de fluorescência é capaz, para a enzima TcGAPDH, de detectar atividade enzimática de compostos com característica de fragmentos;
- A concepção de fragmentos moleculares é útil para identificação de compostos de alta eficiência de ligante ($LE > 0,3$) mesmo para sítio ativo bastante amplo como é o caso da TcGAPDH;
- A simulação por MD foi capaz de identificar o composto mais potente observado nos ensaios enzimáticos;
- O possível modo de interação do fragmento mais ativo envolve a abertura de uma nova cavidade na enzima TcGAPDH, por meio de uma interação de hidrogênio com o Asp210;
- O Asp210 pode servir como um ponto de seletividade em relação à enzima humana, já que nesta, o resíduo de aminoácido que se encontra nessa posição é uma Leucina.

Pelos estudos teóricos baseados em informações estruturais e energéticas obtidas por meio de simulação por MD, é possível inferir que:

a. Para os compostos de segunda geração:

- a posição de substituição no anel central do fragmento ativo está fortemente relacionada à estabilidade da molécula no sítio e à sua energia de interação com a mesma;
 - a substituição em *para* em um benzeno leva a uma perda na energia de interação do ligante com a estrutura proteica;
 - a substituição 2,3 em um anel furano melhora consideravelmente a estabilidade da molécula no sítio e também sua energia de interação com a enzima;

- uma extensão da molécula após o anel pirrolidina leva a uma diminuição na atividade inibitória do composto ou modificação do possível modo de interação da molécula;
 - modificação ou retirada do anel da pirrolidina com a manutenção do átomo de nitrogênio na mesma posição não leva a uma perda da atividade;
 - grupo carboxílico aparenta ser importante para a estabilização da molécula no sítio ativo da proteína;
- b. Para o comportamento do NAD^+ e complexo proteico durante uma simulação por MD na presença de diferentes ligantes:
- a presença dos ligantes e do NAD^+ no sítio da proteína não causa grandes modificações conformacionais na cadeia principal da enzima, ao menos durante o tempo de simulação realizado;
 - a estabilidade do NAD^+ no sítio está provavelmente relacionada à capacidade do grupo nicotinamida em interagir com os resíduos de aminoácidos Asn335 e Glu336. Perda nessa ligação leva a grande variação na posição do NAD^+ no sítio da proteína;
 - a saída do grupo adenina do NAD^+ de sua cavidade de interação exerce pouca influência na estabilidade da molécula no sítio.

5 Conclusões

Os resultados obtidos por este trabalho mostram os benefícios de se utilizar a técnica de fragmentos moleculares em conjunto com métodos de química computacional, como docagem molecular, similaridade química e simulação por dinâmica molecular, para seleção de moléculas e predição de modo de interação com o sítio ativo da enzima, e com espectroscopia de fluorescência para determinação de K_i de pequenas moléculas frente a um alvo biomacromolecular. Por meio do uso conjunto dessas técnicas, foi possível selecionar dentre um banco de dados de moléculas de mais de 500 mil compostos, cinco moléculas para ensaios enzimáticos *in vitro* e encontrar uma molécula ativa. O alto índice de acerto entre moléculas testadas e moléculas ativas pode ser creditado ao uso integrado de técnicas baseadas na estrutura do ligante e na estrutura do alvo macromolecular e à extensa e rigorosa inspeção visual das moléculas mais bem classificadas pela etapa consensual para seleção de moléculas para compra.

O fragmento ativo possuía uma baixa massa molecular e K_i elevado, de 425,1 μM , típicos de moléculas com características fragmento-similares, mas com uma elevada eficiência do ligante, igual a 0,33. Essa eficiência do ligante, caso seja mantida até o final do planejamento molecular, pode levar um composto-matriz, com massa molecular já próximo de 500 Da, com K_i da ordem de baixo nanomolar. Deste modo, a abordagem de fragmentos moleculares tem a vantagem de iniciar o desenvolvimento com moléculas de baixa massa molecular e baixa complexidade química, o que possibilita o planejamento da molécula levando em consideração parâmetros farmacocinéticos e farmacodinâmicos desde o princípio, reduzindo a chance de encontrar efeitos indesejados, como por exemplo toxidez e absorção inadequada, com a molécula em fases mais avançadas da gênese.

Além disso, a ferramenta de simulação por Dinâmica Molecular mostrou-se bastante útil para predizer a pose de interação do fragmento no sítio ativo da enzima, conseguindo, inclusive, classificar o fragmento ativo na primeira posição na classificação de energia de interação, com uma energia de interação consideravelmente mais favorável em relação ao fragmento classificado na segunda posição. A simulação por dinâmica molecular permitiu a observação de uma interação preferencial de nitrogênio capaz de atuar como doador de ligações de hidrogênio com os resíduos de aminoácidos Ser224 e Asp210, como acontece para a molécula NEQUIMED 293. Na ausência de um grupo com essa característica, a Ser224 e o Asp210 apresentam uma interação de

hidrogênio entre si. Uma vez o fragmento com um nitrogênio protonado se aproxima dessa região, o Asp210 interage preferencialmente com esse grupo e abre uma nova cavidade na enzima. Uma comparação com a enzima humana sugere que esse pode ser um ponto de seletividade entre as duas enzimas, já que a GAPDH de humanos possui no lugar do aspartato, uma leucina, resíduo de aminoácido incapaz de realizar ligações de hidrogênio.

A realização de uma relação estrutura-atividade teórica por meio de simulação por Dinâmica Molecular para as moléculas de segunda geração mostrou que a posição de substituição do anel central do fragmento ativo pode tanto aumentar quanto diminuir a energia de interação do fragmento com a enzima. Além disso, foi possível notar que a substituição da pirrolidina por um grupo mais volumoso leva à diminuição da energia de interação ou modificação do modo de interação do fragmento com o sítio ativo da enzima, indicando que a enzima não suporta um grupo mais volumoso naquela posição. Também se observou que o que de fato importa no anel pirrolidina é o nitrogênio capaz de doar ligações de hidrogênio, já que uma molécula sem o anel mas com o nitrogênio teve uma energia de interação predita similar à molécula ativa.

Desta forma, a simulação por MD aparece como uma importante técnica computacional para o planejamento de moléculas bioativas, capaz de prever energia de interação relativa entre as moléculas e também prever a posição de interação do ligante no interior do sítio de ligação do alvo macromolecular.

6 Perspectivas

Uma vez identificado um fragmento ativo e com boa eficiência de ligante contra a TcGAPDH e feito o estudo de predição de modo de interação da molécula com a enzima e de relação estrutura-atividade para o mesmo, a próxima etapa do trabalho é validar experimentalmente essas informações.

Inicialmente, faz-se necessário medir o K_i da molécula NEQUIMED 293 em diferentes concentrações da mesma (700 e 300 μM) de modo a confirmar o seu valor de K_i e obter o modo de inibição que a molécula causa na TcGAPDH (competitivo, incompetitivo, não-competitivo). Após isso, faz-se necessário o ensaio enzimático *in vitro* das moléculas de segunda geração para validação da relação estrutura-atividade predita por meio da simulação por Dinâmica Molecular. A molécula que possuir o melhor K_i passará à etapa de crescimento do fragmento de modo a ocupar os sítios adjacentes e aumentar a potência, sem perder eficiência de ligante.

De modo a confirmar a predição da pose obtida pela simulação por Dinâmica Molecular, a cristalografia por Raios-X será utilizada para determinação da posição do fragmento e suas interações com o alvo. Trabalho este que já foi iniciado por outro membro do grupo.

Paralelamente a isto, ensaios enzimáticos *in vitro* com as moléculas selecionadas para o sítio de interação do cofator devem ser realizados na busca de fragmento com elevada eficiência de ligante em uma cavidade próxima, para uma tentativa de construir um composto-matriz por ligação de fragmentos.

m

Referências

1. NEVES, D. P. **Parasitologia humana**. 10. ed. São Paulo: Editora Atheneu, 2000.
2. SCHMUNIS, G. A. Epidemiology of Chagas disease in non-endemic countries: The role of international migration. **Memorias do Instituto Oswaldo Cruz**, v. 102, n. SUPPL. 1, p. 75-85, 2007.
3. PARKER, E. R.; SETHI, A. Chagas disease: Coming to a place near you. **Dermatologic Clinics**, v. 29, n. 1, p. 53-62, 2011.
4. PÉREZ-LÓPEZ, F. R.; CHEDRAUI, P. Chagas disease in pregnancy: A non-endemic problem in a globalized world. **Archives of Gynecology and Obstetrics**, v. 282, n. 6, p. 595-599, 2010.
5. COURA, J. R.; BORGES-PEREIRA, J. Chagas disease: 100 years after its discovery. A systemic review. **Acta Tropica**, v. 115, n. 1-2, p. 5-13, 2010.
6. KIERSZENBAUM, F. Where do we stand on the autoimmunity hypothesis of Chagas disease? **Trends in Parasitology**, v. 21, n. 11, p. 513-516, 2005.
7. WORLD HEALTH ORGANIZATION: Chagas Disease (American trypanosomiasis). Disponível em:
<<http://www.who.int/mediacentre/factsheets/fs340/en/index.html>>. Acesso em: 2011, 15 de dez.
8. COURA, J. R.; VÍAS, P. A. Chagas disease: A new worldwide challenge. **Nature**, v. 465, n. 7301 SUPPL., p. S6-S7, 2010.
9. CHIRAC, P.; TORREELE, E. Global framework on essential health R&D. **Lancet**, v. 367, n. 9522, p. 1560-1561, 2006.
10. HUPPATZ, C.; DURRHEIM, D. N. Control of neglected tropical diseases [10]. **New England Journal of Medicine**, v. 357, n. 23, p. 2407, 2007.
11. COURA, J. R. Present situation and new strategies for chagas disease chemotherapy - A proposal. **Memorias do Instituto Oswaldo Cruz**, v. 104, n. 4, p. 549-554, 2009.

-
12. CANÇADO, J. R. Criteria of Chagas disease cure. **Memorias do Instituto Oswaldo Cruz**, v. 94, n. SUPPL. 1, p. 331-335, 1999.
13. CANÇADO, J. R. Long term evaluation of etiological treatment of Chagas disease with benznidazole. **Revista do Instituto de Medicina Tropical de Sao Paulo**, v. 44, n. 1, p. 29-37, 2002.
14. URBINA, J. A. Specific chemotherapy of Chagas disease: Relevance, current limitations and new approaches. **Acta Tropica**, v. 115, n. 1-2, p. 55-68, 2010.
15. MAYA, J. D.; CASSELS, B. K.; ITURRIAGA-VÁSQUEZ, P.; FERREIRA, J.; FAÚNDEZ, M.; GALANTI, N.; FERREIRA, A.; MORELLO, A. Mode of action of natural and synthetic drugs against *Trypanosoma cruzi* and their interaction with the mammalian host. **Comparative Biochemistry and Physiology – A Molecular and Integrative Physiology**, v. 146, n. 4, p. 601-620, 2007.
16. DIAZ DE TORANZO, E. G.; CASTRO, J. A.; FRANKE DE CAZZULO, B. M.; CAZZULO, J. J. Interaction of benznidazole reactive metabolites with nuclear and kinetoplastic DNA, proteins and lipids from *Trypanosoma cruzi*. **Experientia**, v. 44, n. 10, p. 880-881, 1988.
17. KUBATA, B. K.; KABUTUTU, Z.; NOZAKI, T.; MUNDAY, C. J.; FUKUZUMI, S.; OHKUBO, K.; LAZARUS, M.; MARUYAMA, T.; MARTIN, S. K.; DUSZENKO, M.; URADE, Y. A key role for old yellow enzyme in the metabolism of drugs by *Trypanosoma cruzi*. **Journal of Experimental Medicine**, v. 196, n. 9, p. 1241-1251, 2002.
18. BUCKNER, F. S.; NAVABI, N. Advances in Chagas disease drug development: 2009-2010. **Current Opinion in Infectious Diseases**, v. 23, n. 6, p. 609-616, 2010.
19. RENSLO, A. R.; MCKERROW, J. H. Drug discovery and development for neglected parasitic diseases. **Nature Chemical Biology**, v. 2, n. 12, p. 701-710, 2006.
20. MCKERROW, J. H.; DOYLE, P. S.; ENGEL, J. C.; PODUST, L. M.; ROBERTSON, S. A.; FERREIRA, R.; SAXTON, T.; ARKIN, M.; KERR, I. D.; BRINEN, L. S.; CRAIK, C. S. Two approaches to discovering and developing new drugs for Chagas disease. **Memorias do Instituto Oswaldo Cruz**, v. 104, n. SUPPL. 1, p. 263-269, 2009.
21. LESLIE, M. Drug developers finally take aim at a neglected Disease. **Science**, v. 333, n. 6045, p. 933-935, 2011.

-
22. R&D execs paint bleak picture of industry productivity: \$1.4 bil. per NME. **The Pink Sheet**, v. 6, n. 18 August, 2003.
23. WERMUTH, C. G. **The practice of medicinal chemistry**. 3. ed. London: Academic Press, 2008.
24. LISTON, H. L.; MARKOWITZ, J. S.; DEVANE, C. L. Drug glucuronidation in clinical psychopharmacology. **Journal of Clinical Psychopharmacology**, v. 21, n. 5, p. 500-515, 2001.
25. MONTANARI, C. A. **Química medicinal: Metodos e fundamentos em planejamento de fármacos**. 1. ed. São Paulo: EdUSP, 2011.
26. LOMBARDINO, J. G.; LOWE III, J. A. The role of the medicinal chemist in drug discovery – Then and now. **Nature Reviews Drug Discovery**, v. 3, n. 10, p. 853-862, 2004.
27. KNOWLES, J.; GROMO, G. Target selection in drug discovery. **Nature Reviews Drug Discovery**, v. 2, n. 1, p. 63-69, 2003.
28. KROGSGAARD-LARSEN, P.; PELLICIARI, R.; DE SOUZA, N.; TIMMERMAN, H.; TRIGGLE, D. J.; VAN BOECKEL, C. A. A.; WASLEY, J. Medicinal chemistry education: What is needed and where is it going? **Drug Development Research**, v. 66, n. 1, p. 1-8, 2005.
29. BLEICHER, K. H.; BÖHM, H. J.; MÜLLER, K.; ALANINE, A. I. Hit and lead generation: Beyond high-throughput screening. **Nature Reviews Drug Discovery**, v. 2, n. 5, p. 369-378, 2003.
30. DREWS, J. Drug discovery today – and tomorrow. **Drug Discovery Today**, v. 5, n. 1, p. 2-4, 2000.
31. HOPKINS, A. L.; GROOM, C. R. The druggable genome. **Nature Reviews Drug Discovery**, v. 1, n. 9, p. 727-730, 2002.
32. KESERU, G. M.; MAKARA, G. M. Hit discovery and hit-to-lead approaches. **Drug Discovery Today**, v. 11, n. 15-16, p. 741-748, 2006.
33. TAMIMI, N. A. M.; ELLIS, P. Drug development: From concept to marketing! **Nephron – Clinical Practice**, v. 113, n. 3, p. c125-c131, 2009.

-
34. VERLINDE, C. L. M. J.; HANNAERT, V.; BLONSKI, C.; WILLSON, M.; PÉRIÉ, J. J.; FOTHERGILL-GILMORE, L. A.; OPPERDOES, F. R.; GELB, M. H.; HOL, W. G. J.; MICHELS, P. A. M. Glycolysis as a target for the design of new anti-trypanosome drugs. **Drug Resistance Updates**, v. 4, n. 1, p. 50-65, 2001.
35. COURA, J. R.; DE CASTRO, S. L. A critical review on chagas disease chemotherapy. **Memorias do Instituto Oswaldo Cruz**, v. 97, n. 1, p. 3-24, 2002.
36. OPPERDOES, F. R. Compartmentation of carbohydrate metabolism in trypanosomes. **Annual Review of Microbiology**, v. 41, p. 127-151, 1987.
37. ENGEL, J. C.; FRANKE DE CAZZULO, B. M.; STOPPANI, A. O. M.; CANNATA, J. J. B.; CAZZULO, J. J. Aerobic glucose fermentation by *Trypanosoma cruzi* axenic culture amastigote-like forms during growth and differentiation to epimastigotes. **Molecular and Biochemical Parasitology**, v. 26, n. 1-2, p. 1-10, 1987.
38. TIELENS, A. G. M.; VAN HELLEMOND, J. J. Differences in energy metabolism between Trypanosomatidae. **Parasitology Today**, v. 14, n. 7, p. 265-271, 1998.
39. TIELENS, A. G. M.; VAN HELLEMOND, J. J. Surprising variety in energy metabolism within Trypanosomatidae. **Trends in Parasitology**, v. 25, n. 10, p. 482-490, 2009.
40. SILBER, A. M.; TONELLI, R. R.; LOPES, C. G.; CUNHA-E-SILVA, N.; TORRECILHAS, A. C. T.; SCHUMACHER, R. I.; COLLI, W.; ALVES, M. J. M. Glucose uptake in the mammalian stages of *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 168, n. 1, p. 102-108, 2009.
41. BAKKER, B. M.; MICHELS, P. A. M.; OPPERDOES, F. R.; WESTERHOFF, H. V. What controls glycolysis in bloodstream form *Trypanosoma brucei*? **Journal of Biological Chemistry**, v. 274, n. 21, p. 14551-14559, 1999.
42. FREITAS, R. F.; PROKOPCZYK, I. M.; ZOTTIS, A.; OLIVA, G.; ANDRICOPULO, A. D.; TREVISAN, M. T. S.; VILEGAS, W.; SILVA, M. G. V.; MONTANARI, C. A. Discovery of novel *Trypanosoma cruzi* glyceraldehyde-3-phosphate dehydrogenase inhibitors. **Bioorganic and Medicinal Chemistry**, v. 17, n. 6, p. 2476-2482, 2009.
43. ROSE, P. W.; BERAN, B.; BI, C.; BLUHM, W. F.; DIMITROPOULOS, D.; GOODSSELL,

- D. S.; PRLIĆ, A.; QUESADA, M.; QUINN, G. B.; WESTBROOK, J. D.; YOUNG, J.; YUKICH, B.; ZARDECKI, C.; BERMAN, H. M.; BOURNE, P. E. The RCSB Protein Data Bank: Redesigned web site and web services. **Nucleic Acids Research**, v. 39, n. SUPPL. 1, p. D392-D401, 2011.
44. MORAES, V. R. D. S.; TOMAZELA, D. M.; FERRACIN, R. J.; GARCIA, C. F.; SANNOMIYA, M.; SORIANO, M. D. P. C.; DA SILVA, M. F. D. G. F.; VIETRA, P. C.; FERNANDES, J. B.; RODRIGUES FILHO, E.; MAGALHÃES, E. G.; MAGALHÃES, A. F.; PIMENTA, E. F.; DE SOUZA, D. H. F.; OLIVA, G. Enzymatic inhibition studies of selected flavonoids and chemosystematic significance of polymethoxylated flavonoids and quinoline alkaloids in *Neoraputia* (Rutaceae). **Journal of the Brazilian Chemical Society**, v. 14, n. 3, p. 380-387, 2003.
45. DE MARCHI, A. A.; CASTILHO, M. S.; NASCIMENTO, P. G. B.; ARCHANJO, F. C.; DEL PONTE, G.; OLIVA, G.; PUPO, M. T. New 3-piperonylcoumarins as inhibitors of glycosomal glyceraldehyde-3- phosphate dehydrogenase (gGAPDH) from *Trypanosoma cruzi*. **Bioorganic and Medicinal Chemistry**, v. 12, n. 18, p. 4823-4833, 2004.
46. PEREIRA, J. M.; SEVERINO, R. P.; VIEIRA, P. C.; FERNANDES, J. B.; DA SILVA, M. F. G. F.; ZOTTIS, A.; ANDRICOPULO, A. D.; OLIVA, G.; CORRÊA, A. G. Anacardic acid derivatives as inhibitors of glyceraldehyde-3-phosphate dehydrogenase from *Trypanosoma cruzi*. **Bioorganic and Medicinal Chemistry**, v. 16, n. 19, p. 8889-8895, 2008.
47. DE MACEDO, E. M. S.; WIGGERS, H. J.; SILVA, M. G. V.; BRAZ-FILHO, R.; ANDRICOPULO, A. D.; MONTANARI, C. A. A new bianthron glycoside as inhibitor of *Trypanosoma cruzi* glyceraldehyde 3-phosphate dehydrogenase activity. **Journal of the Brazilian Chemical Society**, v. 20, n. 5, p. 947-953, 2009.
48. LADAME, S.; CASTILHO, M. S.; SILVA, C. H. T. P.; DENIER, C.; HANNAERT, V.; PÉRIÉ, J.; OLIVA, G.; WILLSON, M. Crystal structure of *Trypanosoma cruzi* glyceraldehyde-3-phosphate dehydrogenase complexed with an analogue of 1,3-bisphospho-D-glyceric acid: Selective inhibition by structure-based design. **European Journal of Biochemistry**, v. 270, n. 22, p. 4574-4586, 2003.
49. KENNEDY, K. J.; BRESSI, J. C.; GELB, M. H. A disubstituted NAD⁺ analogue is a nanomolar inhibitor of trypanosomal glyceraldehyde-3-phosphate dehydrogenase. **Bioorganic & Medicinal Chemistry Letters**, v. 11, n. 2, p. 95-98, 2001.

-
50. SOARES, F. A. *Ensaio bioquímico e celular para a identificação de novos inibidores de TcGAPDH seletivos à HsGAPDH* Dissertação (Mestrado em Química Orgânica) – Departamento de Química, Universidade Federal de São Carlos, São Carlos, 2011.
51. LADAME, S.; FAURÉ, R.; DENIER, C.; LAKHDAR-GHAZAL, F.; WILLSON, M. Selective inhibition of Trypanosoma cruzi GAPDH by "bi-substrate" analogues. **Organic and Biomolecular Chemistry**, v. 3, n. 11, p. 2070–2072, 2005.
52. FREITAS, R. F. *Integração de métodos em quiminformática e biocalorimetria para o planejamento de inibidores da enzima gliceraldeído-3-fosfato desidrogenase de Trypanosoma cruzi*. Tese (Doutorado em Físico-Química) – Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2009.
53. 2010 Drug Launches: a year of firsts offers hope for a rebound. **The Pink Sheet**, v. 11, n. 10, 2011.
54. TRIGGLE, D. J. The chemist as astronaut: Searching for biologically useful space in the chemical universe. **Biochemical Pharmacology**, v. 78, n. 3, p. 217–223, 2009.
55. BOHACEK, R. S.; MCMARTIN, C.; GUIDA, W. C. The art and practice of structure-based drug design: A molecular modeling perspective. **Medicinal Research Reviews**, v. 16, n. 1, p. 3–50, 1996.
56. VARNEK, A.; BASKIN, I. I. Chemoinformatics as a theoretical chemistry discipline. **Molecular Informatics**, v. 30, n. 1, p. 20–32, 2011.
57. LIPINSKI, C.; HOPKINS, A. Navigating chemical space for biology and medicine. **Nature**, v. 432, n. 7019, p. 855–861, 2004.
58. CAS: CHEMICAL ABSTRACT SERVICE. Disponível em: <<http://www.cas.org/>>. Acesso em: 20 de dez., 2011
59. JACOBY, E.; SCHUFFENHAUER, A.; POPOV, M.; AZZAQUI, K.; HAVILL, B.; SCHOPFER, U.; ENGELCH, C.; STANEK, J.; ACKLIN, P.; RIGOLLI, P.; STOLL, F.; KOCH, G.; MEIER, P.; ORAIN, D.; GIGER, R.; HINRICHS, J.; MALAGU, K.; ZIMMERMANN, J.; ROTH, H. J. Key aspects of the Novartis compound collection enhancement project for the compilation of a comprehensive chemogenomics drug discovery screening collection. **Current Topics in Medicinal Chemistry**, v. 5,

n. 4, p. 397-411, 2005.

60. WIGGERS, H. J.; ROCHA, J. R.; CHELESKI, J.; MONTANARI, C. A. Integration of ligand- and target-based virtual screening for the discovery of cruzain inhibitors. **Molecular Informatics**, v. 30, n. 6-7, p. 565-578, 2011.

61. HAJDUK, P. J.; GREER, J. A decade of fragment-based drug design: Strategic advances and lessons learned. **Nature Reviews Drug Discovery**, v. 6, n. 3, p. 211-219, 2007.

62. SHUKER, S. B.; HAJDUK, P. J.; MEADOWS, R. P.; FESIK, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. **Science**, v. 274, n. 5292, p. 1531-1534, 1996.

63. SUN, C.; PETROS, A. M.; HAJDUK, P. J. Fragment-based lead discovery: Challenges and opportunities. **Journal of Computer-Aided Molecular Design**, v. 25, n. 7, p. 607-610, 2011.

64. CHESSARI, G.; WOODHEAD, A. J. From fragment to clinical candidate—a historical perspective. **Drug Discovery Today**, v. 14, n. 13-14, p. 668-675, 2009.

65. ANDREWS, P. R.; CRAIK, D. J.; MARTIN, J. L. Functional group contributions to drug-receptor interactions. **Journal of Medicinal Chemistry**, v. 27, n. 12, p. 1648-1657, 1984.

66. KUNTZ, I. D.; CHEN, K.; SHARP, K. A.; KOLLMAN, P. A. The maximal affinity of ligands. **Proceedings of the National Academy of Sciences of the United States of America**, v. 96, n. 18, p. 9997-10002, 1999.

67. HOPKINS, A. L.; GROOM, C. R.; ALEX, A. Ligand efficiency: A useful metric for lead selection. **Drug Discovery Today**, v. 9, n. 10, p. 430-431, 2004.

68. HANN, M. M.; LEACH, A. R.; HARPER, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. **Journal of Chemical Information and Computer Sciences**, v. 41, n. 3, p. 856-864, 2001.

69. ZOETE, V.; GROSDIDIER, A.; MICHIELIN, O. Docking, virtual high throughput screening and in silico fragment-based drug design. **Journal of Cellular and Molecular Medicine**, v. 13, n. 2, p. 238-248, 2009.

-
70. LIPINSKI, C. A.; LOMBARDO, F.; DOMINY, B. W.; FEENEY, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews**, v. 23, n. 1-3, p. 3-25, 1997.
71. CONGREVE, M.; CARR, R.; MURRAY, C.; JHOTI, H. A 'Rule of Three' for fragment-based lead discovery? **Drug Discovery Today**, v. 8, n. 19, p. 876-877, 2003.
72. BRENK, R.; SCHIPANI, A.; JAMES, D.; KRASOWSKI, A.; GILBERT, I. H.; FREARSON, J.; WYATT, P. G. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. **ChemMedChem**, v. 3, n. 3, p. 435-444, 2008.
73. HUBBARD, R. E.; MURRAY, J. B. Experiences in fragment-based lead discovery. v. 4932011. p. 509-531.
74. CIULLI, A.; ABELL, C. Fragment-based approaches to enzyme inhibition. **Current Opinion in Biotechnology**, v. 18, n. 6, p. 489-496, 2007.
75. OPREA, T. I.; DAVIS, A. M.; TEAGUE, S. J.; LEESON, P. D. Is There a Difference between Leads and Drugs? A Historical Perspective. **Journal of Chemical Information and Computer Sciences**, v. 41, n. 3-6, p. 1308-1315, 2001.
76. HAJDUK, P. J. Fragment-based drug design: How big is too big? **Journal of Medicinal Chemistry**, v. 49, n. 24, p. 6972-6976, 2006.
77. SIEGAL, G.; AB, E.; SCHULTZ, J. Integration of fragment screening and library design. **Drug Discovery Today**, v. 12, n. 23-24, p. 1032-1039, 2007.
78. CARR, R. A. E.; CONGREVE, M.; MURRAY, C. W.; REES, D. C. Fragment-based lead discovery: Leads by design. **Drug Discovery Today**, v. 10, n. 14, p. 987-992, 2005.
79. 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.
80. FILTER, versão 2.0.2. Santa Fe: OPENEYE SCIENTIFIC SOFTWARE, I., 2010, 1 CD-ROM

-
81. CRUCIANI, G.; CRIVORI, P.; CARRUPT, P. A.; TESTA, B. Molecular fields in quantitative structure-permeation relationships: The VolSurf approach. **Journal of Molecular Structure: THEOCHEM**, v. 503, n. 1-2, p. 17-30, 2000.
82. RAREY, M.; KRAMER, B.; LENGAUER, T.; KLEBE, G. A fast flexible docking method using an incremental construction algorithm. **Journal of Molecular Biology**, v. 261, n. 3, p. 470-489, 1996.
83. GLIDE, versão 5.5. New York: SCHRÖDINGER, L., 2010, 1 DVD
84. OMEGA, versão 2.3.2. Santa Fe: OPENEYE SCIENTIFIC SOFTWARE, I., 2011, 1 CD-ROM
85. ROCS, versão 3.1.1. Santa Fe: OPENEYE SCIENTIFIC SOFTWARE, I., 2011, 1 CD-ROM
86. EON, versão 2.0.1. Santa Fe: OPENEYE SCIENTIFIC SOFTWARE, I., 2011, 1 CD-ROM
87. MAESTRO, versão 7.5. New York: SCHRÖDINGER, L., 2010, 1 CD-ROM
88. VIDA, versão 4.0.0. Santa Fe: OPENEYE SCIENTIFIC SOFTWARE, I., 2010, 1 CD-ROM
89. PETTERSEN, E. F.; GODDARD, T. D.; HUANG, C. C.; COUCH, G. S.; GREENBLATT, D. M.; MENG, E. C.; FERRIN, T. E. UCSF Chimera - A visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605-1612, 2004.
90. CASE, D. A.; CHEATHAM III, T. E.; DARDEN, T.; GOHLKE, H.; LUO, R.; MERZ JR, K. M.; ONUFRIEV, A.; SIMMERLING, C.; WANG, B.; WOODS, R. J. The Amber biomolecular simulation programs. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1668-1688, 2005.
91. CHELESKI, J.; FREITAS, R. F.; WIGGERS, H. J.; ROCHA, J. R.; DE ARAÚJO, A. P. U.; MONTANARI, C. A. Expression, purification and kinetic characterization of His-tagged glyceraldehyde-3-phosphate dehydrogenase from *Trypanosoma cruzi*. **Protein Expression and Purification**, v. 76, n. 2, p. 190-196, 2011.
92. LEACH, A. R.; SHOICHET, B. K.; PEISHOFF, C. E. Prediction of protein-ligand interactions. Docking and scoring: Successes and gaps. **Journal of**

- Medicinal Chemistry**, v. 49, n. 20, p. 5851-5855, 2006.
93. KONTOYIANNI, M.; MCCLELLAN, L. M.; SOKOL, G. S. Evaluation of docking performance: comparative data on docking algorithms. **Journal of Medicinal Chemistry**, v. 47, n. 3, p. 558-565, 2004.
94. SHOICHET, B. K. Virtual screening of chemical libraries. **Nature**, v. 432, n. 7019, p. 862-865, 2004.
95. YURIEV, E.; AGOSTINO, M.; RAMSLAND, P. A. Challenges and advances in computational docking: 2009 in review. **Journal of Molecular Recognition**, v. 24, n. 2, p. 149-164, 2011.
96. SIMONSON, T.; ARCHONTIS, G.; KARPLUS, M. Free energy simulations come of age: Protein-ligand recognition. **Accounts of Chemical Research**, v. 35, n. 6, p. 430-437, 2002.
97. MUEGGE, I.; MARTIN, Y. C. A general and fast scoring function for protein-ligand interactions: A simplified potential approach. **Journal of Medicinal Chemistry**, v. 42, n. 5, p. 791-804, 1999.
98. KRAMER, B.; RAREY, M.; LENGAUER, T. Evaluation of the FlexX incremental construction algorithm for protein-ligand docking. **Proteins: Structure, Function and Genetics**, v. 37, n. 2, p. 228-241, 1999.
99. ONODERA, K.; SATOU, K.; HIROTA, H. Evaluations of molecular docking programs for virtual screening. **Journal of Chemical Information and Modeling**, v. 47, n. 4, p. 1609-1618, 2007.
100. SKJÆRVEN, L.; REUTER, N.; MARTINEZ, A. Dynamics, flexibility and ligand-induced conformational changes in biological macromolecules: A computational approach. **Future Medicinal Chemistry**, v. 3, n. 16, p. 2079-2100, 2011.
101. DURRANT, J. D.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. **BMC Biology**, v. 9, 2011.
102. GALEAZZI, R. Molecular dynamics as a tool in rational drug design: Current status and some major applications. **Current Computer-Aided Drug Design**, v. 5, n. 4, p. 225-240, 2009.
103. AMBER11 BENCHMARK: NVIDIA GPU Acceleration Support. Disponível em:

<<http://ambermd.org/gpus/benchmarks.htm#Benchmarks/>>. Acesso em: 20 de jan., 2012

104. NIKOLOVA, N.; JAWORSKA, J. Approaches to measure chemical similarity – a review. **QSAR and Combinatorial Science**, v. 22, n. 9-10, p. 1006-1026, 2004.

105. LAKOWICZ, J. R.; SZMACINSKI, H.; NOWACZYK, K.; JOHNSON, M. L. Fluorescence lifetime imaging of free and protein-bound NADH. **Proceedings of the National Academy of Sciences of the United States of America**, v. 89, n. 4, p. 1271-1275, 1992.

106. GORDON, J. C.; MYERS, J. B.; FOLTA, T.; SHOJA, V.; HEATH, L. S.; ONUFRIEV, A. H⁺⁺: A server for estimating pK_as and adding missing hydrogens to macromolecules. **Nucleic Acids Research**, v. 33, n. SUPPL. 2, p. W368-W371, 2005.

107. JORGENSEN, W. L.; CHANDRASEKHAR, J.; MADURA, J. D.; IMPEY, R. W.; KLEIN, M. L. Comparison of simple potential functions for simulating liquid water. **The Journal of Chemical Physics**, v. 79, n. 2, p. 926-935, 1983.

108. GAUSSIAN, versão 03, Revision C.02. Wallingford: GAUSSIAN, I., 2004, 1 CD ROM

109. EMOLECULES: Chemical structure drawing search. Disponível em: <<http://www.emolecules.com/>>. Acesso em: 5-jan-2011, 2011

Apêndice 1 – Parâmetros de filtro das moléculas

```

MIN_MOLWT    150    "Minimum molecular weight"
MAX_MOLWT    300    "Maximum molecular weight"
MIN_ROT_BONDS 0     "Minimum number of rotatable bonds"
MAX_ROT_BONDS 3     "Maximum number of rotatable bonds"
MIN_HBOND_DONORS 0   "Minimum number of hydrogen-bond donors"
MAX_HBOND_DONORS 3   "Maximum number of hydrogen-bond donors"
MIN_HBOND_ACCEPTORS 0 "Minimum number of hydrogen-bond acceptors"
MAX_HBOND_ACCEPTORS 3 "Maximum number of hydrogen-bond acceptors"
MIN_XLOGP    -30000.0 "Minimum XLogP"
MAX_XLOGP     3.0    "Maximum XLogP"
MIN_2D_PSA   0.0     "Minimum 2-Dimensional (SMILES) Polar Surface Area"
MAX_2D_PSA   60.0    "Maximum 2-Dimensional (SMILES) Polar Surface Area"
PSA_USE_SandP true    "Count S and P as polar atoms"
ALLOWED_ELEMENTS H,C,N,O,F,S,Cl,Br,P
ELIMINATE_METALS Sc,Ti,V,Cr,Mn,Fe,Co,Ni,Cu,Zn,Y,Zr,Nb,Mo,Tc,Ru,Rh,Pd,Ag,Cd
MIN_SOLUBILITY moderately "Minimum solubility"
PSA_USE_SandP false    "Count S and P as polar atoms"
MIN_2D_PSA   0.0     "Minimum 2-Dimensional (SMILES) Polar Surface Area"
MAX_2D_PSA   150.0    "Maximum 2-Dimensional (SMILES) Polar Surface Area"
AGGREGATORS  true    "Eliminate known aggregators"
PRED_AGG     true    "Eliminate predicted aggregators"
#acceptable molecules must have <= instances of each of the patterns below
#specific, undesirable functional groups
RULE 0 quinone
RULE 0 pentafluorophenyl_esters
RULE 0 paranitrophenyl_esters
RULE 0 HOBT_esters
RULE 0 triflates
RULE 0 lawesson_s_reagent
RULE 0 phosphoramides
RULE 0 beta_carbonyl_quat_nitrogen
RULE 0 acylhydrazide
RULE 0 cation_C_Cl_I_P_or_S
RULE 0 phosphoryl
RULE 0 alkyl_phosphate
RULE 0 phosphinic_acid
RULE 0 phosphanes
RULE 0 phosphoranes
RULE 0 imidoyl_chlorides
RULE 0 nitroso
RULE 0 N_P_S_Halides
RULE 0 carbodiimide
RULE 0 isonitrile
RULE 0 triacyloxime
RULE 0 cyanohydrins
RULE 0 acyl_cyanides
RULE 0 sulfonylnitrile
RULE 0 phosphorylnitrile
RULE 0 azocyanamides
RULE 0 beta_azo_carbonyl
RULE 0 polyenes
RULE 0 saponin_derivatives
RULE 1 cytochalasin_derivatives
RULE 4 cycloheximide_derivatives
RULE 1 monensin_derivatives
RULE 1 squalastatin_derivatives
#functional groups which often eliminate compounds from consideration
RULE 0 acid_halide
RULE 0 aldehyde
RULE 0 alkyl_halide
RULE 0 anhydride
RULE 0 azide
RULE 0 azo

```

RULE 0 di_peptide
RULE 0 michael_acceptor
RULE 0 beta_halo_carbonyl
RULE 0 nitro
RULE 0 oxygen_cation
RULE 0 peroxide
RULE 0 phosphonic_acid
RULE 0 phosphonic_ester
RULE 0 phosphoric_acid
RULE 0 phosphoric_ester
RULE 0 sulfonic_acid
RULE 0 sulfonic_ester
RULE 0 tricarbo_phosphene
RULE 0 epoxide
RULE 0 sulfonyl_halide
RULE 0 halopyrimidine
RULE 0 perhalo_ketone
RULE 0 aziridine
RULE 1 oxalyl
RULE 0 alphahalo_amine
RULE 0 halo_amine
RULE 0 halo_alkene
RULE 0 acyclic_NCN
RULE 0 acyclic_NS
RULE 0 SCN2
RULE 0 terminal_vinyl
RULE 0 hetero_hetero
RULE 0 hydrazine
RULE 0 N_methoyl
RULE 0 NS_beta_haloethyl
RULE 0 propiolactones
RULE 0 nitroso
RULE 0 iodoso
RULE 0 iodoxy
RULE 0 noxide
#groups of molecules
RULE 0 dye
#functional groups which are allowed, but may not be wanted in high quantities
#common functional groups
RULE 6 alcohol
RULE 4 alkene
RULE 4 amide
RULE 4 amino_acid
RULE 2 amine
RULE 4 primary_amine
RULE 4 secondary_amine
RULE 4 tertiary_amine
RULE 2 carboxylic_acid
RULE 6 halide
RULE 0 iodine
RULE 2 ketone
RULE 4 phenol
RULE 1 imine
RULE 1 methyl_ketone
RULE 1 alkylaniline
RULE 4 sulfonamide
RULE 1 sulfonylurea
RULE 0 phosphonamide
RULE 0 alphahalo_ketone
RULE 0 oxaziridine
RULE 1 cyclopropyl
RULE 2 guanidine
RULE 0 sulfonimine
RULE 0 sulfinimine
RULE 1 hydroxamic_acid
RULE 0 phosphoryl
RULE 0 sulfinylthio

RULE 0 disulfide
RULE 0 enol_ether
RULE 0 enamine
RULE 0 organometallic
RULE 0 dithioacetal
RULE 1 oxime
RULE 0 isothiocyanate
RULE 0 isocyanate
RULE 3 lactone
RULE 3 lactam
RULE 1 thioester
RULE 1 carbonate
RULE 0 carbamic_acid
RULE 1 thiocarbamate
RULE 0 triazine
RULE 1 malonic
#other functional groups
RULE 2 alkyne
RULE 4 aniline
RULE 4 aryl_halide
RULE 2 carbamate
RULE 3 ester
RULE 5 ether
RULE 1 hydrazone
RULE 0 nonacylhydrazone
RULE 1 hydroxylamine
RULE 2 nitrile
RULE 2 sulfide
RULE 2 sulfone
RULE 2 sulfoxide
RULE 0 thiourea
RULE 1 thioamide
RULE 1 thiol
RULE 2 urea
RULE 0 hemiketal
RULE 0 hemiacetal
RULE 0 ketal
RULE 1 acetal
RULE 0 aминаl
RULE 0 hemiaminal
#protecting groups
RULE 0 benzyloxycarbonyl_CBZ
RULE 0 t_butoxycarbonyl_tBOC
RULE 0 fluorenylmethoxycarbonyl_Fmoc
RULE 1 dioxolane_5MR
RULE 1 dioxane_6MR
RULE 1 tetrahydropyran_THP
RULE 1 methoxyethoxymethyl_MEM
RULE 2 benzyl_ether
RULE 2 t_butyl_ether
RULE 0 trimethylsilyl_TMS
RULE 0 t_butyldimethylsilyl_TBMDS
RULE 0 triisopropylsilyl_TIPS
RULE 0 t_butyldiphenylsilyl_TBDPS
RULE 1 phthalimides_PHT
RULE 2 arenesulfonyl

Apêndice 2 - Parâmetros utilizados para simulação por Dinâmica Molecular

A: Minimização de solvente e íons

1QXS_complex: initial minimization solvent + ions

```
&cntrl
imin = 1,
maxcyc = 2000,
ncyc = 1000,
ntb = 1,
ntr = 1,
cut = 10.0
/
Hold the 1QXS fixed
500.0
RES 1 1437
END
END
```

B: Minimização de todo o sistema

1QXS_complex: initial minimization whole system

```
&cntrl
imin = 1,
maxcyc = 3500,
ncyc = 2000,
ntb = 1,
ntr = 0,
cut = 10.0
/
```

C: Aquecimento do sistema de 0 a 300 K

Heating 1QXS at V constant

```
&cntrl
ntx=1,
irest=0,
ntr=1,
ntc=2,
ntf=2,
nstlim=25000,
ntpr=1000,
ntwx=1000,
ntwr=1000,
dt=0.002,
cut=10.0,
ntt=3,
gamma_ln=2.0,
tautp=10.0,
tempi=0.0,
temp0=300.0,
ntb=1,
ntp=0,
/
Keep 1QXS fixed with weak restraints
10.0
RES 1 1438
END
END
```

D: Equilíbrio da densidade do sistema

Density 1QXS

```
&cntrl
imin=0,
irest=1,
ntx=5,
nstlim=100000,
dt=0.002,
```

```
ntc=2,  
ntf=2,  
cut=8.0,  
ntb=2,  
ntp=1,  
taup=1.0,  
ntpr=1000,  
ntwx=1000,  
ntt=3,  
gamma_ln=4.0,  
temp0=300.0,  
/  

```

E: Equilíbrio pré-dinâmica do sistema

```
equi 1QXS  
&cntrl  
imin=0,  
irest=1,  
ntx=5,  
nstlim=100000,  
dt=0.002,  
ntc=2,  
ntf=2,  
cut=8.0,  
ntb=2,  
ntp=1,  
taup=1.0,  
ntpr=1000,  
ntwx=1000,  
ntt=3,  
gamma_ln=4.0,  
temp0=300.0,  
/  

```

F: Dinâmica molecular do complexo

```
MD 1QXS  
&cntrl  
imin=0,  
irest=1,  
ntx=5,  
nstlim=500000,  
dt=0.002,  
ntc=2,  
ntf=2,  
cut=8.0,  
ntb=2,  
ntp=1,  
taup=1.0,  
ntpr=10000,  
ntwx=10000,  
ntt=3,  
gamma_ln=4.0,  
temp0=300.0,  
/  

```

G: Arrefecimento simulado

```
170ps fragments simulated annealing  
&cntrl  
imin = 0, irest = 1, ntx=5,  
ntc=2, ntf=2,  
ntpr=500, ntwx=1000,  
ntb = 2, cut = 10.0, rgbmax=10.0,  
nstlim = 340000, nscm= 0,  
dt = 0.0005,  
ntt = 1, ntp=1,  
temp0=310,  
nmropt=1,
```

```
tautp=1.0
/
&wt type='TEMP0', istep1=0,istep2=20000,
value1=310.0, value2=400.0,
/
&wt type='TEMP0', istep1=20001,istep2=40000,
value1=400.0, value2=500.0,
/
&wt type='TEMP0', istep1=40001,istep2=140000,
value1=500.0, value2=500.0,
/
&wt type='TEMP0', istep1=140001,istep2=160000,
value1=500.0, value2=400.0,
/
&wt type='TEMP0', istep1=160001,istep2=180000,
value1=400.0, value2=300.0,
/
&wt type='TEMP0', istep1=180001,istep2=200000,
value1=300.0, value2=200.0,
/
&wt type='TEMP0', istep1=200001,istep2=300000,
value1=200.0, value2=200.0,
/
&wt type='TEMP0', istep1=300001,istep2=320000,
value1=200.0, value2=310.0,
/
&wt type='TEMP0', istep1=320001,istep2=340000,
value1=310.0, value2=310.0,
/
&wt type='TAUTP', istep1=0,istep2=40000,
value1=1.0, value2=0.5,
/
&wt type='TAUTP', istep1=40001,istep2=140000,
value1=4.0, value2=4.0,
/
&wt type='TAUTP', istep1=140001,istep2=200000,
value1=0.5, value2=0.5,
/
&wt type='TAUTP', istep1=200001,istep2=300000,
value1=4.0, value2=4.0,
/
&wt type='TAUTP', istep1=300001,istep2=320000,
value1=0.5, value2=0.5,
/
&wt type='TAUTP', istep1=320001,istep2=340000,
value1=0.5, value2=1.0,
/
&wt type='END'
/
LISTOUT=POUT
```

Apêndice 3 – Gráficos de simulação por Dinâmica Molecular

NEQUIMED 293

NEQUIMED 295-1

NEQUIMED 295-2

NEQUIMED 295-3

NEQUIMED 291

NEQUIMED 294-1

NEQUIMED 294-3

NEQUIMED 358T-1

NEQUIMED 354T-1

NEQUIMED 355T-1

NEQUIMED 359T-1

NEQUIMED 350-1

NEQUIMED 360T-1

NEQUIMED 356T

NEQUIMED 361T