

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

**ESTUDO COMPUTACIONAL DA ENZIMA gGAPDH do
*TRYPANOSOMA CRUZI***

OSMAIR VITAL DE OLIVEIRA*

Tese apresentada como parte dos
requisitos para obtenção do título de
DOUTOR EM CIÊNCIAS (área de
concentração FÍSICO-QUÍMICA)

Orientador: Prof. Dr. Luiz Carlos Gomide Freitas

*Bolsista CNPq

São Carlos – SP

2009

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária/UFSCar**

O48ec

Oliveira, Osmair Vital de.

Estudo computacional da enzima gGAPDH do
Trypanosoma cruzi / Osmair Vital de Oliveira. -- São Carlos :
UFSCar, 2009.
197 f.

Tese (Doutorado) -- Universidade Federal de São Carlos,
2009.

1. Química quântica. 2. Dinâmica molecular. 3. Docking 4.
GAPDH. 5. *Trypanosoma cruzi*. I. Título.

CDD: 541.28 (20^a)

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Departamento de Química
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA
Curso de Doutorado

*Assinaturas dos membros da banca examinadora que avaliaram e aprovaram a defesa de tese de doutorado do candidato **Osmair Vital de Oliveira** realizada em 05 de junho de 2009:*




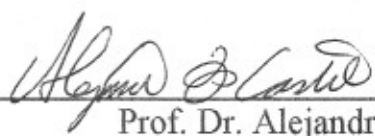
Prof. Dr. Luiz Carlos Gomide Freitas


Prof. Dr. Leo Degreve



Prof. Dr. Adriano Defini Andricopulo


Prof. Dr. Mauro Carlos Costa Ribeiro


Prof. Dr. Alejandro López Castilho

*Dedico este trabalho a minha esposa, a todos meus familiares e
aqueles que acreditaram na realização deste.*

AGRADECIMENTOS

Ao Prof. Dr. Luiz Carlos Gomide Freitas pela orientação, amizade e liberdade concedida ao longo da realização deste trabalho.

Ao Dr. Roberto Dias Lins pela orientação, amizade e incentivo ao longo do meu doutorado sanduíche realizado nos Estados Unidos.

Ao Prof. Dr. André Farias de Moura pela amizade, pelas sugestões e discussões, as quais enriqueceram este trabalho.

Aos amigos e colegas do Laboratório de Química Teórica: Daniella, Eduardo Franca, Eduardo Lopes, Kalil, Marcello, Odair, Raphael, Renata e Sara.

Aos amigos do Departamento de Química e LIEC, em especial ao Edcarlos, Fernando, Renato e Zimer.

Ao meu grande amigo carioca Luciano pelas conversas de “cafezinhos” e momentos de descontração.

As novas amizades realizadas em Richland – WA – USA: Thereza, Jason McDermott, Tjerk Straatsma, Cris, Daynne e Rom.

Ao Roberto e Thereza, que em momentos difíceis eles estavam ao meu lado.

Ao CNPq pela concessão da bolsa de estudo, tanto para os trabalhos realizados no Brasil e no exterior.

À FAPESP pelo financiamento dos recursos computacionais utilizados.

Ao PNNL (*Pacific Northwest National Laboratory*) pela concessão do espaço e recursos computacionais para a realização do doutorado sanduíche nos USA.

Ao Departamento de Química e ao Programa de Pós-Graduação em Química da UFSCar pelas instalações e recursos oferecidos.

Às secretárias do PPGQ-UFSCar pela eficiência e prontidão no atendimento.

LISTA DE ABREVIATURAS

ADT: AutoDockTools

AM1: *Austin Method 1*

1,3-BPG: 1,3-bifósforo glicerato

G3P: gliceraldeído 3-fosfato (do inglês, *glyceraldehyde 3-phosphate*)

gGAPDH: Gliceraldeído-3-fosfato desidrogenase glicossomal (do inglês, *Glycosomal Glyceraldehyde-3-Phosphate Dehydrogenase*)

DFT: Teoria do Funcional de Densidade (do inglês, *Density Functional Theory*)

DM: Dinâmica Molecular

d.p.: desvio padrão

ΔH_f : Calor de formação

$\Delta\Delta H_f$: Calor de formação relativo

LMOs: Orbitais moleculares localizados (do inglês, *Localized Molecular Orbitals*)

LSCF: Campo autoconsistente local (do inglês, *Local Self Consistent Field*)

MNDO: *Modified Neglect of Diatomic Overlap*

MOs: orbitais moleculares (do inglês, *Molecular Orbital*)

NAD⁺: Nicotinamida Adenina Dinucleotídeo

NDDO: *Neglect of Diatomic Differential Overlap*

PM3: *Parametric Method 3*

PM6: *Parametric Method 6*

PME: *Particle Mesh Ewald*

QM: Mecânica Quântica (do inglês, *Quantum Mechanics*)

QM/MM: Mecânica Quântica/Mecânica Molecular

QSAR: Relação quantitativa entre estrutura química e atividade

SCF: Campo autoconsistente (do inglês, *Self Consistent Field*)

ZDO: *Zero Differential Overlap*

LISTA DE TABELAS

- TABELA 3.1 - Comparação entre as diferentes subunidades e formas de gGAPDH, as estruturas obtidas a 20,0 ns de simulação foram consideradas (valores de RMSD estão em nm).....50
- TABELA 3.2 - Energia de interação (em $\text{kJ}\cdot\text{mol}^{-1}$) entre as subunidades. Em parênteses, o desvio padrão.....52
- TABELA 3.3 - Valores médios (em graus) do ângulo de abertura formados entre os domínios catalítico e de ligação. Em parênteses, o desvio padrão.....53
- TABELA 3.4 - Constante de inibição (K_i em μM) e em parênteses, a energia livre de ligação (kcal/mol) estimada pelo programa AutoDock4.....58
- TABELA 3.5 - Valores de K_i estimado e IC_{50} (em micromolar) (VIEIRA *et al.*, 2001) para três derivados da chalepina.....59
- TABELA 3.6 - Valores de K_i estimados e IC_{50} (em micromolar) (PEREIRA *et al.*, 2008) para três derivados da ácido anacárdico.....60
- TABELA 3.7 - Valores de constante de inibição e energia livre de ligação estimados nos cálculos de docking para a G3P.....63
- TABELA 3.8. Calor de formação (em $\text{kcal}\cdot\text{mol}^{-1}$) para o mecanismo proposto na literatura e o proposto neste trabalho.....81

LISTA DE FIGURAS

- FIGURA 1.1 - Distribuição global (em vermelho) da doença de Chagas.
(<http://www.nature.com/nrmicro/journal/v1/n1/full/nrmicro735.html>).....04
- FIGURA 1.2 - Drogas utilizadas no tratamento da doença de Chagas.....05
- FIGURA 1.3 - Representação esquemática da via glicolítica.
(http://www.biocarta.com/pathfiles/h_GLYCOLYSISPATHWAY.asp).....06
- FIGURA 1.4 - Representação do mecanismo da conversão da G3P em 1,3-BPG pela enzima gGAPDH. Figura adaptado do site: <http://home.inje.ac.kr/~biochemi/Lecture/1st/Biochem-II/ch02-1.pdf>.....07
- FIGURA 1.5 - Estrutura da enzima gGAPDH do *T. cruzi*. Figura gerada pelo programa Pymol (DeLano, 2002).....08
- FIGURA 2.1 - Diagrama das etapas previstas e as metodologias que serão utilizadas.....10
- FIGURA 2.2 - Caixa simples de simulação.....19
- FIGURA 2.3 - Condições periódicas de contorno em um sistema bidimensional.....20

- FIGURA 2.4 - Representação do raio de corte esférico para um sistema periódico bidimensional.....21
- FIGURA 2.5 - Algoritmo clássico de simulação.....22
- FIGURA 2.6 - Representação esquemática de docking de um ligante em uma proteína, formando o complexo proteína-ligante.....23
- FIGURA 2.7 - Modelo de ensemble docking ou múltiplas conformações. Em vermelho, o ligante chalepina.....26
- FIGURA 2.8 - Esquema do cálculo do SCF convencional (STEWART, 1996).....31
- FIGURA 2.9 - Esquema do cálculo do SCF utilizando orbitais moleculares localizados (STEWART, 1996).....33
- FIGURA 3.1 - Caixa inicial de simulação: 1 tetrâmero (representado em ribbons) + 36.100 águas (cor vermelha) + 20 Cl^{-1} (esferas em cor verde).....35
- FIGURA 3.2 - RMSD dos átomos $\text{C}\alpha$ da enzima gGAPDH na forma holo. A) Todo os átomos $\text{C}\alpha$ foram considerados; B) Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).....37
- FIGURA 3.3 - 1A) Sobreposição da gGAPDH. 2A, 2B, 2C e 2D) Sobreposição das subunidades A, B, C e D, respectivamente. Cada seta indica a posição do NAD^{+} no sítio ativo. Na parte superior direita estão

os valores de RMSD das respectivas subunidades. As estruturas em vermelho e ciano representam, respectivamente, a enzima inicial (0 ns) e final (20,0 ns) da simulação.....38

FIGURA 3.4 - Sobreposição das subunidades A, B, C e D a 20,0 ns com suas respectivas a 0,0 ns de simulação. NAD⁺ a 20,0 ns em cor ciano e a 0,0 ns em cor amarelo. Para melhor visualização, apenas a enzima a 0,0 ns (cor verde) foi considerada.....39

FIGURA 3.5 - RMSD dos átomos C α para o domínio de ligação do NAD⁺ (resíduos 1-150 e 331-359, linha cinza) e para o domínio catalítico (resíduos 151-330, linha preta) das subunidades A, B, C e D da enzima gGADPH.....40

FIGURA 3.6 - Ligações de hidrogênio entre resíduos da enzima e o cofator NAD⁺ ao longo do tempo de simulação.....41

FIGURA 3.7 - Distância mínima de contato entre Cys166-NAD⁺ (linha cinza) e Cys166-His194 (linha preta).....43

FIGURA 3.8 - Sobreposição das subunidades A, B, C e D a 20,0 ns (cor ciano) com suas respectivas subunidades a 0,0 ns (cor amarelo) de simulação.....44

FIGURA 3.9 - Energia de interação entre as subunidades. A-B (preto), C-D (marron), A-D (verde), B-C (azul), A-C (vermelho) e B-D (amarelo).....45

- FIGURA 3.10 - RMSD dos átomos $C\alpha$ da enzima gGAPDH na forma apo. A) Todo os átomos $C\alpha$ foram considerados; B) Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).....47
- FIGURA 3.11 - RMSD dos átomos $C\alpha$ da enzima gGAPDH na forma holo e apo. Topo: todos os átomos $C\alpha$ da enzima foram considerados; Abaixo: subunidades A, B, C e D. gGAPDH na forma holo em cor preta e na forma apo em cor cinza.....48
- FIGURA 3.12 - RMSD dos átomos $C\alpha$ do domínio catalítico (a esquerda, resíduos 151-330) do domínio de ligação do NAD^+ (a direita, resíduos 1-150 e 331-359) e das subunidades A, B, C e D da enzima gGADPH referente a estrutura inicial. Em cor preta, forma holo e em cinza, forma apo.....49
- FIGURA 3.13 - Energia de interação entre as subunidades da gGAPDH na forma apo. A-B (preto), C-D (marron), A-D (verde), B-C (azul), A-C (vermelho) e B-D (amarelo).....51
- FIGURA 3.14 - Representação do ângulo de abertura. Os átomos $C\alpha$ são representados pelas esferas.....53
- FIGURA 3.15 – Região utilizada como critério para selecionar o ligante a partir dos cálculos de docking. A) para a forma apo e B) para a forma holo.....56
- FIGURA 3.16 - Estrutura molecular da chalepina e ácido anacárdico, respectivamente.....57

FIGURA 3.17 - Derivados da chalepina considerados nos cálculos de docking.....	59
FIGURA 3.18 - Derivados do ácido anacárdico considerados nos cálculos de docking.....	60
FIGURA 3.19 - Primeira etapa do processo catalítico da gGAPDH.....	63
FIGURA 3.20 - Interação do G3P com o NAD ⁺ e gGAPDH, obtidos para a estrutura com melhor valor de K _i	64
FIGURA 3.21 - RMSD dos átomos C α da enzima gGAPDH do sistema gGAPDH-NAD ⁺ -G3P. Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).....	65
FIGURA 3.22 - Distância entre o centro de massa do G3P e da parte nicotinamida do cofator NAD ⁺ , Cys166 e His194 ao longo do tempo de simulação. Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).....	66
FIGURA 3.23 - <i>Snapshot</i> obtido a 1.000 ps de simulação. Para melhor visualização, os hidrogênios apolares foram removidos.....	67
FIGURA 3.24 - Ligações de hidrogênio entre resíduos da enzima e o G3P ao longo do tempo de simulação.....	67

- FIGURA 3.25 - *Snapshot* obtido a 1000 ps de simulação mostrando os resíduos que interagem via ligação de hidrogênio com o G3P. Para melhor visualização, os hidrogênios apolares foram removidos.....68
- FIGURA 3.26 - Alinhamento sequencial das enzimas GAPDH do T. Cruzi (1K3T_A), T. Brucei (1GGA_O) e L. Mexicana (1GYP_A). Em asterístico, os resíduos conservados nos diferentes organismos.....70
- FIGURA 3.27 1- Segunda etapa do mecanismo enzimático da gGAPDH.....72
- FIGURA 3.28 2- *Snapshot* obtido a 719 ps de simulação por dinâmica molecular é utilizado para representar o mecanismo enzimático da gGAPDH proposto na literatura.....73
- FIGURA 3.29 - Região selecionada para o cálculo de química quântica. Parte fixa (344 átomos) está em vermelho. Parte flexível (354 átomos): resíduos (verde), G3P (amarelo) e NAD⁺ (magenta).....75
- FIGURA 3.30 - Calor de formação em função da coordenada de reação. MOZYME (linha cinza) e método convencional (linha preta).....76
- FIGURA 3.31 - Estrutura inicial otimizada (I), intermediário na transição Cys166-H-His194 (II), transferência completa do H da Cys166 para a His194 (III) e estrutura obtida na final da coordenada de reação (IV).....77

- FIGURA 3.32 - *Snapshot* obtido a 530 ps de simulação por dinâmica molecular e mecanismo enzimático da gGAPDH proposto no presente trabalho.....78
- FIGURA 3.33 - Calor de formação em função da coordenada de reação: MOZYME (linha cinza) e método convencional (linha preta).....79
- FIGURA 3.34 - Estrutura inicial otimizada (I), transferência do H da Cys166 para His194 (II) e intermediário tiohemiacetal formado (III)...80

RESUMO

ESTUDO COMPUTACIONAL DA ENZIMA gGAPDH DO *TRYPANOSOMA CRUZI*.

Métodos de química computacional foram utilizados para estudar a enzima gGAPDH do *Trypanosoma cruzi*, protozoário responsável pela doença de Chagas. Simulações por dinâmica molecular foram realizadas para obter a evolução temporal da enzima gGAPDH na forma holo (com o cofator NAD^+) e apo (sem o cofator) em solução. Os cálculos foram efetuados no *ensemble* NpT a $T = 300\text{K}$ e $p = 1\text{bar}$. Nessas simulações ficaram evidenciadas pequenas mudanças conformacionais entre as formas holo e apo no tempo de simulação de 20,0 ns. Para posicionar ou encaixar algumas moléculas candidatas a fármacos no sítio ativo da enzima, cálculos de *docking* foram realizados utilizando a metodologia de *ensemble docking*, utilizando múltiplas conformações obtidas a cada intervalo de 2,0 ns de simulação da enzima na forma holo e apo. Essa metodologia foi adotada para levar em consideração a flexibilidade da enzima. Os resultados desses cálculos mostraram que a melhor estratégia para elaborar compostos candidatos a fármacos é considerar as duas formas da enzima (holo e apo). Com esta estratégia, um composto inibiria a forma apo e outro, a forma holo. Para caracterizar o mecanismo enzimático, o gliceraldeído 3-fosfato (G3P) foi colocado no sítio ativo (via cálculos de *docking*) da enzima gGAPDH na forma holo, a qual foi obtida a 20,0 ns de simulação. Nesse sistema (gGAPDH- NAD^+ -G3P), foi realizado 1,0 ns de simulação por dinâmica molecular para a relaxação e equilibração do sistema. Ao longo dessa simulação, duas estruturas foram escolhidas para a realização de cálculos de química quântica relacionados com a atividade enzimática. Em uma dessas estruturas, um dos mecanismos propostos na literatura foi caracterizado a partir da coordenada de reação de transferência do próton da Cys166 para a His194. Estes cálculos foram efetuados utilizando um procedimento que combina a utilização da metodologia

MOZYME com o cálculo SCF usual implementado no programa MOPAC2009. Nestes cálculos, a parametrização PM6 foi utilizada. A partir de uma segunda estrutura obtida da dinâmica molecular, um novo mecanismo para a enzima gGAPDH foi proposto e caracterizado via cálculos quânticos. Nesse mecanismo, o grupo hidroxila da G3P age como uma ponte para a transferência do próton da Cys166 para a His194.

ABSTRACT

COMPUTATIONAL STUDY OF THE gGAPDH ENZYME FROM *TRYPANOSOMA CRUZI*.

Theoretical methods in computational chemistry were used to study the gGAPDH enzyme from *Trypanosoma cruzi*. This protozoan is responsible by the Chagas' diseases. Molecular dynamics simulations were performed to obtain the time evolution of the gGAPDH enzyme in the holo (with the cofactor NAD^+) and apo (without the cofactor) forms in aqueous solution. The calculations were performed in the NpT ensemble with $T = 300\text{K}$ e $p = 1\text{bar}$. In these simulations little conformational changes were observed in both holo and apo forms of the enzyme along 20.0 ns simulation time. Docking calculations were carried out to fit some drugs in enzyme active site using the *ensemble docking* methodology. Therefore, multiple enzyme conformations of both holo and apo forms were obtained at time intervals of 2.0 ns along the molecular dynamics simulation. This procedure was used to take in into account the flexibility of the enzyme. The results from these calculations indicate that the best way to develop a drug molecule is to consider both enzyme forms (holo and apo forms). In this way, one drug will inhibit the apo form and other, the holo form. To characterize the enzymatic mechanism, the glyceraldehyde 3-phosphate (G3P) was placed in the active site (via docking calculations) of the holo form conformation obtained in the end of the 20.0 ns trajectory. A 1.0 ns molecular dynamic simulation was performed in the (gGAPDH- NAD^+ -G3P) system. Quantum chemical calculations were performed to study reactive process in the enzyme catalytic site. A convenient model was built using about 698 atoms carefully chosen to represent the active site and its surroundings. The calculations were performed using a combination of MOZYME and the usual SCF procedure implemented in the MOPAC2009 program. The calculations were performed at the PM6 level. One of the reaction mechanisms proposed in the literature was characterized

calculating the energy profile along the reaction path to transfer a proton from Cys166 to His194. Using a second structure obtained from molecular dynamics, a new mechanism for the gGAPDH enzyme was proposed and characterized by a similar MOPAC2009 calculation. In this mechanism, the hydroxyl group of the G3P acts as a bridge to transfer the proton from the Cys166 to His194 residue.

SUMÁRIO

Capítulo 1 – Introdução

1.1 - Química Computacional.....	01
1.2 - Doença de Chagas.....	03
1.3 – Objetivos.....	09

Capítulo 2 – Metodologia

2.1 - Dinâmica Molecular.....	11
2.1.1 – Energia Potencial.....	11
2.1.1.1 - Potencial Harmônico de Estiramento.....	13
2.1.1.2 - Potencial Harmônico Angular.....	13
2.1.1.3 - Potencial Torsional	14
2.1.1.4 - Potencial de Lennard-Jones.....	14
2.1.1.5 - Potencial de Coulomb.....	15
2.1.2 - Algoritmo de Integração.....	16
2.1.3 - Controle da Temperatura e da Pressão.....	17
2.1.4 - Condições Periódicas de Contorno.....	18
2.1.5 – Algoritmo de Dinâmica Molecular.....	21
2.2 – <i>Docking</i>	23
2.3 - Química Quântica.....	27

Capítulo 3 – Resultados e Discussão

3.1 - Etapa 1: Dinâmica Molecular.....	34
3.1.1 - Sistema gGAPDH-NAD ⁺	34
3.1.1.1 - Análise estrutural.....	36

3.1.1.2 - Interação gGAPDH-NAD ⁺	41
3.1.2 - Sistema gGAPDH.....	46
3.1.2.1 - Análise estrutural.....	47
3.2 - Etapa 1: <i>Docking</i>	55
3.3 - Etapa 2: <i>Docking</i>	63
3.4 - Etapa 3: Dinâmica Molecular.....	65
3.5 - Etapa 4: Química Quântica.....	72
Capítulo 4 – Conclusões.....	82
Capítulo 5 - Sugestões para Futuros Trabalhos.....	85
Capítulo 6 – Resumo do Doutorado Sanduíche.....	87
Capítulo 7 – Referências Bibliográficas.....	89

CAPÍTULO 1

INTRODUÇÃO

1.1 - Química Computacional

As enzimas são de fundamental importância para o funcionamento dos seres vivos. Portanto, compreender as relações entre estrutura e função de enzimas é fundamental para entender a intrincada teia dos processos vitais. Entre os vários processos vitais, estudar o mecanismo das reações enzimáticas é uma tarefa importante e decisiva para a compreensão dos mecanismos moleculares em sistemas biológicos. Em reações enzimáticas estes estudos implicam conhecer o entendimento das interações entre o substrato e o sítio catalítico da enzima, bem como de outros fatores que influem nesta interação. Entre estes fatores pode-se destacar a composição e estrutura do sítio ativo e efeitos dinâmicos do sistema. Ressalta-se que entender a sinergia destes processos é de real importância para o desenvolvimento de estratégias de intervenção em processos patológicos. Entre as intervenções cabe destacar o desenvolvimento de novos fármacos. Para o melhor entendimento dos processos enzimáticos é necessário tratar o sistema ao nível atômico e molecular. Entretanto, são raras as técnicas experimentais que revelam detalhes completos da dinâmica destes sistemas em escala atômica. As poucas técnicas existentes possuem custo elevado e dificuldades práticas para a aplicação generalizada, sendo, portanto pouco acessíveis à totalidade da comunidade científica. Deste modo, com o avanço e popularização de computadores cada vez mais velozes e de baixo custo, a química computacional consolida-se gradativamente como uma área de pesquisa fundamental para o entendimento de processos químicos a nível atômico e molecular. A sinergia existente entre o desenvolvimento de hardware & software cria novos horizontes para a aplicação em pesquisa científica. Os

avanços obtidos têm sido de grande valia para os químicos teóricos, proporcionando também o surgimento de interfaces de colaboração direta com a pesquisa experimental.

Dentre as diversas metodologias utilizadas em química teórica na sua vertente conhecida como ‘química computacional’, pode-se destacar métodos mecânico-quânticos: *ab initio*, DFT e semi-empíricos (LEVINE, 1991) e métodos de mecânica estatística (ALLEN, 1987): Dinâmica Molecular (DM) e estocásticos (Monte Carlo). Em geral, a escolha do método a ser utilizado está fortemente relacionada com o conjunto de propriedades que se deseja obter e o com o número de partícula que compõem o modelo utilizado para representar o sistema a ser estudado. Os métodos de química quântica (*ab initio* e DFT) tem sido de grande valia para precisar propriedades eletrônicas e estruturais de moléculas. Entretanto, a utilização destes métodos esta restrita a sistemas com poucas dezenas ou mesmo centenas de átomos. Para facilitar a realização de cálculos quânticos em sistemas com centenas de átomos, métodos semi-empíricos foram desenvolvidos. Com estes métodos, devido às aproximações serem mais drásticas que nos métodos *ab initio* e DFT, ocorre uma natural perda de precisão, mas obtém-se um ganho em escala de tamanho do modelo e em tempo computacional. Apesar desses métodos semi-empíricos fornecerem uma descrição quântica de sistemas relativamente sistemas macromoleculares, tais como proteínas, requer nova uma nova abordagem. Baseado nessa problemática, um novo algoritmo para a resolução de equações nos métodos quânticos semi-empíricos, MOZYME (STEWART, 1996), foi desenvolvido. Este novo método permite tratar sistemas com milhares de átomos, o que tornou viável estudar proteínas ou regiões específicas destas ao nível quântico. Apesar de todos esses avanços, o efeito do solvente, que é de grande importância (*e.g.* em sistemas biológicos), tem que ser desprezado ou introduzido via aproximações (*e.g.* método de solvente contínuo) nos métodos quânticos. Por outro lado, os métodos de mecânica estatística que utilizam campos de força são muito úteis

para a obtenção de propriedades termodinâmicas e estruturais de líquidos e macromoléculas. A eficiência das metodologias permite que um número relativamente grande de partículas seja convenientemente tratado. Porém, a grande desvantagem destes métodos é que eles não podem fornecer uma descrição de propriedades eletrônicas, o que impede aplicações em estudos de reações químicas, espectroscopia, processos catalíticos e outros. Portanto, uma solução é encontrar formas adequadas para a utilização simultânea e/ou seqüencial de métodos clássicos e quânticos. Assim, utilizando um conjunto híbrido de técnicas computacionais seria possível mapear um estudo completo de um problema químico específico. Dentro deste escopo, no presente trabalho utilizaremos métodos de química computacional para: i) estudar a enzima gGAPDH (*glycosomal glyceraldehyde-3-phosphate dehydrogenase*) do protozoário *Trypanosoma cruzi*, causador da doença de Chagas; ii) caracterizar o comportamento dinâmico desta enzima na presença e ausência do cofator NAD^+ ; iii) utilizar este conjunto de informações para entender a interação desta enzima com moléculas bloqueadoras conhecidas; iv) sugerir e investigar novas estratégias para a proposição de novos bloqueadores eficientes.

1.2 - Doença de Chagas

A cada ano, 200.000 pessoas são contaminadas pela doença de Chagas e 40 milhões de pessoas encontram-se em risco de contrair a doença. Desse total, cerca de 13.000 pessoas morrem em decorrência de complicações da doença (WHO, 2002; MONCAYO, 2003; URBINA *et al.*, 2003; MONCAYO, *et al.*, 2006). A doença de Chagas é causada pelo protozoário *Trypanosoma cruzi*, o qual é transmitido principalmente pelo inseto-vetor hematófago conhecido popularmente como “barbeiro”. Esse inseto é classificado na ordem *Hemiptera* e família *Reduviidae*, sendo encontrado com frequência em frestas de paredes de barro ou de madeira de lares na zona rural e

outras. Uma das maneiras mais eficientes que tem sido utilizado no controle da doença de Chagas é a eliminação de ambientes favoráveis a procriação do inseto vetor, bem como a eliminação deste por meio de inseticidas.

Outras formas de transmissão da doença ocorrem por transfusão de sangue e transmissão congênita (WHO, 2002). A figura a seguir apresenta as áreas de incidências e de riscos da doença de Chagas.



FIGURA 1.1 - Distribuição global (em vermelho) da doença de Chagas.
(<http://www.nature.com/nrmicro/journal/v1/n1/full/nrmicro735.html>).

Esta doença foi descoberta em 1909 pelo médico sanitário Carlos Chagas. Apesar de bem conhecida há de 100 anos, até os dias atuais, existem apenas dois fármacos contra a doença: benzonidazol e nifurtimox. Entretanto, estes apresentam baixa eficiência e fortes efeitos colaterais e, (COURA, *et al.*, 2002), além disso, apresentam atividade inibitória apenas na fase aguda da doença. No Brasil, a comercialização do composto nifurtimox é proibida (FAIRLAMB, 1993; SCHOFIELD, *et al.*, 2006). Estes fármacos são apresentados na figura a seguir.

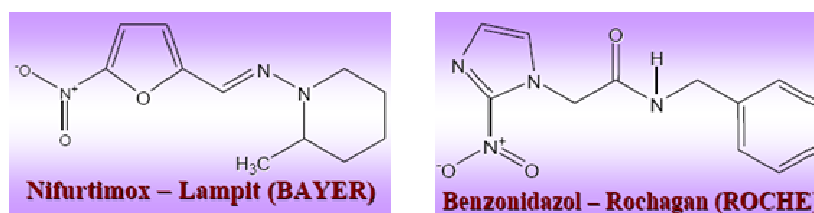


FIGURA 1.2 - Drogas utilizadas no tratamento da doença de Chagas.

Vários processos bioquímicos têm sido apontados como alvos terapêuticos potenciais (RODRIGUEZ, 2001; DOCAMPO, 2001), dentre eles temos: a enzima tripanotona redutase (SALMON, *et al.*, 1998); enzimas envolvidas na importação e interconversão de purinas (FAIRLAMB, 1999); metabolismo de folato e pteridina (FAIRLAMB, 1999); biossíntese de RNA mensageiro (FAIRLAMB, 1999); biossíntese de lipídios; transialidase (FAIRLAMB, 1999); diidrofolato redutase (ZUCCOTTO *et al.*, 1999); enzimas do glicossomo, envolvidas no metabolismo energético, destacando-se a gliceraldeído-3-fosfato desidrogenase (SOUZA, *et al.*, 1998) e outros. Devido à alta dependência das formas infectantes de *T. cruzi* da via glicolítica para a produção de ATP (ENGEL, *et al.*, 1987), esta tem sido tornado como alvo atrativo para o desenvolvimento de novos fármacos. Na figura a seguir temos a representação da via glicolítica.

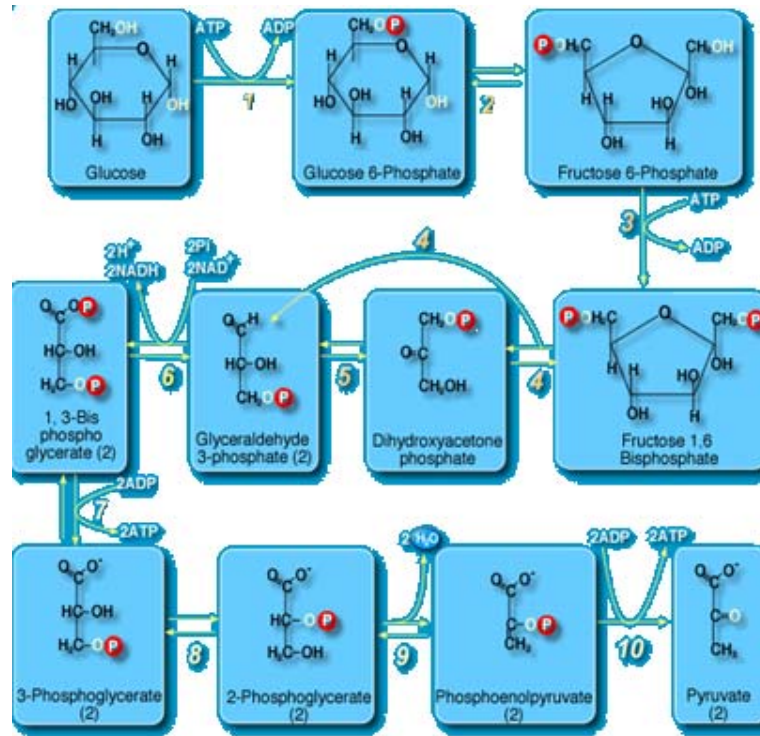


FIGURA 1.3 - Representação esquemática da via glicolítica.

(http://www.biocarta.com/pathfiles/h_GLYCOLYSISPATHWAY.asp).

A gGAPDH do *T. cruzi*, FIGURA 1.5, é um homotetrâmero constituído de 359 aminoácidos e é uma enzima da via glicolítica que catalisa a fosforilação oxidativa de gliceraldeído 3-fosfato em 1,3-bisfosfoglicerato, na presença de fosfato inorgânico e NADH^+ como coenzima. Na figura a seguir temos um mecanismo enzimático que ocorre na enzima gGAPDH.

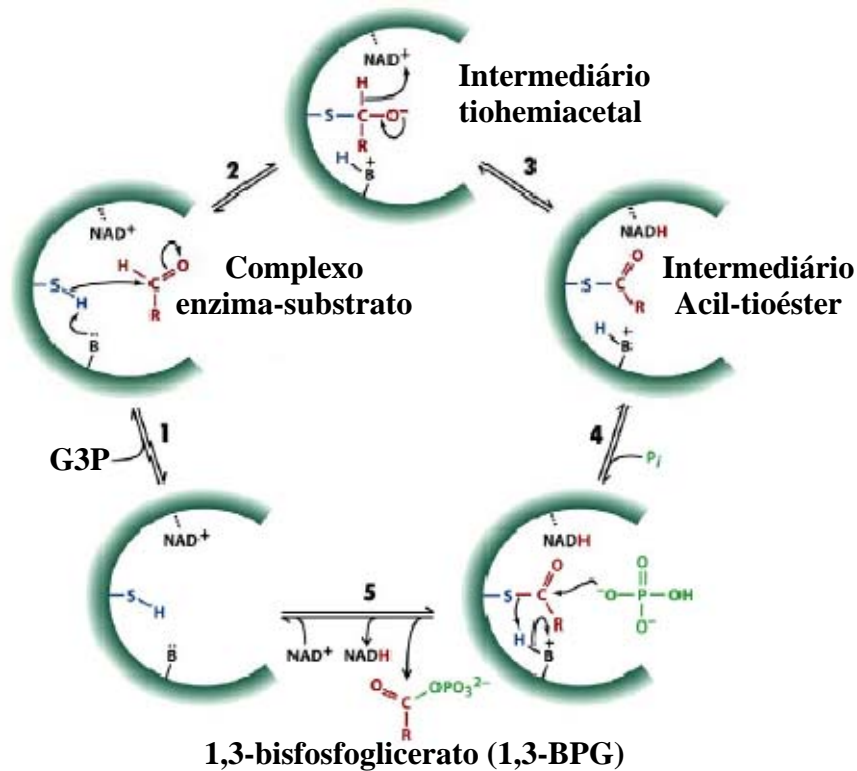


FIGURA 1.4 - Representação do mecanismo da conversão da G3P em 1,3-BPG pela enzima gGAPDH. Figura adaptado do site:

<http://home.inje.ac.kr/~biochemi/Lecture/1st/Biochem-II/ch02-1.pdf>

A primeira estrutura cristalográfica da enzima gGAPDH do *T. cruzi* foi determinada por SOUZA (SOUZA *et al.*, 1998). A partir da análise e comparação dessa com a homóloga humana, diferenças estruturais importantes, principalmente no sítio de ligação do cofator NAD^+ e no sítio ativo, ficaram evidenciadas (SOUZA *et al.* 1998). Deste modo, estas regiões são atraentes para o desenvolvimento de novos inibidores específicos, que não apresentem afinidade pela enzima humana. Trabalhos promissores têm mostrado que compostos flavonóides (TOMAZELA *et al.*, 2000) e cumarínicos (VIEIRA, *et al.*, 2001) apresentam atividade inibitória da enzima gGAPDH. Até a presente data, existem três estruturas cristalográfica da enzima gGAPDH determinadas e depositadas no *Protein Data Bank* (www.rcsb.org) com código PDB: 1QXS (CASTILHO *et al.* 2003), 1ML3 (CASTITLHO *et al.*, 2003) e 1K3T (PAVÃO

et al., 2002). A seguir, temos a estrutura da enzima gGAPDH obtida a partir do código PDB:1K3T.

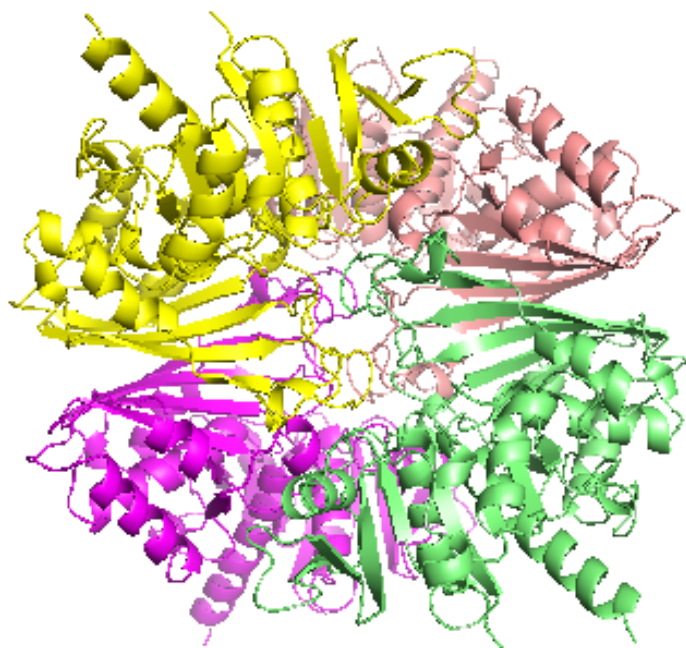


FIGURA 1.5 - Estrutura da enzima gGAPDH do *T. cruzi*. Figura gerada pelo programa Pymol (DeLano, 2002).

Vários trabalhos experimentais têm sido reportados na literatura focalizando a enzima gGAPDH como alvo de novos inibidores do parasita. Dentre destes novos candidatos a inibidores podemos citar: compostos derivados da cumarina (JUNIOR *et al.*, 2005; MARCHI *et al.*, 2004); complexos de rutênio (SILVA *et al.*, 2007; SILVA *et al.*, 2009); derivados do ácido anacárdico (PEREIRA *et al.*, 2008, FREITAS *et al.*, 2009); análogos de *bi-substrate* (LADAME *et al.*, 2005); derivados de adenosina (BRESSI *et al.*, 2001) e outros. Por outro lado, vários trabalhos empregando metodologias teóricas têm sido utilizados para a elaboração de novos compostos candidatos a fármacos (LEITÃO *et al.*, 2004; GUIDO *et al.*, 2008; FREITAS *et al.*, 2009).

Apesar de todos esses estudos, nenhum composto candidato a fármaco que apresentasse valor de IC_{50} (concentração para inibir 50% da enzima) na escala de nanomolar foi encontrado. Este quesito é importante para

que a realização posterior de testes *in vivo*. O insucesso dessas pesquisas pode ser atribuído, principalmente, à dificuldade em entender, ao nível atomístico, o comportamento dinâmico da enzima em solução, principalmente na região do sítio ativo. Procurando preencher esta lacuna, na presente tese de doutorado, métodos de química computacional foram utilizados para investigar o comportamento estrutural e energético da gGAPDH em solução ao nível atomístico.

1.3 – Objetivos

Na presente tese de doutorado, métodos de química computacional foram utilizados para fazer um estudo detalhado da enzima gGAPDH em solução. Esta enzima é importante para a geração de energia do *T. cruzi*, responsável pela doença de Chagas. Utilizando o método de dinâmica molecular com campo de força clássico, propriedades serão determinadas a enzima na forma *apo* e *holo*, ou seja, com e sem o cofator NAD^+ no sítio ativo da enzima. Um conjunto de estruturas obtido ao longo da trajetória será utilizado para a realização de cálculos de *docking*, uma abordagem que leva em consideração a flexibilidade da enzima. Esta tarefa é árdua devido ao alto custo computacional em cálculos de *docking*. Visando considerar o efeito eletrônico no mecanismo e caminho de reação que ocorre no sítio ativo da enzima, métodos de química quântica serão utilizados.

CAPÍTULO 2

METODOLOGIA

Na figura a seguir apresentamos um diagrama mostrando a sequência de etapas que serão realizadas e bem com o tipo de metodologia que será utilizada em cada uma delas.

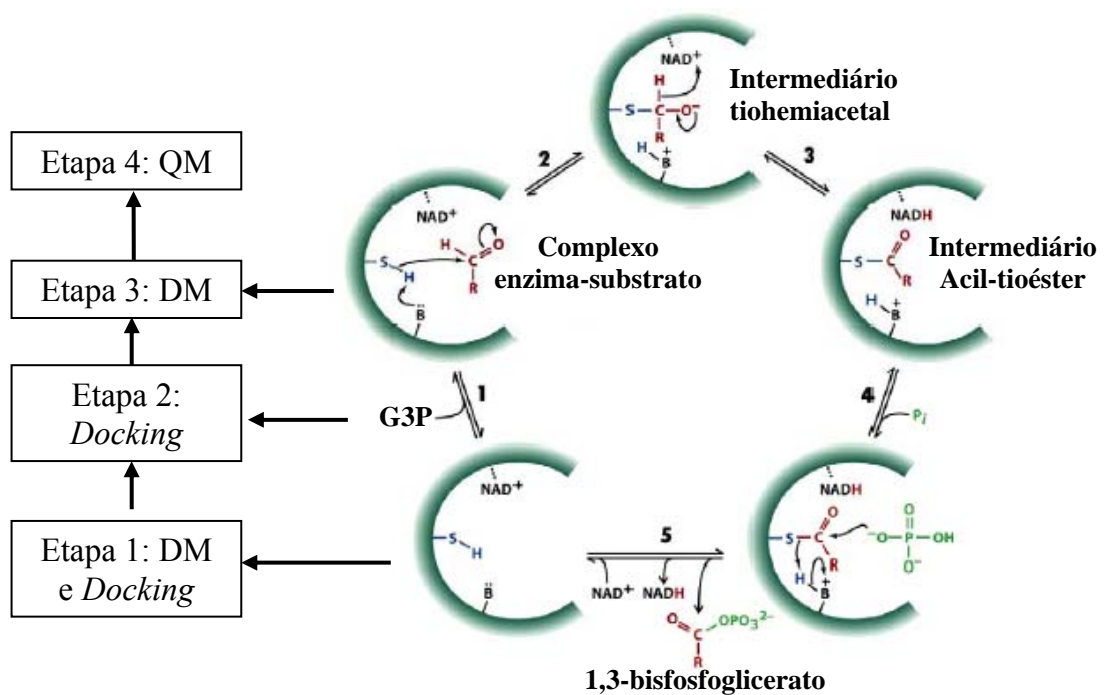


FIGURA 2.1 - Diagrama das etapas previstas e as metodologias que serão utilizadas.

2.1 - Dinâmica Molecular

Dinâmica molecular (DM) é uma metodologia utilizada para resolver problemas em mecânica estatística na impossibilidade da solução analítica das equações. Em sua forma geral, a DM é utilizada para simular a evolução temporal de um sistema sob a ação de forças conhecidas. Esta metodologia recorre à integração das equações de movimento clássicas para determinar aceleração, velocidade e posição temporal das partículas do sistema (ALLEN, 1987).

Dado a superfície de potencial, a força atuando em cada partícula do sistema pode ser calculada a partir da derivada deste potencial e as equações de movimento podem ser estabelecidas segundo as leis de Newton (ALLEN, 1987). A integração destas equações fornece as coordenadas das partículas em função do tempo, ou seja, as trajetórias. Este conjunto de trajetórias é posteriormente utilizado para calcular valores médios de propriedades mecânicas do sistema que podem ser escritas como função da configuração espacial dos componentes.

2.1.1 – Energia Potencial

Em geral, as moléculas que compõem um sistema químico complexo não podem ser consideradas rígidas: vibrações moleculares, torções em ângulos de ligação e diédricos precisam ser incluídas para uma representação realística do sistema. Na maioria dos campos de força, estas contribuições para a energia total do sistema são representadas por funções harmônicas (estiramentos de ligação, ângulos de ligação, diedro impróprios) e por séries de Fourier (energias de torção de ângulos diédricos). Potenciais de Lennard-Jones e Coulombicos são utilizados para representar interações não ligantes (LEACH, 1996). Deste modo, para um sistema constituído por N átomos com posições r_i ,

r_2, \dots, r_n , a força sobre um átomo i , e aceleração resultante da mesma, podem ser calculadas em função do potencial total de interação, $V(r_1, r_2, \dots, r_n)$:

$$F_i = - \frac{\partial V(r_1, r_2, \dots, r_n)}{\partial r_i} \quad (1)$$

$$\frac{d^2 r_i(t)}{dt^2} = \frac{F_i}{m_i} \quad (2)$$

onde t é o tempo, F_i , a força sobre o átomo i , e m_i a sua massa.

Um dos quesitos fundamentais em dinâmica molecular é definir o potencial $V(r_1, r_2, \dots, r_n)$, da equação anterior. Em geral, este potencial é dividido em partes, cada qual construída e parametrizada para representar aspectos da realidade física do problema. A seguir oferecemos uma descrição da função potencial total.

$$V_{total} = V_{ligante} + V_{não-ligante} \quad (3)$$

com mais detalhes,

$$V_{total} = V_d + V_\theta + V\varphi + V_{LJ} + V_C \quad (4)$$

Sendo:

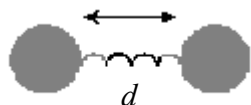
$$\left. \begin{array}{l} V_d : \text{potencial harmônico estiramento} \\ V_\theta : \text{potencial angular} \\ V\varphi : \text{potencial torsional} \end{array} \right\} V_{ligante}$$

$$\left. \begin{array}{l} V_{LJ} : \text{potencial de Lennard-Jones} \\ V_C : \text{potencial de Coulomb} \end{array} \right\} V_{não-ligante}$$

A conceituação básica dos termos de potencial apresentados anteriormente é apresentada a seguir.

2.1.1.1 - Potencial Harmônico de Estiramento

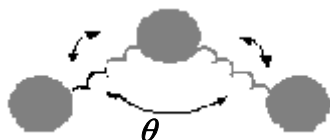
Ou potencial da ligação química, V_d , descreve a energia associada ao desvio da distância de equilíbrio d_0 entre dois átomos ligados covalentemente,

$$V_d = \frac{1}{2} k_d (d - d_0)^2 \longrightarrow \text{Diagrama de dois átomos ligados com distância } d \text{ e constante elástica } k_d \text{ (5)}$$


Onde d é o comprimento da ligação entre dois átomos i e j , d_0 é o comprimento de equilíbrio e k_d é a constante elástica da força. Logo, o potencial V_d mede a distorção da ligação em relação à distância de equilíbrio d_0 .

2.1.1.2 - Potencial Harmônico Angular

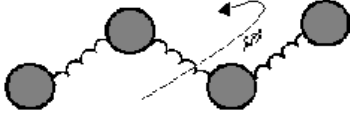
Este potencial, V_θ , está associado a desvios do ângulo de equilíbrio θ_0 , formado entre três átomos, i - j - k , covalentes e consecutivamente ligados,

$$V_\theta = \frac{1}{2} k_\theta (\theta_n - \theta_0)^2 \longrightarrow \text{Diagrama de três átomos ligados formando um ângulo } \theta \text{ e constante elástica } k_\theta \text{ (6)}$$


Onde θ é o ângulo definido pela ligação i - j - k , θ_0 é o ângulo de equilíbrio e k_θ é a constante elástica da força.

2.1.1.3 - Potencial Torsional


O potencial torsional V_φ representa explicitamente a torção própria de ligações covalentes entre dois átomos. Para um conjunto de átomos ligados i, j, k, l , o potencial torsional para rotação em torno de um eixo definido pela ligação $j-k$ pode ser definido por:

$$V_\varphi = k_\varphi [1 + \cos(p\varphi + \delta)] \rightarrow \text{Diagrama de rotação} \quad (7)$$


Onde k_φ é a constante que define a altura da barreira de rotação, p é o número de mínimos para a torção de uma ligação química específica, φ é o ângulo diedral para a ligação central em uma seqüência de quatro átomos e δ é a defasagem no ângulo diedral que pode colocar na posição φ igual a zero um ponto máximo ou de mínimo.

2.1.1.4 - Potencial de Lennard-Jones

O potencial de Lennard-Jones que representa a interação entre dois átomos i e j não ligados e é dado por:

$$V_{LJ}(r_{ij}) = \sum_{i < j} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \rightarrow \text{Diagrama de interação} \quad (8)$$


Onde r_{ij} é a distância entre os átomos i e j e A_{ij} e B_{ij} são parâmetros que dependem do tipo de átomo envolvido na interação, os quais são definidos por: $A_{ij} = 4\varepsilon_{ij}\sigma_{ij}^{12}$ e $B_{ij} = 4\varepsilon_{ij}\sigma_{ij}^6$, e o par (ε, σ) deve ser explicitado para cada sítio i e j . O termo para a interação $i-j$ é calculado utilizando regras de cruzamento. É conveniente ressaltar que este potencial é utilizado geralmente para descrever a interação de átomos separados por três ou mais ligações covalentes, bem como

calcular contribuições para interações intermoleculares. Neste potencial, o primeiro termo descreve a repulsão que ocorre quando há sobreposição de densidade eletrônica com a aproximação dos dois átomos i - j , e segundo termo representa interações devido à formação de dipolos induzidos, instantâneos, originados por flutuações na densidade de carga eletrônica dos átomos.

2.1.1.5 - Potencial de Coulomb

Devido a possíveis assimetrias na distribuição de carga, a interação eletrostática entre dois átomos i e j é representada pelo potencial de Coulomb:

$$V_C = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \rightarrow \begin{array}{c} \text{+} \\ \text{---} \\ \text{-} \end{array} \quad (9)$$

Onde r_{ij} é a distância entre os átomos i e j , q_i e q_j são cargas parciais, ϵ_0 é a permissividade dielétrica do vácuo.

Usualmente, os parâmetros empregados na definição de cada termo do potencial são determinados utilizando cálculos de mecânica quântica e/ou por ajustes a dados experimentais (cristalografia de raios-X, espectroscopia de infravermelho, densidade de líquidos, calores de vaporização, ressonância magnética nuclear, etc). O conjunto desses parâmetros é chamado de campo de força. A qualidade do campo de forças influencia diretamente os resultados da simulação.

2.1.2 - Algoritmo de Integração

Dada a energia potencial total do sistema, é necessário calcular as forças e resolver as equações de movimento de Newton para obter a evolução temporal do sistema. Entretanto, trata-se da resolução de uma equação diferencial acoplada de segunda ordem. Como esta equação não possui solução analítica, o problema é resolvido utilizando-se uma aproximação numérica. Em geral, os métodos numéricos para a solução desta classe de equações diferenciais de segunda ordem estão fundamentados em técnicas de diferenciação finita (GEAR, 1971). Portanto, em dinâmica molecular a integração das equações de movimento é realizada em pequenos intervalos de tempo δt , cujo valor deve ser adequado para simular os movimentos de maior frequência do sistema, que muitas vezes são as vibrações das ligações.

A força total em uma dada partícula, numa determinada configuração do sistema, no tempo t , é calculada como o vetor soma de suas interações com as demais partículas. Obtida a força, determina-se a aceleração resultante. Esta é combinada com as posições e velocidades no tempo t para calcular as posições e velocidades no tempo $t + \delta t$. Assume-se que durante o intervalo de tempo δt , a força seja constante. A nova força sobre outra partícula é então determinada levando às novas posições e velocidades no tempo $t + 2\delta t$, e assim consecutivamente. Deste modo, cada partícula do sistema é movimentada, sucessivamente, gerando uma nova conformação do mesmo.

Existem inúmeros algoritmos para a integração das equações de movimento que utilizam o método de diferenciação finita (VAN GUNSTEREN e BERENDSEN, 1977; BERENDSEN e VAN GUNSTEREN, 1986). Todos assumem que propriedades dinâmicas (velocidades, acelerações, etc.) podem ser aproximadas por expansão em séries de Taylor. Os métodos mais utilizados para a integração das equações de movimento em simulações de dinâmica molecular são: métodos de Verlet (VERLET, 1967), Beeman (BEEMAN, 1976) e *Leap-*

Frog (HOCKNEY *et al.*, 1974). O método de integração utilizado neste trabalho é o *Leap-Frog*, algoritmo este considerado no programa GROMACS 3.3.3. O algoritmo *Leap-Frog* calcula as velocidades na metade do intervalo de integração e usa estas velocidades para calcular as novas posições (HOCKNEY, 1974), como pode-se vê pelas seguintes equações:

$$v(t + \frac{1}{2} \delta t) = v(t - \frac{1}{2} \delta t) + \delta t a(t) \quad (10)$$

$$r(t + \delta t) = r(t) + \delta t v(t + \frac{1}{2} \delta t) \quad (11)$$

A variação do tempo (δt) utilizado para integrar as equações acima em simulações por dinâmica molecular é, geralmente, da ordem de 1-5 femtosegundos ($1-5 \times 10^{-15}$ segundos). Essa escala de tempo é adequada para considerar as vibrações dos átomos de hidrogênio.

2.1.3 - Controle da Temperatura e da Pressão

Em simulação por dinâmica molecular no *ensemble* NpT requer o controle da temperatura e da pressão. Estes controles são obtidos utilizando algoritmos para controle de temperatura (termostato) e pressão (barostato). No presente trabalho foram utilizados os procedimentos propostos de Berendsen (BERENDSEN, 1984) que para manter a temperatura fixa. Ao longo de uma simulação as velocidades são re-escaladas a cada passo de integração, a fim de ajustar a energia cinética do sistema até a temperatura selecionada (BERENDSEN, 1984). Outros métodos mais sofisticados como os termostatos de Nosé-Hoover (NOSÉ, 1984; HOOVER, 1985) e Nosé-Poincaré (HOOVER, 1985) também são utilizados. Os barostatos mais utilizados são os de Berendsen (BERENDSEN, 1984) e Parrinello-Rahaman (PARRINELLO, 1981). Nas equações a seguir temos os fatores λ e μ utilizados para re-escalonar as

velocidades e coordenadas em cada passo da simulação pelo método de Berendsen.

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T} - 1 \right) \right]^{1/2} \quad (12)$$

$$\mu = \left[1 + \frac{\Delta t}{\tau_P} (P_0 - P) \right]^{1/2} \quad (13)$$

Onde, T_0 e P_0 são os valores de temperatura e pressão, respectivamente, desejadas durante a simulação. Os valores λ_T e μ_P são parâmetros que determinam o acoplamento do termostato e barostato, respectivamente.

2.1.4 - Condições Periódicas de Contorno

Um dos grandes problemas em simulação molecular é considerar um sistema com elevado número de partículas, uma necessidade, uma vez que se pretende obter propriedades macroscópicas ($\sim 10^{23}$ partículas). Porém, para simular tal número de partículas, seria necessário um tempo inviável de cálculo computacional. Além disso, deve-se considerar os efeitos de superfície quando o sistema é pequeno. Na figura a seguir estão ilustrados estes dois tipos de problemas.

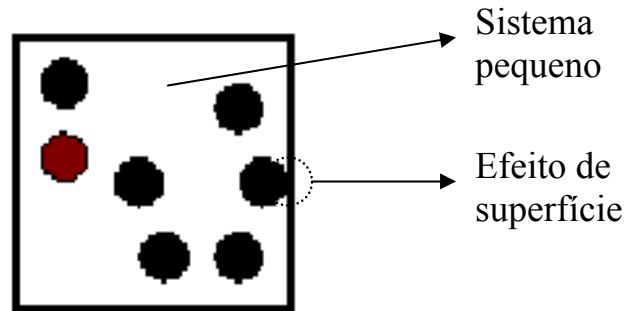


FIGURA 2.2 - Caixa simples de simulação.

Para contornar estes problemas, utiliza-se condição periódica de contorno (ALLEN, 1987). Esta técnica possibilita a realização de simulações utilizando-se um número relativamente pequeno de partículas. Efeitos de superfície são eliminados e as partículas consideradas sofrem o efeito das forças como se estivessem no interior do sistema.

Utilizando condição periódica de contorno, os átomos do sistema que está sendo simulado são colocados em uma caixa, a qual é recoberta por imagens (réplicas) da caixa original. Isto é equivalente a replicar periodicamente a caixa original em todas as direções do espaço, sem necessariamente definir as coordenadas para os átomos que estão nas caixas obtidas por replicação. No decorrer da simulação, quando um átomo move-se na caixa original, sua imagem periódica, em cada uma das caixas imagens, move-se em sincronia. Quando um átomo sai da caixa original, atravessando uma de suas seis faces, a imagem correspondente entra pela face oposta, com velocidade idêntica. Desta forma, o número total de átomos na caixa central e no sistema como um todo, é conservado. Na FIGURA 2.3 é apresentado um desenho esquemático de um sistema periódico em duas dimensões.

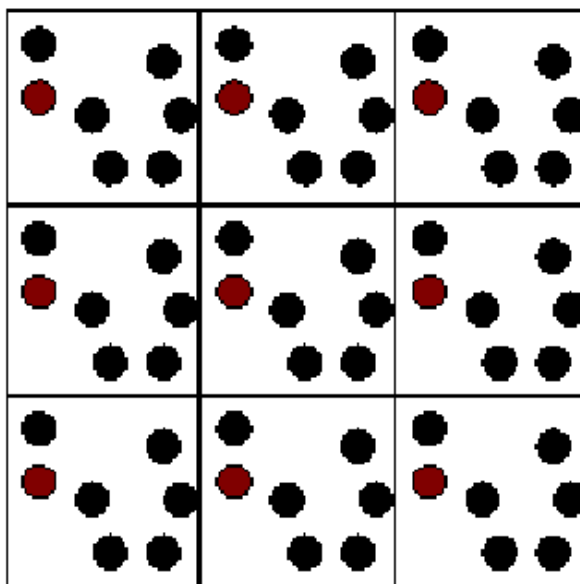


FIGURA 2.3 - Condições periódicas de contorno em um sistema bidimensional

Em uma simulação molecular, o cálculo das interações não ligantes consiste na fase mais trabalhosa da definição da superfície de potencial. Como apresentando anteriormente, as funções potenciais geralmente utilizadas para a definição das interações não ligantes são a função potencial Lennard-Jones e a função Coulomb.

A função potencial Lennard-Jones define interações de curto alcance ($V \sim r^{-6}$), pois o potencial de interação cai rapidamente com o aumento da distância entre as partículas. Para o tratamento das interações de curto alcance usualmente é empregado o truncamento do potencial em um raio de corte esférico (FIGURA 2.4) menor que $L/2$, sendo L o comprimento da célula ou caixa de simulação (ALLEN, 1987). Essa condição é imposta para não permitir que uma partícula interaja com sua própria imagem ou com a mesma molécula duas vezes.

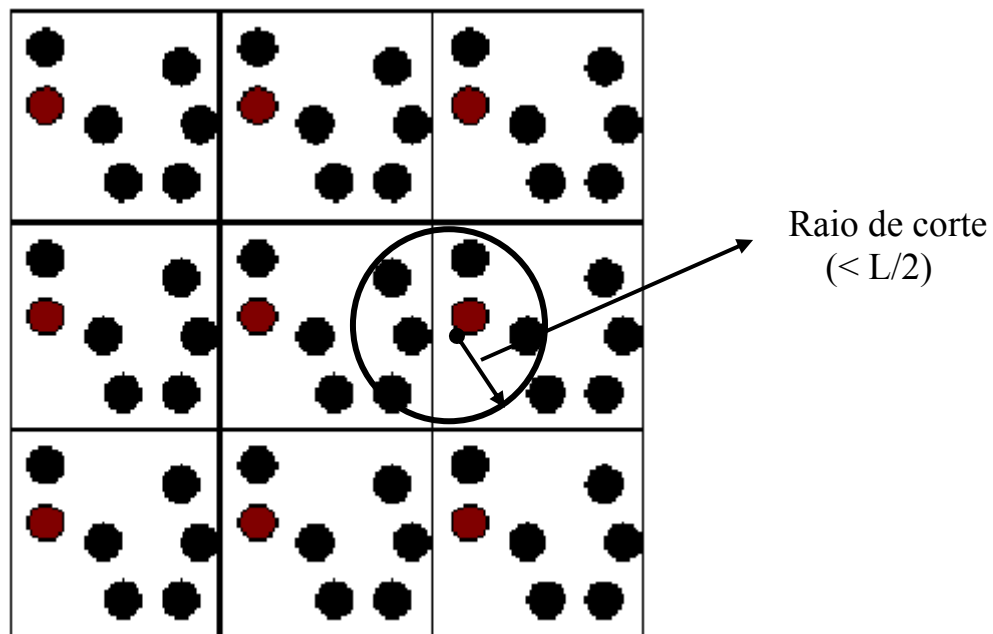


FIGURA 2.4 - Representação do raio de corte esférico para um sistema periódico bidimensional.

2.1.5 – Algoritmo de Dinâmica Molecular

Na figura a seguir temos a representação esquemática de um algoritmo de dinâmica molecular.

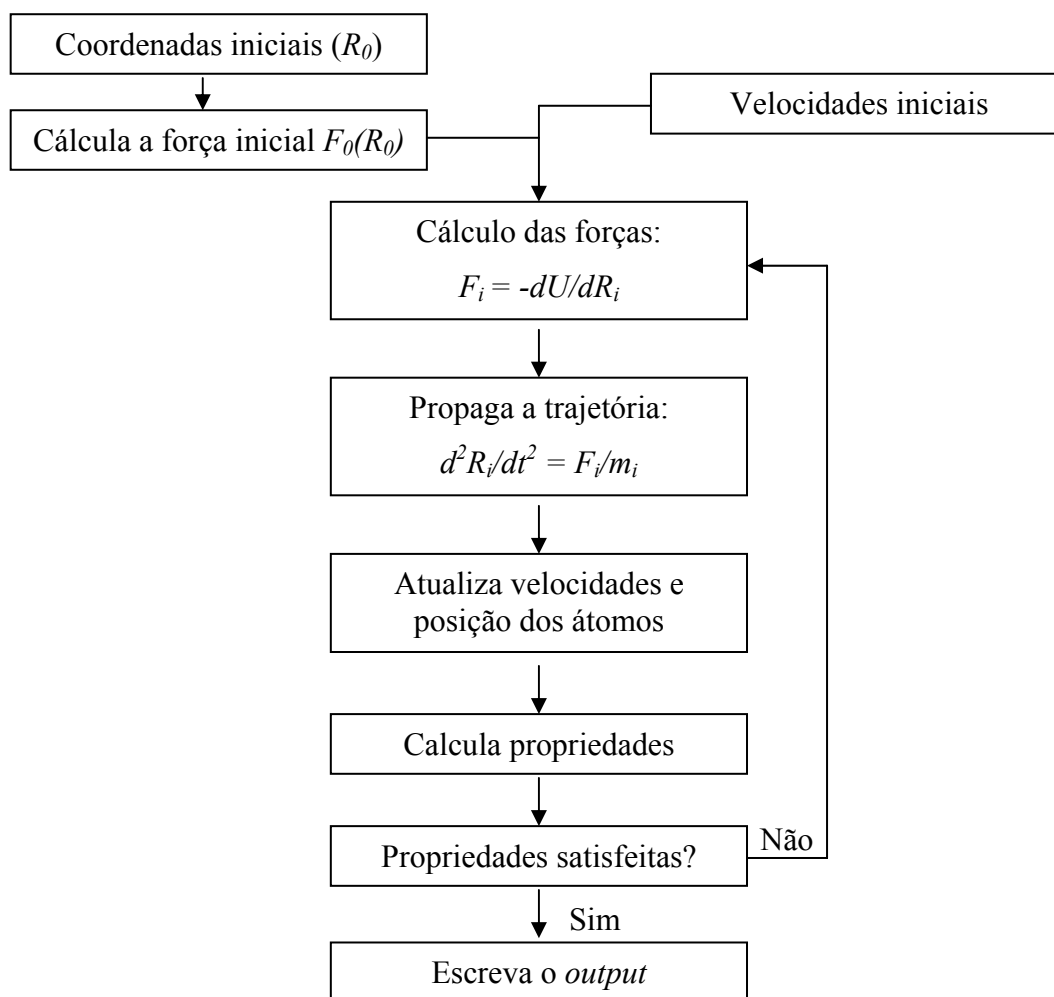


FIGURA 2.5 - Algoritmo clássico de simulação.

Todos os cálculos de dinâmica molecular apresentado neste trabalho foram realizados com o program computacional GROMACS 3.3.3 (BERENDSEN *et al.*, 1995; LINDAHL *et al.*, 2001), conhecido pela eficiência numérica e velocidade de processamento.

As condições de simulação, tais como número de partículas, modelo molecular usado, temperatura e outros parâmetros serão descritos na parte de resultados e discussão.

2.2 - Docking

A idéia central envolvida no desenho de novas estruturas está relacionada com a predição do modo de ligação e a afinidade de interação de composto no sítio de ligação de uma proteína (VERDONK *et al.*, 2003). Uma metodologia muito utilizada ao longo dos anos para estudar a interação proteína-ligante são os métodos de *docking*. Esses métodos são baseados no modelo chave-fechadura (JORGENSEN, 1991) e o docking é realizado via movimentos de rotação e translação do ligante em um sítio específico, geralmente o sítio ativo, da enzima. Na figura a seguir temos a representação esquemática de *docking*.

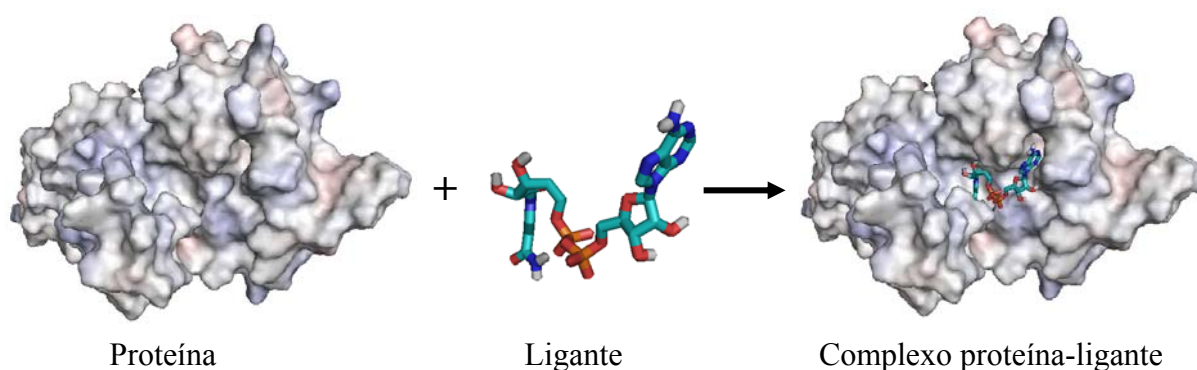


FIGURA 2.6 - Representação esquemática de *docking* de um ligante em uma proteína, formando o complexo proteína-ligante.

Em geral, nos cálculos de *docking* a energia de interação entre enzima e ligante é obtida de uma somatória do potencial de Coulomb e Lennard Jones, utilizando os parâmetros do campo de força escolhido. O objetivo dos programas de *docking* é gerar um máximo possível de complexos ligante-receptor. Em seguida, cabe à análise humana determinar quais complexos são relevantes ao estudo de interesse. Até a presente data existem mais de 60 programas de *docking*, dentre os mais populares podemos citar: AutoDock

(MORRIS *et al.*, 1998), GOLD (JONES *et al.*, 1997), FlexX (RAREY *et al.*, 1996) e Dock (KUNTZ *et al.*, 1982).

Os algoritmos de *docking* são elaborados de tal modo que esses sejam eficientes, rápidos e de fácil uso. Os algoritmos podem ser divididos em três tipos: *docking* rígido (proteína e ligante são tratados rígidos), semi-flexível (proteína é rígida e ligante é flexível) e flexível (proteína e ligante são considerados flexíveis). Os cálculos de *docking* flexíveis são mais eficientes, mas exigem um alto o custo computacional. Geralmente nesses tipos de *docking*, apenas alguns resíduos localizados no sítio ativo ou de interesse da proteína são considerados flexíveis.

Na presente tese, todos o cálculos de *docking* foram realizados no programa AutoDock4 (MORRIS *et al.*, 1998). Programa é distribuído segundo a filosofia de software livre do Projeto GNU.

O AutoDock utiliza uma função empírica para aproximar a energia livre de ligação.

$$\Delta G = \Delta G_{vdw} + \Delta G_{hbond} + \Delta G_{ele} + \Delta G_{conf} + \Delta G_{sol} \quad (14)$$

Onde ΔG_{vdw} é a contribuição de van der Waals, ΔG_{hbond} trata das contribuições energéticas das ligações de hidrogênio, ΔG_{ele} e a contribuição eletrostática, ΔG_{conf} contribuições energéticas conformacionais e ΔG_{sol} trata da solvatação do complexo.

Para a busca conformacional e encaixe do ligante no sítio ativo da enzima, o programa AutoDock utiliza o algoritmo genético. Esse algoritmo é mimetiza a teoria da evolução, na qual as mudanças adquiridas por um ser durante sua vida eram transmitidas para seus descendentes. Nos algoritmos genéticos, uma população inicial de indivíduos é requerida e, em geral, gerada aleatoriamente. Cada indivíduo é representado através de 23 cromossomos, que é a estrutura que contém toda a informação genética de um indivíduo e que são

representados computacionalmente através de uma estrutura de dados conveniente, comumente um vetor. Cada posição dessa matriz ou vetor é ocupada por uma cadeia e essas cadeias são os genes desse cromossomo. Os quais correspondem ao valores de certas variáveis ou conjunto de variáveis. No caso *docking*, as variáveis são os ângulos de translação, rotação e de torção do ligante que são reunidos no cromossomo. Na presente tese, o algoritmo genético Lamarckiano do programa AutoDock foi utilizado. O algoritmo genético usado no AutoDock define um ligante (cromossomo) como tendo 7 genes: três para as coordenadas cartesianas e quatro para variáveis especificando as orientações. Estrutura covalente é mantida intacta, entretanto genes adicionais representando a flexibilidade de certos ângulos torsionais podem ser definidos se necessário. Definidos os genes, o algoritmo genético inicia criando uma população inicial randômica de indivíduos confinados dentro de uma região específica da enzima. Para orientação dos indivíduos na enzima, movimentos translacionais (*e.g.* passos de 0,2 Å) e rotacionais (*e.g.* passos de 5°) são realizados, seguidos pelo cálculo da energia da nova conformação. Se essa nova conformação tiver menor energia de interação que a antecessor, uma pesquisa local é repetida até que a energia diminua ou até que o conjunto de passos tenha sido finalizado (MORRIS *et al.*, 1998).

Embora os cálculos de *docking* sejam relativamente eficientes, esses apresentam algumas limitações: o efeito explícito do solvente não é considerado, a flexibilidade da proteína, quando considerada, é restringida a apenas alguns resíduos e nada garante que o número de cálculos de EM seja suficiente para amostrar todas as conformações de menor energia do complexo proteína-ligante. Para contornar essas limitações, trabalhos recentes na literatura propõem uma nova metodologia corretiva. Essa nova abordagem do problema é conhecida como *ensemble docking*. Um *ensemble docking* é um conjunto de múltiplas conformações de uma proteína, as quais podem ser obtidas via experimentos (*e.g.* ressonância magnética nuclear) ou via modelagem molecular

(e.g. dinâmica molecular e Monte Carlo) (HUANG *et al.*, 2006). Desse modo, o efeito do solvente e a flexibilidade da proteína são explicitamente considerados. Já existem na literatura alguns programas computacionais que realizam esses novos tipos de *docking*, entretanto, esses programas não de livre acesso e geralmente são caros financeiramente. Dentre estes, temos o Fleksy (NABUURS *et al.*, 2007) e o FITTED (CORBEIL *et al.*, 2007).

Na presente tese, esses programas não serão utilizados devido ao seu alto custo, entretanto um *script* foi desenvolvido em combinação com o programa AutoDock para a realização dos cálculos de *docking* levando em consideração múltiplas conformações da enzima. Deste modo, o conjunto de conformações utilizado nesse trabalho foi obtido pelos cálculos de dinâmica molecular. Cálculos de *docking* foram realizados para 10 estruturas obtidas com um intervalo de 2,0 ns de simulação por dinâmica molecular. A figura a seguir mostra um esquema desse modelo *ensemble*.

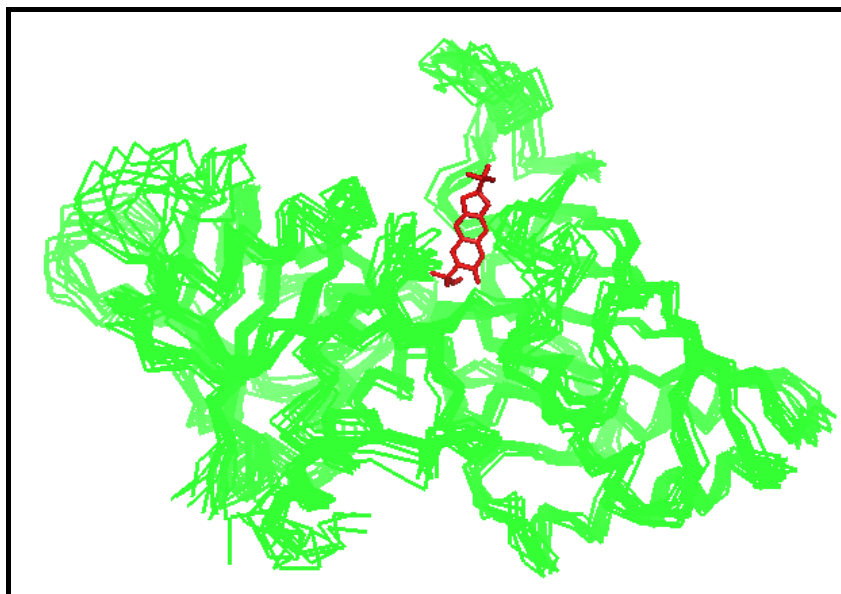


FIGURA 2.7 - Modelo de *ensemble docking* ou múltiplas conformações. Em vermelho, o ligante chalepina.

2.3 - Química Quântica

Os métodos de química quântica são desenvolvidos para resolver a equação de Schrödinger, que na sua forma independente do tempo é dada pela equação (LEVINE, 1991):

$$\hat{H}\Psi = E\Psi \quad (15)$$

Onde, a função de onda Ψ é a autofunção do operador hamiltoniano (\hat{H}) correspondente ao autovalor energia do sistema (E). O hamiltoniano \hat{H} é dado pela equação a seguir.

$$\begin{aligned} \hat{H} = & -\sum_{A=1}^N \frac{\hbar^2}{2M_A} \nabla_A^2 - \sum_{i=1}^n \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_A \sum_i \frac{e^2}{4\pi\epsilon_0 r_{Ai}} Z_A + \\ & \sum_{A<B} \frac{e^2}{4\pi\epsilon_0 r_{AB}} Z_A Z_B + \sum_{i<j} \frac{e^2}{4\pi\epsilon_0 r_{ij}} \end{aligned} \quad (16)$$

Onde M_A é a massa do núcleo A , m_e é massa e e a carga do elétron, Z é a carga nuclear, r_{ij} é a distância entre as partículas, \hbar é a constante de Planck dividido por 2π e ∇^2 é o operador laplaciano.

Na equação acima, os dois primeiros termos representam a energia cinética dos núcleos e elétrons, respectivamente. O terceiro termo é o potencial de interação elétron-núcleo; quarto termo, interação núcleo-núcleo e o quinto termo, interação elétron-elétron

A resolução exata da equação acima dá uma descrição completa de um sistema, onde todas as propriedades desse podem ser calculadas. Entretanto, a resolução da equação de Schrödinger não é trivial, e a solução analítica da mesma é obtida para o átomo de hidrogênio e cátions hidrogenóides. Para

átomos polieletrônicos e moléculas a resolução dessa equação necessita de várias aproximações. A primeira aproximação para a resolução dessa equação foi oferecida por Born e Oppenheimer. Nessa aproximação, os movimentos nucleares e eletrônicos podem ser separados, uma vez que os núcleos são pesados e se movem lentamente comparados aos elétrons. Desse modo, os núcleos são considerados fixos no espaço, e as energias cinética e de repulsão núcleo-núcleo são consideradas constantes. Resta resolver a parte eletrônica. Para isto, ao longo do tempo foram sendo propostas alternativas, em geral, que utilizavam os avanços conceituais e das inovações anteriores: método de Hartree-Fock, Hartree-Fock-Roothaan, interação de configurações, etc. Outra vertente importante foi construída ao longo da proposta inicial de Fermi para utilizar diretamente a densidade eletrônica, dando origem às metodologias do Funcional de Densidade (DFT) (FREITAS, 1999; MORGON e COUTINHO, 2007). Com esses métodos, tem sido possível ao longo dos anos obter soluções para a equação de Schrödinger, possibilitando enormes avanços no estudo de sistemas atômicos e moleculares polieletrônicos. Entretanto, devido ao alto custo computacional, os métodos apresentados acima não podem ser aplicados corriqueiramente para sistemas com centenas de átomos.

Uma solução para tratar sistemas com centenas de átomos é a utilização de métodos semi-empíricos, os quais utilizam parâmetros que podem ser obtidos experimentalmente ou a partir de cálculos *ab initio*, e restrições matemáticas mais drásticas do que aquelas utilizadas nos métodos *ab initio*. Os métodos semi-empíricos mais utilizados são baseados na aproximação ZDO do inglês “*zero differential overlap*”. Nessa aproximação, o produto de dois orbitais atômicos diferentes é considerado nulo, simplificando assim as equações de Hartree-Fock. Dentre os vários métodos pertencentes à família ZDO, podemos citar: MNDO (Dewar *et al.*, 1977), AM1 (Dewar *et al.*, 1985), PM3 (STEWART, 1989), PM6 (STEWART, 2007) e outros.

Na presente tese o método semi-empírico PM6 (*Parametric Method 6*) (STEWART, 2007), o qual é derivado do método MNDO e que por sua vez é derivado do método NDDO (*Neglect of Diatomic Differential Overlap*) (POPLE *et al.*, 1965; POPLE *et al.*, 1967) será utilizado. Na aproximação MNDO (*Modified Neglect of Diatomic Overlap*) as contribuições de elétrons do mesmo átomo (constituente de uma molécula grande) são obtidas através de valores experimentais oriundos de átomos isolados, enquanto as integrais de dois centros são consideradas como parâmetros ajustáveis do método. Estes parâmetros são refinados para reproduzir geometrias moleculares, calor de formação, pontencial de ionização, etc, obtidos experimentalmente. Em geral, um método semi-empírico produz bons valores para uma dada propriedade, mas falha em outras previsões. A grande popularidade do método MNDO se deve à sua capacidade de reproduzir algumas propriedades, tais como geometrias. Entretanto esse método falha totalmente para descrever ligações de hidrogênio, um quesito fundamental para obter propriedades de muitas moléculas, incluindo as essenciais para os sistemas vivos. Assim, o método PM6 e outros derivados do MNDO, corrigiram essa falha. As principais diferenças no método PM6 relativo ao MNDO são: inclusão de orbitais *d*; modificações das interações caroço-caroço (para reproduzir ligações de hidrogênio) e reparametrização completa para 70 átomos (STEWART, 2007). Na presente tese, o formalismo geral destes métodos não será discutido, pois estes podem ser encontrados na literatura especializada.

Embora seja possível realizar cálculos de química quântica com métodos semi-empíricos para sistemas relativamente grandes, na aplicação para estudar sistemas ainda maiores, (*e.g.* proteínas), o custo computacional é ainda excessivo. Em parte, a dificuldade reside na resolução das equações que levam à auto-consistência (SCF). A teoria do campo autoconsistente (SCF) não considera as interações instantâneas entre os pares de elétrons do sistema, e

supõe que o potencial atuando sobre cada elétron é devido aos núcleos e à distribuição média da densidade de carga dos outros elétrons.

Nos métodos semiempíricos, o problema SCF é convencionalmente apresentado como:

$$\Psi | H - E | \Psi^T = 0 \quad (16)$$

Onde Ψ é o conjunto ortogonal de orbitais moleculares, H é a matriz de Fock e E são os autovalores da matrix de Fock. Os orbitais moleculares, Ψ_i , podem ser descritos como uma combinação linear dos orbitais atômicos, ϕ_λ :

$$\Psi_i = \sum_A \sum_{\lambda \in A} c_{\lambda i} \phi_\lambda \quad (17)$$

Onde $c_{\lambda i}$ corresponde os coeficientes das combinações linearer que representarão os orbitais moleculares.

Como consequência dos orbitais moleculares (MOs) dependerem da matriz de Fock, a equações SCF devem ser resolvidas iterativamente. Na figura a seguir apresentamos um algoritmo convencional para o cálculo das equações SCF.

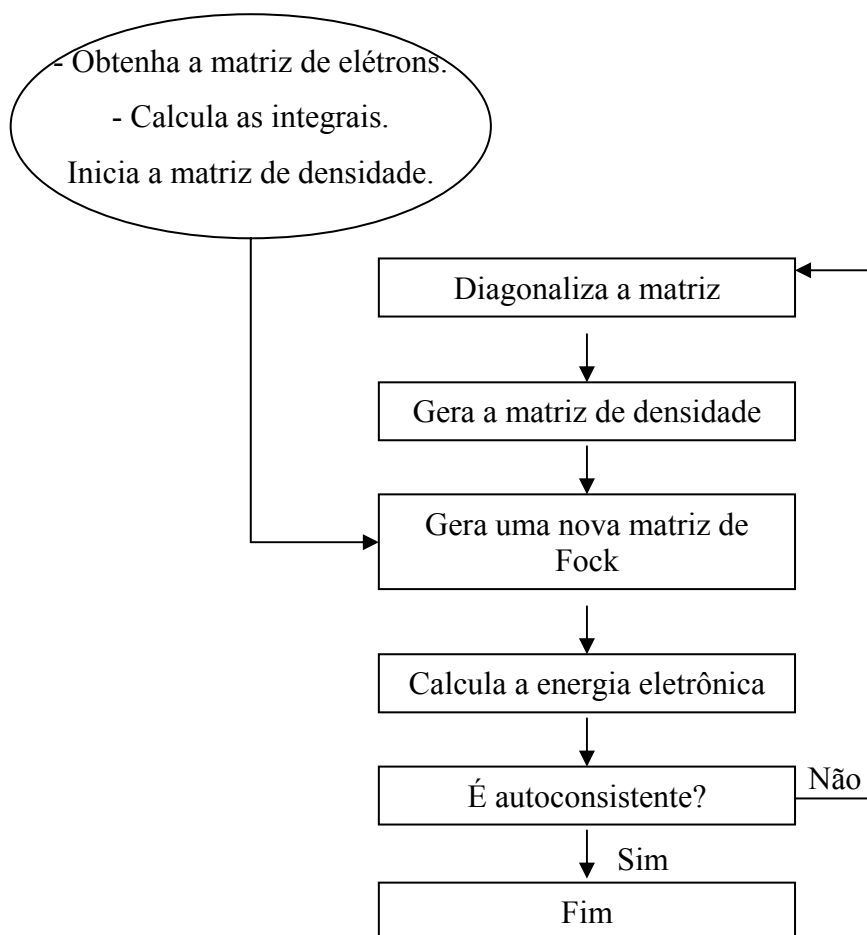


FIGURA 2.8 - Esquema do cálculo do SCF convencional (STEWART, 1996).

A matriz de densidade, P , construída a partir de orbitais moleculares é dada pela equação:

$$P_{\lambda\delta} = 2 \sum_{i=1}^{N_{occ}} c_{\lambda i} c_{\delta i} \quad (18)$$

Para um sistema com N átomos, o número de orbitais atômicos e o número de orbitais moleculares ocupados aumentam na proporção de N . Portanto, desde que o número de elementos na matriz densidade aumenta da ordem de N^2 e o número de operações requeridas no cálculo de cada elemento da matriz também aumenta da ordem de N^2 , então o número total de operações necessárias para o cálculo da matriz densidade varia com N^3 . Deste modo, o fato

do “esforço de cálculo” nos métodos semiempíricos crescer com o cubo do número de átomos envolvidos, torna impraticável tratar sistemas grande, tais como enzimas. Baseado nessa problemática, um novo procedimento para a resolução das equações SCF foi desenvolvido, a qual é chamada de MOZYME (STEWART, 1996). Nessa metodologia, uma alternativa para gerar os MOs é iniciar a partir dos orbitais moleculares localizados (LMOs) que correspondem às estruturas eletrônicas do diagrama de Lewis na estrutura molecular. O fato dos LMOs serem altamente localizados no espaço reduz os cálculos à região próxima da posição do núcleo. Em contrapartida, no método SCF convencional, os MOs do sistema são, em princípio, estendidos à molécula como um todo. Para realizar o cálculo SCF, todos os LMOs precisam ser expandidos em um somatório. Uma consequência disso, é que o espaço ocupado pelos LMOs é, em geral, muito grande. Deste modo, para sistemas grandes, tais como proteínas, o tamanho de um LMO seria pequeno comparado com o sistema total. A seguir temos um resumo do por que nessa nova metodologia é possível tratar sistemas grandes:

- 1 – Todas as interações MOs ocupado-virtual envolvendo LMOs separados por uma grande distância são automaticamente zerados ou aniquilados.

- 2 – O cálculo da matriz densidade é limitada apenas aos elementos da matriz que são representados por um LMO.

- 3 – Cálculo das energias dos LMOs e energias de interação MOs ocupado-virtual podem ser limitadas apenas aos átomos envolvidos.

Em cada caso acima, ao contrário do SCF tradicional, o esforço computacional é linear com o tamanho do sistema, o que torna possível efetuar cálculos para sistemas maiores. A seguir, apresentamos o fluxograma geral utilizado no SCF com LMOs.

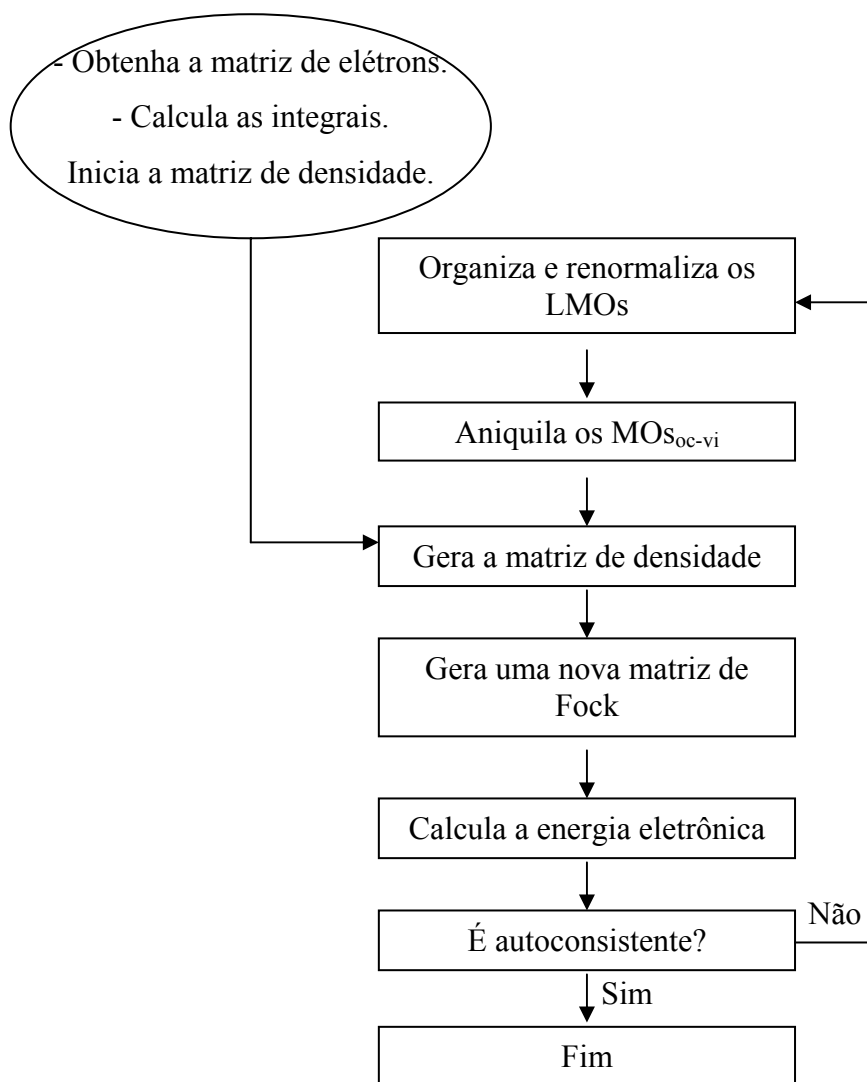


FIGURA 2.9 - Esquema do cálculo do SCF utilizando orbitais moleculares localizados (STEWART, 1996).

A principal diferença entre o cálculo SCF pelo procedimento de LMOs e o convencional, é a replicação da etapa de diagonalização no convencional por aniquilação das interações MOs ocupado-virtual no procedimento LMOs. Nesse procedimento há uma etapa adicional, organizar e renormalizar os LMOs.

Devido a capacidade dessa nova abordagem (MOZYME) de resolver o SCF para sistemas grandes, essa será utilizada juntamente com o método sem-empírico PM6 para a caracterização quântica do sítio ativo da enzima gGAPDH.

CAPÍTULO 3

RESULTADOS E DISCUSSÃO

3.1 - Etapa 1: Dinâmica Molecular

Nessa primeira etapa da presente tese, a simulação por dinâmica molecular foi efetuada para a enzima gGAPDH na forma holo e apo em solução aquosa para entender o efeito do solvente, temperatura, na conformação estrutural dessa enzima.

3.1.1 - Sistema gGAPDH-NAD⁺

Na presente simulação, o tetrâmero da gGAPDH com 4 cofatores NAD⁺ complexados em suas respectivas subunidades foram considerados. Todos os cálculos de DM foram realizados com o programa GROMACS 3.3.3 (BERENDSEN, *et al.*, 1995; LINDAHL, *et al.*, 2001). Para a elaboração do modelo a ser estudado na forma holo (enzima complexada com os cofatores), a estrutura cristalográfica do tetrâmero da enzima gGAPDH foi obtida no *Protein Data Bank* (www.rcsb.org) com código PDB:1K3T (estrutura essa sem os cofatores NAD⁺) foi sobreposto com a estrutura 1QXS (gGAPDH complexada com 4 cofatores da NAD⁺). A partir dessa sobreposição, as coordenadas cristalográficas dos cofatores NAD⁺ foram colocadas dentro da estrutura 1K3T. Esse procedimento foi necessário porque na estrutura 1QXS alguns átomos do resíduo 359 não estavam definidos na estrutura cristalográfica. O complexo 1K3T-NAD⁺ foi solvatado com 36.100 moléculas de água e foram adicionados 20 íons Cl¹⁻ para manter a neutralidade do sistema. O sistema total ou caixa de simulação ficou com 122.648 átomos e com aresta de 11,0 nm em cada direção x, y, z. Condições periódicas de contorno foram adicionadas e um raio de corte

de 1,0 nm foi utilizado para as interações de Coulomb e Van der Waals. O *ensemble* NpT (número de átomos, pressão e temperatura constantes) foi utilizado. A pressão e a temperatura do sistema adotada foi de 1 bar e 298 K, respectivamente, com constantes de tempo de 1,0 ps e 0,1 ps, respectivamente, utilizadas no algoritmo termostato e barostato de Berendsen (BERENDSEN, 1984). Para descrever o solvente, o modelo SPC/flexível (BERENDSEN *et al.*, 1981) foi utilizado; para a enzima, NAD⁺, G3P e átomos de cloro, os parâmetros do campo de força GROMOS 53A6 (OOSTENBRINK *et al.*, 2004) foram considerados.

Para minimizar o efeito de contato ou sobreposição de átomos entre as moléculas foi realizado a minimização da energia do sistema com os algoritmos *steepest descent* (ARFKEN, 1985) e *conjugate gradient* (HESTENES *et al.*, 1952), ambos implementados no GROMACS. Nessas etapas, 1000 passos de minimização foram utilizados em ambos algoritmos. Após essa etapa, uma trajetória de 300 ps foi gerada, mas restringindo os movimentos internos na enzima adotando constantes de força de 1000 kJ.mol⁻¹.nm⁻². O objetivo dessa etapa é relaxar e termalizar o solvente na temperatura desejada. Finalizada essa etapa, uma trajetória de 20,0 ns de simulação foi gerada. Na figura a seguir apresentamos uma visão da caixa inicial de simulação.

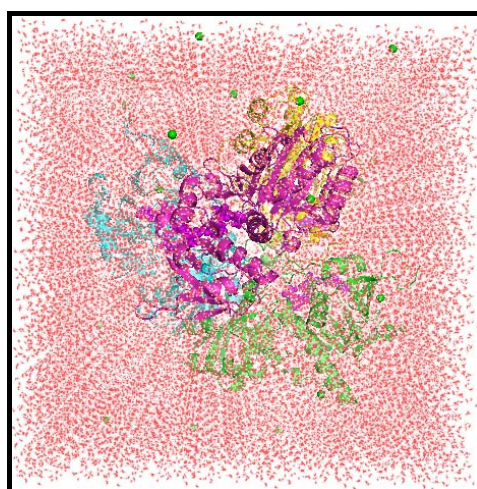


FIGURA 3.1 - Caixa inicial de simulação: 1 tetrâmero (representado em *ribbons*) + 36.100 águas (cor vermelha) + 20 Cl⁻¹ (esferas em cor verde).

3.1.1.1 - Análise estrutural

Sem dúvida, uma medida importante para analisar uma simulação por dinâmica molecular em sistemas biológicos é o cálculo do desvio médio quadrático ou RMSD (*Root Mean Square Deviation*). O RMSD fornece a medida do desvio da estrutura de uma dada molécula ao longo do tempo. Para calculá-lo, a cada intervalo de tempo executa-se uma sobreposição da estrutura atual com a inicial para obter o desvio entre ambas. Matematicamente, o RMSD para um sistema com N átomos pode ser calculado segundo a equação,

$$RMSD(t_i, t_0) = \left[\frac{1}{N} \sum_{i=1}^N \| r_i(t_i) - r_0(t_0) \|^2 \right]^{1/2} \quad (19)$$

Onde r_i e t_i é a posição do átomo i no tempo t_i e t_0 é refere-se à estrutura da molécula no tempo zero. Logo, para uma estrutura constante, o RMSD tende para zero.

Deste modo, para entender o comportamento estrutural da enzima gGAPDH na forma holo (complexada com os cofatores), medidas do RMSD foram realizadas e, para isso, todos os átomos C_α da enzima foram considerados. Na figura a seguir apresentamos valores de RMSD calculados para esse sistema.

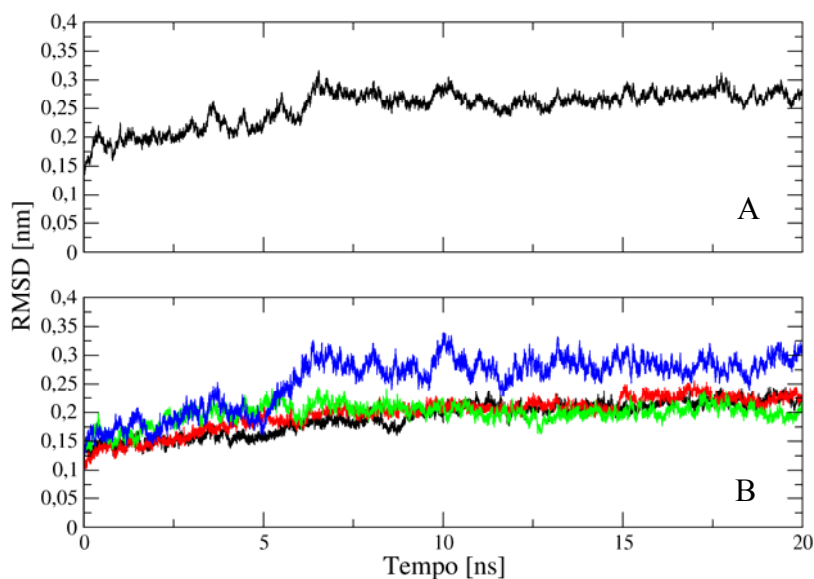


FIGURA 3.2 - RMSD dos átomos C α da enzima gGAPDH na forma holo. A) Todo os átomos C α foram considerados; B) Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).

A FIGURA 3.2A mostra que os valores de RMSD aumentam nos primeiros 6,0 ns de simulação de 0 para 0,26 nm. Isso mostra que a gGAPDH perde sua conformação inicial. Depois desse tempo, o RMSD mantém média de 0,27 nm com desvio padrão (d.p.) de 0,01 nm ao longo do tempo. De forma geral, esse valor de RMSD mostra que a enzima na forma holo como um todo não tem mudanças conformacionais drásticas nesse tempo de 20,0 ns de simulação. Para entender o comportamento conformacional das subunidades da gGAPDH, na FIGURA 3.2B temos o RMSD para as subunidades A, B, C e D. Nessa figura, podemos observar claramente que a subunidade D tem maior valor de RMSD, mostrando que essa subunidade sofre maior mudança estrutural que as subunidades A, B e C. Essas subunidades mantêm valores de RMSD de 0,20 nm (d.p. de 0,02 nm), enquanto que a subunidade D, mantém RMSD de 0,28 nm (d.p. de 0,02 nm) depois de 6,0 ns. Esse alto valor de RMSD, relativo as outras subunidades, se deve ao fator de que o NAD⁺ se deslocar do sítio ativo da enzima, principalmente a parte adenina do cofator. Esse deslocamento do NAD⁺

induz mudanças conformacionais nos *loops* Lys89-Trp99 e His194-Trp211, gerando-se assim, um maior valor de RMSD. É interessante observar que a curva de RMSD da subunidade D é similar com o da enzima total. Desse modo, o comportamento do RMSD da enzima como um todo é influenciado pelas mudanças ocorridas na subunidade D.

Na figura a seguir, temos a sobreposição da estrutura obtida a 20,0 ns de simulação com a estrutura inicial (no tempo zero de simulação).

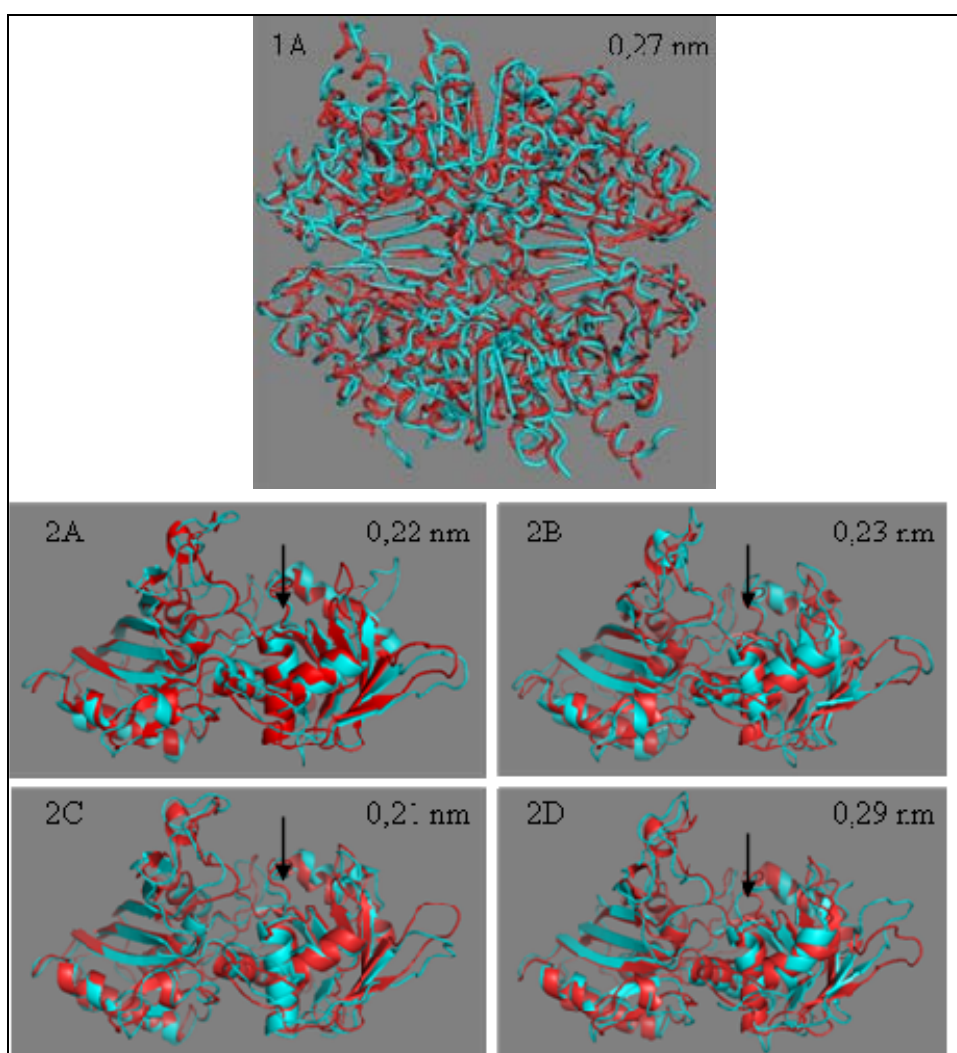


FIGURA 3.3 - 1A) Sobreposição da gGAPDH. 2A, 2B, 2C e 2D) Sobreposição das subunidades A, B, C e D, respectivamente. Cada seta indica a posição do NAD^+ no sítio ativo. Na parte superior direita estão os valores de RMSD das respectivas subunidades. As estruturas em vermelho e ciano representam, respectivamente, a enzima inicial (0 ns) e final (20,0 ns) da simulação.

Para a melhor visualização do NAD^+ no sítio ativo, na figura a seguir apresentamos a sobreposição das subunidades a 20,0 ns com as correspondentes em 0,0 ns de simulação.

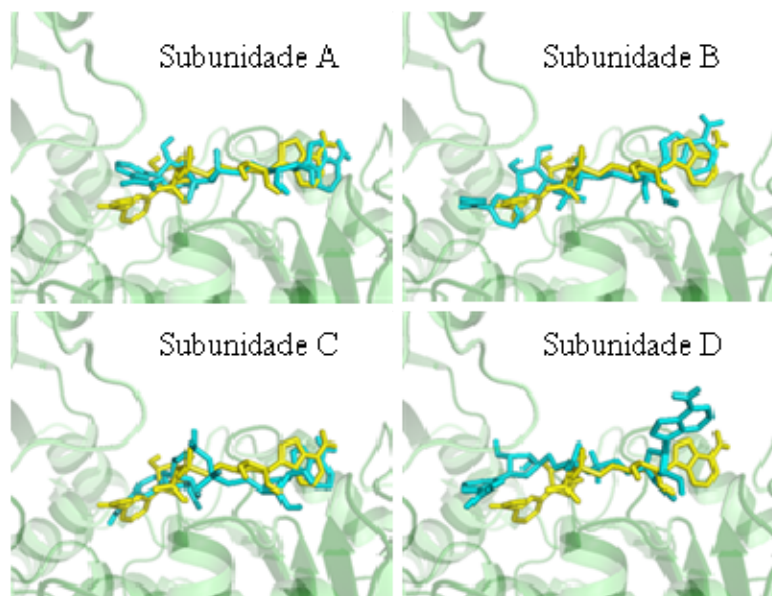


FIGURA 3.4 - Sobreposição das subunidades A, B, C e D a 20,0 ns com suas respectivas a 0,0 ns de simulação. NAD^+ a 20,0 ns em cor ciano e a 0,0 ns em cor amarelo. Para melhor visualização, apenas a enzima a 0,0 ns (cor verde) foi considerada.

Cada subunidade da enzima gGAPDH é composta por dois domínios: domínio de ligação do NAD^+ (resíduos 1-150 e 331-359) e de um domínio catalítico (resíduos 151-330). Para entender o comportamento estrutural desses domínios, na figura a seguir mostramos as curvas de RMSD para cada subunidade da enzima gGAPDH.

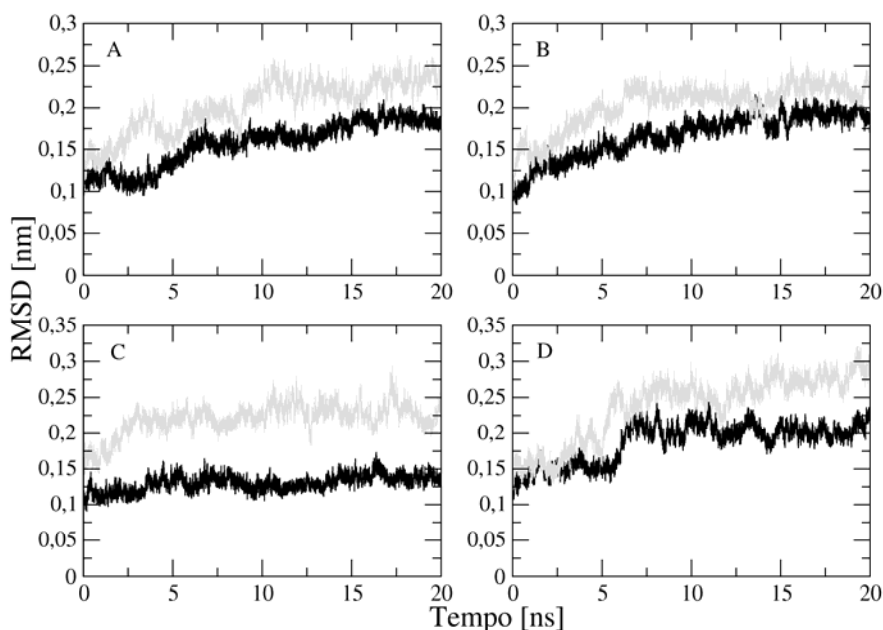


FIGURA 3.5 - RMSD dos átomos C α para o domínio de ligação do NAD⁺ (resíduos 1-150 e 331-359, linha cinza) e para o domínio catalítico (resíduos 151-330, linha preta) das subunidades A, B, C e D da enzima gGADPH.

Na figura acima, pode-se observar que o domínio de ligação do NAD⁺ apresenta valores de RMSD maiores que a do domínio catalítico. Isso mostra que o domínio de ligação do NAD⁺ é mais flexível do que o do catalítico. Esse resultado se deve ao fato de que o domínio de ligação do NAD⁺ ser mais exposto ao solvente do que o do catalítico e bem como, a presença do N- e C-terminal neste domínio, que são flexíveis. Por outro lado, a maioria dos resíduos que são compostos pelo domínio catalítico estão localizado na interface das subunidades, consequentemente, esses possuem pouca flexibilidade devido a interação dessas com os resíduos das outras subunidades.

3.1.1.2 - Interação gGAPDH-NAD⁺

Para entender quais resíduos da gGAPDH preferencialmente com o cofator NAD⁺, na figura a mostramos as ligações de hidrogênio entre o cofator e resíduos da enzima ao longo do tempo de simulação. Ligações de hidrogênio foram definidas pela distância máxima entre doador e receptor de 0,35 nm.

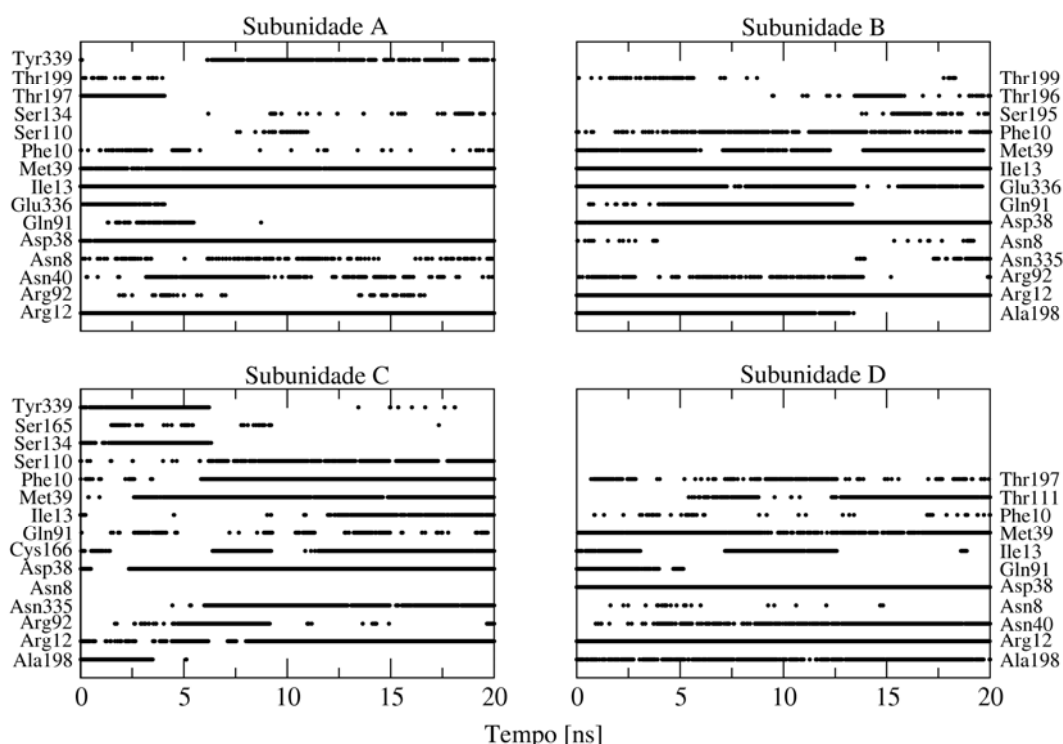


FIGURA 3.6 - Ligações de hidrogênio entre resíduos da enzima e o cofator NAD⁺ ao longo do tempo de simulação.

Observa-se que a subunidade D possui menos aminoácidos interagindo com o cofator. Isso ocorre, como explicado na seção anterior, por que a parte do grupo adenina e nicotinamida do cofator se deslocam do sítio catalítico. Baseado nesse resultado, a partir desse momento a subunidade D não será considerada em futuras análises. Assim, analisando apenas as subunidades A, B e C, conclui-se que os resíduos Arg12, Arg92, Asn8, Asp38, Gln91, Ile13, Met39, Phe10 possuem em comum interação com o cofator. Deste modo, esses

resíduos mostram-se importantes para a estabilização do cofator no sítio ativo da gGAPDH. Destes aminoácidos, Arg12, Asp38, Met39 e Ile13 são os que mais interagem via ligação de hidrogênio com o cofator e essas interações ocorrem com diferentes partes da estrutura deste. Em detalhes, Arg12 interage com o grupo nicotinamida; Ile13 interage com o grupo fosfato; Met39 e Asp38 interagem com o anel ribose. É interessante ressaltar que o resíduo Asp38 é altamente conservado na GAPDH de tripanosomatídeos (SURESH *et al.*, 2001). Essas observações sugerem que se essas ligações de hidrogênio sofrerem interferências, a interação entre o cofator e enzima será comprometida e, conseqüentemente, a atividade enzimática poderá ser diminuída. Embora esses aminoácidos não participem diretamente na reação enzimática, eles são importantes para a estabilização e acomodação do cofator no sítio catalítico. Essas suposições são muito importantes para a elaboração de inibidores, uma vez que, compostos podem ser elaborados para interferir nessas ligações de hidrogênio.

Análises do sítio ativo de ligação do cofator na gGAPDH do *T. cruzi* e o conhecimento do mecanismo bioquímico evidenciaram a importância do resíduo Arg249 para a atividade catalítica da enzima. Localizado na porção central do sítio de ligação do cofator, o grupo guanidina da Arg249, que em pH fisiológico encontra-se na forma protonada, poderia ser explorado como grupo essencial para a interação com moléculas de inibidores que apresentassem grupos aniônicos. Essa suposição se baseia no fato de que na forma apo, o resíduo Arg249 forma *salt bridge* com o Asp210 (SOUZA *et al.*, 1998) enquanto que na forma holo (complexada com um análogo do tioacetil), o Arg249 aponta para os resíduos Thr197 e Thr199 (CASTILHO *et al.*, 2003). A presente simulação confirma esses dados, embora o resíduo Arg249 não aponte diretamente para os resíduos Thr, mas o Arg249 se desloca na direção do cofator NAD^+ , perdendo assim sua interação com o Asp210. Neste estágio da fase da

simulação do complexo gGAPDH-NAD⁺ em solução aquosa, a importância desse deslocamento da Arg249 não foi evidenciada.

A literatura cita os resíduos His194 e Cys166 como os responsáveis pela atividade catalítica da gGAPDH (SOUZA *et al.*, 1998). Assim, na figura a seguir mostramos a distância mínima de contato entre Cys166-His194 e Cys166-NAD⁺. Por definição, distância mínima significa que pelos um dos átomos de um grupo (*e.g.* Cys166) está em contato com um segundo grupo de átomos (*e.g.* NAD⁺).

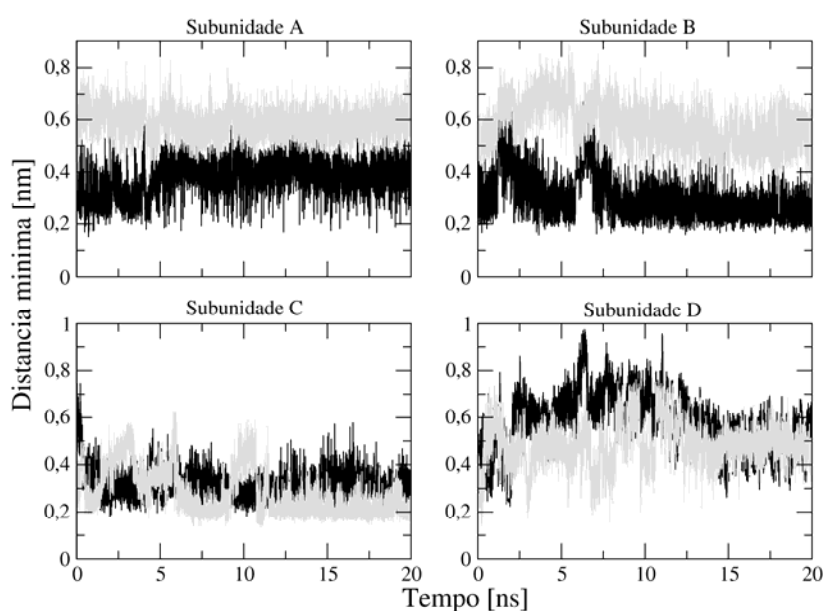


FIGURA 3.7 - Distância mínima de contato entre Cys166-NAD⁺ (linha cinza) e Cys166-His194 (linha preta).

Pela figura acima, observa-se que as distâncias, tanto para Cys166-NAD⁺ e Cys166-His194 são bem comportadas ao longo do tempo de simulação para as subunidades A e B. Nessas subunidades, a distância mínima de contato entre Cys166-NAD⁺ é maior que a entre Cys166-His194. Para a subunidade A, a distância mínima de contato é de 0,38 nm (d.p. de 0,07 nm) e 0,56 nm (d.p. de 0,06 nm) para Cys166-His194 e Cys166-NAD⁺, respectivamente. Já para a

subunidade B, a distância é de 0,30 nm (d.p. de 0,09 nm) e de 0,56 nm (d.p. de 0,09 nm). Em contraste, para as subunidades C e D, as distâncias Cys166-His194 são maiores que Cys166-NAD⁺. Para a subunidade C, a distância mínima de contato Cys166-His194 é de 0,31 nm com d.p. de 0,06 nm e Cys166-NAD⁺ de 0,28 nm com d.p. de 0,09 nm. Para a subunidade D, a distância Cys166-His194 foi de 0,58 nm (d.p. de 0,1 nm) (dp=desvio padrão) e para Cys166-NAD⁺ foi de 0,48 nm (d.p. de 0,09 nm). Analisando os resultados acima, podemos concluir que para todas as subunidades, exceto para a subunidade D, a distância Cys166-His194 está em concordância com dados experimentais. Além disso essa distância mínima de contato seria o suficiente para a transferência do próton da Cys166 para a His194, de acordo com o mecanismo aceito na literatura. A pequena distância entre Cys166 e NAD⁺ na subunidade C já era de se esperar uma vez que, a Cys166 forma ligação de hidrogênio com o NAD⁺ (FIGURA 3.1.1.2.1). Na Figura a seguir, mostramos a orientação dos resíduos Cys166 e His194 em torno do cofator NAD⁺.

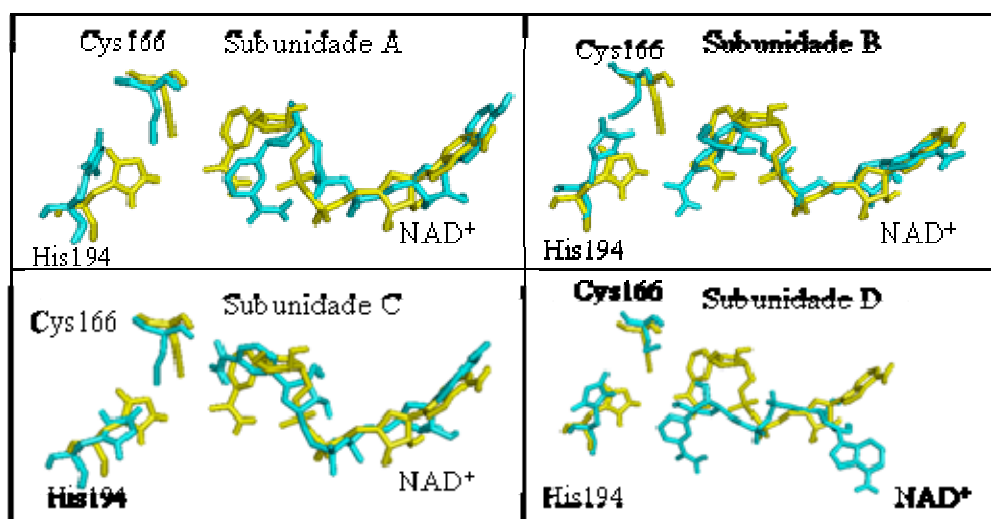


FIGURA 3.8 - Sobreposição das subunidades A, B, C e D a 20,0 ns (cor ciano) com suas respectivas subunidades a 0,0 ns (cor amarelo) de simulação.

Pela figura acima observa-se que, exceto para a subunidade C, o anel da histidina se desloca de um ângulo de aproximadamente 90° . Esse deslocamento deve favorecer a acomodação do substrato G3P no sítio ativo e bem como o melhor alinhamento da His194 na direção da cisteína. Além disso, essa reorientação favorece a um aumento do volume molecular no sítio ativo e isso é muito importante para a entrada do gliceraldeído-3-fosfato, que é essencial a atividade catalítica da gGAPDH, como discutido na introdução.

Para entender o quanto as subunidades interagem, na figura a seguir mostramos a energia de interação (eletrostática + van der Waals) entre as subunidades obtidas nas simulações por DM.

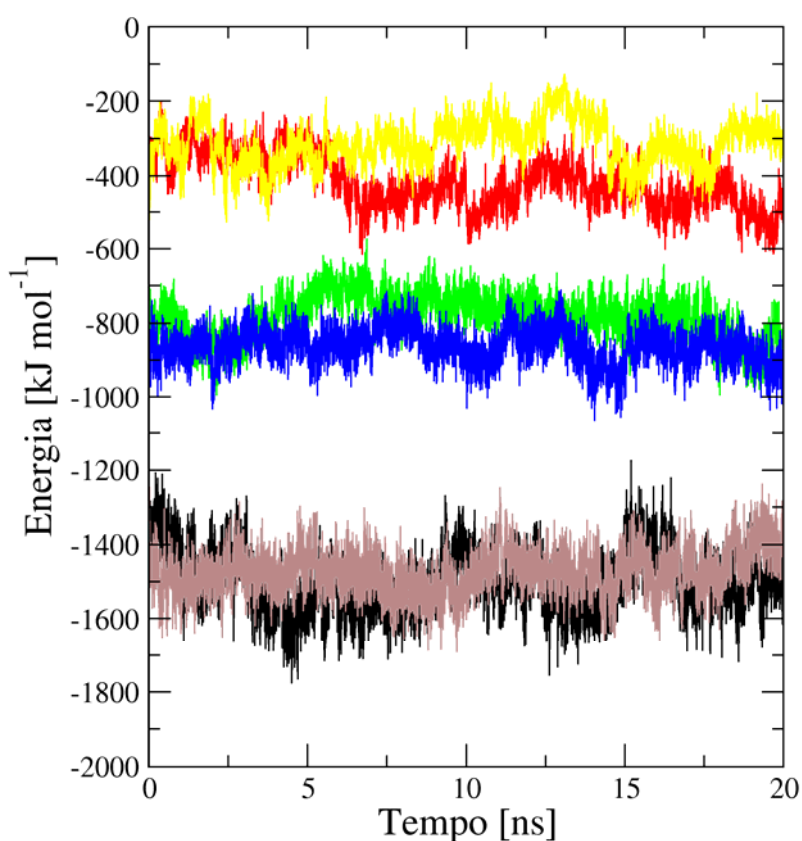


FIGURA 3.9 - Energia de interação entre as subunidades. A-B (preto), C-D (marron), A-D (verde), B-C (azul), A-C (vermelho) e B-D (amarelo).

A grande interação entre as subunidades A-B ($-1.498,6 \text{ kJ.mol}^{-1}$ e d.p. de $83,0 \text{ kJ.mol}^{-1}$) e C-D ($-1.473,5 \text{ kJ.mol}^{-1}$ e d.p. de $59,9 \text{ kJ.mol}^{-1}$) deve-se a à interação entre os respectivos domínios catalíticos, os quais são compostos pelos resíduos 151-330. Dentre estes resíduos, 48 (de cada subunidade) estão em contato em um raio de corte de $0,4 \text{ nm}$ na interface dessas subunidades ao longo de $20,0 \text{ ns}$ de simulação. Já para a interação das subunidades A-D ($-783,9 \text{ kJ.mol}^{-1}$ e d.p. de $59,1 \text{ kJ.mol}^{-1}$) e B-C ($-863,2 \text{ kJ.mol}^{-1}$ e d.p. de $59,9 \text{ kJ.mol}^{-1}$) a interação ocorre devido o contato entre parte dos resíduos do domínio catalítico e de ligação a NAD^+ , no raio de corte de $0,4 \text{ nm}$, tem-se 32 resíduos de cada subunidade interagindo. Finalmente, a baixa interação das subunidades A-C ($-416,4 \text{ kJ.mol}^{-1}$ e d.p. de $71,7 \text{ kJ.mol}^{-1}$) e B-D ($-312,6 \text{ kJ.mol}^{-1}$ e d.p. de $61,3 \text{ kJ.mol}^{-1}$) se deve ao pequeno número de resíduos (apenas 13 resíduos de cada subunidade interagindo uma com a outra no raio de corte de $0,4 \text{ nm}$) na interface dessas subunidades. Esses resíduos são compostos pelos domínios catalítico e de ligação a NAD^+ . As energias médias das interações acima foram calculadas ao longo do tempo de $20,0 \text{ ns}$ de simulação.

3.1.2 - Sistema gGAPDH

Para melhor entender o comportamento estrutural da enzima gGAPDH na forma apo (sem os cofatores) em solução aquosa, $20,0 \text{ ns}$ de simulação por dinâmica molecular foram realizados. As condições e protocolos de simulação utilizados aqui foram os mesmos para a forma holo, exceto pela remoção dos cofatores NAD^+ e bem como a adição de quatro íons Cl^- para manter a neutralidade do sistema.

3.1.2.1 - Análise estrutural

À semelhança do sistema holo, na figura a seguir apresentamos valores de RMSD calculados para a enzima na forma apo.

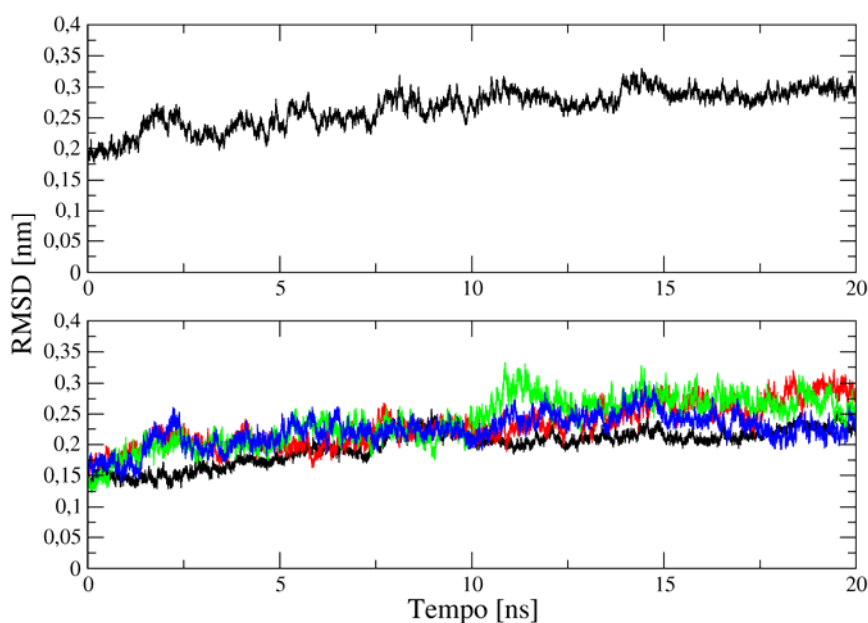


FIGURA 3.10 - RMSD dos átomos C α da enzima gGAPDH na forma apo. A) Todo os átomos C α foram considerados; B) Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).

De modo geral, os valores de RMSD acima estão abaixo de 0,30 nm e isso mostra que assim como para a forma holo da enzima, a estrutura do tetrâmero na forma apo e suas subunidades não apresentam mudanças conformacionais drásticas no tempo de simulação considerado.

Para entender as diferenças entre as forma holo e apo, na figura a seguir comparamos os valores de RMSD para a forma apo e holo, e bem como para suas respectivas subunidades.

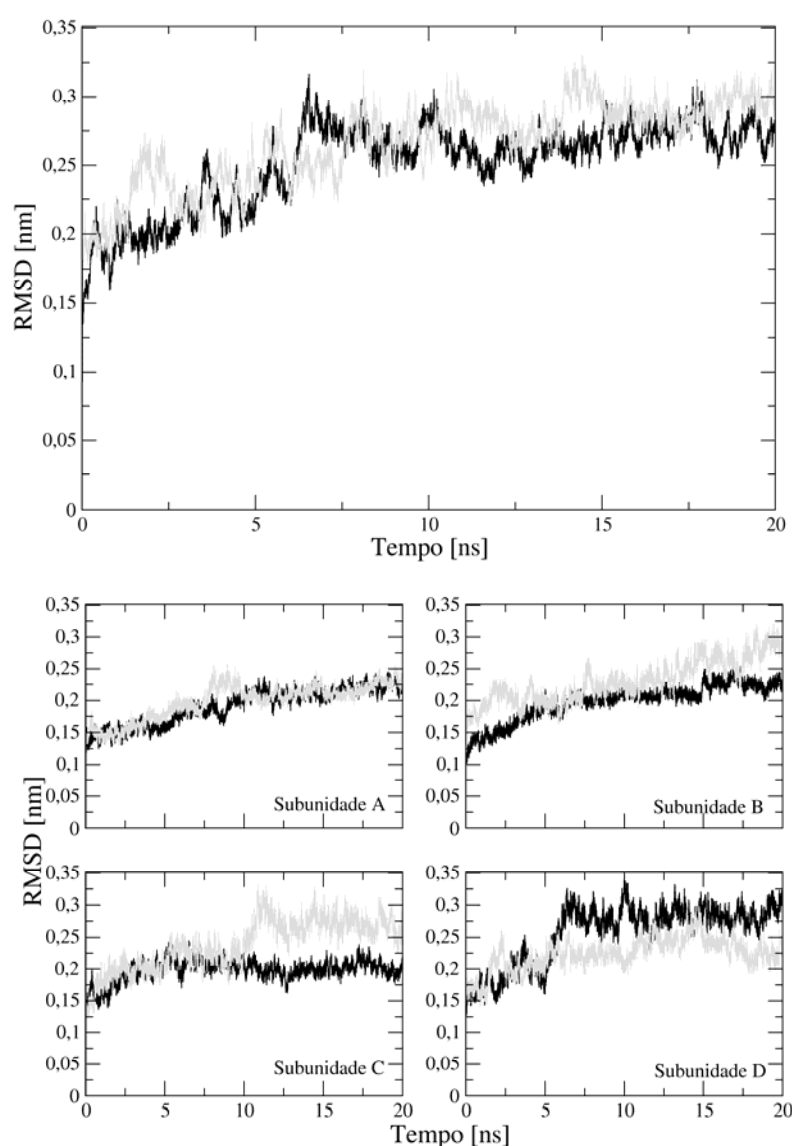


FIGURA 3.11 - RMSD dos átomos $C\alpha$ da enzima gGAPDH na forma holo e apo. Topo: todos os átomos $C\alpha$ da enzima foram considerados; Abaixo: subunidades A, B, C e D. gGAPDH na forma holo em cor preta e na forma apo em cor cinza.

Pela FIGURA 3.11 (no topo), todos os átomos $C\alpha$ da gGAPDH na forma holo e apo foram considerados, observa-se que na forma apo a enzima é ligeiramente mais flexível que na forma holo. O mesmo é observado quando se compara as subunidades B e C da forma apo (FIGURA 3.11, para as subunidades), ambas apresentam maior flexibilidade do que suas respectivas subunidades na forma holo. Contrário, na subunidade D, a forma holo apresenta

maior mudança conformacional que a respectiva subunidade na forma apo. Como dito anteriormente, essa mudança na subunidade D na forma holo se deve ao deslocamento da NAD^+ do sítio ativo da enzima, gerando assim uma mudança conformacional nos *loops* Lys89-Trp99 e His194-Trp211. É interessante observar que para a subunidade A não há diferenças significativas nas curvas de RMSD para as formas apo e holo.

Para entender o comportamento estrutural dos domínios catalítico e de ligação a NAD^+ , na figura a seguir mostramos os valores de RMSD para esses domínios. Para melhor comparação, os valores de RMSD da forma holo também estão presentes nessa figura.

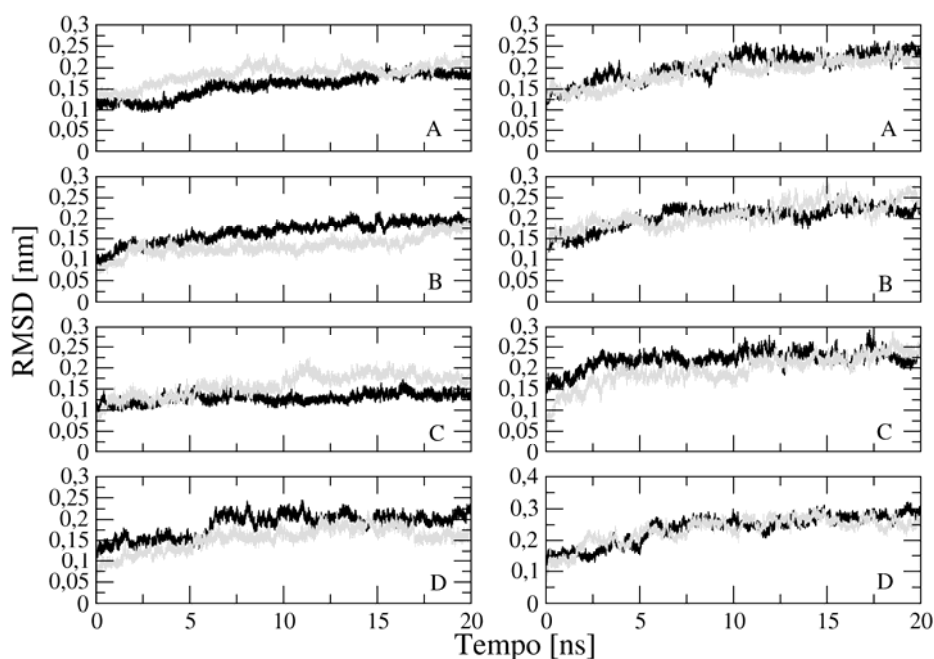


FIGURA 3.12 - RMSD dos átomos $\text{C}\alpha$ do domínio catalítico (a esquerda, resíduos 151-330) do domínio de ligação do NAD^+ (a direita, resíduos 1-150 e 331-359) e das subunidades A, B, C e D da enzima gGADPH referente a estrutura inicial. Em cor preta, forma holo e em cinza, forma apo.

Assim como na forma holo, para a forma apo o domínio catalítico apresenta menos flutuações (menor valor de RMSD) na sua estrutura molecular do que o domínio de ligação a NAD^+ . A razão dessas flutuações são as mesmas

já discutidas para a forma holo anteriormente. De modo geral, para o domínio de ligação a NAD^+ , não se observa diferenças significativas nos valores do RMSD entre a forma holo e apo da gGAPDH. Isso mostra que estruturalmente essa enzima é “indiferente” a presença do cofator no seu sítio ativo. Já para o domínio catalítico, observa-se algumas diferenças nos valores de RMSD, entretanto estas são pequenas.

Os resultados de RMSD apresentados nas figuras anteriores estão resumidos na TABELA 3.1. Para o cálculo dos valores de RMSD, todos os átomos da enzima gGAPDH foram considerados.

TABELA 3.1 - Comparação entre as diferentes subunidades e formas de gGAPDH, as estruturas obtidas a 20,0 ns de simulação foram consideradas (valores de RMSD estão em nm).

Forma	A/B	A/C	A/D	B/C	B/D	C/D
apo/apo	0,30	0,26	0,26	0,18	0,18	0,22
holo/holo	0,22	0,22	0,30	0,26	0,26	0,27
	A/A	B/B	C/C	D/D		
apo/holo	0,22	0,25	0,23	0,22		

Pela tabela acima, podemos observar que dois dos maiores valores de RMSD na comparação holo/holo são atribuídos a A/D e B/C. Nessas subunidades, os cofatores NAD^+ estão “de frente” um para o outro no sítio catalítico da gGAPDH.

Assim como para a forma holo, as energias de interação entre as subunidades foram calculadas para a forma apo, as quais são apresentadas na figura a seguir.

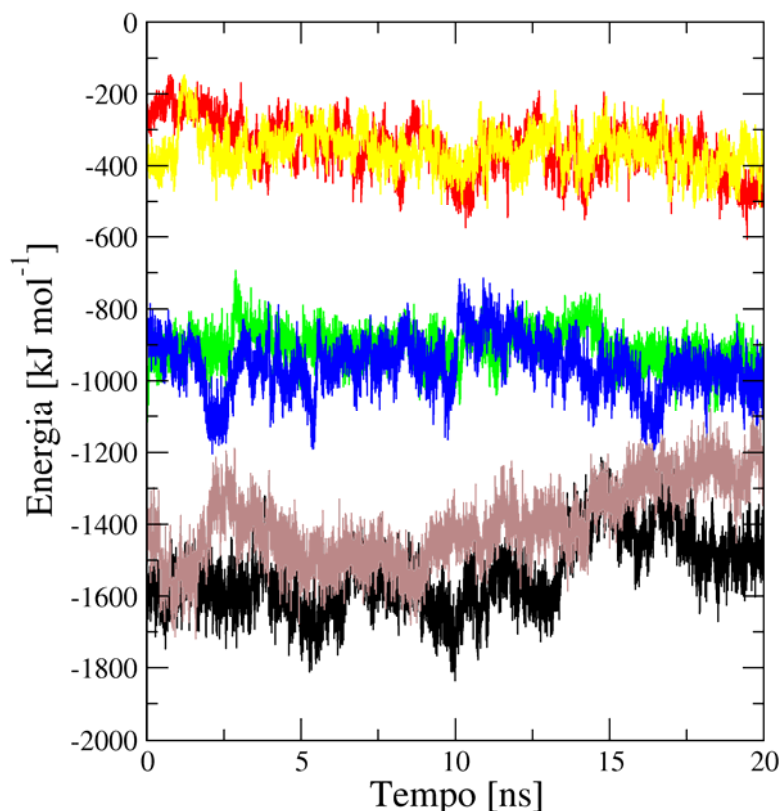


FIGURA 3.13 - Energia de interação entre as subunidades da gGAPDH na forma apo. A-B (preto), C-D (marron), A-D (verde), B-C (azul), A-C (vermelho) e B-D (amarelo).

Na acima é interessante observar que a energia de interação das subunidades A-D, B-C, A-C e B-D tende a diminuir, enquanto que para as subunidades A-B e C-D tende a aumentar ao longo do tempo de simulação. É bom salientar que a interação nas subunidades A-B e C-D se deve ao contato apenas dos seus respectivos domínios catalíticos, enquanto que nas outras subunidades ocorre a interação de ambos domínios. Assim, os dados apresentados na FIGURA 3.13 sugere que com o desfavorecimento da energia de interação entre os domínios catalíticos, há um favorecimento da interação entre os domínios de ligação do NAD^+ .

Para melhor análise e comparação energética entre a forma holo e apo, na tabela a seguir apresentamos o sumário das médias dessas energias calculadas nos últimos 10,0 ns de simulação.

TABELA 3.2 - Energia de interação (em $\text{kJ}\cdot\text{mol}^{-1}$) entre as subunidades. Em parênteses, o desvio padrão.

Forma	A-B	C-D	A-D	B-C	A-C	B-D
Apo	-1.481,1 (101,0)	-1.333,4 (89,5)	-912,3 (53,8)	-953,2 (79,5)	-382,0 (64,3)	-366,8 (52,0)
Holo	-1.494,4 (75,7)	-1.457,6 (60,7)	-794,2 (54,5)	-873,9 (72,9)	-450,7 (56,5)	-296,0 (63,4)

Comparando os dados da tabela acima com os valores de RMSD (TABELA 3.1), pode-se observar que o valor da energia de interação entre as subunidades da mesma forma, exceto para C-D, estão relacionados com os valores de RMSD. Em outras palavras, quanto menor o valor de RMSD, maior a energia de interação entre as subunidades. Isso mostra que largas flutuações nas estruturas das subunidades tendem a minimizar as interações entre as respectivas subunidades. Para melhor clareza, podemos citar como exemplo, a interação da subunidade A-B, na forma holo: a energia de interação é mais favorável e possui menor valor de RMSD do que a forma apo.

Para entender a mudança conformacional na gGAPDH induzida pela presença do cofator no sítio ativo, na TABELA 3.3 apresentamos os valores médios do ângulo de abertura dos domínios catalítico e de ligação. O ângulo formado pelos átomos C_{α} dos resíduos Arg92, Cys166 e Val208 foram utilizados para o cálculo do ângulo de abertura. Os últimos 10,0 ns de simulação foram utilizados para o cálculo médio do ângulo de abertura. Para melhor visualização desse ângulo, na figura a seguir mostramos a estrutura da subunidade A obtida a 20,0 ns de simulação da enzima na forma holo.

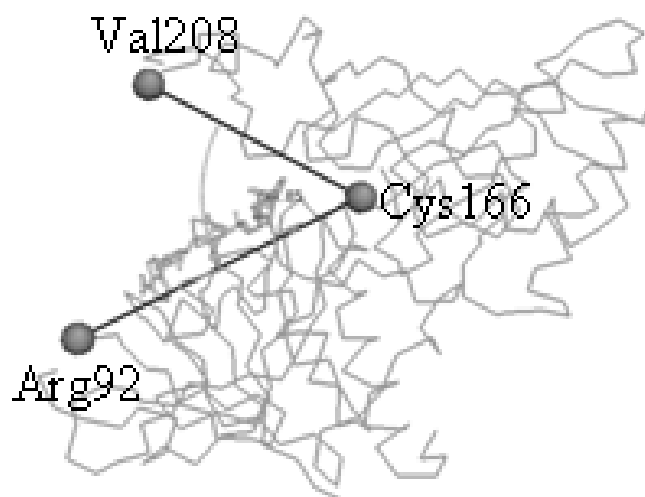


FIGURA 3.14 - Representação do ângulo de abertura. Os átomos C α são representados pelas esferas.

TABELA 3.3 - Valores médios (em graus) do ângulo de abertura formados entre os domínios catalítico e de ligação. Em parênteses, o desvio padrão.

Subunidade	apo	holo	1K3T* apo	1QXS* holo
A	55,6 (3,0)	59,4 (3,0)	56,3	54,1
B	57,6 (7,1)	52,8 (3,4)	58,0	65,3
C	62,9 (3,3)	62,8 (4,0)	55,2	53,5
D	56,2 (3,7)	62,2 (3,5)	57,8	53,6

* Enzima com código PDB obtido do *Protein Data Bank*.

Comparando os valores dos ângulos de abertura da forma holo com a cristalográfica na forma apo complexada com a chalepina (1K3T), pode-se observar que, exceto para a subunidade B, o ângulo aumenta em 6%, 14% e 8% para as subunidades A, C e D, respectivamente. Em contraste, para a subunidade B, o ângulo diminui em 9%. É bom salientar que, como descrito na metodologia, a 1K3T foi a estrutura inicial utilizada nas simulações de dinâmica molecular. Comparando os ângulos de abertura da forma holo com a 1QXS (estrutura na forma holo e com um substrato), a percentagem do aumento desse ângulo é ainda maior. Para a subunidade C, as medidas dos valores do ângulo de abertura,

obtidos pela MD, são similares para a forma holo e apo. Esse resultado mostra que este ângulo não é afetado pela presença do cofator NAD^+ . Já para as subunidades A e D, a transição de apo para holo leva a um aumento do ângulo de abertura de 7% e 11% nas subunidades A e D, respectivamente. Para a subunidade B, o contrário ocorre, nessa transição o ângulo de abertura diminui em 8%. Pelas análises acima, foi possível correlacionar os valores médio dos ângulos de abertura na forma holo com os obtidas a partir dos dados cristalográficos, entretanto não foi possível correlacionar a forma apo com a holo obtidas por simulação. Deste modo, os dados obtidos por dinâmica molecular mostram que na transição apo para holo, pode ocorrer o fechamento, abertura e conservação do sítio ativo da enzima gGAPH. Indo além, os resultados apresentados acima mostram que a posição e arranjo estrutural do cofator NAD^+ no sítio ativo é indiferente a esse ângulo de abertura. Uma vez que o NAD^+ apresenta conformação e posição similar no sítio ativo ao longo da simulação em todas as subunidades, exceto para a subunidade D (FIGURA 3.8).

3.2 - Etapa 1: *Docking*

Nessa etapa, cálculos de *docking* foram utilizados para buscar moléculas candidatas a fármacos e bem como para validar nosso modelo de *ensemble docking* ora proposto para a gGAPDH do *T. cruzi*.

Todos os cálculos de *docking* foram realizados no programa AutoDock4 (MORRIS *et al.*, 1998). O programa ADT (SANNER, 2005) foi utilizado para preparar o arquivos de entrada para os cálculos de *docking*.

Para os cálculos de *docking*, as moléculas de águas e contra-íons das estruturas obtidas pela dinâmica molecular foram removidas. Para os ligantes em estudo e cofator NAD⁺, as cargas dos átomos foram calculadas pelo método de Gasteiger (GASTEIGER *et al.*, 1980) através do programa ADT. Para a enzima, parâmetros do compo de força AMBER (CORNELL *et al.*, 1995) foram utilizados.

Nos cálculos de *docking*, as energias foram calculadas em uma série de grades centradas nos sítio ativo da enzima. Para a realização da busca conformacional do ligante, o algoritmo lamarckiano foi usado. Para permitir que o ligante experimentasse todas as possíveis conformações, o método de busca de Solis e Wets (SOLIS *et al.*, 1981) foi considerado. A grade de interações eletrostática entre a enzima e as moléculas do ligante foram determinadas pelo programa AutoGrid4 (MORRIS *et al.*, 1998). As dimensões da grade foram fixadas em 79x65x55 pontos e com espaçamento de 0,375Å. Esta grade é suficiente para cobrir todos os átomos da enzima próximos ao sítio ativo e para calcular todas as interações entre enzima e ligante.

No cálculo de *docking* usando o programa AutoDock4, as múltiplas conformações da enzima foram consideradas rígidas, entretanto, os ligantes foram considerados flexíveis. Os parâmetros utilizados nos cálculos de *docking* foram os padronizados do programa, exceto para o número de execuções, que foram de 50 para cada ligante. As estruturas dos ligantes foram previamente

otimizados pelo programa MOPAC2009 (STEWART, 2008) usando o método semi-empírico PM6 (STEWART, 2007).

Para a forma apo, a estrutura obtida via *docking* escolhida não foi necessariamente aquela com menor valor de K_i (ou menor energia livre) e sim, aquela que estava com menor K_i e que estivesse entre os resíduos His194, Cys166 e Arg12, como mostra a região em vermelho na FIGURA 3.2.1A. Essa região foi escolhida baseado no fato de que o NAD^+ , principalmente o grupo nicotinamida desse, se liga nessa região. Além disso, o resíduo Arg12 é essencial para a interação via ligação de hidrogênio com o grupo nicotinamida do cofator NAD^+ , como pode-se vê na FIGURA 3.1.1.2.1 da parte 1: Dinâmica Molecular. Por outro lado, para a forma apo, a estrutura obtida via *docking* escolhida foi a que estava próxima e entre os resíduos His194, Cys166 e o cofator NAD^+ (no grupo nicotinamida) e com menor valor de K_i , como mostra a FIGURA 3.2.1B. Para essa forma, o Arg12 não foi utilizado como critério uma vez que esse já está interagindo com o NAD^+ , como visto anteriormente. É bom salientar que é nesta região está localizado o substrato gliceraldeído-3-fosfato, que é essencial ao ciclo da via glicolítica da gGAPDH.

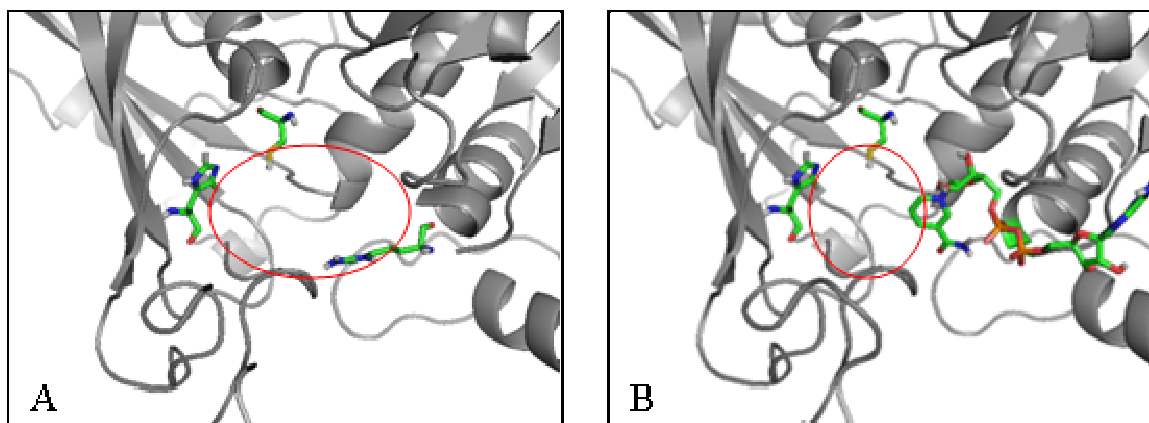


FIGURA 3.15 – Região utilizada como critério para selecionar o ligante a partir dos cálculos de *docking*. A) para a forma apo e B) para a forma holo.

Para a realização dos cálculos de *docking*, duas moléculas foram consideradas (FIGURA 3.16): chalepina e ácido anacárdico. As quais foram submetidas a cálculos de *docking* em um conjunto de 10 estruturas da gGAPDH obtidas a um intervalo de 2 ns de simulação. Esses compostos foram escolhidas por serem bastante conhecidas e estudadas na literatura (VIEIRA *et al.*, 2001; FREITAS *et al.*, 2009). A chalepina foi o primeiro substrato a ser cristalizado no sítio ativo da enzima gGAPDH (PAVÃO *et al.*, 2002). Esse composto apresenta valor de IC_{50} de $64\mu\text{M}$ (VIEIRA *et al.*, 2001). Por definição, IC_{50} é a concentração necessária para inibir 50% da atividade enzimática. Por outro lado, estudos têm mostrado que o ácido anacárdico tem valor IC_{50} de $38\mu\text{M}$ (TAKETA *et al.*, 2004). Na figura a seguir apresentamos as estruturas desses compostos.

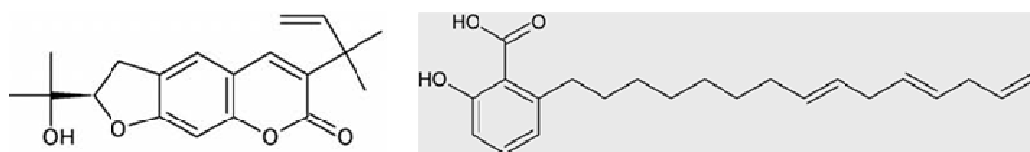


FIGURA 3.16 - Estrutura molecular da chalepina e ácido anacárdico, respectivamente.

Na tabela a seguir apresentamos resultados de cálculos de *docking* para a chalepina e ácido anacárdico obtidos a partir do *ensemble* de estruturas fornecidas pela dinâmica molecular.

TABELA 3.4 - Constante de inibição (K_i em μM) e em parênteses, a energia livre de ligação (kcal/mol) estimada pelo programa AutoDock4.

Estruturas	Chalepina ($\text{IC}_{50} = 64\mu\text{M}$)		Ácido anacárdico ($\text{IC}_{50} = 38\mu\text{M}$)	
	Apo	Holo	Apo	Holo
1K3T	5,27 (-7,20)	25,72 (-6,26)	22,08 (-6,35)	658,85* (-8,43)
0 ns	49,82 (-5,87)	104,39 (-5,43)	197,07 (-5,06)	4,86 (-7,25)
2 ns	82,35 (-5,57)	168,89 (-5,15)	90,68 (-5,51)	1,55 (-7,93)
4 ns	55,16 (-5,81)	130,89 (-5,30)	176,33 (-5,12)	3,84 (-7,39)
6 ns	37,58 (-6,04)	39,89 (-6,00)	210,06 (-5,02)	9,07 (-6,88)
8 ns	18,88 (-6,44)	96,31 (-5,48)	238,93 (-4,94)	3,37 (-7,47)
10 ns	39,53 (-6,01)	13,81 (-6,63)	99,44 (-5,46)	3,33 (-7,47)
12 ns	8,45 (-6,92)	27,01 (-6,23)	51,93 (-5,85)	2,45 (-7,66)
14 ns	59,64 (-5,76)	14,41 (-6,60)	174,68 (-5,13)	13,49 (-6,64)
16 ns	36,52 (-6,05)	144,74 (-5,24)	81,37 (-5,58)	6,63 (-7,07)
18 ns	10,02 (-6,82)	30,69 (-6,16)	13,98 (-6,62)	3,94 (-7,37)
20 ns	12,02 (-6,71)	53,35 (-5,43)	26,35 (-6,25)	1,24 (-8,06)

* Valor em nanomolar.

Para obtenção da estrutura 1K3T na forma holo, essa foi sobreposta com a estrutura 1QXS (estrutura cristalográfica da gGAPDH complexada com 4 cofatores da NAD^+) e as coordenadas estruturais dos cofatores NAD^+ da 1QXS foram salvos dentro do sítio ativo da estrutura 1K3T. Os resultados dos cálculos de *docking* apresentados na tabela anterior nos mostra o quanto os valores de K_i podem variar a partir da estrutura cristalográfica, da estrutura minimizada (0 ns) e ao longo do tempo de simulação. Cálculos de *docking* foram realizados na estrutura cristalográfica da gGAPDH na tabela anterior devido ao fato que a maioria dos cálculos de *docking* reportados na literatura são realizados a partir da estrutura cristalográfica sem a minimização da estrutura enzimática.

Embora não seja possível comparar os valores de K_i com o IC_{50} , os valores de K_i estimados pelos cálculos de *docking* para a chalepina e ácido anacárdico podem ser correlacionados. Observa-se que o ácido anacárdico apresenta menores valores de K_i e IC_{50} , enquanto que para a chalepina, o contrário é observado.

Na tabela anterior (TABELA 3.4), é interessante observar que para a chalepina os melhores valores de K_i são os obtidos para a forma apo. Portanto, esse resultado sugere que a chalepina tende a se ligar melhor na forma apo do que na holo. Isso está de acordo com o observado experimentalmente via cristalografia de raio-X, onde a chalepina está complexada com a gGAPDH sem o cofator NAD^+ (PAVÃO *et al.*, 2002). Por outro lado, para o ácido anacárdico, a melhor energia de interação é a obtida para a forma holo.

Para investigar a interação de outros derivados da chalepina com a gGAPDH, cálculos de *docking* foram realizados para três derivados desta molécula (Fig. 2) e os resultados são apresentados na TABELA 3.2.2. Nesses cálculos de *docking*, apenas as estruturas a 0,0 ns e a obtida em 20,0 ns de simulação para a forma apo e holo foram consideradas.

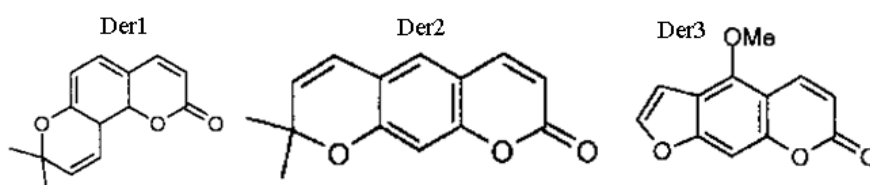


FIGURA 3.17 - Derivados da chalepina considerados nos cálculos de *docking*.

TABELA 3.5 - Valores de K_i estimado e IC_{50} (em micromolar) (VIEIRA *et al.*, 2001) para três derivados da chalepina.

	Der1 ($IC_{50}=123$)		Der2 ($IC_{50}=175$)		Der3 ($IC_{50}=347$)	
<i>Frames</i>	Apo	Holo	Apo	Holo	Apo	Holo
0 ns	37,37	63,31	62,68	99,01	57.92	264.01
20 ns	28,84	29,31	57,38	58,45	151.98	166.97

Na TABELA 3.5, observa-se claramente, principalmente a 0 ns de simulação, que para todos os derivados da chalepina na forma apo apresentam menores valores de K_i e esses resultados estão de acordo com a hipótese ora proposta nesse trabalho.

Na figura a seguir mostramos os três derivados do ácido anacárdico que foram utilizados nos cálculos de *docking*.

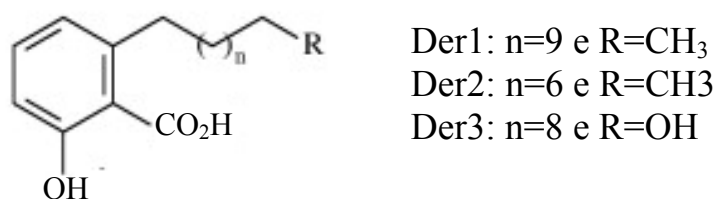


FIGURA 3.18 - Derivados do ácido anacárdico considerados nos cálculos de *docking*.

Na tabela a seguir são apresentados os valores de K_i estimados pelo programa AutoDock4 para os três derivados do ácido anacárdicos apresentados na figura acima.

TABELA 3.6 - Valores de K_i estimados e IC_{50} (em micromolar) (PEREIRA *et al.*, 2008) para três derivados da ácido anacárdico.

Frames	Der1 ($IC_{50}=55\pm5$)		Der2 ($IC_{50}=220\pm17$)		Der3 ($IC_{50}=240\pm22$)	
	Apo	Holo	Apo	Holo	Apo	Holo
0 ns	274,67	392,24	172,37	501,05	198,97	1,19*
20 ns	41,38	234,31	25,56	129,80	68,75	465,48

* Em milimolar.

Pela tabela acima observa-se que os valores de K_i dos derivados são bem maiores que para o ácido anacárdico (veja TABELA 3.4). Pelas suposições anteriores, esperar-se-ia, que da mesma forma que o ácido anacárdico, os seus derivados apresentariam melhores valores de K_i na forma holo. Entretanto, o

contrário foi observado, a forma apo apresenta melhores valores de K_i . Novamente, é bom ressaltar o quanto os valores de K_i diferem quando a estrutura utilizada é a obtida a 20 ns de simulação em comparação com a cristalográfica, a 0 ns de simulação. A partir dos valores de IC_{50} obtidos experimentalmente, seria esperado que os valores de K_i do derivado 2 (Der2) apresentassem K_i maior que o Der1 e menor que o Der3, entretanto isso não ocorre. Assim, esses resultados nos mostram que é preciso muita cautela para tentar correlacionar valores de IC_{50} com K_i .

Fazendo uma média de valores de K_i para o ácido anacárdico a partir da TABELA 3.4, obtemos uma média de 3,33 μ M. Nestes cálculos, os valores obtidos utilizando as estruturas obtida a 6 e 14 ns não foram considerados, devido o seu alto valor em relação as outras. Entretanto, mesmo considerando essas estruturas, o valor de K_i é ainda baixo, 4,78. Essa análise sugere que para considerar um candidato a fármaco derivado do ácido anacárdico mais eficiente que este, o valor de K_i estimado deve ser menor que 3,33 μ M na forma holo da enzima.

Apesar da pequena quantidade de compostos e *ensemble* de estruturas considerados nos cálculos de *docking* como descrito anteriormente, esses resultados preliminares nos indicam um novo caminho para o estudo de compostos que possam ter atividade inibitória. Assim, uma estratégia eficiente na elaboração compostos candidatos a fármacos em onde há a presença de cofatores, é considerar as diferentes formas da enzima: apo e holo. Deste modo, por exemplo, pode-se elaborar um candidato a fármaco que possa inibir a forma apo e um outro para a forma holo. Assim, a combinação desses dois candidatos podem levar a uma maior diminuição ou inibição da atividade da enzima de interesse. É interessante ressaltar que nos estudo de *docking* na enzima, reportados na literatura, apenas a forma apo da enzima tem sido considerada. Assim, a dificuldade para propor novos inibidores cada vez mais potentes pode ser atribuída a essa estratégia usada nesses trabalhos.

Embora a quantidade de cálculos de *docking* realizados aqui seja pequena para um tratamento estatístico adequado, comprova-se o quanto é importante levar em consideração a flexibilidade da enzima em solução. Essa nova abordagem de cálculos de *docking* realizada aqui poderá ser uma grande aliada no desenvolvimento de novas drogas capazes de inibir a enzima gGAPDH. Para refinamentos, estudos mais aprofundados e com mais *ensemble* de estruturas obtidas por dinâmica molecular devem ser utilizados. Dentro desses estudos, cálculos de QSAR (*Quantitative Structure-Activity Relationship*) e *virtual screening* seriam de grande valia para a elaboração de novos ligantes com atividade inibitória da gGAPDH. À partir desses estudos, testes *in vitro* podem ser realizados.

3.3 - Etapa 2: *Docking*

Nessa etapa foi realizado cálculos de *docking* para inserir o G3P no sítio ativo da enzima gGAPDH, que seria a primeira etapa do processo catalítico, como mostra o esquema a seguir.

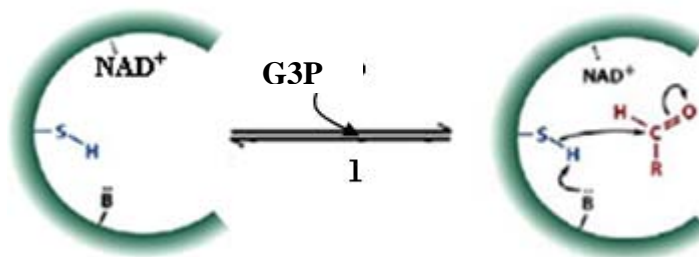


FIGURA 3.19 - Primeira etapa do processo catalítico da gGAPDH.

A estrutura da gGAPDH a ser considerada nos presentes cálculos foi a obtida por dinâmica molecular do sistema gGAPDH-NAD⁺ (forma holo) a 20,0 ns de simulação. Para os cálculos de *docking*, todas as moléculas de águas e contra-íons foram removidos. O procedimento para a preparação dos arquivos de entrada e parâmetros para os cálculos de *docking* foram os mesmos utilizados na etapa anterior (Etapa 1: *Docking*). Na tabela a seguir alguns dados obtidos desses cálculos.

TABELA 3.7 - Valores de constante de inibição e energia livre de ligação estimados nos cálculos de *docking* para a G3P.

Subunidade	K _i (em μM)	Energia livre de ligação (kcal.mol ⁻¹)
A	74,9	-5,63
B	54,0	-5,84
C	20,0	-6,41
D	30,4	-6,16

Pela tabela acima, pode-se observar que os melhores valores de K_i e energia livre de ligação foram obtidos para as subunidades C e D.

Na figura a seguir temos a representação dos complexos formados nos cálculos de *docking*.

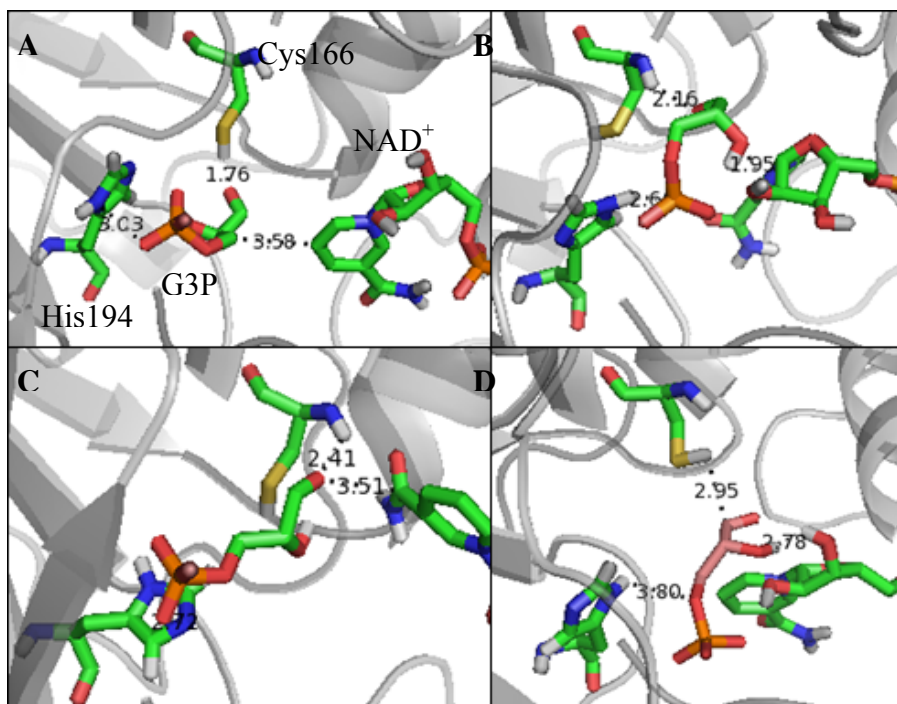


FIGURA 3.20 - Interação do G3P com o NAD⁺ e gGAPDH, obtidos para a estrutura melhor com melhor valor de K_i .

Pela figuras acima, é possível perceber que o grupo aldeído do G3P esta orientado para a Cys166, enquanto que o grupo fosfato esta orientado para a H194. A partir destes dados para o arranjo molecular, compostos podem ser elaborados com as características estruturais e funcionais da G3P para inibir a atividade da enzima gGAPDH.

3.4 - Etapa 3: Dinâmica Molecular

Para entender o efeito e como o G3P se comporta no sítio ativo da enzima gGAPDH em solução aquosa, 1000 ps de simulação foram realizados. O sistema a ser estudado aqui foi o obtido a partir dos cálculos de *docking* da Etapa 2. Para tal, as subunidades com os melhores G3P posicionados no sítio de ativo foram sobrepostos com suas respectivas subunidades do sistema obtido a 20,0 ns de simulação da forma holo. Nessa sobreposição, as moléculas de águas que estavam a uma distância de 0,2 nm da G3P foram removidas. Para manter a neutralidade do sistema 8 íons Cl^- foram removidos da solução, uma vez que cada G3P tem duas cargas negativas. Exceto por essas diferenças do sistema holo da Etapa 1: Dinâmica Molecular, os outros parâmetros e protocolos de simulação foram os mesmos da Etapa 1. Na figura a seguir, apresentamos os valores de RMSD para os sistema gGAPDH-NAD⁺-G3P.

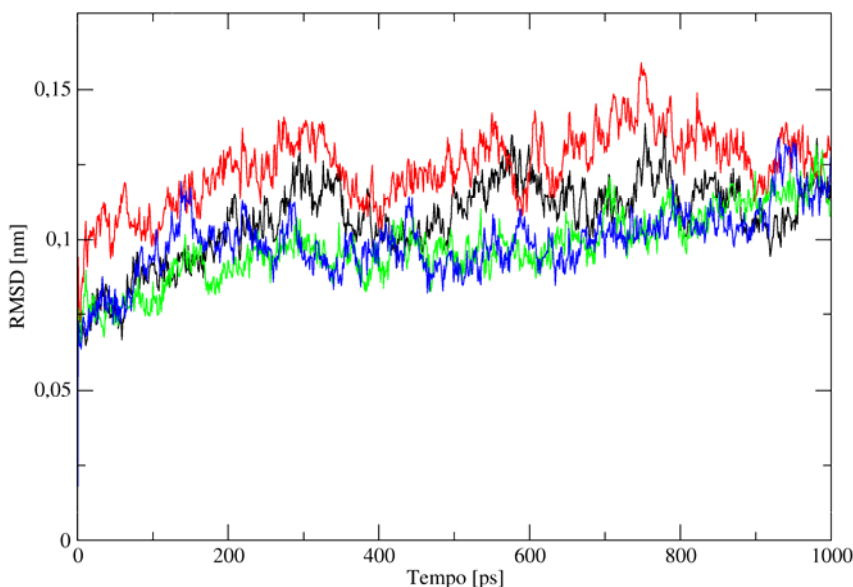


FIGURA 3.21 - RMSD dos átomos C α da enzima gGAPDH do sistema gGAPDH-NAD⁺-G3P. Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).

Pela figura acima, observa-se que para todas subunidades da enzima os valores de RMSD estão abaixo de 0,15 nm. Esse resultado mostra que a enzima complexada com o G3P sofre pequenas mudanças conformacionais ao longo de 1.000 ps de simulação.

Baseado no mecanismo proposto na literatura, a melhor posição do G3P no sítio ativo da enzima gGAPDH deveria ser entre o cofator NAD^+ e os resíduos Cys166 e His194. Baseado nessa observação, a distância entre o centro de massa do G3P e da parte nicotinamida do cofator NAD^+ , Cys166 e His194 foi calculada para cada subunidade, e os dados são mostrados na figura a seguir.

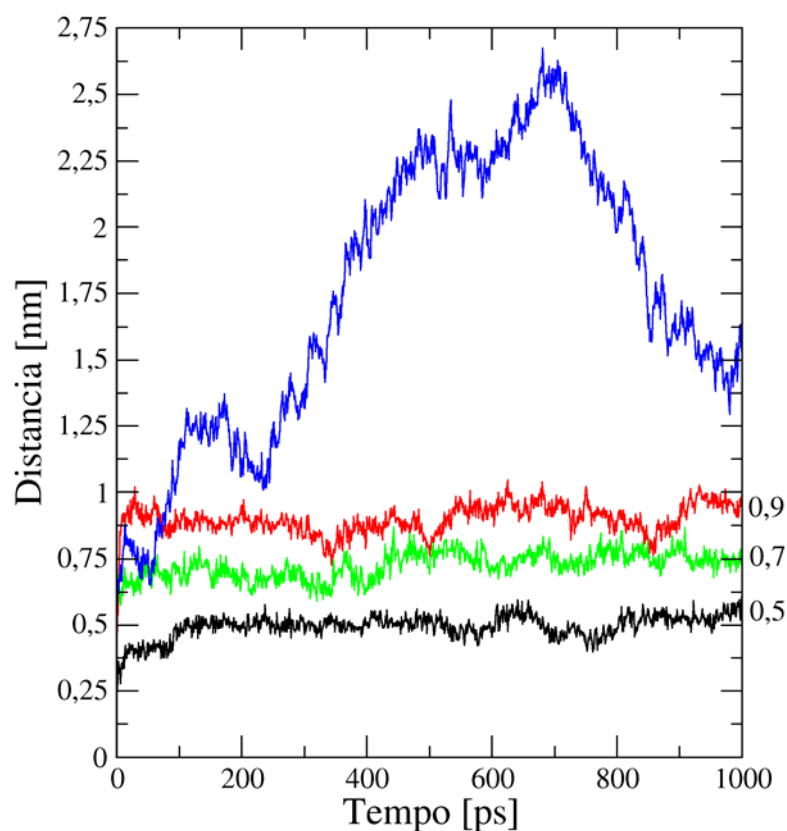


FIGURA 3.22 - Distância entre o centro de massa do G3P e da parte nicotinamida do cofator NAD^+ , Cys166 e His194 ao longo do tempo de simulação. Subunidade A (linha preta), subunidade B (linha vermelha), subunidade C (linha verde) e subunidade D (linha azul).

Pela figura acima, observa-se claramente que o G3P se desloca e se mantém longe da parte nicotinamida do cofator NAD^+ , Cys166 e His194 na

subunidade D. Em contrário, na subunidade A o G3P se matêm próximo a esse grupo de átomos com uma distância de $\sim 0,5$ nm. Esta distância é suficiente para que a reação enzimática possa ocorrer. Para as subunidades B e C, as distâncias se mantêm relativamente constantes ao longo tempo, mas não favorecem um arranjo adequado para o processo enzimático ocorrer. Na figura a seguir, temos o G3P, NAD^+ e os resíduos Cy166 e His194 obtidos a 1.000 ps de simulação.

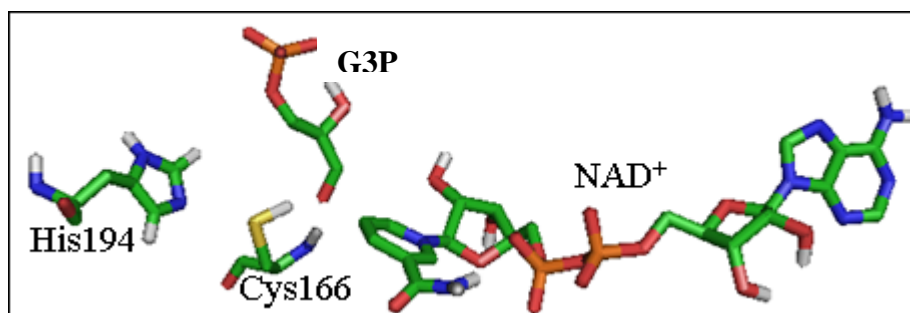


FIGURA 3.23 - *Snapshot* obtido a 1.000 ps de simulação. Para melhor visualização, os hidrogênios apolares foram removidos.

Para identificar quais resíduos estavam interagindo com o G3P via ligação de hidrogênio ao longo do tempo de simulação, na figura a seguir apresentamos essa tipo de análise.

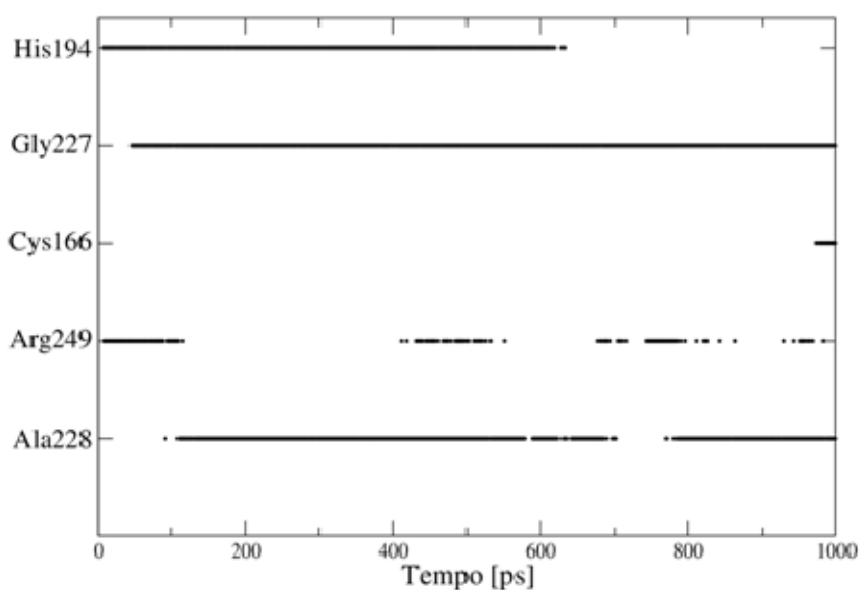


FIGURA 3.24 - Ligações de hidrogênio entre resíduos da enzima e o G3P ao longo do tempo de simulação.

Pela figura acima, pode-se observar que apenas 5 resíduos da enzima gGAPDH participam de ligações de hidrogênio com o G3P ao longo de 1000,0 ps de simulação. É interessante observar que os dois resíduos catalíticos, His194 e Cys166, mantêm ligações de hidrogênio com o G3P. De modo geral, a Fig. 3.4.4 3 mostra o quão esses resíduos são importantes para a estabilização do G3P no sítio ativo da enzima gGAPDH. Assim, na elaboração de candidatos a fármacos, a importância desses resíduos deve ser considerada. Embora os resíduos Gly227 e Ala228 sejam apolares, esses fazem ligação de hidrogênio com G3P a partir do átomo de nitrogênio da ligação peptídica amídica. Na figura a seguir, temos o *snapshot* obtido a 1000 ps de simulação mostrando esses resíduos.

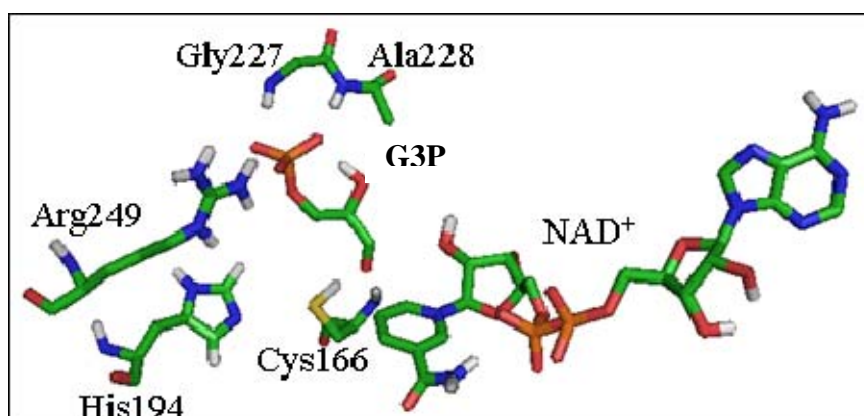


FIGURA 3.25 - *Snapshot* obtido a 1000 ps de simulação mostrando os resíduos que interagem via ligação de hidrogênio com o G3P. Para melhor visualização, os hidrogênios apolares foram removidos.

A partir da figura acima, uma sugestão para a elaboração de compostos candidatos a fármacos seria que esses possuíssem grupos doadores ou aceptores de próton para interagirem com o Arg249, His194 e Cys166. Para a interação com os resíduos Gly227 e Ala228, seria interessante grupos apolares nesses compostos. Assim, esses candidatos a fármacos, pelas análises acima,

deveria ter um parte polar e outra apolar para uma melhor interação na cavidade que o G3P ocupa no sítio ativo.

Nesse ponto da discussão é interessante ressaltar que na Etapa 1: Dinâmica Molecular para a forma holo da enzima, o resíduo Arg249 se deslocou para o centro do sítio catalítico. Naquela parte da discussão a importância desse rearranjo não foi evidenciada. Entretanto, na presente discussão, a importância desse deslocamento ocorre para que o Arg249 possa interagir (FIGURA 3.24) com o G3P (principalmente o grupo fosfato), essa interação estabiliza e orienta o grupo aldeído do G3P ao longo dos resíduos His194 e Cys166, como pode ser observado na FIGURA 3.25. Deste modo, uma boa estratégia para desenvolver moléculas com propriedades inibidoras seria elaborar compostos que possam interferir na interação entre o Arg249 e G3P.

A partir das análises acima, uma hipótese razoável seria a que se um fármaco fosse capaz de inibir a enzima gGAPDH do *T. cruzi*, esse fármaco poderia inibir outras enzimas GAPDH de tripanosomatídeos. Baseado nessa hipótese, alinhamento sequencial da enzima gGAPDH do *T. cruzi* com duas outras enzimas GAPDH de tripanosomatídeos foi realizado. Para isso, a enzima GAPDH do *Trypanosoma brucei* (código PDB: 1GGA) (VELLIEUX *et al.*, 1993) e da *Leishmania mexicana* (código PDB: 1GYP) (KIM *et al.*, 1995) foram consideradas e o programa ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>) foi utilizado para o alinhamento sequencial. O nível de similaridade obtido foi de 90% entre o *T. cruzi* e *T. brucei*, e entre *T. cruzi* e *L. mexicana* foi de 84% de similaridade. Na figura a seguir apresentamos o alinhamento sequencial dessas enzimas obtido pelo programa ClustalW2.

mexicana, onde vários derivados da adenosina foram utilizados como candidatos a inibidores. Nessas medidas, a maioria dos valores de IC₅₀ obtidos para o *T. cruzi* e *T. brucei* foram similares. Por outro lado, para a *L. mexicana* os valores de IC₅₀ foram diferentes com relação ao do *T. cruzi* e *T. brucei*. Entretanto, segundo os autores, essa diferença se deve à presença do resíduo Ser40 não conservado na enzima GAPDH da *L. mexicana*, o qual é replicado pelo Asn no *T. cruzi* e *T. brucei* na região de ligação do grupo adenosina (do cofator NAD⁺) no sítio ativo da enzima (BRESSI *et al.*, 2001). Assim, existem grandes chances dos valores de IC₅₀ ou valores de K_i estimados por cálculos de *docking* sejam similares nos três parasitas na região de ligação do G3P. Ressaltamos que todos os resíduos segeridos aqui como alvos para a elaboração de candidatos a fármacos são conservados na enzima GAPDH destes três parasitas.

Finalizada essa parte do trabalho, na próxima etapa, o sítio ativo da subunidade A do sistema gGAPDH-NAD⁺-G3P será estudado e caracterizado utilizando métodos de química quântica.

3.5 - Etapa 4: Química Quântica

Metodologia de química quântica será utilizada para caracterizar o processo eletrônico que ocorre no sítio ativo da enzima gGAPDH. Essa etapa se faz necessário, uma vez que nos métodos utilizados anteriormente, Dinâmica Molecular e *Docking*, os efeitos eletrônicos não são levados em consideração explicitamente, o que impede investigar mecanismo de reação.

Na figura a seguir temos a segunda etapa de um mecanismo da conversão do G3P em 1,3-BPG aceito na literatura.

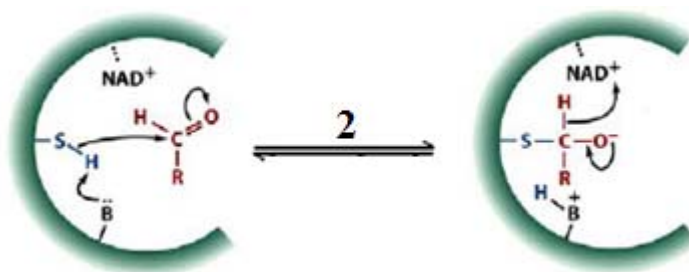


FIGURA 3.27 4- Segunda etapa do mecanismo enzimático da gGAPDH.

Nesse mecanismo, a base (His194) aceita o próton da Cys166, posteriormente, o S¹⁻ da Cys166 ataca o carbono do grupo aldeído do G3P, formando o intermediário tiohemiacetal. Baseado nesse mecanismo, ao longo da simulação, obtivemos vários *snapshots* com a mesma conformação da figura acima, o qual é mostrado na figura a seguir.

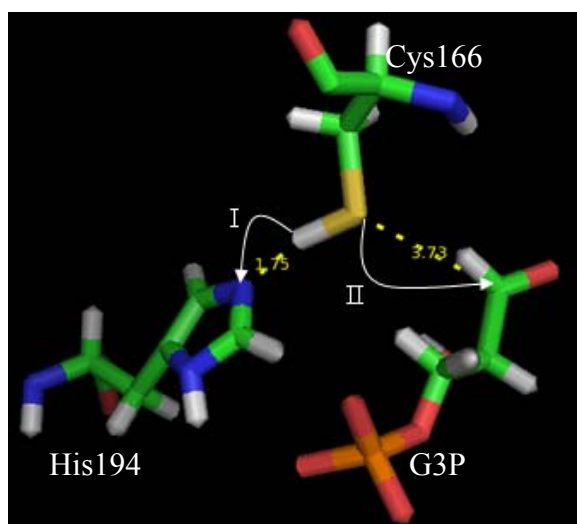


FIGURA 3.28 - *Snapshot* obtido a 719 ps de simulação por dinâmica molecular é utilizado para representar o mecanismo enzimático da gGAPDH proposto na literatura.

Na figura acima, é interessante observar o quanto esse arranjo estrutural obtido pela dinâmica molecular é similar ao proposto na literatura (FIGURA 3.27). O arranjo estrutural da figura foi utilizado em cálculos de mecânica quântica para caracterizar essa região do sítio ativo da enzima gGAPDH. Para isto, estudou-se a coordenada da reação para a transferência do próton da Cys166 para a His194. O método semi-empírico PM6 e o formalismo do MOZYME foram utilizados. Todos os cálculos reportados nessa etapa foram realizados no vácuo. No método semi-empírico PM6, a média do erro do cálculo de calor de formação comparado com o experimental é de $\pm 4,4 \text{ kcal.mol}^{-1}$ para um conjunto de 1.373 compostos (STEWART, 2007). Nesses cálculos de validação do PM6, o procedimento convencional foi utilizado para resolver as equações do SCF. Devido o alto custo deste procedimento para sistemas grandes, os cálculos de otimização da coordenada de reação foi realizado via MOZYME. Entretanto, na literatura ainda foi não citado o erro de utilizar o procedimento MOZYME para um problema específico, *e.g.* o presente sistema estudado. Além disso, para o sistema estudado aqui, não há valor experimental do calor de formação para que possa ser comparado com os obtidos via cálculo

semi-empírico. Deste modo, para ter uma maior confiabilidade nos valores de energia obtidos pelo MOZYME, em cada estrutura otimizada (pelo MOZYME) durante o cálculo da coordenada de reação, foi realizado posteriormente cálculo *single point* (a estrutura não é otimizada) utilizando o procedimento convencional.

Para preparar o sistema a ser tratado quanticamente, um *snapshot* a 719 ps de simulação por dinâmica molecular do sistema gGAPDH-G3P-NAD⁺ em solução foi considerado. Das quatro subunidades desse complexo, apenas a subunidade A, que melhor representou o mecanismo enzimático proposto na literatura, foi considerada. Portanto, as outras subunidades, contra-íons e moléculas de água foram removidas. Para realizar os cálculos quânticos, os resíduos ao redor do cofator NAD⁺ e G3P foram selecionados em um raio de corte de 0,35 nm. O cofator NAD⁺ e o substrato G3P também foram considerados nos cálculos. Deste modo, um total de 354 átomos incluídos e suas coordenadas otimizados no cálculo quântico. Para evitar que esses resíduos ficassem “soltos” no espaço, seus resíduos adjacentes foram também selecionados, somando um total de 344 átomos. Entretanto, as coordenadas destes resíduos de ‘fixação’ não foram otimizadas durante os cálculos de mecânica quântica. Na figura a seguir mostramos um esquema do modelo utilizado nos cálculos.

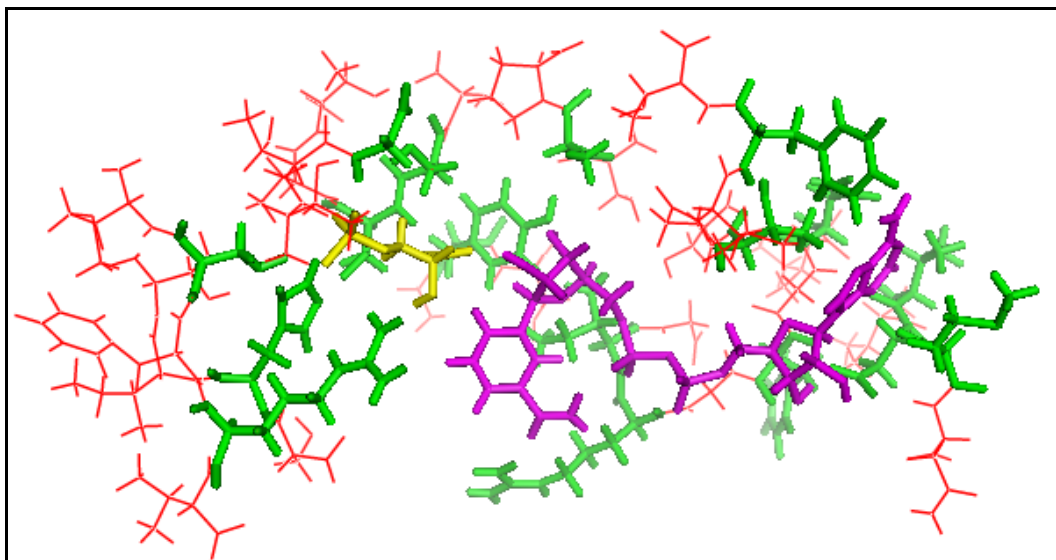


FIGURA 3.29 - Região selecionada para o cálculo de química quântica. Parte fixa (344 átomos) está em vermelho. Parte flexível (354 átomos): resíduos (verde), G3P (amarelo) e NAD^+ (magenta).

Para comparação, um teste de otimização da estrutura do sistema mostrado na FIGURA 3.29 foi realizado utilizando a metodologia MOZYME e o procedimento convencional para a resolução das equações SCF. Neste teste, o tempo de cálculo com o MOZYME (~2 horas) foi cerca 17 vezes menor que o convencional (~35 hrs).

Para gerar a curva de energia *versus* coordenada de reação para a transferência do próton da Cys166 para a His194, a transferência do próton foi “forçada” em 24 etapas, com intervalo de $0,05\text{\AA}$ cada, e as coordenadas dos 354 átomos do modelo foram otimizadas com o procedimento MOZYME. Finalizada essa otimização, cálculo *single point* utilizando o método SCF convencional foi realizado em cada estrutura obtida anteriormente. Na figura a seguir mostramos a variação do calor de formação (ΔH_f) em função da coordenada da reação. Para segurança de análise, incluímos a margem de erro de $\pm 4,4 \text{ kcal.mol}^{-1}$, que desvio para a obtenção do calor de formação via método semi-empírico PM6 com o procedimento convencional de resolução do SCF.

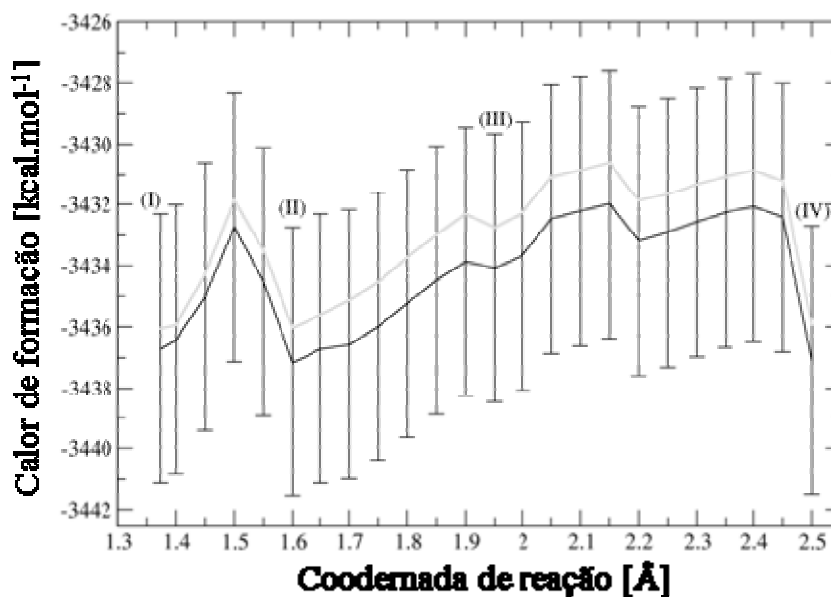


FIGURA 3.30 - Calor de formação em função da coordenada de reação. MOZYME (linha cinza) e método convencional (linha preta).

Pela FIGURA 3.30, percebe-se que as curvas obtidas a partir das duas diferentes metodologias consideradas possuem o mesmo perfil energético. A diferença média do calor de formação ao longo da coordenada de reação entre com o MOZYME e o SCF convencional é de +1,21 kcal.mol⁻¹. Como foi dito anteriormente, não existe uma estimativa para o erro numérico inerente à metodologia MOZYME. Entretanto, no presente estudo as energias obtidas com esta metodologia estão dentro da margem de erro do método convencional, FIGURA 3.30.

Nas figuras a seguir mostramos as estruturas obtidas nos pontos I, II, III e IV destacados na FIGURA 3.30. Para melhor visualização, apenas os resíduos Cys166 e His194, G3P e parte do grupo nicotinamida do NAD⁺ são mostrados.

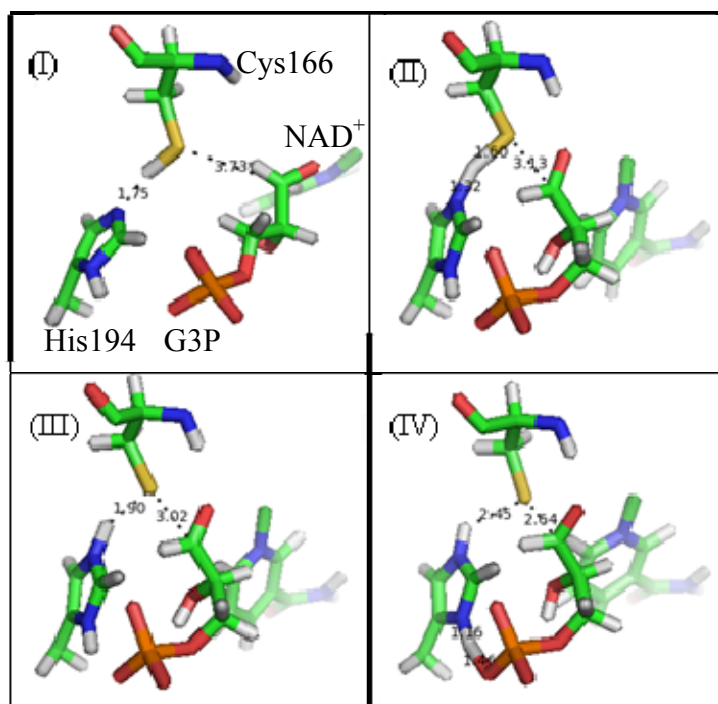


FIGURA 3.31 - Estrutura inicial otimizada (I), intermediário na transição Cys166-H-His194 (II), transferência completa do H da Cys166 para a His194 (III) e estrutura obtida na final da coordenada de reação (IV).

No ponto I (FIGURA 3.30 e FIGURA 3.31), temos a estrutura inicial otimizada no cálculo quântico via MOZYME. O calor de formação (ΔH_f) nesse ponto é de $-3436,05 \text{ kcal.mol}^{-1}$. Em um dos intermediários da coordenada de reação (II), a diferença do calor de formação relativo ($\Delta\Delta H_f$) a estrutura inicial é de $+0,02 \text{ kcal.mol}^{-1}$. Já quando ocorre a transferência total do próton (no ponto III), o $\Delta\Delta H_f$ é de $+3,30 \text{ kcal.mol}^{-1}$. Finalmente, para a última estrutura otimizada na coordenada de reação, o $\Delta\Delta H_f$ é de $-0,31 \text{ kcal.mol}^{-1}$. Embora os valores $\Delta\Delta H_f$ citados acima estejam dentro da margem de erro do cálculo, essas pequenas diferenças de calor de formação mostram que esse mecanismo de transferência do próton é viável. Cartier (CARTIE *et al.*, 1999) tem reportado cálculos de transferência do próton da Cys para a His via cálculos QM/MM na enzima GAPDH da bactéria *Bacillus Stearotherophilus*. Nesse estudo, a parte quântica foi tratada pelo formalismo LSCF (*Local Self Consistent Field*) e o método semi-empírico AM1 foi utilizado. A partir dos cálculos realizados por Cartier, o $\Delta\Delta H_f$ foi de $+11,33 \text{ kcal.mol}^{-1}$. Embora esse resultado

seja maior que o calculado aqui, é bom lembrar que a metodologia usado pelos autores é diferente da usada aqui. Além disso, nos cálculos de Cartier, o substrato G3P não foi considerado. Entretanto, cabe observar que em um processo catalizado por uma enzima, espera-se encontrar valores baixos para as barreiras de energia.

É interessante observar, que ao longo da coordenada de reação a distância entre o enxofre da Cys166 e carbono do grupo aldeído do G3P (onde o S^{1-} da Cys166 se liga) diminui, passando de 0,373 nm (estrutura inicial otimizada) para 0,264 nm (estrutura final otimizada). Distância final essa suficiente para que ocorra o ataque do S^{1-} da Cys166 ao carbono da G3P, formando assim o intermediário tiohemiacetal.

Ao longo das análises e visualizações da trajetória do sistema gGAPDH-G3P-NAD⁺ em solução aquosa, foi encontrado vários *snapshots* com uma outra configuração estável. A partir dessa nova configuração observada, um novo mecanismo foi proposto. A figura a seguir mostra essa conformação, bem como um esquema do mecanismo ora proposto.

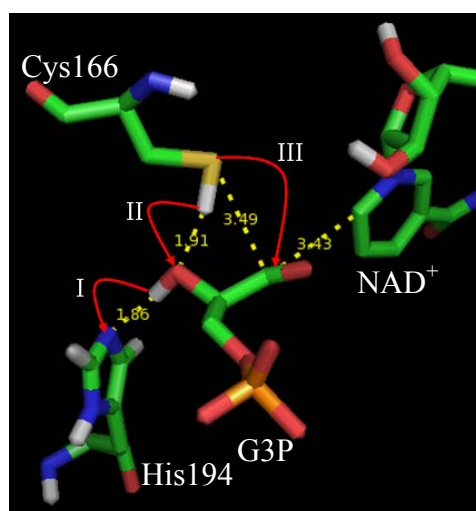


FIGURA 3.32 - *Snapshot* obtido a 530 ps de simulação por dinâmica molecular e mecanismo enzimático da gGAPDH proposto no presente trabalho.

No mecanismo proposto aqui, a partir da figura acima, na primeira etapa (I), ocorre a transferência do próton do G3P para a His194; na segunda etapa (II), o próton do enxofre da Cys166 é transferido para o íon oxigênio do G3P e, finalmente, na terceira etapa (III), a cisteína ativada (Cys^{1-}) ataca o carbono do grupo aldeído do G3P, formando o intermediário tiohemiacetal. Nesse mecanismo, o grupo hidroxila da G3P agiria como um ponte de transferência do próton, uma vez que a distância entre o H da cisteína e o N da His194 é muito grande para que ocorra a transferência direta do próton. Para caracterizar o efeito eletrônico e energético do mecanismo proposto, cálculos semi-empíricos foram realizados. A preparação do sistema a ser tratado quanticamente foi o mesmo para o primeiro mecanismo considerado acima. No novo mecanismo proposto aqui, a coordenada de reação de transferência do próton da G3P para a His194 foi realizada em 26 passos com intervalo de $0,1\text{\AA}$ cada. A figura a seguir mostra o calor de formação ao longo da coordenada de reação obtido pelo cálculo semi-empírico utilizando o MOZYME e cálculos de *single point* com o método SCF convencional.

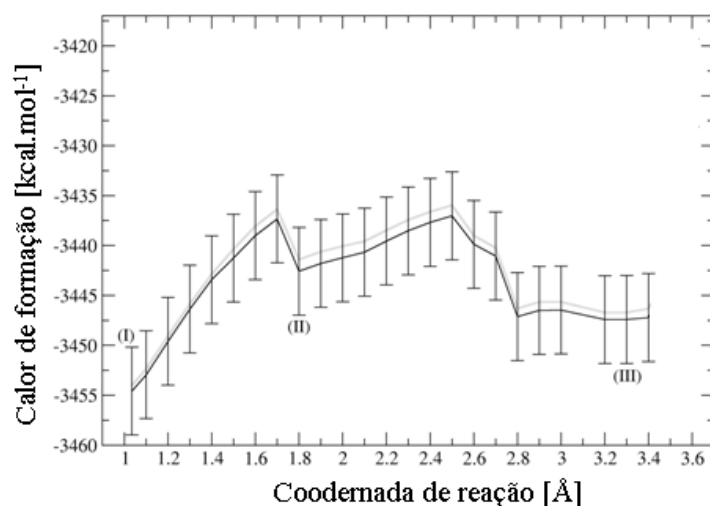


FIGURA 3.33 - Calor de formação em função da coordenada de reação: MOZYME (linha cinza) e método convencional (linha preta).

Assim como para o mecanismo realizado anteriormente, os perfis das curvas obtidas a partir das duas diferentes metodologias são similares. A

média da diferença do calor de formação ao longo da coordenada de reação entre o MOZYME e o convencional foi de $+0.88 \text{ kcal.mol}^{-1}$. Diferença essa menor que quando o mecanismo proposto na literatura foi caracterizado. A curva obtida pelo MOZYME está dentro da margem de erro do método convencional, o que pode ser observado na FIGURA 3.33.

Na figura a seguir detalhamos as estruturas nos pontos I, II e III mostrados na figura acima (FIGURA 3.33).

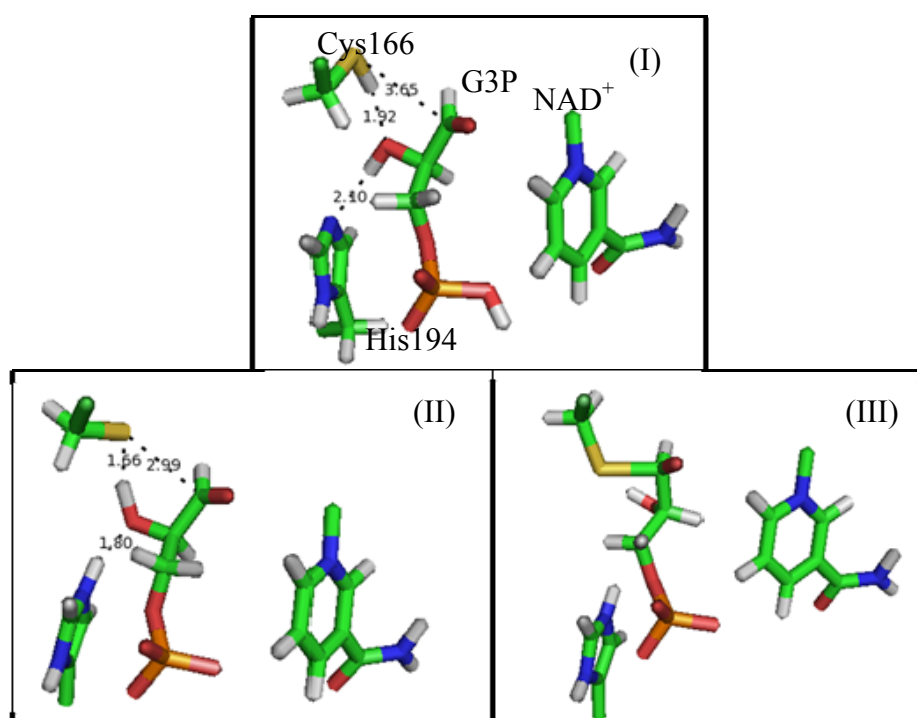


FIGURA 3.34 - Estrutura inicial otimizada (I), transferência do H da Cys166 para His194 (II) e intermediário tiohemiacetal formado (III).

O ΔH_f da estrutura inicial otimizada foi de $-3454,0 \text{ kcal.mol}^{-1}$. Onde, o H do grupo hidroxila da G3P faz ligação de hidrogênio com o N da His194 com distância de $0,185 \text{ nm}$ e H da Cys166 permanece a uma distância de $0,190 \text{ nm}$ em relação ao oxigênio do G3P. No ponto II observou-se a transferência do próton da G3P para a His194 e, espontaneamente, ocorreu a transferência do próton da Cys166 para o íon oxigênio do G3P, exatamente como sugerido no mecanismo proposto aqui. Neste ponto, o $\Delta\Delta H_f$ é de $+12.6 \text{ kcal.mol}^{-1}$. No ponto III, surpreendentemente, ocorre o ataque do S^{1-} da Cys166

no carbono do grupo aldeído do G3P, o que leva a formação espontânea do tiohemiacetal. O calor de formação nesse ponto mostrou-se mais favorável que quando houve a dupla transferência de prótons (ponto II), tendo $\Delta\Delta H_f$ de +7,3 kcal.mol⁻¹. É bom ressaltar que o único “artifício” imposto durante os cálculos semi-empíricos foi “forçar” a coordenada de reação para a aproximação do H da G3P até o N da His194.

Na tabela a seguir ambos mecanismos são comparados.

TABELA 3.8. Calor de formação (em kcal.mol⁻¹) para o mecanismo proposto na literatura e o proposto neste trabalho.

Estrutura	Literatura	Proposto aqui
Inicial otimizada	-3436,05	-3454,01
Transferência do próton	-3432,75	-3441,42
Final otimizada	-3436,36	-3446,71

Pela tabela acima observa-se que para o mecanismo proposto aqui: i) a variação de energia ao longo da coordenada de reação é favorável; ii) os efeitos observados nas coordenadas dos outros átomos é compatível com o mecanismo proposto.

CAPÍTULO 4

CONCLUSÕES

Na presente tese, métodos de química computacional foram utilizados para estudar a enzima gGAPDH do *Trypanosoma cruzi*, o causador da doença de Chagas. Essa enzima faz parte da sexta etapa da via glicolítica, a qual é responsável pela geração energética do parasita. Simulações por dinâmica molecular foram realizadas para obter a evolução temporal da enzima gGAPDH na forma holo (com o cofator NAD^+) e apo (sem o cofator) em solução aquosa. Para melhor representar as condições experimentais em laboratório, o *ensemble* NpT foi utilizado, com temperatura de 298 K e pressão de 1 bar. Nessas simulações, medidas de RMSD indicaram pequenas mudanças conformacionais entre as formas holo e apo no tempo de simulação de 20,0 ns. Na forma holo, o cofator NAD^+ se mantém ao longo do tempo de simulação, no sítio ativo da enzima. Medidas do ângulo de abertura entre o sítio de ligação do NAD^+ e sítio catalítico foram realizadas nas quatro subunidades da enzima. A partir dessas medidas, ficou evidenciado que mudanças ocorridas nesse ângulo não alteram a posição do NAD^+ no sítio de ligação. Medidas de RMSD foram realizadas no sítio de ligação do NAD^+ na forma apo e holo da enzima para entender o efeito do cofator sobre este sítio. Os resultados mostraram que a presença do cofator no sítio de ligação não gera mudanças estruturais drásticas na enzima. Análises de ligação de hidrogênio entre o cofator NAD^+ e enzima foram realizadas ao longo do tempo de simulação. Essas análises mostraram que os resíduos Arg12, Ile13, Asp38 e Met39 são os que mais interagem com a enzima. Isso mostra que esses resíduos são fundamentais para manter o cofator no sítio ativo da enzima. Assim, uma excelente estratégia para a elaboração de candidatos a fármacos seria a proposição de compostos com a capacidade de interferir nessas

interações, bloqueando ou dificultando uma interação efetiva do cofator com a com enzima.

Para posicionar ou encaixar algumas moléculas candidatas a fármacos no sítio ativo da enzima, cálculos de *docking* foram realizados a partir da metodologia de *ensemble docking*. Essa metodologia foi adotada para levar em consideração a flexibilidade da enzima em solução. Múltiplas conformações da enzima foram obtidas a cada intervalo de 2,0 ns de simulação para as formas holo e apo. A chalepina, ácido anacárdico e três derivados de cada um desses compostos foram considerados nos cálculos de *docking*. Esses compostos foram escolhidos por apresentar ótimos valores de constantes de inibição (K_i) da enzima gGAPDH. Valores de K_i estimados nos cálculos de *docking* mostraram que chalepina tende a inibir a forma apo da enzima, enquanto o ácido anacárdico tende a inibir a forma holo. Esses resultados sugerem que a melhor estratégia para elaborar compostos candidatos a fármacos é considerar as duas formas da enzima (holo e apo): um composto inibiria a forma apo e o outro a forma holo. Conseqüentemente, espera-se que um ‘coquetel’ formado por uma combinação desses dois compostos resulte em uma inibição maior da enzima.

Para a caracterização do mecanismo enzimático da enzima gGAPDH, o gliceraldeído 3-fosfato (G3P) foi colocado no sítio ativo (via cálculos de *docking*) da enzima gGAPDH na forma holo, utilizando uma estrutura obtida a 20,0 ns de simulação. Nesses cálculos, os melhores valores de K_i foram obtidos quando o G3P estava interagindo com o cofator NAD^+ e com os resíduos Cys166 e His194. Os melhores valores de K_i foram obtidos nas subunidades C e D da enzima.

Para obter a relaxação e equilibração do sistema gGAPDH- NAD^+ -G3P em solução, foi realizado 1,0 ns de simulação por dinâmica molecular. Esse sistema foi construído a partir do complexo obtido dos cálculos de *docking* anteriormente descritos. Ao longo desse 1,0 ns de simulação, o G3P se deslocou do sítio ativo da enzima nas subunidades B, C e D. Entretanto, este se manteve

no sítio ativo da subunidade A ao longo do tempo de simulação considerado. A estabilização do G3P nessa subunidade se deve as ligações de hidrogênio formadas com os resíduos His194, Cys166, Gly227, Ala228 e Arg249. Deste modo, as características desses resíduos, tais como polaridade, hidrofobicidade e outras, as quais podem ser utilizadas para elaborar moléculas candidatas a fármacos que possam interagir efetivamente, bem como ocupar um volume no sítio ativo da enzima similar ao ocupado pelo G3P

Ao longo da simulação do sistema gGAPDH-NAD⁺-G3P, duas estruturas foram escolhidas para a realização de cálculos de química quântica. Cálculos foram realizados para levar em consideração o efeito eletrônico do sistema, uma vez que nas metodologias de dinâmica molecular e de *docking*, o efeito dos elétrons não é levado em consideração explicitamente. Devido ao alto número de átomos envolvidos, o procedimento MOZYME foi utilizado para a resolução das equações do campo autoconsistente (SCF) em conjunto com o método semi-empírico PM6. O procedimento MOZYME, para o presente sistema, mostrou-se 17 vezes mais rápido na resolução das equações SCF que o método convencional. Em uma das estruturas obtidas da simulação do sistema gGAPDH-NAD⁺, um dos mecanismos propostos na literatura foi caracterizado a partir da coordenada de reação para a transferência do próton da Cys166 para a His194. A partir de uma segunda estrutura obtida da dinâmica molecular, um novo mecanismo para a enzima gGAPDH foi proposto e caracterizado via cálculos quânticos. Nesse mecanismo, o grupo hidroxila da G3P age como uma ponte para a transferência do próton da Cys166 para a His194. Ao longo da coordenada de reação, obtida pelos cálculos semi-empíricos, foi observado a reação química acontecer “naturalmente” como previsto na literatura. Os valores de calor de formação obtidos para o mecanismo ora proposto são similares aos obtidos para o mecanismo proposto na literatura.

CAPÍTULO 5

SUGESTÕES PARA FUTUROS TRABALHOS

Apresentaremos a seguir de forma sucinta alguns temas que julgamos importantes para o futuro da pesquisa nesta área.

- Para entender e determinar o efeito cooperativo ou não da enzima gGAPDH, realizar pelo menos duas novas simulações considerando 1 e 2 cofatores em cada simulação.
- Realizar cálculos de *docking* com mais compostos candidatos a fármacos, bem como considerar mais conformações obtidas via dinâmica molecular.
- Realizar cálculos de *docking* ao longo da região ocupada pelo cofator NAD^+ nos sítio ativo. Para isso, a um *ensemble* de estruturas da forma apo da enzima seria utilizado. Como sugestão, três regiões seriam interessantes: primeira, a região ocupada pelo grupo nicotinamida; segunda, a região ocupado pelo grupo fosfato e terceira, na região do grupo adenina do NAD^+ . Tais compostos podem ser planejados utilizando semelhanças químicas com os grupos do cofator.
- Realizar estudos *virtual screening* para uma busca de ligantes bioativos em bibliotecas virtuais de compostos utilizando um *ensemble* de estruturas obtidas por dinâmica molecular.
- Realizar estudos de QSAR (relação quantitativa entre estrutura química e atividade) usando um conjunto de múltiplas conformações da enzima obtido via dinâmica molecular.

- Realizar cálculos de química computacional para completar o ciclo de etapas envolvidas na conversão do G3P em 1,3-BPG.
- Realizar cálculos de *docking* na região de ligação do G3P nas enzimas GADPH dos parasitas *T. brucei* e *L. mexicana* para validar a suposição que é suficiente utilizar a enzima de apenas um desses parasitas para a realização dos estudos e cálculos.

CAPÍTULO 6

RESUMO DO DOUTORADO SANDUÍCHE

No presente capítulo faremos uma descrição sucinta do doutorado sanduíche que foi desenvolvido em um ano de pesquisa no *Pacific Northwest National Laboratory* (PNNL) localizado na cidade de Richland – WA – USA, tendo como supervisor o Dr. Roberto Dias Lins. O título do projeto de pesquisa do doutorado sanduíche foi: “Estudo computacional da enzima E4: um eficiente conversor de biomassa”. O órgão financiador da bolsa de estudo dessas pesquisas foi o CNPq. Enquanto que, o PNNL foi o responsável pelo fornecimento do espaço e dos recursos computacionais para a execução do projeto.

Os sistemas estudados no exterior foram diferentes dos realizados no Brasil, mas metodologias as empregadas foram as mesmas. Parte dos resultados obtidos foi publicada no *Journal of Molecular Recognition* (OLIVEIRA, O. V.; FREITAS, L. C. G.; STRAATSMA, T. P. & LINS, R. D. “Interaction between the CBM of Cel9A from *Thermobifida fusca* and cellulose fibers”. *Journal of Molecular Recognition*, **22**: 38-45, 2009).

Atualmente dois artigos encontram-se em fase de preparação e correção e em breve serão submetidos para publicação:

- ✓ OLIVEIRA, O. V.; FREITAS, L. C. G.; STRAATSMA, T. P. & LINS, R. D. “Molecular dynamics simulation of wild-type Cel6A cellulase and its D79A and D265A mutations from *Thermobifida fusca*”.
- ✓ OLIVEIRA, O. V.; FREITAS, L. C. G.; STRAATSMA, T. P. & LINS, R. D. “New insights into the mechanistic of the endo/exocellulase Cel9A from *Thermobifida fusca*”.

Devido às condições do acordo de cooperação UFSCar & PNNL, o conteúdo de qualquer texto precisa ser aprovado antes de sua divulgação. Isto ocorre porque estes estudos visam mais a obtenção de patentes que propriamente a divulgação acadêmica dos resultados. Entretanto, o artigo que foi publicado é de livre acesso a comunidade científica e pode ser obtido a partir do DOI:10.1002/jmr.925. Acreditamos que o mesmo contenha as informações necessárias para que a banca julgue a qualidade desta etapa do presente trabalho de doutoramento.

CAPÍTULO 7

REFERÊNCIAS BIBLIOGRÁFICAS

ALLEN, M. P. & TILDESLEY, D. J. *Computer Simulation of Liquids*, Oxford University Press, 1987.

ARFKEN, G. *Mathematical Methods for Physicists*, Academic Press, Orlando, FL, 3rd ed., 1985.

BEEMAN, D. "Some multistep methods use in molecular dynamics calculations". *Comput. Phys.*, **20**: 130, 1976.

BERENDSEN, H. J. C.; POSTAMA, J. P. LM. & VAN GUNSTEREN, W. F. in: B. Pullman (Eds.), *Intermolecular Forces*, Reidel, Dordrecht, 1981.

BERENDSEN, H. J. C.; POSTMA, J. P. M.; DINOLA, A. & HAAK, J. R. "Molecular dynamics with coupling to an external bath". *J. Chem. Phys.*, **81**: 3684, 1984.

BERENDSEN, H. J. C. & VAN GUNSTEREN, W. F. "Practical algorithms for dynamic simulations. Molecular dynamic simulation of statistical mechanical systems". IN: *Proceedings of the Enrico Fermi Summer School*. Ciccotti, G. & Hoover, W. G. (Eds.). Soc. Italiana di Fisica, Bologna, p. 43-65, 1986.

BERENDSEN, H. J. C.; VAN DER SPOEL, D. & VAN DRUNEN, R. "GROMACS: A message-passing parallel molecular dynamics implementation". *Comp. Phys. Comm.*, **91**: 43, 1995.

BRESSI, J. C.; VERLINDE, C. L. M. J.; ARONOV, A. M.; MY LE SHAW; SHIN, S. S.; NGUYEN, L. N.; SURESH, S.; BUCKNER, S.; VAN VOORHIS, W. C.; KUNTZ, I. D.; HOL, W. G. J. & GELB, M. H. "Adenosine analogues as selective inhibitors of glyceraldehyde-3-phosphate dehydrogenase of Trypanosomatidae via structure-based drug design". *J. Med. Chem.*, **44**(13): 2080, 2001.

CASTILHO, M. S.; PAVÃO, F.; OLIVA, G.; LADAME, S.; WILLSON, M. & PERIE, J. "Evidence for the Two Phosphate Binding Sites of an Analogue of the Thioacyl Intermediate for the *Trypanosoma cruzi* Glyceraldehyde-3-

phosphate Dehydrogenase-Catalyzed Reaction, from Its Crystal Structure". *Biochemistry*, **42**: 7143, 2003.

CARTIER, A.; BROWN, D.; MAIGRET, B; BOSCHI-MULLER, S.; RAHUEL-CLERMONT, S. & BRANLANT, G. "Modelling the active site of glyceraldehyde-3-phosphate dehydrogenase with LSCF formalism". *Theor. Chem. Acc.*, **101**: 241, 1999.

CORBEIL, C. R.; ENGLEBIENNE, P. & MOITESSIER, N. "Docking Ligands into Flexible and Solvated Macromolecules. 1. Development and Validation of FITTED 1.0". *J. Chem. Inf. Model.*, **47**: 435, 2007.

CORNELL, W.D.; CIEPLAK, P.; BAYLY, C. I.; GOULD, I. R.; MERZ, K. M. J. R.; FERGUSON, D. M.; SPELLMEYER, D. C.; FOX, T.; CALDWELL, J. W. & KOLLMAN, P. A. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". *J. Am. Chem. Soc.* **117**: 5179, 1995.

COURA, J. R. & CASTRO, S. L. "A critical review on Chagas disease chemotherapy". *Memb. Inst. Oswaldo Cruz*, **97**: 3, 2002.

DE MARCHI, A.; CASTILHO, M.; NASCIMENTO, P.; ARCHANJO, F.; DELPONTE, G.; OLIVA, G. & PUPO, M. "New 3-piperonylcoumarins as inhibitors of glycosomal glyceraldehyde-3-phosphate dehydrogenase (gGAPDH) from *Trypanosoma cruzi*". *Bioorg. & Med. Chem.*, **12**: 4823, 2004.

DeLano, W.L. The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA.

DEWAR, M. J. S. & THIEL, W. "Ground States of Molecules, 38. The MNDO Method. Approximations and Parameters". *J. Am. Chem. Soc.*, **99**: 4899, 1977.

DEWAR, M. J. S.; ZOEBISCH, E. G.; HEALY, E. F. & STEWART, J. J. P. "AM1: A New General Purpose Quantum Mechanical Model". *J. Am. Chem. Soc.*, **107**: 3902, 1985.

DOCAMPO, R. "Recent developments in the chemotherapy of Chagas' Disease". *Curr. Pharm. Des.*, **7**: 1157, 2001.

ENGEL, J. C.; FRANKE, C. B. M.; STOPPANI, A. O.; CANNATTA, J. J. & CAZZULO, J. J. "Aerobic glucose fermentation by *Trypanosoma cruzi* axenic culture amastigote-like forms during growth and differentiation to epimastigotes". *Mol. Biochem. Parasitol*, **26**: 1, 1987.

FAIRLAMB, A. H. "Future prospects for the chemotherapy of Chagas' disease". *Medicina (Buenos Aires)*, **59**: 179, 1999.

FREITAS, L. C. G., "Prêmio Nobel de Química em 1998: Walter Kohn e John A. Pople". *Quím. Nova*, **22**, 293, 1999.

FREITAS, R. F.; PROKOPCZYK, I. M.; ZOTTIS, A.; OLIVA, G.; ANDRICOPULO, A. D.; TREVISAN, M. T. S.; VILLEGAS, W.; SILVA, M. G. V. & MONTANARI, C. A. "Discovery of Novel *Trypanosoma cruzi* Glyceraldehyde-3-Phosphate Dehydrogenase Inhibitors". *Bioorg. & Med. Chem.*, **17**: 001, 2009.

GASTEIGER, J. & MARSILI, M. "Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges" *Tetrahedron*, **36**: 3219, 1980.

GEAR, C. W. *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewoodcliffs, New York, 1974.

GUIDO, R. V. C.; CASTILHO, M. S.; MOTA, S. G. R.; OLIVA, G. & ANDRICOPULO, A. D. "Classical and Hologram QSAR Studies on a Series of Inhibitors of Trypanosomatid Glyceraldehyde-3-Phosphate Dehydrogenase". *QSAR & Combinatorial Science*, **27**: 768, 2008.

GNU (PROJETO GNU) <http://www.gnu.org/>

HESTENES, M. R. & STIEFEL, E. "Methods of Conjugate Gradients for Solving Linear Systems". *J. Res. Nat. Bur. Stand.*, **49**(6): 409, 1952.

HOCKNEY, R. W.; GOEL, S. P. & EASTWOOD, J. W. "Quiet high-resolution computer models of a plasma". *J. Comp. Phys.*, **14**: 148, 1974.

HOOVER, W. G., "Canonical dynamics: Equilibrium phase-space distributions". *Phys. Rev. A*, **31**: 1695, 1985.

HUANG, S. Y. & ZOU, X. "Ensemble docking of multiple protein structures: Considering protein structural variations in molecular docking". *Proteins: Structure, Function, and Bioinformatics*, **62**: 399, 2006.

JONES, G.; WILLETT, P.; GLEN, R. C.; LEACH, A. R. & TAYLOR, R. "Development and validation of a genetic algorithm for flexible docking". *J. Mol. Biol.*, **267**: 727, 1997.

JORGENSEN, W. L. "Rusting of the lock and key model for protein-ligand binding". *Science* **254**(5034): 954, 1991.

JUNIOR, J. A.; DIAS, R. L. A.; CASTILHO, M. S.; OLIVA, G. & CORRÊA, A. G. "Preparation and Evaluation of a Coumarin library towards the inhibitory activity of the enzyme gGAPDH from *Trypanosoma cruzi*". *Braz. Chem. Soc.*, **4**: 763, 2005.

KIM, H.; FEIL, I. K.; VERLINDE, C. L.; PETRA, P. H. & HOL, W. G. "Crystal structure of glycosomal glyceraldehyde-3-phosphate dehydrogenase from *Leishmania mexicana*: implications for structure-based drug design and a new position for the inorganic phosphate binding site". *Biochemistry*, **34**: 14975, 1995.

KUNTZ, I. D.; BLANEY, J. M.; OATLEY, S. J.; LANGRIDGE, R. & FERRIN, T. E. "A geometric approach to macromolecule-ligand interactions". *J. Mol. Biol.* **161**: 269, 1982.

LADAME, S.; FAURÉ, R.; DENIER, C.; LAKHDAR-GHAZAL, F. & WILLSON, M. "Selective inhibition of *Trypanosoma cruzi* GAPDH by "bi-substrate" analogues". *Org. Biom. Chem.*, **3**(11): 2070, 2005.

LADAME, S.; CASTILHO, M. S.; SILVA, C. H., DENIER, C.; HANNAERT, V.; PERIE, J.; OLIVA, G. & WILLSON, M. "Crystal structure of *Trypanosoma cruzi* glyceraldehyde-3-phosphate dehydrogenase complexed with an analogue of 1,3-bisphospho-d-glyceric acid". *Eur. J. Biochem.* **270**: 4574, 2003.

LEACH, A.R. *Molecular Modelling. Principles and Applications*, Addison Wesley Longman Limited, 1996.

LEITÃO, A.; ANDRICOPULO, A. D.; OLIVA, G.; PUPO, M. T.; MARCHI, A. A.; VIEIRA, P. C.; SILVA, M. F. G. F.; FERREIRA, V. F.; SA, M. M. & MONTANARI, C. A. "Structure-Activity Relationships of Novel Inhibitors of Glyceraldehyde-3-Phosphate Dehydrogenase". *Bioorg. & Med. Chem. Letters*, **14**: 2199, 2004.

LEVINE, I. N., *Quantum Chemistry*, Prentice-Hall Inc, New York, 1991.

LINDAHL, E.; HESS, B. & VAN DER SPOEL, D. "GROMACS 3.0: A package for molecular simulation and trajectory analysis", *J. Mol. Mod.*, **7**: 306, 2001.

MONCAYO, A. "Chagas disease: current epidemiological trends after the interruption of vectorial and transfusional transmission in the Southern Cone countries". *Mem. Inst. Oswaldo Cruz*, **98**: 577, 2003.

MONCAYO, A.; ORTIZ YANINE, M.I. "An update on Chagas disease (human American trypanosomiasis)". *Ann. Trop. Med. Parasitol.*, **100**: 663, 2006.

MOPAC2009, STEWART, J. J. P. *Computational Chemistry*, Colorado Springs, CO, USA, [HTTP://OpenMOPAC.net](http://OpenMOPAC.net) (2008).

MORGON, N. H. & COUTINHO, K. *Métodos de Química Teórica e Modelagem Molecular*. Ed. Livraria da Física, São Paulo, 2007.

MORRIS, G. M.; GOODSELL, D. S.; HALLIDAY, R. S.; HUEY, R.; HART, W. E.; BELEW, R. K. & OLSON, A. J. "Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function". *J. Comput. Chem.* **19**: 1639, 1998.

NABUURS, S. B., WAGENER, W. & DE VLIEG, J. "A Flexible Approach to Induced Fit Docking". *J. Med. Chem.*, **50**(26): 6507, 2007.

NOSÉ, S. "A molecular dynamics method for simulations in the canonical ensemble". *Mol. Phys.*, **52**: 255, 1984.

OOSTENBRINK, C.; VILLA, A.; MARK, A. E. & VAN GUNSTEREN, W. F. "A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6". *J. Comp. Chem.* **25**: 656, 2004.

PARRINELLO, M. & RAHMAN, A. "Polymorphic transitions in single crystals: A new molecular dynamics method". *J. Appl. Phys.*, **52**(12): 7182, 1981.

PAVÃO, F.; CASTILHO, M. S.; PUPO, M. T.; DIAS, R. L.; CORRÊA, A. G.; FERNANDES, J. B.; DA SILVA, M. F.; MAFEZOLI, J.; VIEIRA, P. C. & OLIVA, G. "Structure of *Trypanosoma cruzi* glycosomal glyceraldehyde-3-phosphate dehydrogenase complexed with chalepin, a natural product inhibitor, at 1.95 Å resolution". *FEBS Lett.*, **520**: 13, 2002.

PEREIRA, J. M.; SEVERINO, R. P.; VIEIRA, P. C.; FERNANDES, J. B.; SILVA, M. F. G. F.; ZOTTIS, A.; ANDRICOPULO, A. D.; OLIVA, G. & CORREA, A. G. "Anacardic Acid Derivatives as Inhibitors of Glyceraldehyde-3-Phosphate Dehydrogenase from *Trypanosoma cruzi*". *Bioorg. & Med. Chem.*, **16**: 8889, 2008.

POPLE, J. A.; SANTRY, D. P. & SEGAL, G. A. "Approximate Self-Consistent Molecular Orbital Theory I. Invariant Procedures". *J. Chem. Phys.*, **43**: S129, 1965.

POPLE, J. A.; BEVERIDGE, D. L. & DOBOSH, P. A. "Approximate self-consistent molecular-orbital theory. V. Intermediate neglect of differential overlap". *J. Chem. Phys.*, **47**: 2026, 1967.

RAREY, M.; KRAMER, B.; LENGAUER, T. & KLEBE, G. "A fast flexible docking method using an incremental construction algorithm". *J. Mol. Biol.* **261**: 470, 1996.

RODRIGUEZ, J. B. "Specific molecular targets to control tropical diseases". *Curr. Pharm. Des.*, **7**: 125, 2001.

SALMON, L.; LANDRY, V.; MELNYK, O.; MAES, L.; SHERGHERAERT, C. & DAVIOUD-CHARVET, E. A. "A general approach to the synthesis of polyamine linked-monoindolylmaleimides, a new series of trypanothione reductase inhibitors". *Chem. Pharm. Bull.*, **46**: 707, 1998.

SANNER, M. F. "A component-based software environment for visualizing large macromolecular assemblies". *Structure*, **13** (3): 447, 2005.

SCHOFIELD, C. J.; JANNIN, J. & SALVATELLA, R. "The future of Chagas disease control". *Trends Parasitol.*, **22**: 583, 2006.

SILVA, J. J. N.; OSAKABE, A. L.; PAVANELLI, W. R.; SILVA, J. S. & FRANCO, D. W., "In vitro and in vivo antiproliferative and trypanocidal activities of ruthenium NO donors". *Br. J. Pharmacol.*, **152**(1): 112, 2007.

SILVA, J. J. N.; GUEDES, P. M.; ZOTTIS, A.; BALLIANO, T. L.; SILVA, O.; LOPES, L.; ELLENA, J. A.; OLIVA, G. & ANDRICOPULO, A. D. "In Vitro and In Vivo trypanocidal Activity of Ruthenium Complexes as Inhibitors of Glyceraldehyde-3-Phosphate Dehydrogenase from *Trypanosoma cruzi*". *British J. Clinical Pharmacol.*, **156**: 1, 2009.

SOLIS, F. J. & WETS, J. B. "Minimization by Random Search Techniques". *Math. Oper. Res.*, **6**: 19, 1981.

SOUZA, D. H. F.; GARRATT, R. C.; ARAUJO, A. P. U.; GUIMARAES, B. G.; JESUS, W. D. P.; MICHAELS, P. A. M.; HANNAERT, V. & OLIVA, G. "Trypanosoma cruzi glycosomal glyceraldehyde-3-phosphate dehydrogenase: structure, catalytic mechanism and targeted inhibitor design". *FEBS Letters*, **242**: 131, 1998.

STEWART, J. J. P. "Optimization of Parameters for Semi-Empirical Methods I-Method". *J. Comp. Chem.*, **10**: 209, 1989.

STEWART, J. J. P. "Application of localized molecular orbitals to the solution of semi-empirical self-consistent field equations". *Int. J. Quantum Chem.*, **58**: 133, 1996.

STEWART, J. J. P. "Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements", *J. Mol. Model.*, **13**: 1173, 2007.

SURESH, S.; BRESSI, J. C.; KENNEDY, K. J.; VERLINDE, C. L. M. J.; GELB, M. H. & HOL, W. G. J. "Conformational changes in *Leishmania mexicana* glyceraldehyde-3-phosphate dehydrogenase induced by designed inhibitors". *J. Mol. Biol.* **309**: 423, 2001.

TAKETA, A. T. C.; GNOATTO, S. C. B.; GOSMANN, G.; PIRES, V. S.; SCHENKE, E. P. & GUILLAUME, D. "Triterpenoids from Brazilian *Ilex* species and their in vitro antitrypanosomal activity". *J. Nat. Prod.*, **67**: 1697, 2004.

TOMAZELA, D. M.; PUPO, M. T.; PASSADOR, E. A.; SILVA, M. F. G. F.; VIERA, P. C.; FERNANDES, J. B.; RODRIGUES FO., E.; OLIVA, G. & PIRANI, J. R. "Pyrano chalcones and a flavone from *Neuraputia magnifica* and their *Trypanosoma cruzi* glycosomal glyceraldehyde-3-phosphate dehydrogenase-inhibitory activities". *Phytochemistry*, **55**: 643, 2000.

URBINA, J. A. & DOCAMPO, R., "Specific chemotherapy of Chagas disease: controversies and advances". *Trends Parasitol.*, **11**(11): 495, 2003.

VAN GUNSTEREN, W. F. & BERENDSEN, H. J. C. "Algorithms for macromolecular dynamics and constraint dynamics". *Mol. Phys.*, **34**: 1311, 1977.

VELLIEUX, F.M.; HAJDU, J.; VERLINDE, C. L.; GROENDIJK, H.; READ, R. J.; GREENHOUGH, T. J.; CAMPBELL, J. W.; KALK, K. H., LITTLECHILD, J. A. & WATSON, H. C. "Structure of glycosomal glyceraldehyde-3-phosphate dehydrogenase from *Trypanosoma brucei* determined from Laue data". *Proc. Natl. Acad. Sci. USA*, **90**: 2355, 1993.

VERLET, L. "Computer "experiments" on classical fluids 1: Thermodynamical properties of Lennard Jones Molecules". *Phys. Rev.*, **159**: 98, 1967.

VERDONK, M. L.; COLE, J. C.; HARTSHORN, M. J.; MURRAY, C. W. & TAYLOR, R. D. "Improved protein-ligand docking using GOLD". *Proteins: Structure, Functions and Bioinformatics*, **52**: 609, 2003.

VIEIRA, P. C.; MAFEZOLI, J.; PUPO, M. T.; FERNANDES, J. B.; SILVA, M. F. G. F.; ALBUGUERQUE, S.; OLIVA, G. & PAVÃO, F. "Strategies of the

isolation and identification of trypanocidal compounds from the Rutales”. *Pure Appl. Chem.*, **73**: 617, 2001.

WORLD HEALTH ORGANIZATION. “Control of Chagas disease: second report of the WHO expert committee”. Geneva: WHO, 2002. (Technical Report Series, 905).

ZUCCOTTO, F.; BRUM, R.; PACANOWSKA, D. G.; PEREZ, L. M. R. & GILBERT, I. H., “The structure-based design and synthesis of selective inhibitors of *Trypanosoma cruzi* dihydrofolate reductase”. *Bioorg. Med. Chem. Lett.*, **9**: 1463, 1999.