

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS
DEPARTAMENTO DE FÍSICA E INFORMÁTICA

“Estudos Estruturais e Cinéticos da Enzima Gliceraldeído-3-Fosfato Desidrogenase de
Trypanosoma Cruzi e Mutantes D21OL, D21OI-G213D “

Beatriz Gomes Guimarães

Tese apresentada ao Instituto de Física de São Carlos, da Universidade de São Paulo,
para a obtenção do título de Doutor em Ciências “Física Aplicada”.

JK

Orientador: Prof. Dr. Glaucius Oliva

USP/IFSC/SBI



8-2-001234

São Carlos – São Paulo
1998

Meus documentos/declar-defesa

IFSC-107

RECEBIDA
INSCRIÇÃO

Guimarães, Beatriz Gomes

**Estudos Estruturais e Cinéticos da Enzima Gliceraldeído-3-Fosfato
Desidrogenase de *Trypanosoma cruzi* e Mutantes D210L, D210L-G213D
/Beatriz Gomes Guimarães. -- São Carlos, 1998.**

91p.

Tese (Doutorado) -- Instituto de Física de São Carlos, 1998.

Orientador: Prof. Dr. Glaucius Oliva

1. Cristalografia. 2. GAPDH de *T. cruzi*.
 3. Criocristalografia. 4. Desenho racional de drogas.
- I. Título.



**IFSC UNIVERSIDADE
DE SÃO PAULO**

Instituto de Física de São Carlos
E-mail: wladerez@if.sc.usp.br

Av. Dr. Carlos Botelho, 1465
CEP 13560-250 - São Carlos - SP
Brasil


Fone (016) 273-9333
Fax (016) 272-2218

MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE BEATRIZ GOMES GUIMARÃES APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO CARLOS, DA UNIVERSIDADE DE SÃO PAULO, EM 11 DE SETEMBRO DE 1998.

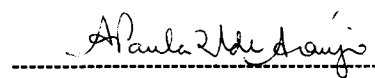
COMISSÃO JULGADORA:



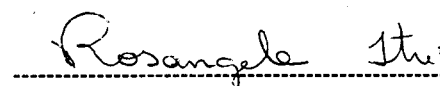
Prof. Dr. Glaucius Oliva/IFSC-USP



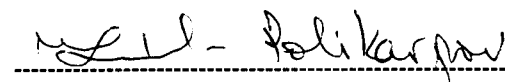
Prof. Dr. Richard Charles Garratt/IFSC-USP



Profa. Dra. Ana Paula Ulian de Araújo/IFSC-USP



Profa. Dra. Rosangela Itri /IF-USP



Dr. Igor Polikarpov/LNLS

*Aos meus pais, Caetano e Ignez e aos
meus irmãos, João Luís e José Eduardo,
pelo amor e incentivo.*

Agradecimentos

Ao Prof. Glaucius Oliva, pelos ensinamentos, pela segurança transmitida e amizade.

Aos Profs. Richard Garratt e Ana Paula Araújo, pela disponibilidade e auxílio constantes no decorrer de todo o trabalho e pela amizade.

Aos Profs. Eduardo Castellano e Odécio Cáceres, pelos ensinamentos.

Ao Prof. Luís Fernando Delboni, pela amizade e aos Profs. do Grupo de Biofísica, pela utilização de seus laboratórios.

Ao Prof. Flávio Henrique Silva e à Ana Maria Baião, da Universidade Federal de São Carlos, pelo trabalho cooperativo.

À Prof. Maria Goreti de Oliveira, da Universidade Federal de Viçosa, pelas discussões a respeito das medidas de cinética enzimática.

Ao Paulo, pelo carinho e apoio incondicionais, tornando tudo sempre muito mais fácil.

À Dulce, pelas discussões e auxílio na realização deste trabalho mas, principalmente, por sua grande amizade e de sua família, que hoje é também um pouco minha.

Aos meus amigos Bianca e Gerard, pela acolhida sincera e presença constante mesmo nos momentos mais difíceis.

À Cristina e ao Ezequiel, pela ajuda constante, troca de experiências e pela amizade tão importante durante esses anos.

Aos meus colegas André, Andressa, Carlos, Cecília, Fernando, Frank, Frederico, Gisele, Humberto, Jabah, João, Jorge, Juan, Marcos, Mônica, Paula, Paulão, Portezani, Regiane, Rodrigo, Rosa, Stefano e Valma, pela ótima convivência.

Aos meus amigos de Belo Horizonte, Ariete, Nivaldo, Carlos e Vado por estarem sempre comigo, mesmo à distância.

A Augusto, Geraldo, Maria Helena, Suzana, Wanda, e Isabel, pela boa vontade em ajudar.

Ao CNPq, pelo suporte financeiro.

Sumário

Lista de Figuras	i
Lista de Tabelas	iv
Lista de Abreviaturas	v
Resumo	vi
Abstract	vii
Prefácio	viii
Capítulo 1: Introdução	1
1.1 - A cristalografia de proteínas e o desenho racional de drogas	1
1.2 - A doença de Chagas	4
1.3 - A enzima gliceraldeído-3-fosfato desidrogenase	6
1.4 - Objetivos do trabalho	9
Capítulo 2: Expressão, Purificação e Cinética Enzimática	11
2.1 - Expressão	11
2.1.1 - Introdução	11
2.1.2 - Expressão da enzima GAPDH de <i>T. cruzi</i> tipo selvagem e mutantes	12
2.2 - Purificação	14
2.2.1 - Introdução	14
2.2.2 - Purificação da enzima GAPDH de <i>T. cruzi</i> tipo selvagem e mutantes	15
2.3 - Cinética enzimática	22
2.3.1 - Introdução	22
2.3.2 - Medidas de cinética enzimática de GAPDH de <i>T. cruzi</i> tipo selvagem e mutantes	24

Capítulo 3: Cristalização, Coleta e Processamento de Dados	29
3.1 - Cristalização	29
3.1.1 - Introdução	29
3.1.2 - Cristalização da enzima GAPDH tipo selvagem e mutante D210L-G213D	31
3.2 - Coleta e processamento dos dados de difração	34
3.2.1 - Introdução	34
3.2.2 - Coleta e processamento dos dados dos cristais de GAPDH tipo selvagem e mutante D210L-G213D	37
Capítulo 4: Resolução e Refinamento das Estruturas	44
4.1 - Resolução das estruturas	44
4.1.1 - Introdução	44
4.1.2 - Resolução das estruturas de GAPDH tipo selvagem e mutante D210L-G213D	48
4.2 - Refinamento das estruturas	50
4.2.1 - Introdução	50
4.2.2 - Refinamento das estruturas de GAPDH tipo selvagem e mutante D210L-G213D	56
4.3 - Avaliação da qualidade dos modelos	58
4.3.1 - Introdução	58
4.3.2 - Avaliação das estruturas de GAPDH tipo selvagem e mutante D210L-G213D	61
Capítulo 5: Análise das Estruturas	71
5.1 - A estrutura de GAPDH de <i>T. cruzi</i> a alta resolução	71
5.2 - Criocristalografia: efeito do resfriamento no empacotamento cristalino	75
5.3 - Mutações sítio dirigidas: papel do resíduo Arg249 no processo catalítico	81
Capítulo 6: Conclusões e Perspectivas	86
Referências Bibliográficas	88

Lista de Figuras

Figura 1.1 - Ciclo do desenho racional de drogas	3
Figura 1.2 - Ciclo de vida do <i>T. cruzi</i>	5
Figura 1.3 - Vista estereográfica da sobreposição dos sítios ativos de GAPDH de <i>T. cruzi</i> e <i>B. stearothermophilus</i>	8
Figura 2.1 - SDS-PAGE 15% - expressão e purificação da GAPDH de <i>T. cruzi</i> tipo selvagem	17
Figura 2.2 - SDS-PAGE 15% - tentativa de purificação do mutante GAPDH D210L	18
Figura 2.3 - SDS-PAGE 15% - nova tentativa de expressão e purificação do mutante GAPDH D210L	20
Figura 2.4 - SDS-PAGE 15% - expressão e purificação do mutante GAPDH D210L	21
Figura 2.5 - Curvas de Michaelis-Menten em função da concentração do substrato G3P	26
Figura 2.6 - Curvas de Michaelis-Menten em função da concentração do cofator NAD^+	27
Figura 3.1 - Cristais de GAPDH de <i>T. cruzi</i> tipo selvagem	32
Figura 3.2 - Cristais de GAPDH de <i>T. cruzi</i> crescidos em ambiente de microgravidade	33
Figura 3.3 - Cristais do mutante GAPDH D210L-G213D	34
Figura 3.4 - Padrão de difração de um cristal de GAPDH de <i>T. cruzi</i> crescido em ambiente de microgravidade	38
Figura 3.5 - Imagens de espalhamento de raios-X por uma solução não crioprotetora e uma solução crioprotetora	39
Figura 4.1 - Região de fitas β na estrutura de GAPDH tipo selvagem, coleta de dados a 277 K	61
Figura 4.2 - Região do sítio de ligação do cofator NAD^+ na estrutura de GAPDH tipo selvagem, coleta de dados a 100 K	62

Figura 4.3 - Região das mutações na estrutura de GAPDH D210L-G213D	62
Figura 4.4 - Diagramas de Ramachandran para um dos monômeros das estruturas de GAPDH tipo selvagem, coletas de dados a 277 e 100 K e mutante D210L-G213D	64
Figura 4.5 - Perfil 3D-1D para a estrutura de GAPDH tipo selvagem, coleta de dados a 277 K	65
Figura 4.6 - Perfil 3D-1D para a estrutura de GAPDH tipo selvagem, coleta de dados a 100 K	66
Figura 4.7 - Perfil 3D-1D para a estrutura do mutante GAPDH D210L-G213D	66
Figura 4.8 - Índice de qualidade de contatos para a estrutura de GAPDH tipo selvagem, coleta de dados a 277 K	67
Figura 4.9 - Índice de qualidade de contatos para a estrutura de GAPDH tipo selvagem, coleta de dados a 100 K	68
Figura 4.10 - Índice de qualidade de contatos para a estrutura do mutante GAPDH D210L-G213D	68
Figura 4.11 - Distribuição dos fatores de temperatura (fatores B) para o monômero A das estruturas de GAPDH tipo selvagem, coletas de dados a 277 e 100 K e mutante D210L-G213D	69
Figura 5.1 - Tetrâmero de GAPDH de <i>T. cruzi</i>	72
Figura 5.2 - Monômero de GAPDH de <i>T. cruzi</i>	72
Figura 5.3 - Perfil de distâncias médias entre os resíduos de um monômero das estruturas sobrepostas de GAPDH obtidas a 2.8 e 2.15 Å de resolução	74
Figura 5.4 - Interação entre os resíduos Asp334 e Asp254	75
Figura 5.5 - Sobreposição das celas unitárias e respectivos conteúdos da unidade assimétrica (um tetrâmero) referentes às coletas de dados a 277 e 100 K	76
Figura 5.6 - Traços de C _α dos tetrâmeros das estruturas a 277 e 100 K, com seus monômeros A sobrepostos	77
Figura 5.7 - Perfil de distâncias médias entre as cadeias laterais ao longo dos quatro monômeros das estruturas a 277 e 100 K sobrepostas	78

Figura 5.8 - Sobreposição dos resíduos da região do sítio ativo das estruturas de GAPDH referentes às coletas de dados a 277 e 100 K	80
Figura 5.9 - Mecanismo enzimático da GAPDH	82
Figura 5.10 - Conformação dos resíduos do sítio ativo da enzima GAPDH de <i>T. cruzi</i>	84
Figura 5.11 - Sobreposição dos resíduos da região do sítio ativo da enzima GAPDH tipo selvagem e do mutante D210L-G213D	85

Lista de Tabelas

Tabela 2.1 - Valores das constantes de Michaelis (K_M), catalítica (k_{cat}) e da razão k_{cat}/K_M para enzima GAPDH tipo selvagem e mutantes	28
Tabela 3.1 - Parâmetros das coletas de dados e das redes cristalinas obtidos para GAPDH tipo selvagem e mutante D210L-G213D	40
Tabela 3.2 - Estatísticas dos dados coletados para os cristais de GAPDH tipo selvagem e mutante D210L-G213D	41
Tabela 3.3 - Conteúdo das celas unitárias dos cristais de GAPDH tipo selvagem e mutante D210L-G213D	43
Tabela 4.1 - Parâmetros de rotação e translação e fatores de correlação obtidos após o refinamento de corpo rígido por AMoRe	49
Tabela 4.2 - Parâmetros finais dos refinamentos	58

Lista de Abreviaturas

ATP	Adenosina trifosfato
DTT	Ditiotreitol
EDTA	Ácido etilenodiaminotetracético
IPTG	Isopropil- β -D tiogalactopiranosídeo
MPD	2-metil-2,4-pentanodiol
NAD ⁺	Nicotinamida adenina dinucleotídeo, forma oxidada
NADH	Nicotinamida adenina dinucleotídeo, forma reduzida
PEG	Polietilenoglicol
PMSF	Fluoreto de α -tolueno sulfonil
SDS	Duodecil sulfato de sódio
SDS-PAGE	Eletroforese em gel de poliacrilamida-SDS
TEA	Trietanolamina
Tris	Tris(hidroximetil)-aminometano

Resumo

A enzima glicossomal gliceraldeído-3-fosfato desidrogenase (GAPDH) de *Trypanosoma cruzi* e os mutantes D210L e D210L-G213D foram expressos em *E. coli*, purificados e submetidos a ensaios de cinética enzimática e de cristalização. A enzima GAPDH tipo selvagem e o mutante D210L-G213D cristalizaram-se no grupo espacial $P2_1$ e os cristais apresentaram padrões de difração de raios-X de boa qualidade. A estrutura cristalográfica da enzima tipo selvagem foi determinada a 2.5 e 2.15 Å de resolução, a partir de coletas de dados realizadas a 277 e 100 K respectivamente. Os fatores R cristalográficos finais dos refinamentos foram de 16.0% para a estrutura a 277 K e 18.8% para a estrutura a 100 K. A estrutura do mutante GAPDH D210L-G213D foi determinada a 2.15 Å de resolução e refinada até um fator R cristalográfico de 18.9%.

A comparação entre as estruturas da enzima tipo selvagem determinadas nas duas temperaturas levou a resultados interessantes no que diz respeito ao empacotamento cristalino. O resfriamento dos cristais provocou uma redução no volume da cela unitária de 10.5%, tendo a maior variação ocorrido no parâmetro de rede a (14.5%). A sobreposição das celas unitárias mostrou uma rotação do conteúdo da unidade assimétrica de cerca de 5 graus em torno de um eixo aproximadamente paralelo a b . Por outro lado, a análise das estruturas da enzima tipo selvagem e mutante, juntamente com os parâmetros cinéticos, permitiram a discussão a respeito de alguns detalhes do mecanismo catalítico da enzima, principalmente no que se refere ao papel do resíduo Arg249. Tal resíduo, que apresenta grande mobilidade conformacional de sua cadeia lateral, parece estar envolvido na etapa de reorientação de um dos intermediários durante o processo catalítico.

Abstract

The glycosomal glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from *Trypanosoma cruzi* and its mutants D210L and D210L-G213D were expressed in *E. coli* and purified, followed by kinetic and crystallization assays. Both wild type enzyme and D210L-G213D mutant were crystallized in $P2_1$ space group and the crystals presented good X-ray diffraction patterns. The three-dimensional structure of the wild type enzyme was determined at 2.5 and 2.15 Å resolution from data collected at 277 and 100 K respectively. The structures were refined to a crystallographic R-factor of 16.0% for data collected at 277 K and 18.8% for those collected at 100 K. Also, the structure of the D210L-G213D mutant was solved at 2.15 Å resolution and refined to a crystallographic R-factor of 18.9% with good geometry indicators.

Comparison between the wild type enzyme structures solved at both temperatures led to interesting results concerning the crystal packing. Flash-cooled crystals presented a 10.5% shrink in the unit cell volume being the major reduction observed in the parameter a (14.5%). Superposition of the unit cells showed a global rotation of the asymmetric unit content of about 5 degrees around an axis approximately parallel to b . On the other hand, the analysis of the wild type and mutant enzyme structures, together with the kinetic parameters, allowed a discussion about some catalytic mechanism details, mainly the role of the Arg249 residue. The results showed that this residue might be involved in the reorientation of one of the intermediates during the catalytic process.

Prefácio

As proteínas compõem um extenso grupo de macromoléculas que desempenham as mais diversas funções nos organismos vivos: transporte e armazenamento, catálise de reações, regulação de atividades biológicas, participam do sistema imunológico dos vertebrados, apresentam funções mecânicas como na contração muscular, entre outras. As propriedades e funções das proteínas estão estreitamente relacionadas ao arranjo espacial de seus átomos, daí a extrema importância em conhecer-se suas estruturas tridimensionais a nível atômico.

Existem duas técnicas comumente utilizadas para a determinação da estrutura de macromoléculas a nível atômico, a difração de raios-X em monocristais e a ressonância magnética nuclear. Entre estas, a difração de raios-X é a que tem sido mais largamente aplicada, já que a segunda restringe-se a macromoléculas de peso molecular inferior a cerca de 30 kDa. Os trabalhos pioneiros na resolução da estrutura tridimensional de proteínas foram realizados por Kendrew e Perutz nos anos 50, quando determinaram as estruturas cristalográficas da mioglobina e hemoglobina. Tais trabalhos renderam-lhes o prêmio Nobel em 1962. A partir de então, observa-se um aumento exponencial no número de estruturas tridimensionais de proteínas resolvidas, número que ultrapassa os 7 000 atualmente.

A crescente utilização da técnica de difração de raios-X para determinação da estrutura tridimensional de proteínas pode ser atribuída a diversos fatores. O desenvolvimento da tecnologia computacional, fornecendo máquinas com capacidade de cálculos e capacidade gráfica cada vez mais avançadas; as técnicas de biologia molecular, que permitem a obtenção em grande quantidade de, teoricamente, qualquer proteína; a construção das fontes de radiação síncrotron, produzindo feixes de raios-X de alta intensidade e o desenvolvimento da técnica de resfriamento dos cristais a temperaturas criogênicas (criocristalografia), que permite a coleta de dados de cristais muito sensíveis à radiação. Além desses fatores, que possibilitaram a utilização da técnica de cristalografia

de raios-X com maior eficácia, o surgimento de novos métodos biotecnológicos, como a mutagênese sítio dirigida e o desenho de drogas baseado em estruturas, contribuíram para o crescimento do interesse na determinação da estrutura tridimensional de um número cada vez maior de proteínas.

Como parte de um projeto amplo que vem sendo desenvolvido no Grupo de Cristalografia de Proteínas e Biologia Estrutural do IFSC-USP, visando o desenvolvimento de uma nova droga contra a doença de Chagas, o presente trabalho aborda aspectos bioquímicos e estruturais da enzima glicossomal gliceraldeído-3-fosfato desidrogenase (GAPDH) de *Trypanosoma cruzi*. As estruturas da enzima GAPDH tipo selvagem e do mutante D210L-G213D foram determinadas a alta resolução pela técnica de difração de raios-X em monocristais.

O processo para se chegar à determinação da estrutura cristalográfica de uma proteína requer a utilização de técnicas de diversas áreas como biologia molecular, bioquímica e cristalografia propriamente dita. Em vista disso, a apresentação deste trabalho será feita de forma que os capítulos que contenham as descrições dos métodos incluam os resultados relacionados, já que tratarão de técnicas tão diversas. No capítulo 1, será feita uma introdução aos temas relacionados ao trabalho: cristalografia de proteínas, a técnica de desenho de drogas baseado em estruturas e a doença de Chagas. O capítulo abordará ainda aspectos importantes das estruturas das enzimas GAPDH de outras espécies e termina com o detalhamento dos objetivos do trabalho. O capítulo 2 tratará dos procedimentos e resultados da expressão, purificação e medidas de atividade da enzima GAPDH tipo selvagem e mutantes D210L e D210L-G213D. No capítulo 3 serão apresentados os métodos de cristalização, coleta e processamento dos dados de difração, bem como seus resultados. O quarto capítulo apresentará os métodos de resolução e refinamento de uma estrutura cristalográfica e suas aplicações ao presente caso. A qualidade dos modelos finais obtidos será também avaliada neste capítulo. Concluído o trabalho de determinação estrutural, os modelos serão analisados no capítulo 5, abordando aspectos como empacotamento cristalino, conformação dos resíduos do sítio ativo e potenciais sítios de ligação de inibidores. Finalmente, o capítulo 6 apresentará uma conclusão geral do trabalho e as perspectivas para sua continuidade.

Capítulo 1

Introdução

A cristalografia tem assumido um papel crescente nas ciências modernas, tendo seu desenvolvimento estimulado pelo avanço de várias áreas do conhecimento como Física, Química, Biologia e Ciência dos Materiais e, por outro lado, contribuindo para esse avanço. O caráter multidisciplinar da cristalografia se faz notar pela sua aplicação no estudo de diversos problemas biológicos. No presente capítulo serão abordados os temas de interesse deste trabalho, a cristalografia de proteínas e o desenho racional de drogas, bem como os problemas alvo, a doença de Chagas e a enzima gliceraldeído-3-fosfato desidrogenase. Serão também detalhados os objetivos do trabalho.

1.1 - A cristalografia de proteínas e o desenho racional de drogas

A cristalografia permite o estudo da estrutura tridimensional de moléculas a nível atômico, através da difração de raios-X em monocristais e, dentro dessa área, a cristalografia de proteínas especializa-se no estudo de macromoléculas biológicas. A utilização dos raios-X para a determinação da estrutura de uma molécula deve-se ao fato de que tal radiação possui comprimento de onda da ordem das distâncias interatômicas (em torno de 1.5 Å). A teoria do espalhamento de raios-X pela matéria e, mais especificamente, sua difração por cristais estão amplamente descritas em diversos livros texto e não serão detalhadas aqui (Blundell; Johnson, 1976; Giacovazzo *et al.*, 1992; Drenth, 1994).

A determinação da estrutura tridimensional de uma proteína por cristalografia de raios-X constitui-se em um processo trabalhoso, cujas várias etapas serão discutidas ao longo desse trabalho e são sumarizadas a seguir:

- 1) purificação da proteína, partindo de sua extração do organismo original ou da expressão recombinante;
- 2) cristalização, uma das etapas mais críticas do processo;
- 3) coleta e processamento dos dados de difração;
- 4) determinação das fases e construção de um modelo inicial da molécula;
- 5) interpretação dos mapas de densidade eletrônica e refinamento do modelo;
- 6) avaliação da qualidade do modelo.

O importante papel que a cristalografia vem assumindo no estudo dos mais variados problemas biológicos deriva da estreita relação existente entre a estrutura e a função das proteínas. Muitos processos celulares têm sido entendidos a partir do conhecimento das estruturas tridimensionais de proteínas envolvidas. Mais recentemente, a técnica denominada desenho de drogas baseado em estruturas vem utilizando os métodos cristalográficos na busca de novas drogas contra doenças como AIDS, doença de Chagas, malária, tuberculose, entre outras.

O processo de desenvolvimento de uma nova droga contra determinada doença começa com a busca de compostos de partida que apresentem significativa atividade inibitória contra o agente patogênico causador da doença em questão. Tal busca pode ser aleatória, através de bancos de dados contendo um número extraordinariamente grande de compostos sintéticos, naturais, meios microbiológicos e extratos de plantas ou dirigida, através do desenho de inibidores específicos contra proteínas alvo escolhidas. Após essa primeira etapa, um longo caminho ainda deve ser percorrido até que se chegue a uma droga comercialmente viável. O composto de partida encontrado pode ser muito tóxico, instável em solução, rapidamente metabolizado pelo organismo humano, difícil de sintetizar, pode apresentar alto custo de produção, entre outros problemas. Dessa forma, tal inibidor deverá sofrer várias modificações visando o melhoramento de suas propriedades farmacológicas, até que possa ser submetido a testes clínicos e, eventualmente, dar origem a uma nova droga (Verlinde; Hol, 1994).

A técnica denominada desenho de drogas baseado em estruturas fundamenta-se no conhecimento da estrutura tridimensional de uma proteína alvo, considerada vital para o agente causador da doença, como ponto de partida para o desenho de um inibidor. Além disso, a determinação da estrutura tridimensional do complexo

proteína-ligante é de grande importância para a otimização desse composto de partida, já que o conhecimento, a nível atômico, do modo de ligação do composto torna mais fácil sua modificação de uma maneira racional. A figura 1.1 mostra um desenho esquemático do ciclo do desenho racional de drogas.

Nos últimos anos, houve um aumento exponencial na utilização dessa técnica, tanto a nível acadêmico quanto industrial, para o desenvolvimento de drogas contra diversas doenças. Entre os casos bem sucedidos, o de maior destaque foi o desenvolvimento do coquetel de drogas contra a AIDS, cuja proteína alvo utilizada foi uma protease do vírus HIV (Lin, 1997).

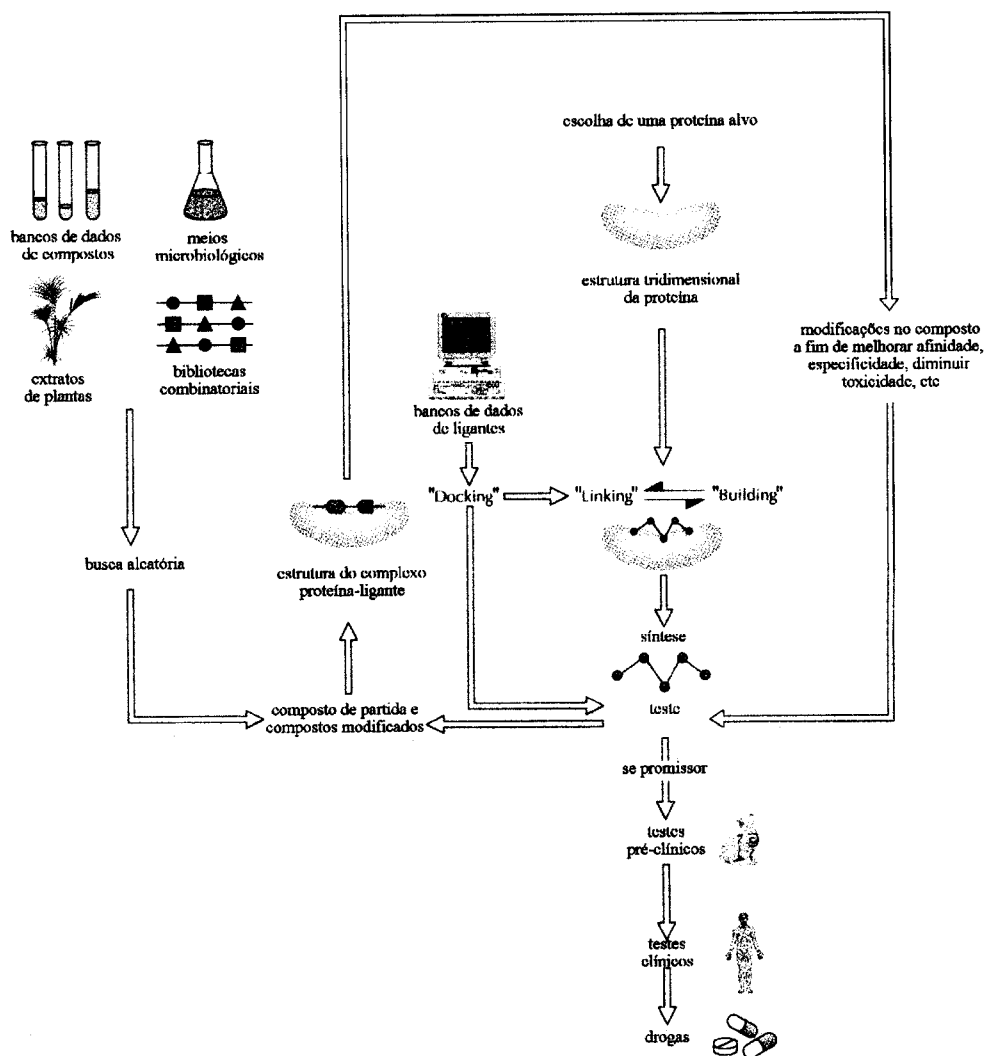


Figura 1.1 - Ciclo do desenho racional de drogas (figura modificada de Verlinde; Hol, 1994).

1.2 - A doença de Chagas

A doença de Chagas constitui ainda hoje um grave problema médico-social no Brasil e em diversos países das Américas do Sul e Central, manifestando-se também em áreas metropolitanas dos Estados Unidos. Segundo a Organização Mundial de Saúde, a doença afeta aproximadamente 18 milhões de pessoas, causando cerca de 45 000 mortes por ano (World Health Organization Statistical Information System Website). O *Trypanosoma cruzi* é o agente causador da doença, descoberta e caracterizada pelo médico Carlos Chagas em 1909.

O mecanismo clássico de transmissão da doença ocorre através dos vetores, insetos da subfamília Triatominae conhecidos popularmente como "barbeiros". A infecção ocorre pela penetração dos parasitas, eliminados nas fezes do inseto durante a picada, em células da pele ou mucosas. Desde que tem sido feito o controle do vetor inseto em áreas endêmicas, outros mecanismos de transmissão como transfusão sanguínea e transmissão congênita tornaram-se significativos.

O *T. cruzi* possui em seu ciclo biológico três formas principais: amastigota, forma esférica ou oval, sem flagelo livre; epimastigota, forma oval alongada com pequena membrana ondulante e flagelo livre; tripomastigota, forma alongada com membrana ondulante e flagelo livre. O ciclo biológico do parasita (figura 1.2) apresenta uma fase no hospedeiro vertebrado (homem e mamíferos pertencentes a 7 ordens diferentes) e outra no inseto vetor (triatomíneos). No início da infecção do vertebrado (fase aguda), a parasitemia é mais elevada, podendo ocorrer a morte do hospedeiro. Na espécie humana, a mortalidade nessa fase da infecção ocorre principalmente em crianças. Quando o hospedeiro desenvolve resposta imune eficaz, a infecção tende a tornar-se crônica, quando o número de parasitas é pequeno na circulação. Na espécie humana, a evolução da doença de Chagas ocorre lentamente. Após a fase aguda, caracterizada por manifestações locais e sintomas como febre, cefaléia, aumento no tamanho do coração e fígado, os sobreviventes passam por um longo período assintomático (10 a 30 anos). Esta é a chamada fase crônica da doença que, em seu estágio final, provoca deficiência no sistema cardiocirculatório e intenso processo inflamatório nos órgãos do sistema digestivo, levando, em muitos casos, o paciente à morte (Neves, 1991).

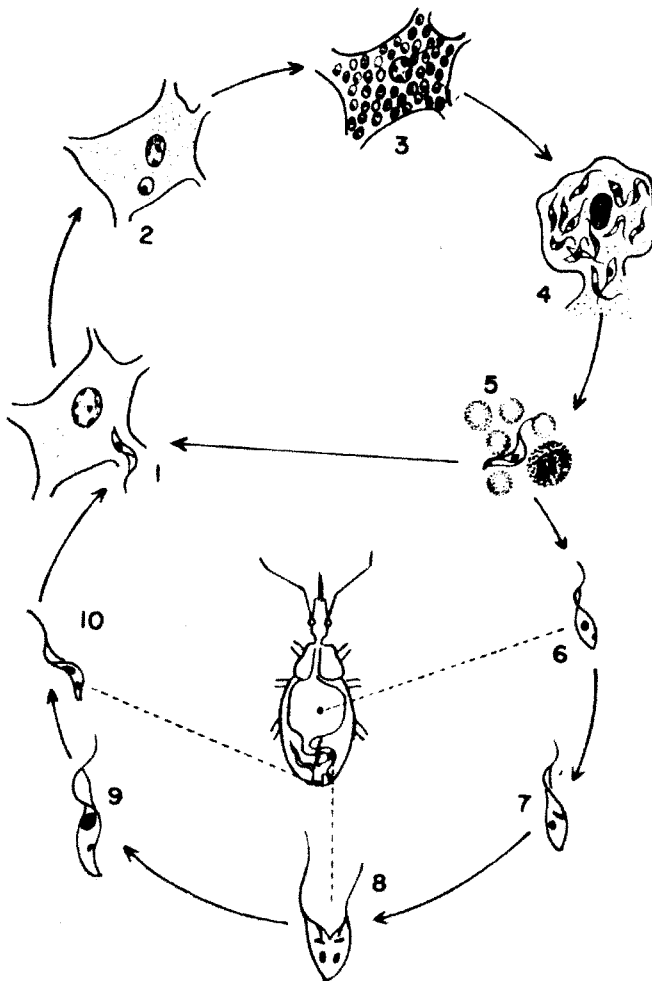


Figura 1.2 - Ciclo de vida do *T. cruzi* (figura extraída de Neves, 1991; cap. 10); 1) penetração tripomastigota em uma célula; 2) transformação do tripomastigota em amastigota; 3) multiplicação por divisão binária e nova diferenciação em tripomastigota; 4) rompimento da célula parasitada liberando tripomastigotas na corrente sanguínea; 5) forma tripomastigota no sangue circulante pode penetrar em outra célula (1) ou ser ingerida pelo triatomíneo (6); 6) forma tripomastigota no estômago do triatomíneo; 7) transformação da forma tripomastigota em epimastigota no intestino posterior do inseto; 8) multiplicação por divisão binária; 9) transformação da forma epimastigota em tripomastigota no reto do inseto; 10) forma tripomastigota, nas fezes do triatomíneo, apta a penetrar no hospedeiro mamífero.

Desde a descoberta da doença, em 1909, até os dias atuais, inúmeras tentativas de tratamento foram realizadas, sem obter-se, entretanto, um medicamento que fosse totalmente eficaz. Muitas pesquisas têm sido realizadas no sentido de desenvolver-se uma nova droga contra a doença de Chagas e, neste contexto, insere-se o presente trabalho.

1.3 - A enzima gliceraldeído-3-fosfato desidrogenase

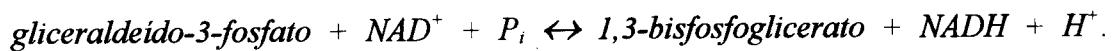
A glicólise é uma via metabólica que consiste de uma seqüência de dez reações catalisadas por enzimas, sendo responsável pela produção de uma parcela significativa da energia utilizada por muitos organismos e pelo processamento da glicose e outros carboidratos para entrada nas demais vias metabólicas. A glicólise converte glicose em piruvato com a liberação de 2 moles de ATP por mol de glicose. Posteriormente, sob condições anaeróbicas, pode ocorrer a fermentação alcoólica ou láctea do piruvato, enquanto sob condições aeróbicas, o piruvato é oxidado via ciclo do ácido tricarboxílico (ou ciclo de Krebs) seguido pela cadeia respiratória (Voet; Voet, 1995).

Várias características bioquímicas peculiares são encontradas nos protozoários da família Trypanosomatidae (Opperdoes, 1985). Em 1977, Opperdoes e Borst descobriram, na forma sanguínea da espécie *Trypanosoma brucei*, a presença de organelas especiais de 0.2 a 0.3 μm de diâmetro, envolvidas por uma membrana fosfolipídica. Tais organelas, denominadas glicossomos, contêm as sete primeiras enzimas da via glicolítica e duas da via de metabolização do glicerol. A compartimentalização dessas enzimas é responsável por um grande aumento na taxa da glicólise e a presença dos glicossomos é observada em todos os organismos da ordem Kinetoplastidae, da qual fazem parte as espécies do gênero *Trypanosoma*. Estudos em *T. brucei* mostraram ainda que a forma sanguínea (tripomastigota) do parasita apresenta a mitocôndria muito reduzida e os ciclos do ácido tricarboxílico e cadeia respiratória suprimidos, sendo o parasita dependente exclusivamente da glicólise como fonte de energia. Esse fato sugere que enzimas da via glicolítica podem constituir-se em bons alvos para o desenho de drogas contra *Trypanosoma*.

As estruturas tridimensionais de três enzimas da via glicolítica do *T. brucei*, causador da doença do sono, foram determinadas como pontos de partida para o desenho racional de drogas: triosefosfato isomerase (Wierenga *et al.*, 1991), gliceraldeído-3-fosfato desidrogenase (Vellieux *et al.*, 1993) e fosfoglicerato quinase (Bernstein *et al.*, 1997). Uma dessas proteínas, a gliceraldeído-3-fosfato desidrogenase, mostra potenciais sítios para ligação de inibidores, apresentando diferenças estruturais significativas quando comparada com a enzima homóloga humana. Alguns inibidores foram desenhados, sintetizados e testados, apresentando

resultados promissores (Verlinde *et al.*, 1994). Tais resultados estimularam o estudo estrutural da enzima homóloga de *T. cruzi*, como proteína alvo para o desenho de drogas contra a doença de Chagas.

A enzima gliceraldeído-3-fosfato desidrogenase (GAPDH) é um tetrâmero com subunidades de aproximadamente 39 kDa e catalisa a reação de fosforilação do gliceraldeído-3-fosfato em 1,3-bisfosfoglicerato, na presença de NAD^+ e fosfato inorgânico:



O mecanismo de reação da enzima tem sido extensivamente estudado (Trentham, 1971; Duggleby; Dennis, 1974; Harris; Waters, 1976; Soukri *et al.*, 1989) e dois resíduos mostraram-se essenciais para o processo catalítico, uma cisteína e uma histidina (Cys166 e His194 em GAPDH de *T. cruzi*).

As estruturas cristalográficas das enzimas GAPDH de várias espécies foram determinadas, entre elas: humana (Watson *et al.*, 1972), *Bacillus stearothermophilus* (Skarzynski *et al.*, 1987), *Trypanosoma brucei* (Vellieux *et al.*, 1993), *Thermotoga maritima* (Korndörfer *et al.*, 1995), *Leishmania mexicana* (Kim *et al.*, 1995) e *Escherichia coli* (Dueé *et al.*, 1996). Algumas dessas estruturas apresentaram dois íons sulfato por monômero, ligados ao sítio catalítico. Acreditava-se que esses íons estariam ocupando as posições dos dois íons fosfato envolvidos na reação, o fosfato do substrato (P_s) e o fosfato inorgânico (P_i), que é adicionado ao substrato durante a reação. A estrutura da GAPDH de *L. mexicana* apresentou dois íons fosfato, mostrando que o sítio de ligação do fosfato do substrato é praticamente o mesmo que o do íon sulfato correspondente, enquanto, para o fosfato inorgânico, foi observado um deslocamento de 2.9 Å em relação ao sítio de ligação do íon sulfato.

Recentemente, a estrutura da enzima GAPDH de *T. cruzi* foi determinada a 2.8 Å de resolução sem a presença de quaisquer íons no sítio catalítico (Souza, 1996; Souza *et al.*, 1998). Observou-se uma grande mudança conformacional do resíduo Arg249, cujos homólogos em outras estruturas interagem com os íons sulfato/fosfato do substrato. Quando comparada com a estrutura de *B. stearothermophilus* (determinada a 1.8 Å de resolução), a mudança resultou em um deslocamento de 5.2 Å na posição do C_ξ , estando o referido resíduo, na ausência de íons sulfato/fosfato,

formando uma ponte salina com o Asp210. Souza e colaboradores (1998) propuseram que o resíduo Arg249 poderia funcionar como uma "porta", regulando a ligação do substrato ao sítio ativo e a liberação do produto após a reação. A conformação "aberta" desse resíduo, observada na estrutura de GAPDH de *T. cruzi*, corresponderia à situação em que o sítio ativo está vazio, antes da entrada do substrato. A figura 1.3 compara a conformação dos resíduos do sítio ativo das estruturas de GAPDH de *T. cruzi* e *B. stearothermophilus*, mostrando a mudança observada para a Arg249. Reforçando a proposição acima, a estrutura de GAPDH de *E. coli*, que também não possui íons ligados ao sítio ativo, apresenta seu resíduo Arg231 (equivalente à Arg249 em *T. cruzi*) em duas conformações possíveis em um dos monômeros da unidade assimétrica, correspondendo às conformações "aberta" e "fechada". Esse fato indica uma alta flexibilidade da cadeia lateral desse resíduo.

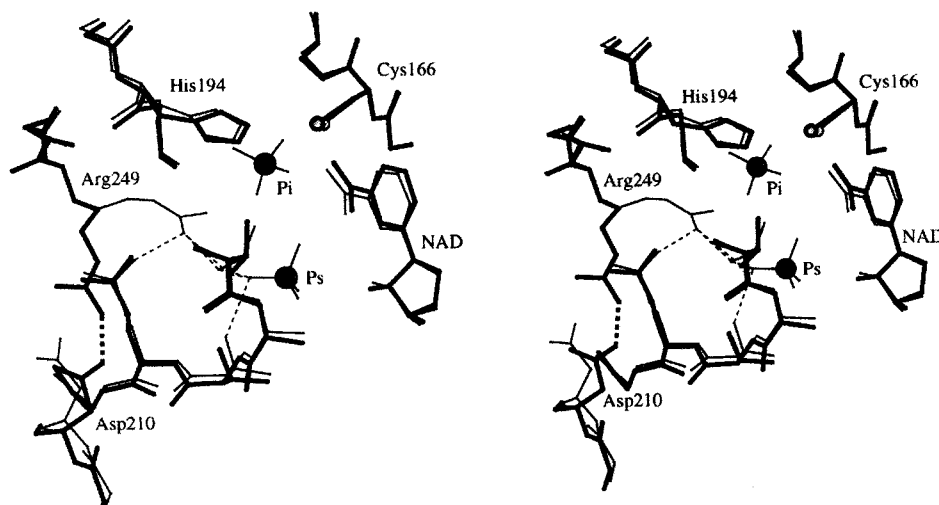


Figura 1.3 - Vista estereográfica da sobreposição dos sítios ativos de GAPDH de *T. cruzi* (em negrito) e *B. stearothermophilus*, mostrando a mudança conformacional do resíduo Arg249. Os íons sulfato encontrados na estrutura de *B. stearothermophilus* correspondem aos íons fosfato inorgânico (P_i) e fosfato do substrato (P_s) (figura extraída de Souza *et al.*, 1998).

A análise das estruturas tridimensionais de várias outras GAPDHs mostra a conservação de um resíduo ácido aspártico na posição do Asp210 em *T. cruzi*. Entretanto, na enzima humana, tal resíduo é substituído por uma leucina juntamente com uma outra substituição, Gly213 em *T. cruzi* por um ácido aspártico em GAPDH

humana (Asp197). Como consequência, um resíduo ácido aspártico é mantido aproximadamente na mesma posição que em *T. cruzi*, sugerindo que Asp197 em humana desempenha o mesmo papel que Asp210 em *T. cruzi*, fazendo uma ponte salina com a arginina na ausência do substrato.

A região do sítio catalítico da enzima GAPDH ainda não tinha sido explorada para o desenho de inibidores, principalmente devido à alta conservação na seqüência e conformação de seus resíduos, observada nas várias espécies. As observações feitas anteriormente indicaram um novo sítio potencial para ligação de inibidores específicos. A substituição do resíduo Asp210 em *T. cruzi* (cadeia lateral carregada) por Leu194 em humana (cadeia lateral neutra) sugere que compostos que ocupem parcialmente o sítio ativo e ao mesmo tempo participem de interações polares com a cadeia lateral de Asp210 poderiam ligar-se seletivamente à enzima do parasita (Souza, 1996; Souza *et al.*, 1998).

1.4 - Objetivos do trabalho

O trabalho aqui apresentado faz parte de um projeto amplo que vem sendo desenvolvido no Grupo de Cristalografia de Proteínas e Biologia Estrutural do IFSC-USP, visando o desenvolvimento de uma nova droga contra a doença de Chagas. Como mencionado anteriormente, a proteína alvo escolhida para o desenho de inibidores específicos foi a GAPDH glicossomal de *T. cruzi*, que teve sua estrutura determinada a 2.8 Å de resolução e apresentada em outra tese de doutoramento do grupo (Souza, 1996; Souza *et al.*, 1998). A fim de chegar-se a conclusões mais precisas sobre o papel proposto para o resíduo Arg249 no mecanismo de entrada e saída do substrato, fez-se o planejamento de dois mutantes da enzima. Um deles, substituindo o resíduo carregado Asp210 por uma leucina e o segundo, reproduzindo a situação encontrada em GAPDH humana, onde a primeira mutação é mantida mas introduz-se um ácido aspártico no lugar da Gly213. Por outro lado, a aplicação da técnica de desenho de drogas baseado em estruturas é mais confiável quando se utiliza estruturas tridimensionais determinadas a alta resolução, já que, nesse caso, maior detalhe atômico é obtido.

Dessa forma, o presente trabalho tem como objetivos a determinação da estrutura tridimensional da enzima GAPDH de *T. cruzi* a mais alta resolução e dos mutantes D210L e D210L-G213D. Além disso, através de medidas de cinética enzimática da enzima tipo selvagem e dos mutantes e sua relação com as estruturas, procurar-se-á tirar conclusões a respeito do papel de alguns resíduos na catálise enzimática.

Capítulo 2

Expressão, Purificação e Cinética Enzimática

A técnica de DNA recombinante permite a obtenção em grande quantidade de, teoricamente, qualquer proteína, constituindo-se em uma importante ferramenta para a cristalografia. Além disso, o sucesso dos experimentos de cristalização de determinada proteína depende da obtenção da mesma com alto grau de pureza. Neste capítulo são descritas a expressão e purificação da enzima GAPDH de *T. cruzi* tipo selvagem e mutantes D210L-G213D e D210L, bem como suas caracterizações através de medidas de cinética enzimática.

2.1 - Expressão

2.1.1 - Introdução

A tecnologia de DNA recombinante tem proporcionado, desde o seu surgimento no início dos anos 70, um avanço extraordinário no entendimento dos mais diversos processos biológicos (Watson *et al.*, 1992). A capacidade dos cientistas em isolar, analisar e manipular genes revolucionou a pesquisa não apenas em biologia fundamental, mas também abriu uma nova área dedicada à clonagem e produção de proteínas de interesse médico e industrial.

No contexto da cristalografia de proteínas, a técnica de DNA recombinante apresenta-se como uma importante ferramenta, já que permite a obtenção em grande quantidade de, teoricamente, qualquer proteína. O processo de cristalização de determinada proteína requer a disponibilidade de uma quantidade significativa da mesma

(da ordem de dezena de miligramas), o que pode ser inviável quando a extração da proteína tem que ser feita a partir de sua fonte natural.

O processo para produção de uma proteína recombinante envolve a clonagem de seu gene e a introdução desse gene em uma célula hospedeira, que será responsável pela síntese ou expressão da proteína de interesse. Em poucas palavras, clonagem significa a separação de um gene específico ou segmento de DNA do seu cromossomo, a introdução desse gene em uma molécula carregadora de DNA (vetor) e a replicação do DNA em milhares de vezes. O organismo mais largamente utilizado para expressão de proteínas recombinantes é a bactéria *E. coli*, embora outros sistemas de expressão têm sido relatados (*B. subtilis*, levedura, células de insetos ou de mamíferos). Os chamados vetores de expressão são responsáveis pela introdução do material genético exógeno na célula hospedeira e regulação da expressão da proteína de interesse (Watson *et al.*, 1992).

O desenvolvimento de sistemas de expressão tem se constituído em uma importante área de pesquisa industrial e acadêmica e tem sido tão eficiente que, em alguns casos, a proteína exógena expressa pode representar até 50% do conteúdo celular protéico (pET System Manual).

2.1.2 - Expressão da enzima GAPDH de *T. cruzi* tipo selvagem e mutantes

O gene que codifica a enzima GAPDH de *T. cruzi* (TcGAP) foi clonado em um plasmídeo pET3a (Hannaert *et al.*, 1995). Tal trabalho foi desenvolvido no Christian de Duve Institute of Cellular Pathology, Bélgica, e o plasmídeo foi gentilmente cedido por Hannaert e colaboradores para o Grupo de Cristalografia de Proteínas do IFSC - USP. Os genes que codificam as enzimas mutantes GAPDH D210L e GAPDH D210L-G213D foram obtidos através da técnica de mutação sítio dirigida por sobreposição e extensão utilizando reação em cadeia de polimerização (PCR) (Ho *et al.*, 1989). Este trabalho foi desenvolvido como colaboração pelo Prof. Dr. Flávio Henrique Silva do Departamento de Genética e Evolução da Universidade Federal de São Carlos.

Células da linhagem BL21(DE3)gap⁻ da bactéria *E. coli* foram transformadas pelos plasmídeos contendo os genes da GAPDH tipo selvagem e mutantes D210L-G213D e D210L. Tal linhagem de bactéria, gentilmente cedida por Hannaert e colaboradores, havia sido manipulada geneticamente de modo a não expressar a enzima GAPDH endógena. Os procedimentos para preparação de células competentes e transformação bacteriana seguiram protocolos amplamente utilizados (Ausubel *et al.*, 1995). A obtenção de colônias transformantes do clone da enzima mutante GAPDH D210L, ao contrário das demais, exigiu a repetição do procedimento de transformação por diversas vezes. As células de *E. coli* transformadas foram armazenadas em 25% de glicerol a -80 °C.

A expressão da enzima GAPDH tipo selvagem e mutantes seguiu o mesmo protocolo. Uma única colônia de *E. coli* BL21(DE3)gap⁻ transformada pelo plasmídeo que contém o gene de interesse foi inoculada em 5 mL de meio LB suplementado (1% de triptona, 0.5% de extrato de levedura, 1% de NaCl, 1% de glicerol, 0.4% de ácido málico e 0.1% de casaminoácidos) adicionado de 100 µg/mL de ampicilina. Após incubação por cerca de 12 horas a 37 °C sob agitação, essa cultura foi utilizada para inocular 500 mL de meio LBS adicionado de 100 µg/mL de ampicilina, permanecendo sob agitação a 37 °C. Quando a cultura atingiu uma densidade ótica a 600 nm de cerca de 0.8, a expressão da GAPDH foi induzida pela adição de 0.4 mM de IPTG. A temperatura de incubação foi então reduzida para 20 °C, permanecendo por um período de aproximadamente 14 horas, quando as células foram separadas por centrifugação a 2 000 × g por 20 minutos a 4 °C.

2.2 - Purificação

2.2.1 - Introdução

O estudo de uma determinada proteína, seja no que diz respeito à sua caracterização bioquímica ou do ponto de vista estrutural, requer a obtenção de tal proteína livre de contaminantes. Por exemplo, estudos detalhados de como uma enzima catalisa determinada reação só serão efetivos se outras enzimas ou substâncias não estiverem presentes ou, o sucesso de experimentos de cristalização de proteínas depende, entre outros fatores, de seu grau de pureza.

A purificação de uma proteína geralmente constitui-se em um processo trabalhoso já que, tipicamente, uma célula contém milhares de substâncias diferentes, muitas das quais com propriedades físico-químicas semelhantes. O processo de purificação consiste em um procedimento de fracionamento. Em uma série de passos independentes, são utilizadas as várias propriedades físico-químicas da proteína de interesse a fim de separá-la, progressivamente, de outras substâncias.

O primeiro passo consiste na extração da proteína, através da ruptura das células por processos mecânicos ou químicos. Uma purificação inicial é feita por centrifugação, quando a fração solúvel é separada das membranas celulares ou outros componentes subcelulares. Em seguida, pode ser utilizado o processo de precipitação fracionada, no qual a solubilidade das proteínas é modificada pela variação na concentração de sal ou pH do meio. Certas proteínas precipitam sob determinadas condições enquanto outras permanecem solúveis.

Outros métodos de purificação são os chamados métodos cromatográficos. De uma maneira geral, a mistura das substâncias a serem fracionadas, dissolvida em um líquido, é denominada de fase móvel. Tal solução é percolada através de uma coluna consistindo de uma matriz sólida porosa, conhecida como fase estacionária. As interações dos solutos individuais que compõem a mistura com a fase estacionária têm o efeito de retardar seu progresso através da matriz, em uma taxa que varia com as propriedades de cada soluto. Os diversos métodos cromatográficos podem ser classificados de acordo

com a natureza da interação dominante entre a fase estacionária e as substâncias que estão sendo separadas. Entre os vários métodos pode-se citar os mais utilizados. A cromatografia de troca iônica, onde matrizes carregadas interagem eletrostaticamente com os solutos, explorando as diferenças de cargas entre as diversas substâncias; a cromatografia de filtração em gel, onde as moléculas são separadas por seu tamanho e forma; a cromatografia de afinidade, que explora o fato de que muitas proteínas têm a propriedade de interagirem especificamente com determinadas moléculas (tal molécula é então covalentemente ligada à matriz da coluna). Finalmente, a cromatografia de interação hidrofóbica, que separa proteínas de acordo com o grau de hidrofobicidade de suas superfícies (Voet; Voet, 1995).

2.2.2- Purificação da enzima GAPDH de *T. cruzi* tipo selvagem e mutantes

A enzima GAPDH tipo selvagem foi purificada conforme descrito anteriormente (Souza, 1996). Os tampões utilizados durante o processo de purificação são relacionados a seguir:

- tampão B: TEA-HCl 0.1 M pH 7.6 contendo NAD^+ 1 mM, DTT 1 mM, EDTA 1 mM, PMSF 1 mM, leupeptina 1 μM e pepstatina 1 μM ;
- tampão A: tampão B adicionado de sulfato de amônio 1 M;
- tampão C: tampão B adicionado de KCl 0.5 M.

Após a expressão, as células bacterianas separadas por centrifugação foram ressuspensas em 30 mL de tampão B e lisadas por sonicação. Aproximadamente 7.5 mL de cultura de células, mantidos em banho de gelo, foram submetidos a pulsos de 50 mA por 5 minutos, divididos em períodos de 1 minuto com intervalos de 10 segundos. A fração solúvel foi recuperada por centrifugação a $20\,000 \times g$ por 20 minutos a 4°C . A primeira etapa de purificação consistiu na precipitação de parte das proteínas contaminantes pela adição de sulfato de amônio até 50% de saturação. A mistura foi mantida sob agitação lenta por 15 minutos a 4°C e o precipitado removido por centrifugação a $20\,000 \times g$ por 20 minutos a 4°C .

As etapas seguintes de purificação consistiram de duas cromatografias. O sobrenadante após a precipitação com sulfato de amônio foi aplicado na coluna de interação hidrofóbica Phenyl Sepharose CL-4B (*Pharmacia*) com cerca de 20 mL de volume e previamente equilibrada com tampão A. O fluxo utilizado foi de 2 mL/min e o procedimento para eluição da amostra constou da aplicação de 60 mL de tampão A, seguido de um gradiente reverso de sulfato de amônio de 1 a 0 M em 100 mL e a lavagem da coluna com 50 mL de tampão B.

As frações eluídas da coluna Phenyl Sepharose contendo a enzima GAPDH ativa (medidas de atividade descritas na próxima seção) foram agrupadas e concentradas em uma câmara de microfiltração (*Amicon*), simultaneamente à troca de tampão. A concentração foi realizada até que o volume final da amostra atingisse 10 mL em tampão B. A amostra foi então aplicada na coluna de troca catiônica Phospho Ultrogel AGR (*IBF Biotechnics*) com cerca de 10 mL de volume e previamente equilibrada com tampão B. Com um fluxo de 1 mL/min, foram aplicados 30 mL de tampão B, um gradiente de 0 a 0.5 M de KCl em 60 mL, seguidos da lavagem da coluna com 30 mL de tampão C.

Todas as etapas de purificação foram conduzidas a 4 °C e acompanhadas por eletroforese em gel de poliacrilamida (Laemmli, 1970). A figura 2.1 mostra um gel de poliacrilamida a 15% em condição desnaturante, ilustrando os processos de expressão e purificação da enzima GAPDH de *T. cruzi* tipo selvagem. Embora as frações eluídas da coluna Phospho Ultrogel ainda apresentassem uma pequena porcentagem de contaminantes, o grau de pureza mostrou-se satisfatório para os experimentos de cristalização da enzima.

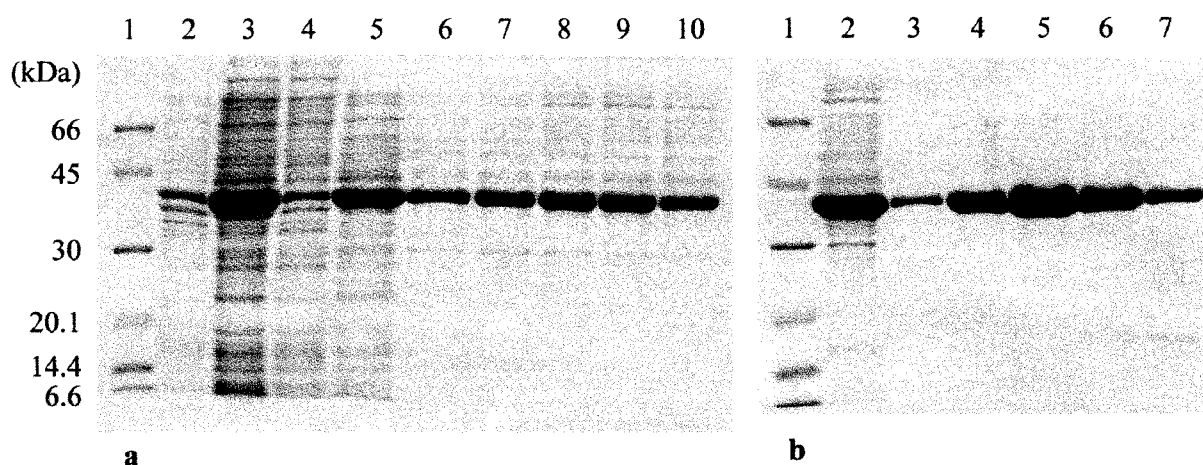


Figura 2.1 - SDS-PAGE 15% - expressão e purificação da GAPDH de *T. cruzi* tipo selvagem. Colunas 1a e 1b: marcadores de peso molecular (66, 45, 30, 20.1, 14.4, 6.6 kDa); colunas 2a e 3a: BL21(DE3)gap⁻ TcGAP lisada - (2a) precipitado, (3a) sobrenadante; colunas 4a e 5a: extrato após precipitação com (NH₄)₂SO₄ - (4a) precipitado, (5a) sobrenadante; colunas 6a a 10a: frações eluídas da Phenyl Sepharose; coluna 2b: amostra concentrada e aplicada na Phospho Ultrogel; colunas 3b a 7b: frações eluídas da Phospho Ultrogel.

A purificação da enzima mutante GAPDH D210L-G213D seguiu o mesmo procedimento descrito para a enzima tipo selvagem e apresentou resultados semelhantes. Entretanto, isso não ocorreu no caso do mutante GAPDH D210L. Após as etapas de precipitação com sulfato de amônio e cromatografia de interação hidrofóbica (coluna Phenyl Sepharose), a amostra foi eluída da coluna Phospho Ultrogel nas primeiras frações após sua aplicação, sem interagir com a coluna e, portanto, com elevada porcentagem de contaminantes. As possibilidades de erros no procedimento como a presença de alta concentração de sal na amostra aplicada na coluna, o não equilíbrio da coluna com o tampão de início do gradiente e até mesmo a preparação equivocada de tampões, possibilidades que poderiam levar a um resultado semelhante, foram verificadas e descartadas. A figura 2.2 mostra um gel de poliácridamida a 15% em condição desnaturante acompanhando a tentativa de purificação do mutante GAPDH D210L, quando foi utilizado o mesmo protocolo que para a enzima tipo selvagem e mutante GAPDH D210L-G213D.

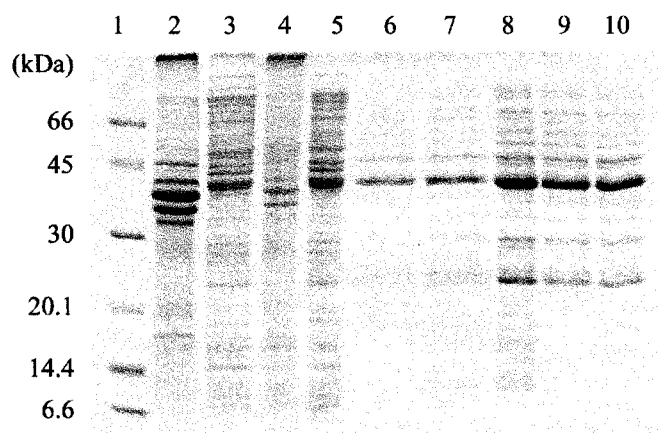


Figura 2.2 - SDS-PAGE 15% - tentativa de purificação do mutante GAPDH D210L. Coluna 1: marcadores de peso molecular (66, 45, 30, 20.1, 14.4, 6.6 kDa); colunas 2 e 3: BL21(DE3)gap⁻ TcGAP210L lisada - (2) precipitado, (3) sobrenadante; colunas 4 e 5: extrato após precipitação com (NH₄)₂SO₄ - (4) precipitado, (5) sobrenadante; colunas 6 e 7: frações eluídas da Phenyl Sepharose; coluna 8: amostra concentrada e aplicada na Phospho Ultrogel; colunas 9 e 10: frações eluídas da Phospho Ultrogel (antes do início do gradiente).

O comportamento diferenciado do mutante GAPDH D210L no processo de purificação não era esperado, já que a substituição de apenas um resíduo na sequência da proteína não deveria alterar sua interação com a coluna. Entretanto, a matriz da coluna Phospho Ultrogel, além de sua característica de trocadora catiônica, apresenta a propriedade de interagir por afinidade com proteínas que ligam íons fosfato, como é o caso da GAPDH. Por outro lado, estruturas cristalográficas de GAPDH de outras espécies mostraram que os sítios de ligação dos íons fosfato podem ser ocupados por íons sulfato, quando estes estão presentes no meio ao invés daqueles (capítulo 1, seção 1.3). Em vista disso e como o resíduo modificado Asp210 interage com a Arg249, que por sua vez interage com íons fosfato/sulfato quando estes estão presentes, levantou-se a hipótese de que a mutação pode ter levado a uma interação mais estável da Arg249 com os íons fosfato/sulfato. Dessa forma, em virtude das primeiras etapas de purificação constituírem-se de uma precipitação com sulfato de amônio e cromatografia de interação hidrofóbica em tampão contendo esse sal, íons sulfato poderiam estar fortemente ligados

nos sítios dos íons fosfato, impedindo, posteriormente, a interação da proteína com a coluna Phospho Ultrogel.

Como procedimento alternativo para a purificação do mutante GAPDH D210L e tendo em vista a hipótese acima, as etapas realizadas em presença de sulfato de amônio (precipitação e cromatografia de interação hidrofóbica) foram substituídas por uma cromatografia de filtração em gel.

Após a lise bacteriana e separação da fração solúvel por centrifugação, o extrato, em tampão B adicionado de NaCl 0.2 M, foi concentrado até o volume de cerca de 3 mL (*CENTRIPREP 30*). A amostra foi então aplicada na coluna de filtração Superdex 200 (*Pharmacia*), previamente equilibrada com tampão B adicionado de NaCl 0.2 M. O fluxo utilizado foi de 0.5 mL/min e a amostra foi eluída com o mesmo tampão. As frações contendo proteína ativa foram agrupadas e concentradas em uma câmara de microfiltração (*Amicon*), simultaneamente à troca de tampão. A concentração foi realizada até que o volume final da amostra atingisse 10 mL em tampão B.

A etapa seguinte foi idêntica àquela descrita para GAPDH tipo selvagem. A amostra foi aplicada na coluna de troca catiônica Phospho Ultrogel AGR (*IBF Biotechnics*) previamente equilibrada com tampão B. Com um fluxo de 1 mL/min, foram aplicados 30 mL de tampão B, um gradiente de 0 a 0.5 M de KCl em 60 mL seguidos da lavagem da coluna com 30 mL de tampão C. Observou-se que parte da proteína novamente não interagiu com a coluna, sendo eluída antes do início do gradiente mas outra parte foi eluída em um volume do gradiente semelhante ao observado na purificação da GAPDH tipo selvagem.

A figura 2.3 ilustra o acompanhamento dos processos de expressão e purificação do mutante GAPDH D210L. O resultado obtido pela eletroforese em gel de poliacrilamida a 15% em condição desnaturante mostra que a utilização da coluna Superdex 200 não foi muito eficaz, talvez sendo este o motivo pelo qual mesmo a parte da amostra que interagiu com a coluna Phospho Ultrogel não apresentou um grau de pureza satisfatório.

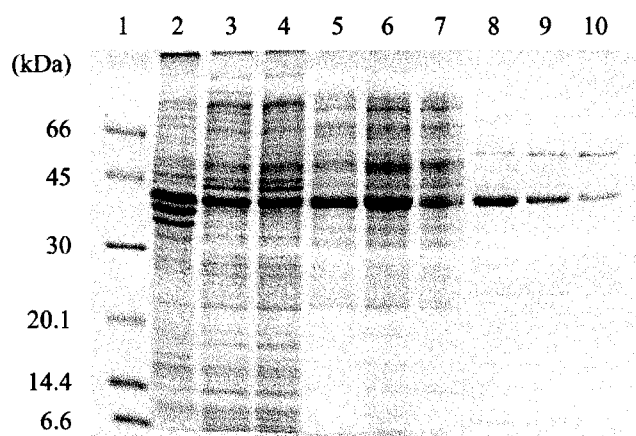


Figura 2.3 - SDS-PAGE 15% - nova tentativa de expressão e purificação do mutante GAPDH D210L. Coluna 1: marcadores de peso molecular (66, 45, 30, 20.1, 14.4, 6.6 kDa); colunas 2 e 3: BL21(DE3)gap⁻ TcGAP210L lisada - (2) precipitado, (3) sobrenadante; coluna 4: amostra concentrada e aplicada na Superdex 200; coluna 5: fração eluída da Superdex 200; coluna 6: amostra concentrada e aplicada na Phospho Ultrogel; coluna 7: fração eluída sem interação com a Phospho Ultrogel; colunas 8 a 10: frações eluídas da Phospho Ultrogel durante o gradiente.

O fato de que a eliminação das etapas de purificação envolvendo a presença de sulfato de amônio devolveu à proteína a capacidade de ligar-se à coluna Phospho Ultrogel vem apoiar a idéia de que a mutação D210L afeta a interação da proteína com íons fosfato/sulfato. Entretanto, apenas esta evidência é insuficiente para que uma análise mais detalhada seja realizada, devendo esta ser feita com base nas estruturas tridimensionais da enzima tipo selvagem e mutantes.

A fim de melhorar o grau de pureza obtido para o mutante GAPDH D210L foram feitas nova expressão e purificação, quando a amostra foi recromatografada na coluna Superdex 200 antes de sua aplicação na coluna Phospho Ultrogel. Os resultados de tal experimento são mostrados na figura 2.4. Novamente a porcentagem de contaminantes foi superior àquelas obtidas na purificação de GAPDH tipo selvagem e mutante D210L-G213D. A amostra purificada do mutante D210L, possivelmente inadequada à cristalização, é suficiente, no entanto, para os estudos cinéticos.

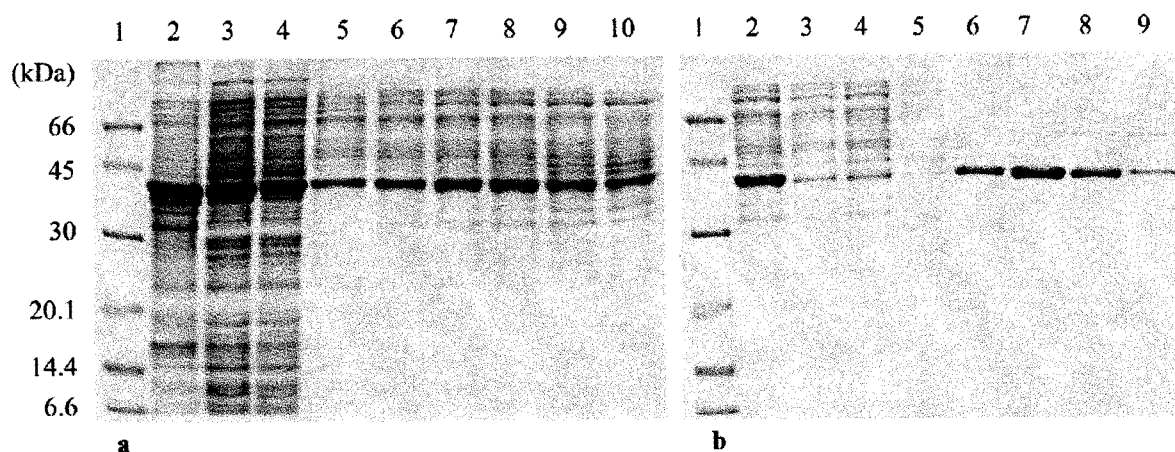


Figura 2.4 - SDS-PAGE 15% - expressão e purificação do mutante GAPDH D210L. Colunas 1a e 1b: marcadores de peso molecular (66, 45, 30, 20.1, 14.4, 6.6 kDa); colunas 2a e 3a: BL21(DE3)gap⁻TcGAP210L lisada - (2a) precipitado, (3a) sobrenadante; coluna 4a: amostra concentrada e aplicada na coluna Superdex 200; colunas 5a a 10a: frações eluídas da Superdex 200; coluna 2b: amostra recromatografada na Superdex 200, concentrada e aplicada na Phospho Ultrogel; colunas 3b a 5b: frações eluídas sem interação com a Phospho Ultrogel; colunas 6b a 9b: frações eluídas da Phospho Ultrogel durante o gradiente.

As amostras purificadas da enzima GAPDH de *T. cruzi* tipo selvagem e mutantes D210L-G213D e D210L tiveram suas concentrações determinadas pelo método de Bradford (Bradford, 1976) e foram utilizadas para medidas de cinética enzimática e experimentos de cristalização. O rendimento dos processos de expressão e purificação da enzima tipo selvagem e do mutante GAPDH D210L-G213D foi de cerca de 14 mg de proteína para 0.5 L de cultura de bactéria. No caso do mutante GAPDH D210L o rendimento foi bem menor, cerca de 1.5 mg de proteína para 0.5 L de cultura de bactéria.

2.3 - Cinética enzimática

2.3.1 - Introdução

No presente contexto, cinética significa o estudo das taxas segundo as quais as reações químicas ocorrem. Mais especificamente, as reações químicas envolvidas em processos biológicos são geralmente mediadas por enzimas, que desempenham a função de catalisadores. Esses notáveis catalisadores são individualmente específicos para determinadas reações e através de medidas de cinética enzimática pode-se determinar a afinidade de ligação entre enzima e substrato e a taxa catalítica máxima de uma enzima. Além disso, a observação da variação da taxa de reação em diferentes condições, combinada com informações químicas e estruturais podem levar à determinação do mecanismo catalítico da enzima (Voet; Voet, 1995).

O primeiro estudo de cinética enzimática foi realizado em 1902, quando Adrian Brown investigou a reação de hidrólise da sacarose catalisada pela enzima invertase (hoje conhecida como β -fructofuranosidase). Seu trabalho propôs que a reação catalisada por uma enzima é composta de duas reações elementares, nas quais o substrato (S) forma um complexo com a enzima (ES) e, posteriormente, tal complexo decompõe-se em produto (P) + enzima (E). Esquemáticamente,



De acordo com tal modelo, quando a concentração do substrato torna-se alta o bastante para converter a enzima completamente para a forma do complexo ES , o segundo passo torna-se o limitante e a taxa da reação global fica insensível a aumentos adicionais na concentração do substrato.

A velocidade da reação, ou velocidade de formação do produto, é dada por:

$$v = \frac{d[P]}{dt} = k_2[ES] \quad (2.2)$$

enquanto a taxa de produção global do complexo ES é a diferença entre as taxas da reação elementar que leva à sua formação e aquelas responsáveis pela sua decomposição:

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] \quad (2.3)$$

Com exceção do estágio inicial da reação, limitado a alguns milisegundos referentes à mistura entre enzima e substrato, a concentração do complexo $[ES]$ é mantida aproximadamente constante até que a quantidade de substrato presente seja rarefeita. Portanto, durante a maior parte da reação, $[ES]$ mantém um estado estacionário e $\frac{d[ES]}{dt} = 0$.

As expressões cinéticas devem ser escritas em termos de quantidades mensuráveis, o que não é o caso de $[E]$ e $[ES]$. Chamando de $[E]_T$ a concentração total da enzima ($[E]_T = [E] + [ES]$) e levando-se em conta a condição de estado estacionário, manipulações simples levam à equação:

$$[ES] = \frac{[E]_T[S]}{K_M + [S]} \quad (2.4)$$

onde $K_M = \frac{k_{-1} + k_2}{k_1}$ é conhecida como constante de Michaelis.

A velocidade inicial da reação pode ser então expressa em termos das quantidades mensuráveis $[E]_T$ e $[S]$:

$$v_0 = \left(\frac{d[P]}{dt} \right)_{t=0} = k_2[ES] = \frac{k_2[E]_T[S]}{K_M + [S]} \quad (2.5)$$

A velocidade máxima de reação, V_{max} , ocorre a altas concentrações de substrato, quando a enzima apresenta-se inteiramente na forma ES . Portanto, $V_{max} = k_2[E]_T$ e obtém-se a chamada equação de Michaelis-Menten,

$$v_0 = \frac{V_{max}[S]}{K_M + [S]} \quad (2.6)$$

equação básica no estudo de cinética enzimática.

A constante de Michaelis K_M é a concentração de substrato para a qual a velocidade de reação corresponde à metade da velocidade máxima e varia largamente de uma enzima para outra, sendo também função da temperatura e do pH. O valor de K_M diminui com o aumento da afinidade da enzima pelo substrato e, portanto, pode ser considerado uma medida dessa afinidade. O parâmetro V_{max} está relacionado com a eficiência catalítica da enzima. Define-se a constante catalítica:

$$k_{cat} = \frac{V_{max}}{[E]_T} \quad (2.7)$$

que corresponde ao número de reações que cada sítio ativo catalisa por unidade de tempo. A razão k_{cat}/K_M é uma medida da eficiência catalítica da enzima.

As medidas das velocidades iniciais da reação para várias concentrações do substrato permitem a construção da curva de Michaelis-Menten e a determinação dos valores de V_{max} e K_M característicos da enzima.

2.3.2 - Medidas de cinética enzimática de GAPDH de *T. cruzi* tipo selvagem e mutantes

Com o objetivo de se determinar a afinidade de ligação entre enzima e substrato e a eficiência catalítica da enzima, medidas de cinética foram realizadas para GAPDH tipo selvagem e mutantes D210L-G213D e D210L. Através da comparação dos resultados pode-se tirar conclusões a respeito da influência das mutações no mecanismo catalítico.

A atividade enzimática foi medida espectrofotometricamente segundo Barbosa e Nakano (1986), com algumas modificações. A enzima era colocada em presença do substrato, DL-gliceraldeído-3-fosfato (G3P), do cofator NAD^+ e de Na_2HAsO_4 , inibidor da reação reversa e media-se a formação do produto NADH.

A solução do substrato DL-gliceraldeído-3-fosfato foi preparada a partir do DL-gliceraldeído-3-fosfato dietilacetal, sal de bário, de acordo com instruções do fabricante (*Sigma*). A concentração da solução foi determinada espectrofotometricamente em presença de excesso de NAD^+ , da enzima e de arseniato de sódio (Dagher; Deal, 1977). A formação de NADH foi acompanhada através de medida de absorção a 340 nm até que

tal leitura se mantivesse constante, indicando, indiretamente, que todo o substrato G3P havia sido transformado. Dessa forma, a concentração do substrato foi determinada utilizando o valor máximo da medida de absorção e o coeficiente de extinção do NADH ($\epsilon = 6.22 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$). A determinação da concentração da solução de NAD^+ foi feita através da medida de absorção a 260 nm, utilizando-se seu coeficiente de extinção molar ($\epsilon = 17.6 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$).

Para as medidas de cinética enzimática, a mistura de ensaio continha em um volume final de 1 mL: tampão Tris-HCl 50 mM pH 8.6 adicionado de EDTA e β -mercaptoetanol 1 mM e arseniato de sódio 30 mM. A concentração do substrato G3P variou de 0.05 até cerca de 2.5 mM, enquanto a concentração do cofator NAD^+ variou de 0.05 até cerca de 2.0 mM. Os parâmetros cinéticos do G3P foram determinados em concentração saturante de NAD^+ (2.7 mM), enquanto os parâmetros cinéticos do NAD^+ foram determinados em concentração saturante de G3P (2.0 mM). A concentração da enzima presente na mistura de reação foi escolhida de modo que a quantidade de produto formado em função do tempo seguisse uma relação linear. Foram utilizados 0.075 $\mu\text{g/mL}$ de enzima para a determinação dos parâmetros cinéticos do G3P e 0.125 $\mu\text{g/mL}$ para determinação dos parâmetros cinéticos do NAD^+ . A reação era iniciada pela adição do substrato, devido à sua instabilidade em soluções aquosas.

Os valores das velocidades iniciais de reação para as diversas concentrações de G3P e NAD^+ foram obtidos através de medidas de absorção a 340 nm, acompanhando-se a formação do produto NADH ao longo de 2 minutos. A razão entre a taxa de formação do produto ($\Delta\text{Abs}/\Delta t$) e o coeficiente de extinção do NADH fornece a velocidade de reação para uma determinada concentração de substrato ou cofator. Os experimentos foram realizados a 25 °C utilizando um espectrofotômetro *Hitachi U2000*. Todas as medidas foram feitas em duplicata e os parâmetros cinéticos foram determinados a partir do ajuste da curva de Michaelis-Menten (equação 2.6) aos dados experimentais.

As figuras 2.5 e 2.6 mostram as curvas de Michaelis-Menten em função da concentração de G3P e NAD^+ para a enzima GAPDH tipo selvagem e mutantes. A partir dos valores obtidos para V_{max} e das concentrações das proteínas na mistura de reação,

pode-se calcular as constantes catalíticas k_{cat} . Os parâmetros cinéticos da enzima tipo selvagem e mutantes em relação ao G3P e ao NAD^+ são apresentados na tabela 2.1.

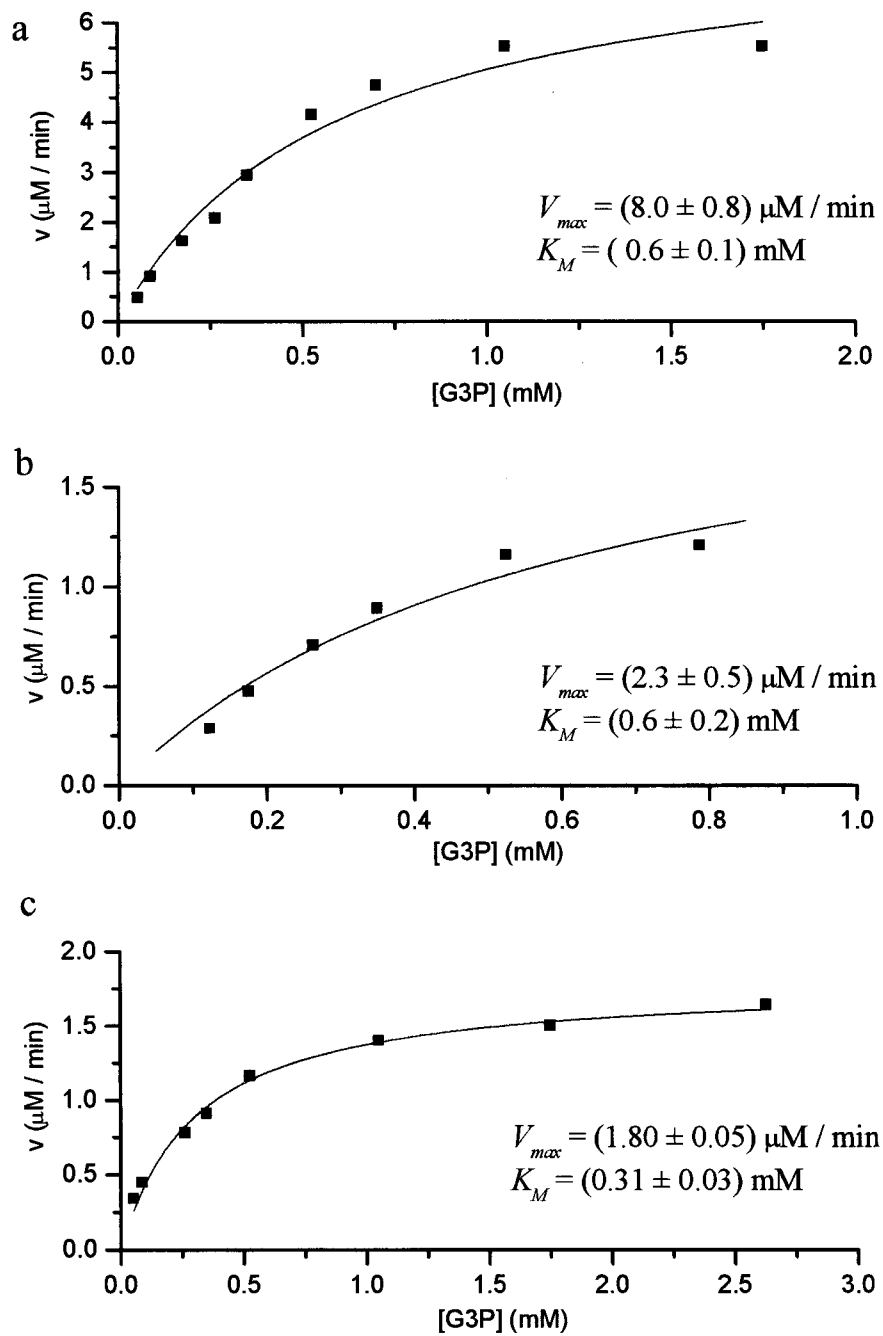


Figura 2.5 - Curvas de Michaelis-Menten em função da concentração do substrato G3P para a enzima a) GAPDH tipo selvagem, b) mutante D210L-G213D e c) mutante D210L.

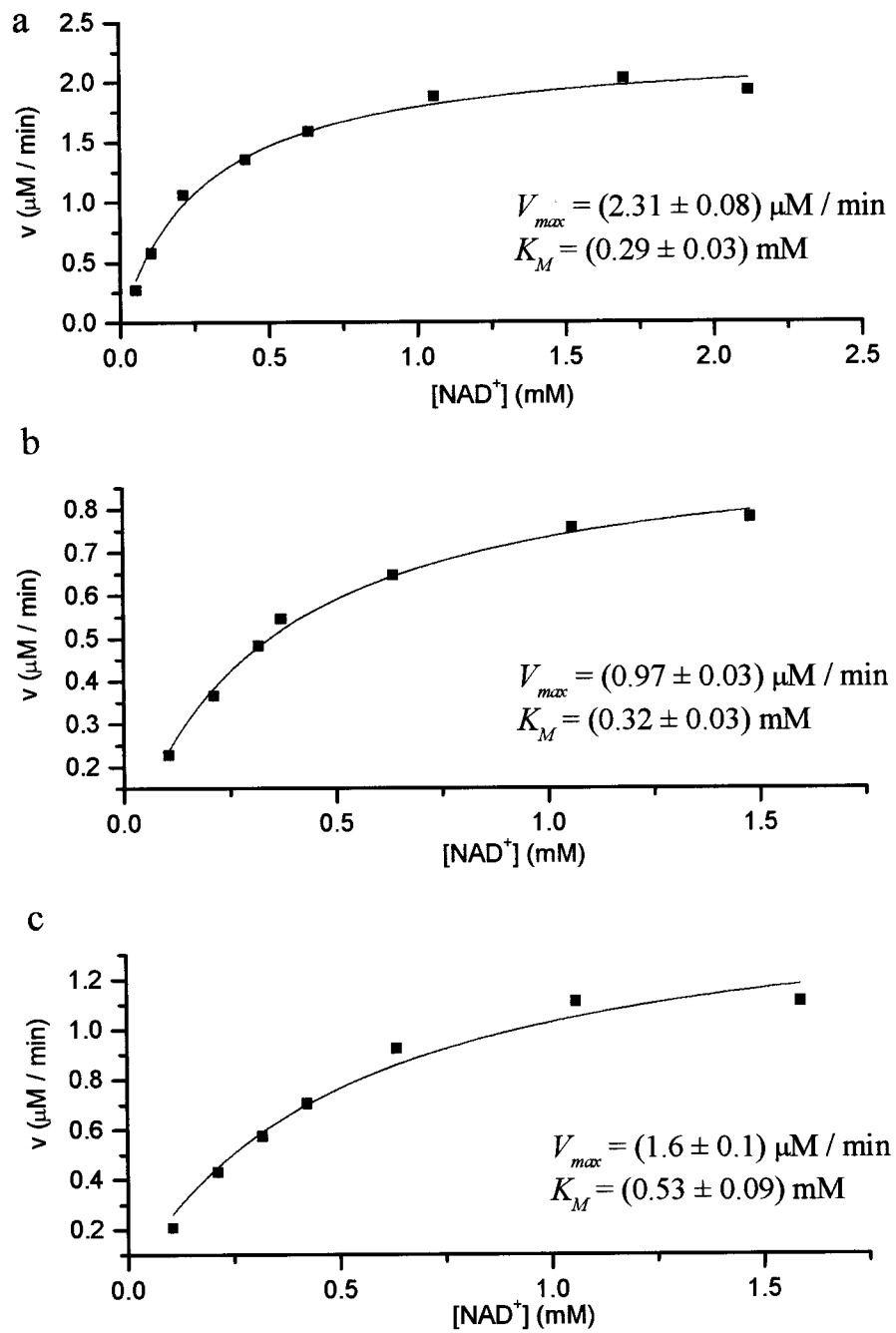


Figura 2.6 - Curvas de Michaelis-Menten em função da concentração do cofator NAD^+ para a enzima a) GAPDH tipo selvagem, b) mutante D210L-G213D e c) mutante D210L.

Tabela 2.1 - Valores das constantes de Michaelis (K_M), catalítica (k_{cat}) e da razão k_{cat}/K_M para enzima GAPDH tipo selvagem e mutantes, em relação ao substrato G3P e ao cofator NAD^+ .

	G3P			NAD^+		
	K_M (mM)	$k_{cat} \times 10^{-3}$ (s ⁻¹)	(k_{cat} / K_M) $\times 10^{-3}$ (s ⁻¹ mM ⁻¹)	K_M (mM)	$k_{cat} \times 10^{-3}$ (s ⁻¹)	(k_{cat} / K_M) $\times 10^{-3}$ (s ⁻¹ mM ⁻¹)
GAPDH tipo selvagem	0.6 ± 0.1	1.8 ± 0.2	3.0 ± 0.8	0.29 ± 0.03	0.31 ± 0.01	1.1 ± 0.1
GAPDH D210L-G213D	0.6 ± 0.2	0.5 ± 0.1	0.8 ± 0.4	0.32 ± 0.03	0.129 ± 0.004	0.41 ± 0.05
GAPDH D210L	0.31 ± 0.03	0.40 ± 0.01	1.3 ± 0.2	0.53 ± 0.09	0.21 ± 0.01	0.40 ± 0.09

Os desvios associados aos valores das constantes são provenientes do ajuste das curvas de Michaelis-Menten aos dados e decorrem de erros experimentais associados às medidas das velocidades iniciais de reação para cada concentração de substrato ou cofator. Erros sistemáticos, como presentes na determinação da concentração da enzima, são difíceis de avaliar, entretanto não devem comprometer análises comparativas.

Os resultados acima mostram que, considerando os desvios associados, não houve diferenças significativas entre as constantes de Michaelis (K_M) da enzima tipo selvagem e dos mutantes, tanto em relação ao substrato quanto em relação ao cofator, indicando que as mutações não afetaram a formação do complexo *ES*. Por outro lado, observa-se uma diminuição significativa na constante catalítica da enzima (k_{cat}) em decorrência das mutações, indicando que as modificações introduzidas afetaram o seu sítio ativo. A influência das mutações sobre o mecanismo catalítico da enzima será discutida mais detalhadamente no capítulo 5, com base na conformação dos resíduos do sítio ativo.

Capítulo 3

Cristalização, Coleta e Processamento de Dados

A determinação da estrutura tridimensional de uma molécula por cristalografia de raios-X requer a obtenção de monocristais da referida molécula, a partir dos quais os dados de difração serão coletados. Vários aspectos teóricos e práticos da cristalização de macromoléculas ainda não estão bem estabelecidos, constituindo-se tal etapa em um dos grandes desafios, senão o maior, no processo de determinação da estrutura tridimensional de uma proteína. O presente capítulo apresenta as metodologias utilizadas na cristalização da enzima GAPDH de *T. cruzi* tipo selvagem e mutantes e na coleta de dados de difração de raios-X, bem como seus resultados.

3.1 - Cristalização

3.1.1 - Introdução

Cristais são arranjos ordenados e periódicos de átomos ou moléculas em três dimensões. A cristalização de moléculas a partir da solução depende de uma combinação de parâmetros físicos e químicos dos solvente e soluto envolvidos e leva o sistema a um estado de energia livre mínima. Sob condições de supersaturação, as moléculas podem agregar-se formando um precipitado amorfo ou podem cristalizar-se. Para sistemas simples, constituídos por íons ou pequenas moléculas, a energia livre em solução, energias de ligação, potenciais ligações no estado sólido e outros fatores são relativamente bem conhecidos e o processo de crescimento de cristais pode ser teoricamente descrito e predito (McPherson, 1982). Por outro lado, para sistemas de macromoléculas, as interações são bem mais complexas, dificilmente descritas e entendidas, fazendo da cristalização um processo altamente empírico.

A estratégia utilizada no processo de cristalização de macromoléculas biológicas consiste em trazer o sistema lentamente para um estado de supersaturação, controlando, simultaneamente, algumas variáveis físicas e químicas do sistema. A dificuldade reside, justamente, no elevado número de parâmetros envolvidos, tais como concentração, temperatura, pH e força iônica, citando apenas alguns. Além disso, as macromoléculas apresentam alta flexibilidade conformacional e alta sensibilidade às condições do meio.

A determinação do conjunto de condições que levam à cristalização de uma proteína não é nem um pouco trivial. Idealmente, dever-se-ia determinar os pontos de precipitação da proteína em diferentes valores de pH para um determinado agente precipitante, repetir o procedimento em diversas temperaturas e examinar os efeitos de diferentes agentes precipitantes. Além disso, existem outras variáveis como concentração da proteína, presença de aditivos, etc. A natureza multi-paramétrica do processo inviabiliza o teste de todas as condições, tendo sido desenvolvidos os chamados métodos fatoriais, a fim de identificar os principais parâmetros e os efeitos de sua variação. O método do fatorial incompleto, desenvolvido por Carter e Carter (1979), é muito utilizado em cristalização de proteínas e constitui-se em um conjunto limitado de soluções, derivadas de condições conhecidas de cristalização de diversas proteínas, considerando a combinação de diferentes pH e agentes precipitantes. A observação de monocristais em alguma das soluções do fatorial, mesmo que de tamanho ou morfologia inadequados, ou mesmo o simples indício de nucleação, é um ponto de partida para que monocristais adequados para experimentos de difração sejam obtidos, através do refinamento das condições iniciais de cristalização.

Há cerca de uma década, teve início uma pesquisa a respeito dos efeitos da ausência de gravidade sobre o crescimento de cristais de proteínas, sugerindo que a cristalização em ambiente de microgravidade pode melhorar a qualidade dos cristais, através da eliminação de correntes de convecção próximas à superfície do cristal em crescimento (De Lucas, 1989). Experimentos de cristalização de proteínas têm sido realizados a bordo de ônibus espaciais, com o objetivo de obter-se cristais internamente mais ordenados, o que é refletido como um padrão de difração até mais alta resolução.

Como dito anteriormente, a cristalização é um dos grandes desafios enfrentados durante o processo de determinação da estrutura tridimensional de uma proteína por difração de raios-X e pode levar dias, meses ou até anos. Utilizando as palavras dos autores: "a cristalização de macromoléculas é quase empírica e, devido à sua imprevisibilidade e freqüente irreprodutibilidade, tem sido considerada mais uma arte que uma ciência" (Ducruix; Giegé, 1992).

3.1.2 - Cristalização da enzima GAPDH tipo selvagem e mutante D210L-G213D

Uma condição de cristalização da enzima GAPDH de *T. cruzi* havia sido previamente determinada (Souza, 1996): solução da enzima (9mg/mL) em Tris-HCl 25 mM pH 7.8 adicionado de sulfato de amônio 0.5 M, NAD⁺ 2mM, azida sódica, DTT e EDTA 1 mM e solução do reservatório contendo cacodilato de sódio 0.1 M pH 6.5, PEG 8000 18% e acetato de cálcio 0.2 M. Em tal condição foram obtidos cristais a 4 °C de cerca de 0.1 × 0.2 × 0.2 mm, pertencentes ao grupo espacial P1 e com poder de difração até 2.8 Å de resolução. A presença sistemática de cristais de sal nas gotas sempre representou uma dificuldade adicional para a obtenção de cristais de melhor qualidade nesta condição.

Com o objetivo de estabelecer-se uma nova condição de cristalização para a proteína, variações foram introduzidas no pH, concentração do agente precipitante e sal utilizado. Também foram feitas tentativas de cristalização com a proteína no próprio tampão de eluição da última coluna cromatográfica. As amostras de proteína purificadas segundo procedimentos descritos anteriormente foram concentradas por centrifugação (*CENTRICON 30*) até uma concentração final de cerca de 9 mg/mL. Os experimentos de cristalização foram realizados a 4 e a 18 °C, utilizando a técnica de difusão de vapor com gotas suspensas (McPherson, 1982).

Cristais da enzima GAPDH de *T. cruzi* de cerca de 0.2 × 0.3 × 0.6 mm foram obtidos após cerca de 10 dias a 4 e a 18 °C (figura 3.1). As condições de cristalização foram: solução de proteína (9 mg/mL) em tampão TEA 0.1 M pH 7.6 contendo NAD⁺ 1 mM, DTT 1 mM, EDTA 1 mM, PMSF 1 mM, leupeptina 1 µM, pepstatina 1 µM e ≈ 0.2 M de KCl e solução do reservatório contendo cacodilato de sódio 0.1 M pH 7.1, PEG 8000 18% e acetato de cálcio 0.1 M. As gotas de cristalização

foram formadas a partir da mistura de 5 μL da solução de proteína e 5 μL da solução do reservatório. Os cristais mostraram-se maiores e morfologicamente superiores aos obtidos por Souza (1996) e, posteriormente, foi obtido um conjunto de dados de difração até mais alta resolução. A determinação dos parâmetros de rede mostrou que os cristais pertenciam ao grupo espacial $P2_1$, possuindo 1 tetrâmero por unidade assimétrica, diferentemente daqueles obtidos nas condições anteriores, pertencentes ao grupo espacial $P1$. Maiores detalhes sobre as coletas de dados serão dados na próxima seção.



Figura 3.1 - Cristais de GAPDH de *T. cruzi* tipo selvagem obtidos a 4 °C.

É importante ressaltar aqui um experimento inédito para o Brasil, a realização de ensaios de cristalização em ambiente de microgravidade. A enzima GAPDH de *T. cruzi* foi uma das proteínas selecionadas para participar de tais experimentos a bordo de um ônibus espacial da agência norte-americana NASA. As missões espaciais ocorreram em abril e julho de 1997.

Cristais de GAPDH foram obtidos a bordo do ônibus espacial em condições um pouco diferentes daquelas da cristalização no laboratório. Solução da enzima (9 mg/mL) em Tris-HCl 25mM pH 7.8 adicionado de sulfato de amônio 0.5 M, NAD^+ 2 mM, azida sódica, DTT e EDTA 1 mM e solução do reservatório contendo cacodilato de sódio 0.1 M pH 7.1, PEG 8000 20% e acetato de sódio 0.1 M. Os cristais apresentaram maior tamanho quando comparados aos obtidos em laboratório.

A figura 3.2 mostra uma fotografia dos cristais tirada na Universidade do Alabama, EUA, algumas horas após o retorno do ônibus espacial, além de uma fotografia tirada três dias mais tarde no IFSC-USP. O transporte dos cristais até o Brasil pode ter danificado os mesmos e causado a precipitação de proteína nas gotas, como observa-se na fotografia tirada no Brasil.

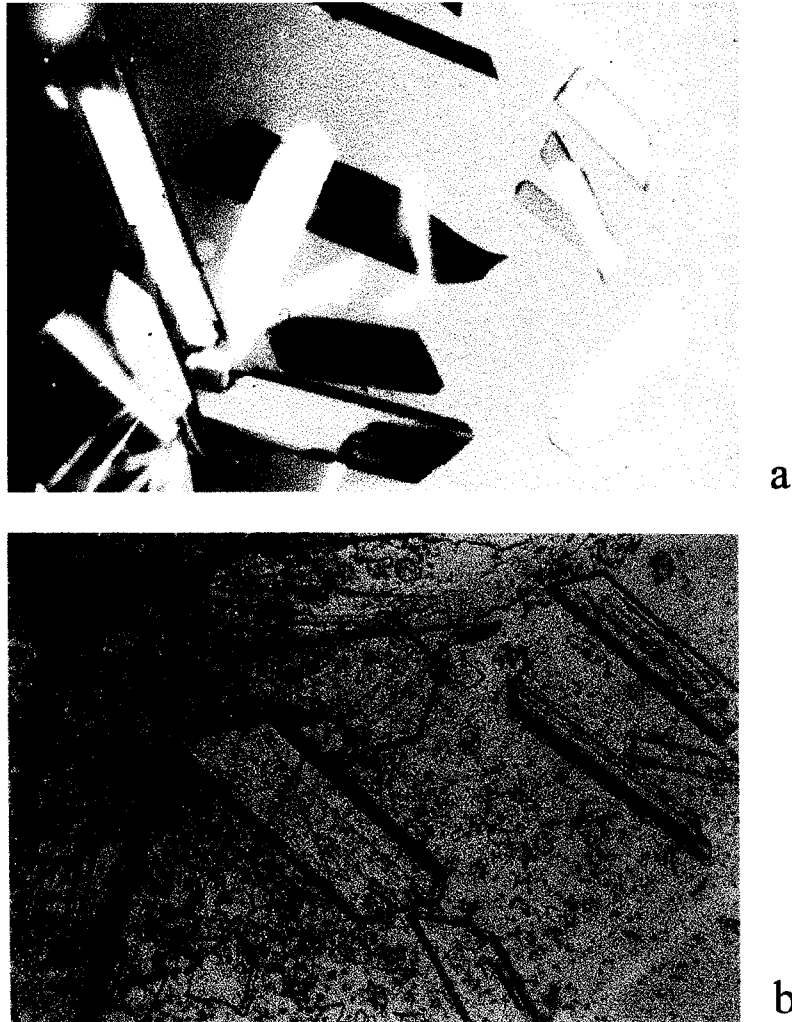


Figura 3.2 - Cristais de GAPDH de *T. cruzi* crescidos em ambiente de microgravidade; a) fotografia tirada na Universidade do Alabama, logo após o retorno do ônibus espacial, b) fotografia tirada no Brasil, três dias mais tarde.

Cristais do mutante GAPDH D210L-G213D de cerca de $0.2 \times 0.3 \times 0.6$ mm foram obtidos após cerca de 10 dias a 4 e a 18 °C (figura 3.3). As condições de

cristalização foram as mesmas daquelas em que foram obtidos cristais da enzima tipo selvagem no laboratório.

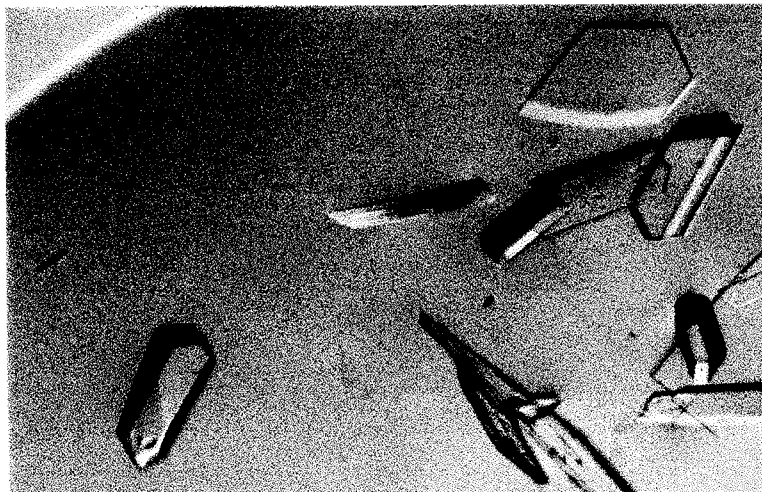


Figura 3.3 - Cristais do mutante GAPDH D210L-G213D obtidos a 4 °C.

As tentativas de cristalização do mutante GAPDH D210L não foram bem sucedidas, provavelmente devido à presença de contaminantes na amostra, como mostrado no capítulo anterior.

Na busca pelas condições de cristalização anteriormente descritas, uma outra forma cristalina foi encontrada nos experimentos a 4 °C, nas mesmas condições dos cristais obtidos no laboratório, com exceção do pH do tampão cacodilato de sódio que mudava de 7.1 para 6.5. Posteriormente, tais cristais apresentaram poder de difração até apenas 3.5 Å de resolução e a determinação dos parâmetros de rede mostrou que os cristais pertenciam ao grupo espacial I4.

3.2 - Coleta e processamento dos dados de difração

3.2.1 - Introdução

O padrão de difração de raios-X produzido por um monocristal contém informações a respeito do conteúdo e simetria de sua cela unitária. O equipamento necessário para a coleta de dados é composto, fundamentalmente, de uma fonte de raios-X, um instrumento mecânico para rotação do cristal, permitindo que os diversos planos cristalinos satisfaçam as condições de difração e um detector para a medição

das posições e intensidades dos feixes difratados. Os cristais de proteínas, por apresentarem celas unitárias relativamente grandes e suas moléculas arranjadas com um certo grau de desordem, espalham raios-X muito mais fracamente se comparados a cristais de pequenas moléculas (Drenth, 1994).

As fontes de raios-X disponíveis são as chamadas fontes convencionais (gerador de tubo selado e de anodo rotatório) e as fontes de radiação síncrotron. Estas últimas vêm sendo cada vez mais utilizadas em cristalografia de proteínas, principalmente devido à alta intensidade do feixe e à possibilidade de variação do comprimento de onda da radiação incidente. Fontes de radiação síncrotron cada vez mais potentes têm sido construídas em vários países, permitindo a realização de experimentos avançados como o estudo cristalográfico de grandes complexos protéicos e experimentos de cristalografia resolvida no tempo. O Brasil é o único país do hemisfério sul que conta com uma fonte de radiação síncrotron, o Laboratório Nacional de Luz Síncrotron (Campinas, SP), cuja linha de cristalografia de proteínas foi aberta aos usuários em setembro de 1997.

Um dos problemas existentes na coleta de dados de difração de um cristal de proteínas é o dano sofrido pelo cristal devido à utilização de uma radiação ionizante para o estudo da amostra. O chamado dano por radiação causa uma diminuição na intensidade dos pontos de difração e é dependente tanto da dose de raios-X recebida pelo cristal quanto do tempo de exposição. O processo ocorre em dois estágios: o primeiro consiste na criação de produtos radiolíticos quando os fótons de raios-X são absorvidos e elétrons são ejetados. O segundo, associado ao elevado conteúdo de solvente presente em cristais de proteína, consiste na difusão e reações desses produtos ao longo da rede cristalina, causando a destruição da ordem interna do cristal. Tem sido observado que o uso de radiação síncrotron fornece um aumento na quantidade de dados coletados antes do decaimento do padrão de difração, já que a alta intensidade do feixe incidente permite a diminuição do tempo de exposição do cristal. Uma outra maneira de aumentar-se o tempo de vida do cristal é através do seu resfriamento a 4 °C ou à temperatura do nitrogênio líquido. Nesse último caso, a difusão de radicais livres através da rede cristalina diminui enormemente e pode até mesmo cessar, mantendo a qualidade do padrão de difração por toda a coleta (Nave, 1995).

O congelamento de um cristal de proteína pode causar sua destruição devido à tensão mecânica induzida pelo aumento do volume específico que acompanha a transição da água para gelo. Tal problema é evitado com a utilização de compostos crioprotetores, que reduzem a taxa de cristalização da água, tanto no interior do cristal quanto na sua superfície, durante o processo de resfriamento. Os crioprotetores são compostos orgânicos de baixo ponto de congelamento entre os quais os mais utilizados são: glicerol, MPD e PEG de baixo peso molecular. A determinação da solução crioprotetora adequada à cada cristal é um processo de tentativa e erro, já que a própria difusão do composto no interior do cristal pode destruir a rede. Os cristais são banhados na solução crioprotetora e rapidamente congelados sob fluxo de N₂ gasoso antes da coleta de dados.

A coleta de dados a temperaturas criogênicas, técnica denominada criocristalografia, tem sido muito utilizada no estudo de cristais de proteínas e métodos experimentais têm sido aprimorados. Uma revisão recente foi feita por Garman e Schneider (1997).

O processamento dos dados de difração consiste na determinação da orientação do cristal e parâmetros da cela unitária, indexação das reflexões, escalonamento e promédio dos dados provenientes de um ou mais cristais. O conjunto dos dados processados, utilizado para a resolução da estrutura cristalográfica, contém uma lista de reflexões independentes com seus índices (*hkl*) associados a conjuntos de planos cristalinos e respectivas intensidades *I*, bem como os erros $\alpha(I)$ associados a cada medida.

Um indicador da qualidade do conjunto de dados coletados é o chamado R_{merge} , que compara as intensidades das reflexões relacionadas por simetria. Tal fator é definido por:

$$R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)} \quad (3.1)$$

onde a primeira soma é feita sobre as *i* reflexões equivalentes. Um valor de R_{merge} de até 15% é considerado aceitável.

3.2.2 - Coleta e processamento dos dados dos cristais de GAPDH tipo selvagem e mutante D210L-G213D

As coletas de dados foram feitas utilizando um difratômetro automático R-AXIS IIC da *Rigaku Corporation*, presente no Laboratório de Cristalografia de Proteínas e Biologia Estrutural do IFSC-USP. Tal equipamento conta com um gerador de anodo rotatório como fonte de raios-X e placas de imagem como detectores. Um sistema para controle de temperatura dos cristais (*Oxford cryosystems 600*) está acoplado ao difratômetro, permitindo a coleta de dados em condições criogênicas. Tal sistema opera entre 80.0 e 375.0 K. O método utilizado para coleta de dados é o de rotação-oscilação (Arndt; Wonacott, 1977), onde o cristal sofre pequenas e sucessivas oscilações em torno do eixo perpendicular à direção do feixe de raios-X incidente.

Uma imagem do padrão de difração de um cristal da enzima GAPDH de *T. cruzi* crescido em ambiente de microgravidade foi obtida no IFSC-USP, três dias após o retorno do ônibus espacial (figura 3.4). Tal imagem apresenta pontos de difração até um limite de 2 Å de resolução, mas o cristal mostrou-se muito sensível à radiação, ocorrendo um rápido decaimento do padrão de difração. Em vista disso, decidiu-se guardar os cristais para a realização das coletas de dados no Laboratório Nacional de Luz Síncrotron, cuja linha de cristalografia de proteínas seria inaugurada 2 meses depois (setembro de 1997). Infelizmente, nessa época os cristais mostraram-se bastante danificados e nenhum conjunto de dados de boa qualidade pôde ser coletado.

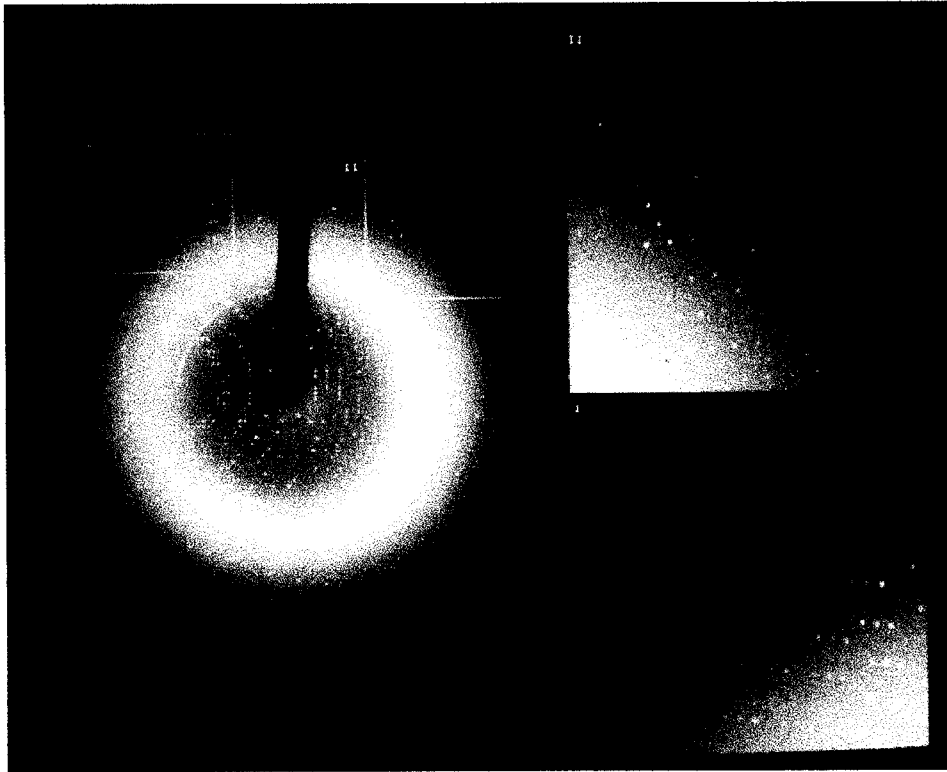


Figura 3.4 - Padrão de difração de um cristal de GAPDH de *T. cruzi* crescido em ambiente de microgravidade. Imagem obtida após 20 min de exposição utilizando fonte de raios-X convencional (gerador de anodo rotatório). As regiões em destaque mostram o limite de difração do cristal, que chega a 2 Å de resolução.

A coleta de dados de um cristal de GAPDH tipo selvagem, obtido no laboratório nas condições descritas na seção anterior, foi feita inicialmente à temperatura de 277 K. Entretanto, devido ao rápido decaimento do seu padrão de difração, um novo conjunto de dados foi coletado, agora em condições criogênicas (100 K). A solução crioprotetora adequada foi determinada após vários testes. Soluções de glicerol em concentrações de 10 a 35% e PEG 400 de 10 a 20% foram adicionadas à solução de cristalização e a mistura colocada sob o fluxo de N₂ a 100 K. A solução era exposta ao feixe de raios-X e a imagem mostrava a presença de anéis característicos de materiais policristalinos quando cristais de gelo eram formados. A figura 3.5 mostra as imagens de espalhamento por uma solução não crioprotetora e uma crioprotetora.

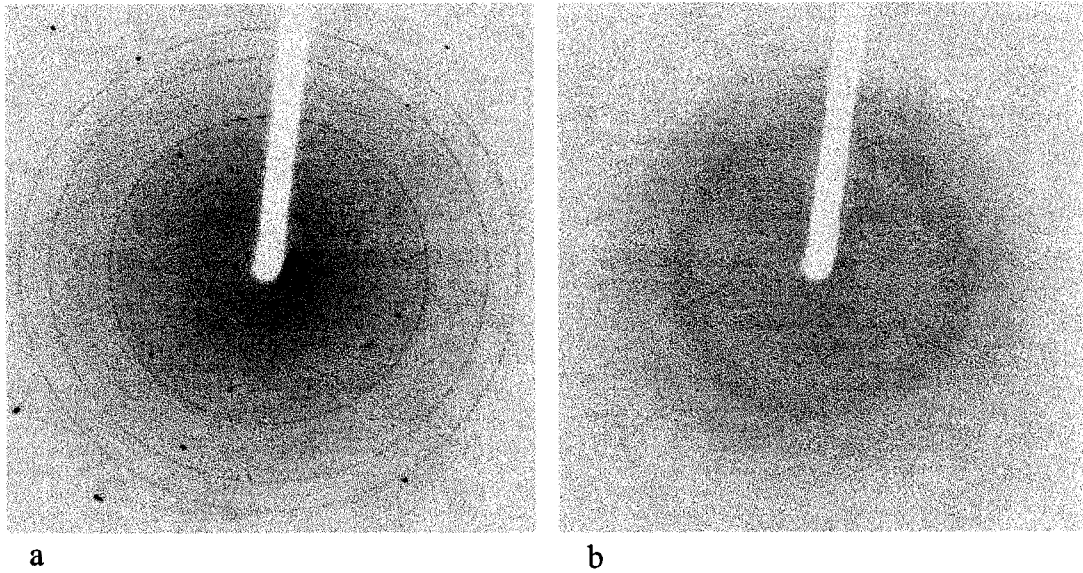


Figura 3.5 - Imagens de espalhamento de raios-X por (a) uma solução não crioprotetora, mostrando anéis decorrentes da formação de policristais de gelo e (b) uma solução crioprotetora.

Além de impedir a formação de cristais de gelo, o composto crioprotetor não deve, por si, destruir a ordem cristalina, fato observado quando da utilização de glicerol. Após algumas tentativas, a solução crioprotetora adequada foi obtida adicionando-se 20% de PEG 400 à composição da solução de cristalização. Os cristais eram deixados submersos nessa solução por cerca de 5 minutos e então colocados sob o fluxo de N_2 a 100 K. Os dados de difração dos cristais da enzima mutante GAPDH D210L-G213D foram coletados apenas em condições criogênicas, utilizando-se também 20% de PEG 400 como composto crioprotetor.

Os conjuntos de dados foram processados e promediados utilizando-se os programas DENZO e SCALEPACK (Otwinowski, 1993; Minor, 1993). Em virtude do grande aumento na mosaicidade dos cristais em decorrência do resfriamento, muitas reflexões foram rejeitadas devido a "overlap", fazendo-se necessária a utilização de mais de um cristal no caso das coletas a 100 K, a fim de se alcançar uma completude de dados satisfatória. A tabela 3.1 mostra os parâmetros das coletas de dados e das redes cristalinas obtidos para GAPDH tipo selvagem (coleta a 277 e a 100 K) e mutante D210L-G213D (coleta a 100 K), enquanto a tabela 3.2 apresenta as estatísticas dos conjuntos de dados.

Tabela 3.1 - Parâmetros das coletas de dados e das redes cristalinas obtidos para GAPDH tipo selvagem e mutante D210L-G213D.

	GAPDH tipo selvagem	GAPDH tipo selvagem	GAPDH D210L-G213D
Temperatura da coleta de dados (K)	277	100	100
Nº de cristais	1	3	2
Mosaicidade dos cristais (°)	0.3	0.9	0.9
Intervalo de oscilação* (°)	2	2, 1, 1	1
Tempo de cada imagem* (min)	50	50, 90, 90	60
Grupo espacial	P2 ₁	P2 ₁	P2 ₁
Parâmetros de rede	$a = 92.27 \text{ \AA}$, $b = 85.17 \text{ \AA}$, $c = 101.61 \text{ \AA}$, $\beta = 98.25^\circ$	$a = 78.85 \text{ \AA}$, $b = 84.93 \text{ \AA}$, $c = 106.16 \text{ \AA}$, $\beta = 95.70^\circ$	$a = 80.08 \text{ \AA}$, $b = 85.34 \text{ \AA}$, $c = 107.33 \text{ \AA}$, $\beta = 96.16^\circ$
Volume da cela unitária (Å ³)	790252	707409	729261

* Os diferentes valores correspondem aos diferentes cristais utilizados na coleta de dados.

Comparando os parâmetros da cela unitária dos cristais de GAPDH tipo selvagem obtidos com as coletas de dados a 277 e 100 K, observa-se uma significativa mudança, principalmente nos valores de a e c . Reduções no volume da cela unitária de 2 a 3% têm sido observadas em experimentos de criocristalografia e um caso de redução de 7.4% no volume foi relatado (Kurinov; Harrison, 1995). A variação aqui observada no parâmetro de rede a foi bastante drástica (14.5%) e uma redução de 10.5% no volume da cela unitária ainda não foi, até onde se sabe, descrita

na literatura. É importante destacar que os cristais utilizados nas coletas de dados a 277 e 100 K foram retirados da mesma gota de cristalização e, a fim de confirmar que a redução no volume da cela unitária foi causada pelo resfriamento, foram feitas imagens de difração de um mesmo cristal a 277 K e posteriormente a 100 K. Os parâmetros de rede determinados a partir de tais imagens revelaram as mudanças na cela unitária. Uma análise detalhada de tais mudanças será feita no capítulo 5, com base no conhecimento das estruturas tridimensionais das moléculas, seus contatos e empacotamentos cristalinos.

Tabela 3.2 - Estatísticas dos dados coletados para os cristais de GAPDH tipo selvagem e mutante D210L-G213D.

	GAPDH tipo selvagem	GAPDH tipo selvagem	GAPDH D210L- G213D
Temperatura da coleta de dados (K)	277	100	100
Nº de reflexões coletadas	107504	241340	283400
Nº de reflexões independentes	49909	74208	74012
Multiplicidade dos dados	2.15	3.3	3.8
R_{merge} (%)	11.8 (35.0 entre 2.6 e 2.5 Å)	8.9 (12.6 entre 2.2 e 2.15 Å)	7.1 (16.9 entre 2.2 e 2.15 Å)
Completeza (%)	92.2 (87.0 entre 2.6 e 2.5 Å)	97.7 (87.4 entre 2.2 e 2.15 Å)	93.8 (82.0 entre 2.2 e 2.15 Å)
Reflexões com $I > 2 \sigma(I)$ (%)	67.1 (39.1 entre 2.6 e 2.5 Å)	91.3 (83.5 entre 2.2 e 2.15 Å)	94.1 (84.9 entre 2.2 e 2.15 Å)
Resolução (Å)	2.5	2.15	2.15

Observou-se que os padrões de difração obtidos para os cristais coletados a 100 K praticamente não sofreram decaimento, o que resultou na mais alta resolução de seus conjuntos de dados. Tentativas de coletas de dados utilizando radiação síncrotron foram feitas no LNLS (Campinas, SP), mas os cristais mostraram-se muito sensíveis à radiação e como tal laboratório ainda não contava com um sistema para resfriamento da amostra, observou-se um decaimento muito rápido do padrão de difração e os dados não puderam ser aproveitados.

Após a determinação dos parâmetros da rede cristalina, uma estimativa do número de moléculas por cela unitária pode ser feita, através de um método proposto por Matthews (1968). Tal método determina que, para a maior parte dos cristais de proteína, a razão V_M entre o volume da cela unitária e o peso molecular tem um valor entre 1.7 e 3.5 Å³/Da, sendo o mais provável em torno de 2.15 Å³/Da. O grupo espacial fornece, através de sua multiplicidade, os números possíveis de moléculas presentes na cela unitária e, conhecendo-se o peso molecular da proteína, calcula-se os possíveis valores do chamado número de Matthews (V_M). A fração do volume da cela unitária ocupada pelas moléculas de proteína é calculada pela razão $V_{prot} = \text{volume específico da proteína} / V_M \times N$, onde N é o número de Avogadro. O volume específico de uma proteína é aproximadamente 0.74 cm³/g e a partir desse valor e do número de Matthews, pode-se calcular a porcentagem de solvente presente na cela unitária do cristal: $V_{solv} = 1 - V_{prot}$.

Considerando-se que a enzima GAPDH de *T. cruzi* apresenta um peso molecular de 39 000 Da por monômero e é ativa como tetrâmero, foram calculados o conteúdo das celas unitárias dos cristais, número de Matthews (V_M) e porcentagem de solvente na cela unitária. Os resultados são apresentados na tabela 3.3.

Tabela 3.3 - Conteúdo das células unitárias dos cristais de GAPDH tipo selvagem e mutante D210L-G213D.

	Conteúdo da célula unitária	V_M (Å ³ /Da)	Porcentagem de solvente (%)
GAPDH tipo selvagem (coleta a 277 K)	2 tetrâmeros	2.53	51
GAPDH tipo selvagem (coleta a 100 K)	2 tetrâmeros	2.27	46
GAPDH D210L-G213D (coleta a 100 K)	2 tetrâmeros	2.34	47

Os dados de difração dos cristais pertencentes ao grupo espacial I4 foram também coletados, embora a resolução fosse bastante inferior. O conjunto de dados foi processado utilizando, como anteriormente, os programas DENZO e SCALEPACK, resultando nos seguintes parâmetros da rede cristalina: $a = b = 307.93$ Å, $c = 119.83$ Å, $\alpha = \beta = \gamma = 90^\circ$. Tal conjunto de dados apresentou um R_{merge} final de 12%, completude de 41% e 3.5 Å de resolução. Apesar de sua baixa qualidade, esses dados foram utilizados para a resolução da estrutura, como será brevemente descrito no capítulo 4.

Capítulo 4

Resolução e Refinamento das Estruturas

A resolução e o refinamento de uma estrutura cristalográfica envolvem cálculos complexos e o crescente avanço da tecnologia computacional tem possibilitado o desenvolvimento de programas cada vez mais sofisticados e automatizados. No presente capítulo são apresentados os métodos utilizados para a resolução e refinamento das estruturas de GAPDH de *T. cruzi* tipo selvagem e mutante D210L-G213D. A seguir, a qualidade dos modelos finais será avaliada.

4.1 - Resolução das estruturas

4.1.1 - Introdução

A resolução da estrutura cristalográfica de determinada molécula consiste na determinação da função densidade eletrônica, que descreve a estrutura tridimensional da molécula no interior da cela unitária do cristal. Tal função pode ser escrita como a soma de Fourier:

$$\rho(x, y, z) = \frac{1}{V} \sum_{hkl} |F(hkl)| \exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)] \quad (4.1)$$

onde $|F(hkl)|$ e $\alpha(hkl)$ constituem o módulo e a fase do chamado fator de estrutura, que descreve a onda espalhada por uma cela unitária do cristal.

Em um experimento de difração, são medidas as intensidades $I(hkl)$ das reflexões, enquanto as fases são perdidas. Portanto, a partir do padrão de difração de um cristal é possível determinar o módulo $|F(hkl)|$ do fator de estrutura ($I(hkl) \propto |F(hkl)|^2$) e o problema de resolução de uma estrutura cristalográfica consiste na determinação das fases $\alpha(hkl)$ das reflexões.

Existem, basicamente, três métodos para resolução da estrutura cristalográfica de proteínas: substituição isomórfica, dispersão anômala a múltiplos comprimentos de onda e substituição molecular (Drenth, 1994). Os dois primeiros requerem a presença de átomos pesados (com número atômico elevado) incorporados ao cristal de proteína. O método da substituição isomórfica utiliza o fato de que átomos pesados contribuem mais para o espalhamento se comparados aos átomos que compõem uma proteína, basicamente C, O e N. Esse fato permite a determinação das posições dos átomos pesados, a partir das quais as posições dos demais átomos são calculadas. O método de dispersão anômala a múltiplos comprimentos de onda (MAD) baseia-se na dependência do espalhamento com o comprimento de onda, observada para elétrons das camadas mais internas dos átomos, na região de sua borda de absorção. Tal efeito é mais pronunciado para átomos mais pesados e sua utilização para a resolução de uma estrutura cristalográfica requer a coleta de dados em dois ou três comprimentos de onda diferentes, em torno da borda de absorção do espalhador anômalo utilizado.

O método de substituição molecular será descrito mais detalhadamente a seguir, já que foi o método utilizado no presente trabalho. Tal método, considerado o mais simples e direto, somente poderá ser aplicado se já for conhecida a estrutura tridimensional de uma proteína com enovelamento semelhante àquela cuja estrutura se quer determinar. A substituição molecular baseia-se no princípio de que proteínas sequencialmente homólogas devem apresentar enovelamento similar e, portanto, as coordenadas atômicas da estrutura conhecida podem fornecer um conjunto inicial de fases para a estrutura que se deseja determinar. O problema passa a ser então a transferência da molécula cuja estrutura é conhecida (molécula modelo) para o arranjo cristalino da molécula de estrutura desconhecida. O posicionamento da molécula modelo na nova cela unitária é feito em duas etapas: uma de rotação, quando a orientação da molécula modelo em relação à molécula desconhecida é determinada e outra de translação, quando as moléculas, já corretamente orientadas, são sobrepostas. Os cálculos das funções de rotação e translação são feitos com base na chamada função de Patterson, cuja definição será apresentada a seguir.

A função de Patterson $P(\mathbf{u})$ ou $P(uvw)$ é definida como:

$$P(\mathbf{u}) = \int_V \rho(\mathbf{r}) \times \rho(\mathbf{r} + \mathbf{u}) \, dV \quad (4.2)$$

onde a integração é feita sobre o volume da cela unitária.

Tal função também pode ser escrita como uma soma de Fourier cujos coeficientes são os módulos do fator de estrutura ao quadrado e fases iguais a zero:

$$P(uvw) = \frac{1}{V} \sum_{hkl} |F(hkl)|^2 \cos[2\pi(hu + kv + lw)] \quad (4.3)$$

Como $\rho(\mathbf{r})$ é uma função atômica, ou seja, com picos bem definidos centrados nas posições de cada átomo, $P(\mathbf{u})$ somente assumirá valores diferentes de zero quando os vetores \mathbf{u} forem iguais a vetores interatômicos. Dessa forma, as coordenadas atômicas não são conhecidas, mas as distâncias vetoriais entre os átomos são obtidas através do mapa de Patterson. Se os pares de átomos pertencem à mesma molécula, os vetores de Patterson correspondentes (vetores de auto-Patterson) são relativamente curtos, enquanto vetores intermoleculares (vetores cruzados de Patterson) resultam em pontos mais distantes da origem no mapa de Patterson.

No método de substituição molecular, a chamada função de rotação sobrepõe as funções de Patterson das estruturas conhecida e desconhecida, considerando apenas os vetores de auto-Patterson. De fato, se apenas vetores *intramoleculares* estão presentes, os mapas de Patterson de duas moléculas homólogas serão muito similares, exceto por uma rotação. A função de rotação é definida como:

$$R(\alpha, \beta, \gamma) = \int_U P(\mathbf{u}) \times P_r(\mathbf{u}_r) \, d\mathbf{u} \quad (4.4)$$

onde $P(\mathbf{u})$ é a função de Patterson da molécula desconhecida e $P_r(\mathbf{u}_r)$ é a função de Patterson da molécula modelo, submetida a uma rotação. A integração é feita sobre o volume U que contém os vetores de auto-Patterson (esfera mínima que contém a molécula modelo).

A função de rotação dependerá dos ângulos (α, β, γ) e terá seu valor máximo correspondente à orientação correta das funções de Patterson. As funções de Patterson $P(\mathbf{u})$ e $P_r(\mathbf{u}_r)$ são expressas como somas de Fourier:

$$P(\mathbf{u}) = 1/V \sum |F(\mathbf{h})|^2 \exp[-2\pi i \mathbf{h} \cdot \mathbf{u}] \quad (4.5a)$$

$$P_r(\mathbf{u}_r) = 1/V \sum |F(\mathbf{h}^*)|^2 \exp[-2\pi i \mathbf{h}^* \cdot [\mathbf{C}]\mathbf{u}] \quad (4.5b)$$

onde $[\mathbf{C}]$ é a matriz de rotação que relaciona os vetores \mathbf{u} e \mathbf{u}_r .

Nos programas atuais que aplicam o método de substituição molecular, as funções de Patterson e a função de rotação são geralmente expressas em termos de harmônicos esféricos e não como somas de Fourier, devido à redução na quantidade dos cálculos.

Após encontrado o conjunto de ângulos (α, β, γ) que determina o valor máximo da função de rotação, passa-se à segunda etapa do método de substituição molecular, a busca por um vetor translação que posicione corretamente a molécula modelo dentro da nova cela unitária. A chamada função de translação é calculada, fornecendo a correlação entre os vetores cruzados de Patterson da molécula modelo e a função de Patterson observada. Os vetores cruzados são aqueles originados de vetores entre átomos de duas moléculas relacionadas por simetria (cristalográfica ou não).

A função de translação é dada por:

$$T(\mathbf{t}) = \int_V P_{1,2}(\mathbf{u}, \mathbf{t}) \times P(\mathbf{u}) \, d\mathbf{u} \quad (4.6)$$

onde $P_{1,2}(\mathbf{u}, \mathbf{t})$ é a função de Patterson, já orientada pelos ângulos de rotação, que leva em conta os vetores cruzados entre duas moléculas modelo relacionadas por simetria e $P(\mathbf{u})$ é a função de Patterson observada. A função $T(\mathbf{t})$ apresentará um máximo quando o vetor \mathbf{t} corresponder a um vetor intermolecular presente no mapa de Patterson observado.

Na busca dos valores máximos das funções de rotação e translação, são calculados fatores de correlação para as diferentes posições da molécula modelo ao longo da nova cela unitária. O fator de correlação é dado por:

$$C = \frac{\sum_{hkl} (|F_{obs}(hkl)|^2 - \overline{|F_{obs}(hkl)|^2}) \times (|F_{calc}(hkl)|^2 - \overline{|F_{calc}(hkl)|^2})}{\left[\sum_{hkl} (|F_{obs}(hkl)|^2 - \overline{|F_{obs}(hkl)|^2})^2 + \sum_{hkl} (|F_{calc}(hkl)|^2 - \overline{|F_{calc}(hkl)|^2})^2 \right]^{1/2}} \quad (4.7)$$

O método de substituição molecular terá encontrado uma solução para o problema das fases quando um conjunto de ângulos de rotação e vetor translação fornecer um fator de correlação que contraste significativamente com os demais.

4.1.2 - Resolução das estruturas de GAPDH tipo selvagem e mutante D210L-G213D

Como mencionado anteriormente, as estruturas de GAPDH tipo selvagem e mutante D210L-G213D foram resolvidas pelo método de substituição molecular. Foi utilizado como modelo de busca o tetrâmero da mesma enzima cristalizada no grupo espacial P1 (Souza, 1996; Souza *et al.*, 1998) e procurava-se também um tetrâmero na unidade assimétrica da estrutura desconhecida. O método foi aplicado também para a coleta de dados a 100 K e para o mutante devido à grande mudança ocorrida na cela unitária em relação aos parâmetros obtidos para os cristais da enzima tipo selvagem com dados coletados a 277 K.

As funções de rotação e translação foram calculadas utilizando o pacote de programas AMoRe (Navaza, 1994). Os arquivos de entrada para o programa contêm: coordenadas atômicas e fatores de temperatura da molécula modelo, parâmetros cristalinos da estrutura que se deseja determinar e o conjunto de dados coletados com índices das reflexões (hkl) e respectivos $|F_{obs}|$. O pacote contém os seguintes programas principais:

- 1) *sorting*: agrupa e ordena a lista de reflexões que serão usadas nos demais programas; é dependente da cela unitária, grupo espacial e limites de resolução utilizados.
- 2) *tabling*: posiciona a molécula modelo no sistema de referência utilizado por AMoRe e calcula os coeficientes de Fourier a partir de suas coordenadas. A origem do sistema de referência para AMoRe corresponde ao centro de massa da molécula e seus eixos principais de inércia são orientados paralelos aos do sistema referencial.

- 3) *rotng*: calcula as funções de rotação. As soluções são ordenadas de acordo com os valores dos fatores de correlação para as várias orientações.
- 4) *traing*: calcula as funções de translação a partir das soluções encontradas para a função de rotação.
- 5) *fitting*: executa um refinamento de corpo rígido. Minimiza uma função da diferença entre $|F_{obs}|$ e $|F_{calc}|$ em relação os fatores de temperatura, escala e parâmetros de rotação e translação.
- 6) *mr2ic*: lê as rotações e translações aplicadas por *tabling* ao modelo de entrada, bem como as rotações e translações refinadas por *fitting* e calcula as matrizes a serem aplicadas ao modelo inicial.
- 7) *cording*: aplica as matrizes de rotação e translação ao modelo inicial, posicionando-o na cela cristalina de interesse.

A tabela 4.1 mostra as soluções finais encontradas (rotação e translação) para o posicionamento da molécula modelo nas novas celas cristalinas no caso de GAPDH tipo selvagem (em ambas as temperaturas) e do mutante D210L-G213D.

Tabela 4.1 - Parâmetros de rotação e translação e fatores de correlação obtidos após o refinamento de corpo rígido por AMoRe.

	α (°)	β (°)	γ (°)	x (Å)	y (Å)	z (Å)	C (%)
GAPDH tipo selvagem coleta a 277 K	118.5	121.4	349.4	0.2878	0.0000	0.2417	75.9
	122.0	58.1	188.7	0.2123	0.0000	0.2585	75.9
	122.1	121.4	171.2	0.2878	0.0000	0.2417	75.7
	118.7	59.1	10.4	0.2123	0.0000	0.2585	76.0
GAPDH tipo selvagem coleta a 100 K	116.6	56.3	189.2	0.2024	0.0000	0.2571	65.9
	116.6	123.3	170.6	0.2976	0.0000	0.2430	66.0
	113.0	123.2	348.7	0.2976	0.0000	0.2432	65.7
	113.1	57.3	11.1	0.2025	0.0000	0.2570	65.8
GAPDH D210L-G213D	117.0	56.3	188.8	0.2053	0.0000	0.2555	66.0
	117.1	123.2	171.1	0.2947	0.0000	0.2446	65.8
	113.6	57.2	10.7	0.2054	0.0000	0.2556	65.9
	113.5	123.2	349.1	0.2948	0.0000	0.2446	65.8

Os quatro conjuntos de rotação e translação associados a cada solução estão relacionados por três eixos de simetria de ordem 2 não cristalográficos, que correspondem à simetria interna do tetrâmero. As soluções para as três estruturas não apresentaram nenhuma ambigüidade, sendo os segundos valores mais altos dos fatores de correlação da ordem de 40%.

O conjunto de dados dos cristais pertencentes ao grupo espacial I4 foi também utilizado em uma tentativa de resolução da estrutura através do programa AMoRe. O modelo de busca foi novamente o tetrâmero da mesma enzima cristalizada no grupo espacial P1 mas, devido à baixa completeza dos dados (41%), o problema de substituição molecular não apresentou uma solução trivial como nos casos anteriores. Foi necessária a modificação de alguns parâmetros do programa, fazendo-se a busca pelo máximo da função de translação a partir dos 500 primeiros picos da função de rotação. De fato, a orientação do terceiro tetrâmero localizado na unidade assimétrica correspondia ao 250° pico da função de rotação. A solução final, que posicionava três tetrâmeros na unidade assimétrica da nova cela unitária, apresentou um fator de correlação de 72.8%.

4.2 - Refinamento das estruturas

4.2.1 - Introdução

A aplicação de um dos métodos de resolução de estrutura fornece um modelo inicial para a molécula, que precisa ser refinado. Tal refinamento consiste no ajuste das coordenadas e fatores de temperatura atômicos, a fim de que o modelo calculado esteja em melhor concordância com os dados experimentais. Um método bastante antigo de ajuste de parâmetros calculados às respectivas medidas experimentais é o de mínimos quadrados, que também pode ser aplicado no refinamento de estruturas de proteínas. Métodos mais modernos são o de dinâmica molecular (Brünger *et al.*, 1987; Brünger *et al.*, 1997) e "maximum likelihood" ou máxima verossimilhança (Murshudov *et al.*, 1997).

Nos métodos que baseiam-se no princípio dos mínimos quadrados, a função a ser minimizada é o quadrado da diferença entre os módulos dos fatores de estrutura observados e calculados a partir do modelo:

$$Q = \sum_{hkl} w(hkl) (|F_{obs}(hkl)| - |F_{calc}(hkl)|)^2 \quad (4.8)$$

O mínimo da função é encontrado variando-se os parâmetros atômicos u_j (coordenadas e fatores de temperatura) e fazendo-se $\partial Q / \partial u_j = 0$. Para uma estrutura de proteína, a razão entre o número de observações e o número de parâmetros a serem refinados é pequena, sendo necessária a introdução de termos restritivos para o refinamento. Tais termos levam em conta características estereoquímicas, que devem ser respeitadas durante os vários ciclos de refinamento. O método de mínimos quadrados pode falhar quando os parâmetros do modelo inicial estão muito longe daqueles da estrutura real. Nesse caso, a estrutura pode ficar presa em um mínimo local da função de minimização, cuja barreira não é ultrapassada.

O método de dinâmica molecular apresenta um raio de convergência maior que o refinamento por mínimos quadrados e pode, em princípio, superar o problema dos mínimos locais. Este método adota um procedimento análogo ao adotado em mecânica clássica, no qual dada a energia potencial de um sistema e um conjunto de velocidades iniciais para corpos materiais, pode-se descrever o comportamento dinâmico do sistema ao longo da coordenada temporal. No presente caso, pode-se definir uma grandeza análoga à energia potencial que descreve as propriedades estereoquímicas da proteína e inclui termos dependentes das observações experimentais. A função análoga à energia potencial total do sistema é definida como a soma de dois termos:

$$E_{pot} = E_{empirica} + E_{efetiva} \quad (4.9)$$

onde o termo de energia empírica leva em conta interações estereoquímicas (energia de ligação covalente, energia de ângulos de ligação e ângulos diédricos) e interações não ligadas (energia eletrostática, de van der Waals e de ligações de hidrogênio). O termo de energia efetiva, ou termo de raios-X, descreve as diferenças entre os fatores de estrutura observados e calculados a partir do modelo e é dado por:

$$E_{efetiva} = w_x \sum_{hkl} (|F_{obs}(hkl)| - |F_{calc}(hkl)|)^2 \quad (4.10)$$

onde w_x é um peso escolhido apropriadamente para o termo de raios-X.

No método de dinâmica molecular a energia potencial total $E_{empirica} - E_{efetiva}$ é minimizada durante o refinamento. O método, também conhecido como "simulated annealing", simula o comportamento dinâmico do sistema atribuindo velocidades iniciais aleatórias aos átomos, caracterizadas por uma certa temperatura. No tempo $t = 0$, os átomos estão em uma configuração inicial e a molécula apresenta um certa energia potencial total E_{pot} . Cada átomo i na posição r_i sofre uma força que é a derivada da energia potencial $F = -\partial E_{pot}/\partial r_i$ e que produz uma aceleração sobre o átomo. Após um curto intervalo de tempo, o processo é repetido com os átomos nas novas posições. Um grande número de passos é necessário até que a energia potencial mínima seja atingida. A idéia básica do método consiste em atribuir um valor elevado para a temperatura inicial do sistema, para que barreiras de energia possam ser vencidas e em seguida resfriá-lo lentamente, até que se atinja o mínimo de energia.

Um método recentemente implementado nos programas de refinamento de estruturas de proteínas é o de máxima verossimilhança, uma generalização do método de mínimos quadrados. No refinamento por máxima verossimilhança os fatores de estrutura não são manipulados como valores, mas como distribuições de probabilidade cujas variâncias podem representar imperfeições do modelo. O método permite o refinamento de parâmetros que modulam as variâncias, parametrizando, assim, as incertezas do modelo.

O método de mínimos quadrados assume que a distribuição de cada $|F_{obs}|$ quando o modelo é conhecido, seja uma gaussiana com valor esperado $|F_{calc}|$ e incertezas conhecidas. Em muitas situações experimentais, o erro entre os valores observados e os calculados seguem uma distribuição de probabilidade gaussiana. Para que isso ocorra é necessário que exista um número suficientemente grande de fontes independentes de erro e que nenhuma delas domine. Entretanto, em cristalografia, a relação entre $|F_{obs}|$ e $|F_{calc}|$ é raramente gaussiana, o que torna o método de mínimos quadrados freqüentemente não apropriado. Isso ocorre porque muitas fontes de erro dos dados cristalográficos têm seu efeito no plano complexo, mas os valores observados são apenas as amplitudes dos fatores de estrutura. Assim, a média sobre

todas as fases possíveis para a obtenção das amplitudes transforma a distribuição de erros de gaussiana para não gaussiana.

O método de máxima verossimilhança não assume uma distribuição de erros gaussiana mas, ao contrário, estima uma distribuição de probabilidade condicional dos dados experimentais quando o modelo é conhecido. Assumindo que existem dados experimentais $|F_{obs}|$ e parâmetros a serem estimados x , o teorema de Bayes, que exprime a relação entre duas probabilidades condicionais, pode ser escrito como:

$$P(x; |F_{obs}|) = p(x)P(|F_{obs}|; x) / P(|F_{obs}|) = p(x)L(x; |F_{obs}|) \quad (4.11)$$

onde $P(x; |F_{obs}|)$ é a distribuição de probabilidade dos parâmetros quando os dados experimentais são conhecidos; $p(x)$ é a distribuição de probabilidade "a priori" dos parâmetros, conhecida antes do experimento e $L(x; |F_{obs}|)$ é a função de verossimilhança, proporcional à distribuição condicional dos dados experimentais quando os parâmetros são conhecidos.

O refinamento dos parâmetros x é feito através da maximização da distribuição P e, para isso, deve-se conhecer as formas da distribuição "a priori" e da função de verossimilhança. Nas equações subseqüentes a notação $|F_{obs}|$ será substituída por $|F^o|$ e os parâmetros x serão descritos por F^c , já que os fatores de estrutura calculados podem ser obtidos diretamente de x . Assim, a função de verossimilhança será escrita como $L(F^c; |F^o|)$.

A distribuição de probabilidade "a priori" incorpora restrições estereoquímicas, ou seja, contém informações a respeito de comprimentos e ângulos de ligação, ângulos torsionais, etc.

Assumindo que os erros em diferentes reflexões são independentes, a distribuição de probabilidade dos dados experimentais quando o modelo é conhecido tem a seguinte forma:

$$P[(|F^o|)^{\text{todas as reflexões}}; (F^c)^{\text{todas as reflexões}}] = \prod_{hkl} P(|F^o(hkl)|; F^c(hkl)) \quad (4.12)$$

Assim, com o objetivo de descrever a função de verossimilhança, a distribuição de probabilidade de cada reflexão é gerada e essas são multiplicadas fornecendo a distribuição de probabilidade condicional conjunta.

Como a maximização de uma função é equivalente à minimização do seu logaritmo multiplicado por -1 , é interessante escrever a equação 4.11 na forma:

$$-\log P[x; [(F^o)^{\text{todas as reflexões}}]] = -\log p(x) - \sum_{hkl} \log L(F^c(hkl); |F^o(hkl)|) \quad (4.13)$$

A expressão para a distribuição de probabilidade condicional dos fatores de estrutura foi desenvolvida considerando que os erros em diferentes átomos são independentes e que a soma de variáveis independentes aleatoriamente distribuídas seguirá a mesma distribuição de probabilidade com o valor médio igual à soma dos valores médios e a variância igual à soma das variâncias (Read, 1990; Pannu; Read, 1996). Tal distribuição é dada por:

$$P[F^o; F^c] = \begin{cases} \frac{1}{\pi(2\sigma_{F^o}^2 + \sum_{wc})} \exp\left(-\frac{|F^o - F_{wc}|^2}{2\sigma_{F^o}^2 + \sum_{wc}}\right) \\ [2\pi(\sigma_{F^o}^2 + \sum_{wc})]^{1/2} \exp\left(-\frac{|F^o - F_{wc}|^2}{2(\sigma_{F^o}^2 + \sum_{wc})}\right) \end{cases} \quad (4.14)$$

onde a primeira igualdade refere-se a reflexões acêntricas, a segunda a reflexões cêntricas e define-se $F_{wc} = \sum_{j=1}^{N_{at.}} D_j F_j^c$; $\sum_{wc} = \varepsilon \sum_{j=1}^{N_{at.}} \sum_j (1 - D_j^2)$, sendo D o centróide da distribuição de F^o dados F^c . A separação dos termos de desvio para F^c , que são escritos como uma soma (\sum_{wc}), permite que diferentes partes da estrutura recebam diferentes tratamentos.

A equação 4.14 descreve a distribuição de probabilidade dos fatores de estrutura considerando seus módulos e fases. Entretanto, a relação de interesse é a distribuição das amplitudes, obtida através da integração sobre todos os valores possíveis de fases:

$$P[|F^o|; F^c] = \int_0^{2\pi} P[|F^o|, \varphi; F^c] d\varphi \quad (4.15)$$

o que resulta em,

$$P[|F^o|; F^c] = \begin{cases} \frac{2|F^o|}{2\sigma_{F^o}^2 + \sum_{wc}} \exp\left(-\frac{|F^o|^2 + |F_{wc}|^2}{2\sigma_{F^o}^2 + \sum_{wc}}\right) I_0\left(\frac{2|F^o||F_{wc}|}{2\sigma_{F^o}^2 + \sum_{wc}}\right) \\ \frac{2}{\pi(\sigma_{F^o}^2 + \sum_{wc})} \exp\left[-\frac{|F^o|^2 + |F_{wc}|^2}{2(\sigma_{F^o}^2 + \sum_{wc})}\right] \cosh\left(\frac{|F^o||F_{wc}|}{\sigma_{F^o}^2 + \sum_{wc}}\right) \end{cases} \quad (4.16)$$

onde, novamente, a primeira igualdade refere-se a reflexões acêntricas e a segunda a reflexões cêntricas.

A função logaritmo da verossimilhança (LLK) pode ser obtida a partir da equação 4.16 e será da forma:

$$\text{LLK} = \sum_{hkl} \text{LLK}(hkl) \quad (4.17)$$

onde,

$$\text{LLK}(hkl) = \begin{cases} c_a - \log(|F^o|) + \log(2\sigma_{F^o}^2 + \sum_{wc}) + \frac{|F^o|^2 + |F_{wc}|^2}{2\sigma_{F^o}^2 + \sum_{wc}} - \log I_0\left(\frac{2|F^o||F_{wc}|}{2\sigma_{F^o}^2 + \sum_{wc}}\right) \\ c_c - \frac{1}{2} \log(|F^o|) + \log(\sigma_{F^o}^2 + \sum_{wc}) + \frac{|F^o|^2 + |F_{wc}|^2}{2\sigma_{F^o}^2 + \sum_{wc}} - \log \cosh\left(\frac{|F^o||F_{wc}|}{\sigma_{F^o}^2 + \sum_{wc}}\right) \end{cases} \quad (4.18)$$

Finalmente, o refinamento é realizado através da minimização de $-\text{LLK}$. O procedimento de minimização utiliza, como usualmente, as derivadas da função em relação aos seus parâmetros. No método de máxima verossimilhança, as incertezas \sum_{wc} também são refinadas.

Um índice utilizado como indicador da convergência do refinamento por qualquer um dos métodos acima apresentados é o chamado fator R cristalográfico, ou R_{factor} , que avalia a concordância entre os fatores de estrutura observados e calculados:

$$R_{factor} = \sum_{hkl} \frac{\| |F_{obs}| - k|F_{calc}| \|}{\sum_{hkl} |F_{obs}|} \times 100\% \quad (4.19)$$

Os valores finais de R_{factor} para estruturas de proteínas geralmente variam entre 15 e 20%, dependendo da qualidade do conjunto de dados de difração.

Um outro índice muito importante a ser calculado durante o refinamento e que funciona como um parâmetro para avaliação da qualidade do modelo é o chamado R_{free} (Brünger, 1992). Para seu cálculo, 5 a 10% das reflexões são retiradas aleatoriamente do conjunto de dados, não sendo utilizadas durante o refinamento. A expressão para o R_{free} é a mesma que aquela para o fator R cristalográfico, levando-se em conta somente as reflexões retiradas do refinamento. Dessa forma, caso a estrutura esteja sendo refinada erroneamente, o valor de R_{free} aumentará embora R_{factor} possa apresentar um valor aceitável.

4.1.2 - Refinamento das estruturas de GAPDH tipo selvagem e mutante D210L-G213D

As estruturas de GAPDH tipo selvagem (coletas de dados a 100 e 277 K) e mutante D210L-G213D foram refinadas utilizando o programa REFMAC (Murshudov *et al.*, 1997) do pacote CCP4 (Collaborative Computational Project Number 4, 1994). Fatores R cristalográficos para os modelos iniciais obtidos por substituição molecular mostraram-se bastante baixos, em decorrência da utilização de um modelo de busca com identidade seqüencial de 100%. A resolução máxima dos conjuntos de dados foi utilizada em todos os ciclos de refinamento, sendo 10% das reflexões selecionadas para o cálculo do R_{free} .

O processo de refinamento das três estruturas seguiram os mesmos passos. Os quatro monômeros foram submetidos a um refinamento de corpo rígido, no qual somente três parâmetros de rotação e três de translação são otimizados para cada um deles. Prosseguiu-se com o cálculo das três operações de simetria não cristalográficas (SNC) que relacionam os quatro monômeros (três eixos de ordem 2) e a partir de então os demais ciclos de refinamento foram conduzidos utilizando SNC restrita entre as subunidades. Quatro moléculas de NAD^+ (uma por monômero) foram incorporadas ao modelo e os ciclos subseqüentes de refinamento utilizando o método de máxima

verossimilhança possibilitaram a construção de mapas de densidade eletrônica de boa qualidade. Tais mapas, calculados através de programas do pacote CCP4, eram resultado da média da densidade eletrônica dos quatro monômeros.

A análise dos mapas de densidade eletrônica promediados $2F_{obs}-F_{calc}$ e $F_{obs}-F_{calc}$ permitiu o ajuste manual de algumas cadeias laterais utilizando o programa gráfico O (Jones *et al.*, 1991; Jones, 1996). No caso da estrutura do mutante GAPDH D210L-G213D as mutações foram introduzidas no modelo nessa etapa. O programa ARP (Lamzin; Wilson, 1993), que posiciona moléculas de água automaticamente na estrutura, foi utilizado nos últimos ciclos de refinamento. A introdução das moléculas de água era posteriormente verificada através da análise dos mapas de densidade eletrônica.

Os modelos finais das estruturas contêm 359 resíduos e uma molécula de NAD^+ por monômero, além das moléculas de água, 111 para a estrutura de GAPDH tipo selvagem a 277 K e 300 para as estruturas de GAPDH tipo selvagem a 100 K e mutante D210L-G213D. O primeiro e dois últimos resíduos de cada monômero não apresentavam densidade eletrônica bem definida para suas cadeias laterais e foram substituídos por alaninas. Os parâmetros finais dos refinamentos são apresentados na tabela 4.2.

A resolução mais alta dos mapas de densidade eletrônica obtidos a partir dos dados coletados a 100 K permitiu a inclusão de um número maior de moléculas de água nos modelos. Além disso, o resfriamento dos cristais pode ter contribuído para o aumento do número de moléculas de água ordenadas. Dentre as 300 moléculas incluídas, 204 ocupam a mesma posição relativa nas estruturas de GAPDH tipo selvagem e mutante D210L-G213D. Por outro lado, das 111 moléculas de água encontradas no mapa de densidade eletrônica da enzima tipo selvagem obtido a partir da coleta de dados a 277 K, 97 estão presentes também na estrutura a 100 K.

Tabela 4.2 - Parâmetros finais dos refinamentos.

	GAPDH tipo selvagem coleta a 277 K	GAPDH tipo selvagem coleta a 100 K	GAPDH D210L-G213D coleta a 100 K
R_{factor} (%)	16.0	18.8	18.9
R_{free} (%)	21.9	24.0	24.3
Desvio médio quadrático para ângulos de ligação (°)	2.8	2.1	1.9
Desvio médio quadrático para comprimentos de ligação (Å)	0.018	0.016	0.015

O modelo da estrutura de GAPDH de *T. cruzi* obtido a partir dos dados dos cristais pertencentes ao grupo espacial I4 foi refinado até uma certa etapa. O promédio do mapa de densidade eletrônica dos 12 monômeros (3 tetrâmeros) permitiu que o problema de baixa completude dos dados fosse em parte superado, mas a precariedade do conjunto de dados impediu que esse modelo fosse utilizado em análises estruturais mais detalhadas. O fator R cristalográfico final obtido foi de 25% e, embora não tenha se podido prosseguir com a análise estrutural, esse trabalho constituiu-se em um problema cristalográfico interessante, apresentado em um congresso internacional (Guimarães *et al.*, 1996).

4.3 - Avaliação da qualidade dos modelos

4.3.1 - Introdução

Uma importante etapa no processo de determinação da estrutura cristalográfica de uma proteína é a avaliação da qualidade do modelo, que mostrará a confiabilidade da estrutura refinada. Algumas estruturas refinadas parcialmente ou inteiramente erradas já foram descritas na literatura, reforçando a necessidade de uma

cuidadosa avaliação dos modelos cristalográficos. Como exemplo, pode-se citar a estrutura da ribulose-1,5-bisfosfato carboxilato oxigenase (Chapman, 1988).

Tal avaliação começa com a análise do mapa de densidade eletrônica, verificando-se sua conectividade nas regiões de cadeias principal e laterais e interpretação da conformação de cadeias laterais. Um parâmetro importante na análise do mapa de densidade eletrônica é o seu nível de contorno, que leva em conta o erro associado ao mapa. Quanto mais alto o nível de contorno ($n\sigma$) do mapa, menor o seu ruído. Os índices residuais R_{factor} e R_{free} definidos anteriormente também são indicativos da qualidade do modelo, mas avaliam a estrutura como um todo e muitas vezes não acusam problemas locais. Uma análise da estereoquímica do modelo, distribuição dos fatores de temperatura e ambientes químicos dos resíduos deve ser também realizada.

Vários são os parâmetros estereoquímicos a serem verificados. O enovelamento da cadeia principal pode ser analisado através do diagrama de Ramachandran (Ramachandran *et al.*, 1963), no qual os ângulos Φ e Ψ para cada resíduo são representados em uma matriz quadrada. Contatos desfavoráveis entre átomos de resíduos adjacentes impedem que os ângulos Φ e Ψ assumam qualquer valor entre -180° e $+180^\circ$, dando origem a regiões não permitidas no diagrama de Ramachandran. Outros parâmetros analisados são comprimentos e ângulos de ligação, planaridade das ligações peptídicas e de algumas cadeias laterais, interações não covalentes, etc. Uma comparação é feita entre a estereoquímica da estrutura a ser verificada e um conjunto padrão de parâmetros derivados de estruturas de proteínas determinadas a alta resolução e estruturas de pequenas moléculas.

Fatores de temperatura elevados são esperados em regiões flexíveis da molécula, por exemplo, cadeias laterais de resíduos expostos e, muitas vezes, nas regiões N e C-terminal. Caso contrário, podem indicar problemas locais na estrutura.

O ambiente químico dos resíduos pode ser analisado através do método de Eisenberg (Bowie *et al.*, 1991). Nesse método, o chamado perfil 3D é comparado com a seqüência de aminoácidos da proteína. Tal perfil é obtido da seguinte maneira: a cada resíduo é associada uma entre seis classes de ambientes, de acordo com a área do resíduo enterrada e a fração da área da cadeia lateral coberta por átomos polares. Cada uma dessas seis classes de ambientes de cadeia lateral é subdividida em outras

três: hélices α , folhas β e outras, o que resulta em 18 classes. A partir de um conjunto de estruturas de proteínas bem refinadas, uma matriz padrão é construída fornecendo uma determinada pontuação para cada tipo de resíduo em cada uma das 18 classes. Tal matriz indica a preferência de cada um dos 20 aminoácidos pelas diferentes classes de ambientes. Por exemplo, é raro encontrar um resíduo carregado enterrado em um ambiente apolar, o que corresponderia a uma baixa pontuação. Assim, determinando a classe de ambiente para uma dada posição na estrutura, é possível atribuir uma pontuação para cada um dos 20 aminoácidos naquela posição, o que fornece as chamadas pontuações 3D-1D para toda a cadeia polipeptídica. O valor global para a compatibilidade entre o modelo e a seqüência é dado pela soma das pontuações de todos os resíduos da cadeia. Uma estrutura correta deve apresentar uma alta pontuação. Além do valor global, um perfil em função do número do resíduo pode ser determinado, com o objetivo de analisar-se problemas locais na estrutura. Tal perfil é calculado a partir da média das pontuações 3D-1D para 21 resíduos de uma janela centrada em um resíduo particular.

Um outro método de análise da vizinhança dos resíduos foi desenvolvido por Vriend e Sander (1993). Os 20 tipos de aminoácidos são divididos em 80 fragmentos. Tais fragmentos são escolhidos de maior tamanho possível de modo a não conterem graus de liberdade torsionais. Os átomos são agrupados em 57 classes, por exemplo, $O_{\epsilon 1}$ e $O_{\epsilon 2}$ do ácido glutâmico são agrupados com $O_{\delta 1}$ e $O_{\delta 2}$ do ácido aspártico em uma classe de átomos. Calcula-se o chamado índice de qualidade de contatos, uma medida da concordância entre as distribuições de átomos em torno de cada fragmento no modelo e distribuições equivalentes obtidas de um banco de dados de estruturas determinadas a alta resolução. Quanto maior a concordância, maior o índice de qualidade. Um índice médio global para a estrutura é calculado, bem como um perfil em função do número do resíduo, no qual leva-se em conta um valor médio dentro de uma janela de 10 resíduos.

4.3.2 - Avaliação das estruturas de GAPDH tipo selvagem e mutante D210L-G213D

Os valores obtidos para os índices R_{factor} e R_{free} , bem como para os desvios quadráticos médios de comprimentos e ângulos de ligação apresentados na tabela 4.2 mostraram-se bastante satisfatórios para as três estruturas.

A presença dos eixos de simetria não cristalográficos relacionando os quatro monômeros permitiu o promédio dos mapas de densidade eletrônica, que foram analisados em vários níveis de contorno. As estruturas de GAPDH tipo selvagem com coletas de dados a 277 e 100 K e do mutante D210L-G213L apresentaram mapas de densidade eletrônica de boa qualidade, como pode ser visto nas figuras 4.1, 4.2 e 4.3. No caso do mutante D210L-G213L foi possível distinguir a densidade eletrônica para os resíduos mutados.

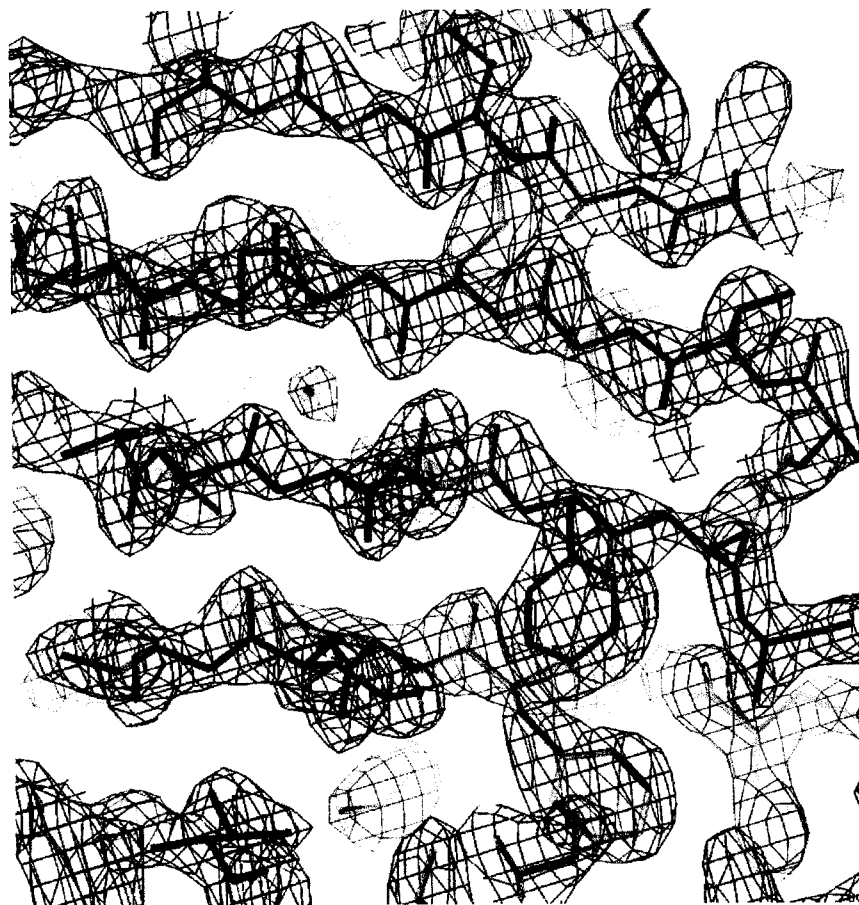


Figura 4.1 - Região de fitas β na estrutura de GAPDH tipo selvagem, coleta de dados a 277 K. Mapa de densidade eletrônica $2F_{obs}-F_{calc}$ promediado e contornado a 2.8σ .

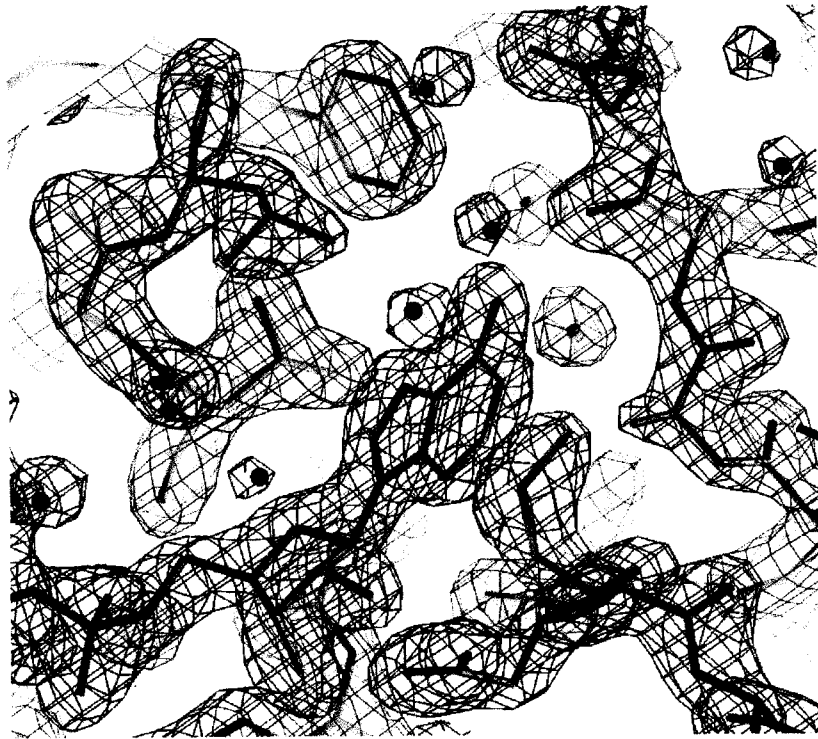


Figura 4.2 - Região do sítio de ligação do cofator NAD^+ na estrutura de GAPDH tipo selvagem, coleta de dados a 100 K. Mapa de densidade eletrônica $2F_{obs}-F_{calc}$ promediado e contornado a 2.8σ .

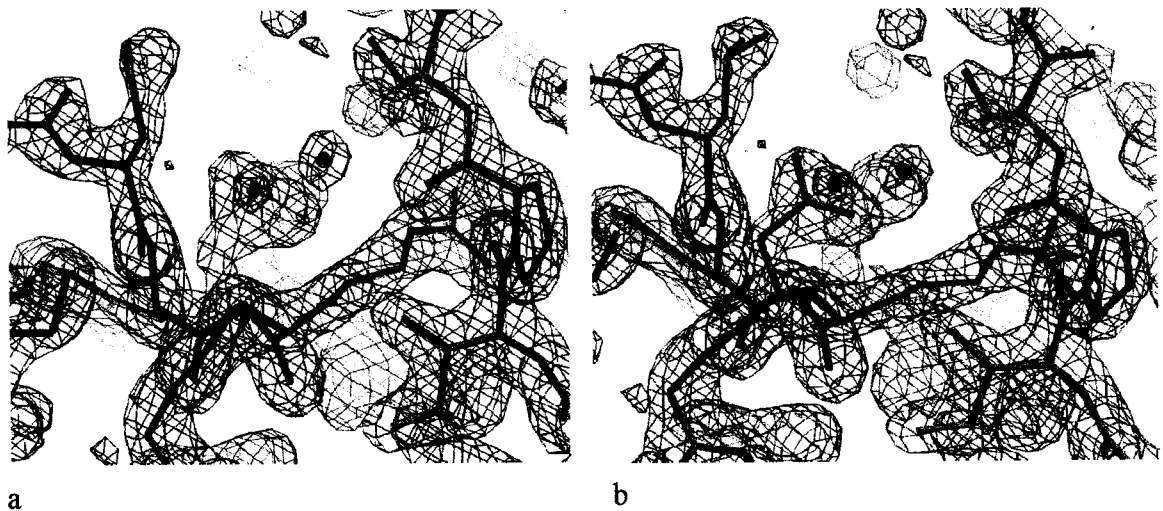


Figura 4.3 - Região das mutações na estrutura de GAPDH D210L-G213D; a) em magenta os resíduos mutados, evidenciando a falta de ajuste à densidade eletrônica; b) mesma vista após a substituição dos resíduos. Mapas de densidade eletrônica $2F_{obs}-F_{calc}$ promediados e contornados a 2.8σ .

Os parâmetros estereoquímicos das estruturas foram analisados utilizando-se o programa PROCHECK (Laskowski *et al.*, 1993). De uma maneira geral, a estereoquímica dos modelos de GAPDH tipo selvagem (coletas de dados a 277 e 100 K) e mutante D210L-G213D mostrou-se satisfatória, alguns parâmetros apresentando valores até mesmo superiores aos encontrados em estruturas à mesma resolução. Os diagramas de Ramachandran para o monômero A das três estruturas são mostrados na figura 4.4. Os gráficos dos demais monômeros são praticamente idênticos.

Nas três estruturas, os resíduos Val255 e Thr283 aparecem em regiões desfavoráveis do diagrama de Ramachandran. O resíduo Val255 é encontrado também em região não permitida do diagrama em estruturas de GAPDH de *T. brucei*, *B. stearothermophilus*, *L. mexicana*, entre outras. Verifica-se que a conformação desfavorável de tal resíduo é necessária para a estabilização dos resíduos catalíticos Cys166 e His194 (Souza *et al.*, 1998). O resíduo Thr283 apresenta-se, nas três estruturas, bem ajustado ao mapa de densidade eletrônica e a conformação de sua cadeia principal é estabilizada por ligações de hidrogênio.

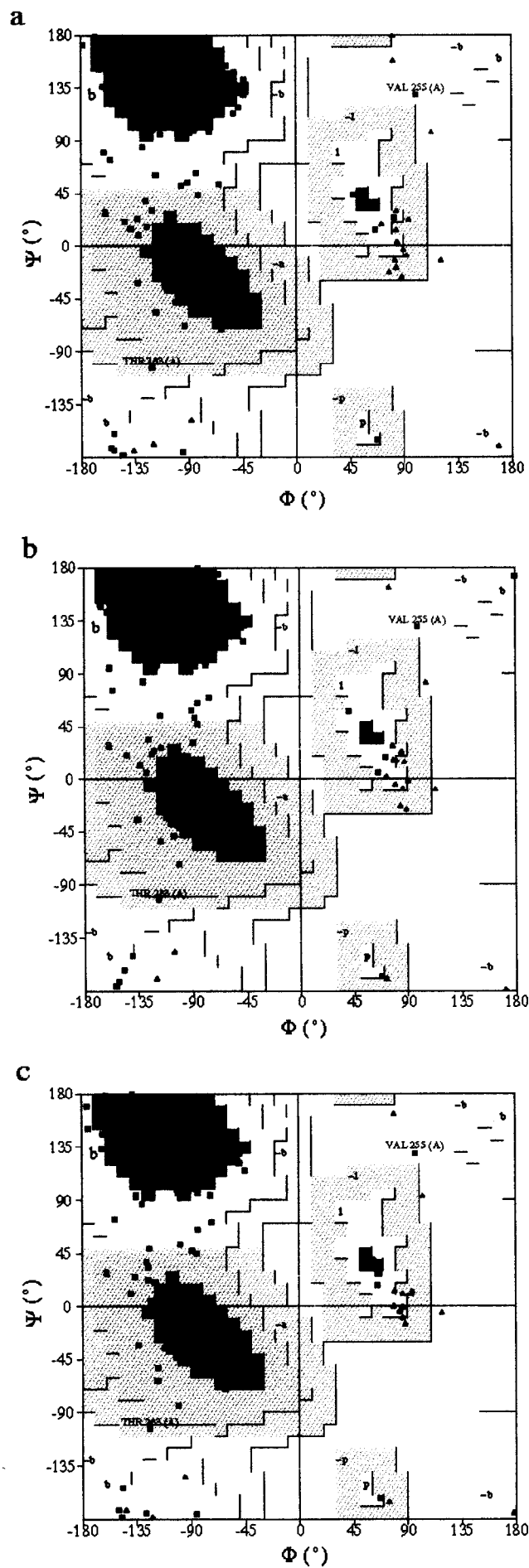


Figura 4.4 - Diagramas de Ramachandran para um dos monômeros das estruturas de a) GAPDH tipo selvagem, coleta de dados a 277 K, b) GAPDH tipo selvagem, coleta de dados a 100 K, c) mutante D210L-G213D. A cor vermelha representa as regiões mais favoráveis, amarelo corresponde às regiões permitidas ou generosamente permitidas e branco corresponde às regiões não permitidas.

O método de Eisenberg foi utilizado para análise do ambiente químico dos resíduos. As figuras 4.5, 4.6 e 4.7 apresentam o perfil 3D-1D em função do número do resíduo para os tetrâmeros das três estruturas. Nenhuma região apresentou valores negativos de pontuação, o que indicaria ambiente químico desfavorável para os resíduos correspondentes.

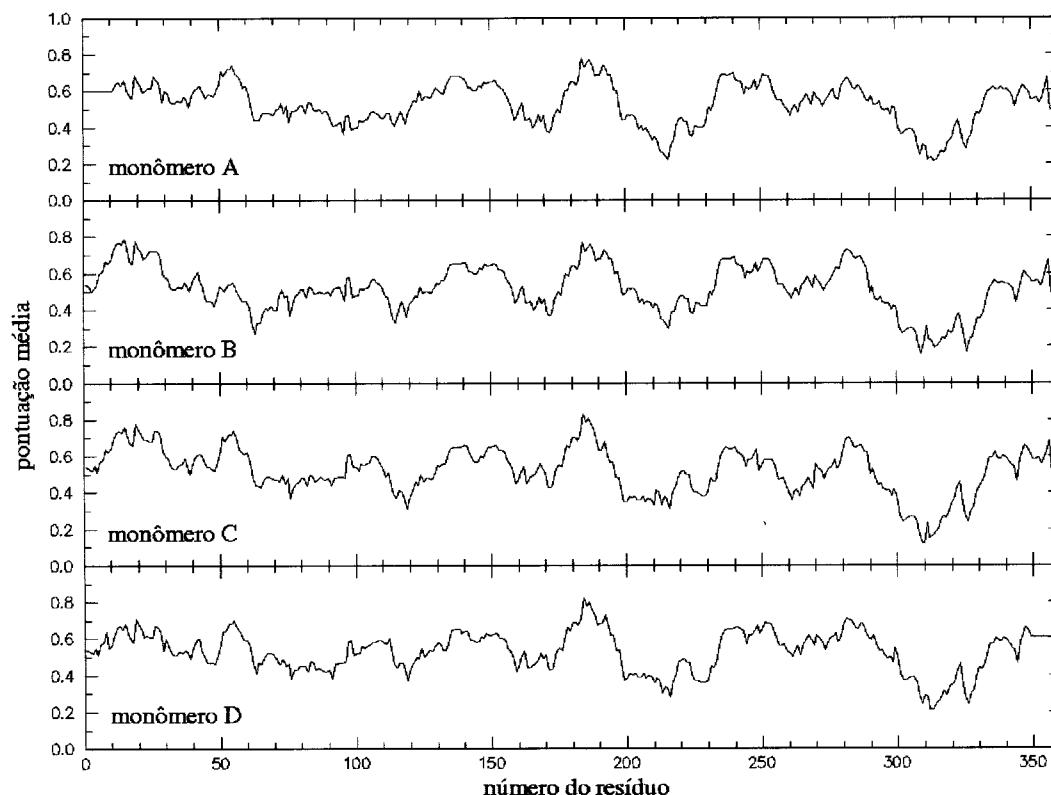


Figura 4.5 - Perfil 3D-1D para a estrutura de GAPDH tipo selvagem, coleta de dados a 277 K. A pontuação por resíduo corresponde à média das pontuações em um intervalo de 21 resíduos centrado no resíduo em questão.

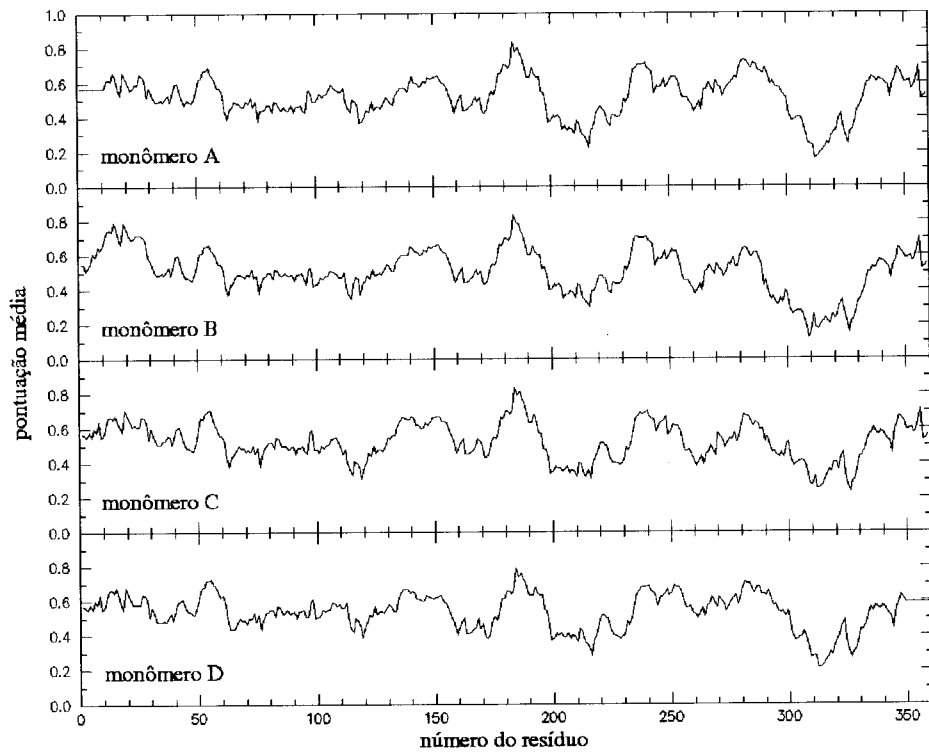


Figura 4.6 - Perfil 3D-1D para a estrutura de GAPDH tipo selvagem, coleta de dados a 100 K. A pontuação por resíduo corresponde à média das pontuações em um intervalo de 21 resíduos centrado no resíduo em questão.

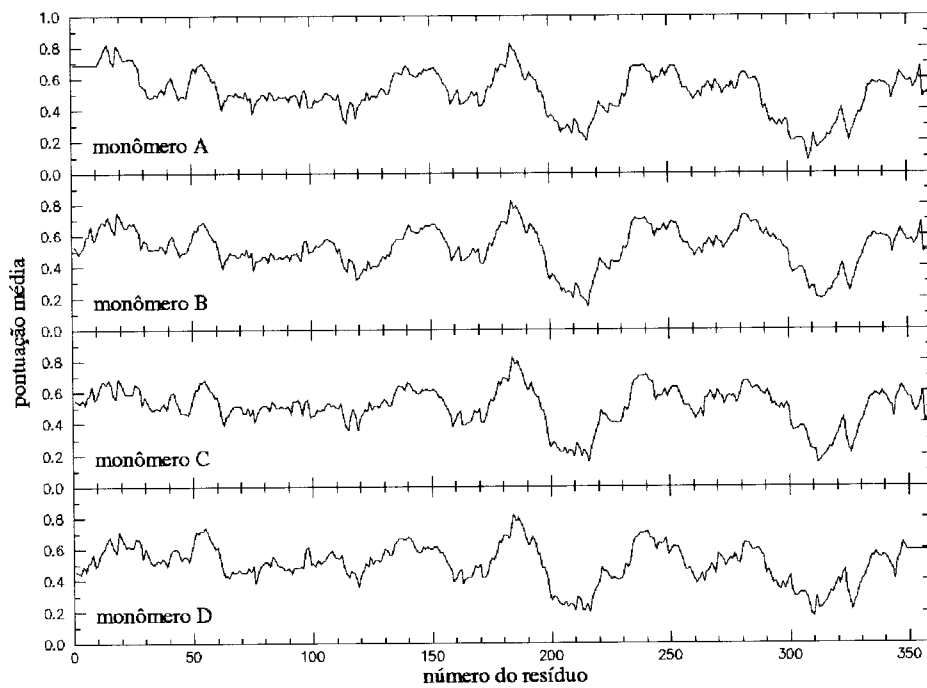


Figura 4.7 - Perfil 3D-1D para a estrutura do mutante GAPDH D210L-G213D. A pontuação por resíduo corresponde à média das pontuações em um intervalo de 21 resíduos centrado no resíduo em questão.

A análise da vizinhança dos resíduos através do índice de qualidade de contatos foi feita utilizando-se a rotina QUALITY do programa WHAT IF (Vriend, 1990; Hooft *et al.*, 1996). Os cálculos foram feitos considerando o tetrâmero como um todo, embora os gráficos sejam apresentados para os monômeros separadamente. Segundo os autores, estruturas apresentando valores entre -2.0 e 2.0 para os índices médios globais são consideradas de boa qualidade e regiões no gráfico com valores inferiores a -1.3 indicam contatos não usuais. As figuras 4.8, 4.9 e 4.10 mostram os gráficos de índices de qualidade em função do número do resíduo para as três estruturas. Os índices médios globais foram de -0.61, -0.15 e 0.03 para as estruturas de GAPDH tipo selvagem com coletas de dados a 277 e 100 K e mutante D210L-G213D, respectivamente.

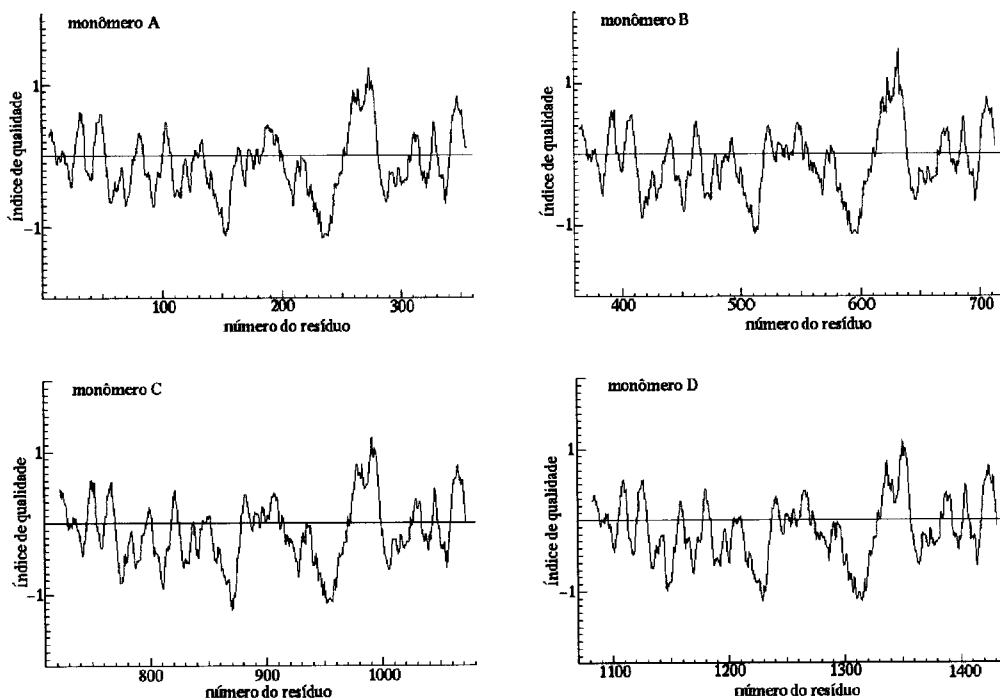


Figura 4.8 - Índice de qualidade de contatos para a estrutura de GAPDH tipo selvagem, coleta de dados a 277 K. O índice por resíduo corresponde à média dos índices em um intervalo de 10 resíduos.

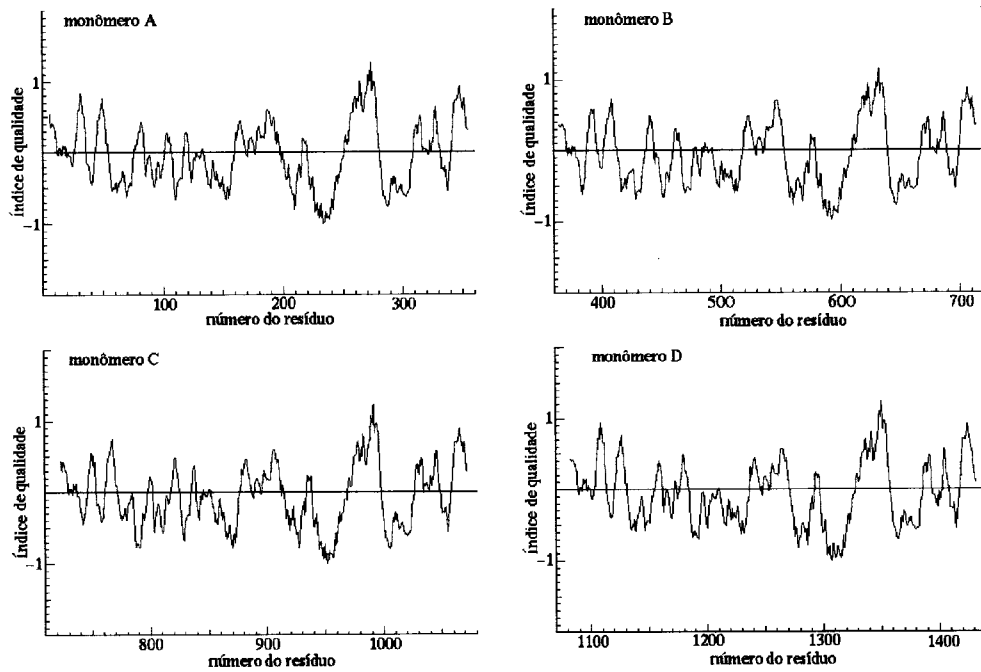


Figura 4.9 - Índice de qualidade de contatos para a estrutura de GAPDH tipo selvagem, coleta de dados a 100 K. O índice por resíduo corresponde à média dos índices em um intervalo de 10 resíduos.

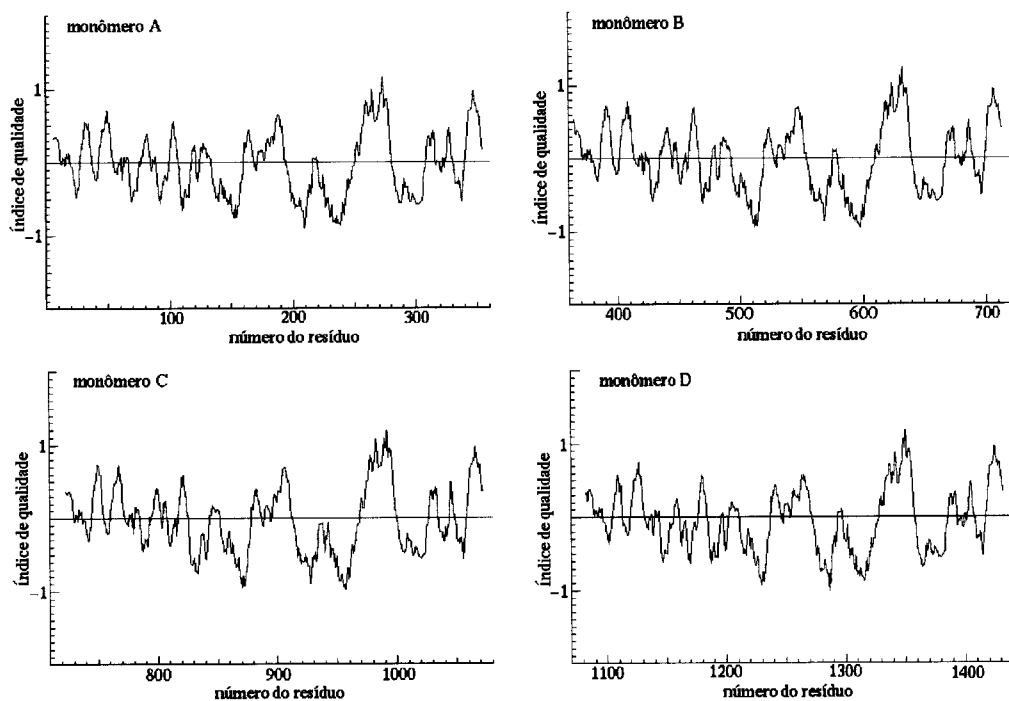


Figura 4.10 - Índice de qualidade de contatos para a estrutura do mutante GAPDH D210L-G213D. O índice por resíduo corresponde à média dos índices em um intervalo de 10 resíduos.

As distribuições dos fatores de temperatura são mostradas na figura 4.11 para as três estruturas, bem como os valores médios globais de B para cadeia principal e todos os átomos. Somente um dos monômeros é apresentado, visto que os valores para os demais são praticamente idênticos.

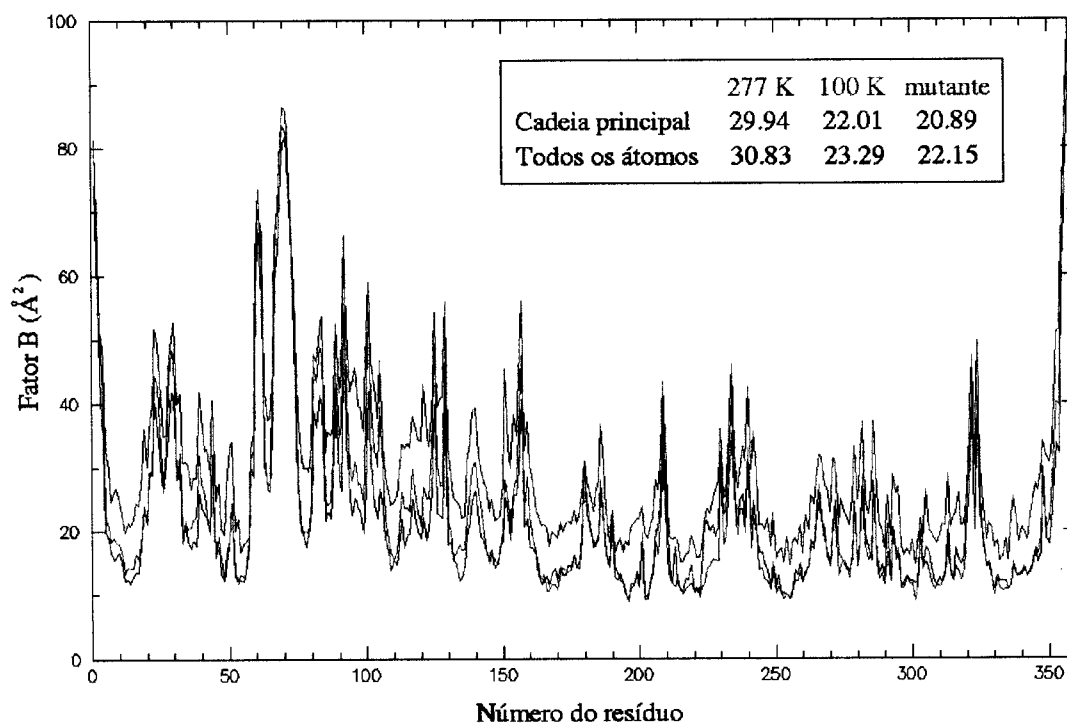


Figura 4.11 - Distribuição dos fatores de temperatura (fatores B) para o monômero A das estruturas de GAPDH tipo selvagem, coleta de dados a 277 K (em vermelho), GAPDH tipo selvagem, coleta de dados a 100 K (em verde) e mutante D210L-G213D (em azul). O valor de B por resíduo corresponde à média para todos os seus átomos e os valores médios globais da cadeia principal e todos os átomos do monômero são apresentados em destaque.

Os valores mais elevados de fatores de temperatura correspondem às regiões N e C-terminal, geralmente flexíveis por participarem de poucas interações com outros resíduos e à região que compreende os resíduos 66 a 76. Tal região forma um "loop" que encontra-se na superfície da molécula, exposto ao solvente e, portanto, bastante flexível. Observa-se que o referido "loop" ocorre somente em GAPDHs glicosomais, ou seja, de espécies da ordem Kinetoplastidae (*T. cruzi*, *T. brucei*, *L. mexicana*, etc).

Todos os critérios utilizados para a avaliação da qualidade dos modelos mostram sua confiabilidade, o que permite a realização de uma análise detalhada das estruturas.

Capítulo 5

Análise das Estruturas

A cristalografia de raios-X é uma técnica sofisticada que permite o estudo da estrutura tridimensional de moléculas a nível atômico. Apesar da determinação da estrutura de uma proteína constituir-se por si em um desafio, na maioria das vezes está longe de ser o objetivo final de um trabalho em cristalografia de proteínas. A análise da conformação de alguns resíduos permite que mecanismos de reação sejam propostos e potenciais sítios de ligação de inibidores sejam sugeridos. Neste capítulo, as estruturas de GAPDH de *T. cruzi* tipo selvagem e mutante D210L-G213D serão analisadas detalhadamente.

5.1 - A estrutura de GAPDH de *T. cruzi* a alta resolução

Como característica comum às GAPDHs de diversas espécies, a estrutura da enzima GAPDH de *T. cruzi* é tetramérica, apresentando suas quatro subunidades relacionadas por três eixos de simetria de ordem 2 (figura 5.1). Cada uma das subunidades é dividida em dois domínios, o domínio de ligação do NAD^+ , na região N-terminal e o domínio catalítico, na região C-terminal dos monômeros. A figura 5.2 apresenta um dos monômeros da estrutura de GAPDH de *T. cruzi* e a molécula de NAD^+ correspondente. Uma descrição detalhada dessa estrutura e seus domínios foi feita anteriormente (Souza, 1996).

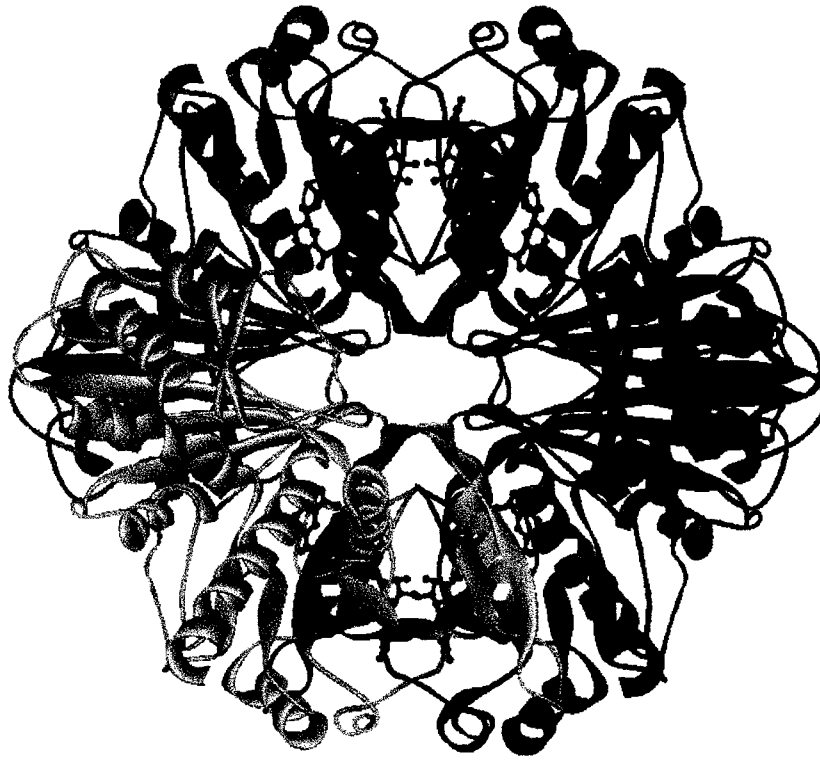


Figura 5.1 - Tetrâmero de GAPDH de *T. cruzi*. Em azul escuro estão representadas as quatro moléculas de NAD^+ .

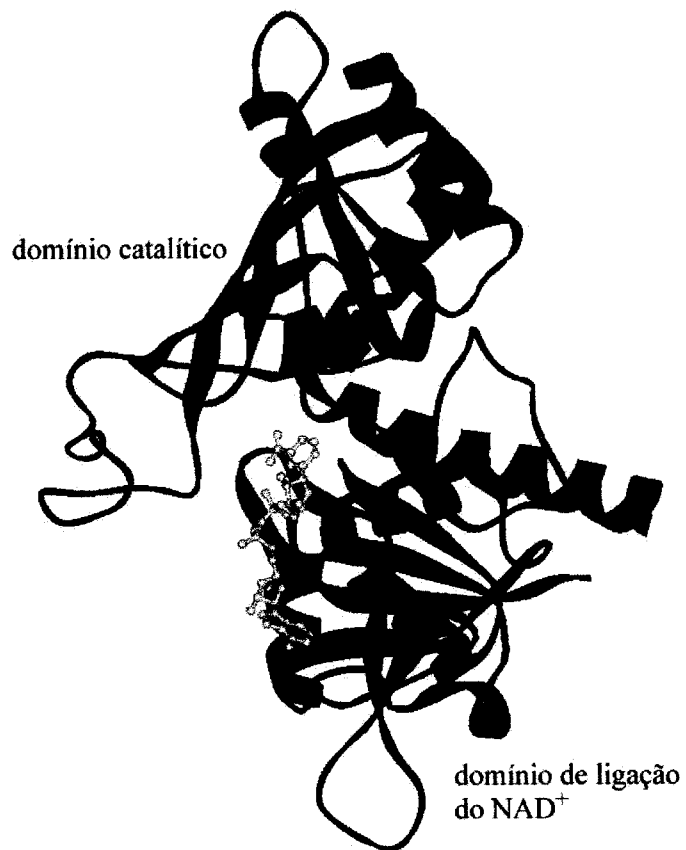


Figura 5.2 - Monômero de GAPDH de *T. cruzi*. Em amarelo está representada a molécula de NAD^+ .

A interpretação dos mapas de densidade eletrônica realizada por Souza (1996) foi dificultada pela baixa completeza e resolução dos dados obtidos (67.1% de completeza a 2.8 Å de resolução), problemas resolvidos no presente trabalho. A obtenção de conjuntos de dados a alta resolução e completeza (capítulo 3, tabela 3.2) resultou em uma melhor definição dos mapas de densidade eletrônica, permitindo que as estruturas sejam analisadas com confiabilidade até maior nível de detalhes. Regiões cujas conformações dos resíduos não puderam ser definidas de maneira satisfatória anteriormente apresentaram mapas de densidade eletrônica de boa qualidade na estrutura a 2.15 Å de resolução, como é o caso do "loop" que compreende os resíduos 66 a 76.

Entretanto, com exceção de certos problemas locais, o modelo da enzima de GAPDH de *T. cruzi* obtido anteriormente concorda com o obtido neste trabalho. A sobreposição dos monômeros das estruturas resolvidas a 2.8 e 2.15 Å de resolução fornece um desvio quadrático médio de 0.33 Å considerando os átomos da cadeia principal e 0.63 Å considerando todos os átomos. Os maiores desvios correspondem a cadeias laterais muito expostas ou envolvidas em contatos cristalinos. A conformação dos resíduos do sítio ativo nas duas estruturas mostram-se bastante coincidentes, com exceção da cadeia lateral da Arg249, fato que será discutido posteriormente. A figura 5.3 mostra um perfil de distâncias médias entre os resíduos de um monômero da estrutura resolvida por Souza (1996) (2.8 Å de resolução, grupo espacial P1) e a obtida neste trabalho (2.15 Å de resolução, grupo espacial P2₁).

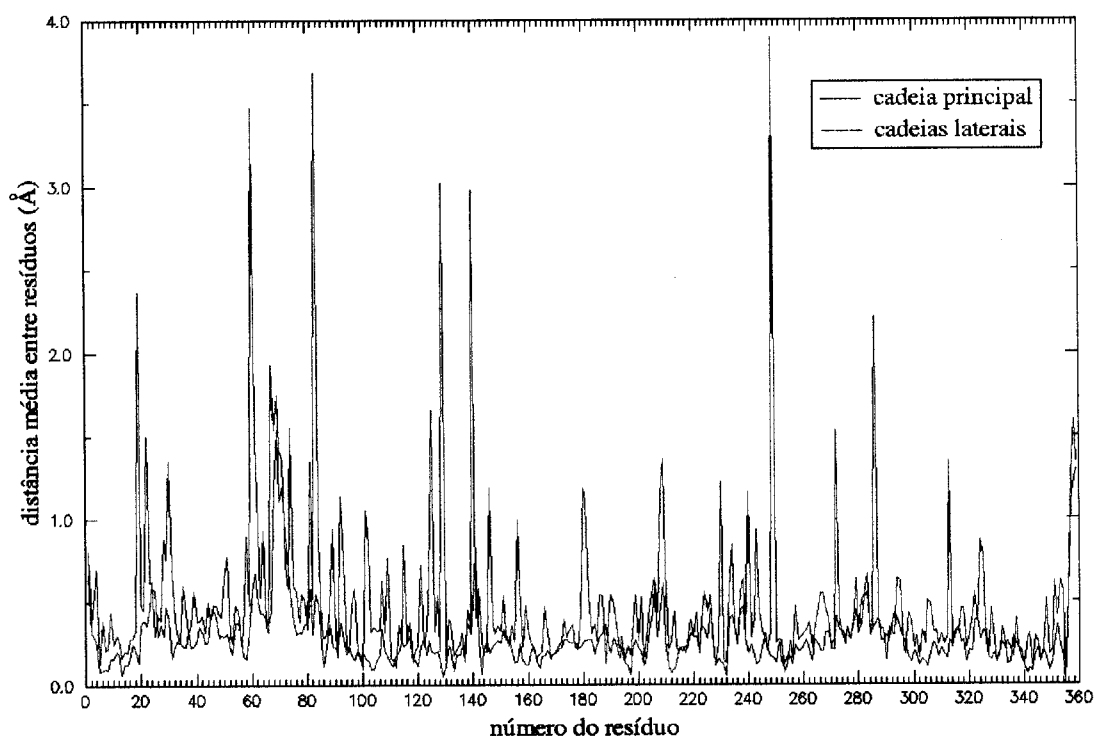


Figura 5.3 - Perfil de distâncias médias entre os resíduos de um monômero das estruturas sobrepostas de GAPDH obtidas a 2.8 Å (Souza, 1996) e 2.15 Å de resolução. O perfil em azul representa a distância média entre os átomos da cadeia principal, enquanto a representação em vermelho leva em conta somente os átomos das cadeias laterais de cada resíduo.

Uma região interessante, analisada com certo cuidado na estrutura obtida a 2.15 Å de resolução, foi a que compreende os resíduos Asp334 e Asp254. Tais resíduos encontram-se inacessíveis ao solvente, situação a princípio desfavorável para dois resíduos carregados. Além disso, formam uma ligação de hidrogênio com uma distância de apenas 2.4 Å entre seus oxigênios, indicando que um dos oxigênios estaria protonado. O programa de avaliação da estereoquímica do modelo (PROCHECK) considerou tal interação como um problema de mau contato, mas a análise do mapa de densidade eletrônica (figura 5.4) e da conformação dos resíduos vizinhos mostrou que a região era bem definida e estabilizada por várias ligações de hidrogênio. Uma situação semelhante ocorre no sítio ativo de áspartico proteases, onde a proximidade entre os átomos de oxigênio de dois ácidos aspárticos indica a presença de um próton entre eles (Pearl; Blundell, 1984). Além disso, uma interação não usual foi também encontrada entre os átomos NZ de duas lisinas (separados por 2.3 Å), na estrutura da ovotransferrina (Dewan *et al.*, 1993).

O resíduo Asp334 apresenta-se conservado em GAPDHs da grande maioria das espécies, enquanto o Asp254 aparece várias vezes substituído por uma asparagina. Apesar de tais resíduos estarem relativamente próximos do sítio ativo (distância de cerca de 8 Å para a cisteína catalítica Cys166), nenhum papel importante foi encontrado para eles.

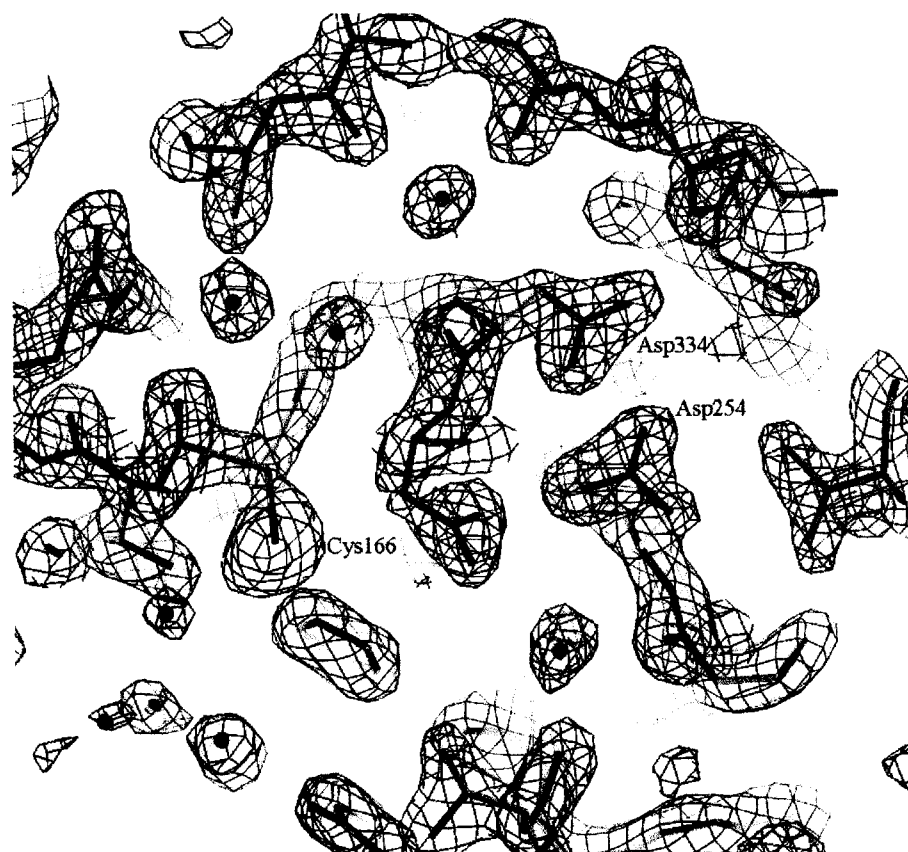


Figura 5.4 - Interação entre os resíduos Asp334 e Asp254. Mapa de densidade eletrônica $2F_{obs}-F_{calc}$ contornado a 2.8σ .

5.2 - Criocristalografia: efeito do resfriamento no empacotamento cristalino

A resolução da estrutura da enzima de GAPDH de *T. cruzi* a partir de coletas de dados realizadas em duas temperaturas diferentes (277 e 100 K) possibilitou a observação de efeitos interessantes do resfriamento do cristal sobre o empacotamento cristalino e mudanças na conformação de alguns resíduos.

Como visto anteriormente no capítulo 3, o resfriamento dos cristais levou a uma redução de 10.5% no volume de suas celas unitárias, ocorrendo uma drástica redução no parâmetro de rede a (14.5%). O conteúdo de solvente foi reduzido de 51% (coleta de dados a 277 K) para 46% (coleta de dados a 100 K). A figura 5.5 mostra a sobreposição das celas unitárias e respectivos conteúdos da unidade assimétrica (um tetrâmero) referentes às duas temperaturas. Observa-se uma rotação do tetrâmero como um todo em cerca de 5° em torno de um eixo aproximadamente paralelo a b , levando a um rearranjo das moléculas e à mudança nos parâmetros da rede cristalina. Segundo os eixos a e c , a orientação dos tetrâmeros permaneceu praticamente a mesma.

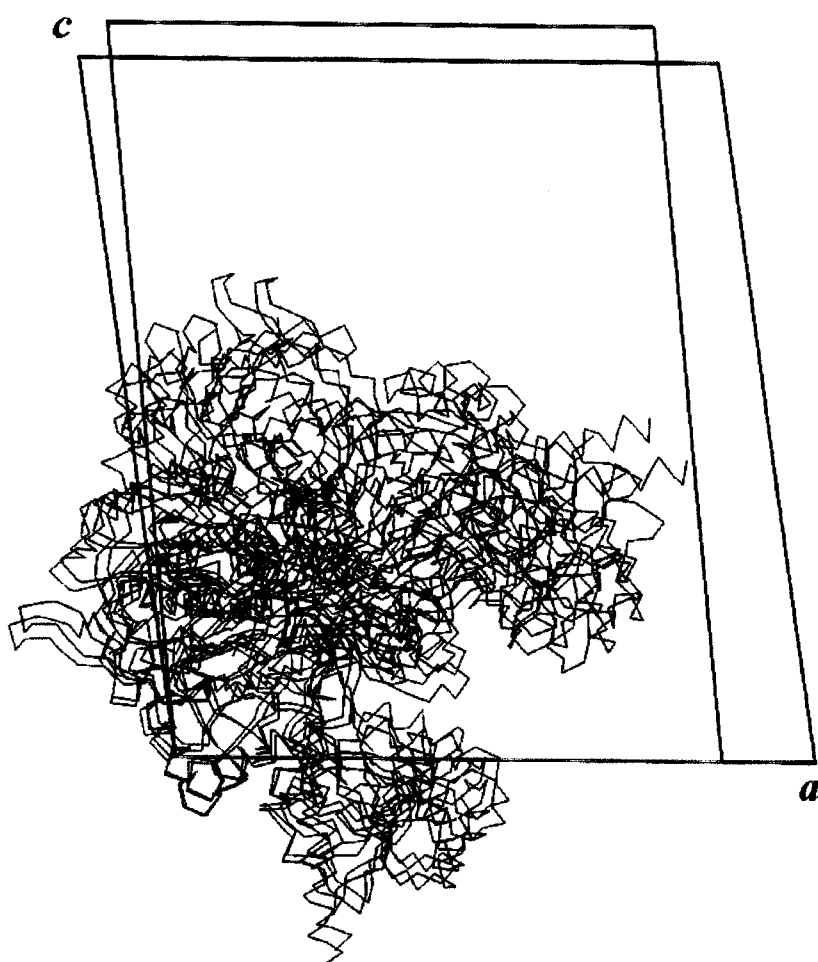


Figura 5.5 - Sobreposição das celas unitárias e respectivos conteúdos da unidade assimétrica (um tetrâmero) referentes às coletas de dados a 277 K (em vermelho) e 100 K (em verde). As direções dos eixos a e b são coincidentes, bem como suas origens. A figura é vista com o eixo b perpendicular ao plano do papel.

Apesar da grande redução ocorrida no volume da cela unitária em decorrência do resfriamento do cristal, o volume do tetrâmero em si não apresentou mudança significativa. Os volumes das moléculas, calculados através do programa VOIDOO (Kleywegt; Jones, 1994), são de $1.434 \times 10^5 \text{ \AA}^3$ para o tetrâmero a 277 K e $1.441 \times 10^5 \text{ \AA}^3$ para o tetrâmero a 100 K. Existe uma pequena mudança na orientação interna dos monômeros, como pode ser observado na figura 5.6, que mostra os dois tetrâmeros vistos a partir da sobreposição de apenas um de seus monômeros.

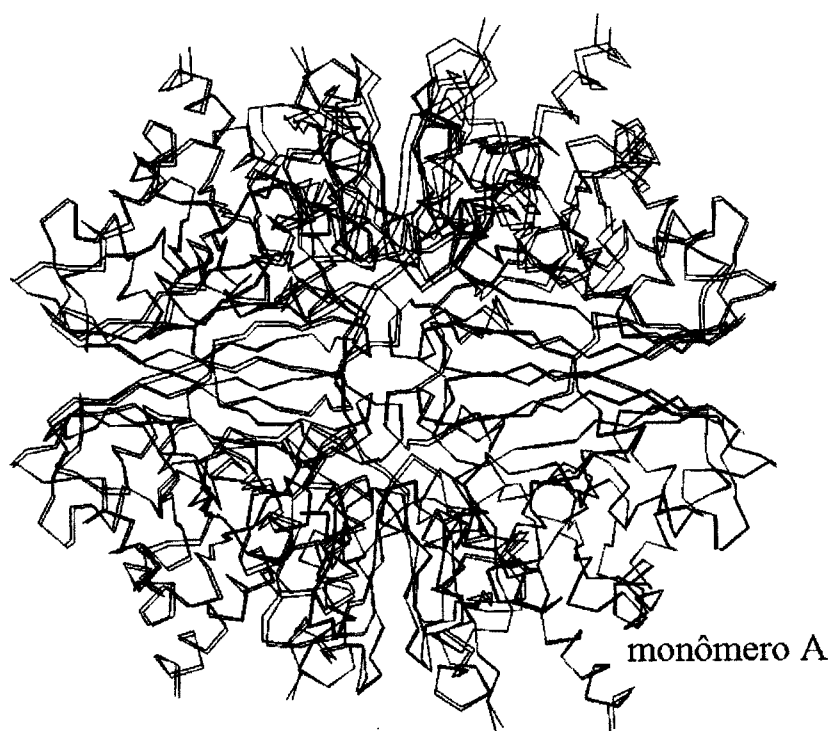


Figura 5.6 - Traços de C_{α} dos tetrâmeros das estruturas a 277 K (em vermelho) e 100 K (em verde), com seus monômeros A sobrepostos.

Entretanto, a estrutura do tetrâmero como um todo é topologicamente a mesma para as duas temperaturas. A sobreposição dos tetrâmeros fornece um desvio quadrático médio de 0.49 \AA considerando os átomos da cadeia principal e 0.66 \AA considerando todos os átomos. As maiores diferenças encontram-se no "loop" flexível que compreende os resíduos 66 a 76, em cadeias laterais expostas e em resíduos envolvidos em contatos cristalinos. A figura 5.7 apresenta um perfil de distâncias médias entre as cadeias laterais das estruturas a 277 e 100 K ao longo dos quatro monômeros, juntamente com a indicação dos resíduos que participam de contatos

cristalinos em ambas as temperaturas. Os resíduos eram considerados envolvidos em contatos cristalinos se a distância entre algum de seus átomos e outro da molécula vizinha fosse igual ou inferior à soma de seus raios de van der Waals acrescida de 0.5 Å. Tais resíduos foram selecionados através do programa WHAT IF (Vriend, 1990). Observa-se uma variação de alguns resíduos que participam de contatos cristalinos em uma estrutura e em outra, o que era esperado devido à mudança no empacotamento cristalino. Nas regiões que não participam de contatos cristalinos, os picos mais proeminentes no perfil de distâncias correspondem às cadeias laterais dos resíduos Lys60, Arg129 e Lys286, que encontram-se bastante expostos ao solvente, daí a sua mobilidade. Além destes, a cadeia lateral da Arg249 e a região entre os resíduos 208 e 210 localizam-se na região do sítio ativo da enzima e sua conformação será discutida mais tarde.

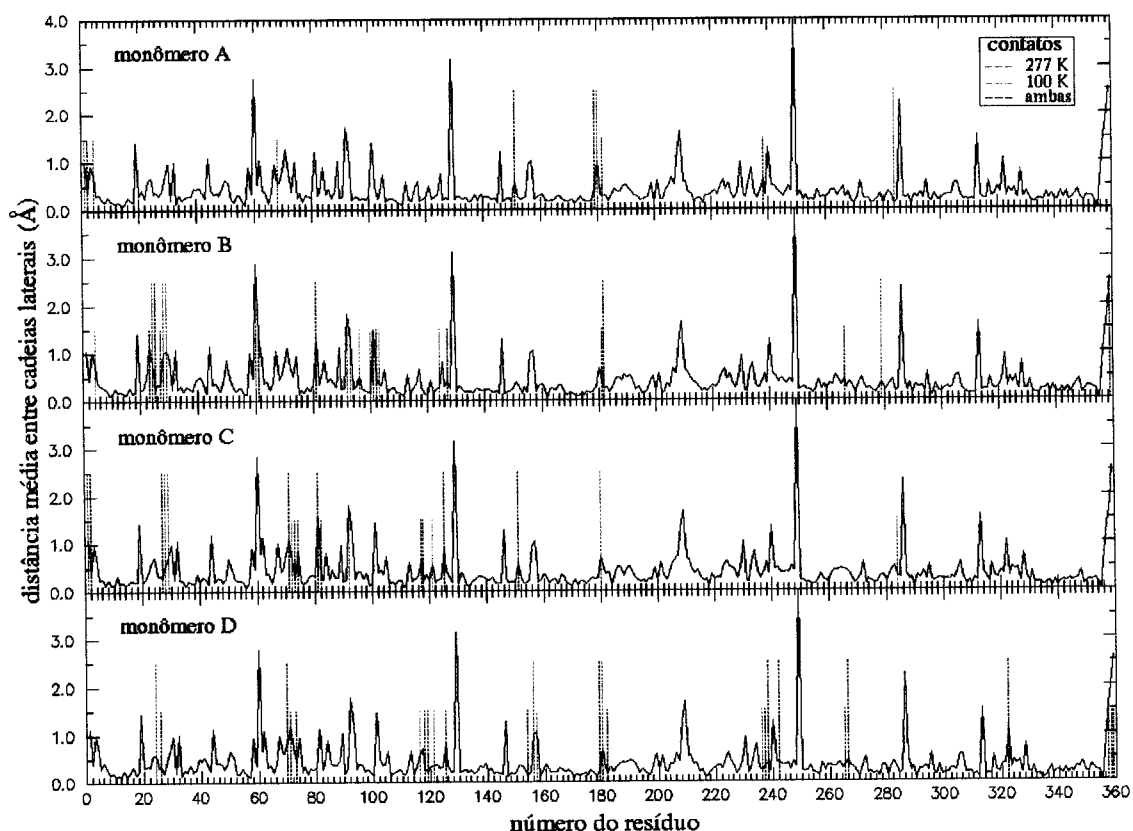


Figura 5.7 - Perfil de distâncias médias entre as cadeias laterais ao longo dos quatro monômeros das estruturas a 277 e 100 K sobrepostas. As linhas verticais pontilhadas indicam os resíduos envolvidos em contatos cristalinos. As linhas vermelhas e verdes correspondem a contatos presentes nas estruturas a 277 e 100 K respectivamente. As linhas azuis e mais altas indicam resíduos que participam de contatos cristalinos em ambas as estruturas.

Estudos cristalográficos da influência da temperatura na estrutura de proteínas foram realizados anteriormente. Tilton e colaboradores (1992) determinaram a estrutura da RNAase em nove temperaturas diferentes (de 98 a 320 K) enquanto Kurinov e Harrison (1995) estudaram a lisozima de 95 a 295 K. A mudança observada no volume da cela unitária foi de 4.7% para os cristais da RNAase e 7.4% para os cristais de lisozima. Em ambos os casos, observou-se uma descontinuidade na mudança dos parâmetros de rede em torno de 200 K, consequência de uma transição de fase associada aos rearranjos estruturais ocorridos nas proteínas. De fato, a evidência de uma transição de fase em sistemas proteína-água no intervalo de 200 a 230 K já havia sido observada por outros métodos (Dodster *et al.*, 1989; Frauenfelder *et al.*, 1988).

Entretanto, as mudanças relatadas anteriormente não foram tão significativas quanto às observadas no presente trabalho. A variação máxima em parâmetro de rede obtida para o cristal de lisozima foi de cerca de 3%, valor muito inferior aos 14.5% observados para o parâmetro *a* do cristal de GAPDH de *T. cruzi*. Além disso, a rotação das moléculas em cerca de 5° foi surpreendente, já que não era de se esperar uma reorientação tão significativa de cerca de 10¹⁴ moléculas no interior do cristal, simultaneamente e quase instantaneamente, já que o resfriamento ocorre em frações de segundo. É possível que haja uma influência da própria solução crioprotetora no rearranjo das moléculas, de modo que este não seja um efeito apenas da mudança de temperatura. Reduções no volume da cela unitária de 2 a 3% são geralmente observadas em experimentos de criocristalografia e a drástica mudança ocorrida no caso dos cristais de GAPDH de *T. cruzi* talvez possa ser atribuída a uma certa fragilidade dos contatos cristalinos característicos dessa estrutura. Relacionado a isso estaria o fato de que cristais da enzima já foram obtidos em três grupos espaciais diferentes, P1, I4 e P2₁, em condições de cristalização muito parecidas, evidenciando uma certa liberdade no empacotamento cristalino das moléculas.

A figura 5.8 mostra a sobreposição dos resíduos da região do sítio ativo das estruturas referentes às coletas de dados a 277 e 100 K. Observa-se que as conformações desses resíduos são bastante coincidentes, com exceção da cadeia lateral da Arg249. O baixo número de interações nas quais a cadeia lateral da Arg249 está envolvida explica sua grande mobilidade: apenas uma ponte salina com o resíduo

Asp210 quando na conformação correspondente à estrutura a 277 K e duas ligações de hidrogênio (com Thr197 e Gln200) quando na conformação correspondente à estrutura a 100 K. A mudança conformacional de tal resíduo deve ter sido um reflexo do rearranjo do tetrâmero. De fato, uma das regiões que apresentam uma diferença significativa nas posições das cadeias laterais nas duas estruturas é a que compreende os resíduos 208 a 210 (figura 5.7), um dos quais (Asp210) interage com a Arg249. A mobilidade desse resíduo pode ter um papel importante no mecanismo catalítico da enzima, como será discutido a seguir.

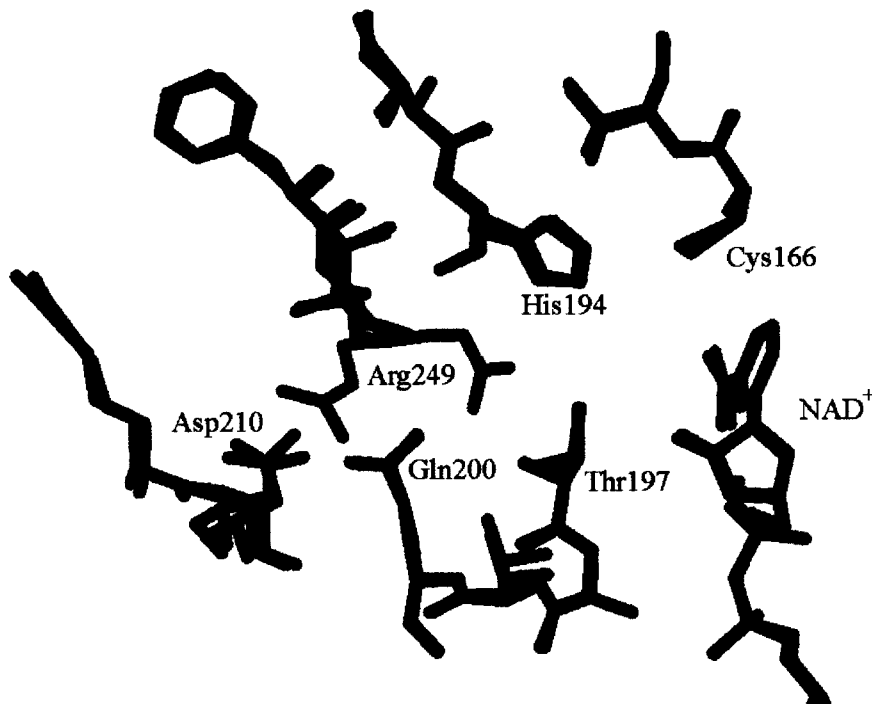


Figura 5.8 - Sobreposição dos resíduos da região do sítio ativo das estruturas de GAPDH referentes às coletas de dados a 277 K (em vermelho) e 100 K (em verde). Observa-se a grande mudança conformacional da cadeia lateral da Arg249.

5.3 - Mutações sítio dirigidas: papel do resíduo Arg249 no processo catalítico

Como dito anteriormente no capítulo 1, a gliceraldeído-3-fosfato desidrogenase (GAPDH) é uma enzima tetramérica que catalisa a reação de fosforilação oxidativa do gliceraldeído-3-fosfato (G3P) em 1,3-bisfosfoglicerato, utilizando NAD^+ como cofator. Dois resíduos essenciais, uma histidina e uma cisteína, estão envolvidos na formação de um intermediário hemitioacetal, quando da ligação enzima-substrato. A figura 5.9 apresenta os diversos passos no mecanismo de ação da GAPDH.

Após a entrada do substrato (G3P) no sítio ativo da enzima, ocorre a formação do complexo hemitioacetal através da ligação covalente do átomo de enxôfre da cisteína com o carbono 1 do G3P. A formação de tal complexo é possível devido à atuação do resíduo histidina como ativador químico, aumentando a reatividade do grupo tiol da cisteína, possivelmente através da formação de um par iônico com o anel imidazólico da histidina. O complexo hemitioacetal doa então um íon hidreto ao anel nicotinamida do cofator NAD^+ reduzindo-o e formando o intermediário tioéster. O cofator reduzido (NADH) é liberado para o meio reacional e outra molécula de NAD^+ entra no sítio ativo da enzima. Ocorre então o ataque nucleofílico do íon fosfato ao complexo tioéster, havendo a formação do produto 1,3-bisfosfoglicerato e a regeneração da enzima livre.

O fato de que uma nova molécula de NAD^+ deve ocupar o sítio ativo da enzima antes que a etapa de fosforilação aconteça parece estar relacionado a mudanças conformacionais que ocorrem no sítio ativo, quando da saída do cofator reduzido. A resolução das estruturas cristalográficas de GAPDH de *Bacillus stearothermophilus* nas formas holo (com NAD^+ ligado) (Skarzynski *et al.*, 1987) e apo (sem NAD^+ ligado) (Skarzynski, Wonacott, 1988) mostrou um aumento da distância entre os resíduos catalíticos cisteína e histidina, na ausência do cofator. Tal aumento, 0.53 Å entre o nitrogênio da histidina e o enxôfre da cisteína, poderia impedir a doação do próton, necessária para a liberação do produto.

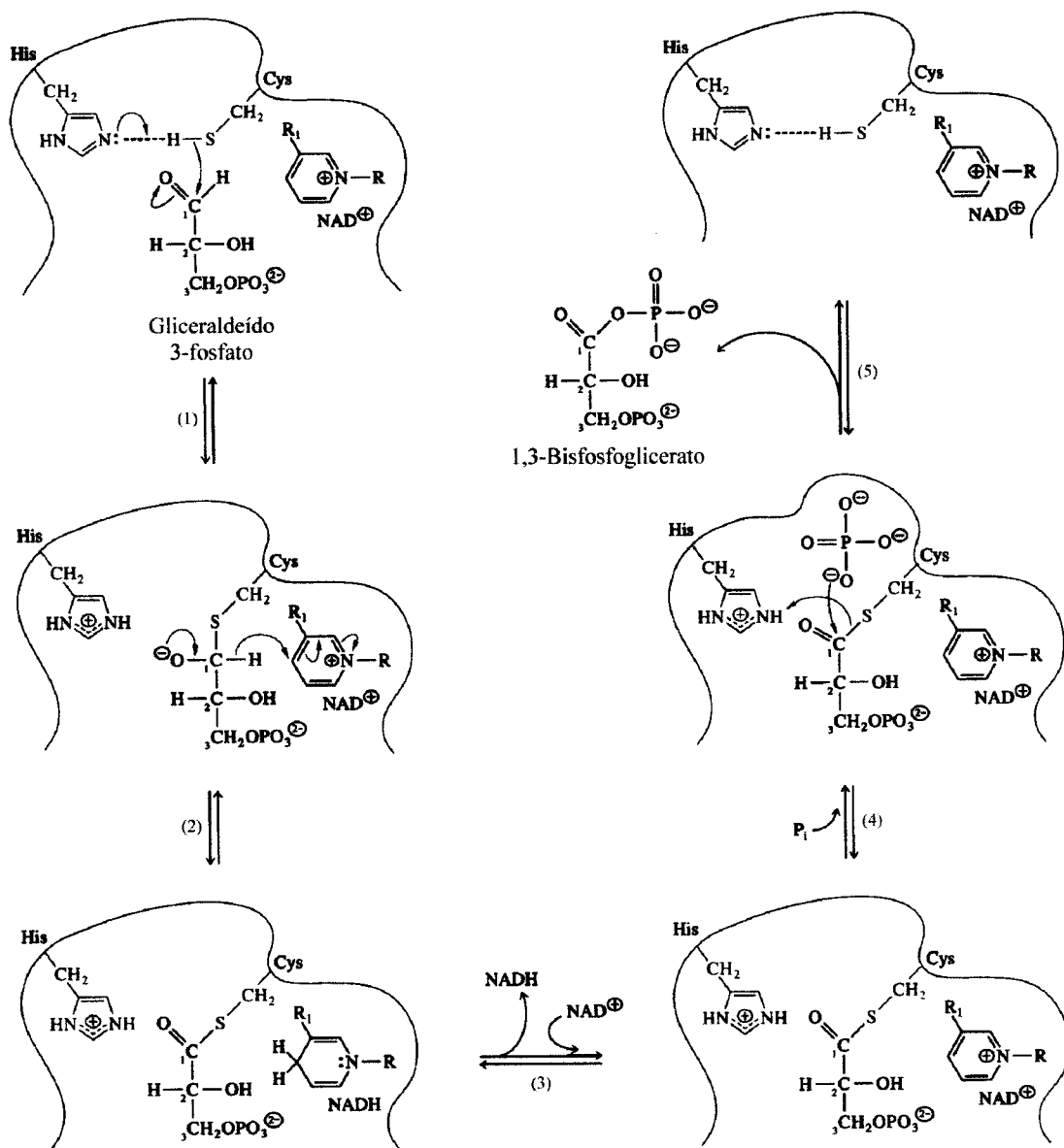


Figura 5.9 - Mecanismo enzimático da GAPDH (figura extraída de Moran *et al.*, 1994; cap. 15).
 1) Grupo sulfidríla da cisteína ionizada ataca o C1 do G3P, resultando na formação do complexo hemitioacetálico; 2) íon hidreto do complexo reduz o NAD⁺, formando o intermediário tioéster; 3) liberação do NADH e entrada de outra molécula de NAD⁺; 4) fosfato inorgânico ataca o grupo carbonila do intermediário tioéster, formando o produto 1,3-bisfosfoglicerato; 5) dissociação do produto.

Estudos mais recentes têm procurado esclarecer o papel dos chamados sítios do fosfato inorgânico (P_i) e fosfato do substrato (P_s) durante o processo catalítico. Com base na análise da estrutura tridimensional da enzima GAPDH de *B.*

stearothermophilus, Skarzynski e colaboradores (1987) propuseram que, embora o sítio P_i deve corresponder à localização do fosfato inorgânico na etapa de fosforilação, ele também poderia exercer algum papel na ligação do C3-P do substrato quando da formação do intermediário hemitioacetil. Dessa forma, tal intermediário seria deslocado durante o processo catalítico.

Em 1994, Corbier e colaboradores realizaram experimentos cinéticos com a enzima GAPDH de *B. stearothermophilus* tipo selvagem e mutantes R195L e R231G. O resíduo Arg231 é equivalente à Arg249 em GAPDH de *T. cruzi* e estão associadas ao sítio do fosfato do substrato. Com base nesses experimentos os autores constataram que a Arg231 não exerce papel significativo na ligação do substrato, mas a sua substituição por uma glicina levou a uma diminuição da eficiência catalítica da enzima em cerca de 500 vezes. Corbier e colaboradores sugeriram que o grupo C3-P do substrato deve ocupar inicialmente o sítio do fosfato inorgânico para a formação do complexo hemitioacetil. Antes da etapa de fosforilação, ocorreria uma isomerização conformacional, quando o C3-P do intermediário tioéster passaria do sítio P_i para o sítio P_s , deixando que a ligação do fosfato inorgânico no C1 ocorresse no sítio P_i . A isomerização conformacional ocorreria durante a troca do cofator NAD^+ , quando o sítio P_s torna-se mais acessível.

Com o objetivo de esclarecer o mecanismo proposto acima, a figura 5.10 mostra a conformação dos resíduos do sítio ativo da enzima GAPDH de *T. cruzi* referente à estrutura a 2.15 Å de resolução obtida a partir da coleta de dados a 100 K. Os sítios P_s e P_i equivalem àqueles determinados na estrutura de GAPDH de *Leishmania mexicana*. O substrato G3P foi manualmente posicionado de modo que seu grupo C3-P ocupasse o sítio do fosfato inorgânico e seu carbono C1 estivesse em localização coerente com a sua ligação ao grupo sulfidril da Cys166. Trata-se apenas de uma simulação da localização do substrato. O resíduo Arg249 (equivalente à Arg231 de *B. stearothermophilus*) estaria associado à isomerização conformacional sugerida anteriormente.

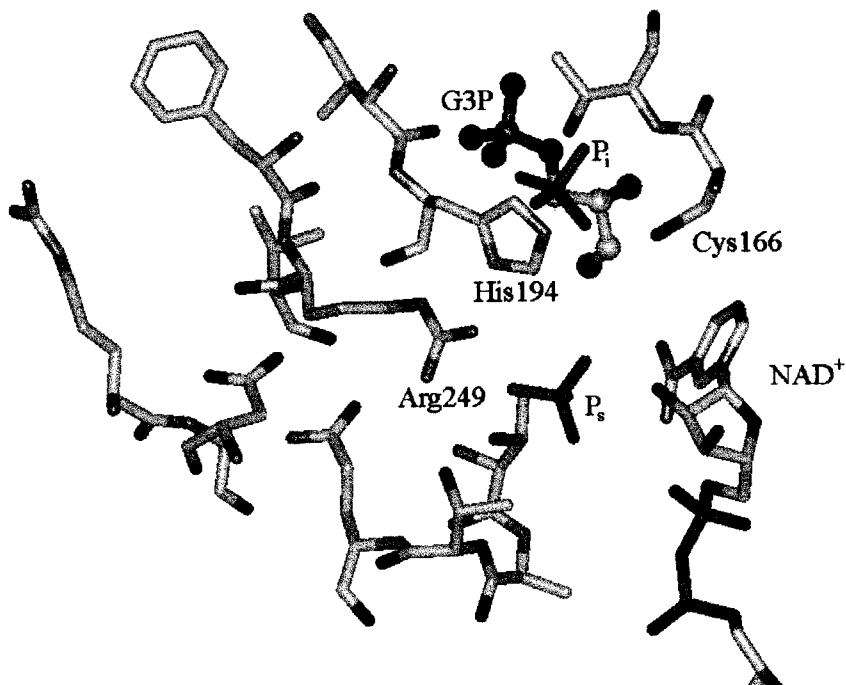


Figura 5.10 - Conformação dos resíduos do sítio ativo da enzima GAPDH de *T. cruzi*. O substrato G3P foi posicionado manualmente de modo a simular sua localização no início do processo catalítico. Os sítios P_i e P_s equivalem às posições do fosfato inorgânico e fosfato do substrato obtidas na estrutura de GAPDH de *Leishmania mexicana*.

Analisando os resultados cinéticos obtidos no presente trabalho, observa-se que as mutações praticamente não alteraram o valor de K_M e, portanto, não influenciaram na ligação do substrato. Entretanto, a constante catalítica k_{cat} diminuiu cerca de 4 vezes, indicando uma redução na taxa de dissociação do complexo enzima-substrato para a liberação do produto ($ES \rightarrow E + P$). Pode-se sugerir que a mobilidade do resíduo Arg249 tenha um importante papel no mecanismo catalítico da enzima, facilitando que a isomerização conformacional do intermediário tioéster ocorra. Durante a isomerização, a Arg249 adotaria sua outra conformação formando uma ponte salina com o resíduo Asp210. Em seguida, voltaria para interagir com o fosfato do complexo tioéster, estabilizando sua conformação até que a etapa de fosforilação ocorresse. As mutações estudadas nesse trabalho, D210L e D210L-G213D, poderiam estar desfavorecendo a movimentação da Arg249, dificultando a isomerização conformacional do intermediário e, portanto, diminuindo a eficiência catalítica da enzima.

A figura 5.11 mostra uma sobreposição dos resíduos da região do sítio ativo da enzima mutante GAPDH D210L-G213D e da enzima tipo selvagem com a Arg249 na conformação de interação com o resíduo Asp210 (coleta de dados a 277 K). A estrutura do mutante, mesmo com a introdução de um ácido aspártico estruturalmente próximo ao Asp210 não consegue recuperar a conformação adequada para a ponte salina entre Arg249 e Asp210. A distância entre o nitrogênio da Arg249 da estrutura da enzima tipo selvagem e o oxigênio do Asp213 da enzima mutante é de apenas 1.6 Å.

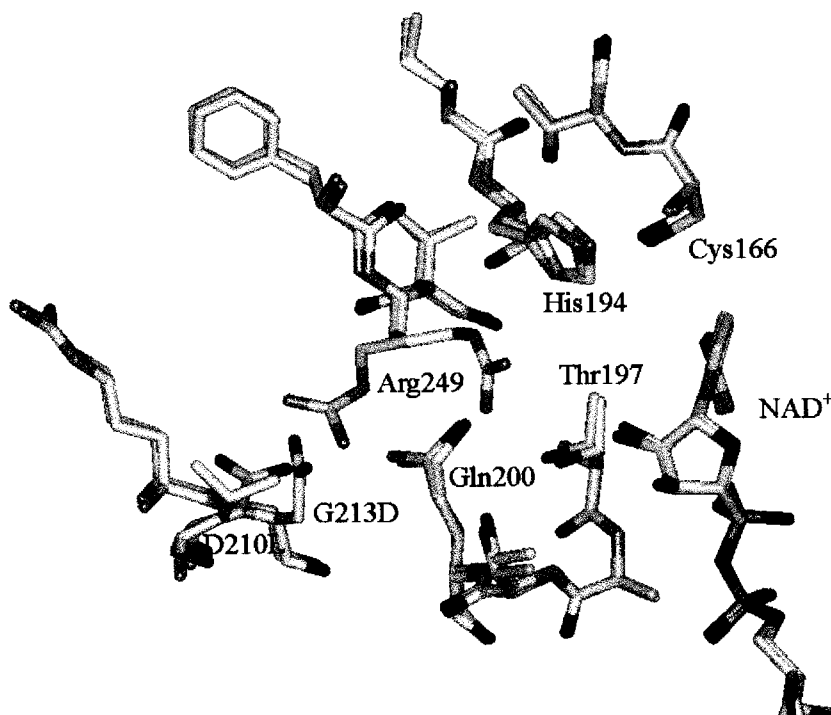


Figura 5.11 - Sobreposição dos resíduos da região do sítio ativo da enzima GAPDH tipo selvagem (átomos de carbono em cinza) e do mutante D210L-G213D (átomos de carbono em amarelo).

Do ponto de vista do desenho de inibidores, a interferência com a mobilidade do resíduo Arg249 parece afetar significativamente a eficiência catalítica da enzima. Dessa forma, como sugerido por Souza e colaboradores (1998), compostos que participem de interações polares com a cadeia lateral de Asp210 (substituído por uma leucina na enzima humana) poderiam ligar-se seletivamente à enzima do parasita, impedindo a mudança conformacional da Arg249, importante no processo catalítico da enzima.

Capítulo 6

Conclusões e Perspectivas

No presente trabalho, foram feitas a expressão, purificação e medidas de cinética enzimática da proteína gliceraldeído-3-fosfato desidrogenase (GAPDH) de *T. cruzi* e dos mutantes D210L-G213D e D210L. As estruturas tridimensionais da enzima tipo selvagem e mutante D210L-G213D foram determinadas por difração de raios-X em monocristais. Tais estruturas, juntamente com os parâmetros cinéticos, permitiram a discussão a respeito de alguns detalhes do mecanismo catalítico da enzima.

Embora a estrutura cristalográfica da enzima tipo selvagem já tivesse sido determinada anteriormente a 2.8 Å de resolução, cristais de melhor qualidade permitiram, no presente trabalho, a obtenção de um modelo a alta resolução (2.15 Å), o que significa maior nível de detalhes atômicos. Por outro lado, a determinação da estrutura em duas temperaturas diferentes (277 e 100 K) levou a resultados interessantes no que diz respeito ao empacotamento cristalino.

O projeto no qual este trabalho está inserido, que visa o planejamento de uma nova droga contra a doença de Chagas, tem-se desenvolvido em várias direções. A partir da estrutura cristalográfica da GAPDH de *T. cruzi* foram desenhados vários compostos que pudessem atuar como inibidores da enzima. Alguns destes compostos já foram sintetizados e testados, apresentando atividade inibitória significativa. Por outro lado, tem-se buscado o isolamento de inibidores a partir de produtos naturais. Extratos de plantas têm sido testados quanto à sua atividade inibitória e fracionados até chegar-se a substâncias puras. O trabalho encontra-se na fase de cristalização de complexos proteína-ligante para que suas estruturas tridimensionais sejam determinadas. Posteriormente à determinação das estruturas dos complexos, os compostos mais promissores devem ser otimizados, prosseguindo-se no ciclo do

desenho racional de drogas (capítulo 1) até que um composto apresente propriedades farmacológicas adequadas.

O processo de obtenção de uma nova droga contra determinada doença leva vários anos e dificilmente poderá ter as suas etapas finais desenvolvidas a nível acadêmico, sendo necessária uma parceria com a indústria farmacêutica. Embora ainda esteja-se distante de uma nova droga contra a doença de Chagas, alguns passos já foram dados nesse sentido.

Referências Bibliográficas

- ARNDT, U.W.; WONACOTT, A.J. *The Rotation Method in Crystallography*. Amsterdam, North-Holland Publishing Company, 1977.
- AUSUBEL, F.M.; BRENT, R.; KINGSTON, R.E.; MOORE, D.D.; SEIDMAN, J.G.; SMITH, J.A.; STRUHL, K. *Short Protocols in Molecular Biology*. 3.ed. USA, John Wiley & Sons Inc., 1995.
- BARBOSA, V.M.; NAKANO, M.. *Comp. Biochem. Physiol.*, v.88B, p.563-8, 1987.
- BERNSTEIN, B.E.; MICHELS, P.A.M.; HOL, W.G.J.. *Nature*, v.385, p.275-8, 1997.
- BRADFORD, M.M.. *Analyt. Biochem.*, v.72, p.248-54, 1976.
- BRÜNGER, A.T.. *Nature*, v.355, p.472-4, 1992.
- BRÜNGER, A.T.; ADAMS, P.D.; RICE, L.M.. *Structure*, v.5, p.325-36, 1997.
- BRÜNGER, A.T.; KURIYAN, J.; KARPLUS, M.. *Science*, v.235, p.458-60, 1987.
- BLUNDELL, T.L.; JOHNSON, L.N. *Protein Crystallography*. London, Academic Press Inc., 1976.
- BOWIE, J.U.; LÜTHY, R.; EISENBERG, D.. *Science*, v.253, p.164-70, 1991.
- CARTER, C.W.JR.; CARTER, C.W.. *J. Biol. Chem.*, v.254, p.12219-23, 1979.
- CHAPMAN, M.S.; SUH, S.W.; CURMI, P.M.G.; CASCIO, D.; SMITH, W.W.; EISENBERG, D.S.. *Science*, v.241, p.71-74, 1988.
- COLLABORATIVE COMPUTATIONAL PROJECT NUMBER 4. *Acta Cryst.*, v.D50, p.760-3, 1994.
- CORBIER, C.; MICHELS, S.; WONACOTT A.J.; BRANLANT, G.. *Biochemistry*, v.33, p.3260-5, 1994.
- DAGHER, S.M.; DEAL, W.C.. *Archs. Biochem. Biophys.*, v.179, p.643-56, 1977.
- DE LUCAS, L. J.. *Science*, v.246, p.651, 1989.
- DEWAN, J.C.; MIKAMI, B.; HIROSE, M.; SACCHETTINI, J.C.. *Biochemistry*, v.32, p.11963-8, 1993.
- DODSTER, W.; CUSACK, S.; PETRY, W.. *Nature*, v.337, p.754-6, 1989.
- DRENTH, J. *Principles of Protein X-ray Crystallography*. New York, Springer-Verlag Inc., 1994.

- DUCRUIX, A.; GIEGÉ, R. *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. New York, Oxford University Press, 1992.
- DUÉE, E.; DEYRIS, L.O.; FANCHON, E.; CORBIER C.; BRANLANT, G.; DIDEBERG, O.. *J. Mol. Biol.*, v.257, p.814-38, 1996.
- DUGGLEBY, R.G.; DENNIS, D.T.. *J. Biol. Chem.*, v.249, p.167-74, 1974.
- FRAUENFELDER, H.; PARAK, F.; YOUNG, R.. *Annu. Rev. Biophys. Biophys. Chem.*, v.17, p.451-9, 1988.
- GARMAN, E.F.; SCHNEIDER, T.R.. *J. Appl. Cryst.*, v.30, p.211-37, 1997.
- GIACOVAZZO, C.; MONACO, H.L.; VITERBO, D.; SCORDARI, F.; GILLI, G.; ZANOTTI, G.; CATTI, M. *Fundamentals of Crystallography*. New York, Oxford University Press Inc., 1992.
- GUIMARÃES, B.G.; SOUZA, D.H.F.; JESUS, W.D.P.; OLIVA, G.. *Acta Cryst.*, v. A52 Supplement, p.C-98, 1996.
- HANNAERT, V.; OPPERDOES, F.R.; MICHELS, P.A.M.. *Prot. Expr. Purif.*, v.6, p.244-50, 1995.
- HARRIS, J.I.; WATERS, M.. *Enzymes*, v.13, p.1-49, 1976.
- HO, S.N.; HUNT, H.D.; HORTON, R.M.; PULLEN, J.K.; PEASE, L.R.. *Gene*, v.77, p.51-9, 1989.
- HOOFT, R.W.W.; VRIEND, G.; SANDER, C.; ABOLA, E.E.. *Nature*, v.381, p.272, 1996.
- JONES, T.A.; ZOU, J.Y.; COWAN, S.W.; KJELDGAARD, M.. *Acta Cryst.*, v.A47, p.110-9, 1991.
- JONES, T.A.; KJELDGAARD, M. *O - the manual version 6.0*. Uppsala, 1996.
- KIM, H.; FEIL, I.K.; VERLINDE, C.L.M.J.; PETRA, P.H.; HOL, W.G.J.. *Biochemistry*, v.34, p.4975-86, 1995.
- KORNDÖRFER, I.; STEIPE, B.; HUBER, R.; TORNSCHY, A.; JAENICKE, R.. *J. Mol. Biol.*, v.246, p.511-21, 1995.
- KURINOV, I.V.; HARRISON, R.W.. *Acta Cryst.*, v.D51, p.98-109, 1995.
- LAEMMLI, U.K.. *Nature*, v.227, p.680-5, 1970.
- LASKOWSKI, R.A.; MACARTHUR, M.W.; THORNTON, J.M.. *J. Appl. Cryst.*, v.26, p.283-91, 1993.
- LAMZIN, V.; WILSON, K.. *Acta Cryst.*, v.D49, p.129-47, 1993.

- LIN, J.H.. *Adv. Drug. Deliver Rev.*, v.27, p.215-33, 1997.
- MATTHEWS, B.W.. *J. Mol. Biol.*, v.33, p.491-7, 1968.
- MCPHERSON, A. *Preparation and Analysis of Protein Crystals*. New York, John Wiley & Sons Inc., 1982.
- MINOR, W. *XDISPLAYF Program*. Purdue University, 1993.
- MORAN, L.A.; SCRIMGEOUR, K.G.; HORTON, H.R.; OCHS, R.S.; RAWN, J.D. *Biochemistry*. USA, Neil Patterson Publishers/Prentice Hall Inc., 1994.
- MURSHUDOV, G.N.; VAGIN, A.A.; DODSON, E.J.. *Acta Cryst.*, v.D53, p.240-55, 1997.
- NAVAZA, J.. *Acta Cryst.*, v.A50, p.157-63, 1994.
- NAVE, C.. *Radiat. Phys. Chem.*, v.45, p.483-90, 1995.
- NEVES, D.P. *Parasitologia Humana*. 8.ed. São Paulo, Livraria Atheneu Editora, 1991.
- OPPERDOES, F.R.. *Br. Med. Bull.*, v.41, p.130-6, 1985.
- OPPERDOES, F.R.; BORST, P.. *FEBS Lett.*, v.80, p.360-4, 1977.
- OTWINOWSKI, Z. *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*. England, 1993. p.56-62.
- PANNU, N.S.; READ, R.J.. *Acta Cryst.*, v.A52, p.659-68, 1996.
- PEARL, L.; BLUNDELL, T.. *FEBS Lett.*, v.174, p.96-101, 1984.
- PET SYSTEM MANUAL. 4.ed. USA, Novagen, 1994.
- RAMACHANDRAN, G.N.; RAMAKRISHNAN, C.; SASISEKHARAN, V.. *J. Mol. Biol.*, v.7, p.95-9, 1963.
- READ, R.J.. *Acta Cryst.*, v.A46, p.900-12, 1990.
- SKARZYNSKI, T.; MOODY, P.C.E.; WONACOTT, A.J.. *J. Mol. Biol.*, v.193, p.171-87, 1987.
- SKARZYNSKI, T.; WONACOTT, A.J.. *J. Mol. Biol.*, v.203, p.1097-118, 1988.
- SOUKRI, A.; MOUGIN, A.; CORBIER, C.; WONACOTT, A.; BRANLANT, C.; BRANLANT, G.. *Biochemistry*, v.28, p.2586-92, 1989.
- SOUZA, D.H.F. *Estrutura Cristalográfica da Enzima Gliceraldeído 3-Fosfato Desidrogenase de T. cruzi: implicações no mecanismo catalítico e potenciais sítios de inibição*. São Carlos, 1996. 75p. Tese (Doutorado) - Instituto de Química de São Carlos, Universidade de São Paulo.

- SOUZA, D.H.F.; GARRATT, R.C.; ARAÚJO, A.P.U.; GUIMARÃES, B.G.; JESUS, W.D.P.; MICHELS, P.A.M.; HANNAERT, V.; OLIVA, G.. *FEBS Lett.*, v.424, p.131-5, 1998.
- TILTON, R.F.JR.; DEWAN, J.C.; PETSKE, G.A.. *Biochemistry*, v.31, p.2469-81, 1992.
- TRENTHAM, D.R.. *Biochem. J.*, v.122, p.71-7, 1971.
- VELLIEUX, F.M.D.; HAJDU, J.; VERLINDE, C.L.M.J.; GROENDIJK, H.; READ, R.J.; GREENHOUGH, T.J.; CAMPBELL, J.W.; KALK, K.H.; LITTLECHILD, J.A.; WATSON, H.C.; HOL, W.G.J.. *Proc. Natl. Acad. Sci. USA*, v.20, p.2355-9, 1993.
- VERLINDE, C.L.M.J.; CALLENS, M.; VAN CALENBERGH, S.; VAN AERSCHOT, A.; HERDEWIJN, P.; HANNAERT, V.; MICHELS, P.A.M.; OPPERDOES, F.R.; HOL, W.G.J.. *J. Med. Chem.*, v.37, p.3605-13, 1994.
- VERLINDE, C.L.M.J.; HOL, W.G.J.. *Structure*, v.15, p.577-87, 1994.
- VOET, D.; VOET, J.G. *Biochemistry*. 2.ed. USA, John Wiley & Sons Inc., 1995.
- VRIEND, G.. *J. Mol. Graph.*, v.8, p.52-6, 1990.
- VRIEND, G.; SANDER, C.. *J. Appl. Cryst.*, v.26, p.47-60, 1993.
- WATSON, H.C.; DUÉE, E.; MERCER, W.D.. *Nature*, v.240, p.130-9, 1972.
- WATSON, J.D.; GILMAN, M.; WITKOWSKI, J.; ZOLLER, M. *Recombinant DNA*. 2.ed. New York, Scientific American Books, 1992.
- WIERENGA, R.K.; NOBLE, M.E.M.; VRIEND, G.; NAUCHE, S.; HOL, W.G.J.. *J. Mol. Biol.*, v.220, p.995-1015, 1991.
- WORLD HEALTH ORGANIZATION STATISTICAL INFORMATION SYSTEM WEBSITE, <http://www.who.ch./whosis/whosis.htm>.