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

ESTRUTURA CRISTALOGRÁFICA
DA BOTHROPSTOXINA-I, UMA MIOTOXINA K49
TIPO FOSFOLIPASE A₂

MARIA TERESA DA SILVA GIOTTO

Tese apresentada ao Instituto de Física
de São Carlos, Universidade de São Paulo,
para obtenção do título de Doutor em
Ciências, área de concentração Física
Aplicada.

ORIENTADOR: RICHARD CHARLES GARRATT

SÃO CARLOS- 1996

Giotto, Maria Teresa da Silva

Estrutura Cristalográfica da Bothropstoxina-I, uma miotoxina K49
tipo fosfolipase A₂. / Maria Teresa da Silva Giotto. -- São Carlos, 1996.
162 p.

Tese (Doutorado) -- Instituto de Física de São Carlos, 1996.

Orientador: Prof. Dr. Richard Charles Garratt

1. Miotoxina. 2. Bothropstoxina-I (BthTX-I). 3. Fosfolipase A₂.
4. Difração de raios-X.

**UNIVERSIDADE
DE SÃO PAULO**

Instituto de Física de São Carlos

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

Fone (016) 272-6222

Fax (016) 272-2218

MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE **MARIA TERESA DA SILVA GIOTTO** APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO CARLOS, UNIVERSIDADE DE SÃO PAULO, EM 13/09/1996.

COMISSÃO JULGADORA:

Prof. Dr. Richard Charles Garratt/IFSC-USP

Prof. Dr. Glaucius Oliva/IFSC-USP

Profa. Dra. Rosemary Sanches/IFSC-USP

Prof. Dr. Richard J. Ward/UNESP-SJRP

Prof. Dr. Raghuvir Krishnaswamy Arni/UNESP-SJRP

Este trabalho teve o apoio financeiro:

CNPq

FINEP

AGRADECIMENTOS

Ao Professor Dr. Richard C. Garratt pela sua orientação, ensinamentos, paciência e sua intensa participação no desenvolvimento deste trabalho.

Ao Professor Dr. Glaucius Oliva pela sua orientação em pontos críticos e decisivos no desenvolvimento deste trabalho.

À Professora Dra. Yvonne P. Mascarenhas pela amizade, ensinamentos e pela introdução nesta área de estudo.

Ao Professor Dr. Giglio e A. C. O. Cintra pelo fornecimento das amostras que tornaram este trabalho possível.

Aos Professores Eduardo Castellano, Júlio Zuckermann- Schpector, Ignez Careccelli, Heloisa S. Selistre de Araujo pelos ensinamentos e amizade.

À técnica Wanda pela introdução e sempre orientação nas técnicas de laboratório.

Aos amigos Valma, Jorge, Paulão, Gisele, Bia e Cris pelas discussões e ensinamentos.

Aos amigos Zac, João, Marcos, Regiane pela amizade.

Aos técnicos Augusto e Geraldo pela amizade e sempre colaboração.

À secretária Maria Helena pela amizade e disposição em ajudar.

Aos amigos Geraldo Nery e Bianca pela participação amigável e carinhosa.

A todos os amigos do grupo que direta ou indiretamente ajudaram na realização deste trabalho.

Aos amigos Chico e Luci pelo constante apoio e carinho.

Um agradecimento especial à minha família.

À amiga Rita pelo seu carinho, amizade e alegria.

Ao amigo Ignácio pela amizade e sabedoria.

SUMÁRIO

Lista de Figuras	v
Lista de Tabelas	ix
Lista de Abreviaturas	xi
Resumo	xii
Abstract	xiii
CAPÍTULO 1 - FOSFOLIPASES A₂	1
1.1 Introdução	1
1.2 Classificação das Fosfolipases	2
1.3 Fosfolipases A₂ (PLA₂)	3
1.3.1 Análise Sequencial das Fosfolipases A₂	5
1.3.1.1 Pontes de Dissulfeto	7
1.3.1.2 N-terminal	9
1.3.1.3 Sítio de Ligação do Ca ⁺²	10
1.3.1.4 Sítio ativo	12
1.3.1.5 Mecanismo Catalítico das Fosfolipases A ₂	12
1.3.1.6 Ativação Interfacial e o Canal Hidrofóbico	16
1.3.1.7 Resíduos do C-terminal	19
1.4 Estrutura Secundária das PLA₂	19
1.5 BthTX-I (Bothropstoxina I)	20
1.6 Monômero x Dímero	22
1.7 Fosfolipases PLA₂ Envolvidas no Processo de Inflamação	25
1.8 Referências Bibliográficas	27
CAPÍTULO 2 - CRISTALIZAÇÃO DA BthTX-I	31
2.1 Introdução	31
2.2 Métodos de Cristalização de Proteínas	35
2.2.1 Difusão de Vapor	36
2.2.1.1 “Hanging drop”	37

2.3 Ensaios de Cristalização	37
2.4 Ensaio de Cristalização da BthTX-I	39
2.4.1 Quantidade de proteína Disponível e Verificação da Pureza	39
2.4.2 Testes iniciais de Precipitação	40
2.5 Referências Bibliográficas	44
CAPÍTULO 3 - COLETA E PROCESSAMENTO DOS DADOS DE DIFRAÇÃO DA BthTX-I.	45
3.1 Introdução	45
3.2 Câmaras e Detetores	45
3.2.1 A Região cega	47
3.2.2 Escolha da Orientação do Cristal para as fotos de oscilação e rotação	48
3.2.3 Ângulo de Oscilação de cada filme para evitar sobreposições	49
3.3 Fatores que afetam as medidas de intensidade	50
3.3.1 A mosaicidade	50
3.3.2 Colimadores e Monocromadores de Feixe de Raios X	51
3.3.3 O efeito do tamanho da cela unitária na medida da intensidade difratada	53
3.3.4 Limitações experimentais da resolução e do número de reflexões medidas	53
3.4 Redução dos dados	55
3.4.1 Fator de Lorentz (L)	55
3.4.2 Fator de Polarização	56
3.4.3 Fator de Absorção	56
3.5 Coleta de dados	59
3.5.1 Geometria e Características Principais do R-AXIS IIC	59
3.5.2 Sistema de Leitura (Detetor de área - Sistema de Placas de Imagem)	60
3.6 Processamento computacional das Imagens Medidas	61
3.6.1 Coordenadas de cada reflexão no filme	62
3.6.2 Cálculo dos vetores base para a determinação da cela unitária e matriz de orientação do cristal.	64
3.6.3. Refinamento da matriz de orientação do cristal	65
3.6.4 Conceitos teóricos envolvidos no Processamento das Imagens de oscilação	67
3.6.4.1 Integração das intensidades de cada reflexão em cada imagem de oscilação	67
3.6.4.2 Determinação do Perfil da Intensidade	68
3.6.4.3 A Integração de cada reflexão pelo método de ajuste do perfil	70

3.6.4.4 Escalonamento relativo das várias imagens de oscilação	71
3.7 Avaliação do conjunto de imagens	73
3.8 Coleta de dados da BthTX-I	73
3.9 Processamento das Imagens da BthTX-I	77
3.10 Referências Bibliográficas	79
CAPÍTULO 4 - ANÁLISE PRELIMINAR E SUBSTITUIÇÃO MOLECULAR DA Bthtx-I	80
4.1 Cálculo do conteúdo da Cella unitária	80
4.2 Função de Patterson e sua Interpretação	82
4.3 A Função de Rotação	84
4.4 A Função de Translação	88
4.5 Substituição Molecular	88
4.5.1 Sorting	90
4.5.2 Tabling	91
4.5.3 Rotating	92
4.5.4 Traing	92
4.5.5 Fitting	93
4.5.6 Cording	94
4.6 Resultados da Função de Auto-rotação para a BthTX-I	94
4.7 Resolução da estrutura da BthTX-I usando o programa AMoRe	96
4.8 Referências Bibliográficas	98
CAPÍTULO 5 - REFINAMENTO DA ESTRUTURA DA Bthtx-I	99
5.1 Introdução	99
5.2 Refinamento utilizando o Programa X-PLOR	100
5.2.1 Etapas do Refinamento	102
5.2.1.1 Rotina Topologia	102
5.2.1.2 Refinamento de corpo rígido	102
5.2.1.3 Rotina Check e a determinação de W_A	103
5.2.1.4 Refinamento por dinâmica Molecular com temperatura controlada	103

(Simulated Annealing)	105
5.2.1.5 Refinamento posicional com minimização de energia	105
5.2.1.6 refinamento dos fatores de temperatura (Individuais ou de grupos de átomos)	106
5.2.1.7 Análise do modelo refinado	106
5.2.1.8 Simetria não-cristalográfica (NCS)	107
5.2.1.9 Colocação das Moléculas de água	108
5.3 Mapas de densidade eletrônica	109
5.4 Avaliação da Qualidade do modelo	111
5.4.1 Refinamento da BthTX-I	111
5.4.2 Estereoquímica do modelo	115
5.4.3 Análise Conformacional do modelo	117
5.5 Referências Bibliográficas	121
CAPÍTULO 6 - O MODELO DA BthTX-I: ANÁLISE E DISCUSSÃO DOS RESULTADOS	123
6.1 Descrição da estrutura	123
6.1.1 Superposição dos monômeros	126
6.1.2 Análise dos Fatores de Temperatura B	128
6.1.3 Colocação das moléculas de água	130
6.2 Questão da Dimerização	130
6.2.1 Apresentação dos resultados das Eletroforeses em gel de 15% de Poliacrilamina	130
6.2.2. Cálculo da Superfície de Interação entre Moléculas	132
6.3 Análise do canal Hidrofóbico	138
6.4 O sítio Ativo	139
6.5 Loop de ligação do Ca^{+2} (loop típico de ligação do primeiro íon Ca^{+2})	140
6.6 Modelo do fosfolipídeo Fosfatidil etanolamina na estrutura da BthTX-I	144
6.7 Resíduos da Interface	148
6.8 Árvores Filogenéticas	152
6.9 Referências Bibliográficas	157
CAPÍTULO 7 - CONCLUSÕES E PERSPECTIVAS FUTURAS	159

LISTA DE FÍGURAS

CAPÍTULO 1

Figura 1.1 - Especificidade das fosfolipases A ₂	2
Figura 1.2 - Mecanismo de atividade de uma fosfolipase A ₂	3
Figura 1.3 - Classificação das pontes dissulfeto	9
Figura 1.4 - Desenho esquemático das interações do N-terminal	10
Figura 1.5 - Representação esquemática do íon Ca ⁺² e seus ligantes	11
Figura 1.6 - Mecanismo catalítico das fosfolipases A ₂	13
Figura 1.7 - Mecanismo catalítico proposto por White et al., 1990	15
Figura 1.8 - Resíduos do canal hidrofóbico/ Naja naja atra	17
Figura 1.9 - Representação esquemática do canal hidrofóbico	18
Figura 1.10 - Estrutura tridimensional da PLA ₂ Pancreática bovina	20
Figura 1.11 - Interações que estabilizam o dímero na Crotalus atrox	23
Figura 1.12 - Estrutura dimérica da Crotalus atrox	24
Figura 1.13 - Dímero da Miotoxina II da Bothrops asper	25
Figura 1.14 - Liberação do ácido araquidônico pela ação das PLA ₂	26

CAPÍTULO 2

Figura 2.1 - Método de Difusão de Vapor “hanging drop”	37
Figura 2.2 - Eletroforese da BthTX-I	39
Figura 2.3 - Condições de cristalização da BthTX-I	41
Figura 2.4 - Fotos dos cristais da BthTX-I nativa	42
Figura 2.5 - Cristais do complexo BthTX-I mais inibidor	43

CAPÍTULO 3

Figura 3.1 - A esfera de Ewald	46
Figura 3.2 - Geometria da câmara de rotação	47
Figura 3.3 - Raio da região cega	48
Figura 3.4 - Projeção ao longo do eixo de rotação	49
Figura 3.5 - Efeito da Mosaicidade na Intensidade da reflexão	51
Figura 3.6 - Resolução Máxima de um filme de raios X	54
Figura 3.7 - Variação do coeficiente de absorção	58
Figura 3.8 - Difratorômetro automático R- AXIS IIC	60
Figura 3.9 - Relação entre as coordenadas do filme e o sistema de coordenadas do cristal.	62
Figura 3.10 - Os perfis das reflexões (regiões)	69
Figura 3.11 - Fotografia do sistema de coleta de dados do R-AXIS IIC	75
Figura 3.12 - Imagem de oscilação do cristal de BthTX-I	76

CAPÍTULO 4

Figura 4.1 - Densidade eletrônica unidimensional; mapa de Patterson	83
Figura 4.2 - Ângulos Eulerianos	86
Figura 4.3 - Ângulos polares esféricos	87
Figura 4.4 - Representação esquemática do método de Subst. Molecular	90
Figura 4.5 - Projeção estereográfica da função de auto-rotação da BthTX-I	95

CAPÍTULO 5

Figura 5.1 (a) - Densidade eletrônica do C-terminal (antes dos omit maps)	113
Figura 5.1 (b) - Densidade eletrônica do C-terminal (mapa atual)	113
Figura 5.2 - Qualidade do mapa de densidade eletrônica resíduos 28 a 33	114

Figura 5.3 - Superposição do monômeros 1 e 2 da BthTX-I	115
Figura 5.4 - Diagrama de Ramachandran da BthTX-I	116
Figura 5.5 - Índices de qualidade program Quality	118
Figura 5.6 - Diagrama dos valores atribuídos pelo programa Verify_3D	119
Figura 5.7 - Gráficos de energia do programa PROSA	120
 CAPÍTULO 6	
Figura 6.1 - Esquema da estrutura secundária da estrutura da BthTX-I	123
Figura 6.2 - Estrutura tridimensional da BthTX-I.	124
Figura 6.3 - As duas moléculas da BthTX-I usadas durante o processo de refinamento	125
Figura 6.4 - Principais diferenças entre os Ca nos dois monômeros da BthTX-I.	126
Figura 6.5 - Diagrama dos fatores de temperatura da BthTX-I	129
Figura 6.6 - Eletroforese da BthTX-I	131
Figura 6.7 - Superposição dos Ca do monômero 1 da BthTX-I e do monômero A da Bothrops asper.	132
Figura 6.8 - Definição do fator de correlação S_c	133
Figura 6.9 - Estruturas diméricas	136
Figura 6.10 - Área de contato entre as subunidades da BthTX-I -GRASP	137
Figura 6.11 - Resíduos envolvidos na interação entre os monômeros da BthTX-I	138
Figura 6.12 - Região do sítio ativo do 2º. monômero da BthTX-I	140
Figura 6.13 - Comparação do loop de ligação do Ca^{+2}	141
Figura 6.14 - Interação da Lys¹²² com o O29 no monômero 1 da BthTX-I	143
Figura 6.15 - Interação da Tyr³¹⁹ com o O229 no monômero 2 da BthTX-I	143
Figura 6.16 - Interação do fosfolípídeo modelado com o monômero 1 da BthTX-I (LIGPLOT).	145
Figura 6.17 - Modelo de Interação entre o fosfolípídeo e a BthTX-I	147

Figura 6.18 (a) Resíduos da interface sugeridos por Salistre de Araujo et al..	149
Figura 6.18 (b) Resíduos da interface sugeridos por Dijkstra e Sigler	150
Figura 6.19 (a) (b) Interação da BthTX-I com a membrana.	151/152
Figura 6.20 (a) Dendrograma baseado em sequências	155
Figura 6.20 (b) Dendrograma baseado em estruturas	156

LISTA DE TABELAS

CAPÍTULO 1

Tabela 1.1 - Os diferentes grupos de fosfolipases A ₂	5
Tabela 1.2 - Alinhamento sequencial de algumas PLA ₂	6
Tabela 1.3 - Porcentagem de identidade entre a BthTX-I e as outras PLA ₂	7
Tabela 1.4 - As sete pontes dissulfeto de uma PLA ₂ do grupo I e uma do grupo II	8

CAPÍTULO 2

Tabela 2.1 - Parâmetros que afetam a cristalização de uma proteína	32
Tabela 2.2 - Principais agentes precipitantes usados na cristalização de proteínas	33

CAPÍTULO 3

Tabela 3.1 - Condições de coleta de dados da BthTX-I utilizando o R-AXIS IIC.	74
Tabela 3.2 - Processamento dos dados da BthTX-I	77
Tabela 3.3 - Estatística dos dados de difração da BthTX-I processados por MOSFLM.	78

CAPÍTULO 4

Tabela 4.1 - Conteúdo de moléculas da cela unitária da BthTx-I	81
Tabela 4.2 - As soluções potenciais obtidas para a estrutura da BthTx-I, programa AMoRe	97

CAPÍTULO 5

Tabela 5.1 - Dados sobre o refinamento e os fatores R e R _{free} da BthTX-I	114
---	-----

CAPÍTULO 6

Tabela 6.1 - Resumo das estruturas secundárias dos dois monômeros da BthTX-I	125
Tabela 6.2 - Classificação das b-turns da BthTX-I	127

Tabela 6.3 - Pontes dissulfeto da BthTX-I	128
Tabela 6.4 - Resultados do programa SPACE (S_c) e as áreas totais WHATIF	135
Tabela 6.5 - Resíduos do canal hidrofóbico	139
Tabela 6.6 - Resíduos da interface de ligação com a membrana.	149

LISTA DE ABREVIATURAS

PLA₂	- Fosfolipase A ₂
BOG	- β- octil glucosídeo
SDS	- Sódio dodecil sulfato
BthTX- I	- Bothropstoxina I
TRIS	- Tris- hidroximetil amino metano
HEPES	- ácido (N-[2-Hidroxietil]piperazina-N'-[2-etanosulfônico])
MES	- ácido (2-[N-Morpholino]etanosulfônico)
SM	- Substituição Molecular
β- MSH	- β- mercaptoetanol

RESUMO

A bothropstoxina I (BthTX-I) é uma miotoxina isolada do veneno da serpente brasileira *Bothrops jararacussu*, a qual é um membro da família das fosfolipases A₂, mas não apresentam atividade catalítica devido á substituição D49K. A proteína for fornecida pelo Prof. Dr. J. R. Giglio e Prof^a. Dra. A. C. O. Cintra do Departamento de Bioquímica da Faculdade de Medicina de Ribeirão Preto e usada em experimentos de cristalização, os quais foram realizados usando a técnica de difusão de vapor “hanging drop” a 18°C. A BthTX-I cristalizou em tampão HEPES 0.1 M, pH variando entre 7.0 e 7.6. O agente precipitante foi o (NH₄)SO₄ em concentrações que variaram de 57% a 62% de saturação.

A coleta de dados foi inicialmente feita utilizando o difratômetro automático R-AXIS IIC da Rigaku Co. do Laboratório de Cristalografia de proteínas do IFSC-USP. Subsequentemente foi realizado uma segunda coleta de dados no SERC Daresbury Laboratory na Inglaterra, usando radiação síncrotron. A BthTX-I cristalizou no grupo espacial P3₁21 com os seguintes parâmetros de rede: a=b=57.58 Å, c= 131.27 Å, α=β=90° e γ=120°.

O processamento de dados foi realizado com o programa MOSFLM, conduzindo a um R_{merge}=6.3% e completeza de 99.6% a uma resolução de 2.1 Å. A estrutura foi resolvida por Substituição Molecular, utilizando o programa AMoRe, onde foi utilizada como modelo inicial a estrutura da miotoxina da serpente *Agkistrodon piscivorus piscivorus* e refinada usando o programa XPLOR que conduziu a um fator R_{final}= 18.7% e R_{free}= 27.4%.

A unidade assimétrica contém dois monômeros, os quais podem ser escolhidos de forma a apresentar interações similares aquelas descritas para a miotoxina II da *Bothrops asper*. A superfície de interface é entretanto, surpreendentemente pequena quando comparada com outras estruturas diméricas e a complementaridade é menor do que o valor esperado.

Um modelo teórico para a ligação do fosfolípídeo na BthTX-I sugere que nenhuma interação direta entre a ligação ester sn-2 e a K49 deve ser esperada de forma a explicar a falta de atividade catalítica.

Foi visto também, que é possível se reproduzir um dendrograma baseado na sequência de amino ácidos, pelo uso de estruturas tridimensionais para os membros da família das PLA₂.

ABSTRACT

Bothropstoxin- I (BthTX-I) is a myotoxin isolated from the Brazilian snake *Bothrops jararacussu* which is a member of the phospholipase A₂ but presents no catalytic activity due to a D49K substitution. Protein was provided from the Departamento de Medicina de Ribeirão Preto by Prof. Dr. J. R. Giglio and Prof^a. Dra. A. C. O. Cintra and used in crystallization experiments which were performed using the vapor diffusion technique in hanging drops at 18°C. The BthTX-I crystallized in 0.1 M HEPES, pH ranging from 7.0 to 7.6. The precipitant was (NH₄)SO₄ in concentrations ranging from 57% to 62%.

The data collection was initially performed using the automatic diffractometer R-Axis IIC from the Rigaku Co. at the Laboratório de Cristalografia of IFSC-USP. Subsequently a second data set was collected at SERC Daresbury Laboratory in England using synchrotron radiation. BthTX-I crystallizes in space group P3₁21 with $a=b= 57.58 \text{ \AA}$, $c=131.27 \text{ \AA}$, $\alpha=\beta=90^\circ$ e $\gamma=120^\circ$.

Processing of the data was performed with the MOSFLM program yielding an $R_{\text{merge}}= 6.3\%$ and completeness of 99.6% at 2.1 $\text{\AA}$ resolution. The structure was solved by Molecular Replacement using the package AMoRe with the *Agkistrodon piscivorus piscivorus* enzyme as search model and refined using the refinement program X-PLOR to a $R_{\text{final}}= 18.7\%$ and $R_{\text{free}}= 27.4\%$.

The asymmetric unit contains two BthTX-I monomers, which can be chosen such that they present similar interactions to those described for the homologous myotoxin II from *Bothrops asper*. The interface is, however, surprisingly small when compared to other dimeric structures and is less complementary than expected.

A theoretical model for phospholipid binding to BthTX-I suggests that no direct interactions between the sn-2 ester bond and K49 would be expected, thus explaining the lack of catalytic activity.

It is shown that it is possible to reproduce a dendrogram based on amino acid sequences and by the use of three dimensional structures for members of the PLA₂ family.

CAPÍTULO 1

FOSFOLIPASES A₂

1.1 Introdução

Diversas técnicas físicas têm sido desenvolvidas para o estudo de materiais biológicos. As proteínas têm recebido especial atenção por formar juntamente com os ácidos nucleicos, carboidratos e lipídeos a base bioquímica macromolecular dos organismos vivos. A estrutura das proteínas tem relação direta com sua função e a determinação de sua estrutura tridimensional possibilita entender qual seu mecanismo de atividade. As técnicas de difração de raios-X e ressonância magnética nuclear têm dado uma importante contribuição nesta área de estudos e tornaram-se áreas de grande importância na enzimologia, biologia molecular e imunologia devido a sua capacidade de produzir modelos estruturais detalhados de macromoléculas em termos das coordenadas tridimensionais de seus átomos. Neste trabalho utilizaremos a técnica de cristalografia de raios-X para o estudo da relação estrutura-função de uma proteína, a bothropstoxina I (BthTX-I), uma miotoxina tipo fosfolipase A₂ (PLA₂) extraída do veneno da cobra *Bothrops jararacussu* (Cintra *et al.*, 1993).

1.2 Classificação das Fosfolipases

As fosfolipases são enzimas que catalisam a hidrólise dos fosfolipídeos. As fosfolipases são agrupadas de acordo com suas especificidades. As ligações hidrolisadas pelas fosfolipases A₁, A₂, C e D são mostradas na Fig. 1.1, onde R₁ e R₂ se referem aos ácidos graxos nas posições sn-1 e sn-2 e X se refere a um grupo polar (Dennis,1983).

Fig. 1.1: Especificidade das fosfolipases mostrando os sítios das fosfolipases A₁,A₂, C e D.

Em geral, as fosfolipases têm especificidade para a posição no fosfolipídeo mais do que para um particular ácido graxo ou grupo polar presente, embora preferências (especialmente para os últimos) têm também sido encontradas. Na sua grande maioria são enzimas digestivas presentes em altas concentrações no suco intestinal, secreções bacterianas e venenos.

1.3 Fosfolipases A₂ (PLA₂)

As fosfolipases A₂ também conhecidas como fosfatidil acil-hidrolases, são enzimas lipolíticas que agem especificamente na catálise da hidrólise da ligação éster na posição 2-acil (sn-2) dos L-fosfolipídeos, como pode ser observado na Fig. 1.2 (Dijkstra *et al.*, 1981).

Fig. 1.2: Mecanismo de atividade de uma fosfolipase A₂.

Estas enzimas pequenas e globulares podem ser classificadas tanto como enzimas extracelulares, exemplificadas pela fosfolipase pancreática bovina e em diversos tipos de venenos, ou como intracelulares, tais como as envolvidas nas inflamações (Holland *et al.*, 1990). As fosfolipases A₂ intracelulares têm interesse particular farmacêutico devido ao seu papel nas numerosas respostas celulares. Entretanto, ao contrário do tipo extracelular, estas enzimas são produzidas em pequenas quantidades e na maioria das vezes são enzimas associadas à membrana tornando difícil a sua separação e caracterização (Holland *et al.*, 1990).

Aqui, daremos especial atenção às enzimas obtidas do pâncreas do porco e do boi e dos venenos de cobra e abelha, porque estas enzimas têm sido estudadas com maior detalhe e estão diretamente relacionadas com a BthTX-I, a toxina de estudo.

O veneno de cobra é um complexo de matérias proteicas com vários efeitos fisiológicos e cada espécie de veneno tem características e toxicidades diferentes.

De acordo com os efeitos fisiopatológicos, os venenos de cobra podem ser distribuídos em dois grandes grupos (Cintra *et al.*,1993):

1. *Venenos das famílias Elapidae e Hydrophidae, os quais são altamente neurotóxicos, causando paralisia e a morte devido a insuficiência respiratória.*

2. *Venenos das famílias Viperidae e Crotalidae, os quais induzem efeitos locais, tais como: hemorragia, necrose e edemas; porém não produzem efeitos neurotóxicos pronunciados. A Bothrops jararacussu pertence à família das Viperidae.*

Embora o estudo dos efeitos neurotóxicos dos venenos das Elapidae e Hydrophidae têm progredido rapidamente nas últimas décadas, pouco é conhecido dos efeitos locais dos venenos Crotalidae e Viperidae. Entre estes efeitos a mionecrose é, particularmente, importante uma vez que produz alterações musculares que podem resultar na morte do tecido.

Muitas fosfolipases são atualmente conhecidas e a maioria delas são fosfolipases A₂ (PLA₂). Apesar das similaridades, há certas diferenças estruturais que permitem a classificação destas enzimas em diferentes grupos. O grupo I compreende as enzimas encontradas principalmente no pâncreas de mamíferos e nos venenos Elapídicos e Hidrofídicos e o grupo II, onde se encontram as enzimas dos venenos Crotalídicos e Viperídicos (Dennis, 1994). Tradicionalmente, estas enzimas extracelulares foram divididas em três grupos principais e alguns subgrupos baseado na sua sequência de amino ácidos. Nos últimos anos novas PLA₂ foram descritas que não se encaixavam em nenhum destes grupos tradicionais, tornou-se, então, necessária uma nova classificação destas enzimas. Os grupos I, II e III são enzimas cujas características principais são o alto conteúdo de pontes dissulfeto, baixo peso molecular e são enzimas que necessitam do ion Ca⁺² para a catálise. O grupo IV é um grupo distinto. Este grupo inclui as enzimas intracelulares de alto peso molecular, dependentes também do cofator Ca⁺² e que são específicas na ativação do ácido araquidônico. A ativação do ácido araquidônio por PLA₂ é o passo inicial para a formação de prostaglandinas e tromboxanas que são mediadores importantes para o desenvolvimento de uma resposta inflamatória.. O grupo V é ainda um grupo

pouco estudado e com características não muito bem definidas. Na tabela 1.1 (Dennis, 1994) encontramos uma classificação dos diferentes grupos de fosfolipases A₂, assim como algumas das suas características, tais como: peso molecular e se o seu mecanismo de atividade é dependente ou não do cofator Ca⁺², pontes dissulfeto, número de cisteínas e número de aminoácidos.

Grupos	Localização	P.M.(kDa)	Ca⁺²	S-S	Cys	# Resíduos
Grupo I A Cobras (Naja) B Pâncreas: Porco/Bov/Hum	Secretada	13 - 15	mM	5 - 7	10 - 14	≈ 125
Grupo II A Cascavéis e Víboras Fluido Sinuvial Humano/ Plaquetas PLA₂ (inflamação)	Secretada	13 - 15	mM	5 - 7	10 - 14	≈ 125
Grupo III Abelhas/Lagartos	Secretada	16 - 18	mM	5 - 7	10 - 14	≈ 125
Grupo IV Plaquetas humanas/U937	Citosólica	≈ 85	μM	0	9	≈ 750
Grupo V Miocárdio Canino/Humano	Citosólica	≈ 40	Não	—	—	—

Tabela 1.1: Os diferentes grupos de fosfolipases A₂ e algumas de suas características.

1.3.1 Análise Seqüencial de Algumas das Fosfolipases A₂

A tabela 1.2, apresenta um alinhamento seqüencial de algumas das fosfolipases A₂ que têm estrutura cristalina determinada e cujas coordenadas estão depositadas no banco de dados, Brookhaven . Estas seqüências representam apenas uma pequena amostragem das estruturas primárias conhecidas (Ogawa *et al.*, 1995) (Davidson; Dennis, 1990) mas são as mais relevantes ao trabalho desenvolvido e descrito nesta tese.

	10	20	30	40
pancreática bovina	-ALWQFNGMI	KCKIPSSEPL	LDFNNYGCYC	GLGSGGTPV- ----DGLDR
pancreática porco	-ALWQFRSMI	KCAIPGSHPL	MDFNNYGCYC	GLGSGGTPV- ----DELDR
Naja naja atra	-NLYQFKNMI	QCTVP-SRSW	WDFADYGCYC	GRGSGGTPV- ----DDLDR
Naja naja naja	-NLYQFKNMI	KCTVP-SRSW	WDFADYGCYC	GRGSGGTPV- ----DDLDR
humana	-NLVNPHRMI	K-LTYGKEAA	LSYGFYGCYC	GVGGRGSPK- ----DATDR
Crotalus atrox	-SLVQFETLI	M-KIAGRSGL	LWYSAYGCYC	GWGGHGLPQ- ----DATDR
Bothrops asper	-SLFELGKMI	L-QETGKNPA	KSYGAYGCNC	GVLGGRGKPK- ----DATDR
BthIX-I	-SLFELGKMI	L-QETGKNPA	KSYGAYGCNC	GVLGGRGKPK- ----DATDR
Aglistrodon	-SVLELGKMI	L-QETGKNAI	TSYGSYGCNC	GWGHRGQPK- ----DATDR
Abelha	-----II	-----	-YPC-TLWC	GHNKSSGPN ELGRFKHTDA
	50	60	70	80
pancreática bovina	CCQTHDNCYK	QAKKLDSECKV	LVDNPHYTNMY	SYSCSNNEIT CSSENNACBA
pancreática porco	CCETHDNCYR	DAKNLDSCKP	LVDNPHYTESY	SYSCSNTEIT CNSKNNACBA
Naja naja atra	CCQVHDNCYN	EAEKISGCW-	----PYFKTY	SYECSQGTLT CKGGNNACAA
Naja naja naja	CCQVHDNCYN	EAEKISGCW-	----PYFKTY	SYECSQGTLT CKGDNNACAA
humana	CCVTHDCCYK	RLEK-RCCG-	----TKFLSY	KFSNNGSRIT C-AKQDSCRS
Crotalus atrox	CCFVHDCCYK	KAT---DCN-	----PKTVSY	TYSEENGELI C-GGDDPCGT
Bothrops asper	CCYVHKCCYK	KLT---GCN-	----PKKORY	SYSWKDKTIV C-GENNSCLK
BthIX-I	CCYVHKCCYK	KLT---GCN-	----PKKORY	SYSWKDKTIV C-GENNPCLK
Aglistrodon	CCFVHKCCYK	KLT---DCN-	----HKTDY	SYSWKKNKAI C-BEKNPCLK
Abelha	CCRTHDMC--	-----	-----PDVM	SAGESKHGLT ---NTASHTR
	100	110		120
pancreática bovina	-ICNCDRNAA	ICFSK-----	VPY-----	---NKEHKNL D-KKN-C---
pancreática porco	-ICNCDRNAA	ICFSK-----	APY-----	---NKEHKNL DTKKY-C---
Naja naja atra	-VCUCDRLAA	ICFAG-----	APY-----	---NDNDYNI NLKAR-C---
Naja naja naja	-VCUCDRLAA	ICFAG-----	APY-----	---NDNDYNI DLKAR-CQ--
humana	-LCECDKAAA	TCFARN---K	TTY-----	---NKK-YQY YSNKH-CRGS
Crotalus atrox	-ICECDKAAA	ICFRDN---I	PSY-----	---DNK-YWL FPPKD-CREE
Bothrops asper	-LCECDKAVA	ICLREN---L	NPY-----	---NKK-YRY YLKPL-CK-K
BthIX-I	-LCECDKAVA	ICLREN---L	GFY-----	---NKK-YRY HLMKF-CK-K
Aglistrodon	-MCECDKAVA	ICLREN---L	DTY-----	---NKK-YKA YEKLK-CK-K
Abelha	-SCDCDDKFY	DCLKNSADTI	SSYFVGKMYF	NLIDTKCYKL EHPVTGCGER
	130	140		
pancreática bovina	-----	-----	-----	
pancreática porco	-----	-----	-----	
Naja naja atra	-----	-----	-----	
Naja naja naja	-----	-----	-----	
humana	TP-RC-----	-----	-----	
Crotalus atrox	PE-PC-----	-----	-----	
Bothrops asper	AD-AC-----	-----	-----	
BthIX-I	AD-PC-----	-----	-----	
Aglistrodon	FD-TC-----	-----	-----	
Abelha	TEGRCLHYTV	DKSKPKVYQW	FDLRK	

Tabela 1.2: Alinhamento sequencial de algumas fosfolipases A₂.

Da análise da tabela 1.2, podemos observar que as fosfolipases A₂ de diversas origens exibem um alto grau de identidade sequencial, sendo que vários resíduos permanecem invariantes. Como pode ser observado a sequência da fosfolipase A₂ do veneno de abelha é a que apresenta menor identidade sequencial, entretanto há conservação dos resíduos funcionais

das PLA₂. Na tabela 1.3, podemos observar as porcentagens de identidade entre a BthTX - I e outras PLA₂.

Proteína	% de Identidade
Pancreática porcina	40.7
<i>Naja naja naja</i> (Elapidae)	39.5
<i>Agkistrodon p.p.</i> (Viperidae)	76.0
<i>Crotalus atrox</i> (Crotalidae)	47.1
Abelha	19.3
<i>Naja naja atra</i> (Elapidae)	39.8
<i>Bothrops asper</i> (Viperidae)	95.9
Pancreática bovina	43.4
Fluído sinovial Humano	51.2

Tabela 1.3: Porcentagem de identidade entre a Bothropstoxina I (a sequência de estudo) e as outras fosfolipases A₂ apresentadas no alinhamento anterior.

Passaremos agora, a analisar em detalhe a sequência de aminoácidos das várias fosfolipases A₂ e avaliar a importância de alguns destes resíduos invariantes.

1.3.1.1 Pontes de Dissulfeto

Como pode ser observado da tabela 1.1 as fosfolipases A₂ do grupo I/II/III possuem entre 10 e 14 cisteínas que formam 5 a 7 pontes de dissulfeto. Na tabela 1.4 são apresentadas as sete pontes dissulfeto da fosfolipase A₂ pancreáticas bovina (grupo I) (Dijkstra *et al.*, 1981) e as sete pontes dissulfeto do veneno da cobra *Crotalus atrox* (grupo II) (Keith *et al.*, 1981), seis das quais são comuns entre os dois grupos.

PLA ₂ -Pancreática bovina	PLA ₂ - <i>Crotalus atrox</i>
Pontes dissulfeto	Pontes dissulfeto
Cys ¹¹ - Cys ⁷⁷	
Cys ²⁷ - Cys ¹²³	Cys ²⁷ - Cys ¹²⁶
Cys ²⁹ - Cys ⁴⁵	Cys ²⁹ - Cys ⁴⁵
Cys ⁴⁴ - Cys ¹⁰⁵	Cys ⁴⁴ - Cys ¹⁰⁵
Cys ⁵¹ - Cys ⁹⁸	Cys ⁵⁰ - Cys ¹³³
Cys ⁶¹ - Cys ⁹¹	Cys ⁵¹ - Cys ⁹⁸
Cys ⁸⁴ - Cys ⁹⁶	Cys ⁶¹ - Cys ⁹¹
	Cys ⁸⁴ - Cys ⁹⁶

Tabela 1.4: As sete pontes de dissulfeto de uma fosfolipase A₂ do grupo I e uma do grupo II. A numeração utilizada aqui é obtida do alinhamento das duas sequências e não do alinhamento apresentado anteriormente.

A diferença estrutural entre estes dois grupos consiste na presença de uma ponte de dissulfeto invariável entre os resíduos Cys¹¹ e Cys⁷⁷ no grupo I, e que está ausente nas PLA₂ do grupo II.

Para formar uma ponte dissulfeto, não apenas a distância entre os Cα das duas cisteínas envolvidas deve estar entre 4 Å e 7.5 Å, mas também os vetores Cα – Cβ devem apresentar o tipo certo de relação (Richardson, J.S.;Richardson, D.C., 1989).

As conformações mais comuns para as pontes dissulfeto são:

I) Espiral de mão esquerda Fig 1.3 (a) a qual tem os vetores Cα – Cβ quase antiparalelos ($C\alpha - C\beta - C\beta - C\alpha \cong 180^\circ$) e o ângulo χ_3 ($C\beta - S - S - C\beta \cong -85.8^\circ \pm 10.7^\circ$).

II) Gancho de mão direita Fig. 1.3 (b) , no qual os vetores Cα – Cβ são aproximadamente perpendiculares e χ_3 $96.8^\circ \pm 14.8^\circ$.

(a)

(b)

Fig. 1.3: Classificação das pontes dissulfeto;
(a) Espiral de mão esquerda
(b) Gancho de mão direita.

1.3.1.2 N-terminal

Vários experimentos têm mostrado a importância da região N-terminal (Díaz *et al.*, 1994) (Kato *et al.*, 1994). A remoção do octapeptídeo NH₂-terminal da miotoxina II de *B. asper* reduziu significativamente o efeito de desorganização da membrana promovida pela toxina bem como a atividade miotóxica desta proteína. Estes trabalhos têm sugerido que esta região está envolvida diretamente no acesso do substrato ao sítio catalítico. Como pode ser observado na tabela do alinhamento sequencial, é uma região de resíduos anfifílicos. Um resíduo altamente conservado é a Met na posição 8 e uma Ile na posição 9.

Na Fig. 1.4 podemos avaliar um diagrama esquemático das interações entre o N-terminal e o sítio ativo na fosfolipase A₂ pancreática bovina (Dijkstra *et al.*, 1981). O grupo $\alpha\text{-NH}_3^+$ da Ala¹ está enterrado e interage com a carbonila da Gln⁴, a carbonila da Asn⁷¹ e uma molécula de água que interage por sua vez com o sítio ativo (His⁴⁸ - Asp⁹⁹). O N-terminal e esta molécula de água parecem importantes para a manutenção da estrutura do sítio ativo, e portanto para a atividade catalítica.

Fig. 1.4: Desenho esquemático das interações entre o N-terminal e o sítio ativo na fosfolipase A₂ pancreática bovina.

1.3.1.3 Sítio de Ligação do Cálcio

O íon cálcio (Ca^{+2}) é um cofator essencial para a atividade da enzima. Os resíduos 25 a 33 fazem parte do loop de ligação do Ca^{+2} . A Fig. 1.5 mostra uma representação esquemática do íon cálcio e seus ligantes na fosfolipase A₂ pancreática bovina (Dijkstra *et al.*, 1981). O Asp⁴⁹ é fundamental para sua ligação devido ao fato de proporcionar uma carga para equilibrar o Ca^{+2} e em fornecer dois átomos de oxigênio que servem como ligantes.

Fig. 1.5: Representação esquemática do íon cálcio e seus ligantes na fosfolipase A₂ pancreática bovina (Dijkstra *et al.*, 1981).

Algumas PLA₂ miotóxicas não neurotóxicas apresentam uma substituição crítica do D49 por uma K49 no sítio de ligação do Ca⁺². A presença do aspártico 49 foi inferida pelos autores como um resíduo essencial para a hidrólise dos fosfolípidos (Van den Berg *et al.*, 1989) (Li *et al.*, 1994). Com esta substituição a enzima fica desprovida de qualquer atividade de fosfolipase A₂ (ou atividade muito baixa) uma vez que a região nas K49 PLA₂ correspondente ao sítio de ligação do cálcio em outras fosfolipases é ocupado pelo conjunto ε -amino da Lys⁴⁹ (Holland *et al.*, 1990), mas elas são muito ativas na indução de necroses. O mecanismo de ação destas miotoxinas é ainda desconhecido. O sítio ativo envolvido na ligação de receptores ou na indução da miotoxicidade ainda não são conhecidos. Existem evidências da existência de dois sítios independentes de atividade destas PLA₂, isto é, um sítio de atividade catalítica e outro sítio diretamente envolvido com o carácter miotóxico destas enzimas (Selistre de Araujo *et al.*, 1996). Outra possibilidade é que ambos os sítios são o mesmo e que a presença de vários resíduos básicos na superfície da molécula induzem o seu carácter miotóxico (Kini; Iwanaga, 1986).

1.3.1.4. Sítio Ativo

O sítio ativo, Fig. 1.4, está localizado numa depressão na superfície da molécula. Embora dissimilaridades ocorram na sequência de aminoácidos das várias fosfolipases A₂ os três resíduos funcionalmente importantes permanecem absolutamente conservados His⁴⁸, Tyr⁷³ e Asp⁹⁹. O par de resíduos Asp⁹⁹ e His⁴⁸ estão diretamente relacionados com a ação catalítica das enzimas cujo processo é semelhante ao das serino proteases mas com uma molécula de água agindo como nucleófilo no lugar da serina catalítica (Dijkstra *et al.*, 1981).

1.3.1.5 Mecanismo Catalítico das Fosfolipases A₂

O mecanismo catalítico da atividade das PLA₂'s de baixo peso molecular continua um assunto de especulação, mas um grande número de estudos envolvendo modificações estruturais e químicas têm sugerido um mecanismo similar ao sistema de revezamento de prótons das enzimas proteolíticas, onde a tríade catalítica é formada pela His⁴⁸ e Asp⁹⁹ e uma molécula de água (Dennis, 1983).

Na ausência de substrato, a His⁴⁸ não está protonada. No entanto, está pronta para aceitar um próton da água. Estudos mostraram que o próton envolvido na ligação de hidrogênio entre a His e Asp fica na His⁴⁸ e que o Asp⁹⁹ permanece com carga negativa (Dennis, 1983). O papel do grupamento -COO⁻ do Asp⁹⁹ é de estabilizar a forma da His com carga positiva no estado de transição. O anel imidizólico com carga positiva é estabilizado eletrostaticamente com o Asp⁹⁹ com carga negativa, como mostrado na Fig. 1.6 (a). Neste mecanismo, a His⁴⁸ e a água participam diretamente na clivagem da ligação do substrato. A hidrólise da ligação com um ataque do átomo de oxigênio da água ao carbono carbonílico da ligação susceptível. A ligação carbono-oxigênio dessa carbonila torna-se uma ligação simples e o átomo de oxigênio adquire uma carga negativa, o oxiânion como pode ser observado na Fig. 1.6 (b). Na última etapa o próton seguro pela forma protonada da His é então doado ao átomo da ligação susceptível que é assim clivada, Fig. 1.6 (c) (Dennis, 1983).

Fig. 1.6: Mecanismo catalítico das fosfolipase A₂.

Uma descrição química da ação das fosfolipases A₂ pôde agora ser melhor entendida a partir da resolução de três estruturas cristalinas a alta resolução: (1^a) a estrutura cristalina da fosfolipase A₂ do veneno da cobra chinesa *Naja naja atra* com um análogo do estado de transição, diC₈(2Ph)PE (White *et al.*, 1990); (2^a) a estrutura cristalina da fosfolipase A₂ do veneno de abelha com um complexo similar ao descrito acima (Scott *et al.*, 1990) e (3^a) a estrutura nativa da fosfolipase A₂ do veneno da cobra *Naja naja atra* (Scott *et al.*, 1990). Apesar das diferenças estruturais entre as fosfolipase A₂ do veneno da cobra *Naja naja atra* e do veneno da abelha, o complexo interage de maneira análoga com o sítio ativo nas duas estruturas. O diC₈(2Ph)PE forma um complexo estável com a PLA₂ pois possui um segundo grupo fosfato

na posição sn-2. Por ter geometria tetraédrica e possuir carga negativa este fosfato imita o estado de transição da reação catalizada.

Nestas estruturas podemos observar a presença de dois íons Ca^{+2} ligados a cada um dos monômeros das estruturas sendo um no sítio convencional e o outro num segundo sítio. Na estrutura cristalina *Naja naja atra* nativa, os dois íons cálcio estão separados por uma distância de 11,1 Å, sendo que na estrutura com o análogo complexo o segundo íon cálcio está a apenas uma distância de 6.6 Å do primeiro íon cálcio no loop de ligação convencional do íon cálcio, como descrito anteriormente. Nestes trabalhos, os autores propõem um mecanismo mais detalhado para as fosfolipases A₂. Fig. 1.7 mostra em detalhe o mecanismo catalítico proposto pelos autores, baseado na ligação do análogo $\text{diC}_8(2\text{Ph})\text{PE}$. O mecanismo de catálise é semelhante ao descrito anteriormente. Devido à ação da molécula de água, observa-se a presença de uma carga negativa no oxianion do estado intermediário, interagindo com o íon Ca^{+2} . Porém formalmente a carga negativa do Asp⁴⁹ e do fosfato sn-3 já neutraliza a carga positiva do íon Ca^{+2} mesmo antes da contribuição do oxianion do intermediário tetraédrico ser considerada.

Os autores propõem que o segundo íon Ca^{+2} é coordenado pelo O29, o oxigênio carbonil da ligação peptídica formada com o NH da G30. Podemos ver que o NH da G30 interage com o oxianion do intermediário tetraédrico e que portanto, a posição do segundo íon Ca^{+2} neste caso, pode neutralizar e estabilizar o sítio oxianico através de polarização desta ligação peptídica. Os dois cálcios estariam relacionados com o processo e mecanismo de catálise através da estabilização do oxianion do intermediário tetraédrico ou diretamente (no caso do primeiro Ca^{+2}) ou indiretamente (no caso do segundo Ca^{+2} assim reduzindo a energia de ativação da reação).

Não existem estudos bioquímicos direcionados para a função ou mesmo a existência deste segundo íon de Ca^{+2} . Se este mecanismo é geral para as PLA₂, deveríamos encontrar um segundo íon Ca^{+2} na estrutura cristalina de qualquer fosfolipase A₂ ou qualquer outra carga positiva nas vizinhanças da ligação peptídica entre o resíduo 29 e 30. Analisando este novo ponto de vista notamos que a estrutura cristalina da fosfolipase A₂ pancreática bovina apresenta uma Lys¹²² interagindo com o O29 através de uma molécula de água, portanto, também em princípio induzindo a polarização da ligação peptídica.

Devemos ressaltar porém, que no caso da estrutura da abelha não foi encontrado uma segunda carga positiva, nenhum metal e nenhum resíduo básico nesta região.

Fig. 1.7: Mecanismo catalítico proposto por *White et al.*, 1990 baseado no complexo da PLA₂ de *Naja naja atra* e o inibidor diC₈(2Ph)PE.

1.3.1.6 Ativação Interfacial e o Canal Hidrofóbico

Um outro aspecto importante a ser analisado está diretamente relacionado com o mecanismo de ativação interfacial de ligação do substrato com a enzima. Vários experimentos mostram que as fosfolipases A₂ têm uma atividade maior em relação a agregados, micelas, filmes ou membranas do que a moléculas simples de fosfolípidos (Van den Berg *et al.*, 1995; Kilby *et al.*, 1995). Este fenômeno de ativação enzimática na interface lípideo-água é chamado de ativação interfacial. Estruturas resolvidas por NMR (Van den Berg *et al.*, 1995) sugerem, ao contrário de estruturas resolvidas cristalográficamente, que mudanças conformacionais da enzima desempenham um papel importante na ativação interfacial da enzima. Vários modelos têm sido sugeridos para explicar este fenômeno de ativação enzimática (Kilby *et al.*, 1995; Scott *et al.*, 1990; Van den Berg *et al.*, 1995).

Da resolução da estrutura da *Naja naja atra* PLA₂ com o complexo análogo, observou-se que os sítios sn-1 e sn-2 estão paralelamente orientados num canal hidrofóbico que se estende por 14 Å aproximadamente, desde o sítio catalítico até uma abertura pouco acima da hélice N-terminal (resíduos 1-12) e no veneno da abelha os resíduos 76 ao 90. A superfície do canal é formada principalmente por resíduos hidrofóbicos altamente conservados. Na parede direita do canal está a Ile⁹ e a invariante Phe⁵ (sn-2). A parede esquerda é formada pela invariante Leu(Val)². A parte superior do canal está o Trp¹⁹. Também na parede esquerda está um resíduo altamente hidrofóbico e bastante flexível. Na enzima *Naja naja atra*, esta parede é o anel da Tyr⁶² que está ligado ao inibidor pelo fosfato sn-3, o qual, por sua vez está firmemente ancorado ao íon Ca⁺². Em alguns membros do grupo I/II o resíduo 62 é uma Lisina, onde neste caso o quarto grupo metileno da cadeia lateral, provavelmente forma a parede hidrofóbica do canal e o ε - amino grupo se liga ao fosfato sn-3. Na Fig.1.8 podemos observar estes resíduos que fazem parte do canal hidrofóbico, no complexo de diC₈(2Ph)PE com o monômero A da estrutura de *Naja naja atra*. No veneno da abelha, esta Tyr/Lys⁶² é substituída pela Arg⁵⁷ e os resíduos do canal hidrofóbico e da superfície de ligação do substrato são: Ile¹, Tyr³, Cys⁹, His¹¹, Thr⁵⁶, Arg⁵⁷, Leu⁵⁹, Val⁸³, Met⁸⁶, Tyr⁸⁷ e Ile⁹¹.

Fig.1.8: Resíduos do canal hidrofóbico do monômero A da estrutura da *Naja naja atra* quando da ligação diC₈(2Ph)PE. Na parede direita do canal hidrofóbico está a Ile⁹ e a Phe⁵. A parede esquerda é formada pela Leu² e o resíduo Tyr⁶². Na parte superior do canal está o Trp¹⁹.

Na Fig. 1.9 vemos uma representação esquemática da interação entre uma PLA₂ e um agregado de substrato. Nesta figura está representado o canal hidrofóbico, a superfície de ativação interfacial assim como os resíduos do sítio ativo (Voet, D.; Voet, J.G., 1995).

Fig. 1.9: Representação esquemática do canal hidrofóbico, da superfície de ativação e de alguns dos resíduos do sítio catalítico (Voet,D.; Voet,J.G., 1995).

Da comparação da estrutura *Naja naja atra* nativa e de sua estrutura com complexo há indicações de que a condição ótima de ligação do substrato e catálise na interface é devido à grande facilidade de difusão do substrato a partir da superfície de ativação em direção ao sítio catalítico, uma vez que não foram observadas variações conformacionais significativas na enzima na forma livre e complexada. Apoiando esta idéia, é interessante notar que a

conformação do análogo no complexo é parecida com aquela encontrada por RMN em membranas (Hauser,Phillips, 1979; Seelig, 1977) e em estruturas cristalográficas (Elder *et al.*, 1977; Hauser *et al.*, 1981; Pearson, Pascher, 1979;). Estes autores concluíram também que não há evidência para a necessidade da formação de dímeros ou maiores oligômeros para a atividade fosfolipídica na interface (Cho *et al.*, 1988; Jain, Berg, 1989; Lombardo, Dennis, 1985; Romero *et al.*, 1987).

1.3.1.7 Resíduos do C-terminal

Esta porção da molécula varia muito entre as PLA₂ de diferentes fontes como pode ser observado na tabela do alinhamento sequencial. Nas estruturas das K49 PLA₂ esta região possui um alto conteúdo de lisinas (Selistre de Araujo *et al.*, 1996) com cadeias laterais flexíveis e estendidas no solvente (Holland *et al.*, 1990). Estas lisinas não são conservadas nas D49 PLA₂ e poderiam estar relacionadas na interação com a membrana e/ou toxicidade (Gopalakrishnakone *et al.*, 1984).

1.4 Estrutura Secundária das PLA₂

As fosfolipases A₂ de diversas origens exibem um alto grau de homologia estrutural caracterizada pelos seguintes elementos: (Holland *et al.*, 1990):

1. N-terminal (uma hélice α de 12 resíduos- hélice 1)
2. Um loop relativamente longo contendo o sítio de ligação do cálcio.
3. Um par de hélices α relativamente longas e paralelas - hélices 2 e 3
4. Duas folhas β pequenas e antiparalelas (asa- β - A e B).
5. C-terminal variável.

A Fig. 1.10 mostra a estrutura tridimensional da PLA₂ pancreática bovina uma D49 PLA₂ (Dijkstra *et al.*, 1981) e esta figura também mostra a distribuição das pontes dissulfeto.

Fig. 1.10: Estrutura tridimensional da PLA₂ pancreática bovina
Também são mostradas as sete pontes dissulfeto e o íon Ca⁺² (esfera branca)

A estrutura da fosfolipase A₂ do veneno de abelha difere consideravelmente das demais estruturas tridimensionais de PLA₂. Embora ela apresente variações significativas estruturais e sequenciais, existe uma conservação dos principais resíduos envolvidos no mecanismo catalítico

1.5 BthTx-I (Bothropstoxina I)

A BthTX-I é uma miotoxina isolada do veneno da cobra *Bothrops jararacussu* (Cintra *et al.*, 1993). As miotoxinas estão presentes em uma série de venenos de cobras e são toxinas capazes de provocar dano aos músculos, efeito este conhecido como miotoxicidade. Muitas delas, além da miotoxicidade apresentam outras atividades biológicas como neurotoxicidade, cardiotoxicidade e atividades inflamatórias (Selistre de Araujo *et al.*, 1996). Algumas delas são

PLA₂ enzimaticamente ativas (PLA₂-D49) e outras são miotoxinas com deficiência ou atividade PLA₂ muito baixa, que pertencem ao grupo de PLA₂ análogas (PLA₂-K49) (Francis *et al.*, 1991; Homs-Brandeburgo *et al.*, 1988; Liu *et al.*, 1990). Até agora estas toxinas só foram descritas para membros da família *Viperidae*. A análise da sequência de aminoácidos, tabela 1.2 mostra que a BthTX-I é uma K49 PLA₂ composta de 121 aminoácidos. Embora desprovida de qualquer atividade fosfolipídica a BthTX-I apresenta homologia sequencial com as D49 PLA₂ e com outras miotoxinas (Cintra *et al.*, 1993). Mutações críticas estão aparentemente envolvidas com a diminuição ou eliminação da atividade PLA₂, particularmente a substituição D49K.

Apesar desta ausência de atividade enzimática de PLA₂ estas proteínas apresentam semelhanças com as PLA₂ quanto às propriedades bioquímicas tais como peso molecular e composição em aminoácidos (Dennis, 1994)

A falta de atividade enzimática de algumas destas PLA₂ miotóxicas torna-se um assunto bastante estudado uma vez que elas apresentam similaridade química e imunológicas com as PLA₂ enzimaticamente ativas (Maraganore; Heinrikson, 1986). A relação estrutura-função de PLA₂ ainda não está totalmente clara. Vários ensaios têm sido utilizados para a identificação de sítios farmacológicos específicos para a neurotoxicidade e miotoxicidade. Estudos de modificações químicas demonstram que é possível dissociar a atividade enzimática das atividades farmacológicas (Verheij *et al.*, 1980). Vários resíduos de interesse biológico têm sido modificados nas PLA₂ usando mutagênese sítio dirigida para se estudar o comportamento catalítico e tóxico destas proteínas (Shen; Cho, 1995; Tzeng *et al.*, 1995).

Num trabalho recentemente desenvolvido (Selistre de Araujo *et al.*, 1996) é feita uma comparação minuciosa das seqüências das várias fosfolipases A₂ D49 e K49 com ou sem atividade miotóxica, para se avaliar quais os resíduos que dariam esta característica miotóxica à enzima assim como avaliar quais os resíduos que estariam diretamente envolvidos na interação da enzima com a membrana. Alguns resíduos indicados são K7, K78, K80, K115, K116, E12, T13, K16 e N17. Possivelmente a falta de atividade fosfolipídica nas K49 é devido a várias destas variações e não somente, à substituição do Asp na posição 49 por uma Lys. Outros autores (Scott *et al.*, 1990; Dijkstra *et al.*, 1981) têm sugerido conjuntos de resíduos provavelmente envolvidos na interação com a interface fosfolipídica nas fosfolipases A₂ de um modo geral. Estes resíduos estão localizados em volta da entrada do sítio ativo e do canal hidrofóbico. Estudos recentes com um peptídeo sintético mostram a importância da região C-terminal para a citotoxicidade em células endoteliais. Esta região ausente em algumas toxinas

não deve ser responsável pela miotoxicidade, porém pode contribuir para a toxicidade em geral. Com a alta densidade de cargas positivas, a região C-terminal parece ser um dos sítios de interação com a heparina, que é capaz de bloquear a ação miotóxica de várias toxinas (Dua; Cho, 1994). Existem fortes evidências de que a atividade catalítica de uma PLA₂ não é necessária para a miotoxicidade.

O mecanismo de ação destas miotoxinas é ainda desconhecido, mas pode ser mediado através da análise da interação com supostos receptores de alta afinidade das PLA₂ recentemente descritos (Lambeau *et al*, 1989; Lambeau *et al.*, 1994; Lambeau *et al.*, 1995; Yen, Tzeng, 1991; Krizaj *et al*, 1994). O sítio molecular das moléculas de PLA₂ envolvidos na ligação destes receptores ou na indução da miotoxicidade é ainda desconhecido.

A existência de uma região catiônica nos resíduos 79-87 foram sugeridos como sendo responsáveis pelo carácter miotóxico das PLA₂, mas nenhum dado experimental foi obtido que comprovasse esta hipótese (Kini; Iwanaga, 1986). Também já foi sugerido que a região N-terminal desempenha um papel importante na miotoxicidade (Diaz *et al.*, 1994).

Foi observado que alguns dos resíduos conservados nas D49 PLA₂ tais como G32 e G33 não são conservados nas K49 PLA₂. Estes dois resíduos estão diretamente envolvidos no loop de ligação do Ca⁺² que é o cofator essencial para a atividade fosfolípida. Como as K49 PLA₂ não tem atividade fosfolípida é de se esperar que a pressão seletiva agindo na conservação da sequência deste loop seja reduzida.

1.6 Monômero X Dímero

Algumas PLA₂ de baixo peso molecular são monoméricas em solução, enquanto outras formam dímeros. Formas triméricas também foram encontradas no veneno da *Naja naja naja* (Fremont *et al.*, 1993).

Alguns trabalhos têm focalizado a controvérsia de que algumas se não todas as fosfolipases A₂ funcionam como dímeros mesmo que elas existam como monômeros em solução. Relacionado com esta controvérsia está a afirmação de que um segundo sítio de ligação do fosfolípido pode servir como função alostérica durante a catálise (Cho *et al.*, 1988; Jain *et al.*, 1989; Lombardo, Dennis, 1985; Romero *et al.*, 1987).

A Fig. 1.11 mostra quais as interações que estabilizam o dímero na *Crotalus atrox* (Keith *et al.*, 1981) que foi cristalizado na ausência de Ca⁺². Este dímero da *Crotalus atrox*, Fig. 1.12 não é um dímero funcional, uma vez que se o Ca⁺² se ligar de maneira similar a da enzima pancreática bovina, ele causará uma reorientação dos grupos C=O da Gly³⁰ e Gly³², a qual por sua vez, leva ao deslocamento das cadeias do Trp³¹ e His³⁴, alterando, portanto, os contatos entre as subunidades.

Fig. 1.11: Interações que estabilizam o dímero na *Crotalus atrox*, onde 'L' se refere á subunidade esquerda e 'R' á subunidade direita..

Fig. 1.12: Estrutura dimérica da PLA₂ de *Crotalus atrox*.

A ligação do Ca⁺² poderá também, por sua vez causar um rompimento mais enfático da interação entre as subunidades. Interação esta formada por uma molécula de água que media a ponte salina entre Asp⁴⁹ que normalmente é um importante ligante do cálcio de uma subunidade e Lys⁶⁹ da outra. Há um excedente de carga positiva, desestabilizando a interação da K69 e desta forma romperia o contacto entre as subunidades.

Uma outra forma de dimerização funcional seria como apresentada na estrutura tridimensional da Miotoxina II do veneno da *Bothrops asper* (Arni; Ward, 1995), mostrada na Fig. 1.13. A miotoxina II apresenta um novo modelo de dimerização. As regiões asa-β e a primeira hélice seriam as regiões de contato das subunidades (resíduos K12, W77, K80 de uma subunidade com K80, Q11, E12 e K80 da segunda subunidade), e desta forma os sítios catalíticos ficariam expostos ao solvente, ao contrário do dímero apresentado na estrutura da *Crotalus atrox*.

Fig. 1.14: Dímero da Miotoxina II da *Bothrops asper*.

1.7 Fosfolipases PLA₂ Envolvidas no Processo de Inflamação

As fosfolipases A₂ (PLA₂) hidrolizam estereoespecificamente a ligação éster na posição sn-2 dos 3-sn-fosfoglicerídeos. A inibição desta reação em certas situações é de grande interesse farmacológico devido à liberação do ácido araquidônico da posição sn-2 dos fosfolípidos que desencadeiam reações em cascata para a produção de eicosanóides (ou prostaglandinas) que são mediadores do processo de inflamação.

Fig. 1.14 mostra o processo de liberação do ácido araquidônico dos fosfolípidos pela ação da fosfolipase A₂ (PLA₂) (Voet, D.;Voet, J.G., 1995).

Fig. 1.14: Liberação do ácido araquidônico pela ação das PLA₂'s.

As enzimas pertencentes a este grupo estão relacionadas com muitas doenças inflamatórias tais como artrite reumatóide, pancreatite e insuficiência renal aguda. Por exemplo, excepcionalmente altos níveis de hnp-PLA₂ (human non-pancreatic secretory) fosfolipases foram encontradas no fluido sinovial das juntas inflamadas em pacientes com artrite reumatóide. Este tipo de fosfolipase é do mesmo grupo da BthTX-I.

A estrutura cristalina desta enzima foi determinada num pH fisiológico na sua forma nativa e na presença de um complexo análogo diC₈(2Ph)PE mostrando que as características químicas da catálise são idênticas aquelas inferidas das outras estruturas de PLA₂'s extracelulares (Scott *et al.*, 1991) (Wery *et al.*, 1991).

Informações estruturais destas hnp-PLA₂'s têm sido usadas para a desenho de inibidores com o objectivo específico de inibir a ação destas fosfolipases A₂ e portanto reduzir a produção do ácido araquidônico (Schevitz *et al.*, 1995). Tais inibidores poderiam ter aplicações no tratamento de doenças inflamatórias.

Um trabalho desenvolvido recentemente na Unicamp (Pereira, 1995) com o objetivo de avaliar o carácter inflamatório (desenvolvimento de um edema) decorrente da picada de *Bothrops jararacussu* assim como a liberação de serotonina (5-HT) de mastócitos e a inibição do edema sugere que a primeira ação irritante de PLA₂ é a liberação de aminas dos mastócitos. O tratamento com cyproheptadina (antagonista de histamina e serotonina) praticamente aboliu a resposta edematogênica da BthTX-I, indicando que o edema decorre da degranulação de mastócitos. Alguns trabalhos (Cirino *et al.*, 1989; Moreno *et al.*, 1992) mostram que a liberação de serotonina e histamina de mastócitos, está diretamente relacionada com a atividade catalítica de PLA₂. Entretanto, o trabalho de Pereira (1995) sugere que a liberação de aminas acontece de forma independente de sua ação fosfolipídica, pois utilizaram no experimento fosfolipases A₂

sem atividade catalítica (BthTX-I) e a BthTX-II com baixa atividade o que portanto apoia a evidência da existência de sítios separados para a expressão das atividades enzimática e farmacológica para estas PLA₂. Porém deve-se lembrar que a questão da ausência completa e total de atividade fosfolipídica das K49's ainda está em aberto.

Estes resultados recentes, enfatizam a importância do melhor conhecimento da atividade das K49 PLA₂ de venenos de cobra e fosfolipases de um modo geral, o tema abordado nesta tese.

1.8 Referências Bibliográficas

- ARNI, R. K., WARD, R. J.; GUTIERREZ, J. M. & TULINSKY, A. *Acta Cryst.*, **D51**, 311-317 (1995).
- CHO, W., TOMASSELLI, A.G., HEINRIKSON, R. L. KEZDY, F.J. *ibid.* **263** 11237 (1988).
- CINTRA, A.C.O., MARANGONI, S., OLIVEIRA, B., GIGLIO, J.R., *J. Protein Chem.*, **12**, 57 - 64 (1993).
- CIRINO, G., PEERS, S. H., WALLACE, J.L. & FLOWER, R.J. *Eur. J. of Pharmacology* **166**, 505-510 (1989).
- DAVIDSON, F. F. & DENNIS, E. A. *J. Mol. Evol.*, **31**, 228-238 (1990).
- DENNIS, E.A.; *J. of Biol. Chem.*, **269**, 18, 13057 - 13060 (1994)
- DENNIS, E.A.; *The Enzymes, Vol. XVI - Academic Press Inc*, London, (1983)
- DIAZ, C., ALAPE, A., LARMONT, B., OLAMENDI, T., GUTIERREZ, J.M.; *Archives of Biochemistry and Biophysics*, **312**, 2 , 336 - 339, (1994)
- DIJKSTRA, B.W., DRENT, J., KALK, K.H.; *Nature*, **289**, 604 - 606, (1981)
- DIJKSTRA, B.W., KALK, K.H., HOL, W.G.J., DRENT, J., *J. Mol. Biol.*, **147**, 97-123, (1981).
- DUA, R. & CHO, W. - *Eur. J. Biochem.*, **221**, 481-490 (1994).

- ELDER, M., HITCHCOCK, P., MASON, R. & SHIPLEY, G. G. *Proc. R. Soc London Ser.*, **A354**, 157 (1977).
- FRANCIS, B., GUTIERREZ, J. M., LAMONTE, B. & KAISER, I. I. *Arch. Biochem. Biophys.* **284**, 352-359 (1991).
- FREDMONT, D. XUONG, N.H., WILSON, J.; *Proc. Nat. Acad. Sci.*, USA, **90**, 342- (1993).
- GOPALAKRISHNAKONE, P., DEMPSTER, D. W., HAWGOOD, B. J. & ELDER, H. Y. *Toxicon* **22**, 85-98 (1984).
- HAUSER, H., PASCHER, I., PEARSON, R. H. & SUNDELL, S. *Biochem. Biophys. Acta* **650**, 21 (1981).
- HAUSER, H. & PHILLIPS, M. C., *Prog. Surf. Membr. Chem.* **13**, 297 (1979).
- HOLLAND, D.R, CLANCY, L.L., MUCHMORE, S.W., RYDE, T.J., EINSPAHR, H.M., FINGEL, B.C., HEINRIKSON, R.L., WATTENPAUGH, K.D.; *J. of Biol. Chem.*, **265**, 29, 17649-17656, (1990)
- HOMSI- BRANDEBURGO, M. I., QUEIROZ, L. S., NETO, H. S., SIMIONI, L. R. & GIGLIO, J. R. *Toxicon* **26** , 615-629 (1988).
- JAIN, M. K. & BERG, O. *Biochem. Biophys. Acta* **1102**, 127 (1989).
- KATO, T., LEE, S., OISHI, O., AOYAGI, H., OHNO, M.; *Biochimica et Biophysica Acta*, **1211**, 215-220, (1994).
- KEITH, C., FELDMAN, D.S., DEGANELLO, S., GLICK, J., WARD, K. JONES, E.O., SIGLER, P.B.; *J. of Biol. Chem.*, **256**, 16, 8602-8607, (1981).
- KILBY, P.M., PRIMROSE, W.U., ROBERTS, G.C.K., *Biochem. J.*, **305**, 935-944, (1995).
- KINI, R.M., & IWANAGA, S. - *Toxicon*, **24**, 895- 905 (1986).
- KRIZAJ, I., DOLLY, J. O. & GUBENSEK, F. *Biochemistry* **33**, 13938 - 13945 (1994).
- LAMBEAU, G., BARHANIN, J., SCHWEITZ, H., QAR, J. & LAZDUNSKI, M. *J. Biol. Chem.* **264**, 11503-11510 (1989).

- LAMBEAU, G., ANCIAN, P., BARHANIN, J. & LAZDUNSKI, M. *J. Biol. Chem.* **269** 1575-1578 (1994).
- LAMBEAU, G., ANCIAN, P., BARHANIN, J., NICOLAS, J. P., BEIBOER, S. H. W., MOINIER, D., VERHEIJ, H. & LAZDUNSKI, M. *J. Biol. Chem.* **270**, 5534-5540 (1995).
- LI, Y., YU, B.Z., ZHU, H., JAIN, M.K., TSAI, M.D., *Biochemistry*, **33**, 14714 - 14722 (1994).
- LOMBARDO, D. & DENNIS, E. A. *J. Biol. Chem.* **260**, 16114 (1985).
- MARAGANORE, J. M. & HEINRIKSON, R. C. *J. Biol. Chem.* **261**, 4797-4804 (1986).
- MORENO, J. J., FERRER, X., ORTEGA, E. & CARGANICO, G. *Agents Actions* **36**, 258-263 (1992).
- OGAWA, T., KITAJIMA, M., NAKASHIMA, K., SAKAKI, Y. & OHNO, M., *J. Mol. Evol.*, **41**, 867-877 (1995).
- PEREIRA, M. F., *Tese de Doutorado em Ciências na Área de Genética*, Unicamp, Campinas - SP, (1995).
- PEARSON, R. H. & PASCHER, I. *Nature* **281**, 499 (1979).
- RICHARDSON, J.S. & RICHARDSON, D.C.; *Prediction of Protein Structure and The Principle of Protein Conformation-Plenium Press*, New York and London, (1989).
- ROMERO, G., THOMPSON, K., BILTONEN, R. L., *inib.* **262**, 13476 (1987).
- SCHEVITZ, R.W., BACH, N.J., CARLSON, D. G., CHIRGADZE, NY, CLAWSON, D.K., DILLARD, R.D., DRAHEIM, S.E., HARTLEY, L.W., JONES, N.D., MIHELICH, E.D., OLKOWSKI, J.L., SNYDER, D.W., SOMMERS, C., WERY, J.P.; *Nature Structural Biology* **2**, **6** , 458-465, (1995).
- SEELIG, J. *Q. Rev. Biophys.* **10**, 353 (1977).
- SELISTRE DE ARAUJO, H.S., WHITE, S.P., OWNBY, C.L.; *Archives of Biochemistry and Biophysics*, **325**, 1, (1996).
- SCOTT, D.L., OTWINOWSKI, Z., GELB, M.H., SIGLER, P.B. *Science*, **250**, 1563-1566 (1990).
- SCOTT, D.L., WHITE, S.P., OTWINOWSKI, Z., YUAN, W., GELB, M.H., SIGLER, P.B., *Science*, **250**, 1541-1546, (1990).

- SCOTT,D.L., WHITE, S.P., BLOWNING, J.L., ROSA, J.J., GELB, M.H., SINGLER, P.B., *Science*,**254**, 1007-1010, (1991).
- SCOTT, D.L., ACHARI, A., VIDAL, J.C., SIGLER, P.B., *J. of Biol. Chem.*, **267**, 31, 22645-22657, (1992).
- SHEN, Z. & CHO, W. *Int. J. Biochem. Cell. Biol.* **27**, 10, 1009-1013 (1995).
- VAN DEN BERG, B., TESSARI, M., BOELEN, R., DIJKMAN, R., KAPTEIN, R., DE HAAS, G.H., VERHEIJ, H.M.; *J. of Biomolecular NMR* **5**, 110-121, (1995).
- VAN DEN BERG, C.J., SLOTBOOM, A. J., VERHEIJ, H.M., DE HAAS, G.H., *J. of Cell. Biochem.*, **39**, 379-390, (1989).
- VAN DEN BERG, B., TESSARI, M., BOELEN, G.C.K., DIJKMAN, R. DE HAAS, G.H., KAPTEIN, R., VERHEIJ, H.M., *Structural Biology* **2**, **5**, 402 -406, (1995).
- VERHEIJ, H. M., VOLWERK, J. J., JASEN, E. H. J. M., PUYK, W. C., DIJKSTRA, B. W., DRENTH, J. & HAAS, G. H., *Biochemistry* **19** ,743- 750 (1980).
- VOET, D., VOET, J.G., *Biochemistry*, John Wiley& Sons, Inc. New York, (1995).
- WERY, J.P., SCHEVITZ, R.W., CLAWSON, D.K., BOBBITT, J.L., DOW, E.R., GAMBOA, G., GOODSON, T., HERMANN, R.B.,KRAMER, R.M., MC CLURE, D.B., MIHELICH, E.D., PUT NAM, J.E., SHARP, J.D., STARK, D.H., TEATER, C., WARRICK, M.W., JONES, N.D., *Nature*, **352**, 79-82, (1991).
- WHITE, S.P., SCOTT, D.L., OTWINOWSKI, Z., GELB, M.H., SIGLER, P.B., *Science* **250**, 1560-1563, (1990).
- TZENG, M-C., YEN, C-H., HSEU, M-J., DUPUREUR, C. M. & TSAI, M. D., *J.Biol. Chem.* **270**, **5**, 2120-2123 (1995).
- YEN, C-H. & TZENJ, M-C *Biochemistry* **30**, 11473-11477 (1991).

CAPÍTULO 2

CRISTALIZAÇÃO DA BthTX-I

2.1 Introdução

A determinação de uma estrutura de proteína utilizando a técnica de difração de raios X, requer como passo inicial a obtenção de monocristais da amostra a ser estudada. Como o número de variáveis envolvidas neste processo é muito grande (tais como concentração de proteína, temperatura, pH, agentes precipitantes e aditivos) o crescimento de cristais de proteína é ainda principalmente um processo de tentativa e erro. A tabela 2.1 ilustra alguns dos parâmetros que afetam a cristalização e/ou solubilidade de uma macromolécula (Ducruix; Giegé, 1992).

Um aspecto inicialmente importante e que deve ser analisado antes de qualquer teste de cristalização é o grau de pureza da proteína. Devido ao fato de que as macromoléculas são extraídas de uma complexa mistura biológica, a purificação desempenha um papel importante na cristalogênese. Uma observação importante é que não somente a proteína deve estar pura, mas também todos os reagentes devem ter um alto grau de pureza. As soluções usadas na cristalização devem ser filtradas.

A cristalização de moléculas em uma solução é um fenômeno de equilíbrio reversível onde os parâmetros termodinâmicos e cinéticos dependem das propriedades físico-químicas do solvente e dos solutos envolvidos.

. Parâmetros físico-químicos

- Supersaturação (concentração da macromolécula e agente precipitante)
- Temperatura, pH (flutuação destes parâmetros)
- Força iônica e pureza (natureza dos precipitantes, tampões e aditivos)
- Efeitos de densidade e viscosidade (diferenças entre o cristal e o líquido mãe)

. Parâmetros bioquímicos e biofísicos

- . Agentes ligantes e cofatores (ligação de ligantes tais como substratos, cofatores, íons metálicos e outros íons)
- . Aditivos específicos (tais como detergentes)
- . Idade das amostras (desnaturação ou degradação)

Tabela 2.1: Parâmetros que afetam a cristalização de uma proteína.

Sob certas condições de supersaturação, o sistema é guiado para um estado de equilíbrio no qual o soluto é particionado entre as fases líquida e sólida. Na cristalização, embora individualmente as moléculas percam sua liberdade rotacional e translacional, diminuindo a entropia do sistema, elas ao mesmo tempo ganham novas ligações químicas estáveis. Isto reduz a energia potencial do sistema e fornece uma força direcionadora para o estado ordenado. Deste modo, o objetivo final para as moléculas cristalizarem é o mesmo que guia todos os sistemas termodinâmicos, ou seja, a minimização da energia livre. Entretanto, nem todos os processos conduzem ao estado cristalino, podendo também agregar-se em um precipitado amorfo (McPherson, 1982).

Para sistemas relativamente simples envolvendo íons e pequenas moléculas, a energia livre em solução, energias de ligação e outros fatores são relativamente bem conhecidos e o crescimento de cristais pode ser tratado como qualquer processo físico, teoricamente

manipulado e predito. Entretanto, para sistemas de macromoléculas devido á sua complexidade o processo é menos facilmente descrito e pobremente entendido.

A estratégia usada para induzir cristalização de macromoléculas é conduzir o sistema lentamente para um estado de solubilidade mínima, e deste modo atingir um grau de supersaturação necessário à precipitação ou cristalização. Este estado de solubilidade mínima pode ser alcançado modificando-se as propriedades do solvente através do equilíbrio com agentes precipitantes (sais ou solventes orgânicos) ou alterando propriedades físicas como pH e temperatura.

A solubilidade de uma molécula de qualquer tipo depende da interação da molécula com o solvente e é essencialmente definida pelas características do solvente como pH, temperatura, sais e aditivos. No caso de soluções supersaturadas, onde não há moléculas de água suficientes para manter a hidratação, as moléculas podem se agregar como um precipitado amorfo ou, então, se cristalizar numa forma cristalina ordenada, sendo o que interessa para o cristalógrafo de proteínas. Portanto, a capacidade de um composto se cristalizar está relacionada com o estado de supersaturação da solução.

Para se determinar um ponto específico de solubilidade mínima, é necessário se obter o ponto de precipitação da molécula em valores de pH seqüenciais, com um dado precipitante, examinando o efeito dos diferentes agentes de precipitação sobre a proteína em estudo. Na tabela 2.2 encontram-se uma lista dos agentes precipitantes mais utilizados na cristalização de macromoléculas (McPherson, 1982).

	Solventes orgânicos	Polímeros (PEG)
Sulfato de amônio ou sódio	Etanol	PEG 200
Sulfato de magnésio	Metanol	PEG 400
Cloreto de sódio/potássio	Dioxano	PEG 600
Cloreto de amônio	Acetona	PEG 1000
Acetato de sódio ou amônio	DMSO	PEG 4000
Formiato de sódio	MPB	PEG 6000
		PEG 8000
		PEG 10000
		PEG 20000

Tabela 2.2: Principais agentes precipitantes mais utilizados na cristalização de macromoléculas.

A adição de solventes orgânicos produz mudanças significativas na solubilidade das proteínas em soluções aquosas. Isto se deve, principalmente, á resultante das constantes dielétricas da mistura água e solvente orgânico. A constante dielétrica da água é igual a 80, mas numa mistura com 34% de etanol em água, a constante dielétrica cai para 60. Esta diminuição da constante dielétrica aumenta a atração Colulombiana entre cargas contrárias da molécula de proteína , o que diminui a solubilidade.

A ação de detergentes é, ás vezes, exigida para solubilizar certas proteínas, como por exemplo, as proteínas de membrana, onde neste caso, torna-se necessário reduzir as interações hidrofóbicas que frequentemente levam á formação de precipitantes amorfos. Alguns dos detergentes mais comuns são BOG (β -octil glucosídeo) e SDS (sódio dodecil sulfato).

O agente precipitante deve ser escolhido com base no comportamento da proteína, levando em consideração fatores como as faixas de pH e temperatura onde a proteína é estável e ativa. Diferentes íons podem afetar a solubilidade de uma proteína de diferentes formas. A eficiência de um sal é proporcional ao número e ao tamanho das cargas carregadas pelos íons que podem ser solvatadas. Os efeitos da concentração e da carga dos íons encontra sua expressão quantitativa na denominada força iônica (Ducruix; Giegé, 1992),

Eq. 2.1.

$$\mu = \frac{1}{2} \sum_i c_i Z_i^2 \quad \text{Eq. 2.1}$$

onde c_i é a concentração de cada íon na solução e Z_i é a sua valência. Por esta razão, íons divalentes (tais como sulfato de amônia) e trivalentes são mais comuns nos experimentos de cristalização, em baixas concentrações. Á medida que a concentração de íons aumenta e consequentemente a força iônica, começa a ocorrer uma competição entre os íons e o soluto (proteína) pelas moléculas de água. A diminuição da solvatação da proteína em consequência da remoção de água, leva a uma diminuição da solubilidade.

O polietileno glicol (PEG) induz a proteína a se cristalizar. Pode-se observar que um grande número de proteínas têm sido resolvidas, usando cristais que cresceram em PEG. A cristalização de proteínas com PEG têm tido mais sucesso quando a força iônica é baixa (Ducruix; Giegé, 1992). Torna-se bastante difícil quando a força iônica é alta. PEG é um precipitante não iônico, não se ioniza na água.

Pode-se modificar a carga líquida de uma proteína adicionando ou removendo prótons, isto é mudando o pH da solução. Geralmente a proteína é mais solúvel quanto mais carga líquida contém. O ponto de menor solubilidade é quando a carga líquida é zero. Isto porque os íons podem empacotar numa forma sólida e interagir eletrostaticamente sem acúmulo de carga líquida. O ponto de carga nula é o ponto isoelétrico (pI). Um controle adequado do pH durante a cristalização necessita da presença de tampões desejáveis. A estabilidade e a atividade da proteína dependem do pH, tornando-o então um fator importante, uma vez que frequentemente se deseja cristalizar a proteína na sua forma ativa.

Muitos dos fatores que governam a solubilidade tais como força iônica e pH têm uma acentuada dependência da temperatura. Generalizações a respeito do efeito da temperatura sobre a solubilidade de proteínas são difíceis, pois depende da proteína e das condições experimentais. Tanto o aumento da temperatura quanto o seu decréscimo podem induzir a supersaturação. Normalmente o processo de cristalização ocorre em câmaras com temperatura controlada para reduzir flutuações que modificam a solubilidade durante a cristalização.

2.2 Métodos de Cristalização de Proteínas

Descrevemos os parâmetros que podem ser modificados para trazer a proteína para um estado de supersaturação e/ou solubilidade mínima. Devemos agora considerar as técnicas experimentais, as quais têm sido utilizadas pelos cristalógrafos de proteínas, que permitem a formação de cristais. Existem muitos métodos para cristalizar macromoléculas biológicas, entretanto, todos seguem o mesmo princípio, ajudando a trazer a solução para um estado de supersaturação onde a escolha de qual método utilizar depende entre outros fatores da quantidade de proteína disponível.

Existem quatro tipos de métodos utilizados para reduzir mais ou menos gradualmente a solubilidade de uma macromolécula, são eles (Drenth, 1994):

- *Difusão de Vapor*
- *Diálise*
- *Interface de Difusão*
- *Banhos de Cristalização*

Os três primeiros métodos podem ser conduzidos com quantidades muito pequenas de proteína. Descreveremos apenas o primeiro método que foi o método utilizado para cristalizar a BthTX-I.

2.2.1 Difusão de Vapor

O princípio deste método é o processo de equilíbrio entre duas soluções através da fase de vapor num meio fechado. O solvente é gradativamente transferido através do vapor da solução de proteína para a solução mais concentrada até entrar em equilíbrio.

Na prática existem três variações deste método: *“hanging-drop”* (gota pendurada), *“sitting-drop”* (gota sentada) e *“sandwich-drop”* (gota sanduíche). O método da gota pendurada é o mais utilizado hoje em dia, enquanto a gota sentada se aplica melhor a condições de cristalização contendo detergente.

2.2.1.1 “Hanging-drop”

Esta técnica permite usar pequenos volumes de solução de proteína contra um volume grande de solução precipitante, conforme observado na Fig. 2.1

Figura 2.1 Método de Difusão de Vapor “hanging-drop”.

Uma gota contendo a macromolécula biológica a ser cristalizada, com a solução tampão, agentes precipitantes e aditivos, é equilibrada com um reservatório contendo uma solução do agente precipitante a uma concentração mais alta do que a gota. Se o volume da solução do reservatório é suficientemente grande comparado com o da gota podemos assumir que quando o equilíbrio é atingido, a concentração do precipitante na solução de proteína será igual à concentração inicial na solução do reservatório. Esta condição é escolhida perto do ponto de precipitação assim levando a solução na gota à condição necessária para a formação de cristais. Este foi o método utilizado na cristalização da BthTX-I

2.3 Ensaios de Cristalização

O experimento de cristalização tipicamente consiste em preparar ensaios de cristalização onde se variam condições, tais como a concentração do agente precipitante, pH, temperatura, etc., produzindo assim um conjunto de condições com variações em torno do ponto de

supersaturação com um volume reduzido de amostra. Normalmente, um valor aproximado do ponto de precipitação para determinados precipitantes em certos pH, pode ser encontrado anteriormente através de um ensaio feito com aproximadamente 5µl de amostra.

A busca de parâmetros ideais de cristalização pode ser feita sistematicamente ou usando o método de fatorial incompleto (Carter & Carter, 1979; Ducruix; & Giegé, 1992). No primeiro caso uma matriz de condições é estatisticamente variada, por exemplo concentração numa direção e pH na outra. No caso de fatorial, um conjunto de parâmetros multidimensionais são inicialmente selecionados e testados. Os resultados são analisados para construir novos experimentos de ensaios mais refinados para as condições de cristalização. O método permite análise matemática buscando descobrir quais parâmetros estão influenciando mais no sucesso dos ensaios.

Podem ocorrer vários resultados: (a) Cristais muito pequenos e em grande quantidade; (b) Cristais muito pequenos e em número reduzido; (c) Cristais grandes e de boa qualidade, mas em número reduzido; (d) precipitado (e) solução límpida, etc..

Em muitos casos, a melhora da qualidade e quantidade do cristal é difícil, e várias estratégias diferentes têm que ser adotadas.

Quando os cristais são muito pequenos embora de boa qualidade, a técnica de “*seeding*” (*semeação*) pode ajudar. Dois tipos de seeding são usualmente utilizados: o “*microseeding*” e o “*macroseeding*”. No “microseeding” os cristais de boa morfologia são esmagados na solução mãe e um pequeno volume da mesma (ou uma série de soluções diluídas da mesma) são incluídos no preparo de uma nova gota de cristalização. Esta técnica é muito utilizada quando a nucleação espontânea é difícil. Já o macroseeding é uma técnica muito mais eficiente na produção de cristais maiores, onde se usa como semente um único cristal de boa qualidade que é escolhido de uma gota e lavado numa solução contendo proteína e solução do reservatório, para finalmente ser transferido para uma outra gota (Ducruix; Giegé, 1992). Às vezes, *semeação cruzada* (*cross-seeding*) pode ser usada, quando cristais de uma proteína relacionada estão disponíveis.

2.4 Ensaio de Cristalização da BthTX-I

2.4.1 Quantidade de Proteína Disponível e Verificação da Pureza da BTHtx-I

As amostras de BthTX-I foram fornecidas pelo Departamento de Bioquímica da Faculdade de Medicina de Ribeirão Preto com a colaboração Prof. Dr. J. R. Giglio e da Prof^a. Dr^a. A.C.O. Cintra. Foram separadas inicialmente alíquotas para eletroforese em SDS-gel de 15 % de Poliacrilamida com o objetivo de verificar a pureza e o peso molecular da BthTX-I. Para tal, as amostras foram fervidas durante 5 minutos na presença de SDS e β -MSH antes de ser aplicadas no gel de empilhamento. O resultado desta eletroforese pode ser visto na Fig. 2.2 e por este método de análise pode-se verificar que a proteína está pura tendo um peso molecular de aproximadamente 14.000 (Cintra *et al.*, 1993).

Fig. 2.2: Eletroforese em SDS-gel de 15% de Poliacrilamida da BTHtx-I. Os padrões de referência foram: soro albumina (P.M. 66.000); ovo albumina (P. M. 45.000); Anidrase carb. (P.M. 30.000); inibidor de tripsina de soja (P .M.-20.000); Citocr. C (P.M. 12.400) ; Riboflavina (P.M. 7000).

As raias correspondem respectivamente: 1- Padrão ; 2- BthTX-I c/ β -MSH
3- BthTX-I c/ β -MSH.

A quantidade de proteína disponível (cerca de 30 mg) era suficiente para se realizarem todos os testes necessários para a sua cristalização. A BTHtx-I é uma proteína composta de uma única cadeia polipeptídica com 121 amino ácidos, pM $\approx$ 14000 e pI = 8.2, e os próximos passos estão relacionados com a busca das condições ideais de cristalização.

2.4.2 Testes Iniciais de Precipitação

As soluções tampões inicialmente testadas foram: (1) 0.1M TRIS pH 8.0; (2) 0.1M de HEPES pH 7.0; (3) 0.1M Mes pH 6.0; (4) 0.1M de acetato de sódio pH 5.0 e os agentes precipitantes foram: (a) $(\text{NH}_4)_2\text{SO}_4$ saturado nos pH descritos anteriormente e (b) duas soluções a 4M de NaH_2PO_4 e K_2HPO_4 . Todos os experimentos utilizaram a técnica de difusão de vapor “hanging - drops”. Foram realizados alguns testes usando “sitting-drop” onde foi usado um detergente BOG (β -octil glucosídeo) que não conduziu a resultados positivos. Todos os ensaios de cristalização foram realizados numa sala de cristalização a temperatura controlada a 18 °C.

Foram preparadas inicialmente caixas de cristalização, variando o pH e a concentração de $(\text{NH}_4)_2\text{SO}_4$ e fosfato. No caso do fosfato, o pH pode ser controlado variando as concentrações relativas de NaH_2PO_4 e K_2HPO_4 . As gotas foram formadas com 2.5 μl de solução de proteína, numa concentração de 10 mg/ml em água milli-Q mais 2.5 μl da solução do poço. Os resultados da primeira caixa de cristalização são mostrados na Fig. 2.3, onde a princípio se usou as soluções tampões descritas anteriormente (1), (2), (3) e (4). Como agente precipitante $(\text{NH}_4)_2\text{SO}_4$ saturado nos pH apresentados anteriormente, variando as concentrações de 50% a 65%.

	1	2	3	4	5	6	
	50%	53%	56%	59%	62%	65%	
TRIS 0.1M pH 8.0		*	*	*			A
HEPES 0.1M pH 7.0		*	*	**	**	**	B
MES 0.1M pH 6.0							C
Acet. Na 0.1M pH 5.0							D

Fig. 2.3: Condições de cristalização da BThTX-I, onde (*) significa muitos cristais e pequenos; () poucos cristais e de bom tamanho para difração de raios X.**

Nestas condições, observaram-se cristais nas gotas 2A, 2B, 3A, 3B, 4A, 4B, 5B e 6B. Os cristais das gotas 2A, 2B, 3A, 3B e 4A eram muito pequenos e em grandes quantidades. Nas gotas , 4B, 5B e 6B apareceram poucos cristais e de bom tamanho para a coleta de dados. As demais gotas apresentaram-se límpidas ou com precipitado. Estes cristais apareceram depois de 4 dias de montada a caixa de cristalização. Cristais não foram obtidos utilizando fosfato como tampão/precipitante.

Analisando a caixa de cristalização apresentada, observa-se que as melhores condições de cristalização foram obtidas com a solução de tampão HEPES 0.1 M pH 7.0. Estes cristais já eram suficientemente bons para uma coleta de dados, mas eram poucos cristais. As próximas tentativas de cristalização foram realizadas com variantes destas condições iniciais. Novas caixas de cristalização foram montadas, variando o pH de 6.7 a 7.6 e as concentrações de $(\text{NH}_4)_2\text{SO}_4$ de 57% a 62% . Nestas condições obtivemos cristais maiores e de boa qualidade para todas as coletas de dados da BthTX-I.

Na Fig. 2.4 podemos observar alguns dos cristais obtidos da BthTX-I nas condições apresentadas.

Fig 2.4 Fotos dos Cristais da BthTX-I nativa. Os maiores cristais possuem dimensões 0.5mm x 0.3mm x 0.3mm.

Foram realizadas tentativas de se cristalizar a BthTX-I com o inibidor β -bromophenacyl bromide (Roberts *et al.*, 1977). Estas tentativas foram realizadas nas mesmas condições de cristalização da proteína nativa utilizando duas técnicas diferentes: 'soaking' e co-cristalização. Para o soaking foram utilizados cristais da proteína nativa nas condições apresentadas anteriormente, onde a nova solução foi preparada numa concentração de 0.005 M e/ou 0.003 M de inibidor.

Para a co-cristalização foram repetidas as condições de cristalização da nativa onde o pH variou de 7.0 a 7.4, a concentração de $(\text{NH}_4)_2\text{SO}_4$ de 59% a 62%. Os cristais complexo/proteína apresentados na Fig. 2.5 demoraram meses para serem observados.

Fig. 2.5: Cristais do complexo BthTX-I mais inibidor β – bromophenacyl bromide (co-cristalização) com 0.1mm na dimensão maior.

Após o soaking verificou-se que não foi ligado o inibidor à toxina utilizando este processo. Os cristais co-cristalizados eram muito pequenos ($\approx 0.1\text{mm}$) e não foi possível se obter um conjunto completo de dados, obtendo-se apenas os parâmetros de rede e o sistema cristalino que serão apresentados mais adiante.

2.6 Referências Bibliográficas

- BLUNDELL, T.L., JOHNSON, L.N., *Protein Crystallography* - Academic Press, New York, (1976).
- CINTRA, A.C.O., MARANGONI, S., OLIVEIRA, B., GIGLIO, J.R. *J. Protein Chem.*, **12**, 57-64, (1993).
- CARTER, JR., C.W. & CARTER, C. W. - *J. Mol. Biol.* **254**, 12219-12223 (1979).
- DRENTH, J., *Principles of Protein X-Ray Crystallography* - Springer-Verlag, New York, (1994).
- DUCRUIX, A., GIERGÉ, R., *Crystallization of Nucleic Acids and Proteins (A Practical Approach)* - Oxford University Press, New York, (1992).
- MC PHERSON, A., *Preparation and Analysis of Protein Crystals*, John Wiley and Sons, Inc., New York, (1982).
- ROBERTS, M. F., DEEMS, R. A., MINCEY, T. C. & DENNIS, E. A. *J. Biol. Chem.*, **252**, 7, 2405-2411 (1977).

CAPÍTULO 3

COLETA E PROCESSAMENTO DOS DADOS DE DIFRAÇÃO DA BthTX-I

3.1 Introdução

O problema experimental que envolve a determinação de uma estrutura molecular por raios-X é obter as amplitudes $F_p(hkl)$ e as fases $\alpha_p(hkl)$ das ondas difratadas pelo cristal. As amplitudes podem ser calculadas a partir de um conjunto tridimensional de intensidades medidas em experimentos específicos de difração enquanto que as informações das fases podem ser determinadas utilizando vários métodos (Blundell; Johnson, 1976), entre eles, o método de Substituição Molecular (SM) o qual foi utilizado para a resolução da estrutura da BthTX-I. Neste capítulo descreveremos inicialmente a geometria e a operação dos instrumentos básicos utilizados para medir tais intensidades. A seguir discutiremos os fatores que governam a coleta de dados de intensidade de raios-X dos cristais de proteína para finalmente discutirmos os métodos utilizados para o processamento das imagens de difração de raios X da BthTX-I.

3.2 Câmaras e Detetores

No caso do uso de radiação monocromática, para que o número maior de pontos da rede recíproca seja trazido à condição de difração, o que é normalmente feito, é rodar a rede, isto é, pela rotação ou oscilação do cristal. Quando um dado ponto da rede recíproca estiver na esfera de Ewald, Fig.3.1, a família de planos correspondente está na condição de difração, determinada pela lei de Bragg, Eq.3.1:

$$\lambda = 2 d_{hkl} \sin \theta_{hkl} \quad \text{Eq.3.1}$$

onde d_{hkl} é a distância interplanar da família de planos cujos índices de Müller são h, k, l , λ é o comprimento de onda e θ_{hkl} é o ângulo de Bragg.

Fig. 3.1: A esfera de Ewald interceptada por um ponto da rede recíproca ,B. s_0 é a direção do feixe incidente, S é o vetor do espaço recíproco, C é o cristal, O é a origem do espaço recíproco e $1/\lambda$ é o raio da Esfera de Ewald.

Atualmente o método mais utilizado para coleta de dados no caso de proteínas é o método de oscilação. O princípio da câmara de oscilação é muito simples. O único movimento é a rotação ϕ do cristal em torno de um eixo. O eixo de rotação do cristal é perpendicular ao feixe de incidência dos raios-X. Quando o cristal é rotacionado em torno de um ângulo ϕ , certos pontos da rede recíproca cruzam a esfera de Ewald e para estes pontos a condição de difração é satisfeita. Na Fig. 3.2 podemos ver a geometria básica de uma câmara de rotação (Blundell; Johnson, 1976).

Fig. 3.2: Geometria da câmara de rotação. A figura mostra parte da rede recíproca e indica a maneira pela qual os pontos da rede recíproca podem ser trazidos à condição de difração.

3.2.1 A Região Cega

Quando o cristal é rotacionado em torno de um eixo de rotação perpendicular ao feixe incidente, esta geometria dá origem ao aparecimento de uma região dentro da qual os pontos da rede recíproca não cruzam a esfera de Ewald e, portanto, tais reflexões não são registradas em um filme. Esta região é conhecida como região cega cujo raio mínimo é dado por $\xi_{\min}$, Fig. 3.3 (Arndt; Wonacott, 1977) e cujo valor é dado pela Eq. 3.2.

$$\xi_{\min} = 1 - (1 - \zeta^2)^{1/2} \quad \text{Eq. 3.2}$$

onde ζ é definido na Fig. 3.3.

Fig. 3.3: Raio da região cega $\epsilon_{\min}$

Se o cristal estiver alinhado por exemplo, com um eixo cristalográfico de ordem dois ao longo do eixo de rotação, naturalmente os volumes do espaço recíproco nas regiões cegas superior e inferior serão relacionadas por simetria e portando de forma a obter um conjunto completo tri-dimensional de dados, torna-se necessário rodar o cristal em torno de uma nova direção axial. Este procedimento nem sempre é realizado, uma vez que o número de reflexões da zona cega não é muito grande, e tipicamente os cristais não estão perfeitamente alinhados.

3.2.2 Escolha da Orientação do Cristal para Fotos de Oscilação ou Rotação

Vários são os fatores que devem ser considerados na escolha da orientação do cristal a ser utilizado para a coleta de um conjunto de dados tridimensionais, a maioria dos quais dependem das características individuais do cristal. Em geral, a coleta de um conjunto de dados requer mais de um cristal. Afim de evitar que as mesmas reflexões sejam coletadas em cristais diferentes e desta forma ter que utilizar um número maior de cristais do que o necessário, é vantajoso ajustar o cristal numa dada orientação e a partir daí definir o ângulo de rotação para cada fotografia e para a rotação total.

A orientação do cristal é tipicamente determinada através da medida de duas fotos “still” com orientações relativas de 90° mas também pode-se determinar esta orientação usando fotos de oscilação.

No caso geral, o ângulo de rotação total necessário de forma a coletar todo o espaço recíproco é dado pela Eq. 3.3 (Blundell; Johnson, 1976).

$$\phi = 180^\circ + 2\theta_{\text{máx}} \quad \text{Eq. 3.3}$$

Na Fig. 3.4 podemos entender melhor o significado desta equação onde a origem do espaço recíproco está localizada no ponto O, o cristal fixo e a esfera de reflexão rodando.

Fig. 3.4 Projeção ao longo do eixo de rotação mostrando a área do espaço recíproco total coletada. O pequeno círculo corresponde ao limite de resolução.

Entretanto, se o cristal tem alguma simetria, então o ângulo total de rotação é menor do que 180° , dependendo da orientação do mesmo.

3.2.3 Ângulo de Oscilação de Cada Filme para Evitar Sobreposição

O ângulo ótimo de oscilação para cada filme é escolhido de modo a permitir que sejam coletados o número máximo de reflexões com o mínimo de sobreposições. O máximo valor do intervalo de rotação $\Delta\phi$, o qual evita a sobreposição das reflexões no plano equatorial, depende da resolução máxima desejada, da separação dos pontos da rede recíproca e portanto da cela unitária e também da orientação da rede (Blundell; Johnson, 1976). Normalmente esta escolha é feita através de simulações das imagens, uma vez determinada a orientação do cristal. Neste

estágio são levados em conta a completeza prevista para uma dada faixa de oscilação e os dados são geralmente coletados com oscilações de $1-3^\circ$. Naturalmente o menor que seja este ângulo de oscilação, menor a probabilidade de sobreposição e menor a intensidade de fundo, mas em compensação o número de reflexões parcialmente medidas será maior e o número de imagens a serem processadas depois também será maior. A escolha do ângulo a ser usado deve, em princípio, levar em conta todos estes fatores.

As imperfeições na rede cristalina (mosaicidade), a falta de monocromaticidade do feixe e problemas na colimação do feixe de raios X, fazem com que os pontos da rede recíproca adquiram dimensões, no sentido de que eles levam um certo tempo para cruzar a esfera de Ewald. Isto faz com que aquelas reflexões presentes nos extremos das faixas de rotação sejam registradas em mais de um filme (depende também do $\Delta\varphi$ escolhido em cada filme) - as reflexões parciais, isto é, reflexões que não cruzaram totalmente a esfera de reflexão durante uma única imagem - afetando assim os valores das intensidades medidas. Entretanto, seus valores individuais podem ser somados ao longo dos filmes para se obter uma intensidade integrada, uma vez que os filmes podem ser escalonados entre si.

3.3 Fatores que Afetam as Medidas de Intensidade

3.3.1 A Mosaicidade

A mosaicidade tem um efeito decisivo no fenômeno de difração. O monocristal real com estrutura de mosaico pode ser considerado como um conjunto de blocos de celas unitárias levemente desorientados um em relação ao outro, conforme ilustrado na Fig. 3.5 (McRee, 1993). O efeito de um aumento na mosaicidade é um alargamento e achatamento da intensidade da reflexão, que influencia o tempo que a reflexão leva para passar pela esfera de Ewald. A mosaicidade para cada cristal é refinada durante o processamento de dados.

Fig. 3.5: Efeito da mosaicidade na intensidade da reflexão.

3.3.2 Colimadores e Monocromadores do Feixe de Raios-X

Os sistemas de colimação podem ser escolhidos de tal forma a otimizar a intensidade máxima no filme e a resolução máxima de separação entre as reflexões. Colimadores circulares são geralmente usados para limitar o feixe incidente, resultando em um feixe circular de raios X com um diâmetro que deve ser menor do que as dimensões do cristal. Como a intensidade difratada depende do volume do cristal e como o feixe incidente é menos homogêneo nas bordas do que no centro, o tamanho do cristal deveria, de fato, ser significativamente menor do que o diâmetro do feixe. Como os colimadores de raios-X têm geralmente diâmetros entre 0.5 mm e 1.0 mm, os cristais usualmente usados na análise de estruturas são geralmente menores que 1.0 mm. Uma estimativa útil para a espessura ideal de um cristal é obtido da aplicação da Eq. 3.4 onde μ é o coeficiente de absorção linear que depende do material usado e do comprimento de onda do feixe incidente, t é a espessura da amostra, k é uma constante de proporcionalidade, I_i e I_o as intensidades difratadas e incidente, respectivamente (Buerger, 1942).

$$I_1 = kI_0 \exp(-\mu t)t^2 \quad \text{Eq. 3.4}$$

$$t = \frac{2}{\mu} \quad (\text{espessura ideal})$$

No R-AXIS IIC (o detetor usado neste estudo e que será discutido abaixo), os colimadores disponíveis são: (a) colimador simples de 0.5 mm e (b) colimador duplo de 0.3 mm de diâmetro.

Para se obter radiação monocromática (K_α) a partir do espectro produzido por um tubo de raios X filtros da radiação K_β ou monocromadores de cristal podem ser usados. Um filtro de níquel de espessura 0.013 mm pode ser usado para reduzir a radiação CuK_β para 2% enquanto a CuK_α é menos atenuada, para 66% da intensidade original. A decisão de se usar um filtro ou um monocromador de cristal não é ainda óbvia. Quase todos os difratômetros atualmente, usam um monocromador que é considerado a melhor forma de se obter radiação monocromática. Entretanto, dois problemas surgem com o uso de monocromadores. O primeiro é uma maior atenuação da intensidade do que com o uso de filtros e a segunda é o perfil do feixe monocromático. Se o cristal monocromador não for de excelente qualidade e se não estiver ajustado com precisão, ele pode dar origem a um perfil do feixe não uniforme, o que pode causar uma fonte de sérios erros nas medidas das intensidades.

No caso dos monocromadores o feixe de raios X se reflete num monocristal colocado em um ângulo de Bragg apropriado para se obter a radiação característica desejada. Para esta finalidade grafite é altamente usado e orientado de tal forma que os seus planos (002) reflitam a radiação no comprimento de onda desejado, isto é, da lei de Bragg e de $d_{002} = 3.35 \text{ \AA}$ do cristal de grafite, temos que o cristal de grafite tem que ser orientado de tal forma que o ângulo de Bragg θ seja igual a $13,3^\circ$ ou $2\theta = 26.6^\circ$ para a radiação característica K_α de um tubo de cobre ($\lambda = 1.5418 \text{ \AA}$).

3.3.3 O Efeito do Tamanho da Cella Unitária na Medida da Intensidade Difrata

Uma das mais importantes relações da cristalografia de proteínas é a relação inversa entre a intensidade difratada e o volume da cela unitária. A maioria das macromoléculas biológicas cristalizam em células unitárias grandes e as intensidades difratadas são correspondentemente fracas. Os cristais de proteína são relativamente fracos espalhadores, por duas razões principais. A primeira, porque eles são formados por átomos leves, tais como C, N e O e a segunda razão está diretamente relacionada com o grande tamanho da sua cela unitária.

Consideremos dois cristais de volume V_{cr} e composição atômica iguais, mas celas unitárias diferentes. A energia total difratada até uma dada resolução é igual em ambos os casos (pois depende do número de espalhadores), mas no caso da maior cela unitária, está distribuída por um maior número de reflexões, que portanto terão intensidades médias inferiores. A intensidade difratada é diretamente proporcional a V_{cr} e inversamente proporcional ao volume da cela unitária, Eq. 3.5.

$$I \approx \frac{V_{cr}}{V} \quad \text{Eq. 3.5}$$

3.3.4 Limitações Experimentais da Resolução e o Número de Reflexões Medidas

Uma das consequências importantes da lei de Bragg é que a resolução máxima, determinada por considerações geométricas é $d_{min} = \lambda / 2$. Na cristalografia de proteínas os termos 6.0 Å (baixa resolução) e 2.0 Å (alta resolução) significa que os dados foram coletados até um limite de 6.0 Å e 2.0 Å de espaçamento interplanar, respectivamente (Blundell; Johnson, 1976). Mais alta resolução, mais detalhes serão vistos no mapa de densidade eletrônica.

Na prática, a resolução máxima alcançada por um cristal de proteína é limitada pela desordem inerente das moléculas devido a vibrações térmicas e o seu empacotamento no cristal (desordem estática), mas existem também restrições práticas na resolução máxima que dependem da geometria da coleta. Para um filme ou detetor planar a situação pode ser analisada a partir da Fig. 3.6, onde D é a distância do cristal ao detetor.

Fig. 3.6: Resolução Máxima de um Filme de Raios X em função da distância do cristal ao detetor.

Da Fig. 3.6 obtemos a Eq. 3.6

$$d_{\min} = \frac{\lambda D \cos \theta}{Y \cos 2\theta} \quad \text{Eq. 3.6}$$

Como podemos observar da Eq. 3.6, quanto maior a distância D do cristal ao detetor (filme) menor será a resolução máxima.

Devido ao fato de que há um ponto da rede recíproca por cela unitária recíproca (cada ponto da rede é compartilhado por 8 celas unitárias), o número de reflexões dentro da esfera limite de resolução é aproximadamente o número de celas unitárias recíprocas dentro desta esfera. Então, o número N de reflexões possíveis é igual ao volume da esfera limite dividido pelo volume V_{recip} de uma cela recíproca, Eq. 3.7 (Blundell, Johnson, 1976).

$$N = \frac{\left(\frac{4\pi}{3}\right) \cdot \left(\frac{2}{\lambda}\right)}{V_{\text{recip.}}} \quad \text{Eq. 3.7}$$

como $V_{\text{recip}} = 1/V$ da cela unitária real então obtemos a Eq. 3.8 para o número de reflexões em função do volume V da cela unitária e do comprimento de onda λ .

$$N = \frac{33.5V}{\lambda^3} \quad \text{Eq. 3.8}$$

A Eq. 3.8 mostra que o número de reflexões possíveis depende apenas do volume da cela unitária V e λ . Em princípio todas estas reflexões teriam que ser medidas até uma dada resolução entretanto, devido à simetria da cela unitária, o número de reflexões independentes pode ser reduzido consideravelmente.

Associando a Eq. 3.6 e 3.8 encontramos a relação entre o número de reflexões e a resolução, Eq. 3.9.

$$N = \frac{33.5V}{d_{\text{min}}^3} \cdot \frac{D^3 \cos^3 \theta}{Y \cos^3 2\theta} \quad \text{Eq. 3.9}$$

3.4 Redução dos Dados

3.4.1 Fator de Lorentz (L)

A intensidade de uma reflexão produzida por um cristal em movimento depende do tempo que um dado ponto da rede recíproca leva para atravessar a superfície da esfera de reflexão (esfera de Ewald). Este tempo não é constante, mas varia com a posição do ponto no espaço recíproco e com a direção com que ele se aproxima da esfera. O fator que descreve esta variação é denominado fator de Lorentz e depende da geometria do equipamento utilizado.

$$I = I_0 \exp(-\mu t) \quad \text{Eq. 3.11}$$

Para um material consistindo de um número N de elementos, o coeficiente de absorção mássico é dado pela Eq. 3.12

$$\mu_m = \sum_{i=1}^N g_i (\mu_m)_i \quad \text{Eq. 3.12}$$

onde g_i é a fração em massa do i th elemento e $(\mu_m)_i$ o coeficiente de absorção mássico do i th elemento. Se ρ é a densidade,

$$\mu = \rho \mu_m$$

O coeficiente de absorção μ_m , varia com o comprimento de onda λ , obedecendo à relação apresentada na Eq. 3.13 fora das arestas de absorção, onde a e b são constantes de proporcionalidade.

$$\mu_m = a\lambda^3 + b\lambda^4 \quad \text{Eq. 3.13}$$

Esta é a famosa relação de Victoreen e o fator λ^3 é fundamentalmente predominante, Fig. 3.7 (Helliwell, 1992).

Fig. 3.7: Variação do coeficiente de absorção μ (mm⁻¹) com o comprimento de onda λ .

A absorção de raios X causam um aquecimento térmico e um consequente aumento dos danos por radiação (Gonzalez et al., 1994) (Gonzalez; Nave, 1994). Benefícios importantes que têm sido relatados sobre o uso de λ menor incluem a redução da absorção, a diminuição dos danos causados por radiação e uma melhora da razão sinal/ruído quando se coletam dados de cristais de proteína, usando comprimentos de onda mais curtos (Arndt, 1984).

A maioria dos processos que ocorrem quando raios X interagem com a matéria (calor, absorção fotoelétrica e espalhamento Compton) causam diretamente ou indiretamente danos por radiação no cristal. Quase todos os cristais de proteína sofrem danos causados pela radiação durante a exposição ao feixe de raios X, embora a sensibilidade dos mesmos varie consideravelmente de proteína para proteína. A alta energia dos fótons de raios X ionizam a proteína e o solvente ocorrendo a formação de radicais livres que são altamente reativos e causam danos no cristal rompendo ligações covalentes. O efeito destes danos se manifestam através do decréscimo das intensidades durante os experimentos de difração. Uma das tentativas de reduzir os efeitos de danos por radiação é de resfriar a amostra usando-se um sistema de baixa temperatura sobre o goniômetro, embora o uso desta técnica possa conduzir a outro problema, que é o aumento da mosaicidade quando o cristal é resfriado, o qual pode conduzir a um sério problema de maior superposição das reflexões (Gonzalez *et al.*, 1994).

3.5 Coleta de Dados

Na prática, uma vez obtido o padrão de difração de raios X de um cristal, ou seja, as intensidades $I(hkl)$ de cada reflexão, deseja-se obter a estrutura da molécula em estudo. A expressão $I(int, hkl)$ se um cristal roda com velocidade angular ω é dada pela equação de Darwin (Darwin, 1914), Eq. 3.14.

$$I(int, hkl) = \frac{\lambda^3}{\omega V^2} \left(\frac{e^2}{mc^2} \right)^2 V_{cr} I_o \times L \times P \times A \times |F(hkl)|^2 \quad \text{Eq. 3.14}$$

onde V_{cr} = volume do cristal, V = volume da cela unitária, I_o = intensidade do feixe incidente, I = Intensidade do feixe difratado, L = Fator de Lorentz, P = Fator de Polarização, A = Fator de absorção, $F(hkl)$ = Fator de estrutura; e , m = carga e massa do elétron respectivamente e c é a velocidade da luz.

A medida das intensidades $I(hkl)$ foi feita em nosso laboratório usando-se um difratômetro Automático R-AXIS IIC (Rigaku Automated X-Ray Imaging System) para a coleta de dados de difração dos cristais de proteína, equipado com um detetor de área bidimensional baseado em sistemas de placas de imagem (IP) (Sato et al., 1992) e um anôdo rotatório RU 200 da Rigaku Co. operando com uma voltagem de 40 kV e uma corrente de 120 mA.

3.5.1 Geometria e Características Principais do R- AXIS IIC

O difratômetro automático R-AXIS IIC é apresentado na Fig. 3.8 (Rigaku Corporation, Instruction Manual for R-AXIX IIC).

Fig. 3.8: Difratorômetro automático R-Axis IIC.

3.5.2 Sistema de Leitura

(Detetor de Área - Sistema das Placas de Imagem)

O sistema de placa de imagens (Sato et al., 1992) é constituído por um suporte plástico flexível coberto por uma camada de brometo de bário dopado com európio: BaFBr:Eu^{+2} . Quando o európio é excitado pelo feixe de raio X $\text{Eu}^{+2} \rightarrow \text{Eu}^{+3}$, ele não reemite a radiação imediatamente e desta forma a imagem vai sendo gravada. Posteriormente, a leitura da placa de imagens é realizada através da excitação do európio por um laser de varredura, $\text{Eu}^{+3} \rightarrow \text{Eu}^{+2}$, que faz com que sejam emitidos fótons na forma visível. Esta luminiscência é amplificada por uma fotomultiplicadora e a intensidade medida é proporcional ao número de fótons de raios X absorvidos.

Um par de placas de imagem é empregado no sistema. Enquanto uma IP está sendo lida, a outra está em posição de exposição registrando a próxima imagem (tempo de leitura de aproximadamente 8 min) .

Após a exposição/leitura, as duas placas de imagens são rotacionadas de 90° e a imagem residual que permanece na IP após a leitura é apagada pelas lâmpadas de deleção. Após a exposição da segunda IP, as placas são rodadas por 90° novamente. Assim, o uso alternativo de duas placas de imagem reduz a perda de tempo devido ao tempo de leitura da IP durante a coleta de dados, tornando a coleta mais contínua e rápida.

3.6 Processamento Computacional da Imagens Medidas

Trataremos agora do processamento computacional das imagens medidas. O processamento de dados de intensidades fornecidas por métodos fotográficos ou por detectores de área envolve a indexação, a integração, a correção por vários fatores (Lorentz e Polarização) e quando necessário, o escalonamento e promediação de conjuntos de dados diferentes.

O processamento dos dados que porvêm do detetor de área R-AXIS IIC, por exemplo consiste em:

1. Das fotos “still”

- . Determinar a orientação do cristal.
- . Indexar as reflexões observadas.
- . Refinar a orientação do cristal.

2. Das fotos de oscilação

- . Medir a intensidade de cada reflexão.
- . Agrupar os dados de intensidade.
- . Calcular os fatores de escala e os fatores de temperatura para os dados.
- . Pós refinamento.
- . Promediar intensidades de reflexões equivalentes.
- . Obter os fatores de estrutura.
- . Avaliar o conjunto de dados coletados.

Após estas etapas de processamento, o arquivo de saída que obtemos contém os índices hkl de cada reflexão, os fatores de estrutura $|F(hkl)|$, os desvios dos $|F(hkl)|$ (σ_{hkl}) das reflexões observadas, a percentagem de dados coletados até uma dada resolução (completeza) e alguns índices que avaliam o conjunto de dados tais como o R_{merge} e o R_{sym} .

Com um bom conjunto de $|F_{\text{obs}}(\text{hkl})|$ os passos seguintes consistem na determinação das fases α_p , para partir para a construção de um modelo inicial da estrutura e um subsequente refinamento da mesma.

3.6.1 Coordenadas de Cada Reflexão no Filme

Fig. 3.9 mostra as relações entre o sistema de coordenadas do filme ou placa de imagem (Y_f, Z_f) e o sistema de coordenadas do cristal (X, Y, Z). O ponto P da rede recíproca está na superfície da esfera e tem coordenadas (x, y, z). A distância do cristal ao filme é $CO' = D$ (Arndt; Wonacott, 1977).

Fig. 3.9: Relação entre as coordenadas do filme (Y_f, Z_f) e o sistema de coordenadas do cristal (X, Y, Z). O ponto P (x, y, z) está sobre a esfera de Ewald.

As coordenadas de cada reflexão P' no filme são lidas diretamente. Consideremos agora o respectivo ponto do espaço recíproco P na superfície da esfera de reflexão. O vetor de

espalhamento $\vec{S}$ é calculado pela conversão das coordenadas da reflexão (Y_f , Z_f) no filme, obtidas durante o procedimento de pesquisa de picos, através da Eq. 3.15.

$$\vec{S} = \begin{pmatrix} D/r - 1 \\ Y_f/r \\ Z_f/r \end{pmatrix} \quad \text{Eq. 3.15}$$

onde $r = (Y_f^2 + Z_f^2 + D^2)^{1/2}$.

Para as fotos de oscilação em uma faixa de oscilação definida por (φ_1 , φ_2) as coordenadas de um ponto da rede recíproca são dadas pela Eq. 3.16 (Higashi, 1990):

$$\vec{x} = \phi \left[-\frac{(\varphi_1 + \varphi_2)}{2} \right] \vec{S} \quad \text{Eq. 3.16}$$

onde se usa a aproximação de que cada ponto do espaço recíproco registrado no intervalo angular (φ_1 , φ_2) cruza a esfera de Ewald em $\varphi = (\varphi_1 + \varphi_2)/2$. Esta aproximação é necessária porque o ângulo φ no qual a condição de difração é satisfeita não é conhecida durante as oscilações e é válida quando o incremento angular $\Delta\varphi$ é pequeno. A matriz de rotação ϕ descreve o ângulo no qual a reflexão é trazida á condição de difração. Nas fotos tipo “still” feitas com o cristal parado, o ângulo ϕ é constante, pois as imagens são coletadas normalmente a 90 ° uma da outra. A partir das coordenadas dos pontos da rede recíproca para cada reflexão da foto, no próximo passo vetores-diferença serão calculados entre todos os vetores $\vec{x}$ e desta forma obtidos os três vetores base para a determinação da cela unitária.

3.6.2 Cálculo dos Vetores Base para a Determinação da Cella Unitária e Matriz de Orientação do Cristal

A partir das coordenadas dos pontos da rede recíproca, que podem ser obtidas tanto das imagens still quando das imagens de oscilação são obtidos para cada reflexão da imagem, vetores diferença entre todos os vetores $\bar{x}$, calculados até uma dada resolução como mostrado na Eq. 3.17 (Higashi, 1990):

$$\bar{x}_{ij} = \bar{x}_i - \bar{x}_j \quad \text{Eq. 3.17}$$

Eles são então ordenados com base nos seus comprimentos e os vetores equivalentes são promediados. A importância da promediação é exatamente devido à existência de erros na medida de ϕ para cada reflexão, uma vez que a reflexão possui dimensão. Esses vetores são ordenados em termos da sua multiplicidade, isto é, o número de vetores diferença utilizados para a promediação.

Três vetores-base $\{v_i\}$ são selecionados a partir dos vetores diferença promediados e ordenados. É Usado o critério (1) e (2) da Eq. 3.18 para a seleção dos três vetores base .

$$\begin{aligned} (1) \quad & (\bar{v}_1, \bar{v}_2) > C_1 \\ (2) \quad & 90^\circ - (\bar{v}_1 \times \bar{v}_2, \bar{v}_3) > C_2 \end{aligned} \quad \text{Eq. 3.18}$$

O valor típico de C_1 é 30° e o valor de C_2 é aproximadamente o valor de $\theta_{\max}$.

Da lista de vetores base $\{v_i\}$ selecionados, o primeiro é normalmente suficiente para estabelecer a matriz de orientação inicial do cristal $\bar{A}$. A idéia básica é rodar-se esse conjunto-base escolhido em relação ao sistema de eixos X,Y,Z de tal forma a se indexar o maior número possível das reflexões de cada imagem *still*. Se o conjunto-base selecionado não for suficiente para se indexar as reflexões, é escolhido automaticamente outro conjunto-base e novamente tenta-se fazer a indexação. Este procedimento é continuado, até se conseguir um conjunto-base que indexe todas ou a grande maioria das reflexões das imagens. Então a matriz $\bar{A}$ pode ser escrita como apresentado na Eq. 3.19

$$\vec{A} = \begin{pmatrix} * & * & * \\ a_x & b_x & c_x \\ * & * & * \\ a_y & b_y & c_y \\ * & * & * \\ a_z & b_z & c_z \end{pmatrix} \quad \text{Eq. 3.19}$$

onde a_x^* é a projeção de a^* sobre X e assim por diante.

Portanto, é possível determinar os índices hkl de cada reflexão, a partir das coordenadas de cada reflexão no filme, do ângulo de rotação ϕ , da faixa de oscilação (ϕ_1, ϕ_2) e da matriz de orientação do cristal como mostrado na Eq. 3.20.

$$x = \vec{A}h \quad \text{e} \quad h = \vec{A}^{-1}x$$

$$\text{onde} \quad \begin{pmatrix} h \\ k \\ l \end{pmatrix} = \begin{pmatrix} * & * & * \\ a_x & a_y & a_z \\ * & * & * \\ b_x & b_y & b_z \\ * & * & * \\ c_x & c_y & c_z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{Eq. 3.20}$$

3.6.3 Refinamento da Matriz de Orientação do Cristal

Durante o processamento das fotos *still*, é possível fazer-se o refinamento dos parâmetros de rede e portanto, da matriz de orientação do cristal, dos parâmetros da câmara e da distância do cristal-detetor (IP). No programa PROCESS, esta matriz de orientação do cristal é calculada apenas uma vez a partir das imagens tipo “still” e utilizada no restante do processamento (imagens de oscilação). Outros programas, permitem que esta matriz de orientação seja recalculada a cada imagem de oscilação levando em consideração neste caso pequenos deslocamentos possíveis do cristal.

O refinamento da matriz de orientação do cristal $\bar{A}$, é baseado no fato de que os índices hkl de cada reflexão devem ser inteiros. Este refinamento é conduzido pela minimização da função ψ apresentada na Eq. 3.21 (Higashi, 1990):

$$\psi = \sum_h \omega |h - h'|^2, \quad h' = \bar{A}^{-1}x \quad \text{Eq. 3.21}$$

onde h é composto pelos valores inteiros mais próximos de h' . O fator peso utilizado aqui é

$$\omega = \frac{1}{d^*} = \frac{1}{|x|}.$$

O procedimento descrito acima é utilizado para se fazer a auto-indexação quando os parâmetros de rede da cela unitária são desconhecidos. Em alguns casos, entretanto, os parâmetros de rede já são conhecidos, e portanto, pode ser utilizado um procedimento alternativo. Neste caso, o método utilizado é semelhante ao método de Kabsch (Kabsch, 1988) que é baseado na comparação de vetores recíprocos observados com aqueles calculados dos parâmetros de cela conhecidos. Neste caso, a matriz de orientação do cristal é definida por três ângulos de rotação $\varphi_x, \varphi_y, \varphi_z$ em relação aos eixos ortogonais X,Y,Z. As contribuições dos parâmetros de cela (A_0) e das operações de rotação (R) para a matriz A podem ser expressas separadamente como apresentadas nas Eq. 3.22 (Higashi, 1990).

$$\bar{A} = \bar{R}\bar{A}_0$$

$$\bar{R} = \bar{R}_z(\varphi_z)\bar{R}_y(\varphi_y)\bar{R}_x(\varphi_x)$$

$$\bar{A}_0 = \lambda \begin{pmatrix} a^* & b^* \cos \gamma^* & c^* \sin \beta^* \\ 0 & b^* \sin \gamma^* & -c^* \sin \beta^* \cos \alpha^* \\ 0 & 0 & \frac{1}{c^*} \end{pmatrix} \quad \text{Eq. 3.22}$$

Como $\vec{A}_0$ é calculada dos parâmetros de rede conhecidos, os ângulos de orientação podem ser obtidos pela resolução destas equações. Uma vez determinada a matriz de orientação $\vec{A}$, o procedimento de auto-indexação é semelhante ao procedimento anterior, descrito para o caso em que não se conhecem os parâmetros de rede.

3.6.4 Conceitos Teóricos Envolvidos no Processamento das Imagens de Oscilação

A integração de intensidades feita pelo programa PROCESS utiliza o ajuste de perfil para o qual divide o filme em cinco regiões e calcula um perfil médio para as reflexões em cada região. Assim, a intensidade de cada reflexão de cada região é determinada descontando a intensidade do 'background'. A correção de Lorentz-polarização também é feita nesta etapa do processamento. Os dados necessários para a resolução de uma estrutura tridimensional de um cristal de proteína quase sempre são coletados utilizando-se vários cristais e/ou provêm de várias imagens de oscilação. Torna-se então necessário escalonar os diferentes conjuntos de dados ou diferentes filmes para um único conjunto, o que é feito com base nas reflexões comuns existentes em cada conjunto individual de reflexões.

3.6.4.1 Integração das Intensidades de Cada Reflexão em Cada Imagem de Oscilação

De forma a realizar a integração da intensidade de uma dada reflexão, torna-se necessário, inicialmente definir uma caixa em volta da reflexão, assim como a a região desta área considerada pico e a região de fundo (background). Na implementação de Rossmann a posição do pixel de densidade ótica dentro da caixa de reflexão é identificada pelas coordenadas (p,q), em passos de varredura.

Então, se ρ é a densidade ótica numa posição (p, q) para uma dada reflexão, a densidade ótica, subtraída o 'background', é dada pela Eq. 3.23, onde a , b e c são constantes e o 'background' é assumido como sendo um plano sobre a área da reflexão (Rossmann, 1985)

$$\rho - ap - bq - c \quad \text{Eq. 3.23}$$

As constantes a, b e c podem ser calculadas minimizando a Eq. 3.24 sobre os n pontos na caixa de reflexão que fazem parte do 'background'.

$$\sum_n (\rho - ap - bq - c) \quad \text{Eq. 3.24}$$

3.6.4.2 Determinação do Perfil da Intensidade

As informações contidas em muitas reflexões vizinhas podem ser usadas para avaliar a intensidade integrada de uma dada reflexão pelo uso de um perfil empírico (Rossmann, 1985). É portanto, razoável assumir que pelo menos para uma dada área do filme, o perfil de cada reflexão seja o mesmo, bastando um fator de escala (J) para descrever o tamanho do pico.

A variação do perfil sobre toda a área do filme ou placa de imagem somente pode ser determinada pela análise de um perfil médio nas diferentes regiões, desta forma, durante a primeira etapa de integração por ajuste de perfil, os perfis são calculados para cinco regiões do filme. Na Fig. 3.10 podemos ver as cinco regiões onde vão ser calculados os perfis médios das reflexões para depois serem realizadas as integrações das intensidades de cada reflexão (Rossmann, 1979).

Fig. 3.10: Os perfis das reflexões são calculados em cinco regiões da imagem de oscilação denotadas de I a V.

Uma reflexão na região I, é mais circular devido ao alto ângulo de incidência do feixe no detetor, enquanto que uma reflexão na borda do filme tem uma forma mais alongada. Na escolha do perfil médio para a integração das intensidades, é importante se levar em consideração estas diferenças.

Seja $P(p, q)$ o valor do perfil padrão em uma dada posição p, q da caixa de reflexão e J o fator de escala necessário para ajustar este perfil padrão a um dada reflexão. Então,

$$JP(p, q) = \rho - ap - bq - c$$

Minimizando o quadrado das diferenças entre $JP(p, q)$ e $\rho - ap - bq - c$ segue-se que o perfil para cada ponto (p, q) dentro da caixa de reflexão é dado pela Eq. 3.25. As somatórias para todas as N reflexões completas e não sobrepostas dentro de uma certa faixa de resolução para a dada área no filme. Os valores das intensidades integradas podem ser usadas como estimativa de J .

$$P(p, q) = \frac{\sum_n JP - p \sum_n Ja - q \sum_n Jb - \sum_n Jc}{\sum_n J^2} \quad \text{Eq. 3.25}$$

3.6.4.3 A Integração de Cada Reflexão pelo Método de Ajuste de Perfil

Partindo da predição exata das coordenadas de cada reflexão no filme e tendo-se uma estimativa de quais reflexões foram coletadas parcialmente ou completamente, as intensidades integradas podem ser determinadas pelo método de ajuste do perfil para as reflexões coletadas completamente e pela simples integração ao longo dos filmes para as coletadas parcialmente.

As constantes a , b e c que definem o melhor ajuste de um plano na região do background devem ser determinadas inicialmente para cada reflexão, como descrito anteriormente. Depois, um ajuste de perfil entre o perfil padrão, $P(p, q)$ quando escalonado pelo fator de escala J , pode ser determinado pela minimização da função E em relação a J apresentada na Eq. 3.26 (Rossmann, 1985) onde P representa $P(p, q)$ e

$$E = \sum_m [JP - (\rho - ap - bq - c)]^2 \quad \text{Eq. 3.26}$$

A soma será feita sobre todos os m pontos (p, q) dentro da área de pico da caixa de cada reflexão. Portanto obtemos a Eq. 3.27.

$$J = \frac{\sum_m \rho P - a \sum_m pP - b \sum_m qP - c \sum_m P}{\sum_m P^2} \quad \text{Eq. 3.27}$$

As quantidades $\sum_m pP$, $\sum_m qP$, $\sum_m P$, $\sum_m P^2$ podem ser determinadas inicialmente se um perfil constante está sendo usado. Entretanto, quando um perfil variável é computado para cada reflexão, estas somatórias devem ser reavaliadas para cada reflexão.

A densidade ótica total integrada sob o pico é então $\sum_m JP$ ou $J \sum_m P$. Então, a intensidade integrada pelo método de ajuste de perfil é dada pela Eq. 3.28, enquanto a intensidade simplesmente integrada seria dada pela Eq. 3.29.

$$I_1 = \frac{(\sum \rho P - a \sum p P - b \sum q P - c \sum P) \sum P}{\sum P^2} \quad \text{Eq. 3.28}$$

$$I_2 = \sum \rho - a \sum p - b \sum q - c_m \quad \text{Eq. 3.29}$$

$I_1 = I_2$ se $P=1$ para toda a área do pico, ou seja a integração equivale o uso de um perfil tipo função delta. Daí a superioridade do uso do perfil.

3.6.4.4 Escalonamento Relativo das Várias Imagens de Oscilação

Dados a alta resolução para a determinação de uma estrutura de proteína quase sempre são coletados através do uso de vários cristais, ou possivelmente de dados provenientes de várias imagens de oscilação. Torna-se necessário escalonar estes diferentes conjuntos de dados com base nas reflexões comuns entre os conjuntos de imagens. O método de Hamilton (Hamilton et al., 1965) é geralmente utilizado para este fim.

Considere $I(h)_i$ como a intensidade da reflexão h para o cristal, no filme i . É necessário calcular-se um conjunto de fatores de escala K_i , para cada um dos cristais ou filmes para ser aplicado a $I(h)_i$ de forma a determinar pelo método de mínimos quadrados o melhor valor de $I(h)$. Nas equações do método de mínimos quadrados é mais conveniente que os parâmetros desconhecidos K_i e $I(h)$ fiquem separados dos parâmetros observados de tal forma que a função de minimização não se altere se cada K_i for multiplicado pelo mesmo fator de escala. Substitui-

se $G_i = 1/K_i$ e resolve-se estas equações em termos dos recíprocos dos fatores de escala (Blundell; Johnson, 1976).

O método utilizado para o escalonamento assume que as equações observacionais $\phi(h)_i$ sejam dadas pela Eq. 3.30, onde $w(h)_i$ é o peso da i -ésima observação e é proporcional ao recíproco do desvio padrão de $I(h)_i$.

$$\phi(h)_i = \sqrt{w(h)_i} [I(h)_i - G_i I(h)] \quad \text{Eq. 3.30}$$

onde a função de escalonamento G_i é definida pela Eq. 3.31, onde C_i é uma constante e B_i é uma fator de dependência com a resolução. O fator $2B_i \sin^2 \theta / \lambda^2$ também é chamado de fator de temperatura, por analogia matemática com o fator de temperatura real. Este fator, como pode ser observado é um fator que varia com a resolução.

$$G_i = [C_i + \exp(\frac{2B_i \sin^2 \theta}{\lambda^2})]^{-1} \quad \text{Eq. 3.31}$$

Deseja-se minimizar ψ , Eq. 3.32, onde N é o número de observações da reflexão h .

$$\Psi = \sum_h \sum_{i=1}^N \phi(h)_i^2 \quad \text{Eq. 3.32}$$

de forma que o melhor valor de $I(h)$ é encontrado quando $\delta\psi/\delta I(h) = 0$, isto é:

$$I(h) = \frac{\sum_{i=1}^N [\omega(h)_i] G_i I(h)_i}{\sum \omega(h)_i G_i^2}$$

A solução das equações observacionais para G_i não é trivial pois são não lineares, e portanto têm que ser resolvidas usando técnicas de mínimos quadrados não lineares.

3.7 Avaliação do Conjunto de Dados

O conjunto de dados formado por diferentes imagens é escalonado com base na condição de que reflexões de mesmo índice ou relacionados por simetria medidas em fotos diferentes devem ter a mesma intensidade. A qualidade do conjunto de dados pode ser medida através de um índice de concordância R_{sym} das reflexões simetricamente equivalentes. A expressão usada para calcular o R_{sym} é dada pela Eq. 3.33.

$$R_{\text{sym}} = \frac{\sum_h \sum_i |I(\vec{h})_i - \langle I(\vec{h}) \rangle|}{\sum_h \sum_i \langle I(\vec{h}) \rangle} \quad \text{Eq. 3.33}$$

onde $\langle I(\vec{h}) \rangle$ é a intensidade média ponderada de todas as medidas de $I(h)_i$.

Um outro parâmetro utilizado na avaliação dos dados processados é a relação entre o número observado e o previsto de reflexões independentes por faixa de resolução. Este parâmetro é denominado completeza dos dados.

3.8 Coleta de Dados da BthTX-I

A coleta de dados de difração de raios X foi feita utilizando o difratômetro automático R-AXIS IIC da Rigaku Corporation do Laboratório de Cristalografia de Proteínas do IFQSC-USP. Para a coleta de dados foram usados 2 cristais da BthTX-I. Com os dados do primeiro cristal deu-se início à caracterização da BthTX-I e com os dados do segundo cristal determinou-se a estrutura da BthTX-I. Estas coletas de dados não foram realizadas no mesmo dia. Foram duas coletas de dados realizadas e processadas independentemente. Na tabela 3.1 estão apresentadas as condições de coleta de dados para estes cristais: (A) imagens still; (B) imagens de oscilação.

A) Imagens tipo still

Cristal	λ (Å)	D (mm)	#Imagens	Tempo (min)
1	1.5418	120	2	15
2	1.5418	120	2	12

B) Imagens de oscilação

Cristal	λ (Å)	D (mm)	Intervalo de Osc (°)	#Imagens	Tempo (min)	Res. (Å)
1	1.5418	130	4	15	40	2.3
2	1.5418	120	2	1-30	35	2.3
				31-45	50	

Tabela 3.1: Condições da coleta de dados da BthTX-I utilizando o difratômetro automático R-AXIS IIC. A) Imagens tipo still; B) Imagens de oscilação

Pode-se observar que a resolução é 2.3 Å em ambos os casos, porém D (mm) é diferente. Esta resolução 2.3 Å significa a resolução até onde foram processados os dados e não o limite teórico de resolução na borda da placa. A mudança do ângulo de oscilação foi realizada para eliminar sobreposições a alta resolução.

As imagens de oscilação foram realizadas a baixas temperaturas (4°C) utilizando um fluxo de nitrogênio gasoso, verticalmente ao capilar, como mostrado na Fig. 3.11.

Fig. 3.11: Fotografia do sistema de coleta de dados R-AXIS IIC, onde é mostrado o sistema de montagem do cristal (cabeça goniométrica) e o sistema utilizado para resfriar a amostra.

O padrão de difração apresentou simetria mínima de ordem 3 e portanto o ângulo de oscilação total de coleta para podermos gravar um conjunto completo de reflexões no espaço recíproco, devido à simetria do cristal foi de 60° com o eixo principal ao longo de ϕ . Os tempos de coleta de dados, do segundo cristal variaram devido ao fato de que depois da 30^a imagem as intensidades tornaram-se mais fracas. Na Fig. 3.12 é mostrada uma das imagens de oscilação do primeiro cristal.

Fig. 3.12: Imagem de oscilação do cristal da BthTX-I proveniente de um experimento de difração de raios X (condições de coleta do 1º. Cristal)

Também foi realizada uma coleta de dados no SERC Daresbury Laboratory na Inglaterra, com o objectivo de se obter dados a mais alta resolução (2.1 Å) e maior completeza para desta forma se melhorar a qualidade dos mapas. A coleta de dados foi realizada com um comprimento de onda $\lambda = 0.88$ Å na estação 9.6 do anel de armazenamento SRS, numa linha de luz fornecida da radiação de um ímã “wiggler” de 5T. Durante a coleta de dados o SRS foi operado com uma energia de 2GeV e uma corrente média de ≈ 200 mA. Os dados foram coletados a temperatura ambiente. O detector usado foi MAR com placa de 30cm e o tempo de exposição de cada imagem foi de 90 segundos. Foram coletadas 72 imagens de oscilação com 1º de oscilação cada uma. A distância do cristal ao detector foi de 300 mm e será a partir do processamento destes dados que a estrutura da BthTX-I vai ser apresentada e discutida

3.9. Processamento das Imagens da BthTX-I

Para o processamento das imagens coletadas com o Difratorômetro automático R-AXIS IIC foi utilizado o método de Higashi (1990), implemado pelo programa PROCESS (Rigaku Corporation, Instruction Manual for R-AXIS IIC).

O processamento das imagens coletadas no síncrotron de Daresbury da Inglaterra, foram processadas usando o o programa MOSFLM (Wonacott et al., 1995) por membros da equipe do próprio laboratório syncrotron.

Na tabela 3.2 apresentamos um resumo do processamento dos dados por PROCESS e por MOSFLM para a BthTX-I.

	BthTX-I	
	PROCESS	MOSFLM
# reflexões totais	12550	91388
# reflexões independentes	8178	15045
Sistema Cristalino	Trigonal	Trigonal
Grupo Espacial	P3 ₁ 21	P3 ₁ 21
Parâmetros de rede	a =b = 57.63 Å c=131.17Å $\gamma=120^\circ$	a =b = 57.58 Å c=131.27Å $\gamma=120^\circ$
R _{merge}	7.47 %	6.3 %
Resolução	2.3 Å	2.1 Å
Completeza	86.4 %	99.6 %

Tabela 3.2: Processamento dos dados da BthTX-I, processados com o programa PROCESS e MOSFLM.

Na tabela 3.3 encontra-se as estatísticas dos dados processados por MOSFLM, que serão os dados utilizados para a resolução da estrutura de BthTX-I.

ResoluçãoÅ	R _{sym}	R _{acumulad} o	I/Sigma	# Refl. Med.	# Refl. Equiv.	Completeza %
8.11	0.037	0.037	15.5	981	261	84.9
5.74	0.040	0.039	13.7	2008	474	92.8
4.69	0.042	0.040	14.0	2717	634	98.4
4.06	0.045	0.042	13.6	3108	726	96.0
3.64	0.046	0.043	13.1	3351	827	97.5
3.32	0.052	0.045	11.6	3946	923	99.4
3.07	0.059	0.046	10.8	4305	1004	99.5
2.87	0.070	0.048	9.3	4477	1054	97.8
2.71	0.082	0.050	8.2	4905	1151	100.0
2.57	0.092	0.052	7.7	4946	1197	99.4
2.45	0.110	0.054	4.0	5151	1262	100.0
2.35	0.122	0.056	5.8	5391	1312	99.9
2.26	0.145	0.058	5.0	5509	1358	99.0
2.17	0.166	0.061	4.0	5777	1395	98.4
2.10	0.201	0.063	3.6	6045	1467	98.6
Total		0.063		62817	15045	98.6

Tabela 3.3: Estatística dos dados de difração de BthTX-I processados por MOSFLM.

O processamento das imagens *still* do complexo inibidor/proteína citado anteriormente forneceu parâmetros de rede respectivamente igual a $a = 60.79 \text{ \AA}$, $b = 91.02 \text{ \AA}$ e $c = 53.03$ e $\beta = 91.16^\circ$. O cristal complexo/inibidor cristalizou no sistema cristalino monoclinico, mas não foi possível coletar dados com os cristais disponíveis devido à baixa qualidade do padrão de difração.

3.9 Referências Bibliográficas

- ARNDT, U. W. & WONACOTT, A. J. - *The Rotational Method in Crystallography*, North- Holland Publishing Company, Amsterdam (1977).
- ARNDT, U.W. - *J. Appl. Cryst.*, **17**, 118-119 (1984).
- BLUNDELL, T. L. & JOHNSON, L. N.; *Protein Crystallography*, Academic Press, New York (1976).
- BÜERGER, M. J. - *X-ray Crystallography* - John Wiley & Sons, Inc., London (1942).
- DRENTH, J. - *Principles of Protein X-ray crystallograohy* - Springer-Verlag (1994).-
- HIGASHI, T.; *J. Appl. Cryst.*, **23**, 253-257 (1990).
- DARWIN, C.G.- *Phil. Mag.* **27**, 315 (1914).
- GONZALEZ, A., DENNY, R., NAVE, C; *Acta. Cryst.*, **D50**, 276-282, (1994).
- GONZALEZ, A. & NAVE, C. - *Acta. Cryst.* **D50**, 874-877 (1994).
- HAMILTON, W. C., ROLLETT, J. S. & SPARKS, R. A.; *Acta Cryst.* **18** 129 (1965).
- HELLIWELL, J.R. -*Macromolecular Crystallography with synchrotron radiation*, Cambridge University Press (1992).
- KIM, S.; *J. Appl. Cryst.*, **22**, 53-60 (1989).
- KABSCH, W. - *J. Appl. Cryst.* **21**, 67-71 (1988).
- RIGAKU CORPORATION, *Instruction Manual* for R-AXIS IIC.
- ROSSMANN, M.; *Methods of Enzymology* vol. **114**, Academic Press (1985).
- ROSSMANN, M. G. - *J. Appl. Cryst.* **12**, 225-238 (1979).
- SATO, M., YAMAMOTO, M., IWADA, K., KATSUBE, Y., TANAKA, N.& HIGASHI, T. - *J. Appl. Cryst.* 348-357 (1992).
- WONACOTT, A. J., DOCKERILL, S. & BUCK, P., - *MOSFLM a Program for the Measurement of Integral Intensities on Oscillation Photographs*, Imperial College, University of London (1985).

CAPÍTULO 4

ANÁLISE PRELIMINAR E SUBSTITUIÇÃO MOLECULAR DA BthTX-I

4.1 Cálculo do Conteúdo da Cella Unitária.

A partir dos dados cristalográficos pode-se calcular o conteúdo da unidade assimétrica **n** (número de monômeros dentro da cela unitária) e o conteúdo do solvente V_w/V , usando as relações das Eq. 4.1 e 4.2 (Matthews, 1968)

$$n = \frac{NV(D_c - D_s)}{M_p[1 - \bar{v}_p D_s]} \quad \text{Eq. 4.1}$$

$$\frac{V_w}{V} = \frac{1}{D_w} (D_c - \frac{nM_p}{NV}) \quad \text{Eq. 4.2}$$

onde

D_c = Densidade do cristal

$D_w = D_s$ = Densidade do solvente (água)

V = Volume da cela unitária

V_p = Volume médio parcial específico de proteínas

M_p = Massa molecular da proteína em estudo

N = Número de Avogadro

V_w = Volume de água

A densidade de um cristal de proteína determinado experimentalmente está na faixa entre 1.10 a 1.30 g/cm³ e o volume parcial específico é aproximadamente igual a 0.74 cm³/g.

Com estas relações, assumindo que a densidade do solvente é unitária pôde-se construir a tabela 4.1 que nos mostra o conteúdo da cela unitária e a fração de solvente variando a densidade do cristal de 1.1 a 1.3 g/cm³, para a BthTX-I cuja M_p = 13720 Daltons, onde n é o número inteiro mais próximo do valor calculado.

D _c	n	V _w /V	D _c (n)
1.10	6	69.0	1.11
1.11	6	69.0	1.11
1.12	7	63.9	1.13
1.13	7	63.9	1.13
1.14	8	58.7	1.15
1.15	8	58.7	1.15
1.16	9	53.5	1.16
1.17	9	53.5	1.16
1.18	10	48.4	1.18
1.19	10	48.4	1.18
1.20	11	43.2	1.20
1.21	12	38.0	1.22
1.22	12	38.0	1.22
1.23	13	32.9	1.24
1.24	13	32.9	1.24
1.25	14	27.7	1.25
1.26	14	27.7	1.25
1.27	15	22.5	1.27
1.28	15	22.5	1.27
1.29	16	17.4	1.29

Tabela 4.1: Conteúdo de Moléculas da Cela Unitária da BTHtx-I e a Fração de Solvente.

A cela unitária, determinada a partir das imagens de difração, para os dados processados por PROCESS foi de a = b = 57.63 Å , c = 131.17 Å , α = β = 90° e γ = 120°, apresentando simetria mínima trigonal e o processamento de dados foi feito em P3. A partir da tabela acima para a determinação do número de moléculas na unidade assimétrica era preciso a determinação definitiva do grupo espacial.

Para a determinação do grupo espacial da BthTX-I foram inicialmente calculados valores de R_{sym} , supondo grupos pontuais diferentes e também a função de auto-rotação, ambos com o intuito de determinar a simetria cristalográfica. Para resolver a ambiguidade na mão do grupo espacial (além da resolução da estrutura) foi usado o método de Substituição Molecular.

Segue uma introdução teórica da função de Patterson, rotação e translação usados neste processo.

4.2 Função de Patterson e sua Interpretação

Em 1935, Patterson publicou um trabalho clássico onde era mostrado que a auto-correlação da densidade eletrônica (Brigham, 1974), ou seja

$$P(\vec{u}) = \int_V \rho(\vec{r})\rho(\vec{r} + \vec{u})dV \quad \text{Eq. 4.3}$$

poderia trazer uma grande contribuição á solução de estruturas cristalinas.

Na Eq. 4.3, $\rho(\vec{r})$ é a densidade eletrônica em uma posição $(x\vec{a} + y\vec{b} + z\vec{c})$ e $\rho(\vec{r} + \vec{u})$ é a densidade eletrônica em $[(x + u)\vec{a} + (y + v)\vec{b} + (z + w)\vec{c}]$. Sendo a função de Patterson resultante do produto de duas densidades eletrônicas, quando qualquer uma delas ou ambas forem iguais a zero (ou muito próximas) a função naquele ponto será praticamente nula. Por outro lado, os máximos ocorrerão quando as duas densidades $\rho(\vec{r})$ e $\rho(\vec{r} + \vec{u})$ forem bastante altas. Como isto só ocorre nas posições atômicas, vemos que a função de Patterson só apresentará máximos quando os vetores $\vec{u}$ representarem vetores interatômicos.

Considerando o exemplo de uma cela unitária unidimensional contendo três átomos com números atômicos Z_1 , Z_2 , e Z_3 localizados nas posições x_1 , x_2 e x_3 , como mostrado na Fig. 4.1.

Fig. 4.1: (a) Densidades Eletrônicas Unidimensionais.
(b) Mapa de Patterson Resultante.

Nesse caso, a função de Patterson pode ser escrita como na Eq. 4.3

$$P(\bar{u}) = \int_a \rho(x)\rho(x + \bar{u})dx$$

o que dá como resultado a Fig. 4.1 (b).

Os picos dessa função ocorrem para os valores de $\bar{u}$ iguais a zero (correspondentes á origem) e a $\pm(x_2 - x_1)$, $\pm(x_3 - x_2)$, $\pm(x_3 - x_1)$. Assim, se temos um pico no ponto uvw do mapa de Patterson, teremos no cristal átomos em x_1, y_1, z_1 e x_2, y_2 e z_2 tais que $u = x_1 - x_2$; $v = y_1 - y_2$ e $w = z_1 - z_2$. A Eq. 4.3 não representa a forma mais útil da função de Patterson na determinação de estruturas, pois necessita do conhecimento prévio da função da densidade eletrônica $\rho(r)$. Porém, pode se mostrar que a Eq. 4.4 representa a mesma função, só que no espaço recíproco.

$$P(\bar{u}) = \frac{1}{V} \sum_{\bar{h}} |F(\bar{h})|^2 \exp[2\pi i \bar{h} \cdot \bar{u}] \quad \text{Eq. 4.4}$$

Analisando esta equação vemos que a função de Patterson é uma série de Fourier que utiliza como coeficientes os módulos dos fatores de estrutura ao quadrado, que são justamente os dados que são conseguidos nos experimentos de difração. Daí se conclui que, fazendo a síntese de Fourier com $|F_o(\vec{h})|^2$, obtemos um mapa tridimensional em que os picos representarão os vetores interatômicos. Para uma molécula contendo N átomos deverão existir N^2 vetores interatômicos. Destes, N serão vetores nulos que se concentrarão na origem, formando um grande pico. Os restantes $N^2 - N$ vetores formarão picos que se distribuirão pela cela.

4.3 A Função de Rotação

A função de rotação é definida como a integral sobre uma região esférica Ω do produto de uma função de Patterson P^1 com a versão rotacionada dela própria (função de auto-rotação) ou de uma outra função P^2 (função de rotação cruzada)(Rossmann & Blow, 1962).

A função de auto-rotação pode ser utilizada quando há mais de uma subunidade idêntica na unidade assimétrica, somente diferindo na orientação e na posição. Assim, é possível determinar a operação geral envolvendo somente rotação, a qual faz com que os vetores de Patterson das diferentes subunidades coincidirem. A função de Patterson rotacionada é a mesma função estacionária da proteína a ser estudada. Logo, a importância da função de auto-rotação é procurar a existência de simetria não cristalográfica, isto é, as operações de simetria que relacionam os monômeros dentro da cela unitária.

No cálculo da função de rotação cruzada, uma molécula modelo é orientada na cela cristalina da estrutura desconhecida, podendo esta conter uma ou mais moléculas similares. No caso de mais moléculas, a função de rotação cruzada deve apresentar vários picos, correspondendo a cada uma das sobreposições da molécula modelo em cada uma destas moléculas separadamente.

Na formulação da função de rotação ênfase é colocada sobre a expansão da função de Patterson em termos de harmônicos esféricos como na “fast rotation function” de Crowther (Crowther, 1972; Navaza, 1987) para uma maior eficiência computacional. A função rotação é definida pela Eq. 4.5.

$$R(\phi, P^{(1)}, P^{(2)}) = \frac{1}{V_{\Omega}} \int P^{(1)}(\vec{r}) [T(\phi) P^{(2)}(\vec{r})] d^3 r \quad \text{Eq. 4.5}$$

onde

$$P^{(t)}(\vec{r}) = \sum_H |\vec{F}_{\vec{H}}^t|^2 \exp(2\pi i \vec{H} \cdot \vec{r})$$

sendo $H(t)$ o conjunto de vetores recíprocos disponíveis experimentalmente, $[T(\phi)P^{(2)}]$ é a função de Patterson rotacionada e P^1 é a função de Patterson observada. O uso das funções de Bessel e dos harmônicos esféricos minimizam o tempo de cálculo da função rotação e são portanto utilizadas como um recurso matemático para facilitar os cálculos.

A função de auto-rotação é utilizada para determinar a orientação relativa de subunidades idênticas na unidade assimétrica. A principal diferença entre a função de auto-rotação e a rotação cruzada, é que no caso da auto-rotação a função de Patterson rotacionada é a própria função de Patterson observada, enquanto que na rotação cruzada a função rotacionada é a do modelo, colocada dentro da cela unitária da estrutura desconhecida.

O programa AMoRe (Navaza, 1994) foi usado para calcular a função de auto-rotação da BthTX-I. O resultado é um mapa de rotação que apresenta picos com intensidades e orientações definidas pelos ângulos Eulerianos (α , β , γ). Esses picos são significativos sempre que uma subunidade se sobrepuser a outra, ou pela ação de simetria cristalográfica ou simetria não cristalográfica.

Os ângulos Eulerianos (α , β , γ) relacionam os eixos rodados X' , Y' e Z' com os eixos ortogonais não rodados X, Y, Z , definidos em relação aos eixos cristalográficos, como pode ser observado na Fig. 4.2 (Blow, 1985). O eixo Z está alinhado ao longo do eixo cristalográfico de maior simetria.

Fig. 4.2: Ângulos Eulerianos, α , β , γ .

Neste sistema, aplica-se primeiro uma rotação α em torno do eixo Z, depois uma rotação β ao redor da nova direção Y após a rotação α e finalmente uma rotação γ ao redor do novo Z. Uma característica peculiar de $R(\alpha, \beta, \gamma)$ são os máximos nas seções $\beta = 0$ (Moss, 1985) que origina-se do fato de que para $\beta = 0$ a rotação α e γ têm o mesmo efeito na rotação, e portanto, todas as rotações com o mesmo $\alpha + \gamma$ são idênticas, isto é, para $\beta = 0$ o que define a rotação é $\alpha + \gamma = \text{constante}$. No caso especial de $\gamma = -\alpha$ (rotação $\alpha, 0, -\alpha$) o resultante é uma rotação nula e portanto todos pontos apresentam o pico na origem. Efeitos equivalentes podem também ser obtidos para combinações das três componentes angulares.

A função de auto-rotação também pode ser representada por uma projeção estereográfica. Neste caso, os ângulos Eulerianos são transformados para o sistema de coordenadas esféricas polares (χ, ω, ϕ) , como mostrados na Fig. 4.3 (Blow, 1985) em que se define a função em termos da rotação χ em torno de um eixo, (chamado de eixo polar) cuja orientação é determinada por ω e ϕ . A busca por eixos de ordem 2 é feita através da análise de picos para ângulos $\chi = 180^\circ$ e pode ser representado por uma projeção estereográfica onde os picos aparecem em posições determinadas por ω e ϕ .

Fig. 4.3: Ângulos Polares Esféricos (χ , ω , ϕ).

Na aplicação do método de rotação é importante que todas as reflexões mais fortes estejam presentes porque o cálculo da função rotação basicamente depende da rotação do mapa de Patterson que é dominado pelas reflexões fortes. Outro ponto a ser considerado é a faixa de resolução do conjunto de dados a ser utilizado no cálculo da função de rotação. Dados a muito baixa resolução podem ser excluídos, pois são dominados pelo arranjo de solvente na proteína. Os dados a mais alta resolução são mais decisivos no cálculo da função de auto-rotação. Outro parâmetro de escolha é a forma e o tamanho da região de integração. Por conveniência assume-se uma região esférica onde o raio de integração pode ser escolhido. A escolha do raio máximo de integração está diretamente relacionado com o diâmetro da molécula a ser estudada. Como para o cálculo da função de rotação interessa somente vetores intra-moleculares, o raio máximo pode ser igual ou um pouco menor do que o diâmetro da molécula. Já o raio mínimo de integração deve ser escolhido de tal forma a eliminar o efeito do pico da origem no mapa de Patterson.

4.4 Função de Translação

A função de translação definida como a superposição das funções de Patterson calculada e observada, para uma dada orientação do modelo (rotação α, β, γ) em função da sua posição é definida pela função $T(\vec{t})$, Eq. 4.6

$$T(t) = \int_V P_{O1}(\vec{u}, \vec{t}) P(\vec{u}) d\vec{u} \quad \text{Eq. 4.6}$$

onde $P(\vec{u})$ representa a função de Patterson do cristal. Quando $\vec{t}$ é igual ao vetor $\vec{t}_0$ intermolecular, os vetores cruzados de Patterson P_{O1} se ajustam com a função de Patterson P , e $T(\vec{t}_0)$ terá um valor positivo grande (Crowther; Blow, 1967).

4.5 Substituição Molecular

Substituição Molecular (SM) é um dos métodos que pode ser utilizado para resolver o problema da fase na determinação de uma estrutura de proteína (Rossmann & Blow, 1962). Este método só pode ser aplicado quando se tem o prévio conhecimento de que uma molécula de estrutura conhecida é similar à estrutura que se quer determinar. Outros métodos largamente usados são MIR (Substituição Isomórfica Multipla) e MAD (dispersão anômala em comprimentos de onda variados).

Substituição Molecular usando um arquivo de coordenadas atômicas como fonte tem se tornado uma técnica padrão na resolução de estruturas para posicionamento de uma estrutura molecular conhecida na cela unitária em estudo. O grande sucesso deste método vem do crescente número de estruturas no banco de dados Brookhaven, aumentando assim a probabilidade de se encontrar um homólogo da proteína de estudo e também a potencia de novos algoritmos de Substituição Molecular tipo empregados em AMoRe. Além da potência dos algoritmos o sucesso de AMoRe é devido entre outros fatores às muitas soluções que podem ser exploradas.

Os seis parâmetros (α , β , γ , x , y , z) que definem a melhor rotação e translação entre as funções de Patterson são afetadas por diversos fatores, tais como: (1) erros experimentais na medida das intensidades e deficiência no modelo; (2) a inevitável superposição das funções auto e cruzada de Patterson dentro da região selecionada para a integração da função de Patterson; (3) limitações no raio de integração e intervalo de resolução das somas de Patterson impostas pela ausência de similaridade entre o modelo e o cristal; (4) correlação intrínseca entre a função rotação e translação que por conveniência computacional são tratados separadamente; (5) frequentemente o baixo tamanho do modelo em relação ao conteúdo da cela unitária.

De uma forma geral, o método de substituição molecular é baseado em três etapas principais: (1) Encontrar a orientação relativa do modelo fonte; (2) Localizar a posição do modelo no sistema cristalino da estrutura desconhecida; (3) Otimizar os parâmetros posicionais através do refinamento de corpo rígido e fazer uma avaliação da qualidade das soluções apresentadas.

A primeira e a segunda etapas são obtidas pelo uso de funções de Patterson. A posição do modelo é variada até sua função de Patterson calculada coincidir com a função de Patterson observada, isto é, da estrutura desconhecida. A orientação é obtida com a função de rotação (RF) e dada a orientação, o deslocamento é obtido através da função de translação (TF). Os parâmetros posicionais (α , β , γ , x , y , z) (Blow, 1985) são otimizados através do refinamento de corpo rígido (RBR).

A Fig. 4.4 representa a idéia geral do Método de Substituição Molecular, onde uma molécula modelo é posicionada na cela cristalina da molécula cuja estrutura é desconhecida, através da aplicação da matriz de rotação e do vetor de translação descritos pelos parâmetros α , β , γ , x , y , z . Uma vez posicionada, o modelo pode ser usado para o cálculo de um conjunto inicial de fases para a estrutura desconhecida.

Substituição Molecular

$$\text{Função de Rotação} \rightarrow R(\alpha, \beta, \gamma) = \int P_1(\vec{r}) P_2([C] \cdot \vec{r}) d\vec{r}$$

$$\text{Função de translação} \rightarrow T(t) = \int P_o(\vec{r}) P_e(\vec{r}, \vec{t}) d\vec{r}$$

Fig. 4.4: Representação esquemática do Método de Substituição Molecular.

Para a determinação da estrutura da BthTX-I pelo método de Substituição Molecular foi utilizado o pacote de programas AMoRe (Automatic Molecular Replacement) desenvolvido por Navaza (Navaza, 1994). AMoRe é um programa baseado em algoritmos fortes que permitem uma avaliação rápida e automática de um grande número de soluções. O pacote pode basicamente ser dividido em cinco etapas principais incorporadas nos seguintes programas: sorting, tabling, roting, traing e fitting.

4.5.1 SORTING

Analisa os arquivos de entrada e ordena os dados de h, k, l e F_{obs} , que serão utilizados nos próximos passos. Devem ser especificados os parâmetros de rede do cristal, grupo espacial, operações de simetria e os limites de resolução.

4.5.2 TABLING

Calcula os coeficientes de Fourier para a molécula modelo. Estes coeficientes são determinados pela transformada de Fourier da densidade eletrônica do modelo, a qual é colocada dentro de uma cela unitária quatro vezes maior que a molécula . A molécula é colocada com o seu centro de massa na origem e com os eixos principais de inércia ao longo dos eixos da cela. Uma vez calculada a transformada de Fourier $f(\mathbf{h})$ do modelo para uma dada rotação $R(\alpha, \beta, \gamma)$ e translação $\mathbf{x}$, o fator de estrutura calculado para o modelo é dado por (Navaza, 1994):

$$F(\alpha, \beta, \gamma, x, y, z)(\vec{H}) = \sum_s f[\vec{H}\vec{M}_s\vec{D}\vec{R}(\alpha, \beta, \gamma)\vec{O}] \exp(2\pi i \vec{H}\vec{t}_s) \exp(2\pi i \vec{H}\vec{M}_s\vec{x})$$

onde

$f(\mathbf{h})$ = a transformada de Fourier da densidade eletrônica do modelo

$\mathbf{H}$ = (H,K,L) o vetor do espaço recíproco

$\mathbf{M}_s$ = matriz correspondente à e-nésima operação de simetria

$\mathbf{D}$ = matriz de transformação de coordenadas ortogonais em fracionárias

$\mathbf{R}(\alpha, \beta, \gamma)$ = matriz de rotação 3x3 parametrizada por ângulos de Euler

$\mathbf{O}$ = matriz de transformação de coordenadas fracionárias em ortogonais

$\mathbf{t}_s$ = vetor de translação correspondente à e-nésima operação de simetria

$\mathbf{x}$ = vetor de translação na cela cristalina.

4.5.3 ROTING

Calcula a função de rotação, como descrita anteriormente no item 4.3.

4.5.4 TRAING

Cálcula a função translação. Para uma dada rotação, o programa faz translações da função de Patterson calculada e em cada posição calcula um fator R_f e um coeficiente de correlação C_c entre as funções de Patterson observada e calculada dados pela Eq. 4.7 e 4.8.

$$R_{\text{fator}} = \sum_H m_H \{ |F_H^{\text{obs}}| - |F_H(\alpha, \beta, \gamma, x, y, z)| \times (\sum_H m_H |F_H^{\text{obs}}|)^{-1} \} \quad \text{Eq. 4.7}$$

$$C_{\text{correl.}} = \sum_{\tilde{H}} m_{\tilde{H}} \Delta |F_{\tilde{H}}^{\text{obs}}| \Delta F_{\tilde{H}}(\alpha, \beta, \gamma, x, y, z) \\ \times \{ \sum_{\tilde{H}} m_{\tilde{H}} (\Delta |F_{\tilde{H}}^{\text{obs}}|)^2 \} \times \{ \sum_{\tilde{H}} m_{\tilde{H}} [\Delta |F_{\tilde{H}}(\alpha, \beta, \gamma, x, y, z)|]^2 \}^{-1/2} \quad \text{Eq. 4.8}$$

com $\Delta |F_H^{\text{obs}}| = |F_H^{\text{obs}}| - \langle F_k^{\text{obs}} \rangle$, e m_H é a multiplicidade da reflexão H .

A análise destes dois fatores ajudam o usuário na escolha da melhor solução.

4.5.5 FITTING

Faz refinamento de corpo rígido das soluções encontradas pela função de translação.

O novo refinamento de corpo rígido é baseado na minimização da expressão $Q(\alpha, \beta, \gamma, x, y, z, \lambda)$ em relação aos seis parâmetros $(\alpha, \beta, \gamma, x, y, z)$, ao fator de escala λ e ao fator de temperatura B (Castellano *et al.*, 1992).

$$Q(\lambda, \alpha, \beta, \gamma, x, y, z, B) = \sum_H m_H \{ |F_{\text{obs}}(H)| \exp(-B) |\vec{H}|^2 - \lambda |F_H(\alpha, \beta, \gamma, x, y, z)| \}^2$$

onde m é a multiplicidade da reflexão $\vec{H}$, (α, β, γ) os parâmetros orientacionais e (x, y, z) os parâmetros translacionais. F_H é calculado através da Eq. 4.9 onde a somatória em $\underline{n}$ representa os $\underline{N}$ possíveis monômeros na unidade assimétrica.

$$\vec{F}_H = \sum_n \sum_s f_n [\vec{H} \vec{M}_s \vec{D} \vec{R}(\alpha_n \beta_n \gamma_n) \vec{O}] \exp(2\pi i \vec{H} \vec{M}_s \vec{X}_n) \quad \text{Eq. 4.9}$$

O procedimento de minimização (mínimos quadrados) envolve derivadas de Q , as quais são calculadas partindo de expressões analíticas. A cada ciclo os valores ótimos de λ e B são determinados e a minimização com relação aos parâmetros posicionais é alternativamente realizada para cada modelo de procura, enquanto que os outros são mantidos fixos.

Os ciclos de minimização param quando o número de interações permitidas é atingido ou quando a norma do gradiente, definida como na Eq. 4.10 se tornar menor do que um dado valor fornecido pelo usuário.

$$N = \left[\frac{1}{6} \sum_{i=1}^6 \left(\frac{\partial Q}{\partial p_i} \right)^2 \right]^{1/2} \quad \text{Eq. 4.10}$$

No arquivo de saída do programa AMoRe são listadas todas as soluções com as respectivas matrizes de rotação (e translação) assim como os fatores de correlação e o fator R.

4.5.6 CORDING

Aplica as matrizes de rotação e translação nas coordenadas do modelo, produzindo desta forma novos conjuntos de coordenadas já na cela da estrutura desconhecida que serão utilizadas na determinação e refinamento da estrutura.

4.6. Resultados da Função de Auto-Rotação para a BthTX-I.

Um dos objectivos do cálculo da função de auto-rotação foi a determinação do grupo pontual do cristal da BthTX-I. O segundo foi a determinação da possível simetria não cristalográfica.

Os resultados foram obtidos para a faixa de resolução de 3 Å a 15 Å, sendo que o raio de integração da esfera de Patterson entre 2 Å e 18 Å, usando o programa AMoRe. Para o cálculo da função de auto-rotação e Substituição Molecular foram utilizados os dados processados por MOSFLM.

Na Fig. 4.5 mostramos a projeção estereográfica da função de auto-rotação ao longo do eixo Z, em seções de χ igual a 180° , mostrando a existência de eixos cristalográficos de ordem 2 na direção de $\vec{a}$ e $\vec{b}$.

Fig. 4.5: Projeção estereográfica da função de auto-rotação ao longo do eixo Z, em seções de $\chi = 180^\circ$.

Da análise destes resultados podemos admitir que os grupos espaciais possíveis para a BTHTx-I são $P321$, $P3_121$, $P3_221$ (todos com a mesma simetria de Patterson $P \bar{3}m1$) e os grupos $P622$, $P6_122$, $P6_522$, $P6_222$, $P6_422$ e $P6_322$ com simetria de Patterson $P6_m mm$.

Para descobrir o grupo espacial da BTHtx-I utilizamos do pacote de programas CCP4 os programas ROTAVATA (Fox; Holmes, 1966) e AGROVATA (Smitt; Wonacott, 1992).

O programa ROTAVATA calcula fatores de escala lineares e exponenciais para as intensidades obtidas de cristais diferentes. O programa AGROVATA aplica os fatores de escala calculados por ROTAVATA, e promedia as reflexões equivalentes. Após este tratamento dos dados obtemos para cada grupo espacial (isto é, para cada simetria de Patterson) um fator R_{sym} que indica a discrepância interna dos dados. Foi exatamente da análise destes destes fatores R_{sym}

para cada simetria de Patterson que a princípio encontramos o grupo espacial. Esses fatores R_{sym} foram: 3% para os grupos P321, 15% para P622 e 48% para P312.

Os três grupos espaciais possíveis foram P321, $P3_121$ e $P3_221$ e devido á observada condição de reflexão, 00l com $l=3n$, os únicos grupos espaciais finais possíveis seriam $P3_121$ e $P3_221$.

A determinação definitiva do grupo espacial foi realizada quando utilizamos o método de Substituição Molecular para a resolução da estrutura da BthTX-I.

4.7 Resolução da Estrutura BthTX-I usando o programa AMoRe

Os grupos espaciais $P3_121$ e $P3_221$ possuem 6 unidades assimétricas por cela unitária. Consultando a tabela 4.1 demonstra-se portanto, que um cristal com duas moléculas de BthTX-I, por unidade assimétrica teria 38% de solvente ou alternativamente 1 molécula na unidade assimétrica e 69% de solvente.

Para a resolução da estrutura da BTHtx-I pelo método de Substituição Molecular (AMoRe) a estrutura utilizada como modelo foi a estrutura de *Agkistrodon piscivorus piscivorus* (Holland *et al.*, 1990) resolvida a 2.0 Å de resolução que cristalizou no grupo espacial $P2_12_12$ contendo um monômero por unidade assimétrica. A procura foi realizada entrando com apenas um monômero da proteína modelo para encontrar duas moléculas da BTHtx-I na unidade assimétrica. Esta procura foi realizada utilizando-se os dois grupos espaciais possíveis $P3_121$ e $P3_221$ e a faixa de resolução foi de 10 a 3 Å. A que nos forneceu melhor fator R e fator de correlação foi a do grupo espacial $P3_121$, Na tabela 4.2 encontram-se duas das soluções potenciais obtidas para a estrutura da BthTX-I no grupo espacial $P3_121$.

solução	molécula	α	β	γ	x	y	z	C_c	R_f
1	1	219.7	84.6	53.0	-25.0	97.8	73.6	62.7	38.7
	2	100.8	74.0	265.1	-13.9	-19.4	86.8		
2	1	219.8	84.6	52.5	-25.0	97.8	73.9	32.3	50.8
	2	120.4	38.3	57.9	79.6	19.8	105.2		

Tabela 4.2 As soluções potenciais obtidas para a estrutura da BthTX-I, programa AMoRe (Navaza, 1987), no grupo espacial P3₁21.

Para cada uma destas soluções foi fornecida uma matriz de rotação e uma de translação, representados pelos parâmetros α , β , γ , x, y e z. A primeira solução é a solução correta e além de apresentar melhores valores para C_c e R_f também se destaca pelo contraste em relação à segunda solução.

Estas duas transformações representadas pela primeira solução foram aplicadas ao monômero da BthTX-I da *Agkistrodon piscivorus piscivorus* e desta forma foi obtido o primeiro modelo para as duas moléculas da BthTX-I na unidade assimétrica do seu cristal, que serão refinadas e analisadas nos próximos capítulos. A BthTX-I cristalizou no grupo espacial P3₁21, contendo, portanto 12 monômeros na cela unitária e da análise da tabela 4.1 um conteúdo de solvente de 38%.

4.8 Referências Bibliográficas

- BLOW, D. M. *Molecular Replacement* - Proceedings of The Daresbury Study Weekend Daresbury Laboratory (1985).
- BRIGHAM, E. D. - *The Fourier Transform*, Prentice-Hall, Inc., New Jersey (1974).
- CROWTHER, R. A. & BLOW, D. M. - *Acta. Cryst.* **23**, 544- 548 (1967).
- CROWTHER, R. A. - *The Molecular Replacement Method*, Gordon-Breach New York (1972).
- CCP4 PROGRAM *Suite Rotation Function* - ALMN - User Documentation Daresbury Laboratory (1984).
- CASTELLANO, E. E.; OLIVA, G. & NAVAZA, J., - *J. Appl. Cryst.* **25**, 281-284 (1992).
- FOX & HOLMES - *Acta Cryst.* **20**, 886 (1966)
- HOLLAND, D.R., CLANCY, L. L., MUCHMORES, S. W., RYDE, T.J., EINSPAHR, H. M., FINZEL, B. C., HEINRIKSON, R. L. & WATENPAUGH, K. D. - *The Journal of Biology chemistry* **265**, 29, 17649-17656 (1990).
- MOSS, D. S. - *Molecular Replacement* - Proceedings of the Daresbury Study Weekend, Daresbury laboratory (1985).
- MATTHEWS, B. W. *J. Mol. Biol.* **33**, 491-497 (1968).
- NAVAZA, J. - *Acta Cryst.* **A50**, 157-163 (1994).
- NAVAZA, J. - *Acta Cryst.* **A43**, 645-653 (1987).
- ROSSMANN, M. G. & BLOW, D. M. - *Acta Cryst.* **15**, 24-31 (1962).
- SMITH, J. M. & WONACOTT, A. J. - Phil Evans / Daresbury Laboratory (1992).
- WONACOTT, A.J., DOCKRILL, S. & BRICK, P. - *MOSFLM A Program for the Measurement of Integral Intensities on Oscillation Photographs*, Imperial College University of London (1985).

CAPÍTULO 5

REFINAMENTO DA ESTRUTURA DA BthTX-I

5.1 Introdução

Se já possuímos um modelo, (ou seja um conjunto de coordenadas atômicas e fatores de temperatura) da estrutura a ser estudada torna-se fácil avaliar a concordância dos dados experimentais e dos fatores de estrutura calculados, obtidos do modelo. A qualidade do processo de refinamento deste modelo é avaliada, normalmente através do cálculo do índice de discordância (Fator R) entre os módulos dos fatores de estrutura observados $|F_{obs}|$ e calculados $|F_{calc}|$, Eq. 5.1:

$$R = \frac{\sum_h \left| |F_{obs}(\vec{h})| - k|F_{calc}(\vec{h})| \right|}{\sum_h |F_{obs}(\vec{h})|} \quad \text{Eq. 5.1}$$

onde k é um fator de escala.

O refinamento cristalográfico se refere a um método onde o modelo inicial é ajustado afim de reduzir a discrepância entre F_{obs} e F_{calc} e portanto, desta forma diminuir o fator R. A queda do fator R, portanto significa que o modelo está reproduzindo melhor os dados experimentais. O ajuste do modelo consiste na variação das coordenadas atômicas e dos fatores de temperatura isotrópicos para todos os átomos (exceto hidrogênios) levando em consideração critérios estereoquímicos. A melhoria nas posições e nos parâmetros térmicos dos átomos deve gerar novas fases para os fatores de estrutura, que passam a ser usadas no cálculo de novos mapas de densidade eletrônica.

Várias metodologias podem ser utilizadas para o refinamento de uma estrutura de proteína. As mais conhecidas são: (1) Refinamento no espaço direto onde o modelo atômico é otimizado na densidade, (por exemplo Diamond, 1971); (2) Minimização de energia na qual simultaneamente fatores de estrutura e energia são minimizados por exemplo EREF (Jack & Levitt, 1978) ; (3) Refinamento de corpo rígido dos fatores de estrutura de grupos rígidos com restrições, por exemplo CORELS (Sussman *et al.*, 1977); (4) PROLSQ (Konnert & Hendrickson, 1980) e RESTRAIN (Haneef *et al.*, 1985) refinamento das posições atômicas com restrições estereoquímicas; (5) Método de Fast Fourier Transform (FFT) exemplo Agarwal (1978) e Tronrud *et al.*, (1978) e (6) Dinâmica Molecular/Simulated annealing exemplo Brünger *et al.*, (1987). Dentre eles o programa X-PLOR (Brünger, 1993) rapidamente ganhou popularidade devido às suas grandes inovações e recursos que de certa forma se somaram para melhorar e avaliar a qualidade do modelo refinado e principalmente automatizar o processo de tal maneira a reduzir a necessidade de intervenção humana no processo. No refinamento, utilizando o programa X-PLOR inclui-se termos de energia, fazendo uma minimização desta durante o processo de refinamento. Também se utiliza a dinâmica molecular, uma vez que durante o processo de refinamento o raio de convergência é limitado. A vantagem do uso de dinâmica molecular através do simulated annealing é que permite passar barreiras energéticas, possibilitando assim, a saída de mínimos locais.

5.2 Refinamento Utilizando o Programa X - PLOR.

O programa X-PLOR é bastante versátil, adequando-se a praticamente todos os problemas encontrados durante o refinamento de estruturas de macromoléculas utilizando tanto dados de difração de raios-X quanto dados de ressonância magnética (NMR), (Brünger, 1992).

No refinamento de uma estrutura macromolecular cristalográfica a partir do modelo inicial (x_j , y_j , z_j , B_j) pode-se calcular $F_c(\mathbf{h})$, Eq. 5.2.

$$F_c(\vec{h}) = \sum_j f_j(\vec{h}) \exp(2\pi i \vec{h} \cdot \vec{r}_j) \exp(-B_j \frac{\sin^2 \theta}{\lambda^2}) \quad \text{Eq. 5.2}$$

onde B_j é o fator de temperatura relacionado com o deslocamento médio quadrático da vibração atômica $\langle \mu \rangle^2$ dada por $B = 8\pi^2 \langle \mu \rangle^2$.

O método dos mínimos quadrados em refinamento funciona bem para moléculas pequenas, pois temos muitas observações por parâmetro a ser refinado (aproximadamente 10 observações para cada parâmetro, em caso típicos). No caso de proteínas temos tipicamente de 2 a 4 observações por parâmetro, portanto devemos levar em consideração outras informações tais como estereoquímicas para aumentar o número de dados e/ou reduzir o número de parâmetros.

O programa X-PLOR (Brünger, 1992) faz uma minimização da função energia E_{total} , Eq. 5.3 que pode ser separada em dois grupos diferentes:

$$E_{Total} = E_{Empirico} + E_{Efetivo} \quad \text{Eq. 5.3}$$

Os termos de energia empírica incluem as energias das ligações covalentes, dos ângulos de ligação, ângulos diédricos, interações de van der Waals e interações eletrostáticos. O termo efetivo é o termo cristalográfico que é usualmente dado pela Eq. 5.4, onde N_A é o fator de normalização, ω_h é a função peso individual para cada reflexão, W_A é um peso adicional que relaciona o termo cristalográfico com o termo empírico.

$$E_{Efetivo} = \left(\frac{W_A}{N_A} \right) \sum_{\vec{h}} \omega_{\vec{h}} [|F_{obs}(\vec{h})| - k |F_{calc}(\vec{h})|]^2 \quad \text{Eq. 5.4}$$

O fator de escala k entre os fatores de estrutura observados e calculados é usualmente dado pela Eq. 5.5.

$$k = \frac{\sum_{\vec{h}} w_h |F_{obs}(\vec{h})| |F_c(\vec{h})|}{\sum_{\vec{h}} \omega_{\vec{h}} F_c^2(\vec{h})} \quad \text{Eq. 5.5}$$

O refinamento segue no sentido de minimizar a energia total, calculando as variações nas coordenadas atômicas e parâmetros térmicos que conduzam ao mais baixo valor de energia.

5.2.1 Etapas do Refinamento.

5.2.1.1 Rotina Topologia.

Uma vez obtido o modelo inicial da estrutura a ser determinada gera-se inicialmente toda a informação molecular da estrutura, utilizando-se a rotina TOPOLOGIA do programa X-PLOR (Brünger, 1993).

A rotina topologia fornece informações sobre os resíduos que constituem uma determinada macromolécula, ou sobre os átomos que compoem os substratos, cofatores e moléculas de água. Estes dados incluem descrições sobre a massa, carga dos átomos, ângulos e ângulos diédricos e outras informações.

A rotina topologia cria um arquivo de coordenadas atômicas com formato .pdb (protein data bank) onde são incluídos todos os hidrogênios e também um arquivo .psf (protein structure file) que contém todas as informações estruturais e moleculares da estrutura em estudo.

5.2.1.2 Refinamento de Corpo Rígido

Este tipo de refinamento trata a molécula como um grupo rígido e pode-se refinar sua posição e orientação dentro da cela unitária como um todo. Este é normalmente o primeiro passo no procedimento de refinamento particularmente em casos de modelos que vêm do procedimento de Substituição Molecular e serve para orientar a molécula dentro da cela unitária. Todas as variações nas coordenadas são aplicadas ao corpo como um todo. Para cada grupo rígido é possível refinar três parâmetros rotacionais e três parâmetros translacionais.

A função E_{total} é usada no processo de refinamento. Os parâmetros rotacionais são os três ângulos de Euler α , β , e γ para uma rotação em torno do centro geométrico do grupo rígido. Partes da molécula que não estiverem especificadas dentro de qualquer grupo rígido serão mantidas fixas.

5.2.1.3 Rotina Check e a Determinação do Peso W_A

Este procedimento checa o valor inicial de R e determina o peso ideal W_A entre o termo cristalográfico E_{Efetivo} e os termos da $E_{\text{empírico}}$ que levam em conta a estereoquímica da estrutura. A rotina CHECK deve ser realizada antes do refinamento propriamente dito. O valor de W_A é obtido através da rotina check que faz uma breve dinâmica molecular sem incluir o termo de E_{Efetiva} .

Normalmente, uma vez obtido o valor de W_A recomendado pela rotina check é utilizada uma porcentagem maior ou menor, dependendo do peso que se quer dar aos termos de raio X e aos termos empíricos. Os valores utilizados para W_A para o termo cristalográfico estão normalmente entre 30-80% do valor obtido da rotina check.

5.2.1.4 Refinamento por Dinâmica Molecular com Temperatura

Controlada (Simulated Annealing)

No refinamento convencional por mínimos quadrados a função a ser minimizada converge facilmente para um valor mínimo se o modelo inicial não for muito diferente da estrutura real. Entretanto, se as distâncias entre os átomos no modelo e na estrutura real são muito grandes, o refinamento pode ser conduzido a um mínimo local, ao invés de alcançar um mínimo global ou mínimo verdadeiro. Para evitar esta situação utiliza-se a dinâmica molecular com temperatura controlada (simulated annealing) cujo princípio básico é simular uma situação onde a proteína é levada a altas temperaturas, ($\approx 3000\text{K}$), seguida de um processo de resfriamento até a temperatura ambiente, 300K. Durante este processo, a molécula adquire energia suficiente para superar as barreiras e escapar dos mínimos locais. Esta técnica foi utilizada pela primeira vez por Brünger (Brünger *et al.*, 1987) e tem sido usada com sucesso na cristalografia de macromoléculas (Kuriyan *et al.*, 1989, Brünger *et al.*, 1989 e Brünger *et al.*, 1990).

Na dinâmica molecular o comportamento dinâmico de um sistema de partículas é simulado. Esta simulação conduz a um conjunto de estruturas que são energeticamente permitidas para uma dada temperatura e pressão: a distribuição de energias das estruturas do conjunto obedece à lei de Boltzmann, segundo a qual o número de estruturas com energia

potencial E_{pot} é proporcional a $\exp[-E_{\text{pot}} / kT]$, onde k é a constante de Boltzmann, T é a temperatura absoluta e a energia potencial depende da posição relativa dos átomos. No tempo $t=0$, os átomos estão numa configuração que possui uma dada energia potencial E_{pot} para toda a molécula. Para cada átomo i numa posição $\vec{r}_i$, a força que está atuando é a derivada da energia potencial Eq. 5.6.

$$\vec{F}_i = - \frac{\partial E_{\text{pot}}}{\partial \vec{r}_i} \quad \text{Eq. 5.6}$$

Com a equação de Newton Eq. 5.7, a aceleração $\frac{\partial^2 \vec{r}_i}{\partial t^2}$ para cada átomo i com massa m_i pode ser calculada e então aplicada.

$$\vec{F}_i = m_i \left(\frac{\partial^2 \vec{r}_i}{\partial t^2} \right) \quad \text{Eq. 5.7}$$

Depois de um pequeno intervalo de tempo Δt ($1\text{fseg} = 10^{-15}$ segundos) o processo é repetido com os átomos na nova posição.

Se o número de passos são suficientes, a energia E_{pot} mínima pode ser alcançada e a informação sobre o comportamento dinâmico dos átomos na molécula é obtido. Se a temperatura do sistema é aumentada para um valor muito alto, mais átomos terão uma velocidade maior e uma alta energia cinética e podem vencer barreiras de energia maiores. A idéia básica da dinâmica molecular como uma técnica de refinamento é aumentar a temperatura de tal forma a encontrar o mínimo global de uma função $f(x) = f(x_1, x_2, \dots, x_n)$ onde x_1 , x_2 e x_n são parâmetros, tais como: posições atômicas e fatores de temperatura para que os átomos vençam barreiras de energia e escapem dos mínimos locais.

Com o uso do simulated annealing se torna possível corrigir erros maiores nos modelos, sendo refinados sem necessidade de intervenção humana. Por exemplo, uma cadeia lateral inicialmente construída numa conformação errada poderia passar para a conformação correta

através da simulação dinâmica. Uma vez nesta conformação haverá uma tendência de se manter nesta posição segurado pelo termo cristalográfico na função de energia.

A temperatura T de simulação não deve ser entendido como uma temperatura física, mas apenas como um parâmetro de controle para que o sistema escape de mínimos locais.

A simulação da dinâmica molecular envolve a solução simultânea das equações clássicas de movimento para todos os átomos i da macromolécula onde as forças são derivadas da energia potencial, E_{pot} , Eq. 5.8.

$$m_i \frac{\partial^2 \vec{r}_i}{\partial t^2} = -\nabla_i E_{\text{pot}} \quad \text{Eq. 5.8}$$

As quantidades $\vec{r}_i$ e m_i são as coordenadas e as massas do átomo i, respectivamente. A solução destas equações de movimento são realizadas no programa X-PLOR, numericamente usando o algoritmo de Verlet (Verlet, 1967), onde a E_{pot} é dada pela Eq. 5.3.

O sucesso do SA (Simulated Annealing) depende da escolha conveniente de alguns parâmetros, tais como: a temperatura e os pesos relativos dos vários componentes da função E_{Total} .

5.2.1.5 Refinamento Posicional com Minimização de Energia.

O programa X-PLOR usa o método de gradiente conjugado de minimização de energia desenvolvido por Powell (Powell, 1977) para fazer o refinamento das coordenadas dos átomos. A estrutura molecular, parâmetros e as coordenadas atômicas devem estar definidas. Este refinamento é realizado em vários ciclos.

5.2.1.6 Refinamento dos Fatores de Temperatura (Individuais ou de Grupos de Átomos).

O refinamento dos fatores de temperatura é normalmente realizado após o refinamento das coordenadas atômicas de cada átomo. Este refinamento é feito usando o método de gradiente conjugado de minimização de energia utilizando como energia potencial não a E_{total} , mas somente minimizando o termo E_{Efeito} . Para o caso de se refinar os B's isotrópicos para cada átomo, a variável é somente o fator B e $F_{calc}(\vec{h})$ é substituído pela Eq. 5.9,

$$\exp\left(\frac{-S^2 B}{4}\right) F_{calc}(\vec{h}) \quad \text{Eq. 5.9}$$

onde $S = \frac{2 \sin \theta}{\lambda}$ e B é o fator de temperatura.

Devido á relação número de observações/parâmetros a serem refinados é muitas vezes impossível se refinar os fatores B's isotrópicos e muitas vezes torna-se necessário refinar apenas um fator B para a cadeia principal e B's para as cadeias laterais.

5.2.1.7 Análise do Modelo Refinado

A determinação de uma estrutura por cristalografia envolve o ajuste das coordenadas atômicas do modelo com os dados de difração experimentais. A medida tradicional da qualidade deste ajuste e, provavelmente, da precisão do modelo é o fator R, apresentado na Eq. 5.1.

Em 1992, Brünger (Brünger, 1992) fez um outro grande avanço na análise do procedimento de refinamento, criando um novo e importante critério de avaliação, o R_{free} . O programa X-PLOR divide os dados cristalográficos em dois conjuntos de dados: (1) um conjunto de trabalho (A) que contém 90-95% dos dados e que será utilizado durante o processo de refinamento e (2) um conjunto teste (T) com usualmente 5-10% dos dados, retirados aleatoriamente do conjunto de dados total. Os dados deste conjunto teste não serão incluídos no refinamento.

O refinamento segue seu protocolo habitual e ao final do refinamento teremos dois fatores R. O fator R(A) é o fator da Eq. 5.1 e o fator R_{Free} é o fator R do conjunto de dados não refinados, e é dado pela Eq. 5.10.

$$R_{\text{T}}^{\text{Free}} = \frac{\sum_{\text{hkl} \in \text{T}} |F_{\text{obs}}(\text{hkl}) - k|F_{\text{calc}}(\text{hkl})|}{\sum_{\text{hkl} \in \text{T}} |F_{\text{obs}}(\text{hkl})|} \quad \text{Eq. 5.10}$$

A importância do R_{Free} tem-se tornado objeto de estudo. Kleywegt e Jones, (1995) avaliaram protocolos de refinamento e analisaram estes dois fatores R como fatores que avaliam a qualidade do procedimento de refinamento. Se o refinamento estiver minimizando E_{total} e o modelo for fisicamente incorreto, o valor de R pode até ter um valor baixo, mas o valor de R_{Free} aumentará durante o processo. A monitorização do R_{Free} é um indicador de confiabilidade do processo de refinamento do modelo. Um aumento ou falta de queda do R_{free} é um indicio de ‘overfitting’ dos dados, provavelmente uma consequência do uso de mais parâmetros do que justificados.

5.2.1.8 Simetria Não-Cristalográfica (NCS)

Das primeiras caracterizações do cristal de proteína, fica normalmente claro quantas moléculas de proteína ou subunidades estão contidas na unidade assimétrica. Se há mais de um monômero na unidade assimétrica diz-se que a estrutura pode possuir simetria não cristalográfica, ou seja, pode haver uma operação de simetria não cristalográfica relacionando as subunidades. Através do método de Substituição Molecular pode-se obter a posição e a orientação relativa destas subunidades.

Durante o processo de refinamento, pode-se tratar a simetria não cristalográfica (NCS) de diferentes modos: (1) se as moléculas forem consideradas iguais, utiliza-se apenas um modelo e utilizam-se matrizes que relacionam as subunidades, refinando desta forma apenas um monômero e reduzindo portanto, o número de parâmetros a serem refinados. Este é o chamado refinamento utilizando NCS estrita; (2) no refinamento com simetria não cristalográfica restrita, os monômeros são considerados independentes com apenas algumas regiões fixas consideradas iguais. Este vem do fato de que mesmo os monômeros sendo iguais sequencialmente, eles estão em ambientes químicos diferentes dentro do cristal e portanto pode haver pequenas diferenças estruturais entre si. Por esta razão, às vezes é interessante considerar apenas parte da estrutura idêntica (por exemplo, as regiões de estrutura secundária) e liberamos as regiões mais flexíveis da estrutura, tal como loops e regiões expostas ao solvente.

5.2.1.9 Colocação das Moléculas de Água

A água influencia profundamente todas as interações moleculares dos sistemas biológicos. A polaridade e a capacidade de formação de pontes de hidrogênio da água tornaram-na um componente importante na determinação de estruturas de proteína, enovelamento de proteínas e reconhecimento molecular. A água é um excelente solvente para moléculas polares. Ela enfraquece enormemente as forças eletrostáticas e as pontes de hidrogênio entre moléculas polares.

A localização de moléculas de água que estão estruturadas internamente na proteína e/ou nas camadas de hidratação que envolvem a proteína faz parte da resolução de uma estrutura de proteína.

Tipicamente águas são incluídas na proporção de aproximadamente 5 a 10% do número de átomos da estrutura, dependendo da qualidade e resolução dos dados.

5.3 Mapas de Densidade Eletrônica

Durante o processo de refinamento da estrutura, mesmo usando a metodologia mais automatizada utilizada pelo XPLOR, precisa-se interromper o processo em várias etapas para visualmente inspecionar mapas de densidade eletrônica e ajustar manualmente a estrutura de acordo com o mesmo.

Para o cálculo da densidade eletrônica é necessário conhecer os $F(hkl)$ em termos de amplitude e fase. Vimos anteriormente que do padrão de difração são medidas apenas as intensidades do feixe difratado, (e deste modo as amplitudes) sendo que toda a informação a respeito da fase é perdida. Uma vez resolvido o problema da fase, por um dos métodos conhecidos descritos anteriormente os $F(hkl)$ passam então a ter estimativas das fases. A densidade eletrônica pode ser calculada para cada ponto (x,y,z) da cela unitária usando a Eq. 5.11 (Drenth, 1994).

$$\rho(x,y,z) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F(hkl) \exp[-2\pi i(hx + ky + lz)] \quad \text{Eq. 5.11}$$

Os tipos de mapas de densidade eletrônica, normalmente utilizados durante a resolução de uma estrutura de proteína são:

(a) Fourier Diferença $|F_0 - F_c|$, usados para encontrar partes da estrutura que estão faltando no modelo, identificar regiões mal modeladas e procurar as moléculas de água. Através da análise destes mapas de densidade podemos detectar picos negativos que nos revelam a inclusão no modelo de átomos em posições erradas. Já os picos positivos correspondem a partes da estrutura ainda faltando no modelo. Densidades positivas e negativas lado a lado é um indicio da necessidade de uma remodelagem da região.

$$\Delta\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l (|F_0| - |F_c|) \exp(i\alpha_{\text{calc}}) \exp[-2\pi(hx + ky + lz)]$$

(b) Mapa $2|F_o| - |F_c|$. Estes mapas são os mais usados para inspecionar o refinamento do modelo e para reconstruir o modelo dentro da densidade. É usado o fator 2 $|F_o|$ porque queremos ressaltar a diferença entre o modelo calculado e o modelo real da estrutura.

$$\Delta\rho(x,y,z) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} (2|F_o| - |F_c|) \exp(i\alpha_{calc}) \exp[-2\pi(hx, ky, lz)]$$

A qualidade do mapa $2|F_o| - |F_c|$ depende da qualidade das fases e variações de $2|F_o| - |F_c|$, tais como $3|F_o| - 2|F_c|$ e $5|F_o| - 4|F_c|$ são algumas vezes usadas. Estas outras alternativas ressaltam mais a diferença entre $|F_{obs}|$ e $|F_{calc}|$ e dão mais ênfase às informações que vêm de $|F_{obs}|$, uma vez que as fases calculadas do modelo, ϕ_{calc} tendenciam o cálculo de $|F_{calc}|$.

(c) Mapas OMIT

$$\Delta\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l (|F_{obs}| - |F_{calc}|) \exp(i\alpha_{calc}) \exp[-2\pi(hx + ky + lz)]$$

onde os F_{calc} são fatores de estrutura parciais, uma vez que alguns fragmentos do modelo não entram neste cálculo. Este tipo de mapa de densidade eletrônica é comumente usado quando alguma região do mapa está pobremente definida; também pode ser usada com $2|F_o| - |F_c|$. A vantagem é que em omitir regiões duvidosas do modelo estas posições atômicas não entrando no cálculo das fases e portanto, os F_{calc} não serão tendenciosos para uma conformação possivelmente incorreto. Supondo que o restante da estrutura seja suficientemente próxima da estrutura verdadeira, a sua contribuição para o cálculo das fases deve ser suficiente para reproduzir densidade eletrônica na região de interesse permitindo uma reinterpretação da estrutura sem influência da conformação do antigo modelo.

Para a realização de todos os mapas, usamos FFT (Fast Fourier Transform) do pacote de programas CCP4 que faz uma transformação de Fourier rápida e/ou o pacote de programas X-PLOR.

5.4 Avaliação da Qualidade do modelo

5.4.1 Refinamento da BthTX-I

O modelo inicial, obtido a partir da resolução da estrutura pelo método de Substituição Molecular, contém dois monômeros na unidade assimétrica ainda com a sequência de aminoácidos da estrutura da *Agkistrodon piscivorus piscivorus* (Holland *et al.*, 1990). Os primeiros mapas de densidade eletrônica $2F_{\text{obs}} - F_{\text{calc}}$ para o dímero foram analisados utilizando o programa “O” (Jones; Kjeldgaard, 1993). Uma análise detalhada para cada resíduo foi realizada e sempre que possível modificou-se o modelo, onde as coordenadas eram ajustadas, da melhor maneira possível à densidade eletrônica, e em seguida, os resíduos diferentes foram trocados, baseado na sequência da BthTX-I.

Foi realizado um refinamento de corpo rígido do modelo inicial e depois um refinamento por dinâmica molecular, seguido do refinamento dos fatores de temperatura B isotrópicos. No refinamento de corpo rígido os dois monômeros da BthTX-I foram tratados independentemente. Neste ponto do refinamento, o fator de discordância R estava em 37.1%. Em seguida, foram visualizados os mapas, com o programa “O” (Jones; Kjeldgaard, 1993) e foram feitas as modificações necessárias dos resíduos que eram diferentes e alguns ajustes da cadeia no mapa de densidade eletrônica. O refinamento por dinâmica molecular com temperatura controlada (simulated annealing) apresenta uma fase de “preparação”, onde os maus contatos da estrutura são abrandados; depois, iniciou-se o refinamento a 3000K e o resfriamento do sistema foi feito em passos de 50K até a temperatura ambiente de 300K. Nesta etapa, o fator R caiu para 25.6%. Nesta época ainda não trabalhávamos com o fator R_{free} . Com a etapa de refinamento posicional e dos fatores B's isotrópicos o fator R decaiu para 22.9%. Novos mapas de densidade foram relizados e o processo de ajuste das cadeias na densidade eletrônica continuado. Este processo de refinamento foi realizado em vários ciclos sendo incluído agora o fator R_{free} ($R = 22.1\%$ e $R_{\text{free}} = 32.2\%$). Com os dados processados por MOSFLM estes ciclos de refinamento conduziram

a fatores $R = 21.9\%$ e $R_{\text{free}} = 31.5$. Foram realizados ainda vários ciclos de refinamento que conduziram a um $R = 19.5\%$ e $R_{\text{free}} = 28.2\%$. As moléculas de água foram colocadas com o programa 'Pickwater' do pacote de programas CCP4, que lista os picos mais intensos encontrados no mapa de Fourier diferença. Foram incluídas 114 moléculas de água, onde só foram aceitas águas que faziam pelo menos uma ligação de hidrogênio e fatores B's inferiores a $50 \text{ \AA}^2$. Ao final desta etapa o fator R era de 18.7% e o fator R_{free} de 27.4% .

A análise dos mapas de densidade eletrônica durante as primeiras etapas de refinamento indicou que a densidade eletrônica apresentava-se com diferenças significativas nos loops de ligação do Ca^{+2} resíduos 28 ao 33 e de forma mais acentuada na região do C-terminal (resíduos 115 ao 126).

Devido a estas diferenças não foi feito o refinamento utilizando simetria não cristalográfica (NCS) e portanto, os dois monômeros foram sempre tratados independentemente. A maior dificuldade durante os vários ciclos de refinamento foi a interpretação do loop C-terminal, Fig.5.1(a). Para resolver este problema foram realizados vários mapas omit (omitindo um, dois ou mais resíduos de cada vez) seguidos de vários ciclos de refinamento e desta forma podemos reconstruir lentamente toda a densidade eletrônica do C-terminal, como pode ser observado na Fig. 5.1 (b).

Fig. 5.1: Densidade eletrônica do C-terminal (resíduo 124 e 125): (a) antes de realizado o mapa omit, apresentando um mapa de densidade pobremente defenido.

Fig. 5.1: Densidade eletrônica do C-terminal (resíduo 124 e 125): (b) mapa atual depois de realizados vários mapas omit e ciclos de refinamento.

A tabela 5.1 mostra que tanto o fator R quanto o R_{free} decaíram substancialmente durante os estágios do refinamento. O comportamento do R_{free} indica que não houve ‘overfitting’ dos dados durante o processo de refinamento.

Etapas do Refinamento	Fator R (%)	Fator R_{free} (%)
Corpo rígido	37.1	-
SA	25.6	-
Posição/ Fatores B's	22.9	-
Vários ciclos	22.1	32.2
MOSFLM	21.9	31.5
Vários ciclos	19.5	28.2
Colocação da águas	18.7	27.4

Tabela 5.1: Dados sobre o refinamento e os fatores R e R_{free} nas etapas relevantes de cada ciclo.

Na Fig. 5.2 podemos avaliar visualmente a qualidade do mapa final $2F_{\text{obs}} - F_{\text{calc}}$ dos resíduos que compõem o “loop de ligação do Ca^{+2} ” no monômero 1 e 2 respectivamente.

Fig. 5.2: Qualidade do mapa de densidade eletrônica dos resíduos 28 a 33 na BthTX-I. Esta região faz parte do sítio de ligação de Ca^{+2} nas PLA_2 tipo D49, no monômero 1 e 2 da BthTX-I.

Na figura 5.3 é possível analisar-se a superposição dos dois monômeros da BthTX-I. O modelo final, contém dois monômeros com 121 aminoácidos cada um, composto de 1906 átomos (sem átomos de hidrogênio) e 114 moléculas de água. Este modelo final foi refinado a 2.1 Å de resolução e é decorrente dos dados coletados no sincrotron de Daresbury na Inglaterra e processados utilizando o programa MOSFLM. É interessante, observar nesta figura as duas regiões citadas anteriormente que apresentam maiores diferenças, particularmente a região do C-terminal.

Fig. 5.3: Superposição dos dois monômeros da BthTX-I. Nesta figura, pode ser observada as duas regiões essencialmente diferentes: a região do “loop de ligação de Ca^{+2} ” e particularmente a região do C-terminal.

5.1.2: Estereoquímica do Modelo

Os valores finais do fator R e R_{free} (para 10% dos dados) foram 0.187 e 0.274 respectivamente para dados a 2.1 Å de resolução. A análise dos desvios da idealidade para comprimentos de ligação, ângulos de ligação e ângulos diédricos foram obtidos pelo programa XPLOR. Os desvios encontrados para o modelo da BthTX-I fo de 0.008 Å para os comprimentos de ligação, 1.49° para os ângulos entre as ligações, 22.53° para os ângulos diédricos e 1.27° para os ângulos impróprios, o que está dentro dos valores esperados.

As restrições para os ângulos torcionais ou diédricos que determinam a conformação de uma proteína devem ser também analisados. A cadeia principal da proteína é caracterizada pelos ângulos diedros ϕ (formado pelos átomos C_{i-1} , N_i , $C_{\alpha i}$ e C_i), ψ (formado pelos átomos N_i , $C_{\alpha i}$, C_i e N_{i+1}) e ω (formado pelos átomos $C_{\alpha i}$, C_i , N_{i+1} e $C_{\alpha i+1}$).

A análise estereoquímica completa foi feita com o programa PROCHECK (Laskowski et al., 1994). O programa compara os parâmetros do modelo estrutural com parâmetros equivalentes de estruturas resolvidas a alta resolução. Na Fig. 5.4 está apresentado o diagrama de Ramachandran para o modelo final da BthTX-I.

Fig. 5.4: Diagrama de Ramachandran do modelo da BthTX-I para os dois monômeros.

A análise indica que 88.8% dos resíduos estão em regiões favoráveis, 10.2% em regiões permitidas, 0.5% nas regiões generosamente permitidas e 0.5% nas regiões não permitidas. O resíduo Glu⁸⁷ do primeiro monômero está numa região não permitida. Esta região do loop apresenta uma densidade pobremente definida e, devido a este fato, foram realizados mapas “omit” mas não se obteve melhor definição. Já no segundo monômero, o Glu²⁸⁷ não apresentou o mesmo problema. Em geral, a estereoquímica da estrutura está muito boa e de acordo com o que se espera para uma estrutura a 2.1 Å.

5.1.3: Análise conformacional do modelo

A estrutura final também foi examinada por vários outros critérios de qualidade. Entre eles **OOPS** (Kleywegt; Jones, (1994 e 1995) para identificar conformações incomuns, **QUALITY** (Vriend; Sande, 1993) analisar contatos atômicos, **VERIFY-3D**(Lüthy et al., 1991) analisa ambientes químicos dos resíduos e **PROSA** (Sippl, 1993) energia local.

Na Fig. 5.5 podemos analisar a distribuição dos índices de qualidade dos valores para cada resíduo da BthTX-I, atribuídos pelo programa **QUALITY**. Valores abaixo de -5.0 devem ser reexaminados. A média do índice de qualidade das estruturas do banco de dados é de -0.59 com um desvio padrão de 0.42. Para o dímero da BthTX-I o valor encontrado foi de -0.94, o que significa que a estrutura, na média não apresenta contatos atômicos incomuns, um indicio da sua qualidade.

Fig. 5.5: Distribuição dos índices de qualidade dos valores para cada resíduo da BthTX-I, atribuídos pelo programa Quality.

Na Fig. 5.6 é mostrado o diagrama dos valores médios de qualidade de ambientes químicos atribuídos pelo programa **Verify-3D** dentro de uma janela de 21 resíduos, para a BthTX-I. O programa avalia o modelo de duas maneiras: a primeira calcula um índice para o modelo todo, que depende do comprimento da sequência. Este valor pode ser comparado com o valor ideal para um dado comprimento da cadeia polipeptídica (Lüthy, 1991). Para o dímero da BthTX-I este valor foi de 99.17 enquanto que o valor ideal foi de 110.26. A segunda análise, indica possíveis regiões mal construídas. As regiões que estiverem abaixo de zero representam regiões provavelmente mal construídas pois apresentam ambientes químicos incomuns para os resíduos (por exemplo cargas num ambiente hidrofóbico), e precisam ser reexaminados. Novamente a análise feita com esta metodologia comprova a qualidade da estrutura apresentada.

Fig. 5.6: Diagrama dos valores atribuídos com o programa Verify-3D para a BthTX-I.

As análises realizadas pelo programa **OOPS** para conformações incomuns não apresentaram resultados fora da normalidade.

A análise final do modelo feita pelo programa **PROSA** está representada na Fig. 5.7. Duas janelas de tamanhos diferentes foram usadas para comparação, uma de 10 e a outra de 50 resíduos. Em ambos os casos a energia é plotada em função da sequência. Esta análise da distribuição de energia pode ser útil em identificar regiões de estrutura duvidosa (alta energia local). No caso da BthTx-I os gráficos são típicos de uma estrutura bem determinada, apresentando poucos picos positivos no gráfico correspondente á uma janela de 10 e nenhum pico positivo para o caso da janela de 50 resíduos (Sippl, 1993).

Fig. 5.7: Gráficos de energia calculados pelo programa PROSA II para cada monômero da BthTX-I. A diferença entre eles também é apresentada. Nesta figura, obj1 se refere ao monômero 1 e obj2 ao monômero 2. Duas janelas de tamanhos diferentes foram usados para comparação. Linha fina, uma janela de 10 e a outra de 50 resíduos.

No próximo capítulo faremos uma análise e discussão da estrutura da BthTX-I.

5.5 Referências Bibliográficas

- AGARWAL, R. *Acta Cryst.* **A34**, 791-809 (1978).
- BRÜNGER, A. T., KURIYAN, J. & KARPLUS, M. - *Science* **235**, 458-460 (1987).
- BRÜNGER, A. T. - X-PLOR , *A System for X-ray Crystallography and NMR* - Yale University Press, New Haven/ London (1992).
- BRÜNGER, A. T., KARPLUS, M. & PETSKE, G. A. - *Acta Cryst.* **A45**, 50-61 (1989).
- BRÜNGER, A. T. & KRUKOWSKI, A. - *Acta Cryst.* **A46**, 585-593 (1990).
- BRÜNGER, A. T. *Nature* **355**, 472-474 (1992).
- CCP4 Programa Suite - User Documentation, Daresbury Laboratory (1984).
- DRENTH, J. - *Principles of Proteins X-ray Crystallography*, Springer-Verlag New York (1994).
- DIAMAND, R. *Acta Cryst.* **A27**, 436-452 (1971).
- HANNEF, I., MOSS, D., STANFORD, M. & BORKAKOTI, N. *Acta Cryst.* **A41**, 426-433 (1985).
- JACK, A. & LEVITT, M. *Acta Cryst.* **A34**, 931-935 (1978).
- JONES, T. A. & KJELDGAARD, M. - "O" *The Manual* - 5.9 version (1993).
- KLEYWEGT, G. J. & JONES, T. A. - *Structure* **3**, 6, 535-540 (1995).
- KURIYAN, J., BRÜNGER, A. T., KARPLUS, M. & HENDRICKSON, W. A. - *Acta Cryst.* **A45**, 396-409 (1989).
- KONNERT, J. & HENDRICKSON, W. *Acta Cryst.* **A36**, 344-350 (1980).
- LASKOWSKI, R. A., MACARTHUR, M. W., SMITH, D. K., JONES, D.T., HUTCHINSON, E.G., MORRIS, A. L., NAYLOR, D., MOSS, D. S. & THORNTON, J. M. - *PROCHECK v.3.0 - Program to check the stereochem. quality of Protein Structures* - Operating instructions (1994),

- LÜTHY, R., McLACHLAN, A. D. & EISENBERG, D., *Proteins* **10**, 229 - 239 (1991).
- POWELL, M. J. D. *Mathematical Programming* **12**, 241 (1977).
- SIPPL, M. J. - *Proteins: Structure, function and genetics* **17**, 355-362 (1993).
- SUSSMANN, J. , HOLBROOK, S., CHURCH, G. & KIM, S. *Acta Cryst.* **A33**, 800-804 (1977).
- TRONRUD, D., TEN E. L. & MATTHEWS, B. *Acta Cryst.* **A43**, 489-501 (1987).
- VERLET, L. *Phys. Rev* **159**, 98-103 (1967).
- VRIEND & SANDEL *J. Appl. Cryst.* **26**, 47-60 (1993).

CAPÍTULO 6

O MODELO DA BthTX-I: ANÁLISE E DISCUSSÃO DOS RESULTADOS

6.1 Descrição da Estrutura

Na Fig. 6.1 encontra-se o esquema da estrutura secundária da BthTX-I, assim como as respectivas ligações de hidrogênio, fornecidas pelo programa HERA (Hutchinson & Thornton, 1990). Na Fig. 6.2 vemos que a estrutura é composta basicamente de 4 hélices α (numeradas de 1 a 4) sendo as hélices 3 e 4 as principais. Entre a hélice 3 e 4 existe uma região de folha β composta de duas fitas A e B, formando uma estrutura conhecida como asa- β .

Fig. 6.1: Esquema da estrutura secundária da estrutura da BthTX-I, fornecida pelo programa HERA (monômero 1 e 2)

Fig. 6.2: Estrutura tridimensional da BthTX-I, mostrando as quatro hélices α (numeradas de 1 a 4), as duas folhas β formando a região conhecida por asa- β .

Na tabela 6.1 podemos resumir estes resultados para os dois monômeros da estrutura de BthTX-I e identificar algumas regiões diferentes. O monômero 2 possui uma pequena região de folha β antiparalela entre as hélices 2 e 3 e duas hélices α de uma volta na região C-terminal. Estas diferenças podem ser atribuídas ao empacotamento diferenciado para as duas moléculas cristalograficamente independentes, Fig. 6.3.

Fig. 6.3: As duas moléculas da BthTX-I, usadas durante todo o processo de refinamento.

BthTX-I	Hélices α	Folhas β
Monômero 1	(1) 1S 14G	
	(2) 17N 23G	
	(3) 39D 53K	
	(4) 89N 109N	(A) 74S 78K (B) 81T 86G
Monômero 2	(1) 201S 214G	
	(2) 217N 223G	
	(3) 239D 254K	224A 225Y 229C 231V
	(4) 289N 309N	(A) 274S 278K (B) 281T 286G
	310L 313Y	
	314N 317Y	

Tabela 6.1: Resumo das estruturas secundárias dos dois monômeros da BthTX-I.

6.1.1 Superposição dos dois monômeros

Como já foi descrito anteriormente, os dois monômeros da BthTX-I apresentam diferenças significativas na região do loop característico de ligação do Ca^{+2} das PLA_2 e no C-terminal. Estas regiões são variáveis em outras estruturas de PLA_2 e frequentemente apresentam fatores de temperatura B's altos (Arni; Ward, 1995). Na Fig.6.4 estão apresentadas as distâncias entre os $\text{C}\alpha$ equivalentes dos dois monômeros e como pode ser observado na mesma figura estas diferenças ocorrem exatamente nas regiões anteriormente descritas. Estas diferenças estruturais entre os dois monômeros podem ser justificadas exatamente devido aos diferentes contatos cristalinos entre as subunidades.

Fig. 6.4: Principais diferenças entre os $\text{C}\alpha$ nos dois monômeros da BthTX-I.

Em parte, o menor conteúdo de estrutura secundária no monômero 1 é compensado pelo maior número de voltas β (ver tabela 6.2). Em geral as voltas apresentam a mesma conformação nos dois monômeros com exceção dos resíduos 67-71 que fazem parte do loop entre a hélice 3 e a asa- β , e formam duas voltas consecutivas. Os tipos de volta são trocados nos dois monômeros.

Monômero 1		Monômero 2	
Resíduos	Tipo voltas β	Resíduos	Tipo voltas β
22 - 25	IV	222 - 225	IV
25 - 28	II'	225 - 228	II'
26 - 29	IV	226 - 229	IV
55 - 61	IV	255 - 261	IV
67 - 70	IV	267 - 270	I
68 - 71	I	268 - 271	IV
77 - 80	IV		
78 - 81	IV		
117 - 120	IV		

Tabela 6.2: Classificação das voltas β para cada um dos monômeros da BThTX-I, segundo o programa PROMOTIF (Hutchinson; Thornton, 1996)

A estabilidade da estrutura 3D das PLA_2 se deve em boa parte á presença das sete pontes dissulfeto. Na tabela 6.3 encontra-se a classificação das pontes dissulfeto de ambos os monômeros, assim como as distâncias entre S-S , $C\alpha$ - $C\alpha$ e o ângulo χ_3 .

Monômero 1				
P. dissulfeto	S-S (Å)	Cα-Cα (Å)	$\chi_3(^{\circ})$	Classificação
Cys27 - Cys126	2.02	6.1	-99.5	Espiral
Cys29 - Cys45	2.03	5.1	91.6	Gancho
Cys44 - Cys105	2.03	5.8	-78.3	Espiral
Cys50 - Cys133	2.01	5.6	-76.3	Espiral
Cys51 - Cys98	2.02	5.9	-85.1	Espiral
Cys61 - Cys91	2.03	6.0	-78.3	Espiral
Cys84 - Cys96	2.03	6.0	-87.3	Espiral
Monômero 2				
P. dissulfeto	S-S (Å)	Cα-Cα (Å)	$\chi_3(^{\circ})$	Classificação
Cys27 - Cys126	2.02	6.0	-96.9	Espiral
Cys29 - Cys45	2.02	5.1	92.3	Gancho
Cys44 - Cys105	2.02	5.8	-76.0	Espiral
Cys50 - Cys133	2.03	5.3	-75.8	Espiral
Cys51 - Cys98	2.01	5.8	-80.6	Espiral
Cys61 - Cys91	2.03	6.1	-84.7	Espiral
Cys84 - Cys96	2.03	6.0	-89.5	Espiral

Tabela 6.3: Pontes dissulfeto do modelo 1 da BthTX-I (classificação, distâncias e ângulos), fornecidas pelo programa PROMOTIF.

A conformação de cada ponte é igual nos dois monômeros sendo a maioria do tipo espiral de mão esquerda (Richardson & Richardson, 1989). A densidade eletrônica é bem definida para todas as pontes e os parâmetros apresentados na tabela 6.3 estão dentro dos limites normais, sendo que a única ponte em conformação gancho (χ_3 positivo) também apresenta a menor distância Cα-Cα.

6.1.2 Análise dos Fatores de Temperatura B

A observação dos fatores de vibração térmica, nos permite identificar no modelo de uma estrutura, as regiões de maior desordem estática e/ou dinâmica inclusive trechos da molécula onde o modelo é duvidoso.

Valores de **B** aceitáveis para átomos da cadeia principal encontram-se na faixa de 5 a 35 Å² e para a cadeia lateral em torno de 5 a 60 Å². Valores acima de 60 Å², em geral, indicam desordem ou localização errada do resíduo. Os valores médios para a cadeia principal e a cadeia lateral foram 21.76 Å² e 25.63 Å² para o monômero 1 e 27.58 Å² e 30.75 Å² para o monômero 2.

Na Fig. 6.5 (a) e (b) encontramos o diagrama dos fatores de temperatura para a cadeia principal e lateral da BthTX-I, respectivamente.

(a) cadeia principal

(b) cadeia lateral

Fig. 6.5 : Diagrama dos fatores de temperatura para a cadeia principal e lateral da BthTX-I (dois monômeros).

Como pode ser observado, o monômero 2 tem fatores de temperatura um pouco mais elevados que o monômero 1, que pode ser atribuído a diferenças no empacotamento das moléculas na rede cristalina.

Normalmente, as regiões mais flexíveis e expostas ao solvente tais como Lys apresentam uma maior mobilidade estrutural e por isso possuem valores de B's mais altos (C-terminal).

6.1.3 Colocação das Moléculas de água

Para a colocação das moléculas das 114 moléculas de água foi utilizada a rotina do **Pickwater** do pacote de programas CCP4 e foram calculados mapas de Fourier diferença $F_{obs} - F_{calc}$. As moléculas de água selecionadas apresentavam distâncias entre 2.5 e 3.5 Å e densidade bem definida e faziam pelo menos uma ligação de Hidrogênio. Algumas destas águas localizadas são importantes localizando-se no sítio ativo e serão posteriormente apresentadas, Fig 6.12. As moléculas de água que apresentavam valores de B's acima de 50 Å² foram eliminadas.

6.2 Questão da Dimerização

6.2.1 Apresentação do Resultados das Eletroforeses em Gel de 15% de Poliacrilamida

Como foi apresentado no capítulo de introdução, algumas PLA₂ de baixo peso molecular (enzimas pancreáticas) são monoméricas em solução, enquanto outras formam dímeros (PLA₂ de *Bothrops asper*). Formas triméricas foram descritas em venenos e postula-se que este estado oligomérico não tem participação no processo catalítico, mas possa ocorrer *in vivo* como um mecanismo para manter a proteína em uma forma inativa nas glândulas de venenos de serpentes,

onde são encontradas em altas concentrações. Estas enzimas permaneceriam então neste estado “inativo” até sua atividade ser requerida (Mebs; Ownby, 1990).

Para avaliar esta questão da dimerização foram feitas várias eletroforeses em SDS-Gel de 15% de Poliacrilamida da BthTX-I com e sem β - mercaptoetanol. As amostras foram fervidas durante 5 minutos na presença de SDS com e sem β -MSH antes de aplicadas no gel de empinhamento.

Na Fig. 6.6 podemos observar que com β - mercaptoetanol a proteína se apresenta no estado monomérico (P.M. $\cong$ 14.000) e sem β - mercaptoetanol aparece um rastro entre dois pesos moleculares aqui, sugerindo um equilíbrio entre o dímero e o monômero.

Já que a BthTX-I é uma proteína bastante estável, devido á presença das sete pontes dissulfeto é de se esperar uma certa dificuldade de desenovelamento total da molécula (mesmo na presença de SDS) sem β - mercaptoetanol, indicando a existência mesmo de um dímero.

Fig. 6.6: Eletroforese em gel de 15% de poliacrilamida da BthTX-I, com e sem β - mercaptoetanol, mostrando um rastro entre um dímero e monômero. As raias correspondem respectivamente: 1- Padrão ; 2- BthTX-I c/ β -MSH ; 3- BthTX-I c/ β -MSH ; 4-5 BthTX-I s/ β -MSH ; 6- Padrão.

O dímero escolhido para definir a unidade assimétrica daqui para frente chamado de modelo 1, Fig. 6.3 não apresentou nenhum contato atômico e tinha orientações relativas entre os monômeros bastante diferentes da *Bothrops asper* que também é dimérica e possui alto grau de identidade sequencial. A avaliação do empacotamento cristalino revelou a presença de um dímero parecido daqui para a frente chamado de modelo 2, com o dímero da *Bothrops asper* e resolvemos portanto, avaliar os contatos neste segundo caso.

6.2.2 Cálculo da Superfície de Interação entre Moléculas

A superposição dos C α do monômero 1 da BthTX-I com o monômero A da estrutura dimérica da *Bothrops asper* pode ser visto na Fig. 6.7. Da análise desta figura podemos observar que a estrutura dimérica da BthTx-I está numa orientação ligeiramente diferente da estrutura dimérica da *B. asper* (Arni; Ward, 1995). Da superposição dos dois monômeros vemos que não existem diferenças significativas.

Fig. 6.7: Superposição dos C α do monômero 1 da BthTX-I com o monômero A da estrutura dimérica da *B. asper*, mostrando uma orientação ligeiramente diferente entre eles.

Para avaliar esta forma dimérica e compará-la com outras estruturas diméricas, utilizamos um programa desenvolvido por Laurence & Colman (1993) para calcular o parâmetro S_c que analisa a complementaridade na interface de duas moléculas proteicas..

O fator de correlação de forma é definido, como mostrado na Fig. 6.8.

Fig. 6.8: Definição do fator de correlação de forma S_c .

A e B são duas moléculas que interagem. P_A é a porção da molécula A que interage com a molécula B e P_B a superfície de B que interage com A. Seja x_A um ponto de P_A com n_A sendo um vetor normal saindo da superfície de P_A , x'_A o ponto de P_B mais próximo de x_A e n'_A seu vetor normal entrando na superfície P_B .

A função escalar

$$S^{A \rightarrow B}(x_A) = (n_A \cdot n'_A) \exp[-w \left(\|x_A - x'_A\| \right)^2]$$

pode ser definida para cada ponto em P_A , onde w é uma constante, colocada neste estudo como $0.5 \text{ \AA}^2$. De modo similar definimos

$$S^{B \rightarrow A}(x_B) = (n_B \cdot n'_B) \exp[-w \left(\|x_B - x'_B\| \right)^2]$$

O fator de correlação de forma S_c é então definido como na Eq. 6.1.

$$S_c = \frac{\left(\left\{ S^{A \rightarrow B} \right\} + \left\{ S^{B \rightarrow A} \right\} \right)}{2} \quad \text{Eq. 6.1}$$

Para proteínas oligoméricas e superfícies de interação proteína/inibidor, S_c varia entre 0.70 e 0.76, enquanto que para superfícies de interação entre anticorpo/antígeno, S_c varia de 0.64 a 0.68 (Lawrence; Colman, 1993). Isto significa que anticorpo/antígeno têm uma forma significativamente mais pobre de complementaridade do que proteínas oligoméricas ou mesmo complexos proteína/inibidor. É de se esperar que para simples contatos cristalinos entre duas moléculas o valor de S_c não deve apresentar valores elevados.

Para avaliar a questão da dimerização da BthTX-I foram calculados os valores de S_c e a área total interrada usando o programa SPACE para todas as estruturas diméricas conhecidas (incluindo dímeros não funcionais) e também alguns obtidos apenas a partir da aplicação das operações de simetria, ou seja contatos cristalinos que não devem ter nenhum sentido fisiológico.

O valor da área total foi também calculada utilizando o programa DSSP e o programa WHATIF usando o método de Lee & Richards com uma sonda de raio = 1.4 Å). Neste caso, apresentaremos apenas os valores obtidos com WHATIF, uma vez que os dois cálculos não apresentaram diferenças significativas, apesar de diferenças no algoritmo.

Na tabela 6.4 estão apresentados os valores encontrados para esta análise.

	Estrutura	Área total enterrada WHATIF (Å ²)	S _c
(2)	Modelo 2 (Bthtx-I)	1020	0.59
(2)	B. asper	750	0.57
(2)	Crotalus atrox	2523	0.68
(2)	Crotalus-sym	635	0.65
(1)	Bovina-sym	154	0.20
(1)	Humana-sym	562	0.30
(1)	Porco-sym	1194	0.60
(2)	Atra-sym	1081	0.65
(1)	App-sym	459	0.60
(2)	Humana	604	0.75
(2)	Porco-sym	1229	0.67

Tabela 6.4 Resultados da análise do fator de correlação de forma gerada pelo programa S_c . Também são apresentadas as áreas totais (programa WHATIF).

Estudos de contatos entre subunidades em proteínas diméricas apresentam áreas tipicamente acima de 2000 Å² enterrados. O menor valor de uma interface destas descrito num artigo recente foi de 1340 Å² no caso da superóxido dismutase (janil et al., 1988). Tanto no dímero (modelo 2) da BthTX-I quanto na toxina K49 de *Bothrops asper* (miotoxina II), os valores equivalentes são consideravelmente menores. O grau de complementaridade entre as duas superfícies em contato no dímero, (avaliado pelo valor de S_c) também está abaixo daquilo esperado para uma molécula oligométrica (≈ 0.58 comparado com ≈ 0.70). A tabela 6.4 mostra que em muitos casos um simples contato cristalino apresenta maior complementaridade e área enterrada do que os dímeros propostos para a BthTX-I e *B. asper*.

Estes resultados quando comparados com os pesos moleculares determinados por eletroforese são, no mínimo, intrigantes, e levanta a questão da existência ou não do dímero em solução e a sua relevância na atividade biológica. Os resultados de Ward *et al.* (comunicação pessoal), sugerem que mesmo na presença de β - MSH e SDS o dímero de BthTX-I permanece intacto em eletroforese, mas os nossos resultados são discordantes desta afirmação. A explicação disso pode estar relacionada com diferenças no preparo das amostras.

Certamente, a existência da mesma interface dimérica nas moléculas de duas espécies diferentes é um forte indicio da sua relevância biológica. Se existisse em solução ambos os sítios ativos (veja no próximo item) estariam orientados para o mesmo lado da molécula permitindo interação simultânea com a membrana. Isto não é o caso para o dímero inicialmente descrito para a toxina D49 de *Crotalus atrox*. Na Fig. 6.9 estão representadas estas estruturas diméricas.

Fig 6.9 : Estruturas diméricas. Nesta figura estão representados as duas moléculas na unidade assimétrica da BthTX-I (modelo 1); a estrutura dimérica da BthTX-I tipo *B. asper* (modelo 2); a estrutura dimérica da *B. asper* e a estrutura dimérica da *Crotalus atrox*.

A aparentemente pobre complementaridade superficial entre as subunidades do dímero e a baixa área de contato, Fig. 6.10 permanecem peculiaridades do dímero se existisse em solução. Talvez o dímero não seja tão estável quanto se imaginava ou se forma apenas na hora da interação com a membrana. Na Fig. 6.11 estão representadas as interações entre os dois monômeros na estrutura dimérica da BthTX-I. A interface é bastante hidrofílica envolvendo uma ponte salina entre E12 de uma subunidade e K80 da outra, e pontes de hidrogênio entre os resíduos 77 e 11, 80 e 80 e 14 e 7. Contactos hidrofóbicos são feitos pelos resíduos Trp⁷⁷ e Val¹⁰ com seus parceiros simétricos.

Fig. 6.10: Área de contato entre as subunidades da estrutura dimérica da BthTX-I (em vermelho) calculada pelo programa GRASP (Nicholls, 1992).

Fig. 6.11: Resíduos envolvidos na interação entre os monômeros na estrutura dimérica da BThTX-I. Também são mostradas as distâncias em Å.

6.3 Análise do Canal Hidrofóbico

Como foi descrito anteriormente, algumas estruturas de fosfolipases A_2 , entre elas a estrutura da (1) *Naja naja atra* (White *et al.*, 1990); (2) da PLA_2 do veneno de abelha (Scott *et al.*, 1990) e (3) a estrutura da $hnps-PLA_2$ (human non-pancreatic secretory PLA_2) (Scott *et al.*, 1991; Wery *et al.*, 1991) foram resolvidas com um análogo do estado de transição, $diC_8(2Ph)PE$. Em todas estas estruturas foi comparada a região de ligação do análogo com a proteína, que revela uma conservação de resíduos hidrofóbicos que formam um canal de ligação. Na tabela 6.5 estão destacados os resíduos do canal de ligação do $diC_8(2Ph)PE$ na estrutura humana e os resíduos conservados na BthTX-I, resultantes da superposição da estrutura da molécula humana (regiões mais conservadas).

Resíduos do canal hidrofóbico (hnps-PLA ₂)	Resíduos conservados na BthTX-I
Leu ² Phe ⁵ His ⁶ Ile ⁹ Ala ¹⁷ Ala ¹⁸ Tyr ²¹ Gly ²² Phe ²³ Cys ²⁸ Gly ²⁹ Val ³⁰ Cys ⁴⁴ Tyr ⁵¹ Lys ⁶² Phe ⁹⁸	Leu ² Leu ⁵ Gly ⁶ Ile ⁹ Pro ¹⁸ Ala ¹⁹ Tyr ²² Gly ²³ Ala ²⁴ Cys ²⁹ Gly ³⁰ Val ³¹ Cys ⁴⁵ Tyr ⁵² Lys ⁶⁹ Leu ¹⁰⁶

Tabela 6.5: Resíduos do canal hidrofóbico da estrutura hnps-PLA₂ e os resíduos da BthTx-I, (sobrepondo as regiões conservadas).

6.4 “O Sítio Ativo”

Fig. 6.12 mostra a região do sítio ativo do segundo monômero da BthTX-I. Nesta figura estão também representadas as duas moléculas de água diretamente envolvidas no mecanismo catalítico de uma típica PLA₂, sendo uma delas a água catalítica (mais próxima da His⁴⁸) e a outra fazendo a ligação do sítio ativo com o N-terminal, Asp⁹⁹, Pro⁶⁸ e Tyr⁵². As respectivas distâncias em Å desta malha de ligações também estão representadas. No primeiro monômero a situação é idêntica.

Fig. 6.12: Região do sítio ativo do 2^o. monômero da BthTX-I. Duas moléculas de água e as respectivas distâncias também se encontram assinaladas.

Como pode ser observado os resíduos do sítio ativo são altamente conservados, quando comparados com outras PLA₂, principalmente a Tyr⁵², Tyr⁷³, Asp⁹⁹ e His⁴⁸. Esta alta conservação de sequência e estrutura levanta a questão da falta de atividade catalítica que será tratado mais adiante.

6.5 Loop de Ligação do Ca⁺² (loop típico de ligação do primeiro íon Ca⁺²).

Na Fig. 6.13 é realizada uma comparação do “loop de ligação do Ca⁺²” (resíduos 28 ao 33) de algumas das estruturas de fosfolipases A₂. Entre elas, (1) PLA₂ pancreática bovina- D49; (2) *Agkistrodon p.p.* - K49; (3) os dois monômeros da BthTX-I - K49 e (4) os dois monômeros da *Bothrops asper*- K49, apenas a primeira, sendo uma PLA₂ tipo D49, possui o íon Ca⁺² ligado.

Nas demais estruturas, incluindo as duas moléculas de BthTX-I a posição do Ca^{+2} está ocupado pelo N_ξ da K49.

Fig. 6.13: Comparação do loop de ligação de Ca^{+2} (resíduos 28 a 33) em algumas das estruturas de PLA_2 .

Como pode ser notado, observa-se uma grande variabilidade desta região. Nas PLA_2 o Ca^{+2} é coordenado pelas carbonilas dos resíduos Tyr^{28} , Gly^{30} , Gly^{32} , $\text{O}\delta_1$ e $\text{O}\delta_2$ do Asp^{49} e duas moléculas de água. Nas K49 o N_ξ da lisina forma pontes de hidrogênio com uma ou mais carbonilas dos resíduos 28, 30, 32 ou 33, como pode ser visto na figura. No caso da BthTX-I, na molécula 1 o N_ξ da lisina é coordenado pelas carbonilas da Tyr^{28} e Gly^{30} enquanto na molécula 2 é também coordenado pelas carbonilas da Tyr^{28} , Gly^{30} e adicionalmente Gly^{33} .

Sigler, sugeriu que nas K49 esta região da molécula seria bastante flexível. As diferenças no loop encontradas na BthTX-I são provenientes de contatos cristalinos que diferem nas duas moléculas.

Na estrutura da BthTX-I não foi encontrado um segundo íon Ca^{+2} como encontrado na resolução da estrutura da *Naja naja atra* (White *et al.*, 1990) onde os autores propõem que este segundo íon de cálcio é coordenado pelo O29, isto é, o oxigênio carbonil da ligação peptídica formada com o NH da G30. Este NH da G30 por sua vez interage com o oxiânion do intermediário tetraédrico e portanto, a posição deste segundo íon cálcio pode agir indiretamente na estabilização do mesmo. Esta hipótese foi levantada numa tentativa de explicar o verdadeiro mecanismo de estabilização da carga negativa do oxiânion em fosfolipases ativas uma vez que a carga do cálcio seria formalmente neutralizada pelo D49 e pelo sn-3 fosfato do fosfolípido, uma vez complexado. O segundo íon cálcio poderia polarizar a ligação peptídica entre os resíduos 29 e 30, deixando uma carga parcial no NH da G30 que interage diretamente com o oxiânion.

Foi sugerido também que uma Lys^{122} interagindo com o O29 através de uma molécula de água poderia agir da mesma maneira nas fosfolipases que não possuem este segundo íon Ca^{+2} . Este tipo de interação foi encontrada na PLA_2 pancreática bovina e também no monômero 1 da BthTX-I, (Fig 6.14). No segundo monômero a Lys^{322} está voltada para o solvente e a Tyr^{319} está voltada para o O229, (Fig. 6.15). É curiosa a presença desta interação na BthTX-I uma vez que ela é uma PLA_2 tipo K49 que não apresenta atividade catalítica. Scott *et al* propuseram um modelo para a complexação do fosfolípido com a PLA_2 K49 inativa de *Agkistrodon piscivorus piscivorus*, no qual a orientação da ligação peptídica 29-30 não permite a estabilização do oxiânion através do NH da G30. Sugeriram que a ausência desta interação poderia explicar a falta de atividade nestas toxinas. Estas observações estimularam um estudo de modelagem do complexo BthTX-I com um fosfolípido.

Fig. 6.14: Intereção da Lys¹²² com o O29 no monômero 1 da BthTX-I. Esta figura mostra também a qualidade do mapa na região C-terminal, que é diferente nos dois monômeros da BthTX-I.

Fig. 6.15: Interação da Tyr³¹⁹ com o O229. Esta figura mostra também a qualidade do mapa na região C-terminal, que é diferente nos dois monômeros da BthTX-I.

6.6 Modelo do fosfolípídeo Fosfatidil etanolamina na estrutura da BthTX-I

Como nenhuma estrutura de um complexo com uma K49 Pla₂ foi determinada até hoje, imaginamos que um complexo ligaria neste canal de uma maneira similar ao substrato de PLA₂ D49 e portanto, isto nos estimulou a fazer uma modelagem de um fosfolípídeo com a BthTX-I.

Para a realização deste modelo, inicialmente sobreposemos a estrutura da molécula humana com a BthTX-I, usando as regiões mais conservadas na família como um todo e identificados por Scott *et al.*, 1990. Em seguida, modificamos o análogo diC₈(2Ph)PE para produzir o fosfolípídeo Fosfatidil etanolamina, ao invés do intermediário tetraédrico. Ajustamos manualmente as conformações dos ácidos graxos e cadeia lateral da K69, para favorecer a interação entre a proteína e o fosfolípídeo. Finalmente uma minimização de energia foi feita usando o programa INSIGHT (User guide - Biosym Technologies, version 2.3.0).

Fig. 6.16 mostra uma figura esquemática das interações entre o fosfolípídeo modelado e a BthTX-I. Vários contatos hidrofóbicos são feitos com resíduos que participam do canal hidrofóbico e três ligações de hidrogênio estão presentes com a Lys⁴⁹, Lys⁶⁹ e o NH da Gly³⁰.

Fig. 6.16: Interações do fosfolípídeo Fosfatidil etanolamina (modelado) com o monômero 1 da BthTX-I , baseado no complexo diC₈(2Ph)PE da estrutura humana utilizando o programa LIGPLOT (McDonald & Thornton, 1994).

A estrutura cristalográfica da BthTX-I apresenta algumas diferenças importantes entre os dois monômeros na região do loop de ligação de Ca^{+2} , como descrito anteriormente. Na molécula 2 o N_ξ da K49 forma três ligações de hidrogênio com carbonilas do loop além de uma molécula de água. Em contrapartida a molécula 1 forma apenas duas ligações de hidrogênio com carbonilas dos resíduos 28 e 30. A terceira (resíduo 33) não se forma devido às diferenças na conformação do loop. As diferenças foram atribuídas aos contatos cristalinos e parece que a conformação representada pela segunda molécula seria a mais estável em solução na ausência do fosfolipídeo. No modelo do complexo, propomos que o fosfato do fosfolipídeo interage através de um dos seus oxigênios com o N_ξ da K49. Na molécula 1, esta interação pode facilmente acomodada, porém na molécula 2, para haver complexação necessitaria de uma mudança conformacional do loop.

Propomos, portanto, que a flexibilidade do loop de ligação do Ca^{+2} é fundamental para a ligação do fosfolipídeo. Em solução e inicialmente na conformação encontrada na molécula 2 a BthTX-I sofreria uma mudança de conformação do loop na hora da complexação, com o oxigênio do fosfato deslocando a carbonila do resíduo 33. Carbonilas 28 e 32 complementaríamos a coordenação do N_ξ da lisina 49. Sugerimos portanto, que as moléculas 1 e 2 da BthTX-I na forma cristalina aqui estudada poderiam servir como modelos para a forma complexada e não complexada respectivamente da toxina com o fosfolipídeo.

Ambas as moléculas possuem o grupo NH da G30 apontando para a carbonila da ligação ester na posição sn-2 do fosfolipídeo, (Fig 6.17) ao contrário daquilo observado para a forma cristalina 1 (CF1) descrito por Scott et al., 1992. Estes autores descrevem um modelo no qual a carbonila de 29 ao invés do NH da G30 estaria apontando para o sítio do oxiânion e atribuem a falta de atividade fosfolipídica a esta diferença. No caso não haveria como estabilizar o intermediário tetraédrico da reação. Porém os mesmos autores descrevem também uma segunda forma cristalina da toxina de *Agkistrodon p.p.* (CF2) onde a ligação peptídica está virada e portanto o NH da glicina 30 estaria apontando para o sítio de ligação do oxiânion, como numa fosfolipase ativa. Neste caso, é difícil explicar a falta de atividade catalítica usando o mesmo argumento. As coordenadas atômicas das duas formas cristalinas ainda não foram depositadas no banco de dados.

Fig. 6.17: Modelo da interação entre o fosfolípido e BthTX-I (monômero 1).

A Fig. 6.16 do nosso modelo demonstra a viabilidade da formação de uma ligação de hidrogênio entre o grupo carbonila na posição sn-2 do fosfolípido e o NH da G30, simultaneamente com a interação do fosfato na posição sn-3 com a K49. A orientação da ligação peptídica entre resíduos 29 e 30 parece particularmente bem determinada devido á interação do CO 29 com K122 no monômero 1 ou Tyr319 através de uma molécula de água no monômero 2.

Qual então seria a origem da falta de atividade catalítica ? A explicação mais provável seria mesmo a não estabilização do oxiânion mas não devido a inviabilidade de interagir com o NH da G30 mas sim com o N_ε da K49 que substitui o cálcio. Nas fosfolipases D49 o cálcio possui uma heptacoordenação incluindo duas moléculas de água. Na formação do complexo com o substrato as águas são deslocadas por um oxigênio do sn-3 fosfato e pelo oxigênio na posição sn-2.

Durante o ataque nucleofílico pela água catalítica, este último oxigênio adquire uma carga negativa (o oxiânion) estabilizada simultaneamente pelo Ca^{+2} , e o NH da G30, como descrito acima. O nosso modelo sugere que não haveria como estabilizar esta carga (e portanto reduzir a barreira energética) uma vez que as ligações de hidrogênio potenciais da lisina 49 estão completas. No complexo N_ξ está previsto a interação com as carbonilas 28, 30 e o fosfato.

Seria, então a falta de uma água trocável que poderia ser deslocada pelo oxiânion e o fato que as interações da K49 já estão completas que é a base fundamental da falta de atividade.

Já que existem diferenças na sequência do loop de cálcio, nas proteínas K49, nada impede que haja diferentes motivos pela falta de atividade em moléculas diferentes. Notavelmente na posição 33 da molécula *Agkistrodon p.p.* há uma Histidina, provavelmente responsável pela conformação incomum descrita por Scott *et al.*, 1992 o que levou estes autores à proposta descrita anteriormente. Porém o fato que a segunda forma cristalina (CF2) e a estrutura da mesma toxina determinado por um segundo grupo (Holland *et al.*, 1990) apresentam o NH da G30 na mesma orientação descrita aqui para BthTX-I, indica que seria possível uma interação normal entre G30 e o oxiânion. Acreditamos que a proposta acima detalhada seria mais geral.

6.7 Resíduos da Interface

Vários são os trabalhos que analisam os resíduos da superfície de interface de ligação da fosfolipase com a membrana. Na tabela 6.5 encontra-se um resumo dos resíduos citados por alguns autores.

[1] Selistre de Araujo, et al., (1996)	[2] White et al., (1990)	[3] Dijkstra et al., (1981)	BthTxI
K7 E12 T13 K16 N17 K78 K80 K118 K117	L2 F5 Y6 I9 W19 G31 Y75	W3 N6 L19 L31 Y69	L2 W3 G6 I9 K20 L32 Y75 A19 V31 K69 K115 K116 R118 K122 K126 K128

Tabela 6.5: Resíduos da interface de ligação com a Membrana sugeridos por vários autores. No caso da BthTx-I autores sugerem que os resíduos poderiam estar relacionados com miotoxicidade de K49's, mas não necessariamente via interação com a membrana.

Além dos resíduos citados acima incluímos os demais resíduos básicos do C-terminal.

Na Fig. 6.18 (a) estão os resíduos da interface sugeridos no trabalho [1], enquanto que na Fig 6.18 (b) estão os resíduos sugeridos nos trabalhos [2] e [3], mapeados na estrutura dimérica (modelo 2) da BthTX-I.

Fig. 6.18 (a) Resíduos da interface na estrutura dimérica da BthTX-I (modelo 2) sugeridos pelo trabalho [1]. Os resíduos em branco, ou rosa, foram identificados como conservados nas miotoxinas K49 e/ou miotoxinas/neurotóxicas D49. Os resíduos em vermelho são resíduos básicos do C-terminal.

Fig. 6.18 (b) Resíduos da interface na estrutura dimérica da BthTX-I (modelo 2) sugeridos pelos trabalhos [2] e [3].

Na Fig. 6.18 (a) os resíduos em branco , ou rosa, foram identificados como conservados nas miotoxinas K49 e/ou miotoxinas/neurotóxicas (D49). A maioria destes resíduos estão apontando para a interface entre os monômeros, sugerindo que a motivação da sua conservação seria a manutenção desta interface. Este resultado sugere que é improvável que estes resíduos estejam envolvidos diretamente na interação com a membrana, mas dá um certo apoio á ideia da importância do dímero em miotoxinas desta classe. Uma alternativa, é que após a interação com a membrana o dímero se desfaz expondo estes resíduos, permitindo expressão de uma segunda função. Os resíduos em vermelho são resíduos básicos encontrados no C-terminal. A figura mostra a possibilidade destes resíduos estejam envolvidos numa interação direta com o fosfato das cabeças polares dos fosfolípeidos que compõe a membrana. A Fig. 6.18 (b) demonstra também a viabilidade dos resíduos propostos nos artigos de Dijkstra e Scott em exercer um papel no reconhecimento da membrana.

Na Fig. 6.19(a) e (b) é sugerido um modelo esquemático para a interação da estrutura dimérica da BthTx-I com a membrana. Desta forma sugerida, os resíduos dos sítios ativos estariam expostos para a interação simultânea com a membrana. Estas figuras servem apenas para demonstrar que os dois canais hidrofóbicos que levam ao sítio ativo poderiam encontrar uma superfície (membrana, micela, etc.) simultaneamente. É possível que na realidade a interação seja bem mais complexa, podendo até envolver um receptor proteico na membrana, como tem sido sugerido. Na primeira figura, os fosfolipídeos que eventualmente entrarem nos canais hidrofóbicos da BthTX-I estão destacados da membrana, somente para facilitar a visualização do processo.

A)

B)

Fig. 6.19 (a) e (b) Interação da estrutura dimérica da BthTx-I com a membrana Estas figuras servem apenas para demonstrar que os dois canais hidrofóbicos que levam ao sítio ativo poderiam encontrar uma superfície simultaneamente.

6.8. Árvores filogenéticas

A relação entre as sequências e/ou estruturas tridimensionais de um grupo de proteínas funcionalmente relacionadas pode ser expressada quantitativamente em termos de árvores filogenéticas ou dendogramas. A topologia das árvores dá uma indicação de como tais sequências (ou estruturas) devem ser agrupadas; o tamanho dos ramos fornece algum dado sobre as distâncias evolucionárias entre as espécies, ou moléculas. No caso, quando se compara proteínas da mesma função de espécies diferentes esta informação pode ser usada para traçar a história evolucionária das espécies.

Tradicionalmente, tais árvores ou dendogramas (no caso de proteínas parálogas) são calculadas a partir de alinhamentos de sequências (Methods in enzymology, vol. 183, 1990). Mais recentemente métodos foram desritos para o uso de estruturas com a mesma finalidade. O uso de estruturas tem certas vantagens e desvantagens. Para estudos de eventos antigos na evolução , ou para estudos de proteínas que evoluíram rapidamente, estruturas 3D trazem a vantagem de conservar mais ao longo do processo evolutivo. Isto acontece porque mutações ocorrem principalmente nas regiões da proteína expostas ao solvente, não alterando desta forma a estrutura tridimensional da proteína e também os seus principais resíduos do sítio ativo. Frequentemente, proteínas homólogas (evolucionariamente relacionadas) se diferem tanto ao nível de estrutura primária que mesmo o alinhamento das sequências se torna difícil. Nestes casos, através da sobreposição de estruturas tridimensionais a informação perdida pode ser mais facilmente recuperada (Bajaj & Blundell, 1984). Em contrapartida, o uso de estruturas tridimensionais é evidentemente mais complexo devido às diversas maneiras de definir similaridade, as diferentes qualidades das estruturas usadas (medida de resolução, R_{fator} ou outros critérios de estereoquímica) e os diversos métodos de determinação da estrutura. Neste último caso, vale lembrar que estruturas de média resolução determinadas por Substituição Molecular podem ser viciadas pela estrutura usada como modelo.

Dendogramas foram calculados incluindo todas as fosfolipases A_2 de estrutura e sequência conhecida. Para as sequências o alinhamento apresentado na tabela 1.2 foi utilizado. Dois métodos Neighbor-joining e Fitch -Margoliash, 1967 foram utilizados para o cálculo tanto no caso de sequências quanto de estruturas. Mas como os resultados foram praticamente iguais, apresentaremos somente o resultado do método Fitch -Margoliash. Neste dendograma foram incluídas: a PLA_2 do fluido sinovial humano; duas PLA_2 pancreáticas (bovina/porco); os venenos das cobras *Naja naja atra*, *Naja naja naja*, *Crotalus atrox*, *Bothrops asper* e *Bothrops jararacussu* e *Agkistrodon p.p*; e a PLA_2 do veneno da abelha. A construção de todos os dendogramas foi feita com o programa PHILIP (Felsenstein, 1993). No caso de sequências o significado estatístico das ramificações foram calculadas através do método Bootstrap (Felsenstein, 1993). No caso das estruturas sobreposições de moléculas foram feitas usando inicialmente as regiões identificadas como modelo comum (Scott *et al.*, 1990) e melhorados usando a opção LSQ-improve do program "O".

Para o cálculo da distância evolucionária no caso de estruturas usamos o método que faz uma comparação par a par de corpo rígido (Methods in enzymology vol 183, 1990). As informações obtidas são: a equivalência topológica e o rms das distâncias sobre as posições dos carbonos alfa equivalentes. O número de posições topologicamente equivalentes PFTE é convertido em uma topologia fracional pela divisão do número de posições equivalentes pelo comprimento da menor estrutura. O rms é a distância medida e é convertida num score similar, o SRMS, calculado como $1 - \text{rms} (\text{\AA}) / 3.5 \text{\AA}$. A distância métrica D é calculada como na Eq. 6.2.

$$D = -100 \ln(w_1 \text{PFTE} + w_2 \text{SRMS}) \quad \text{Eq. 6.2}$$

onde

$$w_1 = [(1 - \text{PFTE}) + (1 - \text{SRMS})] / 2$$

$$w_2 = (\text{PFTE} + \text{SRMS}) / 2$$

com

$$w_1 + w_2 = 1$$

Quanto mais relacionadas forem as sequências, o fator PFTE não as diferencia. Entretanto, a distancia rms fornece uma boa medida das diferenças entre elas. Como resultado, w_1 e w_2 , são usados para inversamente modular a contribuição do PFTE e do SRMS no cálculo de D.

Os dois dendrogramas estão apresentados na Fig. 20 (a) e (b). A topologia é idêntica nos dois casos, mostrando a possibilidade de reproduzir fielmente um dendrograma a partir da estrutura. Este caso, representa um teste interessante no sentido de que o dendrograma calculado pelas sequências é estatística e razoavelmente estável. Portanto, podemos acreditar que representa bem as relações de similaridade entre as proteínas. É notável então, a reprodução da mesma topologia pelo estudo de estruturas.

Mais importante é a aparência no dendrograma da BthTX-I, como parente próximo da miotoxina II da *B. asper*. Este resultado é notável pois foi a estrutura da proteína de *Agkistrodon p.p.* que foi usado como modelo na Substituição Molecular. O dendrograma demonstra que o procedimento de refinamento da BthTX-I foi suficiente para eliminar o vício da estrutura inicial.

Fig. 6.20 (a): Dendrograma baseado no alinhamento de seqüências, utilizando o método de Fitch. Nesta figura estão incluídas as seqüências da PLA₂ do fluido sinovial humano; duas PLA₂ pancreáticas (bovina/porco); os venenos das cobras *Naja naja atra*, *Naja naja naja*, *Crotalus atrox*, *Agkistrodon p.p.*, *Bothrops asper* e *Bothrops jararacussu*; e a PLA₂ do veneno da abelha.

Fig. 6.20 (b) : Dendrograma baseado na estrutura tridimensional de algumas fosfolipases A₂ utilizando o método de Fitch. Nesta figura estão incluídas: a PLA₂ do fluído sinovial humano; duas PLA₂ pancreáticas (bovian/porco); os venenos das cobras *Naja naja atra*, *Naja naja naja*, *Crotalus atrox*, *Agkistrodon p.p.*, *Bothrops asper* e *Bothrops jararacussu*; e a PLA₂ do veneno da abelha.

6.9 Referências Bibliográficas

- ARNI, R. K. & WARD, R. J.; GUTIERREZ, J. M. & TULINSKY, A. - *Acta Cryst.* **D51**, 311-317 (1995).
- BAJAJ, M. & BLUNDELL, T.L. *Annu. Rev. Biophys. Bioeng.* **13**, 453 (1984).
- DIJKSTRA, B.W., DRENTH, J. & KALK, K. H. *Nature* **289**, 604-606 (1981).
- CCP4 - User Documentation , Daresbury Laboratory (1984).
- FELSENSTEIN, J. - *Phylogeny Inference Package* - version 3.5 (1993).
- FITCH, W. M. & MAGOLIASH, E. *Science* **155**, 279 (1967).
- HUTCHINSON, E. G. & THORNTON, J. M. *Protein Science* **5**, 212-220 (1996).
- HUTCHINSON, E. G. & THORNTON, J. M. - *HERA* (A Program to Draw Schematic Diagrams of Protein Structures. University College, London (1993).
- INSIGHT - User Guide Biosym Technologies Version 2.3.0 (1993).
- JANIL, J., MILLER, S. & CHOTHIA, C., *J. Mol. Biol.* **204**, 155-164 (1988)
- LAWRENCE, M. C. & COLMAN, P. M.- *J. Mol. Biol.* **234**, 946-950 (1993).
- McDONALD, I. K. & THORNTON, J. M. - *J. Mol. Biol.* **238**, 777-793 (1994).
- MEBS, O. & OWNBY, C. L. *Pharmacol. Ther.* **48**, 223-236 (1990).

- Methods in Enzymology , vol **183** - *Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences* - Academic Press (1990).

- NICHOLLS, A. - GRASP (*Graphical representation and Analysis of Surface properties*) Manual for Version 1.0.4 (1992).

- RICHARDSON, J. S. & RICHARDSON, D. C. *Prediction of Protein Structure and The Principle of Protein Conformation* - Plenum Press New York (1989).

- SELISTRE DE ARAUJO, H. S., WHITE, S.P. & OWNBY, C.L. *Archives of biochemistry and Biophysics* **325**, (1996).

- SCOTT, D. L., OTWINOWSKI, A., GELB, M. H. & SIGLER, P. B., *Science* **250**, 1563-1566 (1990).

- SCOTT, D. L., WHITE, S.P., BLOWING, J. L., ROSA, J.J., GELB, M. H. & SIGLER, P. B., *Science* **254**, 1007-1010 (1991).

- SCOTT, D. L., ACHARI, A., VIDAL, J. C. & SIGLER, P. B., *J. of Biol. Chem.* **267**,31, 22645-22657 (1992).

- WHITE, S. P., SCOTT, D. L., OTWISNOWSKI, Z., GELB, M. H. & SIGLER, P. B., *Science* **250**, 1560-1563 (1990).

CAPÍTULO 7

CONCLUSÕES E PERSPECTIVAS FUTURAS

As melhores condições de cristalização da BthTX-I foram obtidas com a solução tampão HEPES 0.1 M pH 7.0 variando a concentração de $(\text{NH}_4)_2\text{SO}_4$ de 50% a 65%. O método de cristalização foi o método de difusão de vapor “hanging drop”.

A BthTX-I cristalizou no sistema trigonal, com os seguintes parâmetros de rede: $a=b=57.58 \text{ \AA}$, $c = 131.27 \text{ \AA}$, $\alpha=\beta=90^\circ$ e $\gamma = 120^\circ$. A unidade assimétrica contém dois monômeros.

A coleta de dados foi inicialmente realizada utilizando o difratômetro automático R-Axis IIC da Rigaku Co. do Laboratório de Cristalografia de Proteínas do IFSC-USP. Também foi realizada uma coleta de dados no SERC Daresbury Laboratory na Inglaterra, com o objetivo de obter dados a mais alta resolução $2.1 \text{ \AA}$ e desta forma se melhorar a qualidade dos mapas. O processamento das imagens coletadas no síncrotron de Daresbury foram processadas usando o programa MOSFLM. A completeza do conjunto de dados foi 99.6 % até $2.1 \text{ \AA}$ com um fator $R_{\text{sym}} = 6.3\%$.

A cela unitária determinada a partir das imagens de difração, apresenta simetria trigonal e o processamento dos dados foi feito em P3. Para a determinação do grupo espacial da BthTX-I foram inicialmente calculados valores de R_{sym} supondo grupos pontuais diferentes e também a função de auto-rotação, ambos com o intuito de determinar a simetria cristalográfica. Para resolver a ambiguidade na mão do grupo espacial (além da resolução da estrutura) foi usado o método de Substituição Molecular. A estrutura foi resolvida por substituição Molecular, com o uso do programa AMoRe. O modelo utilizado foi a estrutura da *Agkistrodon piscivorus piscivorus* resolvida a $2.0 \text{ \AA}$ de resolução. A procura foi realizada entrando com apenas um monômero da App para encontrar duas moléculas da BthTX-I na unidade assimétrica. Foram encontradas duas soluções. Para cada uma destas soluções foi fornecida uma matriz de rotação e uma de translação, sendo a primeira a solução correta com um coeficiente de correlação de 62.7% e um fator $R = 38.7\%$. Estas transformações foram aplicadas ao monômero da App e desta forma obtivemos o primeiro modelo das duas moléculas para a BthTX-I.

O refinamento por dinâmica molecular com temperatura controlada foi realizado com o programa X-PLOR. Os mapas produzidos do modelo final foram de boa qualidade, para uma resolução de 2.1 Å. O modelo final contém 1907 átomos e 114 moléculas de água e possui um fator $R = 18.7\%$ e um $R_{\text{free}} = 27.4\%$ para 14910 reflexões.

Toda a análise estrutural e estereoquímica indica que a estrutura está bem determinada.

A superposição dos dois monômeros apresenta diferenças significativas na região do “loop de ligação do Ca^{+2} ” resíduos 28 a 33 e no C-terminal. Estas diferenças estruturais entre os dois monômeros podem ser justificadas exatamente devido aos diferentes contatos cristalinos entre as subunidades.

O dímero escolhido para definir a unidade assimétrica (modelo 1) não apresentou nenhum contato atômico e tinha orientações relativas entre os monômeros bastante diferentes da *Bothrops asper* que também é dimérica e possui um alto grau de identidade sequencial. A avaliação do empacotamento cristalino revelou a presença de um dímero parecido (modelo 2) com o dímero da *Bothrops asper*.

Foram realizadas várias eletroforeses em gel de 15% de Poliacrilamida com e sem β -MSH que mostraram que a BthTX-I apresenta um rastro sugerindo um equilíbrio entre o dímero e o monômero.

Para avaliar esta forma dimérica utilizamos o programa desenvolvido por Lawrence & Colman para calcular o parâmetro S_c que analisa a complementaridade na interface entre duas moléculas proteicas. O grau de complementaridade entre as duas superfícies em contato no dímero apresentaram um valor abaixo daquilo esperado para uma molécula oligomérica (≈ 0.58 comparado com ≈ 0.70). Estes resultados quando comparados com os pesos moleculares determinados por eletroforese são intrigantes, e levanta a questão da existência ou não do dímero em solução e a sua relevância na atividade biológica.

Certamente, a existência da mesma interface dimérica nas moléculas de duas espécies diferentes (*B. asper* e *B. jararacussu*) é um forte indício da sua relevância biológica. Se existisse em solução ambos os sítios ativos estariam orientados para o mesmo lado da molécula permitindo uma interação simultânea com a membrana.

Foi realizada uma comparação do “loop de ligação do Ca^{+2} ” resíduos 28 ao 33 de algumas das estruturas de fosfolipases A_2 . Desta comparação, observa-se uma grande variabilidade desta região. A estrutura cristalográfica da BthTX-I apresenta diferenças importantes entre os dois monômeros na região do “loop de ligação do Ca^{+2} ”. As diferenças foram atribuídas aos contatos cristalinos e parece que a conformação apresentada pelo monômero 2 seria a mais estável em solução na ausência do fosfolipídeo. No modelo do complexo, propomos que o fosfato do fosfolipídeo interage através de um dos seus oxigênios com o N_ξ da K49. No monômero 1, esta interação parece facilmente acomodada, porém no monômero 2, para haver complexação necessitaria de uma mudança conformacional do loop. Propomos portanto, que a flexibilidade do loop de ligação do Ca^{+2} é fundamental para a ligação do fosfolipídeo.

O modelo teórico para a ligação do fosfolipídeo na BthTX-I sugere que nenhuma interação direta entre o sn2 do fosfato e a K49 deve ser esperada de forma a explicar a falta de atividade catalítica.

A explicação mais provável para a falta de atividade catalítica seria mesmo a não estabilização do oxiânion conforme sugerido, mas não devido a inviabilidade de interagir com o NH da G30, mas sim com o N_ξ da K49, que substitui o Ca^{+2} . Nas fosfolipases D49 o cálcio possui uma heptacoordenação, incluindo duas moléculas de água. Na formação do complexo com o substrato, as águas são deslocadas por um oxigênio do sn3 fosfato e pelo oxigênio na posição sn2. Durante o ataque nucleofílico pela água catalítica, este último oxigênio adquire uma carga negativa (o oxiânion) estabilizada pelo Ca^{+2} e o NH da G30. O nosso modelo sugere que não haveria como estabilizar esta carga, uma vez que as ligações de hidrogênio potenciais da K49 estão completas. Seria, então a falta de uma água trocável que poderia ser deslocada pelo oxiânion e o fato que as interações da K49 já estão completas a base fundamental para a falta de atividade.

Foram analisados também os resíduos sugeridos por vários autores da superfície de interface de ligação da fosfolipase com a membrana e/ou carácter miotóxico, mostrando: a viabilidade dos resíduos propostos nos artigos de Dijkstra e Scott em exercer um papel de reconhecimento da membrana; a possibilidade dos resíduos básicos encontrados na região do C-terminal estarem envolvidos numa interação direta com o fosfato das cabeças polares dos fosfolípeos que compõem a membrana; e os resíduos da interface do dímero sugerindo que é improvável que estes resíduos sejam envolvidos diretamente na interação com a membrana mas dá um certo apoio à ideia da importância do dímero em miotoxinas desta classe.

Dendrogramas foram calculados incluindo todas as fosfolipases A_2 de estrutura e sequência conhecida. A topologia é idêntica nos dois casos mostrando a possibilidade de reproduzir fielmente um dendrograma a partir de estruturas.

Quanto às perspectivas futuras, um dos aspectos iniciais seria o fato de que como não existe nenhuma estrutura de uma K49 PLA_2 com um inibidor fosfolípido ou análogo, seria interessante se realizarem tentativas de cristalização e resolução de uma estrutura, para se poder avaliar exatamente a proposta teórica sugerida aqui.

Outro aspecto é que muitas mutagêneses específicas podem ser realizadas para avaliar a função destas K49 PLA_2 . Uma possibilidade interessante seria então, tentar transformar uma K49 em uma D49 PLA , e comparar com resultados de outras toxinas, para desta forma se examinar a importância deste resíduo em particular. Uma outra possibilidade seriam mutações nos resíduos sugeridos por Selistre de Araujo et al., 1996 para se avaliar o carácter miotóxico e a estabilização da sua forma dimérica. Estudos bioquímicos e cristalográficos poderiam ser feitos para correlacionar estrutura e função de uma K49 PLA_2 .