



UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA MEDICINAL  
E MODELAGEM MOLECULAR

Fábio Jorge de Nazaré Ferreira

**INVESTIGAÇÃO COMPUTACIONAL DE BROMO-ARILOXI-2-ACETAMIDA-  
ETIL-BENZIMIDAZÓLICOS COMO INIBIDORES NÃO-PEPTÍDICOS DA  
PROTEINASE CRUZAÍNA DE TRYPANOSOMA CRUZI**

Belém-Pará

2017

Dados Internacionais de Catalogação na Publicação (CIP)  
Sistema de Bibliotecas da Universidade Federal do Pará  
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

---

- F383i Ferreira, Fábio Jorge de Nazaré  
Investigação computacional de bromo-ariloxi-2-acetamida-etil-benzimidazólicos como inibidores não-peptídicos da proteinase cruzaiña de Trypanosoma cruzi / Fábio Jorge de Nazaré Ferreira. - 2017.  
80 f. : il. color.
- Dissertação (Mestrado) - Programa de Pós-graduação em Química Medicinal e Modelagem Molecular (PPGQM3), Instituto de Ciências da Saúde, Universidade Federal do Pará, Belém, 2017.  
Orientação: Prof. Dr. Agnaldo da Silva Carneiro  
Coorientação: Prof. Dr. Nelson Alberto Nascimento de Alencar.
1. Modelagem molecular. 2. Doença de Chagas. 3. Trypanosoma cruzi. 4. Cruzaína. 5. Dinâmica molecular. I. Carneiro, Agnaldo da Silva, *orient.* II. Título
- 

CDD 547.122

Fábio Jorge de Nazaré Ferreira

**Investigação computacional de  
bromo-ariloxi-2-acetamida-etil-benzimidazólicos  
como inibidores não-peptídicos da proteinase  
cruzaína de *Trypanosoma cruzi***

Belém-Pará

2017



Fábio Jorge de Nazaré Ferreira

**Investigação computacional de  
bromo-ariloxi-2-acetamida-etil-benzimidazólicos como  
inibidores não-peptídicos da proteinase cruzaina de  
*Trypanosoma cruzi***

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Química Medicinal e Modelagem Molecular da Universidade Federal do Pará (PPGQMMM-UFPA) como parte dos requisitos necessários para obtenção do título de Mestre em Química Medicinal e Modelagem Molecular.

Resultado: . Belém-Pará, 17 de outubro de 2017:

---

**Prof. Dr. Agnaldo da Silva Carneiro**  
Orientador

---

**Prof. Dr. Nelson Alberto Nascimento  
de Alencar**  
Coorientador

---

**Prof. Dr. Anderson Henrique Lima e  
Lima**  
Membro Externo

---

**Prof. Dr. Rosivaldo dos Santos Borges**  
Membro Interno

Belém-Pará  
2017



*Não há nada que os seres vivos façam  
que não possa ser compreendido do ponto de vista  
de sua constituição por átomos que se comportam  
segundo as leis da natureza.*

*- Richard Feynman, ganhador do Prêmio Nobel de 1965 em Física*



# Agradecimentos

A Deus por ter me dado saúde e força para superar as dificuldades.

À minha família, por sua capacidade de acreditar e investir em mim.

À minha esposa Sueli e meu filho Fábio de Jesus que são os maiores presentes que Deus poderia ter me dado e sempre farão parte de cada vitória.

À UFPA, pela oportunidade de fazer o curso.

Ao meu orientador Prof. Dr. Agnaldo Carneiro, pela orientação e suporte no pouco tempo que lhe coube, pelas suas correções e também aos incentivos ao longo do programa de mestrado.

Ao meu coorientador Prof. Dr. Nelson Alencar pelas correções, orientações e *scripts* que deram um norte correto ao meu trabalho.

Aos professores Dr. Rosivaldo Borges, Dr. Anderson Lima pela disposição em participar da banca.

Aos amigos do LCSB, Jordano, Mateus, Jean e Camila.

Aos amigos do LACC, Prof. Dr. Carlos Barros, Jorddy e Karoline.

Ao casal 100 %, Claubert Henrique e Amanda Ruslana, pelo companheirismo, pela amizade, por me incentivarem a fazer este mestrado e por serem praticamente meus coorientadores.



# Resumo

A doença de Chagas é uma infecção causada pelo protozoário flagelado *Trypanosoma cruzi*, transmitida por insetos conhecidos no Brasil como barbeiros. Na região amazônica, estudos têm mostrado que a contaminação por via oral tem sido frequente. Os únicos fármacos disponíveis para o tratamento da doença de Chagas - o Benzonidazol (Rochagan®), Roche) e o Nifurtimox (Lampit®, Bayer) - têm demonstrado uma eficiência limitada e severos efeitos colaterais. A cruzaina é uma enzima presente em todos os estágios do ciclo de vida do *T. cruzi* e é a mais abundante da família da papaína de cisteína proteases encontradas no parasito, sendo um alvo enzimático promissor para o planejamento e desenvolvimento de inibidores contra a doença. Inibidores não-peptídicos que ligam não-covalentemente à enzima foram sintetizados e avaliados biologicamente *in vitro* e *in vivo* por Ferreira e colaboradores (2014) análogos do composto líder (cristalografado) **8D** (ou B95), obtendo uma série de compostos ativos, dos quais os mais potentes são: **8K**, **8L** e **8R**. Este trabalho investigou as potenciais interações e energias da cruzaina (código PDB: 3KKU) complexadas com estes quatro ligantes por meio de ferramentas computacionais, a fim de ajudar a elucidar sua potencial atividade na inibição nesta enzima. O protocolo computacional (parâmetros, topologias, coordenadas, minimizações, termalizações e produções) foi o mesmo para cada sistema. Na etapa final de produção de dinâmica molecular (DM), cada sistema foi simulado por um período de 100 ns, aos quais foram analisados os valores de estabilidade do desvio da raiz média quadrática (RMSD) da enzima e a mudança acentuada na conformação do ligante 8L. Também foi avaliada a qualidade da simulação por meio de gráficos de energia potencial, cinética e total, volume e temperatura. Foram analisadas interações de ligações de hidrogênio dos ligantes com alguns resíduos de aminoácidos pertencentes ao sítio catalítico. Tem como destaque, a interação entre o ASP161 e o ligante 8R, sendo ratificado pela decomposição de energia por resíduo mostrando que o ASP161 tem a melhor contribuição. Em termos de energia livre de ligação, o  $\Delta G_{\text{total}}$  segue a tendência experimental, apontando o ligante 8R como o mais favorável à reação apresentando valor teórico de  $-30,04 \text{ kcal.mol}^{-1}$ . Esta espontaneidade da reação é ratificada por meio dos valores obtidos com o método SIE, cujo valor teórico foi de  $-7,54 \text{ kcal.mol}^{-1}$ . Os resultados deste trabalho devem favorecer a otimização do composto 8R ou desenvolvimento de uma série de análogos desta molécula a fim de se utilizar como possível fármaco para o tratamento da doença de Chagas.

**Palavras-chave:** Química medicinal. Modelagem Molecular. Doença de Chagas. *Trypanosoma Cruzi*. Cruzaina. Dinâmica molecular.



# Abstract

Chagas' disease is an infection caused by the *Trypanosoma Cruzi* flagellated protozoan transmitted by insects (gnat) known in Brazil as "barbeiro" (barber). In the Amazon region, studies have shown that oral contamination has been frequent. The only available drugs for the treatment of Chagas' disease - Benzonidazole (Rochagan®), Roche) and Nifurtimox (Lampit®, Bayer) - have shown limited efficiency and severe side effects. Cruzain is an enzyme present at all stages of the life cycle of *T. cruzi* and is the most abundant of the family of papain cysteine proteases found in the parasite, being a promising enzymatic target for the design and development of inhibitors against the disease. Non-peptidic non-covalently bound to the enzyme were synthesized and evaluated biologically *in vitro* and *in vivo* by Ferreira et al. (2014) analogs of the 8D (or B95) leader compound (crystallographic), yielding a series of active compounds, of which the most powerful are: 8K, 8L and 8R. This work investigated the potential interactions and energies of the cruzain (PDB code: 3KKU) complexed with these four ligands by means of computational tools in order to help elucidate their potential inhibition activity in this enzyme. The computational protocol (parameters, topologies, coordinates, minimizations, thermalizations and productions) was the same for each system. In the final stage of molecular dynamics (MD) production, each system was simulated for a period of 100 ns, to which the mean square deviation (RMSD) stability values of the enzyme and the marked change in 8L ligand conformation were analyzed. The quality of the simulation was also evaluated through potential, kinetic and total energy, volume and temperature graphs. Interactions of hydrogen bonds of the ligands with some amino acid residues belonging to the catalytic site were analyzed. The interaction between the ASP161 and the 8R ligand is emphasized, being ratified by the energy decomposition by residue showing that ASP161 has the best contribution. In terms of binding free energy, the  $\Delta G_{total}$  follows the experimental trend, pointing the 8R ligand as the most favorable to the reaction having a theoretical value of -30.04 kcal.mol<sup>-1</sup>. This spontaneity is ratified by means of the values obtained with the SIE method, whose theoretical value was -7.54 kcal.mol<sup>-1</sup>. The results of this work should favor the optimization of compound 8R or development of a series of analogs of this molecule in order to be used as a possible drug for the treatment of Chagas' disease.

**Keywords:** Medical Chemistry. Molecular Modeling. Chagas Disease. *Trypanosoma Cruzi*. Cruzain. Molecular Dynamics.



# Lista de ilustrações

Figura 1 – O chagoma como um sinal clássico da fase aguda da doença de Chagas	19
Figura 2 – Radiografia do tórax mostrando a cardiomegalia e aumento de órgãos digestivos de um paciente chagásico em 20 anos de evolução do estágio crônico da doença. . . . .	20
Figura 3 – Ciclo evolutivo do <i>T. cruzi</i> . . . . .	23
Figura 4 – Estruturas moleculares dos fármacos atualmente disponíveis para o tratamento da doença de Chagas . . . . .	24
Figura 5 – Conjunto de inibidores da cruzaina: (1) vinilsulfonas, (2) fluorometilce-tonas, (3) triazóis, (4) pirimidinas, (5) tiosemicarbazonas, (6) chalconas e (7) benzimidazóis. . . . .	26
Figura 6 – Estrutura cristalográfica da cruzaina. PDB 3KKU. Em roxo, as alfa-hélices. Em azul, folhas beta. Em verde, o sítio ativo. Em amarelo, as pontes dissulfetos. . . . .	27
Figura 7 – Estruturas químicas utilizadas neste trabalho. . . . .	29
Figura 8 – À esquerda, o complexo proteína-ligante, onde a enzima está represen-tada em modelo de fitas e o ligante em bastão. À direita, representação das interações da glicina e do aspartato com o inibidor (bastão). As interações de hidrogênio estão representadas pelas linhas tracejadas. . .	51
Figura 9 – Resumo do protocolo seguido para a realização da simulação. . . . .	53
Figura 10 – Representação gráfica da variação da energias total, cinética e potencial dos sistemas em função do tempo. . . . .	55
Figura 11 – Representação gráfica da variação da temperatura dos sistemas em função do tempo. . . . .	56
Figura 12 – Representação gráfica da variação de volume em função do tempo. . . .	57
Figura 13 – Representação gráfica da variação de densidade em função do tempo. .	57
Figura 14 – RMSD da cruzaina complexada com os ligantes 8D, 8K, 8L E 8R . . .	58
Figura 15 – RMSD em sobreposição dos ligantes 8D, 8K, 8L E 8R . . . . .	59
Figura 16 – Conformação do ligante 8L no RMSD mais estável (a) e no RMSD mais alto (b). . . . .	59
Figura 17 – Representação gráfica dos valores de RMSF em função dos resíduos de aminoácidos. . . . .	60
Figura 18 – Representação 3D da cruzaina. Em vermelho estão destacadas as regiões mais flexíveis da enzima; em azul, as mais rígidas. . . . .	61
Figura 19 – Interação de ligação hidrogênio entre o ligante 8D e o resíduo ASP161 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.	62
Figura 20 – Gráfico das distâncias pelo tempo. Ligante 8D. . . . .	63

Figura 21 – Interação de ligação hidrogênio entre o ligante 8K e o resíduo GLY66 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.	64
Figura 22 – Gráfico das distâncias pelo tempo. Ligante 8K. . . . .	64
Figura 23 – Interação de ligação hidrogênio entre o ligante 8L e o resíduo GLY66 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.	65
Figura 24 – Gráfico das distâncias pelo tempo. Ligante 8L. . . . .	65
Figura 25 – Interação de ligação hidrogênio entre o ligante 8R e o resíduo ASP161 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.	66
Figura 26 – Gráfico das distâncias pelo tempo. Ligante 8R. . . . .	67
Figura 27 – Gráfico da decomposição das energias por resíduo dos ligantes 8D e 8R.	69

# Lista de tabelas

Tabela 1 – Valores de RMSF do sítio ativo. . . . .	61
Tabela 2 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8D. . . . .	62
Tabela 3 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8K. . . . .	63
Tabela 4 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8L. . . . .	66
Tabela 5 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8R. . . . .	67
Tabela 6 – Energia livre de ligação via MMGBSA dos complexos. . . . .	68
Tabela 7 – Comparação dos valores de MMGBSA, SIE e experimental. . . . .	68
Tabela 8 – Decomposição da energia por resíduo, em kcal/mol . . . . .	69



# Lista de abreviaturas e siglas

Å	ângstrom, $10^{-10}$ metros, unidade de medida de distância
AMBER	<i>Assisted Model Building with Energy Refinement</i>
B3LYP	Becke, <i>three-parameter</i> , Lee-Yang-Parr
CHELPG	<i>CHarges from ELectrostatic Potential using a Grid based method</i>
CPU	<i>Central Processing Unit</i>
DFT	Teoria do Funcional Densidade
DM	Dinâmica Molecular
GAFF	<i>General Amber Force Field</i>
GGA	<i>Generalized Gradient Approximation</i>
GPU	<i>Graphics Processing Unit</i>
GTO	<i>Gaussian Type Orbital</i>
HF	Hartree-Fock
HK	Hohenberg e Kohn
HTS	<i>High Throughput Screening</i>
IC <sub>50</sub>	Concentração necessária para inibir 50 % da atividade
IUPAC	União Internacional de Química Pura e Aplicada (da sigla inglesa – <i>International Union of Pure and Applied Chemistry</i> )
K <sub>i</sub>	Constante de inibição
KS	Kohn-Shan
LDA	<i>Local Density Approximation</i>
MMGBSA	<i>Molecular mechanics generalized Born surface area</i>
MMPBSA	<i>Molecular mechanics Poisson-Boltzmann surface area</i>
nM	nanomolar, $10^{-9}$ M, unidade de medida de concentração
NMR	<i>Nuclear Magnetic Resonance</i>

ns	nanossegundo, $10^{-9}$ segundos, unidade de medida de tempo
OMS	Organização Mundial da Saúde
PDB	<i>Protein Data Bank</i>
pH	Potencial Hidrogeniônico
pKa	escala logarítmica da constante de acidez
ps	picosegundo, $10^{-6}$ segundos, unidade de medida de tempo
RESP	<i>Restrained Electrostatic Potential</i>
RMSD	<i>Root Mean Square Deviation</i>
SCF	<i>Self Consistent Field</i>
SESPA	Secretaria de Saúde Pública do Estado do Pará
STO	<i>Slater Type Orbital</i>

# Sumário

	<b>Introdução</b>	<b>19</b>
<b>1</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>23</b>
<b>2</b>	<b>OBJETIVOS</b>	<b>31</b>
2.1	Objetivo geral	31
2.2	Objetivos específicos	31
<b>3</b>	<b>METODOLOGIA</b>	<b>33</b>
3.1	Modelagem molecular	33
3.2	Mecânica molecular	34
3.3	Dinâmica molecular	36
3.4	Método de Hartree-Fock	40
3.5	Teoria do Funcional de Densidade	41
3.6	Conjunto de Funções de Base	43
3.7	O método RESP	45
3.8	RMSD	46
3.9	RMSF	46
3.10	Energia livre de ligação	47
3.11	Decomposição da energia livre por resíduo	49
3.12	Protocolo para as simulações	50
3.12.1	Obtenção da estrutura inicial e preparação do arquivo PDB	50
3.12.2	Otimização geométrica e cargas RESP dos ligantes	51
3.12.3	Preparação de arquivos de parâmetros, topologia e coordenadas	52
3.12.4	Minimização, termalização e produção	52
<b>4</b>	<b>RESULTADOS</b>	<b>55</b>
4.1	Energias, temperatura, volume e densidade	55
4.2	RMSD dos complexos	57
4.3	RMSF dos complexos	60
4.4	Análise das ligações de hidrogênios e distâncias	61
4.5	Energia livre de ligação e energia de interação solvatada (SIE)	67
4.6	Decomposição da energia por resíduos	68
<b>5</b>	<b>CONCLUSÃO</b>	<b>71</b>

**REFERÊNCIAS . . . . . 73**

# Introdução

A Tripanossomíase Americana ou doença de Chagas é uma doença tropical parasitária responsável por cerca de 14 000 mortes por ano na América Latina, segundo a Organização Mundial da Saúde (OMS) (WHO, 2015). A descrição do agente etiológico da doença, o protozoário *Trypanosoma cruzi*, e também o ciclo de vida do parasito, foi realizada pelo médico sanitário brasileiro Carlos Chagas em 1909 (CLAYTON, 2010). Este ciclo envolve uma fase de multiplicação intracelular obrigatória no hospedeiro vertebrado (mamíferos) e extracelular no inseto vetor (triatomíneos) em três estágios de diferenciação celular: o tripomastigota (eliminada pelo inseto vetor e transmissível ao homem) e as formas replicativas amastigota (hospedeiro vertebrado) e epimastigota (inseto vetor).

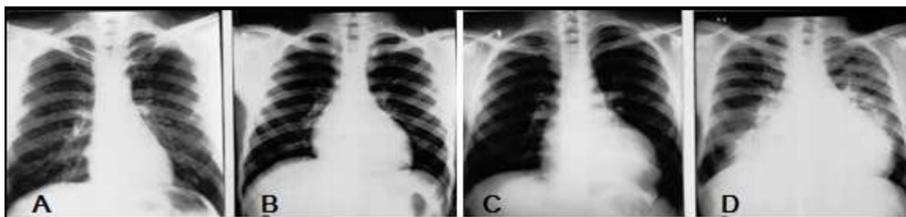
A maioria dos infectados não apresenta manifestações clínicas. Na doença há duas fases: a aguda e a crônica. A fase aguda é geralmente assintomática, e tem uma incubação de uma semana a um mês após a infecção. Se a infecção for por meio de picada do inseto, pode-se desenvolver uma lesão volumosa no local, chamado chagoma (TARTARI; NETO, 1953), Figura 1. A fase crônica sintomática decorre com maior frequência de lesões cardíacas, com cardiomegalia e arritmias (MUNOZ-SARAVIA et al., 2012), conforme apresentado na Figura 2. Este quadro clínico da doença durante a fase crônica é decorrente da ação direta do parasito e indireta da resposta autoimune do hospedeiro que ataca e destrói tecidos saudáveis do corpo, levando a um processo inflamatório autossustentável que persiste mesmo após o tratamento antiparasitário (URBINA, 2010). Desta forma, a evolução da doença diminui drasticamente a qualidade de vida e a capacidade produtiva do paciente chagásico, levando a grande morbidade, mortalidade e vários outros problemas socioeconômicos quando comparados com outras doenças como malária, esquistossomose e leishmaniose (JR; RASSI; MARIN-NETO, 2010).

Figura 1 – O chagoma como um sinal clássico da fase aguda da doença de Chagas



Fonte: Adaptado de (REY, 2002).

Figura 2 – Radiografia do tórax mostrando a cardiomegalia e aumento de órgãos digestivos de um paciente chagásico em 20 anos de evolução do estágio crônico da doença.



Fonte: Adaptado de (MUNOZ-SARAVIA et al., 2012).

A doença de Chagas sofre diferenciação nas contaminações em meio urbano. Na região amazônica, estudos têm mostrado que a contaminação por via oral tem sido frequente. Segundo dados da Secretaria de Saúde Pública do Estado do Pará (SESPA, 2014), em 2014 foram confirmados 156 casos de doença de Chagas aguda em vinte municípios paraenses, sendo que a maioria dos casos ocorreu por transmissão oral. Em 2015, foram registrados 48 casos com dois óbitos. As regiões de saúde que apresentaram maiores números de casos no mesmo ano foram: Tocantins (Igarapé Miri, Abaetetuba, Moju, Barcarena); Metropolitana I (Belém e Ananindeua) com o maior número de casos) e Marajó II (Anajás e Breves).

A forma mais comum dessas contaminações acontecer é por meio do açaí, no momento em que é despulpado em máquina ou manualmente, onde o inseto transmissor mais popular, o barbeiro, estando contaminado pode ser triturado ou as fezes dele podem contaminar o açaí tornando-o impróprio para o consumo.

Desta forma, em virtude das limitações da quimioterapia atual, em relação à doença de Chagas, torna-se necessária a descoberta de novos agentes quimioterápicos para o tratamento seguro e eficaz da doença.

O processo de descoberta e desenvolvimento de fármacos é complexo, longo e de alto custo, tendo suas raízes profundamente ligadas às inovações científicas e tecnológicas (GUIDO; ANDRICOPULO; OLIVA, 2010). Os avanços expressivos da química e biologia e a melhor compreensão de vias bioquímicas, alvos moleculares e de mecanismos que levam ao aparecimento e desenvolvimento de doenças, tornaram possível a descoberta de inovações terapêuticas notáveis, proporcionando melhorias significativas na qualidade de vida das diversas populações no mundo.

A química medicinal, de reconhecido papel central no processo de produção e desenvolvimento de fármacos, caracteriza-se por seu relevante caráter multidisciplinar, abrangendo diversas especialidades, tais como química orgânica, bioquímica, farmacologia, informática, biologia molecular e estrutural, física, entre outras (BARREIRO; FRAGA, 2008). De acordo com a União Internacional de Química Pura e Aplicada (da sigla inglesa Iupac – *International Union of Pure and Applied Chemistry*), a química medicinal envolve a invenção, a descoberta, o planejamento, a identificação, a preparação e a interpretação

do mecanismo de ação molecular de compostos biologicamente ativos. Além da descoberta de moléculas bioativas, a química medicinal também incorpora os estudos do metabolismo e das relações entre a estrutura química e atividade (WERMUTH, 2003). Assim, fica claro o estabelecimento de interfaces fundamentais entre as ciências químicas, biológicas, farmacêuticas, médicas, físicas e computacionais.

A modelagem molecular tem como objetivo principal reproduzir o comportamento real (experimental) de moléculas e sistemas moleculares (SANT et al., 2009). Com o aperfeiçoamento dos microcomputadores, estes se tornaram uma poderosa ferramenta para ser usada no desenvolvimento de fármacos, visualizando as conformações tridimensionais tanto dos ligantes como de seus sítios-ativo, a fim de avaliar suas interações. Baseado nos estudos computacionais, pode-se sintetizar e testar os compostos mais promissores, o que aumenta consideravelmente as chances de descobrir uma molécula biologicamente ativa que possa ser usada como fármaco. A tendência atual é o crescimento cada vez maior das aplicações de softwares, principalmente em centros de pesquisa e universidades, pois as tecnologias *in silico*, quando comparadas com os métodos *in vitro* e *in vivo*, são menos onerosas, consomem menos tempo, possuem maior rendimento, têm maior reprodutibilidade se o mesmo modelo é utilizado, requerem reduzida síntese de compostos, podem submeter-se à otimização constante (adicionando propriedades úteis, descritores ou novos compostos), e têm o potencial para reduzir o uso de animais (GUIDO; ANDRICOPULO; OLIVA, 2010).

Dentro do estudo computacional, a Dinâmica Molecular (DM) é uma técnica de simulação que calcula o movimento dos átomos em uma molécula ou de átomos individuais ou moléculas em sólidos, líquidos e gases, de acordo com as leis do movimento de Newton (VERLI, 2014). A DM permite avaliar o padrão de flexibilidade conformacional de um determinado sistema em função do tempo, o que traz uma aproximação mais realista dos fenômenos biológicos (NAMBA; SILVA; SILVA, 2008). Soma-se a isto, a abordagem das interações dos compostos com os resíduos aminoácidos contribuintes com base na energia livre de ligação, na energia de interação solvatada e na análise de decomposição de energia por resíduo (CELE; RAMESH; SOLIMAN, 2016).

Portanto, o estudo por meio dessas ferramentas computacionais permite elucidar o comportamento de inibidores da cruzaina e é de fundamental importância para a proposição de novas moléculas ativas no combate à doença de Chagas.

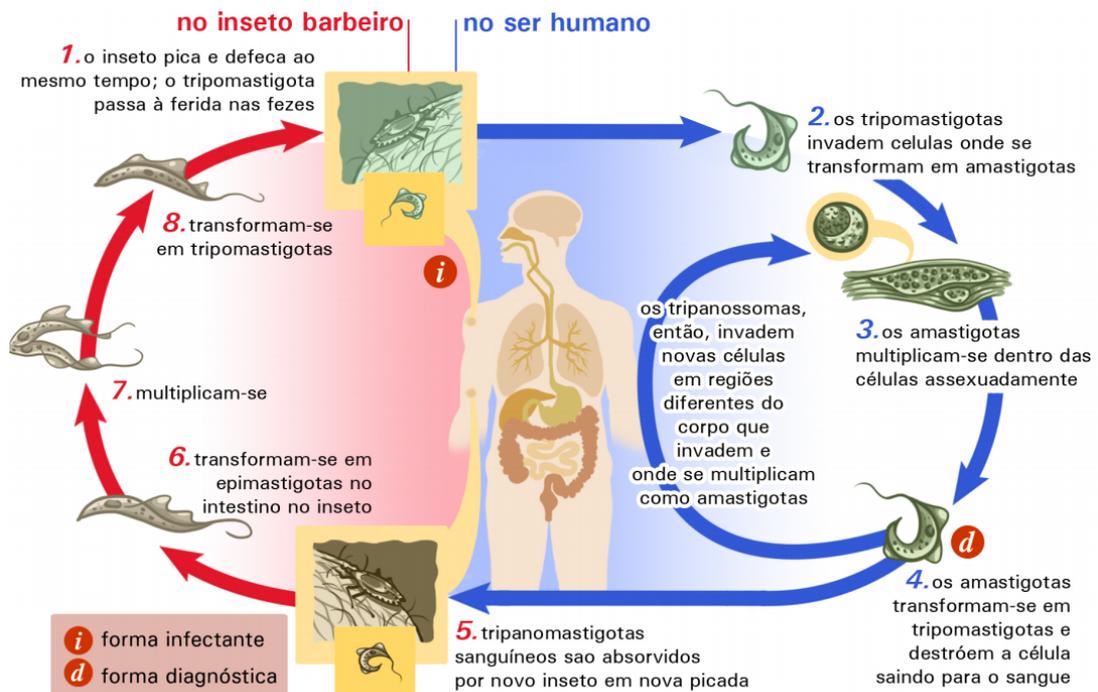


# 1 Fundamentação teórica

A doença de Chagas ou tripanossomíase americana é uma infecção causada pelo protozoário flagelado *Trypanosoma cruzi* (*T. cruzi*). Segundo a Organização Mundial de Saúde (OMS), a doença atinge aproximadamente 6 a 7 milhões de pessoas em todo o mundo, no qual a América Latina é região endêmica da doença, onde milhões de indivíduos vivem em áreas de risco de infecção, causando cerca de 10.000 mortes por ano. Tal doença na sua fase crônica causa deficiências significativas, tendo um alto impacto social e econômico, incluindo desemprego e diminuição de capacidade produtiva (ESPERANDIM et al., 2013).

As formas de transmissão são por via oral, transfusão sanguínea, transplante de órgãos, congênita e a mais conhecida, e também responsável pela maior parte dos casos de infecção em seres humanos e outros vertebrados, ocorre via fezes ou urina do inseto triatomíneo hematófago infectado, conhecido vulgarmente como “barbeiro”, que pode ser inoculada dentro da pele pela picada do inseto ou através de feridas ou escoriações causadas pelo ato de coçar (ESPERANDIM et al., 2013). O ciclo biológico do *T. cruzi* - Figura 3 - envolve um hospedeiro vertebrado, humanos e outros mamíferos, e um invertebrado, geralmente os triatomíneos, que pode ser encontrado nas seguintes formas: amastigotas, epimastigotas e tripomastigotas, onde esta última é a forma infectante da doença (TEIXEIRA et al., 2012).

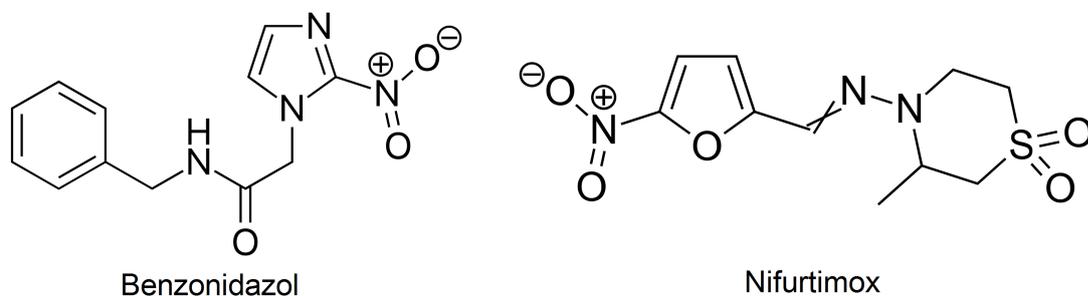
Figura 3 – Ciclo evolutivo do *T. cruzi*



Fonte: (CDC, 2015)

A doença de Chagas foi descoberta pelo cientista brasileiro Carlos Chagas há mais de 100 anos, porém, não existe um único fármaco seguro e eficaz para o tratamento quimioterápico da doença, apesar do genoma (EL-SAYED et al., 2005) e proteoma (ATWOOD et al., 2005) do *T. cruzi* serem conhecidos. É notório o pouco progresso realizado no tratamento da doença, parte devido ao desinteresse das indústrias farmacêuticas que veem poucos atrativos econômicos para o desenvolvimento de novas quimioterapias sendo considerada deste modo, uma doença negligenciada. Atualmente, o tratamento é mais sintomático do que etiológico. Do ponto de vista sintomático, procura-se amenizar as diversas manifestações clínicas da doença (administração de diuréticos no caso de insuficiência cardíaca congestiva). Do ponto de vista etiológico, os dois fármacos (Figura 4) atualmente disponíveis para o tratamento desta doença - o Benzonidazol (Rochagan®, Roche) e Nifurtimox (Lampit®, Bayer) - foram introduzidos há mais de 40 anos, têm demonstrado uma eficiência limitada contra o estágio crônico da doença (URBINA, 2010). A estes fármacos são ainda associados severos efeitos colaterais, grande variabilidade na eficiência terapêutica para pacientes vindos de diferentes regiões geográficas que, aliados ao longo período de tratamento (meses), levam ao frequente abandono da quimioterapia pelos pacientes (CASTRO; DEMECCA; BARTEL, 2006).

Figura 4 – Estruturas moleculares dos fármacos atualmente disponíveis para o tratamento da doença de Chagas



Fonte: (DIAS et al., 2009)

É de elevada importância a identificação de compostos bioativos que atuem em proteínas no conceito atual da Química Medicinal. Quase todos os processos biossintéticos dos organismos vivos são regulados por enzimas. As enzimas são proteínas especializadas na catálise de reações biológicas e estão entre as biomoléculas mais notáveis, devido a sua extraordinária especificidade e poder catalítico, que são muito superiores aos dos catalisadores produzidos pelo homem. Praticamente todas as reações que caracterizam o metabolismo celular são catalisadas por enzimas (FREY; HEGEMAN, 2007). Elas estão presentes nas células dos seres vivos e são de fundamental importância, pois dependendo do interesse pode-se melhorar ou inibir a atividade de uma determinada enzima (FREY; HEGEMAN, 2007). Dentro deste contexto a enzima cisteína protease do *Trypanosoma cruzi* tem-se mostrado um alvo promissor para o planejamento de inibidores contra o

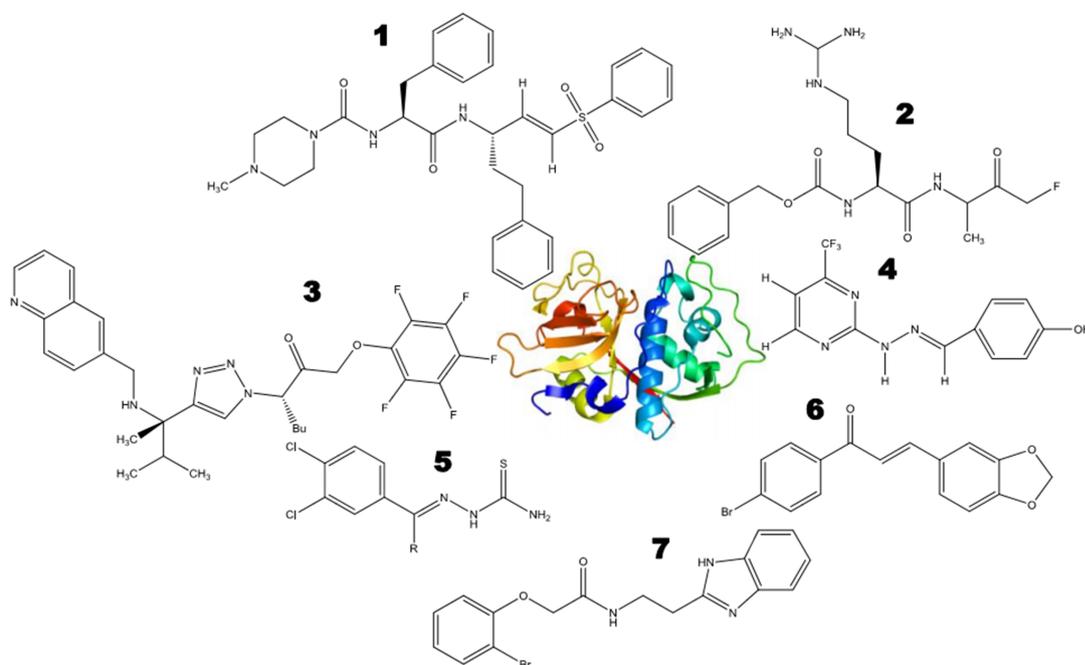
parasito, em particular a cruzaina.

A cruzaina, forma recombinante da cruzipaina (EC 3.4.22.51), é o membro mais abundante da família da papaína de cisteíno proteases encontradas no *T. cruzi*. Esta enzima está envolvida no processo de infecção da célula do hospedeiro, replicação, diferenciação, metabolismo, evasão de mecanismos de defesa sendo essencial na regulação da interação entre parasito e hospedeiro. A vulnerabilidade do parasito a inibidores da cruzaina somados a falta de sistemas de redundância para esta enzima, fazem desta um alvo terapêutico validado importante no combate à doença de Chagas (CAZZULO; STOKA; TURK, 2001; ENGEL et al., 1998; DOYLE et al., 2011). Esta enzima é expressa por diferentes cepas de *T. cruzi* e em todos os seus estágios de desenvolvimento. Nas epimastigotas, a cruzaina encontra-se primordialmente em organelas denominadas reservossomos. É nesta forma que a proteína apresenta sua maior abundância. Para o estágio intracelular amastigota, chave no tratamento da doença, isoformas da protease são encontradas ligadas à face externa da membrana plasmática diretamente em contato com o citoplasma da célula do hospedeiro, enquanto que nas tripomastigotas situa-se nos lisossomos e na bolsa flagelar (SOUTO-PADRON et al., 1990). Portanto, a cruzaina pode operar em dois microambientes químicos: no reservossoma (pH 6,3) e na superfície do parasito (pH 7,4) (ENGEL et al., 2000).

A cruzaina foi validada como alvo por meio de estudos bioquímicos, modelos animais e determinação da estrutura cristalográfica de raios-X. Além disso, novas classes de inibidores da cruzaina foram descobertas ao longo dos anos, e alguns dos modos de ligação foram determinados por meio de cristalografia (GILLMOR; CRAIK; FLETTERICK, 1997; MCGRATH et al., 1995; MOTT et al., 2009; FERREIRA et al., 2010). No entanto, nenhum desses compostos atingiu ensaios clínicos.

A Figura 5 apresenta as principais classes de inibidores da cruzaina (FERREIRA et al., 2010): os derivados peptídicos, como as vinilsulfonas (1) (KERR et al., 2009; BRINEN et al., 2000; BRYANT et al., 2009) e as fluorometilcetonas (2) (GILLMOR; CRAIK; FLETTERICK, 1997); e os derivados não-peptídicos, como os triazóis (3) (BRAK et al., 2008), as pirimidinas (4) (ZANATTA et al., 2008), as tiossemicarbazonas (5) (FUJII et al., 2005), as chalconas (6) (BORCHHARDT et al., 2010) e os benzimidazóis (7) (FERREIRA et al., 2010).

Figura 5 – Conjunto de inibidores da cruzaina: (1) vinilsulfonas, (2) fluorometilcetonas, (3) triazóis, (4) pirimidinas, (5) tiosemicarbazonas, (6) chalconas e (7) benzimidazóis.

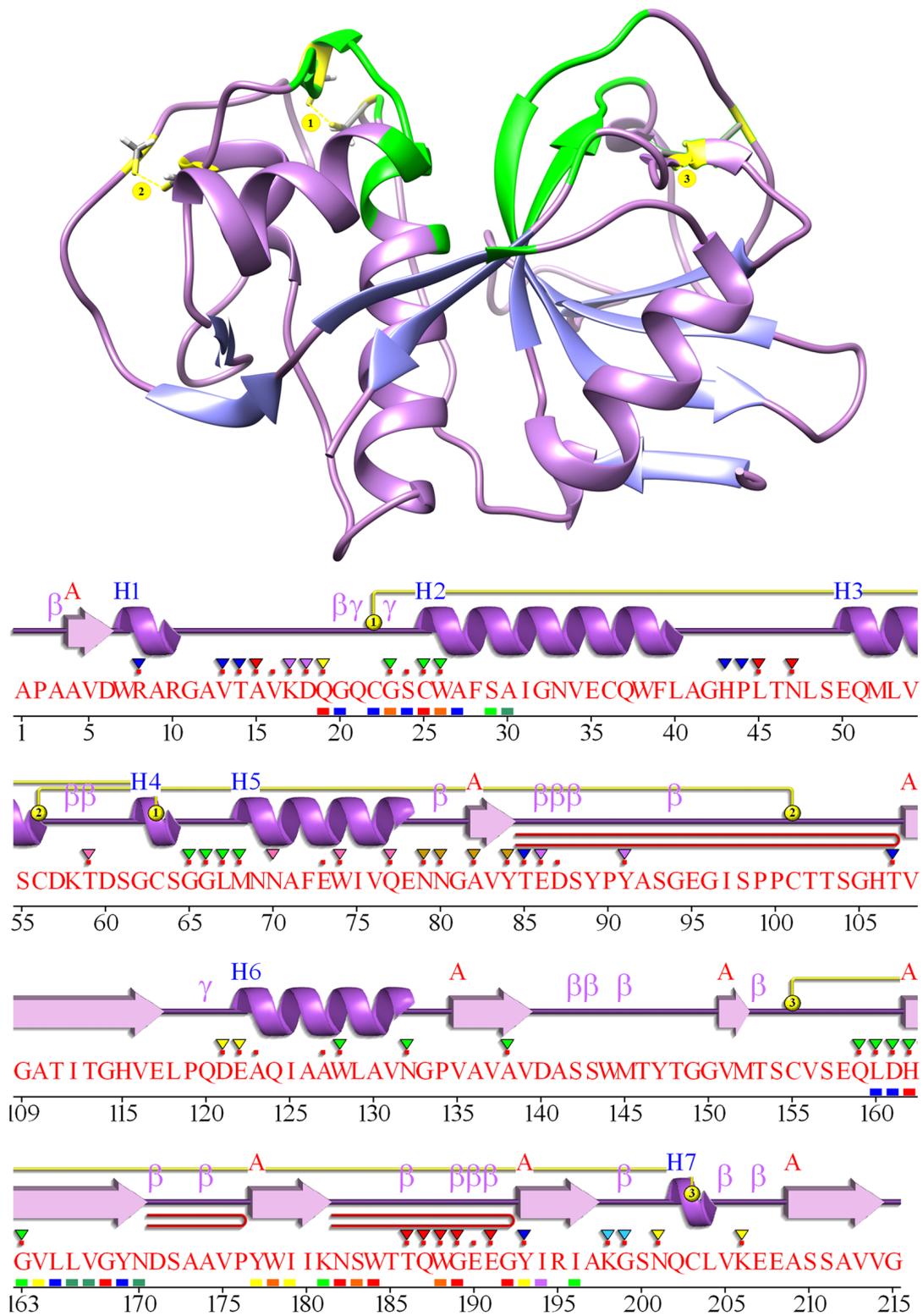


Fonte: (SOUZA, 2012).

A primeira estrutura cristalográfica da cruzaina foi determinada em complexo com um inibidor irreversível do tipo fluorometilcetona (MCGRATH et al., 1995), o qual serviu de base para uma variedade de estudos da enzima. Outras estruturas de inibidores em complexo com a cruzaina estão depositadas no PDB.

Estruturalmente, a cruzaina é composta por uma cadeia polipeptídica de 215 resíduos de aminoácidos, estruturada em dois domínios: um predominantemente formado por alfa-hélices e o outro constituído por interações do tipo folhas beta antiparalelas Figura 6. A tríade catalítica é composta pelos resíduos CYS25, HIS162 e ASN182 (SANTOS et al., ). Tanto a tríade catalítica, como o sítio global de ligação ao substrato, são encontrados em uma fenda entre os dois domínios (MCGRATH et al., 1995). O sítio ativo consiste dos seguintes aminoácidos: Gly 23, Cys 25, Trp 26, Ser 61, Gly 65, Gly 66, GLY133, Gln 159, Leu 160, Asp 161, His 162.

Figura 6 – Estrutura cristalográfica da cruzaina. PDB 3KKU. Em roxo, as alfa-hélices. Em azul, folhas beta. Em verde, o sítio ativo. Em amarelo, as pontes dissulfetos.



Fonte: Adaptado de (LASKOWSKI, 2001).

A elucidação da estrutura cristalográfica do domínio catalítico da cruzaina permitiu o desenho racional de moléculas sintéticas, capazes de agir sobre as cisteíno proteases

como inibidores de alta eficácia e baixa toxicidade (SCHARFSTEIN, 2008).

Na cruzaína, foram investigados os fatores estruturais e energéticos que determinam a atividade catalítica desta enzima envolvendo o substrato Benzoil-Arginina-Alanina-7-amino-4-metilcoumarina (Z-Arg-Ala-AMC), por meio de simulações de dinâmica molecular (DM) e métodos híbridos, que combinam cálculos de mecânica quântica (MQ) e de mecânica molecular (MM) (CARNEIRO, 2012).

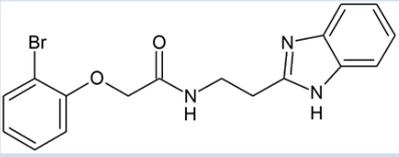
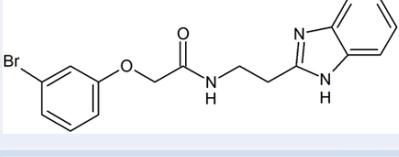
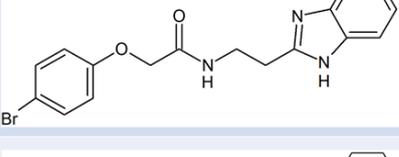
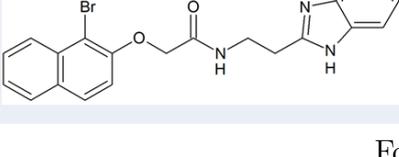
O trabalho de Malvezzi identificou inibidores dessa cisteíno-protease por meio de aplicação de modelos de virtual screening sobre bibliotecas de compostos orgânicos de baixa massa molecular. Para reconhecimento detalhado das características estruturais da enzima e de seu sítio de ligação, utilizou-se dinâmica molecular (MALVEZZI, 2008).

Uma série de moléculas não-peptídicas tem sido alvo de estudos. Recentemente uma série de inibidores não peptídicos, reversíveis foi identificada utilizando métodos *in silico* de triagem, ensaios bioquímicos e cristalografia de raios X, inibindo a cruzaína na faixa sub-micromolar (WIGGERS et al., 2013). Cinco destes inibidores foram ativos nos ensaios celulares contra formas tripomastigotas infectivas do *T. cruzi* exibindo potência comparável ao Benzonidazol, utilizado como controle. Apesar dos inibidores peptídicos apresentarem alta eficácia e baixa toxicidade, os inibidores não-peptídicos são capazes de apresentar maior biodisponibilidade oral. A natureza não peptídica desses inibidores em associação com seu baixo custo de síntese, fazem destes compostos novos candidatos para o desenvolvimento de uma efetiva terapia tripanocida (DU et al., 2002; BRAK et al., 2008).

Em um estudo executado pelo grupo de (FERREIRA et al., 2010), foi cocristalizado o composto B95 (classe de inibidores não covalente) com a cruzaína. Este estudo foi executado por meio de triagem virtual, docagem e HTS e depositado sob o código 3KKU no banco de dados PDB. Este inibidor reversível competitivo caracteriza-se como um importante composto líder para estudos de química medicinal.

Posteriormente, (FERREIRA et al., 2014), sintetizaram e avaliaram biologicamente *in vitro* e *in vivo*, análogos do composto líder **8D** (N-[2-(1H-benzimidazol-2-il)etil]-2-(2-bromofenoxi)acetamida) - no banco de dados PDB, este é denominado como B95 - obtendo uma série compostos ativos, dos quais os mais potentes são: **8K** N-[2-(1H-benzimidazol-2-il)etil]-2-(3-bromofenoxi)acetamida, **8L** (N-[2-(1H-benzimidazol-2-il)etil]-2-(4-bromofenoxi)acetamida) e **8R** (N-[2-(1H-benzimidazol-2-il)etil]-2-[(1-bromonaftaleno-2-il)oxi]acetamida), conforme Figura 7. Estes são os ligantes utilizados neste trabalho. No artigo, destaca-se o análogo 8R por ser o menos citotóxico e o mais potente, sendo um possível agente terapêutico no tratamento da doença de Chagas.

Figura 7 – Estruturas químicas utilizadas neste trabalho.

ESTRUTURA	DESIGNAÇÃO	IC <sub>50</sub> ( $\mu$ M)	K <sub>i</sub> (nM)
	<b>8D</b>	<b>0,8 ± 0,1</b>	<b>800</b>
	<b>8K</b>	<b>0,8 ± 0,5</b>	-
	<b>8L</b>	<b>2,7 ± 0,4</b>	-
	<b>8R</b>	<b>0,21 ± 0,04</b>	<b>82</b>

Fonte: (FERREIRA et al., 2014)

Os dados experimentais mostram que o composto 8R apresenta-se mais favorável no processo inibitório da enzima cruzaina, uma vez que possui um menor IC<sub>50</sub>, entorno de 0,21  $\mu$ M e uma constante de inibição de 82 nM.

No trabalho de (SANTOS et al., ), foi feita investigação do comportamento dinâmico e das energias de um conjunto de inibidores não covalentes das enzimas cruzaina e rodesaina. Estados de protonação diferentes tanto dos resíduos catalíticos das enzimas (His162 e Cys25) quanto do ligante foram investigados por meio de dinâmica molecular e cálculos de energia livre de modo a elucidar o possível modo de ligação do inibidor não-covalente 8D e de uma série de análogos desse inibidor.



## 2 Objetivos

### 2.1 Objetivo geral

Investigar as potenciais interações da cruzaina do *T. cruzi* complexadas com quatro inibidores bromo-ariloxi-2acetamida-etil-benzimidazólicos por meio de ferramentas computacionais.

### 2.2 Objetivos específicos

- Avaliar o comportamento molecular do sistema enzimático por meio de dinâmica molecular.
- Identificar as principais interações intermoleculares que contribuem para o processo de inibição.
- Avaliar quais as interações de ligação de hidrogênio e suas distâncias médias.
- Verificar quais resíduos aminoácidos fazem maior contribuição energética de interação.
- Estimar a afinidade dos ligantes frente ao sítio ativo da enzima cruzaina através de cálculos de energia livre.



## 3 Metodologia

### 3.1 Modelagem molecular

Química teórica pode ser definida como uma descrição matemática de química. Um processo pode ser aplicado em química teórica quando matematicamente bem desenvolvido para que possa ser implementado num programa de computador. Atualmente, muitas propriedades podem ser calculados teoricamente; no entanto, exigem muito mais precisão, mas produzem bem o suficiente para comparar teoria e experimental (YOUNG, 2001).

Há muita confusão entre os termos “química teórica”, “química computacional” e “modelagem molecular”. Química computacional inclui não apenas métodos de química teórica (química quântica), mas também abrange métodos de mecânica molecular (derivadas de física clássica). Modelagem molecular utiliza estes procedimentos de manuseio de estruturas e propriedades das moléculas que dependem dessas estruturas tridimensionais. Estes métodos são capazes de calcular a disposição geométrica correspondente a moléculas estáveis (minimização), simulações, análise conformacional, energias relativas, momento dipolar, polarizabilidade, espectros no infravermelho, espectros Raman, NMR, cargas de Mulliken e estado de transição (LEWARS, 2016).

Existem duas abordagens comumente aplicadas na área de modelagem molecular, que são as aproximações clássica, também denominadas mecânica molecular (MM), e as aproximações baseadas na química quântica (QM). Quando se está interessado em fenômenos que envolvem a quebra ou formação de ligações atômicas, ou a avaliação de polarização e transferência de cargas em moléculas, será necessária a utilização de métodos quânticos (SCHLICK, 2010).

O desenvolvimento de computadores mais poderosos permitiu a expansão de novos métodos teóricos. Até a década de 1950 apenas sistemas hidrogeniônicos pôde ser resolvido. Sistemas com duas partículas foram resolvidos por um sistema de coordenadas envolvendo centro de massa com duas pseudo partículas. Até então, os cálculos precisos de moléculas maiores parecia impossível (LEWARS, 2016).

Atualmente, os sistemas com mais de 100 átomos são facilmente calculados em um computador *desktop* em um bom nível de teoria. Sistemas mais complexos, utilizando os níveis mais robustos da teoria, incluem o desafio da modelagem molecular nos dias de hoje. Mas, com o advento dos processadores multicore mais eficientes, computação acelerada por placas de vídeo (GPU, *Graphics Processing Unit*) juntamente com uma CPU para acelerar aplicativos, estes obstáculos estão sendo superados.

## 3.2 Mecânica molecular

Apesar de não apresentar a mesma precisão dos métodos quânticos, muitos estudos de química teórica são desenvolvidos com o uso de técnicas computacionais baseadas em mecânica molecular, que utilizam as teorias de mecânica clássica para modelar sistemas moleculares.

A mecânica molecular considera as moléculas como um conjunto de partículas com uma determinada massa ligadas entre si por um conjunto de molas. As interações e movimentos são descritos pelas leis da mecânica clássica (newtonianas). Consideram-se os núcleos muito maiores que os elétrons de tal forma que os elétrons se adaptam instantaneamente ao movimento dos núcleos. Assim o átomo é a partícula menor do sistema e as suas propriedades dependem das respectivas distribuições nucleares, mas também da distribuição média dos seus elétrons.

A energia total de um sistema definido por mecânica molecular é calculada com base nas posições dos átomos constituintes e de certos parâmetros que validam a aplicação da equação ao sistema em estudo. Como exemplo, há o campo de forças AMBER que é descrito pela equação 3.1 passível de ser decomposta em cinco termos independentes.

$$E_{\text{total}} = E_{\text{ligação}} + E_{\theta} + E_{\omega} + E_{\text{vdW}} + E_{\text{el}} \quad (3.1)$$

$E_{\text{total}}$  é a energia total do sistema;  $E_{\text{ligação}}$  é a energia associada à variação dos ângulo das ligações;  $E_{\omega}$  é a energia associada à torção dos ângulos diedros;  $E_{\text{vdW}}$  é a energia associada às interações não ligantes independentes da carga;  $E_{\text{el}}$  é a energia associada às interações não ligantes devidas às cargas atômicas.

A energia de ligação pode ser descrita, de forma aproximada, pela lei de Hook (potencial harmônico), onde a ligação química é tomada como uma mola e portanto a energia associada depende do quadrado da elongação  $(l - l_{\text{equilíbrio}})^2$ , e da constante de força  $K_{\text{ligação}}$  (equação 3.2).

$$E_{\text{ligacao}} = \frac{1}{2} K_{\text{ligação}} (l - l_{\text{equilíbrio}})^2 \quad (3.2)$$

$l_{\text{equilíbrio}}$  corresponde ao comprimento da ligação na posição de equilíbrio e  $l$  ao comprimento da ligação num determinado instante.

A energia dos ângulos de ligação é também ela calculada segundo uma aproximação a um potencial harmônico (equação 3.3) tal como a energia de ligação.

$$E_{\theta} = \frac{1}{2} V_{\theta} (\theta - \theta_{\text{equilíbrio}})^2 \quad (3.3)$$

$E_{\theta}$  corresponde à constante de força;  $\theta_{\text{equilíbrio}}$  ao ângulo de equilíbrio entre três átomos; e  $\theta$  ângulo entre esses três átomos num determinado tempo.

Esta aproximação ao potencial harmônico falha quando o ângulo se afasta muito do ângulo na posição de equilíbrio, assim como em sistemas em que existe mais do que um valor para o ângulo no estado de equilíbrio (comum em compostos inorgânicos).

A energia de torção dos ângulos diedros caracteriza-se como sendo a variação de energia inerente à rotação de uma ligação química B-C ( $V_\omega$ ), onde os átomos B e C estão ligados a outros átomos, de tal forma que há o estabelecimento de um ângulo diedro entre os planos ABC (A está ligado a B) e BCD (D está ligado a C). Habitualmente, a energia associada à torção do ângulo diedro varia de forma senoidal (equação 3.4).

$$E_\omega = \frac{1}{2} \sum k_{diedros} (1 + \cos(n\omega - \gamma)) \quad (3.4)$$

$\omega$  corresponde ao ângulo de torção;  $k_{diedros}$  à barreira energética para que a rotação ocorra;  $n$  à multiplicidade (número de mínimos da função) e  $\gamma$  o ponto onde o ângulo de torção é mínimo.

A equação 3.5 descreve matematicamente o comportamento destas interações, designado de potencial de Lennard-Jones (CHIQUITO; ALMEIDA, 1999).

$$E_{VdW}(i, j) = 4\varepsilon_{ij} \left( \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right) \quad (3.5)$$

$i$  e  $j$  representam um par de átomos;  $\varepsilon_{ij}$  o mínimo de energia da função (normalmente obtido por ajuste de propriedades experimentais;  $r_{ij}$  a distância entre os dois átomos;  $\sigma_{ij}$  a distância entre os átomos à qual a energia é mínima;  $4\varepsilon_{ij} \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12}$  corresponde ao termo repulsivo e  $-4\varepsilon_{ij} \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6$  ao termo atrativo.

A energia eletrostática é o último termo da equação para a energia total do sistema e corresponde às interações não ligantes do tipo coulombiana, caracterizada pelas cargas dos átomos ou por dipolos permanentes. A energia destas interações é calculada com base na lei de Coulomb (equação 3.6).

$$E_{el}(i, j) = \frac{1}{4\pi\varepsilon} \frac{q_i q_j}{r_{ij}} \quad (3.6)$$

$i$  e  $j$  representam um par de átomos;  $\varepsilon$  a constante dielétrica do meio;  $r_{ij}$  a distância entre dois átomos,  $q_i$  e  $q_j$  as cargas dos átomos  $i$  e  $j$  respectivamente.

A mecânica molecular, neste caso usando o campo de forças AMBER, permite calcular a energia do sistema em cada ponto no tempo ( $U$ ) segundo a equação (equação 3.1 reorganizada matematicamente para facilitar o cálculo computacional por isolamento e cálculo prévio dos termos constantes).

$$U(\vec{R}) = \sum_{\text{ligações}} K_{\text{ligações}} (l - l_0)^2 + \sum_{\text{ângulos}} K_\theta (\theta - \theta_0)^2 + \sum_{\text{diedros}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \sum_{\text{não-ligantes}} \left( \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\varepsilon R_{ij}} \right) \quad (3.7)$$

$K_{ligacoes}$ ,  $K_\theta$  e  $V_n$  são constantes de força;  $A_{ij}$  e  $B_{ij}$  são os termos Van der Waals repulsivos e atrativos, respetivamente;  $R_{ij}$  é a distância entre dois átomos  $i$  e  $j$ ;  $q$  é a carga dos átomos;  $\varepsilon$  é a constante dielétrica do meio;  $l$ ,  $\theta$  e  $\omega$  representam respetivamente comprimento de ligação, ângulo de ligação e ângulo diedro; índice zero indica a respectiva propriedade do estado de equilíbrio.

### 3.3 Dinâmica molecular

Antes de se iniciar as simulações de DM, o sistema deve ser minimizado para eliminar maus contatos entre os átomos. A minimização de energia, é uma técnica que visa encontrar um conjunto de coordenadas que minimizam a energia potencial do sistema de interesse. O sistema minimizado possui forças pequenas sobre cada átomo e serve, portanto, como estrutura de partida para iniciar as simulações de DM.

Referimo-nos a cálculos clássicos como aproximações onde são válidas as leis de Newton. Para descrever corretamente os sistemas, devemos estabelecer os graus de liberdade que serão explicitamente tratados e aqueles que serão implicitamente tratados e incorporados à função de interação. Dentro de nosso esquema de cálculo, estamos interessados nos graus de liberdade de posição das coordenadas dos átomos. Num sistema clássico o hamiltoniano de um sistema molecular é descrito da seguinte forma:

$$H(p, r, m, g) = V(r, g) + K(p, m) \quad (3.8)$$

onde  $V$  é a energia potencial,  $K$  a energia cinética,  $p$  é o momento das partículas,  $r$  a posição,  $m$  a massa, e  $g$  são os parâmetros presentes no campo de forças.  $V(r, g)$  corresponde à energia potencial que descreve a interação em função das coordenadas das partículas. A energia cinética está representada da seguinte forma:

$$K(p, m) = \sum_{i=1}^N \frac{p_i^2}{2m_i} = \sum_{i=1}^N \frac{m_i v_i^2}{2} \quad (3.9)$$

dependente das massas e velocidades e independente das coordenadas  $r$  da partícula.

De uma forma geral, a energia potencial no campo de forças pode ser escrito da seguinte forma (SPOEL et al., 2005):

$$V(r, g) = V^{\text{ligados}}(r, g) + V^{\text{não-ligados}}(r, g) + V^{\text{especiais}}(r, g) \quad (3.10)$$

onde  $V^{\text{ligados}}$  corresponde aos termos chamados covalentes,  $V^{\text{não-ligados}}$  aos termos não ligados (forças eletrostáticas e de van der Waals) e  $V^{\text{especiais}}$  aos termos especiais, como, por exemplo, as restrições de posição e vínculos em comprimentos e ângulos de ligação.

A mecânica molecular nos permite minimizar a energia calculada a partir da equação 3.10 com o objetivo de obter as configurações do sistema no estado fundamental e

de reduzir as forças iniciais muito grandes que poderiam produzir trajetórias sem sentido físico. Mas atingir o mínimo absoluto é impraticável pelas numerosas conformações que uma proteína pode apresentar. O método *steepest-descent* (COURANT, 1943) permite encontrar o mínimo local mais próximo da superfície de energia potencial, já o método dos Gradientes Conjugados (FLETCHER; REEVES, 1964) encontra o mínimo de energia de maneira mais rápida, aproximando-se mais ao mínimo absoluto.

No método *steepest-descent*, os  $n + 1$  passos de minimização são feitos com base no cálculo das forças  $f(t_n)$  a partir da derivada da equação 3.10 para a configuração  $r(t_n)$  dada, e do cálculo da configuração seguinte  $r(t_{n+1})$  com um passo  $\Delta x$ :

$$r(t_{n+1}) = r(t_0) + N_n \Delta x f(t_n) \quad (3.11)$$

$N_n$  é um fator de normalização. Ao início do cálculo, um valor inicial de  $\Delta x$  é dado. No passos da simulação, se  $V(r)$  decresce, então  $\Delta x$  deve aumentar, se  $V(r)$  aumenta,  $\Delta x$  diminui. Assim, este método permite descer rapidamente para perto de um mínimo local, mas isso não assegura a convergência.

No método dos Gradientes Conjugados, a força (gradiente)  $f(t_{n+1})$  calculada em cada iteração é a conjugada da precedente para determinar o vetor unitário  $p(t_{n+1})$  da direção descendente sobre a hipersuperfície energética:

$$p(t_{n+1}) = f(t_{n+1}) + \beta_n p(t_n) \quad (3.12)$$

O passo de minimização  $s_{min}$  é escolhido levando em conta o gradiente da iteração precedente. Isto é o que permite a otimização do método com relação ao método *steepest-descent*. A nova configuração é obtida por:

$$r(t_{n+1}) = r(t_n) + s_{min} p(t_n) \quad (3.13)$$

O critério de convergência é feito sobre um número, a força máxima que atua por átomo. Geralmente uma minimização de energia é feita primeiro pelo método *steepest-descent* e depois pelo método dos Gradientes Conjugados.

A minimização da energia faz uso somente de uma pequena parte do espaço de configurações. Porém, pelos ajustes nas posições atômicas, ela relaxa as distorções nas ligações químicas, nos ângulos entre ligações, nos comprimentos de ligação e nos ângulos de torção (GOSS; PETRUCCI, 2007). Ela consiste em, primeiramente, relaxar (com restrição de posição) somente o solvente e evitar a sobreposição dos raios de van der Waals das moléculas de água. Em seguida, relaxar (sem restrição) o sistema soluto-solvente, ambos utilizando o algoritmo *steepest-descent*. Então, usa-se para todo o sistema, sem restrição de posição, o método Gradiente Conjugado.

A passagem da mecânica molecular para a dinâmica molecular baseia-se na derivada da energia calculada pela mecânica molecular em ordem à posição de cada átomo, que

resulta na força aplicada em cada átomo em cada instante (equação 3.14).

$$-\frac{dU}{dr} = F(r) \quad (3.14)$$

Pela integração das equações newtonianas do movimento é possível a partir da força calculada (equação 3.14) determinar as velocidades dos átomos e conseqüentemente as suas posições em cada instante. Contudo não é possível determinar uma solução analítica, pelo que tem de ser considerado um passo de integração finito ( $\Delta t$ ) de modo a obter valor numéricos para cada  $\Delta t$  considerado.

Assim, pela segunda lei de Newton calcula-se a aceleração em cada instante com base na massa de cada átomo:

$$\vec{a}(t_0) = \frac{\vec{F}(t_0)}{m} \quad (3.15)$$

Agora é possível determinar a posição de cada átomo no instante  $t_0 + \Delta t$ :

$$\vec{r}(t_0 + \Delta t) = \vec{r}(t_0) + \vec{v}(t_0)\Delta t + \frac{1}{2}\vec{a}^2(t_0)\Delta t^2 \quad (3.16)$$

E a sua respectiva velocidade:

$$\vec{v}(t_0 + \Delta t) = \vec{v}(t_0) + \vec{a}(t_0)\Delta t \quad (3.17)$$

Estes cálculos são repetidos para cada microestado. A duração de cada microestado,  $\Delta t$ , é suficientemente pequena para que a força se mantenha praticamente constante durante esse intervalo de tempo. Considerando esta aproximação, as trajetórias obtidas são tanto mais exatas quando menor for o intervalo de tempo  $\Delta t$ . Assim, teremos de adaptar os passos de integração de modo a estabelecer um compromisso entre o tempo de computação e a exatidão.

O método de integração das equações de movimento tem papel fundamental na conexão entre a mecânica clássica e a dinâmica molecular.

A integração das equações de movimento é realizada através de algoritmos baseados nos métodos das diferenças finitas, nos quais a integração é dividida em pequenos intervalos de tempo (passos de integração) (VERLET, 1967),  $\Delta t$ , permitindo simular os movimentos de maior frequência do sistema, que muitas vezes são as vibrações das ligações.

O primeiro passo é começar com um método simples, que é conhecido como método de Euler. Este método consiste em aproximar as equações por uma tangente.

Expandindo a posição e a velocidade até segunda ordem

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}\Delta t + O(\Delta t^2) \quad (3.18)$$

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \vec{v}\Delta t + O(\Delta t^2). \quad (3.19)$$

Reescrevendo as equações sabendo que  $\vec{r}(t) = \vec{v}(t)$  e  $\vec{v}(t) = \vec{F}(t)/m$

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + O(\Delta t^2) \quad (3.20)$$

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \frac{\vec{F}(t)}{m}\Delta t + O(\Delta t^2), \quad (3.21)$$

e para que a aproximação seja válida, é necessário que o incremento  $\Delta t$  seja muito pequeno. O erro durante o processo de integração é da ordem de  $\Delta t$ . A princípio, somos agora capazes de determinar a trajetória do sistema no espaço de fase. Este método porém, possui uma imprecisão significativa (LEITE et al., 2016). Verifica-se que, a medida em que evoluímos o sistema, algumas quantidades que deveriam ser conservadas sofrem variações, como as de conservação de energia e de amplitude de movimento. variação não está relacionada somente com a imprecisão do método, mas também com a ausência de propriedades das equações de movimento, uma vez que o método não foi construído utilizando os princípios da mecânica clássica (LEITE et al., 2016).

Temos que buscar integradores numéricos que respeitem essas propriedades das equações de movimento. O conjunto de algoritmos com essa característica é chamado de algoritmo simplético. Um exemplo de algoritmo simplético é o conhecido algoritmo de Verlet. Este é um dos algoritmos mais utilizados em dinâmica molecular, pois preserva as propriedades das equações de movimento clássica (LEITE et al., 2016).

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2 + \frac{1}{6}\vec{\dot{a}}(t)\Delta t^3 + O(\Delta t^4) \quad (3.22)$$

$$\vec{r}(t - \Delta t) = \vec{r}(t) - \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2 - \frac{1}{6}\vec{\dot{a}}(t)\Delta t^3 + O(\Delta t^4) \quad (3.23)$$

somando as duas equações, teremos

$$\vec{r}(t + \Delta t) = 2\vec{r}(t) - \vec{r}(t - \Delta t) + \vec{a}(t)\Delta t^2 + O(\Delta t^4) \quad (3.24)$$

onde  $\Delta t$  representa o incremento no tempo. O erro de truncamento deste algoritmo é da ordem de  $\Delta t^4$  e o valor típico do incremento é da ordem de  $\Delta t \approx 10^{-15}$ , resultando em um erro muito pequeno. Isto torna o algoritmo “estável”.

De acordo com a equação 3.24, nota-se que a velocidade não aparece explicitamente no algoritmo. Entretanto, de maneira geral, estamos interessados em calcular alguns observáveis que utilizam a velocidade, como temperatura, pressão ou energia cinética. Sendo assim, uma forma de obter a velocidade é subtrair as equações 3.22 e 3.23

$$\vec{v}(t) = \frac{\vec{r}(t + \Delta t) - \vec{r}(t - \Delta t)}{2\Delta t} \quad (3.25)$$

### 3.4 Método de Hartree-Fock

Não é trivial obter soluções para a equação de Schrodinger nos métodos quânticos. É necessário obter aproximações. Entre os mais populares em uso atualmente, é o método Hartree-Fock (HF) (ATKINS; FRIEDMAN, 2011). Além de ser capaz de oferecer uma boa solução aproximada para mais elétrons dentro de um sistema, o método de HF tem a vantagem de servir como um ponto de partida para outros métodos, como os semi-empíricos (SZABO; OSTLUND, 1996).

O método de HF parte do princípio de que a função de onda que descreve o sistema molecular pode ser descrita pelo determinante de Slater (LEVINE; BUSCH; SHULL, 2009), que pode ser representado na equação 3.26:

$$\phi_0 = \frac{1}{\sqrt{N!}} \begin{bmatrix} X_1(x_1) & X_2(x_1) & \cdots & X_N(x_1) \\ X_1(x_2) & X_2(x_2) & \cdots & X_N(x_2) \\ & : & & : \\ X_1(x_N) & X_2(x_N) & \cdots & X_N(x_N) \end{bmatrix} \quad (3.26)$$

onde os  $X$ 's são funções das coordenadas espaciais e do spin de um único elétron. Estas são chamadas de funções spin-orbitais moleculares. Admitindo-se que eles são ortonormais, o fator  $1/\sqrt{N!}$ , onde  $N$  é o número de elétrons do sistema, é uma constante de normalização de  $\phi_0$ . A simplificação introduzida pela aproximação do determinante único é a substituição do problema de se encontrar uma função de onda que depende das coordenadas de  $N$  elétrons pelo problema de encontrar  $N$  funções de onda de um elétron (as funções spins-orbitais moleculares).

Para sistema com mais de três corpos em movimento interagindo uns com os outros, utiliza-se aproximações. O método de HF é para iniciar o cálculo de uma função de onda tentativa, submetendo-o ao operador Fock (FREITAS, 2010).

No sistema eletrônico da molécula, para o estado fundamental, a solução aproximada da equação 3.26 é obtida pelo Princípio Variacional (SIMONS, 1991).

Como os elétrons de um determinado sistema não interagem entre si e estão sob influência do potencial efetivo  $\hat{V}_{HF}$ , pode-se dizer que o determinante de Slater é uma função de onda exata de um sistema de  $N$  elétrons, em um campo efetivo  $\hat{V}_{HF}$ .

$$\hat{H}_{HF}\psi_{SD} = E_{HF}\psi_{SD} \quad (3.27)$$

Como o operador Hamiltoniano  $\hat{H}_{HF}$  depende dos orbitais, a solução da equação 3.27 é obtida através de aproximações que ocorrem a partir de um processo denominado ciclo SCF (*Self Consistent Field* ou campo autoconsistente) (ATKINS; FRIEDMAN, 2011). Este ciclo consiste em escolher um conjunto de funções de base, propor os coeficientes iniciais para cada uma dessas funções, a partir de então, é feito o cálculo de energia.

O cálculo segue diminuindo a energia do sistema até que aconteça a convergência. A função de onda obtida é descrita como a mais próxima possível da molécula no seu estado fundamental para aquele conjunto de funções e pode ser utilizada para o cálculo de propriedades químicas desta molécula.

O método SCF é dado por equações que resolvem sistemas multieletrônicos considerando cada elétron como uma partícula que é independente de outros elétrons. Desta forma, um sistema de  $N$  elétrons é resolvido com a resolução de  $N$  equações, que representam cada elétron individualmente, até que a distribuição de carga não varie significativamente. Ao término dos procedimentos de cálculo, teremos as energias orbitais, que são as energias dos elétrons individuais resultantes das energias: cinética, de atração com os núcleos, de troca e de repulsão devido aos outros elétrons nas densidades de carga (CARDOSO et al., 2008).

### 3.5 Teoria do Funcional de Densidade

Na Teoria do Funcional Densidade (DFT) não mais se fala em funções de onda, mas em densidade eletrônica. Criada na década de 60, o DFT é um método que une uma extrema simplicidade a uma precisão notável. Com o uso de modernos computadores, é possível estudar sistemas complexos permitindo uma compreensão e predição das propriedades dos átomos, moléculas e sólidos. A formulação do método foi feita por Hohenberg, Kohn e Shan (KOHN; SHAM, 1965). Por esse trabalho, em 1998, foi atribuído o Premio Nobel de Química a Walter Kohn e John Pople.

Sendo uma alternativa ao método de HF, o DFT apresenta melhor concordância com os resultados experimentais e custo computacional reduzido devido a um menor número de integrais a serem resolvidas. Partindo do princípio de que a energia de um sistema de  $N$  elétrons pode ser expressa em termos de densidade eletrônica -  $\rho(\vec{r})$  - desse sistema. Diferente da função de onda, que é um ente matemático, a densidade eletrônica é um observável quântico.

Estes estudos começaram no início do século XX com a teoria dos gases em sólidos de Drude (MARQUES; BOTTI, 2006), sendo aperfeiçoados por Sommerfeld, Thomas, Fermi e Dirac. Este utilizou pela primeira vez o conceito de funcional, que é uma regra que faz corresponder a uma função um número, ou seja, uma função de uma função. Ainda acrescentou, em 1928, um funcional de energia de troca para um gás de elétrons. Porém todos estes modelos são muito simples e imprecisos para reproduzir a estrutura eletrônica de camadas dos átomos ou ligações químicas.

A densidade eletrônica foi tratada como variável básica num artigo de Waler Kohn e Pierre Hohenberg, em 1964. Minimizando a energia que é um funcional de densidade (probabilidade de encontrar um elétron num ponto de coordenada), obtém-se a energia

do estado fundamental do sistema. Este formalismo é sustentado por dois teoremas de Hohenberg e Kohn (HK) que fornecem os fundamentos da DFT.

O primeiro teorema de HK, dos dois que formam a base do DFT, estabelece que o potencial externo deveria ser um funcional de  $\rho(\vec{r})$  além de uma constante de adição. Ou seja, ele demonstra que a densidade eletrônica de um dado sistema determina o potencial externo, o número de elétrons e o Hamiltoniano desse sistema.

Como a energia do sistema é determinada a partir da resolução da equação de Schrödinger,  $\hat{H}\psi = E\psi$ , ela pode ser determinada pela densidade eletrônica:

$$E = E_V[\rho] \quad (3.28)$$

Neste caso,  $V$  indica a dependência com o potencial externo  $V(r)$ .

O segundo teorema de HK afirma que a densidade eletrônica pode ser obtida pelo método variacional onde, a partir de uma densidade eletrônica arbitrária pode-se variá-la de forma sistemática sucessivamente até ela atingir um mínimo de energia. Neste caso, a densidade eletrônica seria a do estado fundamental. Em outras palavras, havendo qualquer aproximação da densidade eletrônica, a energia final será sempre maior ou igual à energia exata do sistema. Sendo assim:

$$F[\rho] = \langle \psi | \hat{T} + \hat{V}_e | \psi \rangle \quad (3.29)$$

onde  $F$  é definido como funcional universal e onde  $\hat{T}$  e  $\hat{V}_e$  aplicam-se a todos os sistemas eletrônicos.

A fim de descrever as camadas eletrônicas determinando a densidade do sistema em interação, Walter Kohn e Lu Shan idealizaram um sistema eletrônico fictício, com partículas independentes, que não interagissem entre si, onde sua densidade fosse igual à densidade do sistema original. Para isso, os elétrons são “mergulhados” nesse sistema imaginário num potencial efetivo, o potencial de Kohn-Shan (KS), escolhido de forma a satisfazer essa condição. Os elétrons, sendo independentes, obedeceriam a uma equação de Schrödinger, a chamada equação KS, muito semelhante à equação para o átomo de hidrogênio (H). A diferença básica é a troca do potencial criado pelo núcleo do átomo de H, pelo potencial de KS que também é um funcional de densidade,  $V_{KS}$ . A expressão pode ser escrita como sendo:

$$V_{KS}[\rho](\vec{r}) = V_{ext}(\vec{r}) + V_{Hartree}[\rho](\vec{r}) + V_{XC}[\rho](\vec{r}) \quad (3.30)$$

onde primeiro termo representa o potencial externo, criado pelos núcleos atômicos, o segundo é o potencial de Hartree, também presente na teoria de Thomas-Fermi leva em consideração a repulsão eletrostática entre os elétrons, ou seja, a interação entre o elétron e a densidade de carga do sistema. Por último, o potencial de troca-correlação,  $V_{XC}$  inclui

todos os efeitos não triviais da interação, isto é, engloba todas as correções entre os dois primeiros termos.

Uma boa aproximação para esse potencial é chamada de aproximação de densidade local (LDA - *Local Density Approximation*) (PARR; YANG, 1989). Nela, KS, basearam suas ideias a partir do gás homogêneo de elétrons, em um sistema onde os núcleos são substituídos por uma carga positiva e constante (KOHN; SHAM, 1965), obtendo um custo computacional menor e se equivalendo à aproximação obtida do método HF para átomos e moléculas. Para estudos que envolvem o estado sólido, a LDA fornece resultados melhores que os demais métodos e tornou-se um padrão para estudos das propriedades mecânicas, eletrônicas, etc. Posteriormente, foram propostas inúmeras aproximações, entre elas a aproximação do gradiente generalizado (GGA - *Generalized Gradient Approximation*) (PARR; YANG, 1989), que consiste em incluir correções do gradiente da densidade eletrônica no funcional, em que se leva em consideração a não homogeneidade local da densidade eletrônica. O seu procedimento é mais preciso, sendo um método eficiente para cálculos que envolvem complexos metálicos. No entanto, para as interações do tipo Van der Waals, uma boa parte das aproximações GGA não fornecem resultados satisfatórios. Por isso, existem muitas modificações das aproximações LDA e GGA, assim como aproximações mistas. Entre os funcionais de troca GGA mais importantes e populares são devido a Perdew e Becke (BECKE, 1983) e os mais importantes, os propostos por Lee, Yang e Parr (LYP)(LEE; YANG; PARR, 1988). O funcional híbrido B3LYP é amplamente usado nos cálculos teóricos envolvendo sistemas moleculares e pode ser justificado pelo bom acordo dos resultados obtidos na parte da otimização geométrica com dados experimentais.

## 3.6 Conjunto de Funções de Base

Um conjunto de funções de base (GODBOUT et al., 1992) consiste na representação das funções os spins-orbitais que são usados para criar orbitais moleculares, que se expandiram a partir da combinação linear de tais funções. Cada orbital representa uma combinação linear de um ou mais orbitais do tipo Slater (STO - *Slater Type Orbital*) ou do tipo Gaussianas (GTO - *Gaussian Type Orbital*).

As STOs são funções que representam as principais características da parte radial de um orbital atômico:

$$STO = \left(\frac{\zeta^a}{\pi}\right)^5 e^{-(\zeta r)} \quad (3.31)$$

onde  $r$  é o raio e  $\zeta$  (sigma) representa o valor da função de base. A maior dificuldade dessas funções está relacionada ao cálculo de suas integrais. Outro fato está relacionado às moléculas poliatômicas em que o custo computacional aumenta devido à inclusão de integrais mais complexas tornando-a indesejável.

As GTOs, por sua vez, são funções usadas em combinações lineares que são utilizadas nos cálculos de sistemas moleculares.

$$GTO = \left(\frac{2\alpha}{\pi}\right)^{0,75} e^{(-\alpha r^2)} \quad (3.32)$$

Esta expressão é muito similar aos STO. A diferença está no termo  $\alpha$  usado no lugar de  $\varsigma$ . A principal razão do uso da GTO como função de base é que ela garante que duas funções em torno de dois átomos diferentes são uma soma finita de gaussianas centradas em um ponto ao longo do eixo de ligação delas, o que reduz o custo computacional.

São encontrados vários tipos de funções de base Mínima, *Split-Valence*, *Double-Zeta*, *Triple-Zeta*, Polarizada e Difusa. São essas funções que ajudam no procedimento computacional dos cálculos das propriedades moleculares.

O primeiro tipo, base mínima (*Minimal*), foi concebido com o objetivo de acelerar os cálculos e são pouco utilizados para fins de estudo. Elas são usadas principalmente para obter as propriedades moleculares. São exemplos as funções de base STO-3G e STO-6G. A letra G representa o número de gaussianas usadas.

O conjunto de base *Split-Valence* considera que os elétrons da camada de valência estão envolvidos nas ligações e reações químicas, ao contrário dos elétrons do núcleo. Esse conjunto é formado pelos orbitais mais externos dos átomos que participam das ligações químicas. O *Split-Valence* possui duas subdivisões: o *double* (duplicado), onde cada orbital de valência é descrito por duas funções, por exemplo, 3-21G e 6-31G, e o *triple* (triplicado), em que cada orbital é descrito por três funções de base, como por exemplo, o 6-311G. Elas descrevem cada orbital atômico ou molecular como uma combinação linear de duas ou mais funções de base, respectivamente para melhorar algumas das propriedades físicas e químicas do sistema em estudo.

Os conjuntos de funções de base polarizada (*Polarized*) são aqueles que permitem uma melhor descrição dos elétrons de um orbital. É através desse conjunto que se consegue uma melhor descrição das deformações sofridas pelos orbitais que participam das ligações químicas. O termo polarização nesse caso refere-se a uma correção da distorção dos orbitais atômicos que ocorre na formação de espécies poliatômicas, ou seja, perda de simetria. Com isso pode-se calcular de maneira satisfatória algumas propriedades químicas. Dentre os conjuntos de bases com funções de polarização temos STO-3G\*, 3-21G\*, 6-31G\*\* e 6-311G\*. O asterisco (\*) indica que as funções de polarização de simetria d estão correlacionadas aos átomos pesados, e a presença de dois asteriscos significa que se adicionaram funções de polarização de simetria p aos átomos de H.

As funções difusas descrevem uma região maior do espaço dos orbitais. Geralmente os elétrons se encontram em maior quantidade próximos ao núcleo do átomo. Em distâncias maiores, a probabilidade de se encontrar um elétron diminui, pois o raio atômico aumenta. Para alguns sistemas, no entanto, usa-se uma base difusa para estender e melhorar a

descrição de sistemas com densidades eletrônicas significativas à longas distâncias. Elas são aplicadas para descrever sistemas metálicos, onde os átomos metálicos têm simetria  $d$ , com característica difusa. Esses conjuntos de funções de base são representados por um  $+$ . Têm-se como exemplos as funções 3-21+G, 6-31+G\* são formados por quatro funções altamente difusas ( $s$ ,  $p_x$ ,  $p_y$  e  $p_z$ ) para cada átomo pesado. Já os conjuntos 3-21++G e 6-31++G\* incluem uma função  $s$  altamente difusa em cada átomo de H.

Quanto mais completa a base, melhores são os resultados. Porém, o simples aumento no número de funções de base não necessariamente resulta em melhoria nos resultados dos cálculos e das propriedades dos sistemas moleculares.

### 3.7 O método RESP

Como resultado de um cálculo quântico, nós obtemos o potencial eletrostático (ElectroStatic Potential - ESP). É a partir do ESP que se gera a distribuição das cargas nas moléculas.

Existem diferentes metodologias para gerar cargas através do ESP, sendo que entre as mais bem sucedidas estão:

- O método CHELPG (CHarges from ELectrostatic Potential using a Grid based method, (BRENEMAN; WIBERG, 1990);
- O método RESP (Restrained Electrostatic Potential) (BAYLY et al., 1993).

No caso de sistemas nos quais existem átomos fortemente blindados (chamados de átomos ocultos) como o ferro no centro do anel porfirínico, o método CHELPG falha. Entre as falhas do método CHELPG estão a superestimação ou subestimação das cargas de átomos ocultos e a variação da distribuição das cargas com a conformação da molécula.

Já o método RESP diminui a variabilidade da distribuição das cargas com a conformação e estima melhor as cargas dos átomos ocultos. Este método envolve uma aproximação em dois níveis de cálculo, onde as cargas dos átomos como os hidrogênios metil são forçadas a ser equivalentes até o segundo nível do cálculo. Neste ponto estas cargas são reajustadas, enquanto que as cargas dos outros átomos são mantidas nos valores do primeiro nível (BAYLY et al., 1993).

Os resultados baseados no método RESP dão valores concordantes com a energia conformacional para pequenas moléculas e também reproduzem corretamente as energias livres de solvatação (WANG; CIEPLAK; KOLLMAN, 2000).

### 3.8 RMSD

O RMSD (desvio da raiz média quadrática - do inglês *Root Mean Square Deviation*) permite comparar conformações de um dado conjunto de pontos, fornecendo uma medida da diferença na posição desses mesmos conjuntos. Quanto maior for o valor de RMSD, maior será a diferença entre as estruturas. Seu objetivo é verificar se as estruturas geradas na dinâmica já se encontram em equilíbrio ou não, fazendo uma comparação entre todas as estruturas obtidas durante o processo de simulação (VERLI, 2009). No geral, para um conjunto de átomos idênticos presentes em duas conformações diferentes, a análise por RMSD fornece uma medida da diferença na posição desses dois conjuntos de átomos. Considerando dois conjuntos de  $n$  pontos  $v$  e  $w$ , o RMSD é dado por:

$$RMSD(v, w) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( (v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2 \right)} \quad (3.33)$$

fornecendo o desvio padrão entre os mesmos átomos na conformação  $v$  e na conformação  $w$ . É expresso em unidades de comprimento, tipicamente Å ( $10^{-10}$ m).

De acordo com (VERLI, 2009), a comparação de estruturas tridimensionais se dá da seguinte forma:

- Estruturas idênticas:  $RMSD = 0$  Å.
- Estruturas similares:  $0$  Å  $<$   $RMSD <$   $3$  Å.
- Estruturas distantes:  $RMSD >$   $3$  Å.

### 3.9 RMSF

A análise de RMSF (flutuação da raiz média quadrática - do inglês *Root Mean Square Fluctuation*), é de grande utilidade em dinâmica molecular. Esta propriedade permite identificar as partes do sistema (resíduos / átomos) cuja posição varia mais em torno da sua posição na estrutura em equilíbrio no decurso de uma dinâmica molecular. Por outras palavras identifica as zonas mais flexíveis do sistema. O uso da análise do RMSF juntamente com a de RMSD permite identificar o grau de mudança conformacional, flexibilidade e equilibração para cada região da macromolécula em estudo (REIS, ).

Para uma partícula  $i$  e um tempo em termos de número de passos de simulação,  $n_{sim}$ , o valor de RMSF é dado por

$$RMSF(x) = \sqrt{\frac{1}{n_{sim}} \sum_{n=1}^{n_{sim}} (x_i - \tilde{x}_i)^2} \quad (3.34)$$

onde  $x_i$  é a posição da partícula no instante de tempo  $t$ , enquanto  $\tilde{x}_i$  a posição média da partícula  $i$  durante a simulação.

## 3.10 Energia livre de ligação

A energia livre é muito importante para o entendimento de processos de solvatação, descrevendo a tendência à associação e a espontaneidade do processo. A capacidade de prever o valor dessa propriedade usando modelos de química teórica é de grande valia, embora a obtenção de estimativa mais acurada envolva um custo computacional muito elevado (GENHEDEN et al., 2010).

Segundo (CLAUSIUS, 1865), a desordem dos componentes de um sistema químico é expressa como entropia, denotada por  $S$ . Qualquer mudança da desordem do sistema é expressa como variação da entropia ( $\Delta S$ ), a qual, por definição, tem um valor positivo quando a desordem aumenta. J. Willard Gibbs, que desenvolveu a teoria da variação de energia durante reações químicas (GIBBS, 1878), mostrou que o índice de energia livre,  $G$ , de qualquer sistema fechado pode ser definido por três quantidades: entalpia  $H$ , refletindo o número e tipos das ligações; entropia,  $S$ ; e a temperatura absoluta,  $T$  (em Kelvin). A definição da energia livre é

$$G = H - TS \quad (3.35)$$

A energia livre é usualmente expressa como a função de Helmholtz ( $A$ ) ou a função de Gibbs ( $G$ ). A energia livre de Helmholtz é apropriada para um sistema com número de partículas, temperatura e volume constantes. A energia livre de Gibbs é apropriada para sistemas onde o número de partículas, a temperatura e a pressão são constantes (LEACH, 2001). As reações ocorridas em nosso organismo acontecem em condições de temperatura e pressão constantes, onde a função de Gibbs é a quantidade de energia livre apropriada.

A variação de energia livre de Gibbs,  $\Delta G$ , é determinada pela mudança de entalpia,  $\Delta H$ , refletindo os tipos e os números de ligações químicas e de interações não-covalentes quebradas e formadas, e pela variação da entropia,  $\Delta S$ , descrevendo a mudança na desordem do sistema:

$$\Delta G = \Delta H - T\Delta S \quad (3.36)$$

A tendência de uma reação química ocorrer pode ser expressa como uma constante de equilíbrio, dada por

$$K_{eq} = \frac{[AB_{eq}]^c}{([A_{eq}]^a[B_{eq}]^b)} \quad (3.37)$$

onde  $[A_{eq}]$  é a concentração de  $A$ ,  $B_{eq}$ , a concentração de  $B$ , e assim por diante, quando o sistema alcança o equilíbrio. Para uma reação onde ocorre a interação de um ligante com um receptor, os reagentes são as duas moléculas livres e o produto é o complexo ligante-proteína. Um valor alto de  $K_{eq}$  significa que a reação tende a prosseguir até que os reagentes estejam quase que completamente convertidos em seus produtos.

Sendo  $\Delta G^0$  uma constante que é característica de cada reação,  $\Delta G$  para qualquer reação química é uma função da mudança padrão da energia livre e um termo que expressa

as concentrações iniciais dos reagentes e dos produtos:

$$\Delta G = \Delta G^0 + RT \ln \frac{[AB_i]^c}{[A_i]^a [B_i]^b} \quad (3.38)$$

onde  $[A_i]$  é a concentração inicial de  $A$ , e de maneira equivalente, para outros elementos;  $R$  é a constante universal dos gases,  $1,987 \text{ cal K}^{-1} \text{ mol}^{-1}$  e  $T = 298,15 \text{ K}$

Quando uma reação alcança o equilíbrio, temos  $\Delta G = 0$ . Para este caso especial,  $[A_i] = [A_{eq}]$ , e assim por diante, para todos os reagentes e produtos, e

$$\frac{[AB_i]^c}{[A_i]^a [B_i]^b} = \frac{[AB_{eq}]^c}{([A_{eq}]^a [B_{eq}]^b)} \quad (3.39)$$

Fazendo as devidas substituições, obtemos a relação

$$\Delta G^0 = -RT \ln K_{eq} \quad (3.40)$$

que é definida pela reação de associação,  $E + I \rightleftharpoons EI$  (GENHEDEN; RYDE, 2015), onde  $E$  é a enzima e  $I$  o inibidor. No entanto, como as constantes de inibição ( $K_i$ ) estão relacionadas com o processo oposto, de dissociação entre o ligante e o receptor,  $EI \rightleftharpoons E + I$ , o cálculo do  $\Delta G$  perde o sinal negativo passando a ser (MORRIS et al., 1998)

$$\Delta G^0 = RT \ln K_i \quad (3.41)$$

Para investigar a contribuição energética de afinidades de ligação proteína-ligante, utiliza-se alguns métodos, dos quais se destacam o MMPBSA (Molecular Mechanics Poisson-Boltzmann Surface Area) e o MMGBSA (Molecular Mechanics Generalized Born Surface Area) (KOLLMAN et al., 2000). MMPBSA usa a equação de Poisson-Boltzmann para calcular a contribuição eletrostática para a energia livre; o MMGBSA usa a aproximação generalizada de Born, que é um tratamento aproximado e mais rápido da equação de Poisson-Boltzmann.

MMPBSA e MMGBSA são duas técnicas estreitamente relacionadas para calcular diferenças aproximadas de energia livre em sistemas (bio)moleculares. Geralmente, são usados para estimar a energia livre de ligação de um ligante a uma proteína. Esses métodos representam um meio-termo entre o *docking* molecular, que é rápido mas muito impreciso, e a energia de interação solvatada (SIE), que oferece os melhores resultados à custa de um custo computacional um pouco maior.

Em MMPBSA e MMGBSA, a energia livre é avaliada como uma soma de um termo de energia conformacional (a parte MM - mecânica molecular) e um termo de solvatação sem energia que é calculado usando eletrostática contínua (tornando estes métodos abordagens implícitas de solvatação). MM refere-se ao tipo de função de energia usada para calcular a energia potencial de uma estrutura molecular. Essas funções, geralmente chamadas

de campos de força, são potenciais clássicos, incluindo termos que descrevem a ligação covalente, as interações de Van der Waals, etc. A outra parte, o termo de solvatação, pode ser adicionalmente escrito como a soma de um componente polar e uma contribuição não polar. Supõe-se que este último dependa linearmente da área de superfície acessível ao solvente (SA).

Em relação à parte polar, pode ser rigorosamente calculada usando a equação de Poisson-Boltzmann (PB). Esta equação dá o potencial eletrostático em torno de um soluto em uma solução. É uma equação diferencial parcial bastante complexa e, como tal, é computacionalmente caro de resolver. A aproximação generalizada de Born (GB) é uma solução aproximada de PB que é mais rápido para calcular, obtendo valores mais precisos em relação ao PB. Assim os métodos de MMPBSA e MMGBSA são eficientes para o estudo de desenvolvimento de fármacos (KOLLMAN et al., 2000).

Em termos de energia livre de ligação ( $\Delta G_{\text{ligação}}$ ), para cada conformação do inibidor da proteína é calculada por meio da equação 3.42:

$$\Delta G_{\text{ligação}} = (G_{\text{complexo}}) - (G_{\text{proteína}}) - (G_{\text{ligante}}) \quad (3.42)$$

A energia livre de ligação é estimada como uma soma de três termos:

$$\Delta G_{\text{ligação}} = \Delta G_{MM} + \Delta G_{\text{solv}} - T\Delta S \quad (3.43)$$

onde,  $\Delta G_{MM}$  é a energia da mecânica molecular dos reagentes em fase gasosa;  $\Delta G_{\text{solv}}$  é a energia livre de solvatação e é calculada com uma contínua representação do solvente para a parte polar;  $T\Delta S$  é a variação de entropia conformacional associada ao ligante.

A energia livre da mecânica molecular ( $\Delta G_{MM}$ ) é ainda dividida em energia de van der Waals ( $\Delta G_{vdw}$ ) e energia eletrostática:

$$\Delta G_{MM} = \Delta G_{ele} + \Delta G_{vdw} \quad (3.44)$$

### 3.11 Decomposição da energia livre por resíduo

O pacote de programas AMBER (Construção Assistida de Modelos com Refinamento de Energia - do inglês *Assisted Model Building with Energy Refinement*) (CASE et al., 2012) é um conjunto de programas que reúne diversos campos de força de mecânica molecular, bastante conhecido e utilizado para realizar simulações de dinâmica molecular. É um pacote de aproximadamente 50 programas, no qual inclui programas licenciados e livres. Ele fornece vários protocolos para decompor as energias livres calculadas em contribuições específicas dos resíduos, usando modelos de solventes implícitos GB ou PB (METZ et al., 2011; GOHLKE; KIEL; CASE, 2003), segundo o trabalho de (GOHLKE; KIEL; CASE, 2003). As interações podem ser decompostas para cada resíduo, incluindo

apenas as interações nas quais um dos átomos do resíduo está envolvido, é a decomposição por resíduo. Alternativamente, as interações podem ser decompostas por pares de resíduos específicos, incluindo apenas as interações nas quais um átomo de cada um dos resíduos analisados está participando, é a chamada decomposição em pares. Esses protocolos de decomposição podem fornecer informações importantes nos cálculos de energia livres (GOHLKE; KIEL; CASE, 2003).

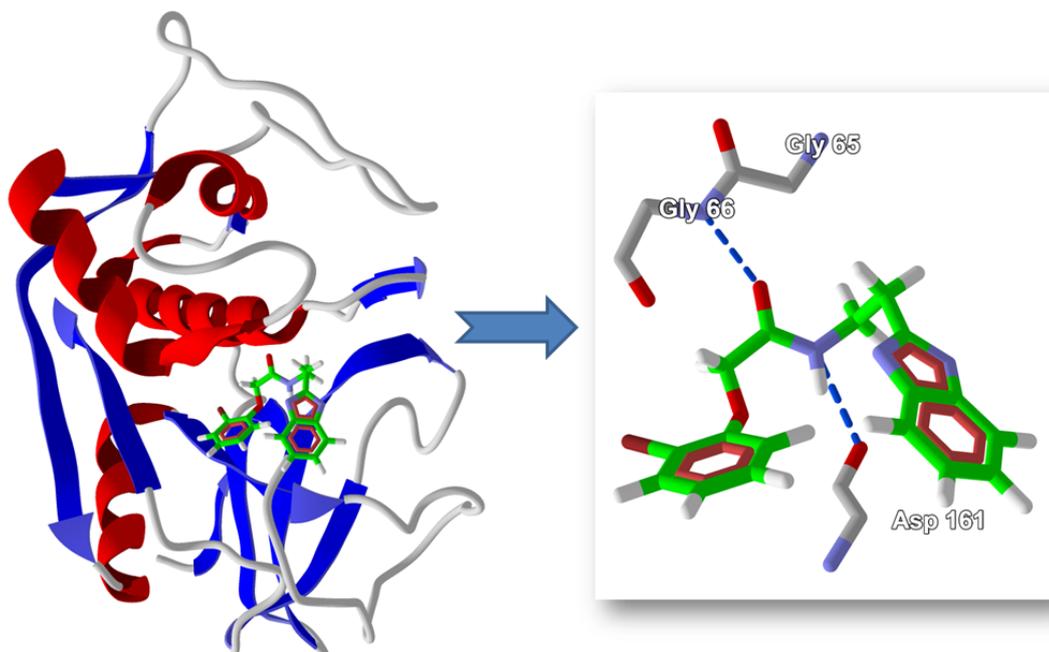
## 3.12 Protocolo para as simulações

### 3.12.1 Obtenção da estrutura inicial e preparação do arquivo PDB

A fim de estudar as estruturas da enzima cruzaina de *T. cruzi*, foram utilizadas e analisadas informações de estruturas cristalográficas depositadas no Protein Data Bank (PDB) (BERMAN et al., 2006) - IDs: 3HD3, 2OZ2, 1U9Q, 1EWL, 1EWM, 1EWO, 1ME3, 1ME4, 1F29, 1F2A, 1F2B, 1F2C, 1EWP, 1AIM, 2AIM, 3IO6, 3IUT, 3KKU e 3LXS - com o auxílio dos programas de visualização molecular *Pymol* (DELANO, 2002) e Chimera (PETTERSEN et al., 2004). As interações entre os complexos receptor-ligante também foram analisadas pela ferramenta “Ligand Explorer”, disponível no PDB. Esta ferramenta permite observar as principais interações entre proteína e inibidor (ligações de hidrogênio, interações hidrofóbicas, entre outras).

Após análise detalhada de todos os complexos, a estrutura de código 3KKU (FERREIRA et al., 2010) (Figura 8) foi selecionada para os estudos. Esta estrutura apresenta alta resolução (1,28 Å) e o ligante em complexo faz apenas interações não covalentes com a proteína.

Figura 8 – À esquerda, o complexo proteína-ligante, onde a enzima está representada em modelo de fitas e o ligante em bastão. À direita, representação das interações da glicina e do aspartato com o inibidor (bastão). As interações de hidrogênio estão representadas pelas linhas tracejadas.



Fonte: o próprio autor.

A estrutura foi preparada removendo todas as linhas correspondentes ao cabeçalho do arquivo PDB, as moléculas de água, o ligante e as duplas conformações dos resíduos da estrutura. Foram adicionados átomos de hidrogênio e checado o estado de protonação dos resíduos His162 (forma protonada) e Cys25 (forma desprotonada) (FERREIRA et al., 2010; FERREIRA et al., 2011; MLADENOVIC et al., 2008). Para checar o estado de protonação, a estrutura foi submetida ao servidor online PDB2PQR Server (DOLINSKY et al., 2004) para realizar o cálculo de pKa.

Também foram analisados os resíduos de cisteína, pois estes podem apresentar-se em forma de ponte dissulfeto, que é uma ligação covalente simples resultante da junção de dois átomos de enxofre (S). A metionina, o outro aminoácido que contém enxofre, é incapaz de formar este tipo de ligação. As pontes dissulfeto são muito importantes em uma enzima, pois são capazes de manter rígida/estável a sua estrutura proteica que apresenta *loops* seguros por estas interações.

### 3.12.2 Otimização geométrica e cargas RESP dos ligantes

O ligante 8D foi separado do arquivo PDB utilizando o programa Chimera. Utilizando o programa Avogadro (HANWELL et al., 2012), as estruturas 8K, 8L e 8R foram desenhadas, plotadas as geometrias iniciais e, então, especificou-se os cálculos de

otimização geométrica e cargas RESP por meio do *Gaussian 03* (FRISCH et al., 2009), que é um programa capaz de prever muitas propriedades de átomos, moléculas e sistemas reativos, isto é, energias, estruturas e frequências vibracionais. Para estes ligantes, a otimização foi realizada com método B3LYP e conjunto de base 6-311G e duas funções polarizadas (\*\*). Após a otimização as estruturas foram submetidas ao cálculo de carga RESP com método Hartree-Fock e conjunto de base 6-31G com uma função polarizada (\*).

### 3.12.3 Preparação de arquivos de parâmetros, topologia e coordenadas

Para iniciar os cálculos de trajetórias, é preciso preparar os arquivos necessários. Para isto, foi carregado no pacote *tleap*, que faz parte do conjunto de programas AMBER, o arquivo PDB juntamente com as bibliotecas para os diversos resíduos de aminoácidos, ff99SB (MAIER et al., 2015); GAFF (*General Amber Force Field*) (WANG; CIEPLAK; KOLLMAN, 2000) que contempla toda a parametrização relativa às ligações, ângulos, diedros, tipos de átomos e interações não ligantes; e os arquivos de parâmetros *.frcmod* e *ligand\_resp.mol2* específicos para os ligantes 8D, 8K, 8L e 8R.

Foi realizada a neutralização das cargas por adição de íons sódio (positivos) para que a carga global do sistema seja nula. No sistema em estudo, resultou a adição de 11 íons sódio, uma vez que o inibidor não é eletricamente carregado e a cruzaina possui um excesso de 11 cargas negativas. Ainda com o *tleap*, procedeu-se à adição de uma 'caixa' de moléculas de água de solvatação do tipo TIP3P (JORGENSEN; JENSON, 1998) (o qual representa a entalpia de vaporização e constante dielétrica da água) em formato octaédrico até uma distância mínima de 9 Å em todas as direções relativas ao complexo proteína-inibidor. A partir de então, foram gerados os arquivos de topologia e de coordenadas, respectivamente, *.top* e *.crd*. Este último arquivo contempla as coordenadas atômicas e as coordenadas da caixa de simulação (para condições periódicas de fronteira), adicionalmente poderá conter velocidades no caso de dinâmicas moleculares.

### 3.12.4 Minimização, termalização e produção

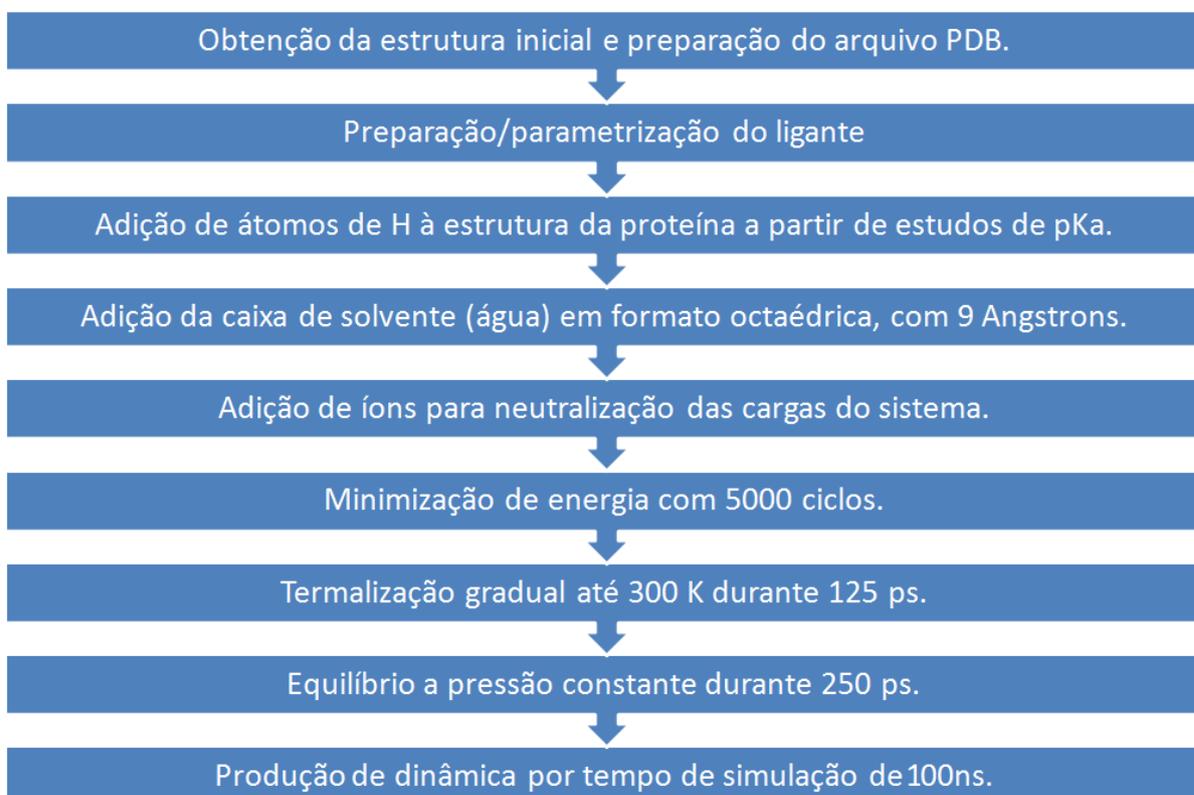
Antes do cálculo das simulações é preciso fazer a minimização energética do sistema. Contudo o processo pretende apenas relaxar as ligações e interações intermoleculares, principalmente as que envolvem os hidrogênios e as moléculas de água adicionadas computacionalmente com o *tleap*. Esta minimização é feita de forma sequencial. No primeiro passo foram minimizadas apenas as geometrias e posições das moléculas de água de solvatação adicionadas, mantendo todo o resto estático. Seguiu-se uma minimização dos átomos de hidrogênio, também adicionados com o *tleap*, ausentes na estrutura inicial do PDB. Por fim, uma última minimização de todos os átomos do sistema. Todas as minimizações englobaram um total de 5000 ciclos, dos quais os primeiros 3000 recorreram ao algoritmo

*steepest-descent* e os últimos, ao algoritmo de Gradiente Conjugado que é mais eficaz para valores de energia potencial próximos do mínimo. Durante as minimizações definiu-se um sistema periódico a volume constante com um *cut-off* para interações intermoleculares de 9,0 Å. Após cada minimização é gerado um arquivo *.rst* que contempla as coordenadas atômicas pós-minimização, sendo usado como coordenadas na minimização seguinte. O último arquivo *.rst* gerado servirá de base ao cálculo subsequente da dinâmica molecular.

Feita a minimização, o sistema foi gradualmente aquecido de 0 para 300 K, em seis passos (0-50, 50-100, 100-150, 150-200, 200-250 e 250-300 K) durante 25 ps (picosegundos) cada, totalizando 175 ps, sempre mantendo o *restraint* (restrição) no sítio ativo. A velocidade de cada passo de integração foi reescalada de acordo com a distribuição de Maxwell-Boltzmann, a fim de ajustar a energia cinética do sistema até a temperatura final. Este período é chamado de termalização e foi realizada a volume constante.

Com o sistema já termalizado, foram retiradas as restrições e a pressão foi alterada para constante antes de executar mais 250 ps de equilíbrio a 300 K. Na produção de dinâmica foi utilizado o campo de força AMBER e tempo de simulação de 100 ns (nanossegundos) com auxílio da ferramenta *Sander*, que faz parte do conjunto de programas AMBER. A Figura 9 mostra o resumo do protocolo utilizado neste trabalho.

Figura 9 – Resumo do protocolo seguido para a realização da simulação.





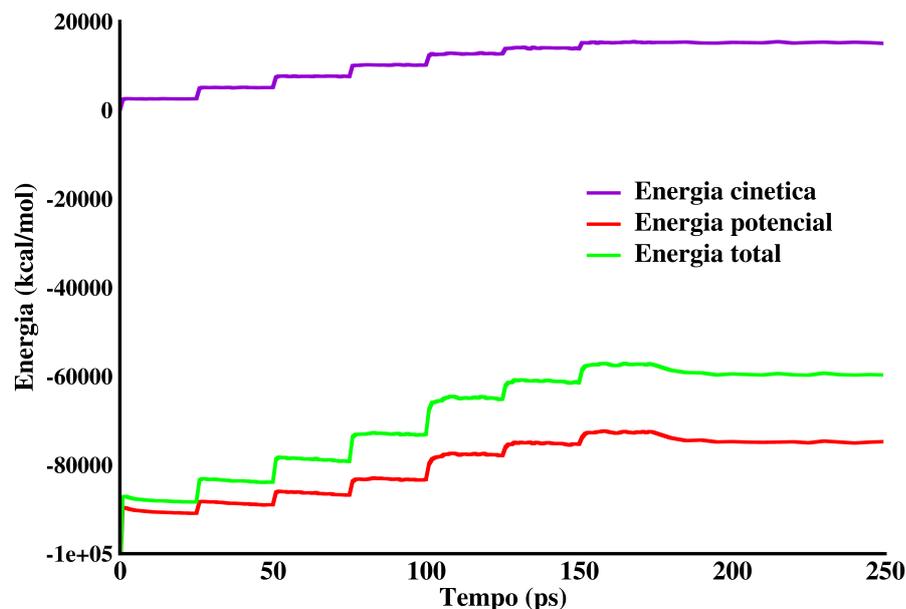
## 4 Resultados

### 4.1 Energias, temperatura, volume e densidade

Com o objetivo de analisar a qualidade dos resultados das etapas de termalização e equilíbrio, foram gerados os gráficos das principais propriedades do sistema, que são: as energias total, potencial e cinética, a temperatura, a densidade e o volume. As informações necessárias para a criação dos gráficos foram extraídas com auxílio da ferramenta *ptraj* (ROE; III, 2013) do AMBER, e os gráficos gerados no *xmgrace* (TURNER, 2005). Estes gráficos representam as propriedades das simulações dos sistemas e são comuns aos quatro ligantes.

As energias potencial, cinética e a total, aumentaram inicialmente e depois se estabilizaram durante a fase de volume constante. Após isto, ocorreu uma pequena diminuição à medida que o sistema vai relaxando quando se tira as restrições e muda-se à pressão constante. As energias, em seguida, permaneceram constante durante o resto da simulação, indicando que o relaxamento foi concluído e que chegou-se a um equilíbrio. Estes comportamentos estão evidenciados na Figura 10.

Figura 10 – Representação gráfica da variação da energias total, cinética e potencial dos sistemas em função do tempo.

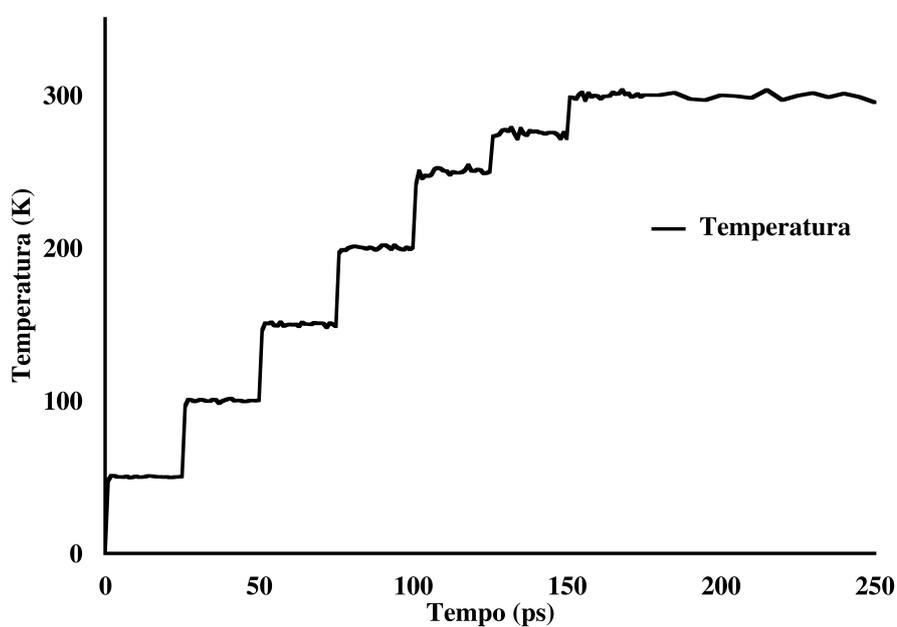


Fonte: o próprio autor.

Verifica-se que nos primeiros 175 ps de termalização há aumento gradativo da temperatura até 300K. A partir de então, há pouca variação em torno desta temperatura,

permanecendo aproximadamente constante ao longo de toda a simulação, Figura 11. Como a temperatura é uma propriedade do sistema que reflete a energia cinética total do mesmo, a variação desta energia se comporta de modo semelhante ao da temperatura.

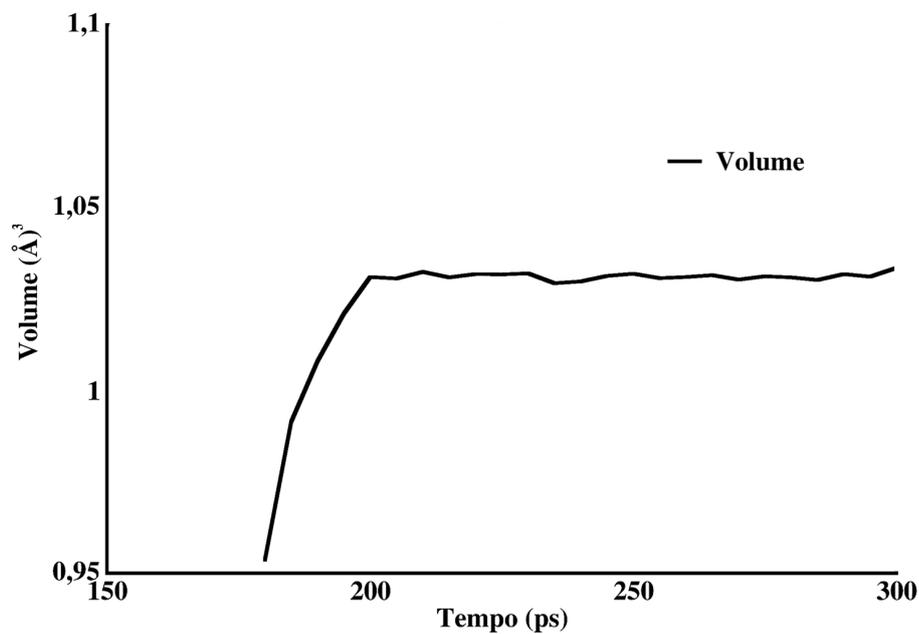
Figura 11 – Representação gráfica da variação da temperatura dos sistemas em função do tempo.



Fonte: o próprio autor.

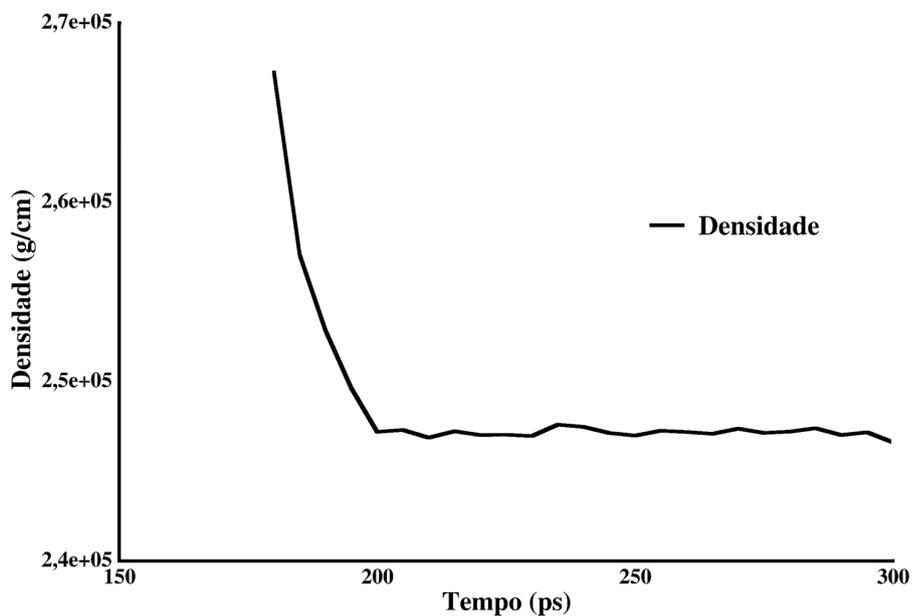
Como o número de moléculas permanece constante no sistema, os valores de densidade e volume (que são inversamente proporcionais) são os resultados esperados no equilíbrio, conforme Figuras 12 e 13. O volume do sistema sofre expansão até que tornar-se constante na etapa de termalização. Com o aumento do volume tem-se por consequência a diminuição da densidade.

Figura 12 – Representação gráfica da variação de volume em função do tempo.



Fonte: o próprio autor.

Figura 13 – Representação gráfica da variação de densidade em função do tempo.



Fonte: o próprio autor.

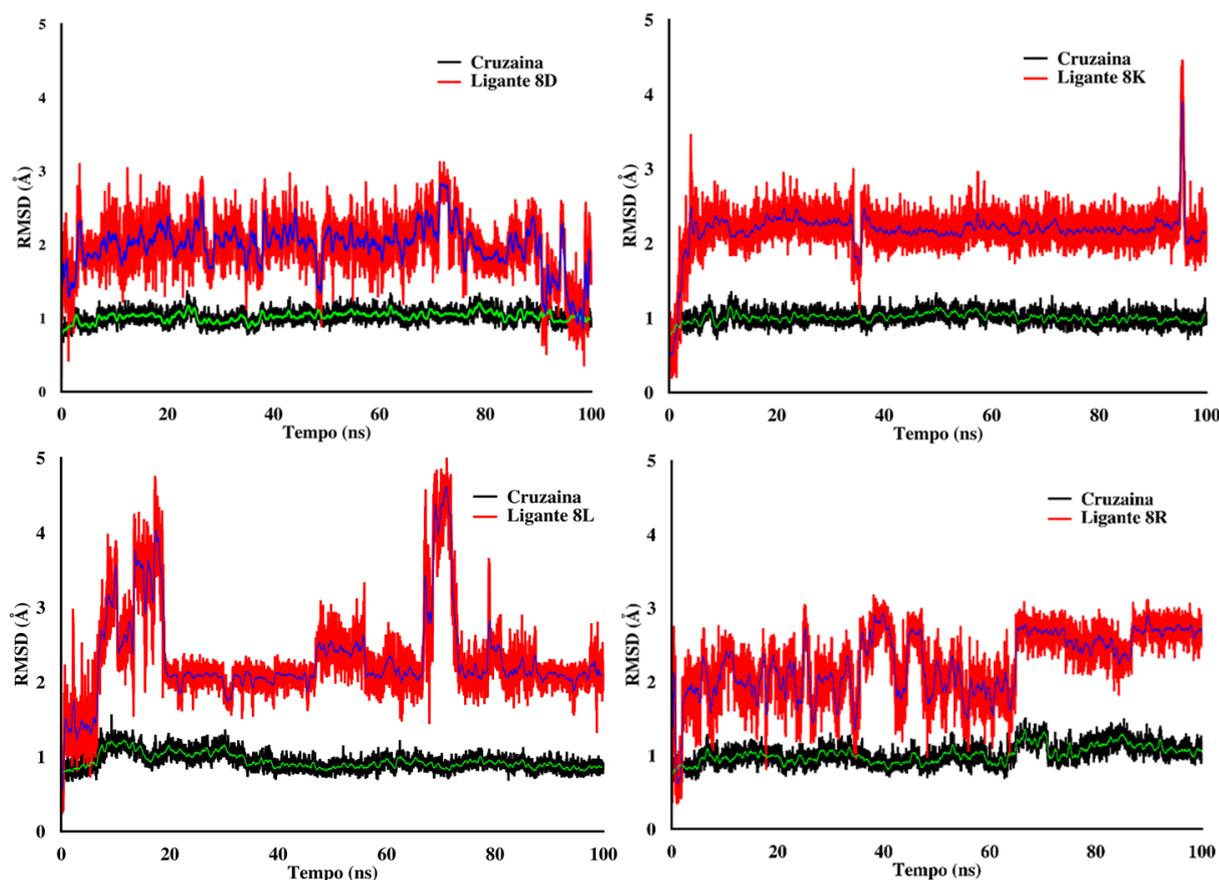
## 4.2 RMSD dos complexos

O RMSD é normalmente expresso em unidades de comprimento, tipicamente Å (ângstrom) e representa o desvio padrão entre a posição dos mesmos átomos ou moléculas nas duas conformações. Esta medida é muito usada também para avaliar se a estrutura

de uma macromolécula gerada por uma DM está em equilíbrio ou não, usando para esse efeito tipicamente a estrutura cristalográfica do sistema como referência.

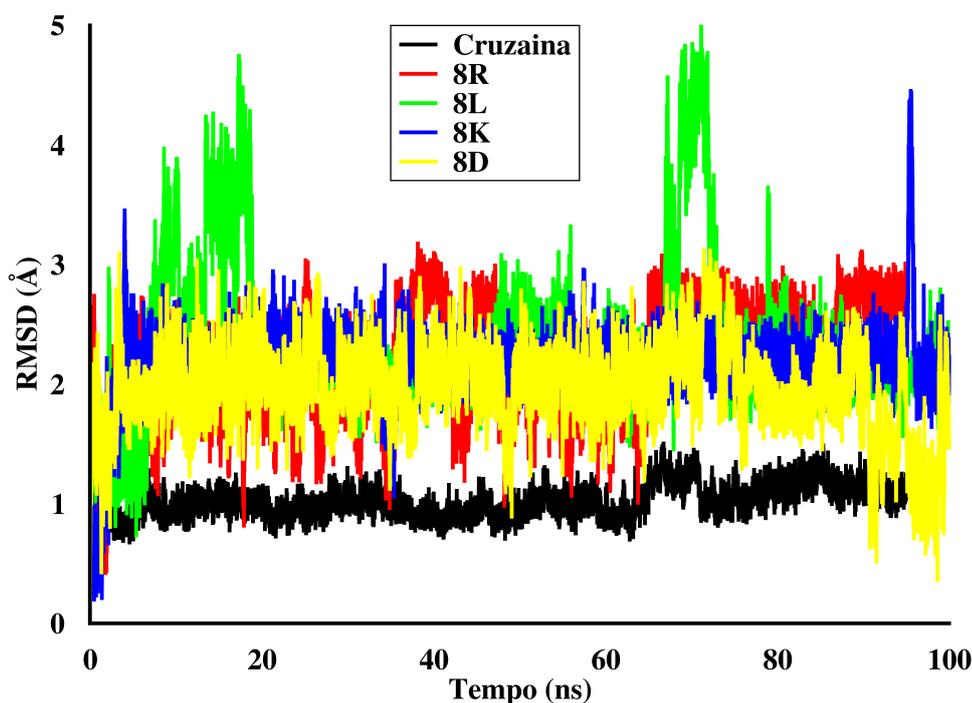
A fim de conferir as variações das posições dos ligantes complexados na cruzaina ao longo da dinâmica, em relação à sua posição inicial, foram plotados os gráficos RMSD dos quatro sistemas na Figura 14. Pode-se perceber que o valor médio do RMSD da cruzaina ficou abaixo de 1,5 Å em todas as simulações, indicando pouca movimentação da enzima durante as dinâmicas.

Figura 14 – RMSD da cruzaina complexada com os ligantes 8D, 8K, 8L E 8R



Fonte: o próprio autor.

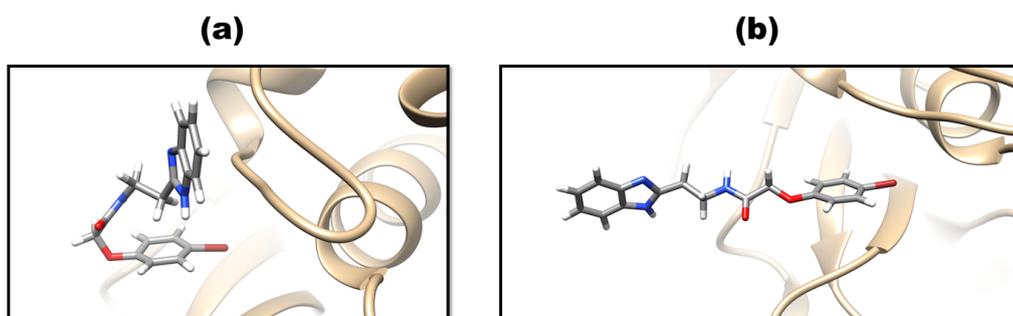
Figura 15 – RMSD em sobreposição dos ligantes 8D, 8K, 8L E 8R



Fonte: O próprio autor.

Assim como na enzima, foi analisado especificamente, a variação da posição dos ligantes ao longo das simulações em relação à sua posição inicial, e plotado em sobreposição o gráfico da Figura 15. Pode-se notar uma variação, em média, abaixo de 3 Å, o que indica que os inibidores pouco se movimentam, estando relativamente presos à proteína. É importante destacar que o ligante 8L se movimenta mais que os demais, chegando a picos de quase 5 Å, porém não prejudicando sua estabilidade, que foi em torno de 2 Å, nos tempos finais da simulação.

Figura 16 – Conformação do ligante 8L no RMSD mais estável (a) e no RMSD mais alto (b).



Fonte: O próprio autor.

Os maiores picos de RMSD do ligante 8L, foram entre 69 ns e 72 ns, cuja conformação está na Figura 16(b). Neste período, o ligante se encontra mais próximo (cerca de 4Å) dos

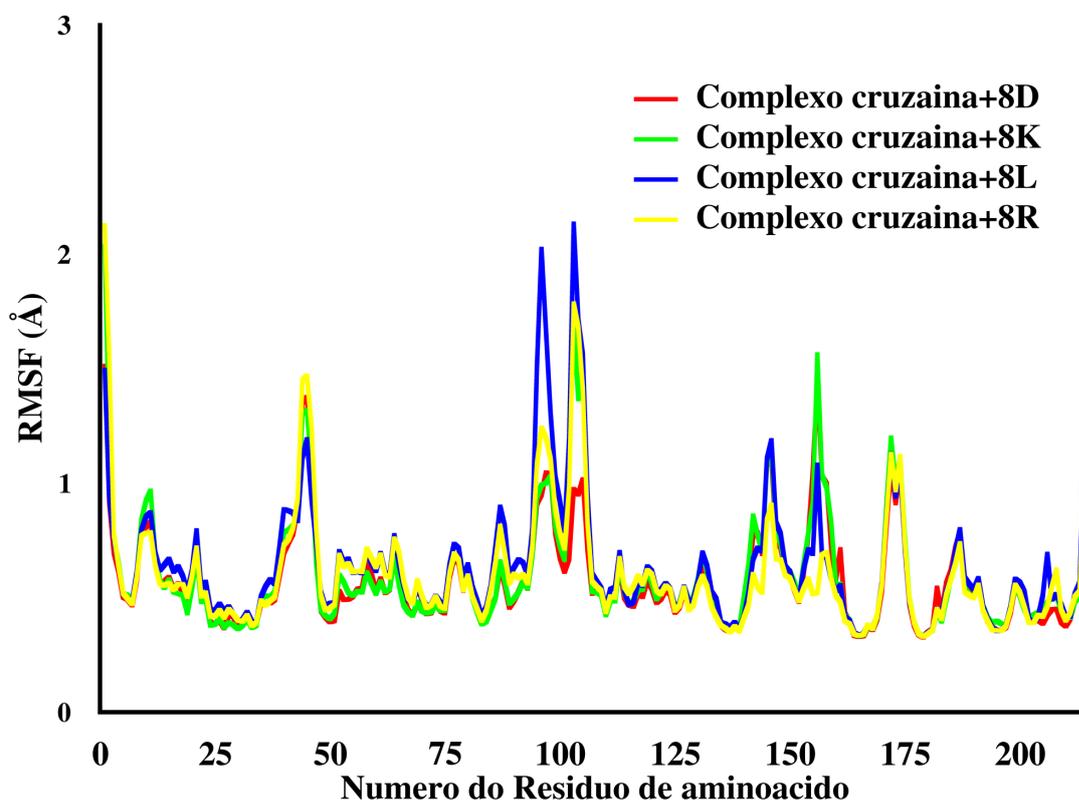
resíduos TRP26, LEU 67, MET68, LEU160 e ASP161. Após isso, ele retorna à sua posição mais estável, Figura 16(a), cujos resíduos mais próximos são: ASH60, SER61, SER64, GLY65, GLY66, LEU67, MET68, ALA138, LEU160 e GLU208.

### 4.3 RMSF dos complexos

Para avaliar a flexibilidade das diferentes zonas da cruzaina, determinou-se o quanto variou a posição de cada resíduo de aminoácido em torno de sua posição de equilíbrio ao longo de toda a simulação.

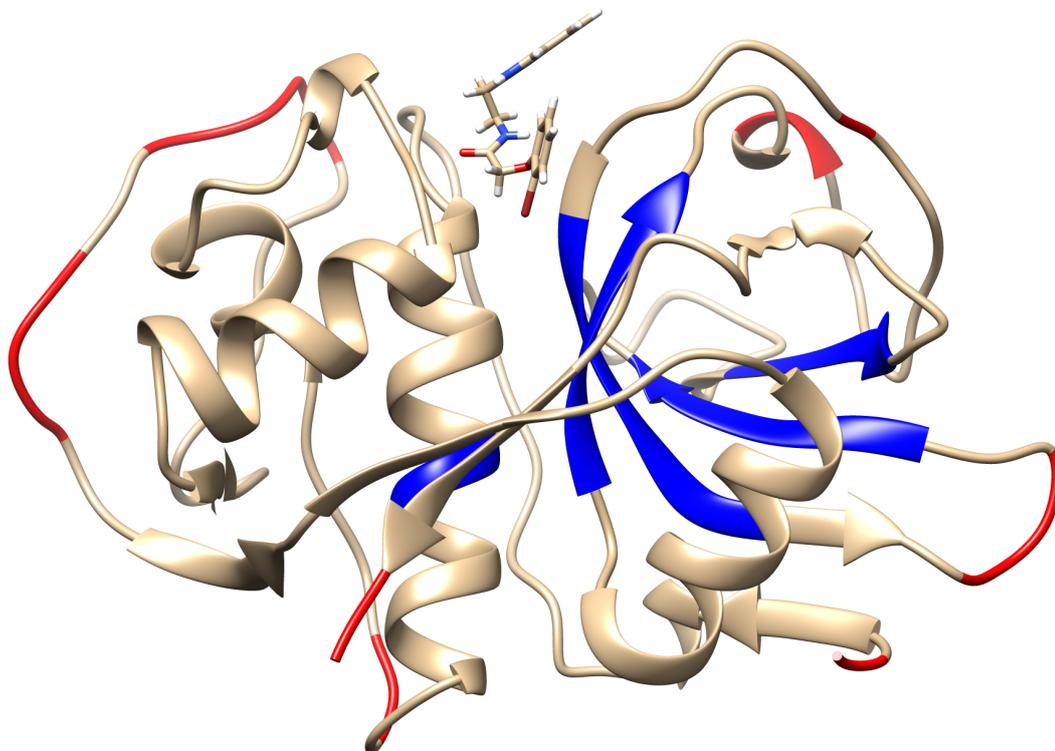
Verificou-se que as regiões mais flexíveis correspondem ao início e fim da enzima, resíduos 1 e 2, 214 e 215. Este comportamento é esperado, uma vez que as regiões N e C-terminais correspondem a terminações livres, e portanto, mais sujeitas a movimento de acordo com as interações do solvente. Soma-se a isto, algumas regiões de *loop* mais expostas ao solvente, experimentam maior flexibilidade, como os resíduos THR103 e GLY96 (na ordem de flutuação). Em contrapartida, regiões mais interiores, como as destacadas em azul (Figura 18), são mais rígidas.

Figura 17 – Representação gráfica dos valores de RMSF em função dos resíduos de aminoácidos.



Fonte: O próprio autor.

Figura 18 – Representação 3D da cruzaina. Em vermelho estão destacadas as regiões mais flexíveis da enzima; em azul, as mais rígidas.



Fonte: O próprio autor.

Tabela 1 – Valores de RMSF do sítio ativo.

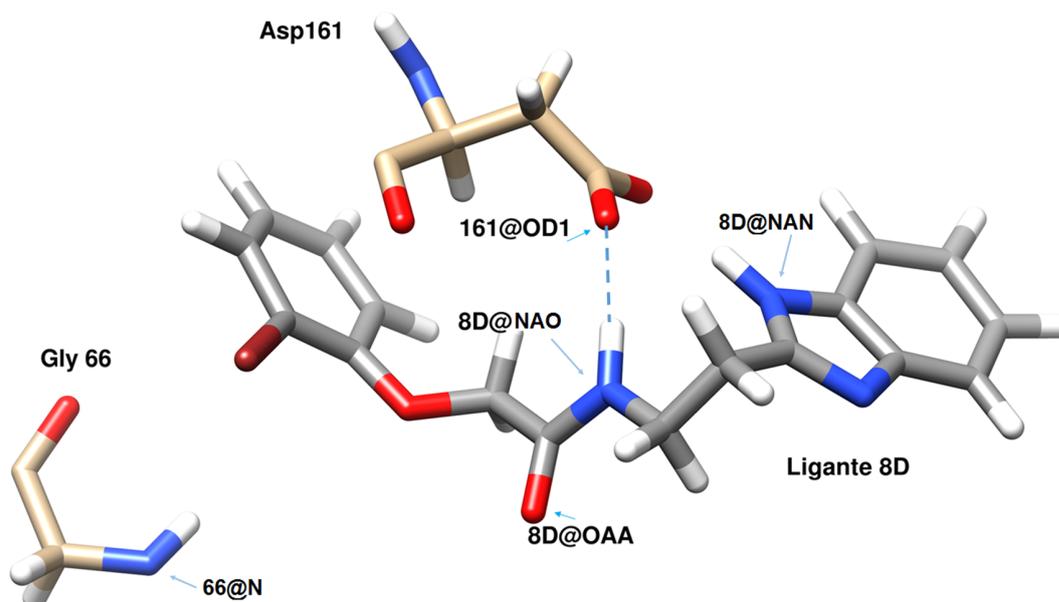
Resíduo	RMSF (Å)
GLY23	0.5007
CYS25	0.3846
TRP26	0.4015
SER61	0.5838
GLY65	0.6014
GLY66	0.4999
GLY133	0.4951
GLN159	0.7208
ASP161	0.7113
HIE162	0.4311

## 4.4 Análise das ligações de hidrogênios e distâncias

As ligações de hidrogênio foram calculadas tendo em vista todas as interações entre grupos -OH e -NH e os respectivos receptores N e O, limitando a distância de interação a 3,50 Å. Considerando todas as pontes de hidrogênio presentes em mais de 50% do tempo de simulação, foram analisadas as 5 mais relevantes em cada complexo.

Durante a simulação, entre o ligante 8D e a cruzaína estabeleceu-se com maior frequência a interação de ligação de hidrogênio entre o oxigênio denominado OD1 do resíduo ASP161 e a nitrila denominada NAO deste ligante, conforme Figura 19. Esta ligação permaneceu em 14,24 ns, tendo como distância média 2,82 Å - esta mesma distância na estrutura inicial do PDB foi de 2,12 Å, Tabela 2. Ainda na estrutura inicial do PDB, há ligação de hidrogênio entre a nitrila N da GLY66 com o oxigênio OAA do ligante 8D com distância de 2,37 Å. Entretanto, esta ligação permaneceu por somente 0,01 ns, com distância média de 2,98 Å. De acordo como a Figura 20, percebe-se que entre 25 e 50 ns houve maior aproximação do ligante com o resíduo ASP161.

Figura 19 – Interação de ligação hidrogênio entre o ligante 8D e o resíduo ASP161 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.

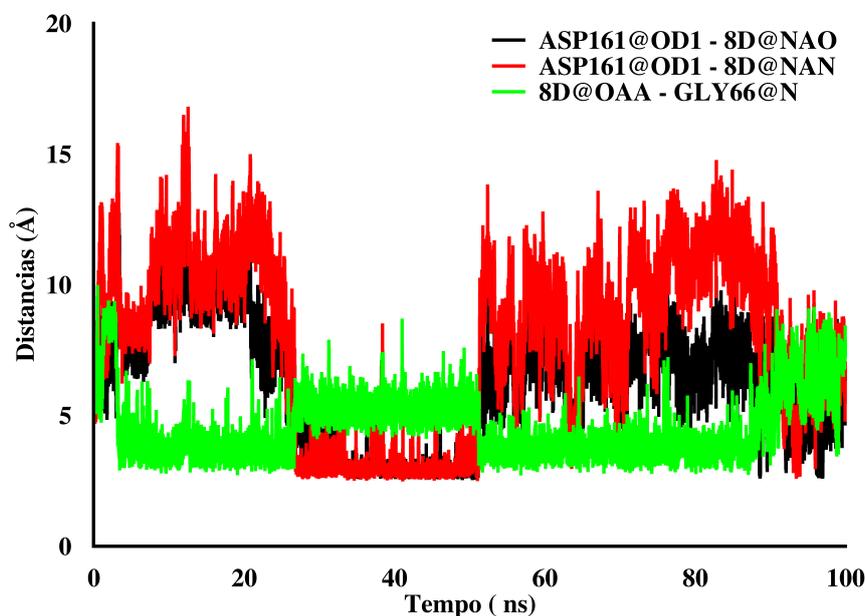


Fonte: O próprio autor.

Tabela 2 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8D.

Interação	Tempo (ns)	$d_{\text{média}}$ (Å)	$d_{\text{estrutura inicial PDB}}$ (Å)
ASP161@OD1 8D@NAO	14,24	2.8198	2.12
ASP161@OD1 8D@NAN	9,48	2.8243	-
8D@OAA GLY66@N	0,01	2.9844	2.37

Figura 20 – Gráfico das distâncias pelo tempo. Ligante 8D.



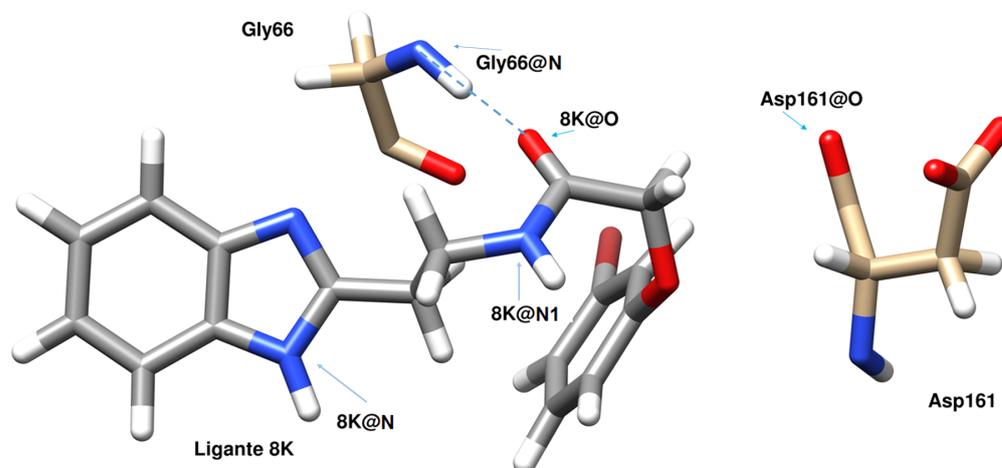
Fonte: O próprio autor.

A interação de ligação de hidrogênio entre o ligante 8K e a cruzaina que se estabeleceu com mais frequência foi entre o hidrogênio H do resíduo GLY66 com o oxigênio O do ligante, Figura 21. Esta ligação permaneceu por 12,88 ns, com distância média 2,89 Å, não tendo dados experimentais de distância para este ligante, Tabela 3. Com menos frequência do que a ligação 8K@O - GLY66@N, está a ligação entre o oxigênio O do ASP161 e a nitrila N1 do ligante 8K, com distância média de 2,86 Å. A Figura 22, que contém o gráfico das distâncias das principais interações pelo tempo, só ratifica a maior permanência da ligação entre a GLY66 e o ligante.

Tabela 3 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8K.

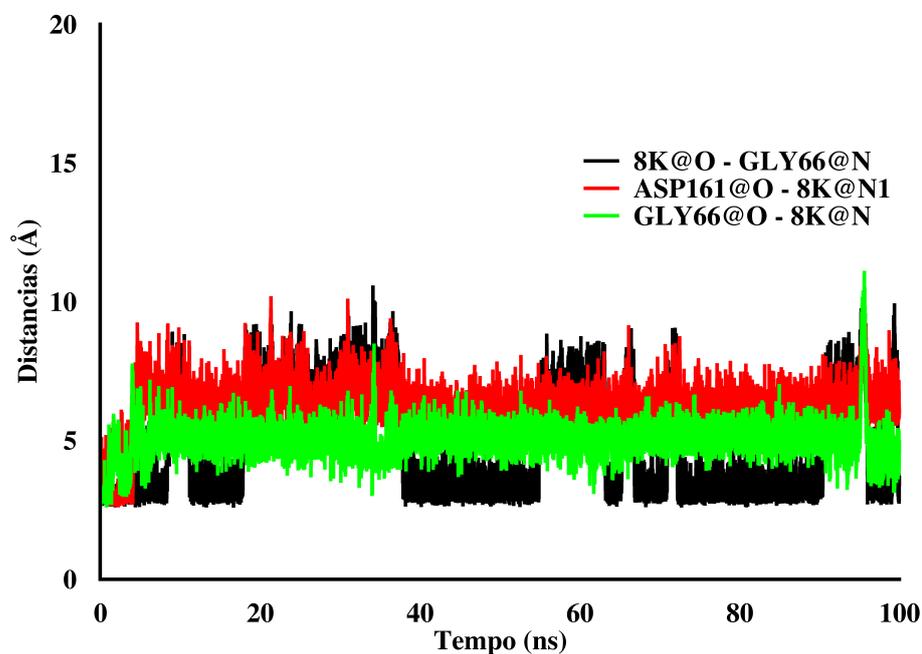
Interação	Tempo (ns)	$d_{\text{média}}$ (Å)
8K@O GLY66@N	12,88	2.8893
ASP161@O 8K@N1	0,89	2.8631
GLY66@O 8K@N	0,71	2.8709

Figura 21 – Interação de ligação hidrogênio entre o ligante 8K e o resíduo GLY66 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.



Fonte: O próprio autor.

Figura 22 – Gráfico das distâncias pelo tempo. Ligante 8K.

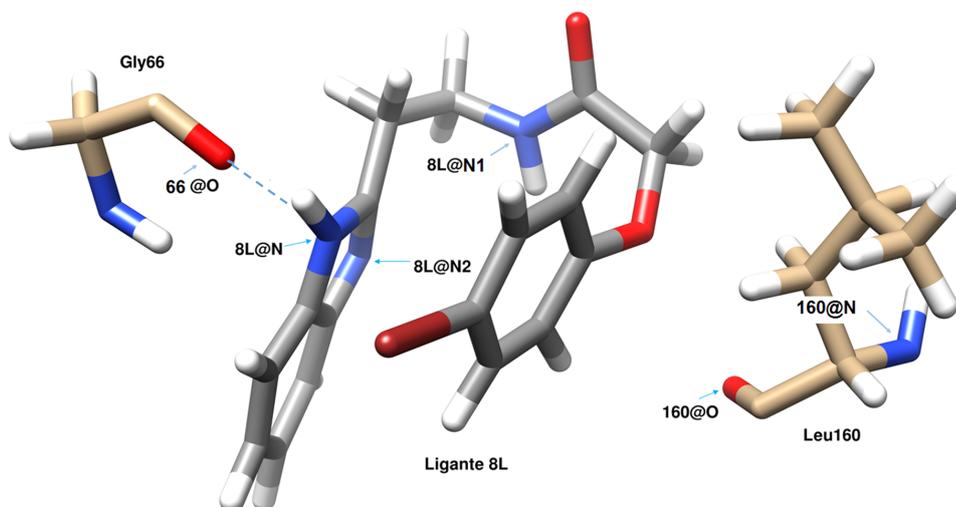


Fonte: O próprio autor.

Na simulação da cruzaina complexada com o ligante 8L, a maior frequência de interação de ligação de hidrogênio foi entre o oxigênio O do resíduo GLY66 com a nitrila N deste ligante, Figura 23. Esta ligação se deu em 19,09 ns, com distância média de 2,8 Å, não tendo dados experimentais para distâncias, Tabela 4.

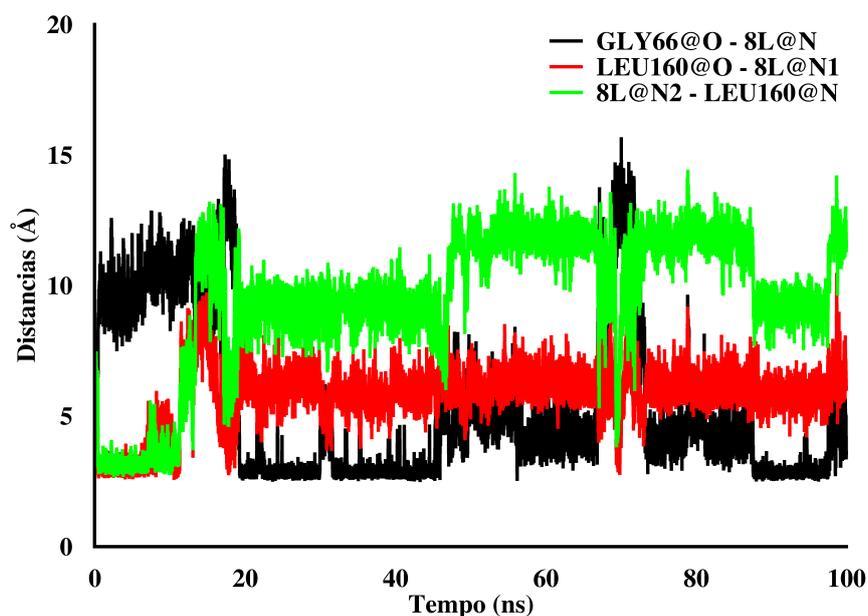
Observa-se na Figura 24, as distâncias de ligação de hidrogênio pelo tempo, onde nos momentos iniciais de simulação há uma aproximação da ligação LEU160@O - 8L@N1 havendo, posteriormente, o afastamento definitivo até o final. A principal ligação, GLY66@O - 8L@N, se estabelece somente a partir de 20 ns.

Figura 23 – Interação de ligação hidrogênio entre o ligante 8L e o resíduo GLY66 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.



Fonte: O próprio autor.

Figura 24 – Gráfico das distâncias pelo tempo. Ligante 8L.



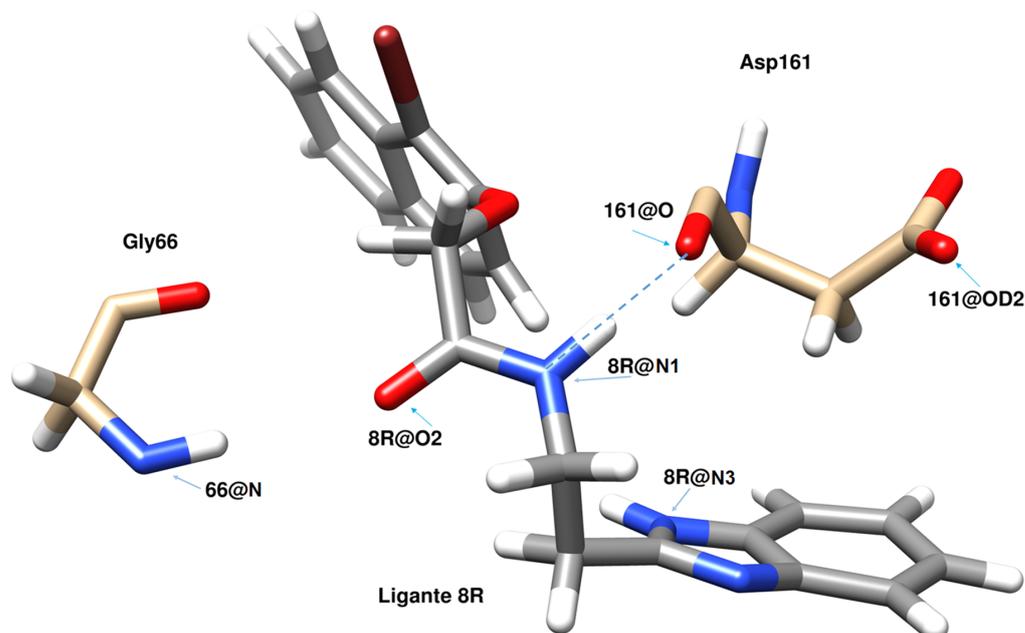
Fonte: O próprio autor.

Tabela 4 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8L.

Interação	Tempo (ns)	$d_{\text{média}}$ (Å)
GLY66@O 8L@N	19,09	2.7999
LEU160@O 8L@N1	4,74	2.8769
8L@N2 LEU160@N	2,07	2.9276

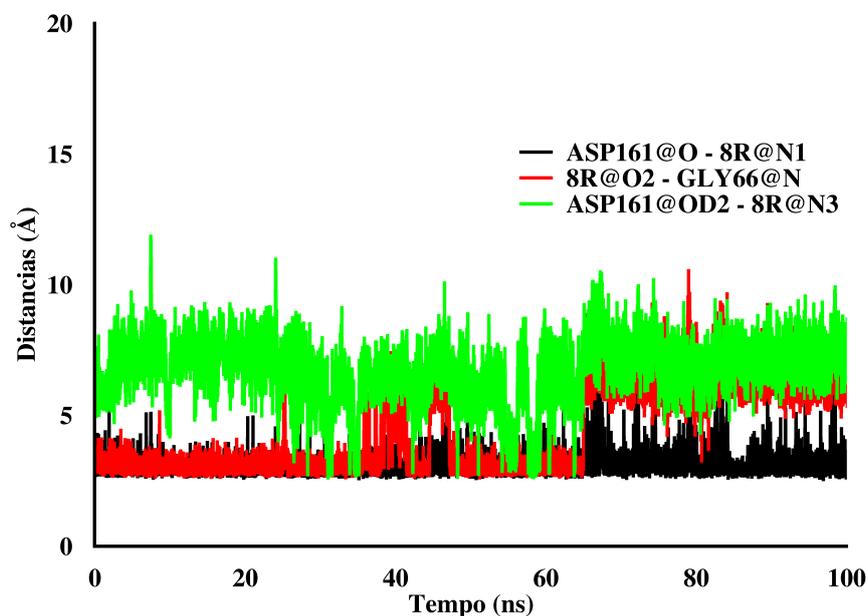
Sendo o ligante 8R mais potente e menos citotóxico da série de análogos (FERREIRA et al., 2014), em complexo com a cruzaina, estabelece por mais vezes a ligação de hidrogênio entre o oxigênio O do ASP161 e a nitrila N1 deste ligante, Figura 25. Esta ligação permanece por 50,79 ns, o que representa, aproximadamente, 3 vezes mais que os ligantes 8D, 8K e 8L, mantendo uma distância média de 2,86 Å, Tabela 5. Apesar de menos frequente que a ligação acima citada e não menos importante, está a ligação 8R@O2 - GLY66N, à qual permanece por 27,22 ns (cerca de 2 vezes mais que os ligantes anteriores), com distância média de 2,88 Å. Entretanto, esta ligação permanece até cerca de 65 ns, se afastando definitivamente até o final da simulação, Figura 26.

Figura 25 – Interação de ligação hidrogênio entre o ligante 8R e o resíduo ASP161 (linha tracejada azul) que mais se estabeleceu durante toda a simulação.



Fonte: O próprio autor.

Figura 26 – Gráfico das distâncias pelo tempo. Ligante 8R.



Fonte: O próprio autor.

Tabela 5 – Distâncias médias e tempo de duração das principais ligações de hidrogênio com o ligante 8R.

Interação	Tempo (ns)	$d_{\text{média}}$ (Å)
ASP161@O 8R@N1	50,79	2.8644
8R@O2 GLY66@N	27,22	2.8818
ASP161@OD2 8R@N3	1,46	2.8420

Pelas análises das interações de ligações de hidrogênio, é possível notar que os ligantes 8D e 8R, têm preferência em se ligar com o resíduo ASP161, enquanto que os ligantes 8K e 8L, com a GLY66. A ligação que permanece disparadamente com mais frequência é a ligação ASP161@O - 8R@N1 (Figura 25), do ligante 8R.

## 4.5 Energia livre de ligação e energia de interação solvatada (SIE)

Foram analisadas as energias livre de ligação dos ligantes 8D, 8K, 8L e 8R e analisadas as contribuições energéticas por meio de decomposição de resíduos dos ligantes mais relevantes, ou seja, 8D (cristalográfico) e 8R (apontado no artigo (FERREIRA et al., 2014) como o mais potente e menos citotóxico).

Por ser mais eficiente e menos dispendioso em termos de custo computacional, foi utilizado o método MMGBSA para obter valores em termos de contribuição energética de afinidades de ligação proteína-ligante. Pela Tabela 6, temos o valor mais negativo,  $\Delta G_{\text{total}} = -31,04$ , para o complexo cruzaína-8R, o que sugere que este complexo tenha mais afinidade de ligação. Quando confrontado com o experimental, vemos que a tendência é que este mesmo ligante tenha mais afinidade com a enzima do que o 8D.

Tabela 6 – Energia livre de ligação via MMGBSA dos complexos.

Complexo	$\Delta G_{\text{total}}$	$\Delta G_{\text{experimental}}$
Cruzaína-8D	-10.6680	-8.317
Cruzaína-8K	-25.3546	-
Cruzaína-8L	-22.5016	-
Cruzaína-8R	-31.0380	-9.666

Também foi calculada a energia livre de ligação utilizando o método de energia de interação solvatada (SIE) (NAÏM et al., 2007). Este método trata o sistema com efeitos de solvatação implicitamente, comum às avaliações proteína-ligante. Por meio deste, os resultados apresentaram uma maior precisão relacionada aos dados experimentais, seguindo a mesma tendência dos valores de MMGBSA apontando para o ligante 8R como o que tem mais afinidade, como mostrado na Figura 7.

Tabela 7 – Comparação dos valores de MMGBSA, SIE e experimental.

Ligante	$\Delta G$ (kcal/mol)		
	MMGBSA	SIE	Experimental
8D	-10.6680	-5.6900	-8.3170
8K	-25.3546	-6.4400	-
8L	-22.5016	-6.2100	-
8R	-31.038	-7.5400	-9.6660

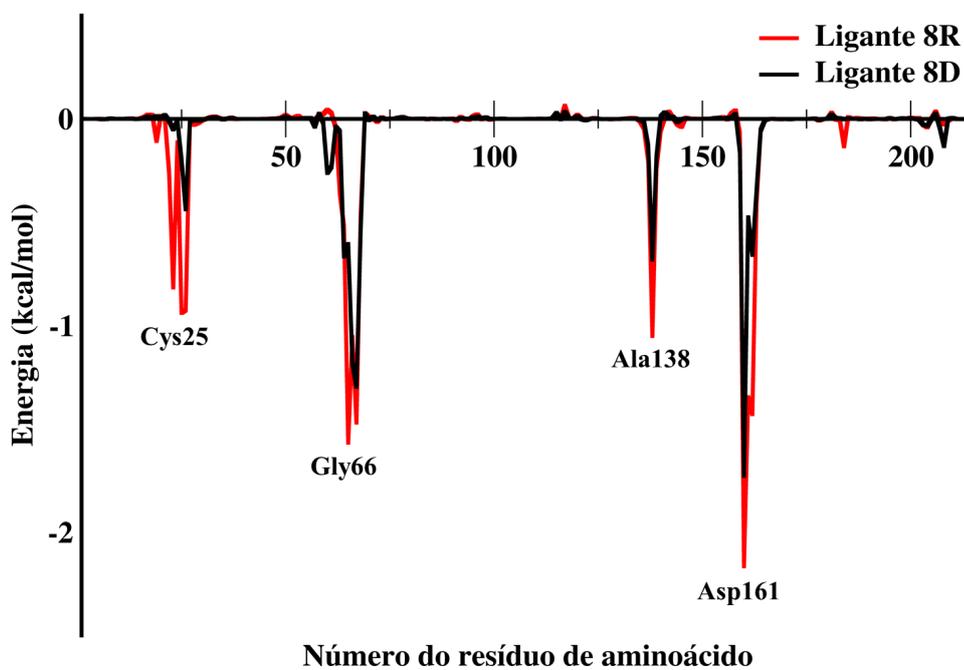
## 4.6 Decomposição da energia por resíduos

Tendo como destaque as importantes interações dos ligantes 8D e 8R com os resíduos de aminoácidos GLY66 e ASP161 pertencentes ao sítio catalítico ao longo das simulações, o resultado da decomposição por resíduo aponta para estes mesmos aminoácidos como os que mais contribuem energeticamente às interações do inibidor com a cruzaína, conforme Tabela 8 e Figura 27.

Tabela 8 – Decomposição da energia por resíduo, em kcal/mol

Sistema	Resíduo	$E_{VdW}$	$E_{Elet}$	$E_{Total}$
Cruzaína-8D	Cys25	-0.013	-0.138	-0.008
	Gly66	-1.162	-0.154	-1.297
	Ala138	-0.758	-0.328	-1.051
	Asp161	-1.540	-0.873	-1.728
Cruzaína-8R	Cys25	-0.636	-0.187	-0.939
	Gly66	-0.654	-2.034	-1.567
	Ala138	-0.654	-2.034	-1.567
	Asp161	-2.183	-1.134	-2.165

Figura 27 – Gráfico da decomposição das energias por resíduo dos ligantes 8D e 8R.



Fonte: O próprio autor.



## 5 Conclusão

Dentro do estudo computacional, a fim de investigar as potenciais interações da cruzaina com alguns inibidores para a proposição de novas moléculas ativas no combate à doença de Chagas, foi realizado DM desta enzima com os ligantes 8D (composto líder), 8K, 8L e com 8R. Seguindo um protocolo rígido para os 4 sistemas, foram realizadas a escolha para obtenção da estrutura inicial e preparação do arquivo PDB, otimização geométrica e cargas RESP dos ligantes, preparação de arquivos de parâmetros, topologia e coordenadas, minimização, termalização e produção de DM.

Foram analisadas as propriedades dos sistemas, as quais são comuns aos quatro: energia cinética, energia potencial, energia total, temperatura, volume e densidade.

Ao analisar o RMSD, vimos que a cruzaina se comporta com pouca movimentação ao ser complexada com os quatro ligantes. Ao serem analisados separadamente, nota-se também pouca movimentação dos ligantes 8D, 8K, 8R (cerca de 2 Å). Entretanto, o 8L apresentou um valor de RMSD altos em torno de 18 ns e em torno de 70 ns, tendo mudança acentuada na conformação.

O RMSF aponta os resíduos THR103, GLY96 (nesta ordem) e as regiões N e C-terminais (terminações livres) como os que mais flutuam ao longo das simulações.

Na investigação das interações de ligação de hidrogênio e suas distâncias, foi possível inferir que os ligantes 8D e 8R têm preferência em ligar-se com o resíduo ASP161, tendo como destaque o 8R, no qual aparece com mais frequência a ligação com este resíduo. Os ligantes 8K e 8L têm preferência em se ligar com o resíduo GLY66.

Em termos de energia livre de ligação, o  $\Delta G_{\text{total}}$  - com o ligante 8R mais negativo, portanto mais favorável à reação - segue a tendência do  $\Delta G_{\text{experimental}}$ , o qual aponta para este mesmo ligante como sendo o mais favorável. O método SIE também aponta o ligante 8R como o mais favorável, cujos valores estão bem próximos do experimental. Feita a decomposição da energia por resíduo, o resultado mostra a melhor contribuição do resíduo ASP161.

As análises pormenorizadas dos resultados das simulações de dinâmica molecular, das energias livres de ligação, das energias de interação solvatada e da decomposição por resíduo, ratificam os resultados experimentais, os quais apontam o composto 8R como o mais ativo e menos tóxico dentre a série de derivados desta classe. Esses resultados devem favorecer a otimização desse composto ou desenvolvimento de uma série de análogos a fim de se utilizar como possível fármaco para o tratamento da doença de Chagas.



## Referências

- ATKINS, P. W.; FRIEDMAN, R. S. *Molecular quantum mechanics*. [S.l.]: Oxford university press, 2011. Citado na página 40.
- ATWOOD, J. et al. The trypanosoma cruzi proteome. *Science*, American Association for the Advancement of Science, v. 309, n. 5733, p. 473–476, 2005. Citado na página 24.
- BARREIRO, E. J.; FRAGA, C. A. *Química medicinal: as bases moleculares da ação dos fármacos*. [S.l.]: Artmed, 2008. Citado na página 20.
- BAYLY, C. I. et al. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *The Journal of Physical Chemistry*, ACS Publications, v. 97, n. 40, p. 10269–10280, 1993. Citado na página 45.
- BECKE, A. Hartree–fock exchange energy of an inhomogeneous electron gas. *International journal of quantum chemistry*, Wiley Online Library, v. 23, n. 6, p. 1915–1922, 1983. Citado na página 43.
- BERMAN, H. M. et al. The protein data bank, 1999–. In: *International Tables for Crystallography Volume F: Crystallography of biological macromolecules*. [S.l.]: Springer, 2006. p. 675–684. Citado na página 50.
- BORCHHARDT, D. M. et al. Biochemical evaluation of a series of synthetic chalcone and hydrazide derivatives as novel inhibitors of cruzain from trypanosoma cruzi. *Journal of the Brazilian Chemical Society*, SciELO Brasil, v. 21, n. 1, p. 142–150, 2010. Citado na página 25.
- BRAK, K. et al. Identification of a new class of nonpeptidic inhibitors of cruzain. *Journal of the American Chemical Society*, ACS Publications, v. 130, n. 20, p. 6404–6410, 2008. Citado 2 vezes nas páginas 25 e 28.
- BRENEMAN, C. M.; WIBERG, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry*, Wiley Online Library, v. 11, n. 3, p. 361–373, 1990. Citado na página 45.
- BRINEN, L. S. et al. A target within the target: probing cruzains p1 site to define structural determinants for the chagas disease protease. *Structure*, Elsevier, v. 8, n. 8, p. 831–840, 2000. Citado na página 25.
- BRYANT, C. et al. Novel non-peptidic vinylsulfones targeting the s2 and s3 subsites of parasite cysteine proteases. *Bioorganic & medicinal chemistry letters*, Elsevier, v. 19, n. 21, p. 6218–6221, 2009. Citado na página 25.
- CARDOSO, F. J. B. et al. A study on antimalarial artemisinin derivatives using mep maps and multivariate qsar. *Journal of molecular modeling*, Springer, v. 14, n. 1, p. 39–48, 2008. Citado na página 41.

- CARNEIRO, A. d. S. *Estudo teórico do mecanismo catalítico e de inibição das enzimas cisteína protease e GAPDH de Trypanosoma cruzi*. Tese (Doutorado) — Universidade Federal do Pará, 2012. Citado na página 28.
- CASE, D. et al. Amber 12. p. 1–826, 2012. Citado na página 49.
- CASTRO, J. A.; DEMECCA, M. M.; BARTEL, L. C. Toxic side effects of drugs used to treat chagas' disease (american trypanosomiasis). *Human & experimental toxicology*, Sage Publications Sage CA: Thousand Oaks, CA, v. 25, n. 8, p. 471–479, 2006. Citado na página 24.
- CAZZULO, J.; STOKA, V.; TURK, V. The major cysteine proteinase of trypanosoma cruzi: a valid target for chemotherapy of chagas disease. *Current pharmaceutical design*, Bentham Science Publishers, v. 7, n. 12, p. 1143–1156, 2001. Citado na página 25.
- CDC. *Parasites - American Trypanosomiasis (also known as Chagas Disease)*. Centers for Disease Control and Prevention, 2015. Disponível em: <<https://www.cdc.gov/parasites/chagas/biology.html>>. Acesso em: 27 fev 2017. Citado na página 23.
- CELE, F. N.; RAMESH, M.; SOLIMAN, M. E. Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: A study for identification of reverse transcriptase inhibitors as potential anti-hiv agents. *Drug design, development and therapy*, Dove Press, v. 10, p. 1365, 2016. Citado na página 21.
- CHIQUITO, A. J.; ALMEIDA, N. G. de. O potencial de lennard-jones: Aplica ção a moléculas diatômicas. *Revista Brasileira de Ensino de Física*, v. 21, n. 2, 1999. Citado na página 35.
- CLAUSIUS, R. Über verschiedene für die anwendung bequeme formen der hauptgleichungen der mechanischen wärmetheorie. *Annalen der Physik*, Wiley Online Library, v. 201, n. 7, p. 353–400, 1865. Citado na página 47.
- CLAYTON, J. Chagas disease 101. *Nature*, Nature Publishing Group, p. S4, 2010. Citado na página 19.
- COURANT, R. Variational methods for the solution of problems of equilibrium and vibrations. *Bulletin of the American mathematical Society*, v. 49, n. 1, p. 1–23, 1943. Citado na página 37.
- DELANO, W. L. The pymol molecular graphics system. <http://pymol.org>, DeLano Scientific, 2002. Citado na página 50.
- DIAS, L. C. et al. Quimioterapia da doença de chagas: estado da arte e perspectivas no desenvolvimento de novos fármacos. *Química Nova*, Sociedade Brasileira de Química, 2009. Citado na página 24.
- DOLINSKY, T. J. et al. Pdb2pqr: an automated pipeline for the setup of poisson–boltzmann electrostatics calculations. *Nucleic acids research*, Oxford University Press, v. 32, n. suppl\_2, p. W665–W667, 2004. Citado na página 51.
- DOYLE, P. S. et al. The trypanosoma cruzi protease cruzain mediates immune evasion. *PLoS pathogens*, Public Library of Science, v. 7, n. 9, p. e1002139, 2011. Citado na página 25.

- DU, X. et al. Synthesis and structure- activity relationship study of potent trypanocidal thio semicarbazone inhibitors of the trypanosomal cysteine protease cruzain. *Journal of medicinal chemistry*, ACS Publications, v. 45, n. 13, p. 2695–2707, 2002. Citado na página 28.
- EL-SAYED, N. M. et al. The genome sequence of trypanosoma cruzi, etiologic agent of chagas disease. *Science*, American Association for the Advancement of Science, v. 309, n. 5733, p. 409–415, 2005. Citado na página 24.
- ENGEL, J. C. et al. Cysteine protease inhibitors cure an experimental trypanosoma cruzi infection. *Journal of Experimental Medicine*, Rockefeller University Press, v. 188, n. 4, p. 725–734, 1998. Citado na página 25.
- ENGEL, J. C. et al. Upregulation of the secretory pathway in cysteine protease inhibitor-resistant trypanosoma cruzi. *J Cell Sci*, The Company of Biologists Ltd, v. 113, n. 8, p. 1345–1354, 2000. Citado na página 25.
- ESPERANDIM, V. R. et al. In vivo infection by trypanosoma cruzi: a morphometric study of tissue changes in mice. *Parasitology research*, Springer, v. 112, n. 1, p. 431–436, 2013. Citado na página 23.
- FERREIRA, R. S. et al. Synthesis, biological evaluation, and structure–activity relationships of potent noncovalent and nonpeptidic cruzain inhibitors as anti-trypanosoma cruzi agents. *Journal of medicinal chemistry*, ACS Publications, v. 57, n. 6, p. 2380–2392, 2014. Citado 4 vezes nas páginas 28, 29, 66 e 67.
- FERREIRA, R. S. et al. Integração das técnicas de triagem virtual e triagem biológica automatizada em alta escala: oportunidades e desafios em p&d de fármacos. *Química Nova*, Sociedade Brasileira de Química-SBQ, v. 34, n. 10, p. 1770–1778, 2011. Citado na página 51.
- FERREIRA, R. S. et al. Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. *Journal of medicinal chemistry*, ACS Publications, v. 53, n. 13, p. 4891–4905, 2010. Citado 4 vezes nas páginas 25, 28, 50 e 51.
- FLETCHER, R.; REEVES, C. M. Function minimization by conjugate gradients. *The computer journal*, Oxford University Press, v. 7, n. 2, p. 149–154, 1964. Citado na página 37.
- FREITAS, A. C. d. Aplicação de métodos de mecânica quântica no estudo do termocromismo de alcóxidos de vanádio (iv) em solução. 2010. Citado na página 40.
- FREY, P. A.; HEGEMAN, A. D. *Enzymatic reaction mechanisms*. [S.l.]: Oxford University Press, 2007. Citado na página 24.
- FRISCH, M. et al. *Gaussian 09, revision D. 01*. [S.l.]: Gaussian, Inc., Wallingford CT, 2009. Citado na página 52.
- FUJII, N. et al. Discovery of potent thiosemicarbazone inhibitors of rhodesain and cruzain. *Bioorganic & medicinal chemistry letters*, Elsevier, v. 15, n. 1, p. 121–123, 2005. Citado na página 25.

- GENHEDEN, S. et al. Nonpolar solvation free energies of protein- ligand complexes. *Journal of chemical theory and computation*, ACS Publications, v. 6, n. 11, p. 3558–3568, 2010. Citado na página 47.
- GENHEDEN, S.; RYDE, U. The mm/pbsa and mm/gbsa methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*, Taylor & Francis, v. 10, n. 5, p. 449–461, 2015. Citado na página 48.
- GIBBS, J. W. Art. lii.–on the equilibrium of heterogeneous substances. *American Journal of Science and Arts (1820-1879)*, American Periodicals Series II, v. 16, n. 96, p. 441, 1878. Citado na página 47.
- GILLMOR, S. A.; CRAIK, C. S.; FLETTERICK, R. J. Structural determinants of specificity in the cysteine protease cruzain. *Protein Science*, Wiley Online Library, v. 6, n. 8, p. 1603–1611, 1997. Citado na página 25.
- GODBOUT, N. et al. Optimization of gaussian-type basis sets for local spin density functional calculations. part i. boron through neon, optimization technique and validation. *Canadian Journal of Chemistry*, NRC Research Press, v. 70, n. 2, p. 560–571, 1992. Citado na página 43.
- GOHLKE, H.; KIEL, C.; CASE, D. A. Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the ras–raf and ras–ralgds complexes. *Journal of molecular biology*, Elsevier, v. 330, n. 4, p. 891–913, 2003. Citado 2 vezes nas páginas 49 e 50.
- GOSS, D. J.; PETRUCCI, R. H. *General Chemistry Principles & Modern Applications*, Petrucci, Harwood, Herring, Madura: Study Guide. [S.l.]: Pearson/Prentice Hall, 2007. Citado na página 37.
- GUIDO, R. V.; ANDRICOPULO, A. D.; OLIVA, G. Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas. *estudos avançados*, SciELO Brasil, v. 24, n. 70, p. 81–98, 2010. Citado 2 vezes nas páginas 20 e 21.
- HANWELL, M. D. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, Springer, v. 4, n. 1, p. 17, 2012. Citado na página 51.
- JORGENSEN, W. L.; JENSON, C. Temperature dependence of tip3p, spc, and tip4p water from npt monte carlo simulations: Seeking temperatures of maximum density. *Journal of Computational Chemistry*, Wiley Online Library, v. 19, n. 10, p. 1179–1186, 1998. Citado na página 52.
- JR, A. R.; RASSI, A.; MARIN-NETO, J. A. Chagas disease. *The Lancet*, Elsevier, v. 375, n. 9723, p. 1388–1402, 2010. Citado na página 19.
- KERR, I. D. et al. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *Journal of Biological Chemistry*, ASBMB, v. 284, n. 38, p. 25697–25703, 2009. Citado na página 25.
- KOHN, W.; SHAM, L. J. Self-consistent equations including exchange and correlation effects. *Physical review*, APS, v. 140, n. 4A, p. A1133, 1965. Citado 2 vezes nas páginas 41 e 43.

- KOLLMAN, P. A. et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research*, ACS Publications, v. 33, n. 12, p. 889–897, 2000. Citado 2 vezes nas páginas 48 e 49.
- LASKOWSKI, R. A. Pdbsum: summaries and analyses of pdb structures. *Nucleic acids research*, Oxford University Press, v. 29, n. 1, p. 221–222, 2001. Citado na página 27.
- LEACH, A. R. *Molecular modelling: principles and applications*. [S.l.]: Pearson education, 2001. Citado na página 47.
- LEE, C.; YANG, W.; PARR, R. G. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, APS, v. 37, n. 2, p. 785, 1988. Citado na página 43.
- LEITE, R. P. et al. The uhlenbeck-ford model: Exact virial coefficients and application as a reference system in fluid-phase free-energy calculations. *The Journal of chemical physics*, AIP Publishing, v. 145, n. 19, p. 194101, 2016. Citado na página 39.
- LEVINE, I. N.; BUSCH, D. H.; SHULL, H. *Quantum chemistry*. [S.l.]: Pearson Prentice Hall Upper Saddle River, NJ, 2009. v. 6. Citado na página 40.
- LEWARS, E. G. *Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics*. [S.l.]: Springer, 2016. Citado na página 33.
- MAIER, J. A. et al. ff14sb: improving the accuracy of protein side chain and backbone parameters from ff99sb. *Journal of chemical theory and computation*, ACS Publications, v. 11, n. 8, p. 3696–3713, 2015. Citado na página 52.
- MALVEZZI, A. *Modelos de virtual screening de inibidores da cruzaina: desenvolvimento e validação experimental*. Tese (Doutorado) — Universidade de São Paulo, 2008. Citado na página 28.
- MARQUES, M. A.; BOTTI, S. O que é e para que serve a teoria dos funcionais da densidade. *Gazeta de física*, v. 29, n. 4, p. 10–15, 2006. Citado na página 41.
- MCGRATH, M. E. et al. The crystal structure of cruzain: a therapeutic target for chagas' disease. *Journal of molecular biology*, Elsevier, v. 247, n. 2, p. 251–259, 1995. Citado 2 vezes nas páginas 25 e 26.
- METZ, A. et al. Hot spots and transient pockets: predicting the determinants of small-molecule binding to a protein–protein interface. *Journal of chemical information and modeling*, ACS Publications, v. 52, n. 1, p. 120–133, 2011. Citado na página 49.
- MLADENOVIC, M. et al. On the origin of the stabilization of the zwitterionic resting state of cysteine proteases: a theoretical study. *Journal of the American Chemical Society*, ACS Publications, v. 130, n. 27, p. 8696–8705, 2008. Citado na página 51.
- MORRIS, G. M. et al. Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*, v. 19, n. 14, p. 1639–1662, 1998. Citado na página 48.
- MOTT, B. T. et al. Identification and optimization of inhibitors of trypanosomal cysteine proteases: cruzain, rhodesain, and tbcab. *Journal of medicinal chemistry*, ACS Publications, v. 53, n. 1, p. 52–60, 2009. Citado na página 25.

- MUNOZ-SARAVIA, S. G. et al. Chronic chagas' heart disease: a disease on its way to becoming a worldwide health problem: epidemiology, etiopathology, treatment, pathogenesis and laboratory medicine. *Heart failure reviews*, Springer, v. 17, n. 1, p. 45–64, 2012. Citado 2 vezes nas páginas 19 e 20.
- NAÏM, M. et al. Solvated interaction energy (sie) for scoring protein- ligand binding affinities. 1. exploring the parameter space. *Journal of chemical information and modeling*, ACS Publications, v. 47, n. 1, p. 122–133, 2007. Citado na página 68.
- NAMBA, A.; SILVA, V. D.; SILVA, C. D. Dinâmica molecular: teoria e aplicações em planejamento de fármacos. *Eclética Química*, Universidade Estadual Paulista Júlio de Mesquita Filho, v. 33, n. 4, 2008. Citado na página 21.
- PARR, R.; YANG, W. *Density Functional Theory of Atoms and Molecules*, New York: Oxford Univ. [S.l.]: Press, 1989. Citado na página 43.
- PETTERSEN, E. F. et al. Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, Wiley Online Library, v. 25, n. 13, p. 1605–1612, 2004. Citado na página 50.
- REIS, R. A. G. *Estudo dinâmico conformacional da proteína calgranulina C (S100A12) mediante interação com íons e receptor RAGE*. Tese (Doutorado) — Universidade de São Paulo. Citado na página 46.
- REY, L. *Bases da parasitologia médica*. [S.l.]: Guanabara Koogan, 2002. Citado na página 19.
- ROE, D. R.; III, T. E. C. Ptraj and cpptraj: software for processing and analysis of molecular dynamics trajectory data. *Journal of chemical theory and computation*, ACS Publications, v. 9, n. 7, p. 3084–3095, 2013. Citado na página 55.
- SANT, C. M. R. et al. Métodos de modelagem molecular para estudo e planejamento de compostos bioativos: Uma introdução. *Revista Virtual de Química*, v. 1, n. 1, p. 49–57, 2009. Citado na página 21.
- SANTOS, L. H. S. d. et al. *Planejamento racional de fármacos aplicado à busca e otimização de inibidores do HIV-1 e doença de Chagas*. Tese (Doutorado). Citado 2 vezes nas páginas 26 e 29.
- SCHARFSTEIN, J. *Ativação do Sistema Cinina: um papel para a cruzipaina na imunopatogênese da doença de Chagas experimental*. Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Brasil., 2008. Disponível em: <<http://www.fiocruz.br/chagas/cgi/cgilua.exe/sys/start.htm-sid=27>>. Acesso em: 26 06 2008. Citado na página 28.
- SCHLICK, T. *Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide*. [S.l.]: Springer Science & Business Media, 2010. v. 21. Citado na página 33.
- SESPA. Belém: Secretaria Executiva de Saúde do Pará, 2014. Disponível em: <<http://www.saude.pa.gov.br>>. Acesso em: 21 fev 2017. Citado na página 20.

- SIMONS, J. An experimental chemist's guide to ab initio quantum chemistry. *The Journal of Physical Chemistry*, ACS Publications, v. 95, n. 3, p. 1017–1029, 1991. Citado na página 40.
- SOUTO-PADRON, T. et al. Cysteine proteinase in trypanosoma cruzi: immunocytochemical localization and involvement in parasite-host cell interaction. *Journal of Cell Science*, The Company of Biologists Ltd, v. 96, n. 3, p. 485–490, 1990. Citado na página 25.
- SOUZA, M. L. d. *Identificação de novos inibidores da enzima cruzaina de Trypanosoma cruzi candidatos a fármacos contra a doença de Chagas*. Tese (Doutorado) — Universidade de São Paulo, 2012. Citado na página 26.
- SPOEL, D. V. D. et al. Gromacs: fast, flexible, and free. *Journal of computational chemistry*, Wiley Online Library, v. 26, n. 16, p. 1701–1718, 2005. Citado na página 36.
- SZABO, A.; OSTLUND, N. S. *Modern quantum chemistry: Intro to advanced electronic structure theory*. Dover Publications, 1996. Citado na página 40.
- TARTARI, J. T. A.; NETO, V. A. O diagnóstico da forma aguda da moléstia de chagas. *Revista de Medicina*, v. 37, n. 206, p. 95–98, 1953. Citado na página 19.
- TEIXEIRA, D. E. et al. Interactive multimedia to teach the life cycle of trypanosoma cruzi, the causative agent of chagas disease. *PLoS neglected tropical diseases*, Public Library of Science, v. 6, n. 8, p. e1749, 2012. Citado na página 23.
- TURNER, P. Xmgrace, version 5.1. 19. *Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR*, 2005. Citado na página 55.
- URBINA, J. A. Specific chemotherapy of chagas disease: relevance, current limitations and new approaches. *Acta tropica*, Elsevier, v. 115, n. 1, p. 55–68, 2010. Citado 2 vezes nas páginas 19 e 24.
- VERLET, L. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical review*, APS, v. 159, n. 1, p. 98, 1967. Citado na página 38.
- VERLI, H. Análise de simulações de dinâmica molecular. *Grupo de Bioinformática Estrutural. Centro de Biotecnologia/Faculdade de Farmácia. Universidade Federal do Rio Grande do Sul. Porto Alegre*, 2009. Citado na página 46.
- VERLI, H. *Bioinformática: da biologia à flexibilidade molecular*. Sociedade Brasileira de Bioquímica e Biologia Molecular, 2014. Citado na página 21.
- WANG, J.; CIEPLAK, P.; KOLLMAN, P. A. How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules? *Journal of computational chemistry*, Wiley Online Library, v. 21, n. 12, p. 1049–1074, 2000. Citado 2 vezes nas páginas 45 e 52.
- WERMUTH, C. G. Strategies in the search for new lead compounds or original working hypotheses. *The practice of medicinal chemistry*, Academic Press, v. 3, p. 126–129, 2003. Citado na página 21.

- WHO. *Chagas disease (American trypanosomiasis)*[Internet]. Geneva: World Health Organization; 2015.[cited 2015 June]. 2015. Citado na página 19.
- WIGGERS, H. J. et al. Non-peptidic cruzain inhibitors with trypanocidal activity discovered by virtual screening and in vitro assay. *PLoS neglected tropical diseases*, Public Library of Science, v. 7, n. 8, p. e2370, 2013. Citado na página 28.
- YOUNG, D. C. *A Practical Guide for Applying Techniques to Real-World Problems*. [S.l.]: Wiley-Interscience New York:, 2001. Citado na página 33.
- ZANATTA, N. et al. Convergent synthesis and cruzain inhibitory activity of novel 2-(n-benzylidenehydrazino)-4-trifluoromethyl-pyrimidines. *Bioorganic & medicinal chemistry*, Elsevier, v. 16, n. 24, p. 10236–10243, 2008. Citado na página 25.