

**Estudos de difração de raios-x a alta resolução da
beta-lactoglobulina bovina**

Autora: Kesley M. Godinho de Oliveira ✓

Orientador: Prof. Dr. Igor Polikarpov

2013.05.18

UNIVERSIDADE ESTADUAL DE CAMPINAS

**ESTUDOS DE DIFRAÇÃO DE RAIOS-X
A ALTA RESOLUÇÃO DA
BETA-LACTOGLOBULINA BOVINA**

POR

KESLEY MORAES GODINHO DE OLIVEIRA

UNIVERSIDADE ESTADUAL DE CAMPINAS,
INSTITUTO DE FÍSICA "GLEB WATAGHIN", UNICAMP

TESE

SUBMETIDA AO INSTITUTO DE FÍSICA
UNIVERSIDADE ESTADUAL DE CAMPINAS
COMO REQUERIMENTO PARA OBTENÇÃO DO
TÍTULO DE MESTRE EM FÍSICA

CAMPINAS, SÃO PAULO, BRASIL

09/03/99

ESTE EXEMPLAR CORRESPONDE À REDAÇÃO FINAL
A TESE DE MESTRADO DEFENDIDA PELA ALUNA
KESLEY MORAES GODINHO DE OLIVEIRA E APROVADA
PELA COMISSÃO JULGADORA.

LNS, 04 de maio 1999

M. D. - Polikarpov

UNIV.	IFGW
Nº	
V.	
PREÇO	38130
	229,99
	x
PREÇO	R\$ 1,00
	20/07/99

IF/1177

CM-00125390-3

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IFGW - UNICAMP

OL4e

Oliveira, Kesley Moraes Godinho de

Estudos de difração de raios-x a alta resolução da beta-lactoglobulina bovina. - Campinas, SP : [s. n.], 1999.

Orientador: Igor Polikarpov

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Física "Gleb Wataghin".

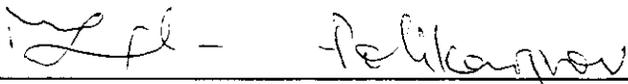
1. Cristalografia. 2. Proteínas. 3. Raios x - Difração. 4. *Beta-lactoglobulina bovina. 5.*Tanford, Transição de. I. Polikarpov, Igor. III. Universidade Estadual de Campinas. Instituto de Física "Gleb Wataghin". IV. Título.

*Ao meu marido, Sandro,
aos meus pais, Rubens e Leonice,
ao meu irmão, Luciano,
aos meus sogros, Paulo e Flora
e ao meu cunhado, Rogério*

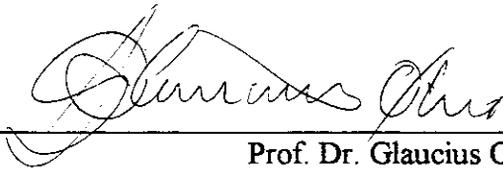
**Instituto de Física "Gleb Wataghin"
UNIVERSIDADE ESTADUAL DE CAMPINAS**

MEMBROS DA COMISSÃO JULGADORA DA TESE DE MESTRADO DE KESLEY MORAES GODINHO DE OLIVEIRA, RA 920751 APRESENTADA E APROVADA AO INSTITUTO DE FÍSICA "GLEB WATAGHIN", DA UNIVERSIDADE ESTADUAL DE CAMPINAS, EM 09 DE MARÇO DE 1999.

COMISSÃO JULGADORA:



Prof. Dr. Igor Polikarpov - (Orientador da Candidata)
Laboratório Nacional de Luz Síncrotron



Prof. Dr. Glaucius Oliva
Instituto de Física-USP/São Carlos



Profa. Dra. Maria Cristina dos Santos
DFESCM/IFGW/UNICAMP

Agradecimentos

Ao prof. Dr. Igor Polikarpov pela oportunidade dada para trabalhar com cristalografia de proteínas.

Às técnicas do laboratório de biologia molecular, Elisabeth, Milena, Viviane, Patrícia, Valéria e aos técnicos da linha de cristalografia de proteínas, Conrado e Brandão pelo apoio dado durante este período e pela amizade.

À Dra. Paula Kuser pela amizade, pelo apoio, pelas discussões e "dicas" na área de cristalografia.

Ao Lee, Júlio, Ricardo Aparício, Luís Augusto, Sandra, Andrea, Gabriela pelo apoio, pelos bons momentos e companheirismo.

Aos funcionários do LNLS pelo empenho em proporcionar todas as condições para o desenvolvimento do projeto e pelas amizades conquistadas.

À Vera, Michele e ao prof. Dr. Sérgio Ferreira pelo trabalho cooperativo na UFRJ com os experimentos de fluorescência com alta pressão. Em especial à Vera por sua amizade.

Aos professores Dr. Paulo Arruda e Dr. Adilson Leite por terem permitido que etapa inicial do projeto fosse realizada no CBMEG.

Ao meu marido pelo amor, apoio incondicional e a segurança transmitida.

Aos meus pais e irmão por tudo que eles representam em minha vida.

Aos meus sogros pelo carinho e apoio.

À CAPES (6 meses) e à FAPESP (22 meses) pelo suporte financeiro e ao Laboratório Nacional de Luz Síncrotron (LNLS) pelo suporte técnico durante o projeto.

Àquele que torna todas as coisas possíveis.

*"Porque, como a terra produz os seus renovos,
e como o horto faz brotar o que nele se semeia,
assim o Senhor Deus fará brotar a justiça
e o louvor perante todas as nações"*
Isaiás 61:11

Resumo	ix
Abstract	xi
Capítulo 1: A beta-lactoglobulina bovina	1
1.1 Introdução.....	1
1.2 Estrutura cristalográfica da beta-lactoglobulina bovina	3
1.3 Objetivos do trabalho	5
1.4 Conteúdo da tese	5
Capítulo 2: Cristalização e coleta de dados de difração de raios-x da beta-lactoglobulina bovina	7
2.1 Cristalização de macromoléculas biológicas.....	7
2.1.1Cristalização em grupo.....	9
2.1.2 Cristalização por diálise	9
2.1.3 Cristalização por difusão através da interface.....	10
2.1.4 Cristalização por difusão de vapor	10
2.1.5 Cristalização da beta-lactoglobulina bovina	11
2.2 Coleta e processamento dos dados de difração	12
2.2.1 Coleta dos dados.....	12
2.2.2 Processamento dos dados.....	14
2.2.3 Coleta e processamento dos dados de difração dos cristais de beta-lactoglobulina bovina	16
2.3 Princípios teóricos de difração de raios-x	18
2.3.1 Periodicidade do cristal e a análise de Fourier	18
2.3.2 Condições de difração	19
2.3.3 Fator de estrutura.....	22
Capítulo 3: Determinação da estrutura cristalográfica.....	25

3.1 Introdução.....	25
3.2 Princípios teóricos da substituição molecular	27
3.2.1 Função de Patterson	27
3.2.2 Função de rotação.....	29
3.2.3 Função de translação	31
3.3 Aplicação da substituição molecular à beta-lactoglobulina	33
Capítulo 4: Refinamento das estruturas cristalográficas	36
4.1 Aspectos básicos do refinamento de estruturas.....	36
4.1.1 Refinamento vinculado	38
4.1.2 Refinamento restrito	38
4.1.3 Simulações de dinâmica molecular com resfriamento lento	44
4.1.4 Índices de confiabilidade.....	44
4.1.5 Sínteses de Fourier	45
4.2 Refinamento das estruturas cristalográficas da beta-lactoglobulina	46
4.3 Qualidade e acurácia dos modelos	48
4.4 Empacotamento cristalino das formas cristalinas X e Y.....	52
Capítulo 5: Comparação entre as estruturas cristalográficas da beta-lactoglobulina.....	53
5.1 Alguns aspectos gerais da BLG	53
5.2 Estrutura da BLG na forma cristalina X a 1.6Å.....	56
5.2.1 Comparação entre estruturas da BLG na forma cristalina X	59
5.3 Estruturas da BLG na forma cristalina Y	61
5.3.1 Análise da interface do dímero cristalográfico das variantes A e B	67
5.4 A transição de Tanford.....	68
5.5 Deposição no PDB das estruturas cristalográficas das variantes A e B.....	69
Conclusões	70
Apêndice A: Proteínas.....	71
A.1 Constituição das proteínas.....	71
A.2 Estrutura primária, secundária, terciária e quaternária das proteínas.....	74
A.3 Estabilidade da conformação de proteínas	76
Referências	78

Lista de figuras

Figura 1.1: Composição aproximada do leite de vaca	2
Figura 1.2: Família das lipocalinas	5
Figura 1.3: Panorama geral na determinação da estrutura cristalográfica de uma proteína.....	6
Figura 2.1: Princípio da cristalização	8
Figura 2.2: Método de cristalização em grupo	9
Figura 2.3: Método de cristalização por diálise	9
Figura 2.4: Método de cristalização por difusão através da interface	10
Figura 2.5: Método de cristalização por difusão de vapor	11
Figura 2.6: Cristais da mistura de isoformas A/B crescidos em pH 7,9	12
Figura 2.7: Preparação do cristal para a coleta de dados de raios-x.....	13
Figura 2.8: Parte experimental	14
Figura 2.9: Separação das contagens de fundo do ponto de reflexão	15
Figura 2.10: Planos de reflexão imaginários.....	22
Figura 2.11: Relação geométrica entre os vetores de onda.....	22
Figura 3.1: Fases da substituição molecular.....	28
Figura 3.2: Estrutura bidimensional de um arranjo de três átomos e o mapa de Patterson correspondente	29
Figura 3.3: Sistema angular usado na matriz [C] da função de rotação.....	31
Figura 3.4: Vetores utilizados na função de translação.....	32
Figura 4.1: Densidade eletrônica ao redor de um resíduo de triptofano	37
Figura 4.2: Alguns parâmetros estereoquímicos.....	39
Figura 4.3: Distância de átomos ao plano ideal	40
Figura 4.4: Volume quiral	40
Figura 4.6: Conformações de uma cadeia polipeptídica definidas em termos dos ângulos diédricos phi e psi.....	49

Figura 4.7: Conformação do resíduo de Tyr 99	49
Figura 4.8: Diagramas de Ramachandram	50
Figura 5.1: Estrutura cristalográfica da beta-lactoglobulina	54
Figura 5.3: Localização espacial dos resíduos de Trp19 e Arg124.....	55
Figura 5.4: Gráfico de diversos parâmetros para a mistura de isoformas A/B	57
Figura 5.5: Contatos na interface do dímero da BLG na forma cristalina X	58
Figura 5.6: Comparações entre estruturas da BLG na forma cristalina X	60
Figura 5.7: Diferenças no resíduo Pro38 nas formas cristalinas X da BLG	60
Figura 5.8: Gráfico do fator de temperatura das variantes A e B.....	61
Figura 5.9: Diferenças nas coordenadas entre as variantes A e B.....	64
Figura 5.10: Gráfico de diversos parâmetros para as variantes A e B	65
Figura 5.11: Comparação entre as estruturas da BLG	66
Figura 5.12: Diferenças nas coordenadas das variantes A e B da BLG normalizadas pelo deslocamento térmico médio.....	67
Figura 5.13: Comparação entre dímeros nas formas cristalinas X e Y	68
Figura A.1: Composição dos aminoácidos e formas quirais. (a) Forma L, (b) Forma D.....	72
Figura A.2: Os vinte tipos de aminoácidos	73
Figura A.3: Conformação da cadeia principal	74
Figura A.4: Formação de uma ponte dissulfeto	75
Figura A.5: Estrutura secundária das proteínas. (a) α -hélices (b) fios beta	75
Figura A.6: Estrutura primária, secundária, terciária e quaternária das proteínas	76
Figura A.7: Exemplo de espectros de fluorescência com diferentes pressões do triptofano ...	77

Lista de Tabelas

Tabela 2.1: Dados estatísticos dos dados de raios-x para mistura de isoformas A e B no pH ácido	16
Tabela 2.2: Dados estatísticos dos dados de raios-x para vários cristais da beta-lactoglobulina bovina crescidos no pH 7,9	17
Tabela 3.1: Resultados da substituição molecular para as estruturas da beta-lactoglobulina..	35
Tabela 3.2: Valores para o coeficiente de correlação e Rfactor ao final do processo de substituição molecular	35
Tabela 4.1: Resultado final do refinamento para as estruturas cristalográficas da BLG na forma cristalina Y	48
Tabela 4.2: Resultados do refinamento para as estruturas da BLG	51

A beta-lactoglobulina (BLG) pertence à família das lipocalinas e é encontrada em um grande número de mamíferos. Desde sua primeira purificação tem sido exaustivamente estudada. Em condições fisiológicas, a BLG existe como dímero onde cada monômero possui 162 aminoácidos correspondendo a aproximadamente 18kDa. Embora as propriedades físico-químicas da BLG sejam notórias, sua função biológica permanece desconhecida.

A BLG sofre mudanças conformacionais próximo ao pH 7,5 constatadas através de experimentos de rotação óptica específica e de curvas de titulação. Esta transição, denominada transição de Tanford, definiu o primeiro objetivo deste trabalho: resolver estruturas cristalográficas da BLG na faixa de pH ácido e básico a fim de investigar as modificações estruturais ocasionadas por esta transição.

O leite de vaca possui muitas variantes genéticas sendo que as variantes A e B são predominantes. A isoforma A difere na seqüência primária em relação à isoforma B nas posições 64 (Asp→Gly) e 118 (Val→Ala). Estas diferenças resultam em propriedades físico-químicas diferentes destas variantes e afetam o processamento do leite e derivados. O segundo objetivo deste trabalho foi a resolução das estruturas cristalográficas destas variantes a fim de explorar possíveis variações de origem estrutural, que viessem a justificar as diferenças físico-químicas entre ambas.

Cristais de beta-lactoglobulina foram crescidos na faixa de pH ácido e básico. Em pH ácido, o cristal foi obtido a partir da mistura de isoformas A/B e pertence ao grupo espacial P1. Em pH básico, os cristais pertencem ao grupo espacial C222₁ e foram crescidos para a mistura de isoformas A/B, variante A e variante B. Os cristais foram expostos a raios-X e os padrões de difração coletados. As estruturas cristalográficas da beta-lactoglobulina foram determinadas por substituição molecular e as coordenadas dos modelos de fases obtidos foram submetidas a refinamento. A mistura de isoformas A/B (P1), em pH ácido, foi determinada a 1,6 Å e refinada até um fator R cristalográfico de 19,8% ($R_{free} = 25,0\%$ e 171 moléculas de água). A correspondente em pH básico foi determinada a 2,0 Å e refinada até um fator R cristalográfico de 19,4% ($R_{free} = 27,2\%$ e 135 moléculas de água). As estruturas das variantes A e B foram determinadas a 2,0 Å e 1,95 Å e refinadas até um fator R cristalográfico de 19,2% ($R_{free} = 28,7\%$ e 123 moléculas de água) e 18,9% ($R_{free} = 25,2\%$ e 107 moléculas de água) respectivamente.

A comparação entre as estruturas resolvidas em diferentes valores de pH possibilitou a visualização do efeito da transição de Tanford: duas alças da estrutura representadas pelos resíduos 84 a 90 e 123 a 131 sofrem uma abertura e tornam expostos os grupos carboxil dos resíduos Glu 89 e Asp 129. Estes resíduos seriam, então, os responsáveis pelo comportamento anômalo observado para a BLG em torno do pH 7,5.

A superposição entre as estruturas das variantes A e B revelou que os resíduos 108 a 117 adotam uma conformação diferente nas variantes. Outro aspecto encontrado foi a

formação de uma cavidade no interior da proteína no resíduo 118. Nesta posição ocorre a troca de um resíduo de Val por Ala nestas variantes. Esta cavidade pode estar associada com a estabilidade destas variantes.

Beta-lactoglobulin (BLG), a very abundant protein found in several species of mammals belongs to the lipocalin family. Under physiological conditions, BLG is a dimer of 162 aminoacids per monomer, with a molecular mass of approximately 18kDa in each subunit. Although BLG has important physicochemical properties, its biological function remains unclear.

BLG undergoes a number of conformational changes near pH 7,5 as indicated by specific optical rotation and titration curve. This transition, termed Tanford transition, defined the first goal of this work: to solve crystallographic structures of BLG near acidic and basic pH, with the purpose of investigating the structural changes caused by the Tanford transition.

Bovine milk has many genetic variants of BLG, but the most prevalents are variants A and B. Variant A differs in the amino acid sequence at positions 64 (Asp→Gly) and 118 (Val→Ala) from variant B. These aminoacid changes cause differences in physicochemical properties of these variants and influence milk processing and milk derivatives. The second goal of this work was to solve the crystallographic structures of variants A and B with the purpose of exploring possible variants of structural origin, which could justify the physicochemical differences between both of them.

Crystals of beta-lactoglobulin were grown under acidic and basic pH. Under acidic pH, the crystal was obtained from a mixture of isoforms A and B, and belongs to space group P1. Under basic pH, the crystals belong to space group C222₁ and were grown from the mixture A/B, variant A and variant B. Crystals were exposed to X-rays and diffraction data were obtained. The crystal structures of beta-lactoglobulin were solved by the molecular replacement method and the initial model obtained was refined. The mixture of isoforms A/B (P1), under acidic pH, was determined at 1,6 Å resolution and refined to a crystallographic R factor of 19,8% ($R_{\text{free}} = 25,0\%$) with 171 solvent molecules. The corresponding structure in basic pH was determined at 2,0 Å resolution and refined to a crystallographic R factor of 19,4% ($R_{\text{free}} = 27,2\%$) with 135 solvent molecules. The crystal structures of variants A and B were determined at 2,0 Å e 1,95 Å resolution and refined to an crystallographic R factor of 19,2% ($R_{\text{free}} = 28,7\%$ and 123 solvent molecules) and 18,9% ($R_{\text{free}} = 25,2\%$ and 107 solvent molecules) respectively.

The comparison between structures solved under different pHs allowed the visualisation of the effect caused by Tanford transition: two loops of the structure formed by residues 84 to 90 and 123 to 131 open up and expose the carbonyl groups of residues Glu 89 and Asp 129. These two residues are considered responsible for the anomalous behaviour observed to BLG when it is in pH 7,5.

The superposition of the structures of variants A and B revealed that residues 108 to 117 adopt a different conformation in the variants. Another aspect that was observed is the

creation of an internal cleft close to residue 118, due to the change of one Valine by an Alanine. This cleft may be related to the stability of these variants.

A beta-lactoglobulina bovina

Neste capítulo serão apresentadas as características gerais da beta-lactoglobulina bovina.

1.1 Introdução

O leite de vaca e seus derivados fazem parte da dieta diária recomendada por nutricionistas devido ao seu valor nutricional. A composição aproximada do leite de vaca é apresentada na figura 1.1. Cerca de 80% das proteínas do leite são caseínas e possuem baixa solubilidade no pH 4,6. As proteínas restantes aparecem no sobrenadante após a precipitação das caseínas e são denominadas, genericamente, proteínas do soro do leite. As proteínas do soro do leite são globulares, mais solúveis em água que as caseínas e são sujeitas à desnaturação térmica. A maior fração destas proteínas, cerca de 9,8% (3,2g/litro de leite), corresponde à beta-lactoglobulina bovina (BLG), isolada pela primeira vez em 1934 (Palmer, 1934).

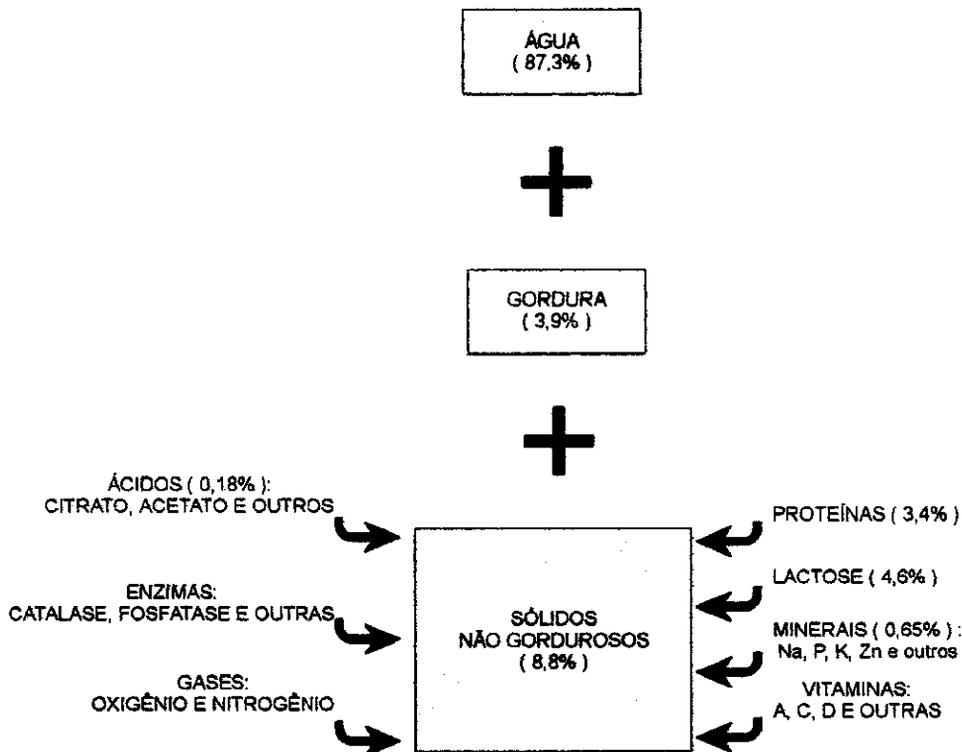


Figura 1.1

Composição aproximada do leite de vaca

A BLG é encontrada em um grande número de mamíferos (Hambling, McAlphine & Sawyer, 1992) e tem sido exaustivamente estudada desde sua primeira purificação. Somente em 1998 e início de 99 foram publicados 410 trabalhos.

Em condições fisiológicas, a BLG existe como dímero onde cada monômero possui 162 aminoácidos correspondendo à aproximadamente 18kDa. Duas pontes dissulfeto estão presentes na estrutura, Cys 66 — Cys 160 e Cys 106 — Cys 119, além de uma cisteína livre, Cys 121.

De acordo com os trabalhos realizados por Tanford (Tanford *et al.*, 1959), a BLG sofre mudanças conformacionais próximo ao pH 7,5 constatadas através de experimentos de rotação óptica específica e de curvas de titulação. Esta transição denominada transição de Tanford é caracterizada pela exposição de um grupo carboxila e modificações nas vizinhanças de um resíduo de tirosina. Também ocorre um aumento na reatividade de um único grupo sulfidril (Dunnill *et al.*, 1965).

Embora as propriedades físico-químicas da BLG sejam notórias, sua função biológica permanece desconhecida. O complexo BLG-retinol foi encontrado em bezerros recém-nascidos (Papiz *et al.*, 1986) indicando a possibilidade de que a BLG estivesse envolvida no transporte de vitamina A. Entretanto, Perez e colaboradores verificaram que as proteínas do leite ligam-se de modo inespecífico com o retinol (Perez *et al.*, 1995). A BLG também liga-se a muitas pequenas moléculas hidrofóbicas (Spector & Fletcher, 1970, Farrell *et al.*, 1987, Prez *et al.*, 1989) e poderia ser responsável pelo transporte das mesmas. Outro aspecto interessante foi relatado em 1996 por Robert Neurath e colaboradores que modificaram a BLG com 3-hidroxyphtalico anidrido (3-HP-BLG) e verificaram que este novo composto bloqueou o vírus HIV ligando-se ao receptor do vírus (Neurath *et al.*, 1996).

Em 1955, Aschaffenburg e Drewry descobriram variantes da BLG que diferiam em mobilidade frente a eletroforese no pH 8,6 e que posteriormente foram nomeadas variantes A e B. Embora muitas outras variantes genéticas existam, as variantes A e B são predominantes no leite de vaca. A isoforma A difere na seqüência primária, em relação a isoforma B, nas posições 64 (Asp→Gly) e 118 (Val→Ala). Estas diferenças resultam em propriedades físico-químicas diferentes destas variantes como estabilidade frente ao aquecimento, por exemplo.

O aquecimento altera as propriedades do leite como o sabor (devido a exposição de grupos sulfridrilas da BLG), diminui a facilidade na produção do queijo e produz mudanças na estabilidade das proteínas do soro do leite (Science of foods, Oregon State University). Assim, é evidente o interesse industrial nas particularidades destas proteínas desde que afetam diretamente o processamento do leite e seus derivados (Hill *et al.*, 1996).

A BLG é também identificada como a causa freqüente de alergia em lactantes (Fällström, 1978). Testes laboratoriais revelaram que a resposta imunológica em ratos é maior para a variante B quando comparada a variante A. A compreensão das propriedades alergênicas destas variantes propiciaria a produção de uma fórmula hipoalergênica para o leite (Malik, 1988).

1.2 Estrutura cristalográfica da beta-lactoglobulina bovina

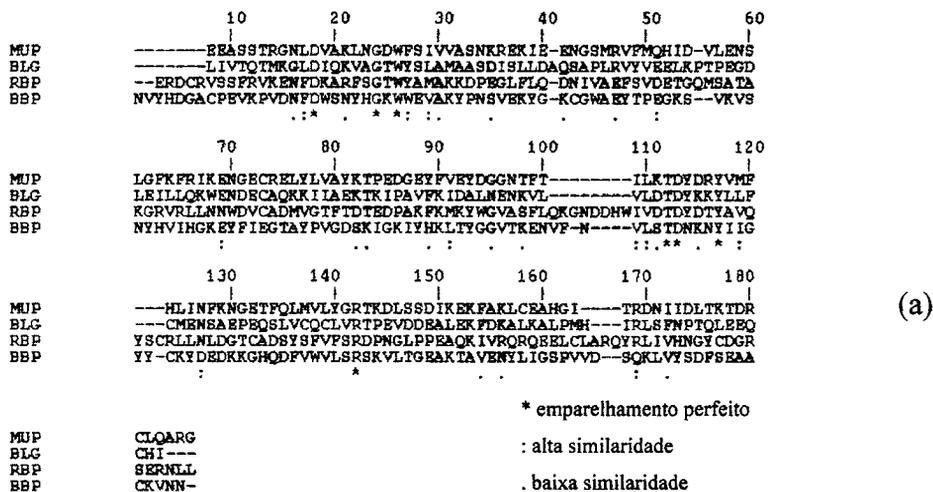
Diferentes formas cristalinas da BLG são conhecidas (Green *et al.*, 1979). Aquelas que crescem em torno do pH fisiológico são denominadas X, Y e Z e pertencem aos grupos

espaciais P1, C222₁ e P3₂21 respectivamente. Cristais na forma cristalina X são obtidos para valores de pH entre 6,0 e 6,9. A forma Y é obtida para valores de pH entre 7,0 e 8,0. A forma Z ocorre nos dois intervalos anteriores.

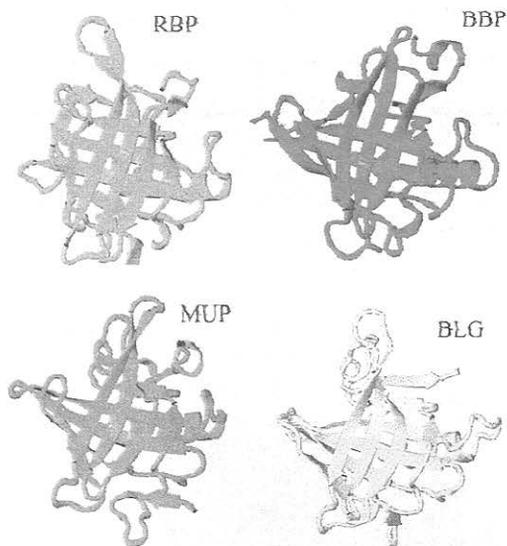
O enovelamento da BLG foi determinado em 1985 por Lindsay Sawyer. No entanto, por um longo período foi difícil refinar sua estrutura cristalográfica e a primeira estrutura refinada foi publicada recentemente para a forma cristalina X a 1,8Å (Brownlow *et al.*, 1997).

A BLG pertence à família das lipocalinas (Brownlow *et al.*, 1997). O nome da família foi adotado devido à similaridade com a proteína RBP ("*Retinol-Binding Protein*") responsável pelo transporte do retinol. A família das lipocalinas é composta por proteínas com diferentes funções e que apresentam baixa identidade sequencial entretanto com alta similaridade estrutural. Além da RBP, pode-se destacar outros membros da família das lipocalinas como a BBP ("*Bilin-Binding Protein*") e a MUP ("*Major Urinary Protein*"). Estas proteínas ligam, respectivamente, o pigmento azul em insetos (usado para camuflagem) e feromônios em ratos. A figura 1.2 mostra em (a) o alinhamento das seqüências da RBP, BBP, MUP e BLG e em (b) as estruturas cristalográficas destas proteínas¹.

A forma estrutural da BLG é descrita (Papiz *et al.*, 1986, Monaco *et al.*, 1987, Brownlow *et al.*, 1997, Qin *et al.*, 1998) como um cálice formado por fios β antiparalelos (figura 1.2 (b)). Recentemente, a estrutura do complexo palmitato-BLG foi resolvida e refinada (Wu *et al.*, 1998). Este trabalho propiciou a evidência direta de que um sítio de ligação para substrato está localizado na cavidade central da proteína.



¹ V. Apêndice A



(b)

Figura 1.2

Família das lipocalinas. (a) Alinhamento entre as seqüências das proteínas RBP, BBP, MUP e BLG. (b) Estrutura cristalográfica destas proteínas.

1.3 Objetivos do trabalho

Duas metas principais determinaram os objetivos do trabalho com a beta-lactoglobulina bovina:

1. Resolver estruturas cristalográficas da BLG na faixa de pH ácido e básico a fim de investigar as modificações estruturais ocasionadas pela transição de Tanford;
2. Resolver estruturas cristalográficas das variantes A e B com a finalidade de explorar possíveis variações de origem estrutural que viessem a justificar as diferenças físico-químicas entre ambas.

1.4 Conteúdo da tese

A determinação das estruturas tridimensionais da BLG foi efetuada através da técnica de difração de raios-X e será tratada ao longo dos quatro primeiros capítulos. Estes capítulos estão distribuídos em função das etapas usadas na resolução da estrutura da BLG. Estas etapas seguem o procedimento geral na determinação da estrutura cristalográfica de uma proteína como ilustrado na figura 1.3. Os primeiros tópicos de cada capítulo são reservados aos

princípios teóricos utilizados e os restantes para os resultados obtidos para a BLG. O quinto e último capítulo traz a análise das estruturas resolvidas, discutindo os aspectos estruturais da transição de Tanford e as diferenças entre as variantes A e B.

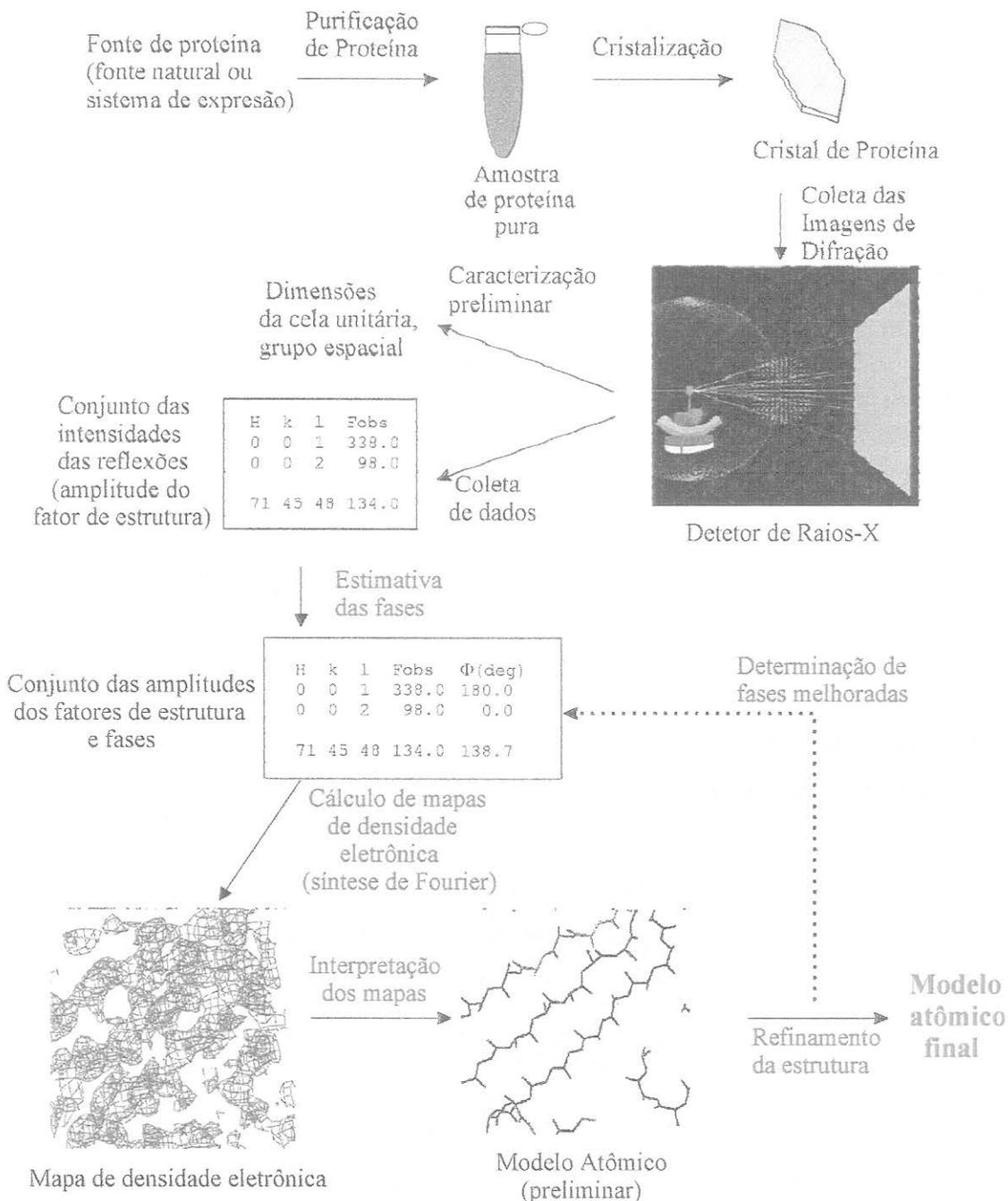


Figura 1.3

Panorama geral na determinação da estrutura cristalográfica de uma proteína (Imagem cedida pela Dr. Paula Kuser através de comunicação pessoal)

Cristalização e coleta de dados de difração de raios-X da beta-lactoglobulina bovina

Para que a estrutura cristalográfica de uma proteína seja determinada necessitamos de amostras no estado cristalino. Obter cristais de macromoléculas é uma verdadeira arte, pois muitos são os parâmetros (e nem todos conhecidos) que governam o mecanismo de cristalização. Conseqüentemente, não há disponível uma teoria abrangente, de forma que o processo é empírico. Neste capítulo serão apresentados os princípios de cristalização, o método utilizado para a beta-lactoglobulina, a coleta de dados de difração de raios-X dos cristais obtidos e, finalmente, uma introdução aos conceitos básicos da teoria de difração a fim de mostrar a relação entre a densidade eletrônica e o fator de estrutura, bem como situar o "*problema das fases*" nas medidas experimentais.

2.1 Cristalização de macromoléculas biológicas

Para cada proteína a ser cristalizada, um protocolo específico de cristalização deve ser encontrado. Devido à grande quantidade de parâmetros que governam a cristalização, abordagens estratégicas foram estabelecidas ao longo do tempo através de experimentação. A proteína é dissolvida em um solvente, normalmente água, que contém um tampão para controlar o pH e um agente precipitante. Em todas as técnicas praticadas esta solução protéica é conduzida à supersaturação e induzida ao retorno para o equilíbrio na expectativa de que as moléculas remanescentes do soluto sejam separadas para o estado cristalino (figura 2.1).

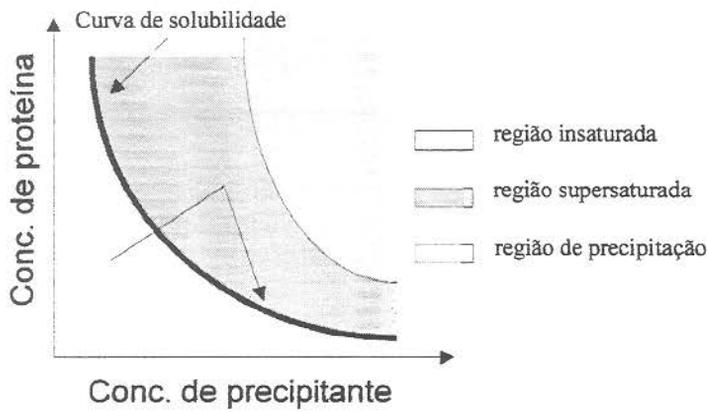


Figura 2.1

Princípio da cristalização

Em geral, a solubilidade de uma dada proteína depende dos fatores :

- pH;
- temperatura;
- composição química do precipitante;
- propriedades da proteína e do solvente.

Normalmente, o precipitante é um sal adicionado à solução. Seus íons competem com as moléculas de proteína pelas moléculas de água, pois ambos necessitam de hidratação para manter a solubilidade. Quando a competição se torna intensa, as moléculas protéicas se associam para satisfazer seus requerimentos eletrostáticos, eliminando assim as moléculas de água e, conseqüentemente, começam a perder solubilidade. Este fenômeno é conhecido por "*salting-out*".

Entretanto, há proteínas cuja solubilidade diminui à medida que a força iônica da solução diminui. Este fato é conhecido por "*salting-in*" e a cristalização pode ocorrer nestas condições. Logo, alterando a concentração do precipitante modificam-se as forças iônicas. Para a mesma força iônica, a variação do pH da solução acarreta mudanças na distribuição de cargas na superfície da molécula que podem viabilizar a cristalização.

Embora todas as técnicas de cristalização possuam o mesmo princípio, a maneira de conduzir o processo é particular a cada uma e favorece a cristalização de um ou outro tipo de proteína. A seguir, serão descritos brevemente os métodos de cristalização de macromoléculas biológicas usualmente aplicados (McPherson, 1994 e Weber, 1997).

2.1.1 Cristalização em grupo ("batch crystallization")

Neste método, todos os componentes para cristalização são colocados em uma única solução, juntamente com a solução de proteína preparada em uma concentração elevada, de modo que a condição inicial parte do estado de supersaturação. O sistema, então, é isolado e deixado em repouso (figura 2.2). O atrativo deste processo é a sua simplicidade e, geralmente, é utilizado para a produção de cristais quando são conhecidas as condições de cristalização.

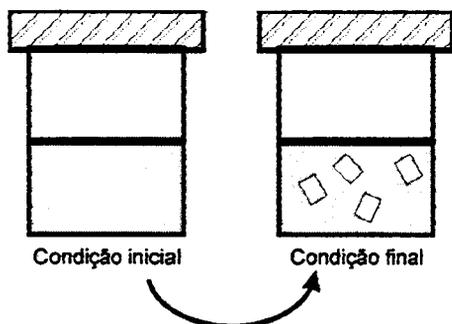


Figura 2.2

Método de cristalização em grupo

2.1.2 Cristalização por diálise

Nesta técnica uma solução de proteína é separada dos agentes cristalizantes, por uma membrana semipermeável (figura 2.3). A concentração da solução de proteína permanece constante, porém o estado de supersaturação é atingido através da difusão de componentes de baixo peso molecular pela membrana que irão alterar a composição da solução. Este método é adequado para cristalizações em condições de baixa força iônica.

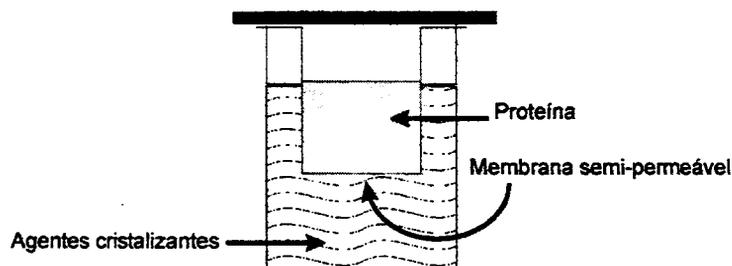


Figura 2.3

Método de cristalização por diálise

2.1.3 Cristalização por difusão através da interface

A finalidade desta abordagem é realizar a mistura entre a solução de proteína e a solução contendo os agentes cristalizantes de forma lenta através da interface entre os líquidos (figura 2.4). Na interface ocorre uma supersaturação local permitindo a nucleação e o subsequente crescimento de cristais. Este procedimento é propício em condições de microgravidade, pois a interdifusão entre as soluções pode ser efetuada de uma maneira mais controlada.

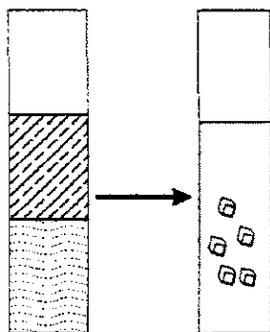
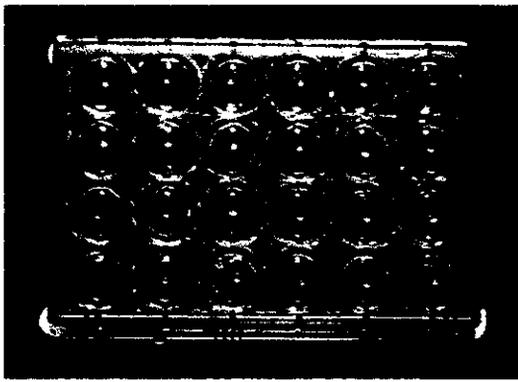


Figura 2.4

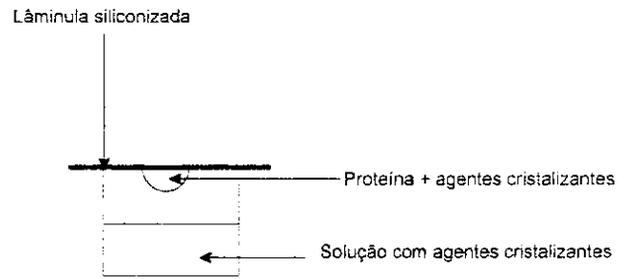
Método de cristalização por difusão através da interface

2.1.4 Cristalização por difusão de vapor

Nesta estratégia, o estado de supersaturação é alcançado através da difusão de vapor entre duas soluções. Na montagem experimental (figuras 2.5 (a) e (b)), volumes iguais das soluções de proteína e do reservatório (que contém agentes cristalizantes) são depositados sobre a própria lamínula de vidro siliconizada (método da gota "suspensa") que será utilizada para isolar o sistema. Se a solução de proteína possui baixa tensão superficial, ao invés de depositar as soluções sobre uma lamínula utiliza-se uma microponte (método da gota "sentada"). A gota depositada possui concentração menor (com relação ao precipitante) do que o reservatório. Esta diferença de concentração favorece a troca de vapores da solução menos concentrada para a mais concentrada. À medida que ocorre a perda de água, a concentração da gota muda favorecendo a cristalização. Esta técnica é amplamente utilizada devido à grande flexibilidade para variação de condições, como concentração do precipitante e pH.



(a)



(b)

Figura 2.5

Método de cristalização por difusão de vapor

2.1.5 Cristalização da beta-lactoglobulina bovina

Na faixa de pH ácido (em torno de 6,5) os cristais foram obtidos através do método de cristalização em grupo usando como agente cristalizante o sal sulfato de amônio e fosfato de sódio e potássio como tampão (experimentos realizados na Inglaterra por I. Polikarpov).

Na faixa básica de pH, os cristais de beta-lactoglobulina foram obtidos por difusão de vapor pelo método da gota "suspensa". As condições de cristalização apropriadas foram determinadas através de ensaios onde foram explorados a variação de pH, no intervalo de 7,0 à 8,5, contra um gradiente de concentração de sulfato de amônio $(\text{NH}_4)_2\text{SO}_4$ de 2,0 M à 3,5 M. A vantagem do $(\text{NH}_4)_2\text{SO}_4$ como precipitante é a sua alta solubilidade e grande força iônica.

As amostras da beta-lactoglobulina cristalizadas nesta faixa de pH foram: mistura de isoformas A/B, isoforma A e isoforma B, todas adquiridas da companhia SIGMA. As condições nas quais houve formação de cristais são listadas abaixo:

- 180 mM de tampão HEPES/TRIZMA (ver próximo parágrafo);
- 2,5 M de sulfato de amônio;
- solução de proteína à 20 mg/mL;
- temperatura a 18° C.

Para a solução de proteína da mistura de isoformas os cristais crescem no período de uma semana no intervalo de pH de 7,0 a 8,5 (experimentos realizados em passos de pH de 0,3). Entretanto, para as isoformas puras este intervalo de tempo é conseguido apenas no pH 7,9; em contrapartida, no pH 7,3 este tempo se estende de dois a quatro meses e em outros valores de pH nenhum resultado satisfatório foi obtido. A figura 2.6 exibe uma foto de alguns cristais beta-lactoglobulina.

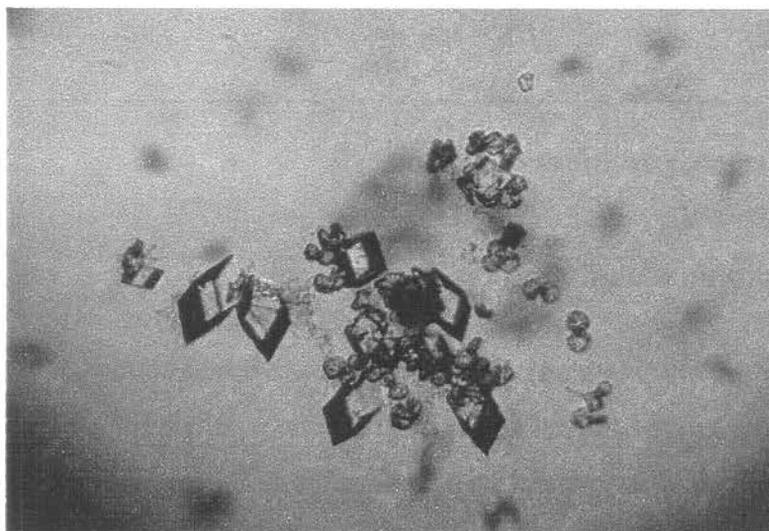


Figura 2.6

Cristais da mistura de isoformas A/B crescidos em pH 7,9

2.2 Coleta e processamento dos dados de difração

2.2.1 Coleta dos dados

Os cristais obtidos são colocados em capilares de quartzo, fixados em uma cabeça goniométrica (figura 2.7 (a) e (b)) e levados para exposição a um fonte de raios-X. A fonte utilizada foi a de radiação síncrotron. O cristal crescido em pH ácido foi coletado no sícrotron DESY, na Alemanha na linha X31. Os cristais crescidos em pH básico foram coletados no Laboratório Nacional de Luz Síncrotron localizado na cidade de Campinas, em São Paulo. As vantagens de coletar dados com uma fonte deste tipo residem na perspectiva de escolha do comprimento de onda mais adequado ao experimento, na redução do tempo de aquisição dos

dados, na obtenção de melhor resolução em relação a fontes convencionais (Longhi *et al.*, 1997).

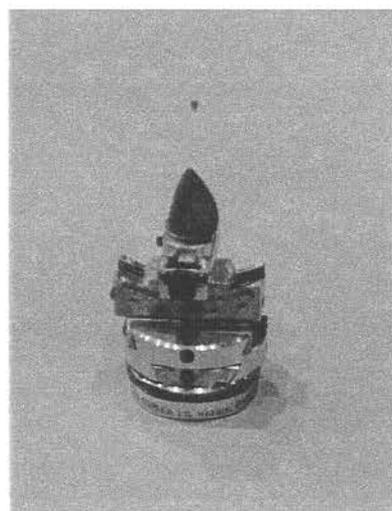
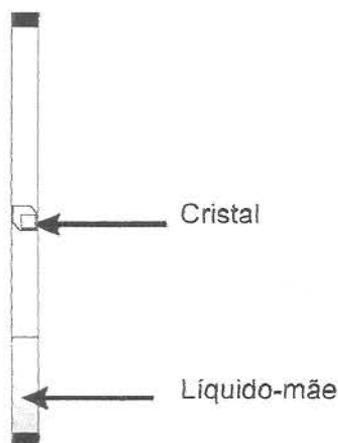
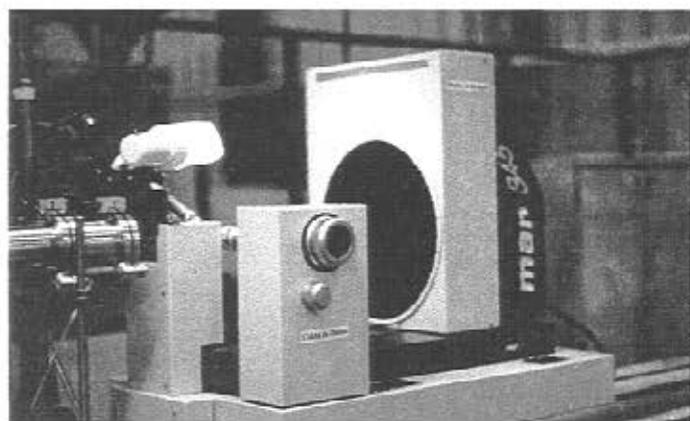


Figura 2.7

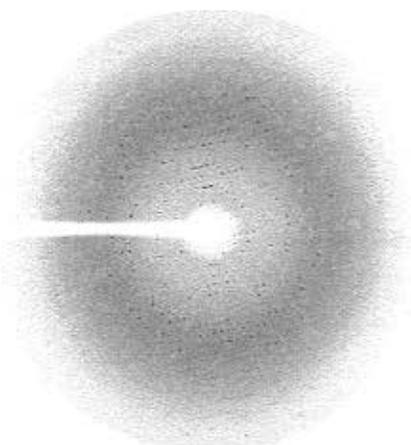
Preparação do cristal para a coleta de dados de raios-x. (a) Montagem do cristal no capilar. O líquido-mãe é colocado para manter a hidratação do cristal. (b) Cabeça goniométrica.

Cada conjunto de reflexões é adquirido girando o cristal em passos angulares fixos para expor os diferentes planos cristalinos à radiação; o intervalo angular total depende da simetria do cristal : quanto maior a simetria, menor o intervalo angular e vice-versa. A exposição pode ser efetuada por tempo ou por dose . Com tempo, o feixe de luz síncrotron perde intensidade e por esta razão a exposição por dose é mais apropriada, pois garante que o mesmo número de fótons alcancem a amostra ao longo de todo o processo de coleta. Os raios difratados são registrados em uma placa de imagem. Este tipo de detetor é cerca de dez vezes mais sensível que um filme convencional e, portanto, reflexões de baixa intensidade podem ser coletadas. O detetor consiste de um material inorgânico à base de fósforo depositado sobre uma base plana. Os fótons que incidem sobre o detetor promovem os elétrons deste tipo de material a níveis energéticos mais altos. A energia armazenada é liberada quando o detetor é iluminado (laser que emite uma linha no vermelho). A intensidade da luz emitida é proporcional ao número de fótons que incidiram no detetor naquela posição. Após a leitura dos dados, o detetor pode ser reutilizado simplesmente iluminando-o com luz branca. A figura

2.8 (a) e (b) mostra, respectivamente, um detetor deste tipo (MAR345) e uma imagem do padrão de difração de um cristal da beta-lactoglobulina coletada com este detetor.



(a)



(b)

Figura 2.8

(a) Placa de imagens da empresa MAR RESEARCH. (b) Imagem dos dados de difração da beta-lactoglobulina bovina coletada com este tipo de placa de imagem

2.2.2 Processamento dos dados

O processamento compreende três etapas : indexação, integração e escalonamento. Na primeira etapa, determina-se a orientação do cristal em relação ao feixe de raios-X, a rede cristalina e os parâmetros da ccla unitária e executa-se a indexação propriamente dita. A indexação consiste em obter três vetores linearmente independentes que descrevam os pontos de difração observados com coeficientes inteiros h, k, l .

Na segunda etapa, avalia-se a intensidade de cada ponto de difração separando as contagens de fundo (*background*) das reflexões. Isto é,

$$I_{h,k,l} = \sum \text{contagens}(\text{ponto de difração}) - \sum \text{contagens}(\text{fundo})$$

Este tipo de separação é exemplificado na figura 2.9 e é feita para todos os pontos de cada imagem.

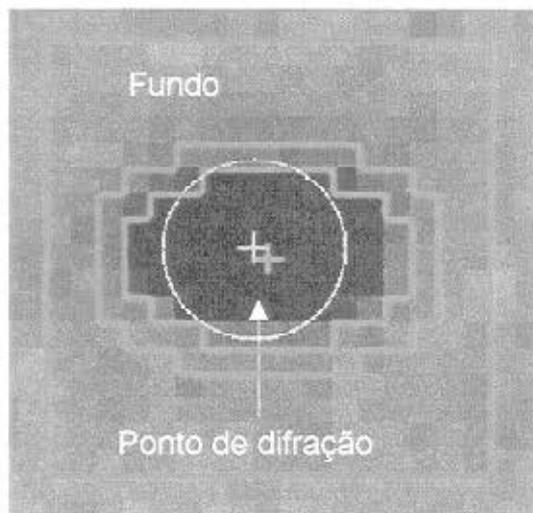


Figura 2.9

Separação das contagens de fundo do ponto de reflexão

Na terceira etapa, é feito o escalonamento entre partes de reflexões comuns a sucessivos pares de imagens e promédio dos dados provenientes de um ou mais cristais, reduzindo as reflexões redundantes a um conjunto de reflexões únicas. O conjunto de dados resultante é formado pelos índices h, k, l , do plano cristalino, pela intensidade, I , correspondente e pelo desvio padrão, $\sigma(I)$, da medida. A escolha da resolução do conjunto de dados é feita nesta fase e consiste em escolher a camada de resolução onde no mínimo 50% das reflexões possuam a razão $I/\sigma(I)$ acima de 2.

Ao final do processamento de dados, devem ser observados a completude e o R_{merge} . O R_{merge} é uma medida da dispersão das reflexões simetricamente relacionadas, matematicamente expresso por:

$$R_{merge} = \frac{\sum_{h,k,l} \sum_i |I_i(h,k,l) - \overline{I(h,k,l)}|}{\sum_{h,k,l} \sum_i I_i(h,k,l)}$$

onde i corresponde ao número de reflexões simétricas.

2.2.3 Coleta e processamento dos dados de difração dos cristais de beta-lactoglobulina bovina

As informações para o conjunto de dados do cristal em pH ácido são para um dímero na unidade assimétrica e estão dispostas na tabela 2.1.

Mistura de isoformas A e B	
Grupo espacial	P1

Parâmetros da cela unitária

a (Å)	37,80	α (°)	68,37
b (Å)	49,88	β (°)	68,03
c (Å)	51,03	γ (°)	76,28

Resolução (Å) nominal	18,4 - 1,38
Número de reflexões únicas	61312
Multiplicidade dos dados	2,1
Completeza (%)	93
Completeza da última camada (%)	77,2
R_{merge} da última camada	72,1
R_{merge} (%)	8,9

Tabela 2.1

Dados estatísticos dos dados de raios-x para mistura de isoformas A e B no pH ácido

Para os cristais em pH básico foram coletados dados em passos de oscilação angular de 1,5° e a exposição feita por dose de raios-x para as variantes A, B e mistura de variantes A/B cujos cristais foram crescidos em torno do pH 7,9. Os dados foram processados com os programas DENZO (Otwinowski, 1996-1997) e SCALEPACK (Otwinowski, 1996-1997). Para estes cristais, a unidade assimétrica contém um monômero. Um sumário das informações estatísticas do processamento destes dados é apresentado na tabela 2.2.

	Variante A	Variante B	Mistura de isoformas A e B
Grupo espacial	C222 ₁ (20)	C222 ₁ (20)	C222 ₁ (20)

Parâmetros da cela unitária

a (Å)	55,47	55,64	55,72
b (Å)	82,11	81,65	82,16
c (Å)	66,76	66,86	66,96

Resolução (Å)	12,92 - 2,00	14,98 - 1,95	15,0 - 2,03
No. de reflexões únicas	9631	9834	9235
Multiplicidade dos dados	2,8	3,0	3,0
Completeza (%)	90,9	86,4	90,3
Completeza da última camada (%)	91,2	90,9	92,0
R_{merge} da última camada	46,3	44,4	10,7
R_{merge} (%)	6,6	8,5	3,4

Tabela 2.2

Dados estatísticos dos dados de raios-x para vários cristais da beta-lactoglobulina bovina crescidos no pH 7,9

2.3 Princípios teóricos de difração de raios-X

2.3.1 Periodicidade do cristal e a análise de Fourier

Um cristal é um arranjo periódico de subunidades idênticas no espaço. Uma subunidade é a menor divisão de um cristal que possui as propriedades que o caracterizam e é denominada cela unitária. Matematicamente, a estrutura cristalina é descrita por uma rede com uma base em cada ponto da rede.

A rede é definida por três vetores \vec{a} , \vec{b} , \vec{c} , de modo que cada ponto de posição, \vec{r} , é descrito por uma combinação linear destes vetores com coeficientes inteiros (definição de uma rede de Bravais), ou seja, $\vec{r} = u\vec{a} + v\vec{b} + w\vec{c}$ (u, v, w inteiros). Isto significa que esses vetores devem ser tais que sob uma translação, \vec{T} , igual a qualquer vetor da rede, o arranjo atômico permanece o mesmo (invariância do cristal).

A base é o conteúdo de cada ponto da rede, ou seja, um átomo, um grupo de átomos ou uma molécula protéica.

A invariância frente a translação do cristal implica invariância das propriedades físicas locais sob a mesma translação. Portanto, dada uma função $n(\vec{r})$ devemos ter:

$$n(\vec{r}) = n(\vec{r} + \vec{T}) \quad (2.1)$$

Devido a esta periodicidade, esta função pode ser expandida em termos de uma série de Fourier. Uma série de Fourier pode ser escrita em termos de senos e cossenos ou de uma forma mais compacta através de ondas planas do tipo $\exp(i\vec{G} \cdot \vec{r})$ e deste modo:

$$n(\vec{r}) = \sum_{\vec{G}} n_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \quad (2.2)$$

onde \vec{G} pertence ao espaço de Fourier (ou espaço recíproco).

Para que a condição (2.1) seja satisfeita os vetores \vec{G} da expansão (2.2) não são quaisquer, mas somente aqueles que mantêm a periodicidade da rede sob uma translação \vec{T} . Isto implica que na expressão $n(\vec{r} + \vec{T}) = \sum_{\vec{G}} n_{\vec{G}} \exp(i\vec{G} \cdot \vec{r}) \exp(i\vec{G} \cdot \vec{T})$ o termo $\exp(i\vec{G} \cdot \vec{T})$ deve ter o valor unitário, isto é, $\vec{G} \cdot \vec{T} = n2\pi$ (2.3). Dado que o vetor de translação é uma combinação linear de um conjunto de vetores com coeficientes inteiros, os vetores \vec{G} devem ser expandidos em uma base do espaço recíproco a coeficientes inteiros a fim de que (2.3) seja aplicável. Tal base de vetores é convenientemente dada por:

$$\vec{a}^* = \frac{2\pi\vec{b} \times \vec{c}}{\vec{a} \cdot \vec{b} \times \vec{c}} \quad \vec{b}^* = \frac{2\pi\vec{c} \times \vec{a}}{\vec{a} \cdot \vec{b} \times \vec{c}} \quad \vec{c}^* = \frac{2\pi\vec{a} \times \vec{b}}{\vec{a} \cdot \vec{b} \times \vec{c}} \quad (2.4)$$

com $\vec{G} = h\vec{a}^* + k\vec{b}^* + l\vec{c}^*$ (h, k, l inteiros).

2.3.2 Condições de difração

A difração de raios-x pode ser analisada sob o ponto de vista que considera cada átomo como um espalhador independente. Os elétrons dos átomos são os responsáveis pelo espalhamento das ondas eletromagnéticas que incidem no cristal. Portanto, é importante o conhecimento da distribuição dos elétrons no cristal. A distribuição dos elétrons no cristal é representada pela densidade eletrônica, a qual pode ser expandida em uma série de Fourier por ser uma propriedade periódica, de acordo com que foi apresentado no tópico anterior. Assim,

$$\rho(\vec{r}) = \sum_{\vec{G}} F(\vec{G}) \exp(i\vec{G} \cdot \vec{r}) \quad (2.5)$$

Da teoria de interferência de ondas eletromagnéticas sabemos que as ondas espalhadas por elementos separados por um vetor \vec{r} diferem em fase por um fator $\exp[i(\mathbf{k} - \mathbf{k}') \cdot \vec{r}]$, onde \mathbf{k} e \mathbf{k}' são os vetores de onda da onda eletromagnética incidente e espalhada

respectivamente. Dependendo da diferença de fase, essas ondas podem interferir construtiva ou destrutivamente.

Deste modo, para determinarmos a amplitude da onda difratada por um cristal devemos levar em consideração a distribuição eletrônica e o fator de fase relativo a um ponto de referência no cristal. Matematicamente, isto é traduzido por:

$$A = \int_{\text{vol. cristal}} \rho(\vec{r}) \exp(-i\Delta\vec{k} \cdot \vec{r}) dV \quad (2.6)$$

onde $\Delta\vec{k} = \vec{k}' - \vec{k}$.

Substituindo as componentes de Fourier da equação (2.5) no lugar de $\rho(\vec{r})$ na equação (2.6) obtemos:

$$A = \sum_{\vec{G}} \int_{\text{vol. cristal}} F(\vec{G}) \exp[i(\vec{G} - \Delta\vec{k}) \cdot \vec{r}] dV \quad (2.7)$$

Quando $\Delta\vec{k} = \vec{G}$ (2.8) a exponencial adquire o valor unitário e a amplitude A assume um valor diferente de zero, caso contrário, para valores de $\Delta\vec{k}$ diferentes de \vec{G} o valor de A é nulo ou desprezível¹. Logo, a difração ocorre quando a diferença entre o vetor de onda eletromagnético da onda espalhada e o da incidente for um vetor da rede recíproca.

Um outro modo de expressar (2.8) é tomar o produto interno de ambos os lados da igualdade pelos vetores \vec{a} , \vec{b} , \vec{c} da rede direta lembrando as relações (2.4) apresentadas no tópico anterior:

$$\begin{aligned} \vec{a} \cdot \Delta\vec{k} &= \vec{a} \cdot \vec{G} = 2\pi h \\ \vec{b} \cdot \Delta\vec{k} &= \vec{b} \cdot \vec{G} = 2\pi k \\ \vec{c} \cdot \Delta\vec{k} &= \vec{c} \cdot \vec{G} = 2\pi l \end{aligned} \quad (2.9)$$

O conjunto de expressões dadas em (2.9) é chamado equações de Laue. Podemos entender (2.9) como “reflexões” por planos imaginários no cristal reescrevendo as equações de Laue convenientemente:

$$\frac{\vec{a}}{h} \cdot \Delta\vec{k} = 2\pi; \quad \frac{\vec{b}}{k} \cdot \Delta\vec{k} = 2\pi; \quad \frac{\vec{c}}{l} \cdot \Delta\vec{k} = 2\pi \quad (2.10)$$

Notemos que os vetores $\frac{\vec{a}}{h}$, $\frac{\vec{b}}{k}$, $\frac{\vec{c}}{l}$ possuem a mesma projeção sobre $\Delta\vec{k}$. Logo, os três pontos definidos por esses vetores formam um plano perpendicular a $\Delta\vec{k}$ caracterizado pelos índices h, k, l (figura 2.10). Considerando o espalhamento elástico, então, $|\vec{k}| = |\vec{k}'|$ e assim verificamos geometricamente que o módulo de $\Delta\vec{k}$ é igual a $2|\vec{k}|\sin\theta$ (figura 2.11). Sendo d a distância entre planos sucessivos e as projeções sobre $\Delta\vec{k}$ iguais a $\frac{2\pi}{|\Delta\vec{k}|}$ concluímos que

$$d = \frac{2\pi}{|\Delta\vec{k}|}, \text{ ou melhor, } d = \frac{2\pi}{2|\vec{k}|\sin\theta} \quad (2.11). \text{ Substituindo } |\vec{k}| = \frac{2\pi}{\lambda} \text{ em (2.11) chegamos}$$

a:

$$2d\sin\theta = \lambda \quad (2.12)$$

que é conhecida como lei de Bragg. A expressão indica que toda vez que a diferença de caminho $2d\sin\theta$ entre ondas de planos sucessivos for um número inteiro de λ , ocorre interferência construtiva. Portanto, descrever o padrão de difração por “reflexões” de ondas ou por teoria de difração é equivalente.

¹ A integral do produto de funções pares é nula se o intervalo de integração corresponder a um número inteiro de períodos.

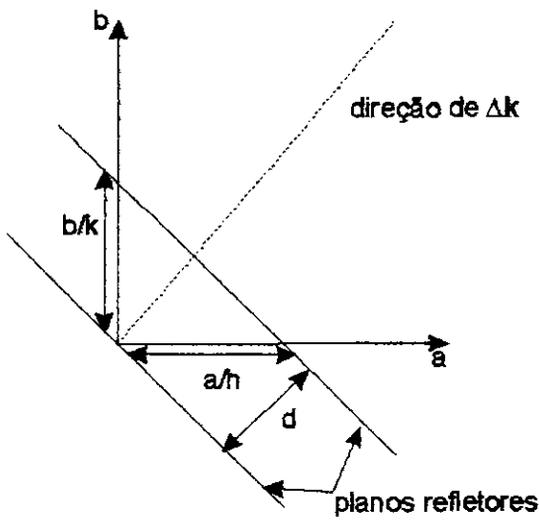


Figura 2.10

Planos de reflexão imaginários que cruzam os eixos da cela unitária em a/h , b/k , c/l (o eixo c é perpendicular a página).

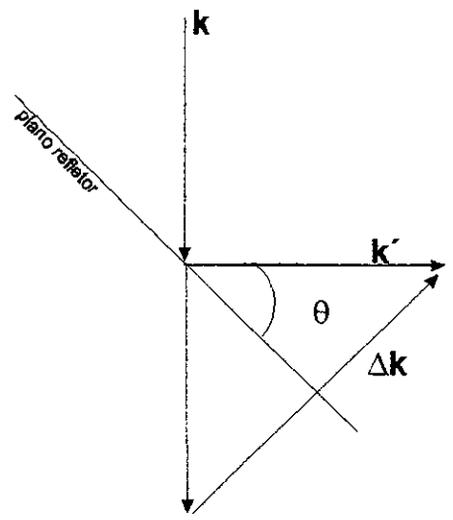


Figura 2.11

Relação geométrica entre os vetores \vec{k} , \vec{k}' e $\Delta\vec{k}$.

2.3.3 Fator de estrutura

Retornando à equação (2.6) podemos reescrevê-la com a condição de difração

$\Delta\vec{k} = \vec{G}$ e considerar um cristal constituído de N celas unitárias:

$$A = N \int_{\text{vol. cela}} \rho(\vec{r}) \exp(-i\vec{G} \cdot \vec{r}) dV = NF(\vec{G}) \quad (2.13)$$

onde $F(\vec{G})$ é chamado fator de estrutura, cuja integral é feita sobre uma cela unitária com $\vec{r} = 0$ em um dos vértices.

A densidade eletrônica escrita de modo apropriado fornece uma interpretação conveniente ao fator de estrutura. Se, em um ponto \vec{r} , ela for escrita como sendo a contribuição da densidade eletrônica, ρ_j , de cada átomo j localizado no ponto \vec{r}_j com relação à origem da célula unitária teremos:

$$\rho(\vec{r}) = \sum_j \rho_j (\vec{r} - \vec{r}_j) \quad (2.14)$$

e o fator de estrutura passa a ser expresso por:

$$F(\vec{G}) = \int \sum_j \rho_j(\vec{r} - \vec{r}_j) \exp(-i\vec{G} \cdot \vec{r}) dV \quad (2.15)$$

ou ainda, se fizermos $\vec{R} = \vec{r} - \vec{r}_j$ teremos :

$$F(\vec{G}) = \sum_j \exp(-i\vec{G} \cdot \vec{r}_j) \int \rho_j(\vec{R}) \exp(-i\vec{G} \cdot \vec{R}) dV \quad (2.16)$$

O termo expresso pela integral acima é denominado fator de forma, f_j , do átomo j .

O fator de forma é uma propriedade atômica e nos dá informação sobre o poder de espalhamento de um átomo. Em sua expressão há dependência da distribuição dos elétrons no átomo através de ρ_j e a dependência do comprimento de onda e do ângulo de espalhamento através do termo exponencial. Combinando $\vec{G} = h\vec{a}^* + k\vec{b}^* + l\vec{c}^*$ com $\vec{r}_j = x_j\vec{a} + y_j\vec{b} + z_j\vec{c}$ ($0 \leq x_j \leq 1, 0 \leq y_j \leq 1, 0 \leq z_j \leq 1$) obtemos para o fator de estrutura:

$$F(h, k, l) = \sum_j f_j \exp[2\pi i(hx_j + ky_j + lz_j)] \quad (2.17)$$

Pela expressão (2.17) observamos que o fator de estrutura é uma série de Fourier onde cada termo representa a contribuição de cada átomo à reflexão h, k, l . A contribuição depende da natureza do átomo (representado por f_j) e de sua posição x_j, y_j, z_j na cela unitária. Em função desta descrição onde o fator de estrutura possui amplitude e fase sua representação, deste ponto em diante, será em negrito usualmente destinada para fasores e vetores.

Correções ao fator de forma surgem devido à vibração atômica a qual diminui a intensidade do espalhamento. A correção é feita através de um fator multiplicativo com

dependência exponencial. Se a vibração for a mesma em todas as direções a vibração é dita isotrópica² e o fator de correção é dado por (Blundell, T.L. & Johnson, L.N., 1976) :

$$\exp\left(-B\frac{\sin^2\theta}{\lambda^2}\right) \quad (2.18)$$

onde $B = 8\pi^2\overline{u^2}$ é chamado de fator de temperatura. O fator de correção está associado a desordem dinâmica dos átomos em consequência da vibração atômica. Entretanto, há um outro tipo de desordem em cristais de proteínas: a desordem estática. Esta classe de desordem surge em função da diferença na orientação e posicionamento de moléculas (ou apenas partes delas) ao longo do cristal e não é possível discernir seu efeito em relação a desordem dinâmica sobre os dados de difração.

Por outro lado, através da equação (2.13) concluímos que a densidade eletrônica é a transformada de Fourier do fator de estrutura e é dada por:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l \mathbf{F}(h, k, l) \exp(2\pi i(hx + ky + lz)) \quad (2.19)$$

A intensidade do espalhamento é proporcional a $|\mathbf{F}(h, k, l)|^2$ e, portanto, a informação sobre a fase do fator de estrutura é perdida. Isto é o chamado "*problema das fases*" em cristalografia. No entanto, em cristalografia de proteínas, existem métodos que permitem o acesso a uma estimativa inicial destas fases, um destes é o assunto do próximo capítulo.

² Caso contrário, a vibração é anisotrópica e o deslocamento quadrático médio dos átomos é descrito por um tensor.

Determinação da estrutura cristalográfica da beta-lactoglobulina bovina por substituição molecular

Estruturas tridimensionais de macromoléculas biológicas são determinadas com o objetivo de compreender a base molecular de suas funções. Determinar uma estrutura significa obter coordenadas x , y , z para os átomos presentes na molécula dentro do arranjo cristalino. Não é possível obter experimentalmente informações sobre as fases dos fatores de estrutura. Assim sendo, é inviável localizar os átomos diretamente. Entretanto, se a macromolécula em questão possui homologia estrutural com uma outra cuja estrutura já tenha sido resolvida, é possível empregar a estrutura conhecida como um modelo inicial de fases para aquela desconhecida. Esta técnica é denominada substituição molecular. Caso o pré-requisito acima não seja satisfeito, outras duas técnicas podem ser aplicadas: substituição isomórfica e dispersão anômala múltipla. Em ambas, átomos com número atômico elevado são incorporados ao cristal de proteína de forma que suas propriedades de espalhamento são utilizadas para calcular a posição dos átomos presentes na proteína.

Neste capítulo serão apresentados os princípios básicos e a aplicação da substituição molecular, no caso da beta-lactoglobulina.

3.1 Introdução

Como descrito no capítulo anterior, a função de densidade eletrônica descreve a distribuição de elétrons dentro do cristal e pode ser representada, como todas as funções periódicas, por uma série de Fourier:

$$\rho(x, y, z) = \frac{1}{V} \sum_{h,k,l} |F(h, k, l)| e^{-2\pi i(hx+ky+lz)+i\alpha(h,k,l)} \quad (3.1)$$

Assim, com a equação 3.1, podemos reconstruir a estrutura cristalina a partir do módulo dos fatores de estrutura, $|\mathbf{F}(h,k,l)|$, e de suas fases correspondentes, $\alpha_{h,k,l}$. Entretanto, os únicos dados experimentais disponíveis são os módulos ao quadrado dos fatores de estrutura, isto é, $|\mathbf{F}(h,k,l)|^2$, e não suas fases¹. Mas felizmente, podemos obter alguma informação sobre a estrutura cristalina graças à função de Patterson, a qual depende somente de $|\mathbf{F}(h,k,l)|^2$ e será apresentada na seção 3.2.1. Isto não significa que esta função supra a necessidade de obtermos um conjunto de fases, ainda que estimadas, para o cálculo de $\rho(x,y,z)$.

Para que a estimativa de fases seja feita com o método de substituição molecular, é necessário que a molécula biológica em estudo possua homologia com uma outra cuja estrutura tenha sido determinada. Este pré-requisito fundamenta-se na observação de que moléculas biológicas com identidade seqüencial possuem similaridade estrutural. Podemos então adotar as fases da molécula cuja estrutura é conhecida como uma aproximação inicial para resolver a estrutura de interesse. Este procedimento de adoção não é simplesmente associativo uma vez que as fases do fator de estrutura dependem da posição dos átomos em relação à cela unitária. Por este motivo, a molécula de estrutura conhecida, escolhida para este fim, é denominada modelo de busca. Portanto, para que as informações de fase provenientes do modelo sejam utilizadas, é necessário que este modelo seja orientado e posicionado no arranjo cristalino da molécula de estrutura desconhecida. Isto envolve o conhecimento de seis variáveis : três de rotação ($\theta_1, \theta_2, \theta_3$) e três de posição (x,y,z).

Computacionalmente, a determinação simultânea de seis parâmetros seria inviável. No entanto, com o uso da função de Patterson podemos agilizar este processo tornando independentes as variáveis de rotação e de translação. Esta função permite separar os vetores interatômicos em intramoleculares e intermoleculares de forma que definições apropriadas de funções de rotação e de translação façam uso das características destes vetores.

¹ Os dados experimentais são as intensidades, I_{hkl} , para cada reflexão h,k,l , onde $|\mathbf{F}(h,k,l)|^2 \propto I$.

Finalmente, após este árduo processo computacional, tem-se um modelo cujas fases podem ser utilizadas como estimativas para o cálculo de $\rho(x, y, z)$ do arranjo cristalino de interesse.

3.2 Princípios teóricos da substituição molecular

Os princípios que serão apresentados nos tópicos a seguir têm o objetivo de esclarecer, de maneira qualitativa, os aspectos matemáticos empregados nas etapas de substituição molecular. A apresentação será feita seguindo os trabalhos de Rossmann e Blow (1962) para a função de rotação e, para a função de translação, os de Crowther e Blow (1967)

Esquemáticamente, a figura 3.1 (a) e (b) representa estes estágios . Em (a), orientação do modelo através de três ângulos, provenientes da função de rotação, que irão definir uma transformação [C]; em (b), posicionamento do modelo determinado pelo vetor, \mathbf{t} , oriundo da função de translação, de acordo com a simetria cristalográfica do arranjo cristalino em estudo.

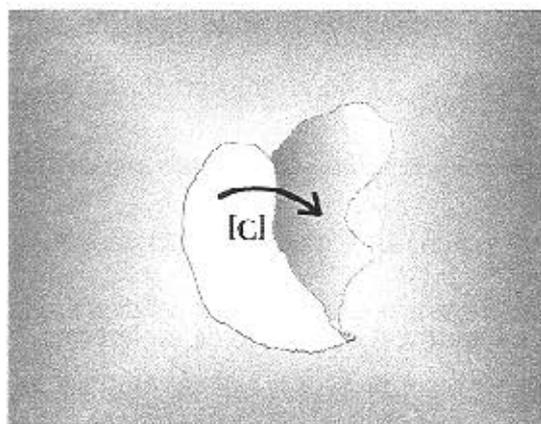
3.2.1 Função de Patterson

A função de Patterson é definida como o produto de densidades eletrônicas calculadas nos extremos de um vetor cujas coordenadas fracionais são u, v', w .

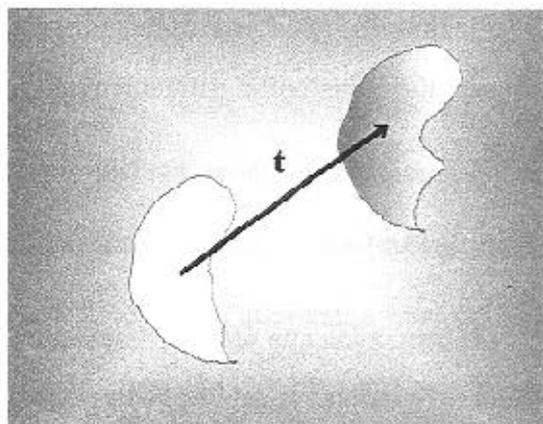
Matematicamente :

$$P(u, v', w) = \int_0^V \rho(x, y, z) \rho(x + u, y + v', z + w) dV \quad (3.2)$$

Pela equação 3.2, a função $P(u, v', w)$ atingirá valores diferentes de zero quando o vetor (u, v', w) representar uma distância entre átomos. Esta expressão relaciona-se somente com $|\mathbf{F}(h, k, l)|^2$ e isto pode ser visto substituindo a equação 3.1 em 3.2 o que resulta em:



(a)



(b)

Figura 3.1

Fases da substituição molecular. (a) rotação, (b) translação

$$P(u, v', w) = \frac{1}{V} \sum_{h,k,l} |F(h, k, l)|^2 e^{2\pi(hu + kv' + lw)} \quad (3.3)$$

Em suma, a função de Patterson fornece as distâncias vetoriais entre os átomos (mas não suas coordenadas) somente com os dados parciais dos quais dispomos. Cada vetor encontrado é colocado em um mapa designado mapa de Patterson. A figura 3.2 (a) ilustra este mapa através de uma estrutura bidimensional composta por três átomos cujos vetores interatômicos são: $\vec{1,2}$, $\vec{1,3}$, $\vec{2,3}$, $\vec{2,1}$, $\vec{3,1}$, $\vec{3,2}$. Abaixo, está o correspondente mapa de Patterson onde podem ser observadas duas características: a primeira, é que o mapa é simétrico em relação à origem e a segunda, que o pico pronunciado na origem ocorre quando $u = 0$, $v' = 0$ e $w = 0$, ou seja, ele é formado pela soma dos quadrados da densidade eletrônica de cada átomo presente na estrutura.

Os vetores classificam-se em intramoleculares e intermoleculares. Os vetores intramoleculares estão próximos à origem do mapa e são semelhantes em estruturas aproximadamente similares. Ao contrário, os vetores intermoleculares não estão tão próximos à origem e relacionam-se com simetria cristalográfica e não-cristalográfica.

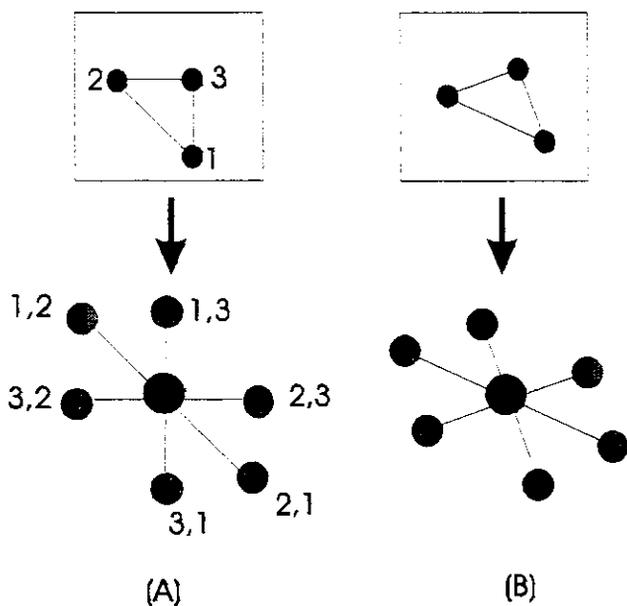


Figura 3.2

Estrutura bidimensional de um arranjo de três átomos e abaixo o mapa de Patterson correspondente : (a) arranjo atômico com uma determinada orientação, (b) o mesmo arranjo rotado de 20°.

Um aspecto importante do mapa de Patterson é a sua relação intrínseca com a orientação do arranjo atômico em relação à cela unitária. Isto significa que o mesmo arranjo pode produzir mapas distintos entre si (figura 3.2 (b)). Contudo, a relação entre estes mapas pode ser encontrada (via comparação) e conseqüentemente, a relação entre as possíveis orientações deste arranjo. Igualmente, em estruturas similares confrontam-se mapas de Patterson com a finalidade de encontrar a orientação entre elas e subseqüentemente a posição relativa a outras moléculas no cristal.

3.2.2 Função de rotação

A função de rotação pode ser utilizada para encontrar a orientação tanto entre moléculas dentro da mesma unidade assimétrica (auto-rotação) do cristal quanto em moléculas pertencentes a diferentes redes cristalinas (rotação cruzada). Posto que os vetores de Patterson intramoleculares de duas moléculas idênticas (ou quase idênticas) relacionam-se por uma operação de rotação, somente estes vetores são necessários neste tipo de análise. Portanto, a procura por esta relação angular é o fundamento desta etapa na substituição molecular.

Seja \mathbf{r} a posição de um ponto em relação a três eixos cristalográficos \mathbf{a} , \mathbf{b} , \mathbf{c} de modo que $\mathbf{r} = x_1\mathbf{a} + x_2\mathbf{b} + x_3\mathbf{c}$. Esta mesma posição relativa a um outro conjunto de eixos \mathbf{a}' , \mathbf{b}' , \mathbf{c}' é expressa por ²:

$$\mathbf{r}' = [C]\mathbf{r} \quad (3.4)$$

onde $[C]$ é a matriz de rotação em torno da origem que transforma estes sistemas entre si.

Assim sendo, por definição, a função de rotação é :

$$R(\theta_1, \theta_2, \theta_3) = \iiint_U P_o(\mathbf{u})P_m(\mathbf{u}')dV \quad (3.5)$$

onde U corresponde ao volume ocupado pelos vetores de Patterson intramoleculares, θ_1 , θ_2 , θ_3 são os ângulos de rotação, P_o representa a função de Patterson construída a partir dos dados experimentais e P_m é a função de Patterson construída por interpolação dos fatores de estrutura do modelo, nos pontos $\mathbf{u}' = [C]\mathbf{u}$ ³.

Pela expressão 3.5, podemos observar que $R(\theta_1, \theta_2, \theta_3)$ atingirá seu valor máximo quando forem encontrados três ângulos que proporcionem a superposição entre os picos correspondentes nos mapas de Patterson do modelo e dos dados experimentais. A figura 3.2 (a) e (b) exemplifica esta situação no caso bidimensional: os mapas de Patterson correlacionam-se por um rotação de 20°, o que implica uma rotação de mesmo valor na estrutura (b) para atingir a mesma orientação apresentada em (a).

Embora o sistema angular polar possa ser visualizado diretamente (figura 3.3 (a)), o sistema euleriano (figura 3.3 (b)) é preferido porque exhibe a simetria da função de rotação de um modo conveniente. A relação entre estes dois sistemas é expressa por:

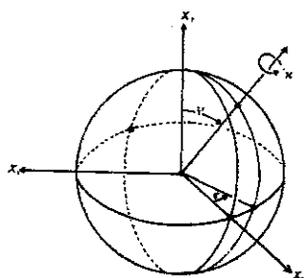
² Por questão de simplicidade o módulo do vetor que representa a translação entre a origem dos dois sistemas está sendo considerado nulo.

³ Rotações de vetores no espaço recíproco resultam normalmente em vetores com coeficientes não inteiros de forma que o valor do fator de estrutura (conhecido apenas para vetores com coeficientes inteiros) necessita ser calculado por interpolação.

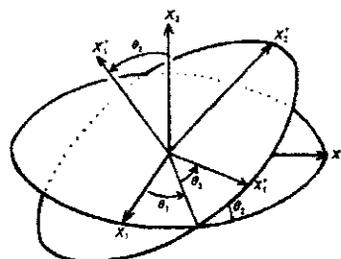
$$\cos\left(\frac{\chi}{2}\right) = \cos\left(\frac{\theta_2}{2}\right) \cos\left(\frac{\theta_1 + \theta_3}{2}\right)$$

$$\tan \varphi = -\cot\left(\frac{\theta_2}{2}\right) \operatorname{sen}\left(\frac{\theta_1 + \theta_3}{2}\right) \sec\left(\frac{\theta_1 - \theta_3}{2}\right)$$

$$\cos \varphi \tan \psi = \cot\left(\frac{\theta_1 - \theta_3}{2}\right)$$



(a)



(b)

Figura 3.3

Sistema angular usado na matriz [C] da função de rotação. (a) polar, (b) euleriano

3.2.3 Função de translação

A função de translação consiste em definir a posição do modelo dentro da cela unitária, considerando as posições ocupadas por moléculas simetricamente relacionadas. Como esta fase depende da posição relativa entre moléculas, somente vetores intermoleculares são essenciais.

Sejam duas moléculas, mol0 e mol1, cujas origens locais estejam nas posições O_1 e O_2 , relacionadas por uma operação de simetria ($Ar + d$). A função de Patterson entre estas duas moléculas é dada por

$$P_{r,s}(\mathbf{u}) = \int \rho_r(\mathbf{r}) \rho_s(\mathbf{r} + \mathbf{u}) d\mathbf{r} \quad (3.6)$$

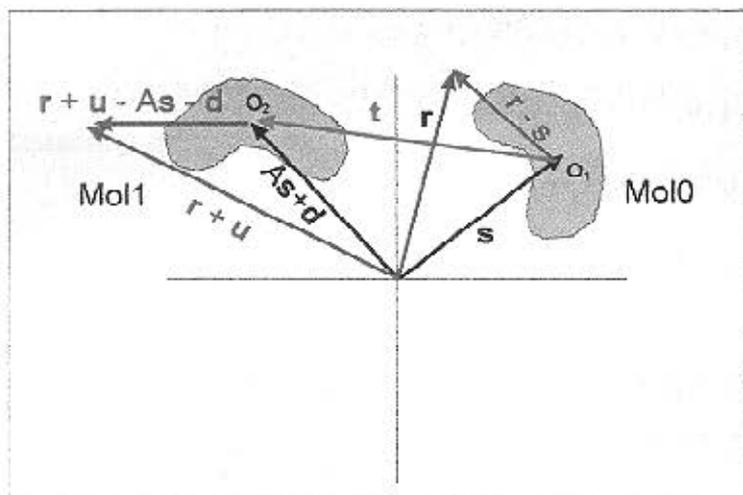


Figura 3.4

Vetores utilizados na função de translação

Se as densidades $\rho_r(\mathbf{r})$ e $\rho_s(\mathbf{r})$ forem expressas em termos da densidade eletrônica, ρ_M , da primeira molécula (figura 3.4) cuja origem local é O_1 , a expressão 3.6 transforma-se em:

$$P_{r,s}(\mathbf{u}, \mathbf{s}) = \int_V \rho_M(\mathbf{r} - \mathbf{s}) \rho_M[\mathbf{A}^{-1}(\mathbf{r} + \mathbf{u} - \mathbf{A}\mathbf{s} - \mathbf{d})] d\mathbf{r} \quad (3.7)$$

Expandindo ρ_M como uma série de Fourier em termos dos fatores de estrutura F_M tem-se:

$$P_{r,s}(\mathbf{u}, \mathbf{s}) = \sum_{\mathbf{H}} F_M(\mathbf{H}) F_M(-\mathbf{H}\mathbf{A}) \exp[2\pi i \mathbf{H} \cdot (\mathbf{s} - \mathbf{A}\mathbf{s} - \mathbf{d})] \exp(2\pi i \mathbf{H} \cdot \mathbf{u}) \quad (3.8)$$

definindo $\mathbf{t} = -\mathbf{s} + \mathbf{A}\mathbf{s} + \mathbf{d}$ e que $F_M(-\mathbf{H}\mathbf{A}) = F_M^*(\mathbf{H}\mathbf{A})$ teremos:

$$P_{r,s}(\mathbf{u}, \mathbf{t}) = \sum_{\mathbf{H}} F_M(\mathbf{H}) F_M^*(\mathbf{H}\mathbf{A}) \exp(-2\pi i \mathbf{H} \cdot \mathbf{t}) \exp(2\pi i \mathbf{H} \cdot \mathbf{u}) \quad (3.9)$$

A função de translação é então definida por :

$$T(\mathbf{t}) = \int_V P_{r,s}(\mathbf{u}, \mathbf{t}) P_o(\mathbf{u}) d\mathbf{u} \quad (3.10)$$

onde $P_{r,s}(\mathbf{u}, \mathbf{t})$ é a função de Patterson do modelo de referência e as moléculas simetricamente relacionadas e $P_o(\mathbf{u})$ é a função de Patterson observada. O máximo da função 3.10 ocorrerá quando \mathbf{t} for igual a um vetor intermolecular.

A substituição molecular é monitorada pelo coeficiente de correlação e pelo fator R definidos abaixo:

$$C = \frac{\sum_{h,k,l} (|\mathbf{F}_{obs}|^2 - \overline{|\mathbf{F}_{obs}|^2})(|\mathbf{F}_{calc}|^2 - \overline{|\mathbf{F}_{calc}|^2})}{[\sum_{h,k,l} (|\mathbf{F}_{obs}|^2 - \overline{|\mathbf{F}_{obs}|^2})^2 \sum_{h,k,l} (|\mathbf{F}_{calc}|^2 - \overline{|\mathbf{F}_{calc}|^2})^2]^{1/2}}$$

e

$$R_{factor} = \frac{\sum_{h,k,l} ||F_{obs}| - k|F_{calc}||}{\sum_{h,k,l} |F_{obs}|}$$

onde $|F_{calc}|$ e $|F_{obs}|$ são respectivamente os módulos dos fatores de estrutura do modelo (na posição relacionada com \mathbf{t}) e experimental.

3.3 Aplicação da substituição molecular à beta-lactoglobulina

O método de substituição molecular foi aplicado, com o programa AMORE (CCP4), ao conjunto de dados abaixo enumerado:

- 1) Mistura de isoformas A e B, forma cristalina X (P1), pH 6,5
- 2) Isoforma B, forma cristalina Y (C222₁), pH 7,9
- 3) Isoforma A, forma cristalina Y (C222₁), pH 7,9

Os modelos de busca utilizados neste procedimento foram :

I. Mistura de isoformas A e B da beta-lactoglobulina resolvida a 1,8Å, cristal crescido na região de pH ácido com um dímero na unidade assimétrica, grupo espacial P1 (Brownlow *et al.*, 1997).

Parâmetros de cela $a = 37,8 \text{ \AA}$, $b = 49,6 \text{ \AA}$, $c = 56,6 \text{ \AA}$, $\alpha = 123,4^\circ$, $\beta = 97,3^\circ$, $\gamma = 103,7^\circ$;

II. Mistura de isoformas A e B da beta-lactoglobulina resolvida a 2,0Å, cristal crescido na região de pH básico com um monômero na unidade assimétrica, grupo espacial C222₁ (IP, dados não publicados).

Parâmetros de cela $a = 82,4 \text{ \AA}$, $b = 55,4 \text{ \AA}$, $c = 66,4 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$, $\gamma = 90^\circ$;

III. Modelo incompletamente refinado da isoforma B da beta-lactoglobulina de resolução 1,95 Å, cristal crescido em torno do pH 7,9 com um monômero na unidade assimétrica, grupo espacial C222₁.

Parâmetros de cela $a = 55,64 \text{ \AA}$, $b = 81,65 \text{ \AA}$, $c = 66,86 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$, $\gamma = 90^\circ$

Para o conjunto de dados 1 foram utilizadas reflexões entre 3,3 Å e 10 Å e um raio de Patterson de 19 Å para o cálculo da função de rotação onde I foi usado como modelo de busca. Os valores angulares encontrados estão dispostos na tabela 3.1. Neste caso não houve a necessidade de translação por se tratar do grupo espacial P1.

Para os dados 2 e 3 foram empregadas reflexões entre 8,0 Å e 2,0 Å e um raio de patterson de 20 Å para calcular as funções de rotação e translação. Os respectivos modelos de busca foram II e III. Os picos das respectivas funções de rotação foram usados na função de translação e os resultados estão apresentados na tabela 3.1

Após a substituição molecular todos os modelos foram submetidos a refinamento de corpo rígido (quinta etapa do programa denominada "*fiting*") para os parâmetros θ_1 , θ_2 , θ_3 , x , y , z e fatores de temperatura. Os valores finais encontrados, após esta etapa, para o coeficiente de correlação e R_{factor} são mostrados na tabela 3.2.

Nota: O presente projeto envolve quatro estruturas. Três foram resolvidas por substituição molecular e apresentadas neste capítulo. A quarta estrutura corresponde a um

crystal da mistura de isoformas A e B crescido em torno do pH 7,9 cujo grupo espacial é C222₁. Os parâmetros de rede são a = 55.72Å, b = 82,16Å, c = 66.96Å. Devido à semelhança entre os parâmetros de cela do conjunto de dados e do candidato a modelo de fases (modelo II, apresentado no tópico 3.3) foi possível, ao invés de substituição molecular, efetuar algumas mudanças apropriadas no arquivo das coordenadas do modelo e no arquivo de reflexões dos dados. As modificações consistiram em inverter os índices h e k (h→k e k→h) no arquivo de dados com o programa DATAMAN (Kleywegt & Jones, 1996), retirar os parâmetros de cela bem como a matriz de fracionalização presentes no cabeçalho do modelo e substituí-los pelos novos valores correspondentes aos dos dados experimentais.

Conjunto de dados	modelo	resolução	θ_1	θ_2	θ_3	pico da solução		r.m.s	T_x	T_y	T_z	pico da solução		r.m.s
						1o.	2o.					1o.	2o.	
1	I	10,0 - 3,3	76,81	64,35	351,65	68,6	64,9	3,8	-	-	-	-	-	-
2	II	8,0 - 2,0	25,44	95,27	210,65	33,1	13,6	3,6	0,26	0,34	0,36	10,6	6,2	1,4
3	III	8,0 - 2,0	147,76	72,69	37,51	28,6	15,6	3,5	0,24	0,34	0,13	10,4	7,0	1,4

Tabela 3.1

Resultados da substituição molecular para: 1) mistura de isoformas A e B, forma cristalina X (P1), pH 6,5; 2) isoforma B, forma cristalina Y (C222₁), pH 7,9; 3) isoforma A, forma cristalina Y (C222₁), pH 7,9

	Coefficiente de correlação (%)	R_{factor} (%)
1	81,1	27,9
2	87,8	28,8
3	91,7	27,3

Tabela 3.2

Valores para o coeficiente de correlação e R_{factor} ao final do processo de substituição molecular

Refinamento das estruturas cristalográficas da beta-lactoglobulina bovina

O refinamento de uma estrutura cristalográfica consiste em ajustar as coordenadas x_i, y_i, z_i e o fator térmico de vibração, B_i , de cada átomo i do modelo inicial proposto, a fim de explicar o conjunto de dados de difração. Este procedimento é realizado através de processos interativos e exigem grande demanda computacional. Essencialmente, minimizam-se as diferenças entre os módulos dos fatores de estrutura observado e calculado até que o processo satisfaça o critério de convergência. Neste capítulo serão apresentados os aspectos básicos do refinamento bem como os protocolos aplicados no refinamento das estruturas cristalográficas da beta-lactoglobulina bovina.

4.1 Aspectos básicos do refinamento de estruturas

Na fase de refinamento, as coordenadas do modelo são modificadas a fim de minimizar a diferença entre os módulos dos fatores de estrutura observado, $|F_{obs}|$, e calculado a partir do modelo, $|F_{calc}|$. Isto equivale a dizer que estamos procurando por um modelo que explique adequadamente a densidade eletrônica do cristal. Este procedimento se faz por pequenos incrementos nas coordenadas atômicas do modelo e, normalmente, é necessária a intervenção manual no ajuste de conjuntos de átomos à densidade eletrônica.

O ajuste manual de átomos é o resultado da interpretação de um mapa de densidade eletrônica. Este tipo de mapa é construído a partir dos módulos dos fatores de estrutura observados e da fase estimada, α_{calc} , neste caso por substituição molecular, segundo a expressão abaixo:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l w_{hkl} |F_{obs}(h, k, l)| \exp[-2\pi i(hx + ky + lz - \alpha_{calc}(h, k, l))] \quad (4.1)$$

onde w_{hkl} é o peso dado a cada termo da série e que avalia a confiabilidade da fase associada.

Interpretar um mapa de densidade eletrônica assim construído significa incorporar o traçado de uma cadeia polipeptídica como ilustrado na figura 4.1.

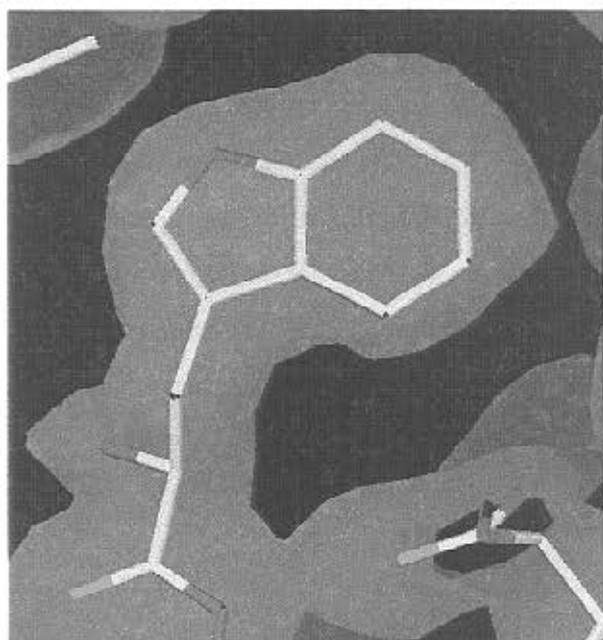


Figura 4.1

Densidade eletrônica ao redor de um resíduo de triptofano

O método de refinamento que será descrito a seguir baseia-se no princípio dos mínimos quadrados (Hendrickson, W.A.,1985). Embora atualmente os programas de refinamento apliquem o método da máxima verossimilhança, o método dos mínimos quadrados foi escolhido por possuir a vantagem de esclarecer os objetivos do refinamento.

A função a ser minimizada é definida por:

$$\phi_0 = \frac{1}{\sigma_f^2(hkl)} [|F_{hkl}^{obs}| - |F_{hkl}^{calc}|]^2 \quad (4.2)$$

e denominada termo de raio-x. O fator de estrutura do modelo, $|F_{calc}|$, presente na equação (4.2) é calculado pela equação abaixo:

$$F_{hkl}^{calc} = \sum_j^{átomos} f_j \exp[2\pi i(hx_j + ky_j + lz_j)] \quad (4.3)$$

Em geral, o número de reflexões únicas disponíveis não são suficientes para que os parâmetros envolvidos (quatro por átomo) sejam bem determinados. Este problema é contornado adicionando-se informações estereoquímicas (provenientes do estudo de pequenas moléculas) sobre os resíduos de modo que a proporção dados/parâmetros aumente. Estas informações entram no processo de minimização ou na forma de distâncias ou na forma de energia pontencial como será descrito adiante.

Neste ponto o refinamento pode ser vinculado ("*constraint*") ou restrito ("*restraint*"). No primeiro, atribuem-se valores fixos a esses novos termos e no segundo, eles são restritos a um intervalo de valores mais próximos à situação real.

4.1.1 Refinamento vinculado

No refinamento vinculado os comprimentos de ligação e ângulos de ligação são fixos ao passo que os ângulos diédricos podem variar (figura 4.2). Um método que se classifica dentro deste grupo é o refinamento de corpo rígido, onde os átomos são tratados como pertencentes a um corpo rígido. Esta técnica é uma opção, no início do refinamento, para se obter um ajuste primário de orientação e posição da estrutura com relação à cela unitária.

4.1.2 Refinamento restrito

No refinamento restrito todo termo incluso na forma de distância tem a forma geral:

$$\phi_i(x) = \sum_h W_h [g_h^{obs} - g_h^{calc}(x)] \quad (4.4)$$

onde w_h é o inverso do desvio padrão ao quadrado (variância) da distribuição de valores esperados para g_h^{obs} . Os valores de g_h^{calc} são calculados a partir do modelo.

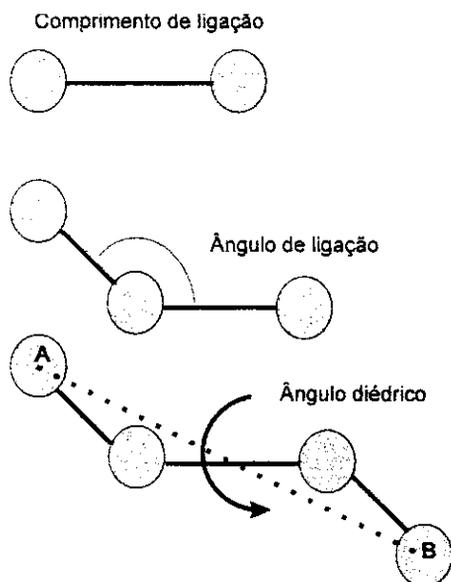


Figura 4.2

Alguns parâmetros estereoquímicos

As restrições incluídas neste procedimento podem ser:

- Restrições sobre as posições atômicas:

1) Distâncias de ligação:

incluem comprimentos de ligação e angulares bem como ângulos diédricos. As relações angulares são tratadas através de distâncias interatômicas (descritas na base de vetores da rede cristalina em questão) como mostrado na figura 4.2 por meio das linhas tracejadas, isto é, $d_{ab}^{modelo} = |\mathbf{r}_a - \mathbf{r}_b|$. A função de observação é dada por :

$$\phi_1 = \sum_j^{distâncias} \frac{1}{\sigma_d^2(j)} [d_j^{ideal} - d_j^{modelo}]^2$$

2) Planaridade de grupos:

considera a variação dos átomos em relação ao plano ideal. Na figura 4.3, para se calcular a distância de um átomo na posição r_{ik} ao plano ideal dois parâmetros são considerados: d_k , que representa a distância do plano ideal à origem e m_k , o vetor unitário normal ao plano ideal. Assim, o desvio é dado por $m_k \cdot r_{ik} - d_k$ e conduz à função de observação:

$$\phi_2 = \sum_k^{\text{planos}} \sum_i^{\text{átomos coplanares}} \frac{1}{\sigma_p^2(i,k)} [m_k \cdot r_{ik} - d_k]^2$$

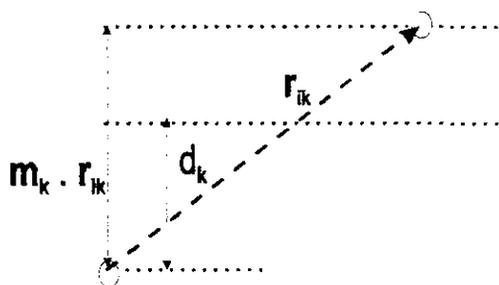


Figura 4.3

Distância de átomos ao plano ideal

3) Quiralidade:

restringe a configuração dos átomos para a forma quiral correta. Este termo é controlado pelo volume quiral. Este parâmetro é definido pelo produto escalar triplo entre os vetores 1, 2, 3 formados entre o átomo localizado em O e os três outros átomos ligados a ele, como ilustrado na figura 4.4. A função de observação apropriada é:

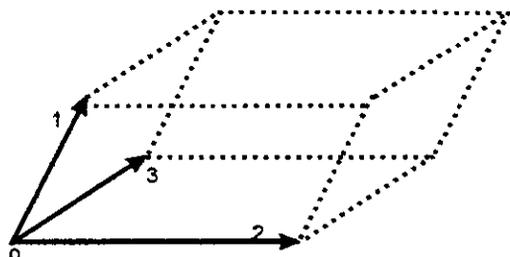


Figura 4.4

Volume quiral

$$\phi_3 = \sum_l^{\text{centros quirais}} \frac{1}{\sigma_q^2(l)} [V_l^{\text{ideal}} - V_l^{\text{modelo}}]^2$$

4) Contatos entre átomos não ligados:

tem por objetivo não permitir que átomos não ligados fiquem a uma distância menor que a distância d_{\min} permitida em termos de energia potencial. A energia potencial (parte repulsiva do potencial de Lennard-Jones) é descrita por $U(d) - U(d^{\min}) \cong \frac{1}{\sigma^4} [d - d^{\min}]^4$, $d < d_{\min}$.

Assim, a função de observação é:

$$\phi_4 = \sum_m^{\text{átomos não ligados}} \frac{1}{\sigma_N^4(m)} [d_m^{\min} - d_m^{\text{modelo}}]^4$$

5) Ângulos de torsão:

são usados para distinguir torsões conformacionais de resíduos que determinam a conformação da cadeia principal inteira. A energia potencial é função do ângulo de torsão e está presente na função de observação na seguinte forma:

$$\phi_5 = \sum_t^{\text{ângulos de torsão conformacional}} \frac{1}{\sigma_T^2(t)} [\chi_t^{\text{ideal}} - \chi_t^{\text{modelo}}]^2$$

- Restrições sobre o movimento atômico:

6) Fator de temperatura isotrópico¹:

¹ Há também restrições sobre o fator de temperatura anisotrópico que devido a sua complexidade não será tratado aqui.

contempla o movimento térmico atômico, B , o qual relaciona-se com a variância do deslocamento, u , dos átomos, isto é, $B = 8\pi u^2$. A função de observação relevante considera a diferença entre os fatores B de dois átomos unidos por uma distância de ligação, ou seja:

$$\phi_6 = \sum_j^{\text{distâncias}} \frac{1}{\sigma_B^2(j)} [B_j^{\text{pri}} - B_j^{\text{seg}}]^2$$

onde B_j^{pri} e B_j^{seg} são os parâmetros térmicos de um dado par de átomos.

- Restrições sobre simetria não-cristalográfica:

7) Posições de objetos simetricamente relacionados:

ocorre quando mais de uma molécula (no mínimo similar) está presente na unidade assimétrica sendo que a operação de simetria existente para cada átomo i pertencente ao objeto k é $r'_{i,k} = R_k r_{i,k} + t_k$, onde R_k é uma matriz de rotação e t_k de translação. Então, definindo-se posições médias $\overline{r'_{i,k}}$ minimiza-se a função:

$$\phi_7 = \sum_i^{\text{átomos únicos}} \frac{1}{\sigma_{sp}^2(i)} \sum_k^{\text{cópias}} [r'_{i,k} - \overline{r'_{i,k}}]^2$$

8) Fator de temperatura de objetos simetricamente relacionados:

a contribuição deste termo assume que objetos simetricamente relacionados possuem fator de temperatura similar. Assim, a função envolvida é:

$$\phi_8 = \sum_i^{\text{átomos únicos}} \frac{1}{\sigma_{sb}^2(i)} \sum_k^{\text{cópias}} (B_{i,k} - \overline{B_{i,k}})^2$$

- Restrições adicionais:

9) Resistência a deslocamentos excessivos:

é recomendável o uso desta restrição quando partes da estrutura não estão bem determinadas (ainda que parâmetros estereoquímicos sejam colocados) para que a instabilidade, que pode ocorrer no processo de minimização, seja controlada. A função aplicada é:

$$\phi_8 = \sum_i^{\text{átomos}} \left[\frac{1}{\sigma_{EP}^2} (r_i - r_i^0)^2 + \frac{1}{\sigma_{EB}^2} (B_i - B_i^0)^2 \right]$$

onde r_i^0 e B_i^0 são parâmetros correntes posicionais e térmicos respectivamente.

10) Fatores de ocupação:

a restrição sobre ocupação aplica-se quando pode-se atribuir a um grupo de átomos que têm ocupação variável a mesma ocupação. Isto ocorre com moléculas de solvente, substratos não covalentemente ligados, inibidores e outros. A função associada é:

$$\phi_9 = \sum_k^{\text{grupos}} \frac{1}{\sigma_Q^2(k)} \sum_i^{\text{átomos no grupo}} [Q_{i,k} - \bar{Q}_{i,k}]^2$$

onde $Q_{i,k}$ é a ocupação do i -ésimo átomo no grupo k e $\bar{Q}_{i,k}$ é a média dos fatores de ocupação no grupo.

A função ϕ_0 pode ser minimizada juntamente com uma função de energia potencial² expressa na forma:

$$E_{total} = E_{empírica} + E_{efetiva}$$

² Este método de refinamento foi proposto inicialmente por Jack e Levitt (Jack & Levitt, 1978).

proposta por Brünger (Brünger, 1988-1992) no pacote computacional X-PLOR³. A energia empírica descreve a energia das moléculas através de funções de energia empíricas como função das coordenadas atômicas. A energia efetiva compreende termos de energia restritos⁴ que usam informações experimentais ou outro tipo de informação.

4.1.3 Simulações de dinâmica molecular com resfriamento lento

Durante o processo de minimização a função $\phi = \sum_{i=0}^N \phi_i$ (N-1 termos estereoquímicos

adicionados) pode atingir um indesejável mínimo local e para transpor a barreira de potencial efetuam-se simulações de dinâmica molecular. Simulações deste tipo são efetuadas submetendo o sistema de partículas a um banho térmico, inicialmente a alta temperatura, e calculando as posições atômicas através da resolução das equações de movimento newtonianas (atribui-se a distribuição de Maxwell para a velocidades dos átomos).

Lentamente, resfria-se o sistema para conduzi-lo ao mínimo tanto de energia potencial quanto de ϕ_0 (Brünger, 1988-1992)³.

4.1.4 Índices de confiabilidade

Ao longo do processo de refinamento existem índices de confiabilidade que indicam a convergência do refinamento e a acurácia do modelo atômico. Estes índices são respectivamente:

$$R_{factor} = \frac{\sum_{hkl} \left| |F_{obs}| - k|F_{calc}| \right|}{\sum_{hkl} |F_{obs}|} \times 100\%$$

e

³ Manual "on-line" com várias referências científicas na página: <http://lance.if.sc.usp.br/XPLOR/htmlman.html>

⁴ V. p. 39

$$R_{free} = \frac{\sum_{hkl \subset T} \left| |F_{obs}| - k|F_{calc}| \right|}{\sum_{hkl \subset T} |F_{obs}|} \times 100\%$$

No caso do R_{free} (Brünger, 1993) são escolhidas aleatoriamente 10% (ou 5%) das reflexões do conjunto de dados que serão excluídas do refinamento e constituem o conjunto de teste, T. O restante das reflexões é denominado conjunto de trabalho, W.

4.1.5 Sínteses de Fourier

Em 1961, Ramachandran e Srinivasan relataram a importância da fase na interpretação do mapa de densidade eletrônica mesclando informações de duas estruturas distintas para construir um mapa de densidade eletrônica. O fator de estrutura calculado possuía amplitude de uma das estruturas e as fases da outra. O mapa resultante apresentava picos nas posições esperadas para a estrutura que cedeu as fases. Como relatado no tópico 4.1.1, construímos mapas tomando as amplitudes observadas experimentalmente e as fases de um modelo inicialmente rudimentar. Para minimizar a interferência que fases do modelo exercerão sobre o mapa de densidade eletrônica construímos mapas segundo a equação abaixo:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l (nm|F_{obs}| - D|F_{calc}|) \exp[-2\pi i(hx + ky + lz - \alpha_{calc})] \quad (4.5)$$

onde D e m são frutos de análises estatísticas que contemplam erros que afetam as coordenadas atômicas e as fases respectivamente.

A interpretação do mapa para $n = 1$ fornece informações a respeito dos erros do modelo, pois este mapa possui valores positivos e negativos. As regiões positivas indicam que há maior contribuição dos dados observados em relação aos calculados; assim o modelo deve ser modificado no sentido de preencher esta região. Caso contrário, se a densidade assumir valor negativo significa que a presença do modelo nesta região deve ser diminuída, pois está

superestimando a densidade eletrônica local. Para $n = 2$, o mapa indica o contorno que a cadeia polipeptídica deve seguir, como exemplificado na figura 4.1.

As modificações são inseridas manualmente por meio de programas que permitem rotações e translações de resíduos do modelo molecular.

4.2 Refinamento das estruturas cristalográficas da beta-lactoglobulina

As estruturas cristalográficas da beta-lactoglobulina resolvidas e refinadas foram:

- 1) Mistura de isoformas A/B, forma cristalina X;
- 2) Mistura de isoformas A/B, forma cristalina Y;
- 3) Isoforma A, forma cristalina Y;
- 4) Isoforma B, forma cristalina Y.

Para a mistura de isoformas A/B na forma cristalina X (grupo espacial P1), o refinamento foi realizado em duas etapas. A primeira etapa foi realizada com o programa XPLOR (Brünger, A.T., 1992) aplicando-se, inicialmente, refinamento de corpo rígido. Simulações de dinâmica molecular com resfriamento lento não resultaram em valores satisfatórios para os índices R e R_{free} . Correções aos módulos, $|F_{obs}|$, foram feitas através da avaliação da anisotropia geral nos dados de difração (Brünger, A.T., 1992, Sheriff & Hendrickson, 1987) resultando na diminuição de 4% e 5% nos valores de R e R_{free} respectivamente. A segunda etapa, feita com o programa REFMAC (CCP4), compreendeu o refinamento restrito das coordenadas atômicas considerando fatores de temperatura isotrópicos individuais para os átomos, simetria não-cristalográfica e escalonamento entre $|F_{obs}|$ e $|F_{calc}|$ estimando a contribuição do solvente (*bulk solvent correction*).

A cada ciclo de refinamento houve a inspeção dos mapas ponderados de densidade eletrônica ($2m|F_{obs}| - D|F_{calc}|$ e $m|F_{obs}| - D|F_{calc}|$) e, quando necessário, o ajuste manual de resíduos com o programa O (Jones, T.A., *et al.*, 1991). A localização de moléculas de água foi efetuada com o programa ARP (Lamzin, V.S., *et al.*, 1993). Até este ponto empregou-se o refinamento contra todas as reflexões até 1,8 Å sendo que 10% foram separadas para o cálculo do fator R_{free} .

Os fatores de confiabilidade R_{factor} e R_{free} atingiram os valores de 21,5% e 25,7% respectivamente e estagnaram. A análise do gráfico de Luzzati demonstrou que para camadas de resolução mais altas do que 1,6 Å não havia convergência do refinamento. Examinando-se o conjunto de dados com o programa DATAMAN (Kleywegt, G.J., *et al.*, 1996) a resolução de 1,6Å foi indicada como a resolução efetiva do conjunto de dados. Assim, a porção final do refinamento prosseguiu contra reflexões até 1,6 Å. A restrição de simetria não-cristalográfica foi retirada e nova correção foi aplicada a $|F_{\text{obs}}|$ com um programa desenvolvido por Igor Polikarpov (Polikarpov *et al.*, 1995). Os valores finais para os índices de confiabilidade foram $R_{\text{factor}} = 19,8\%$ e $R_{\text{free}} = 25,0\%$ com a localização de 171 moléculas de água.

Para as estruturas na forma cristalina Y (grupo espacial C222₁) o protocolo de refinamento utilizado foi idêntico para todas. O refinamento feito com o programa REFMAC usou reflexões no intervalo entre 9,5 Å e o limite superior de resolução correspondente a cada conjunto de dados: 2,0 Å para a mistura de isoformas, 2,0 Å para a variante A e 1,95 Å para a variante B. Foram separadas 5% das reflexões para o cálculo de R_{free} .

Para que o escalonamento entre $|F_{\text{obs}}|$ e $|F_{\text{calc}}|$ fosse efetuado apropriadamente (*bulk solvent correction*), o limite a baixa resolução foi estabelecido após o estudo da intensidade contra resolução dos conjuntos de dados com o programa DATAMAN. Também foram considerados fatores de temperatura isotrópicos individuais para os átomos.

A inspeção dos mapas ponderados de densidade eletrônica revelaram uma densidade clara para a maioria dos resíduos, entretanto não foi possível observá-la para os dez últimos resíduos (para a isoforma A, os oito últimos) e, no caso da mistura de isoformas A/B, também a região de resíduos 110 a 115. Por esta razão, estes resíduos não foram incluídos nos respectivos modelos. O refinamento prosseguiu com o ajuste manual de resíduos (com o programa O) e o refinamento restrito. As moléculas de água foram inseridas com o programa ARP. Os resultados finais obtidos para as estruturas na forma cristalina Y são apresentados na tabela 4.1.

	Região representada pelo modelo num total de 162 resíduos	Resolução (Å)	R _{factor} (%)	R _{free} (%)	No. de moléculas de água
Mistura de isoformas A/B	1-109 e 116-152	2,0	19,4	27,2	135
Isoforma A	1-154	2,0	19,2	28,7	123
Isoforma B	1-152	1,95	18,9	25,2	107

Tabela 4.1

Resultado final do refinamento para as estruturas cristalográficas da BLG na forma cristalina Y

4.3 Qualidade e acurácia dos modelos

Todos os modelos foram submetidos a análise com o programa PROCHECK (Laskowski, R.A., 1993). Para todos os critérios usados pelo PROCHECK os modelos estavam dentro dos limites típicos para estruturas resolvidas na resolução correspondente.

Os ângulos ϕ e ψ para a cadeia principal (figura 4.6) estão apresentados na figura 4.8 em diagramas de Ramachandram. Em todos os diagramas da figura 4.7 o resíduo de Tyr 99 está em uma região desfavorável. Como relatado por Brownlow *et al.* o resíduo Tyr 99 faz parte de uma " γ -turn"⁵ (figura 4.7) que é conservada em todas as estruturas de lipocalinas conhecidas com a seqüência Thr-Asp-Tyr. Na figura 4.8 (a) o resíduo Ala 34 nos dois monômeros também estão em uma conformação não permitida. O resíduo Ala 34 se encontra na interface do dímero e participa de uma ponte de hidrogênio, N Ala 34→OD1 Asp33, com o resíduo de ácido aspártico Asp pertencente a outra subunidade do dímero.

Outros resultados sobre o refinamento das quatro estruturas estão apresentados na tabela 4.2.

⁵ "Gamma-turn" é definida por três resíduos onde i e $i+2$ interagem através de uma ponte de hidrogênio e $i+1$ possui ϕ_{i+1} entre 35° e 115° e ψ_{i+1} entre -19° e -109°

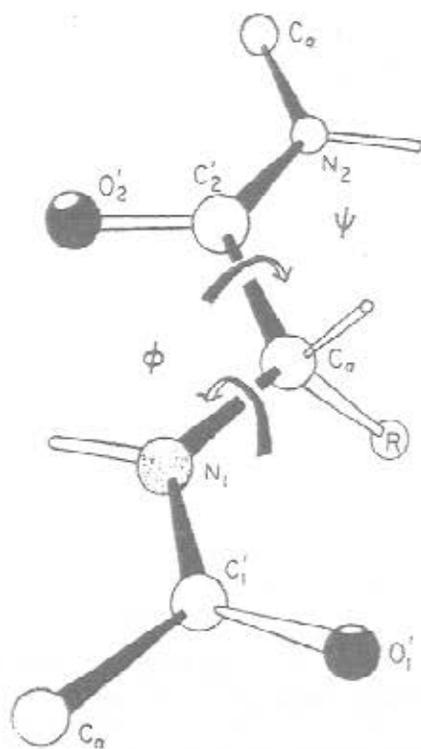


Figura 4.6

Conformações de uma cadeia polipeptídica definidas em termos dos ângulos diédricos ϕ e ψ

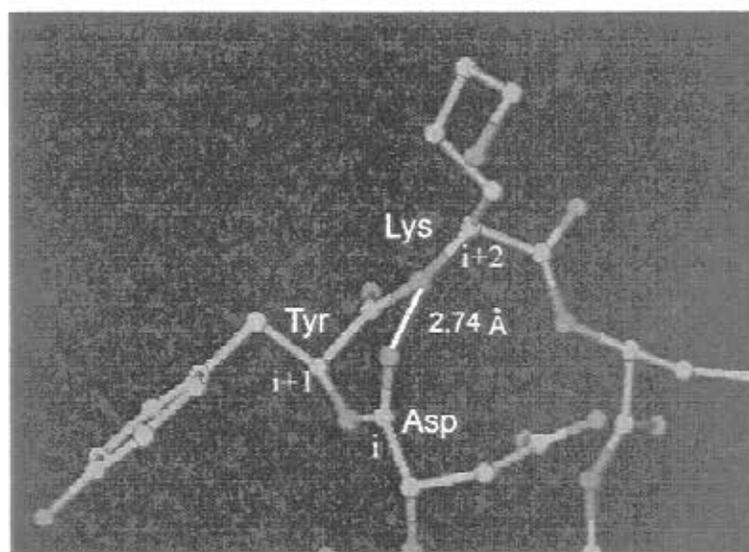
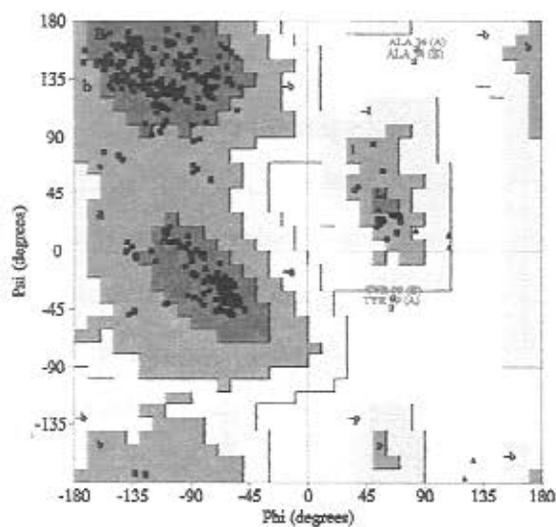
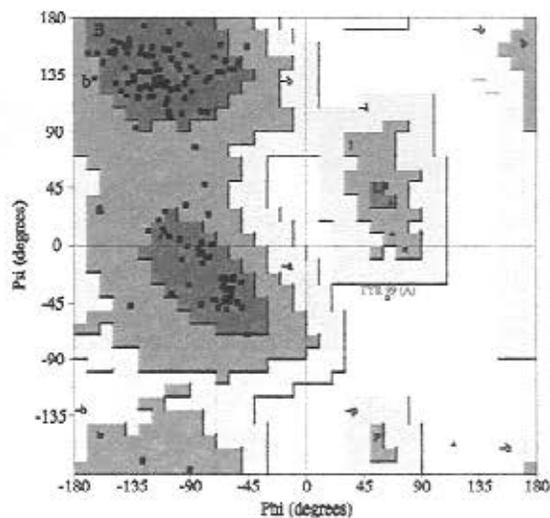


Figura 4.7

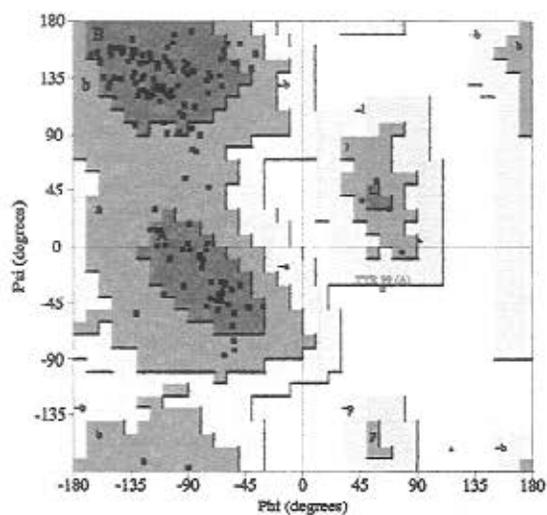
Conformação do resíduo de Tyr 99



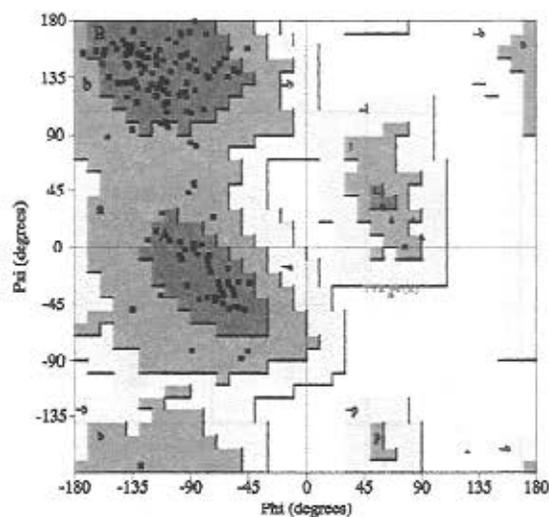
(a)



(b)



(c)



(d)

Figura 4.8

Diagramas de Ramachandram. (a) mistura de isoformas A/B, forma cristalina X, (b) mistura de isoformas A/B, forma cristalina Y, (c) isoforma A, forma cristalina Y, (d) isoforma B, forma cristalina Y. As regiões mais favoráveis encontram-se nas regiões mais escuras dos diagramas.

desvio da idealidade

(*)

(**)

(***)

	comprimento das ligações (Å)	ângulos de ligação (°)	erro nas coordenadas (Å)	figura de mérito	coeficiente de correlação
Mistura de isoformas A/B, forma cristalina X	0,04	3,54	0,14	0,9	0,95
Mistura de isoformas A/B, forma cristalina Y	0,15	4,52	0,15	0,86	0,82
isoforma A, forma cristalina Y	0,07	4,56	0,15	0,89	0,78
isoforma B, forma cristalina Y	0,03	3,55	0,21	0,86	0,82

(*) Calculado com X-PLOR;

(**) Calculado com Sigmaa (CCP4);

(***) Calculado com O.

Tabela 4.2

Resultados do refinamento para as estruturas da BLG

4.4 Empacotamento cristalino das formas cristalinas X e Y

A análise do empacotamento cristalino de todas as estruturas foi feito com o programa O. Este tipo de avaliação indica a ocorrência de contatos impróprios entre moléculas simetricamente relacionadas. As posições destas moléculas são determinadas pelas operações de simetria do grupo espacial correspondente. Tanto para a forma cristalina X (grupo espacial $P1$) quanto para a forma cristalina Y (grupo espacial $C222_1$) nenhum contato inadequado foi encontrado.

Comparação entre as estruturas cristalográficas da beta-lactoglobulina

Neste capítulo serão apresentados os resultados obtidos através de estudos comparativos com estruturas da beta-lactoglobulina (BLG) resolvidas neste projeto e as disponíveis até o presente momento no banco de dados PDB ("*Protein Data Bank*").

5.1 Alguns aspectos gerais da BLG

Como apresentado no capítulo 1, a forma estrutural da BLG, é descrita com um cálice formado por fios β antiparalelos. Os fios β são nomeados de A a I segundo o trabalho de Sawyer *et al.* em 1985 e as alças que os une, BC, CD e assim por diante. As α -hélices foram nomeadas α_1 , α_2 , α_3 e α_4 (figura 5.1).

Certas alças da BLG correspondem a regiões de grande mobilidade e algumas destas dependem da forma cristalina. Na forma X, destaca-se a alça CD (Brownlow *et al.*, 1997). Na forma Z, sobressaem-se as alças: CD e GH (Qin *et al.*, 1998). Na forma Y, evidenciam-se: α_3 B, CD e GH.

A BLG apresenta duas cistinas, Cys66 — Cys 160 e Cys106 — Cys119, e uma cisteína livre, Cys 121¹. A ponte dissulfeto Cys66 — Cys 160 ata os resíduos do C-terminal ao restante da molécula. Entretanto em algumas formas cristalinas, como a forma Y, a porção final da molécula é difícil de ser visualizada nos mapas de densidade eletrônica.

¹ V. Apêndice A

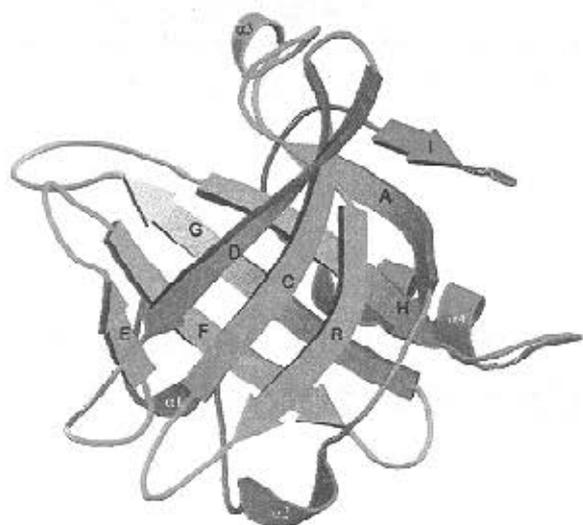


Figura 5.1

Estrutura cristalográfica da beta-lactoglobulina

A família das lipocalinas possui uma região altamente conservada determinada pela seqüência Gly-X-Trp (onde X é um resíduo qualquer) sendo que o resíduo de triptofano é o único invariante (figura 1.2). Na BLG, este resíduo corresponde ao Trp19.

Como dito no primeiro capítulo, proteínas desta família possuem habilidade em ligar pequenas moléculas hidrofóbicas e sua forma tridimensional é semelhante a um cálice (ou barril). Baseados na conservabilidade deste resíduo de triptofano, Katakura e colaboradores (1994) investigaram a importância do Trp19 na formação da estrutura da BLG e o seu papel na ligação com o retinol. As análises foram conduzidas baseando-se nos efeitos da substituição do resíduo Trp19 por um resíduo de tirosina. Os resultados encontrados indicaram que o Trp19 não é essencial para a ligação com o retinol. Entretanto, é importante para estabilizar a molécula ligada e protegê-la de degradação mantendo a estrutura local e impedindo o acesso do solvente. A proteína mutante também se mostrou menos estável frente ao desnaturante Gdn-HCl. Isto revela que o Trp19 é relevante não só na manutenção da estrutura local mas também na estabilidade geral da proteína. Katakura *et al.* assumiram que a perda de pontes de hidrogênio ao redor do Trp19 devido a mutação, seria uma das causas principais para a perda de estabilidade. Analisando as interações ao redor do Trp19 para as quatro estruturas cristalográficas resolvidas foi possível identificar duas pontes de hidrogênio: N Trp19 → O Glu 44 e NE1 Trp19 → O Val15 (figura 5.2).

Outro aspecto interessante a respeito do resíduo Trp19 se refere aos estudos na emissão de fluorescência com pressão relatados por Valente-Mesquita *et al.*(1998). Nestes estudos foram relatados um aumento na intensidade da emissão de fluorescência causado pela pressão. Este comportamento pode ser explicado através dos dados cristalográficos investigando-se as vizinhanças dos resíduos de triptofano. A BLG possui dois resíduos de triptofano: Trp19 e Trp61. Próximo ao resíduo Trp19 (distância menor que 4Å) encontra-se o resíduo Arg124 (figura 5.3). Este resíduo é conhecido por seus efeitos supressivos na emissão de fluorescência (Yu *et al.*, 1992). Dado que os espectros de fluorescência da BLG sob pressão ou na presença do desnaturante Gdn-HCl foram similares, concluiu-se que a pressão promoveu a desnaturação da proteína. Assim, conforme ocorre o desenovelamento da proteína, os resíduos Trp19 e Arg124 distanciam-se e o aumento na intensidade pode ser explicado.

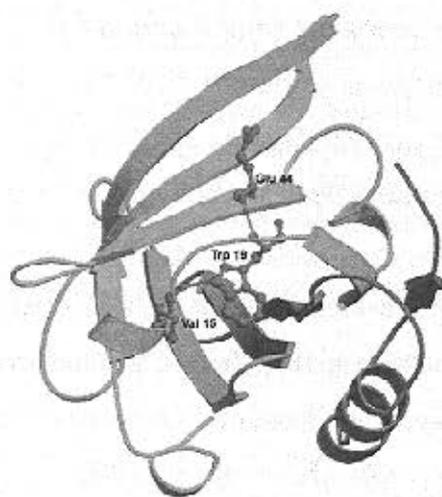


Figura 5.2

Pontes de hidrogênio com o Trp19. Este resíduo encontra-se na ponta do cálice formado pelos fios β na BLG

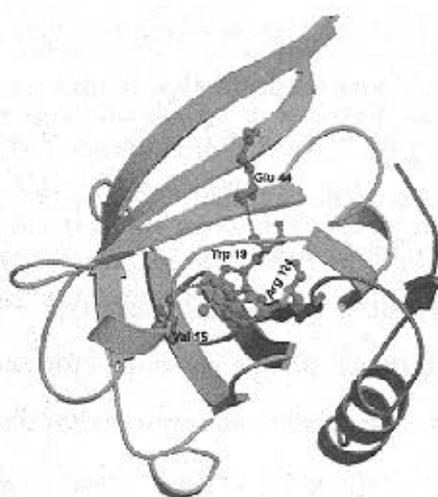


Figura 5.3

Localização espacial dos resíduos Trp19 e Arg124

5.2 Estrutura da BLG na forma cristalina X a 1.6Å

Esta estrutura da BLG nesta forma cristalina representa um dímero na unidade assimétrica de um cristal da mistura de isoformas A e B crescido em pH ácido. A figura 5.4 mostra a relação entre a estrutura secundária e parâmetros como fator de temperatura, coeficiente de correlação e acessibilidade ao solvente. As regiões na estrutura da BLG que possuem um fator de temperatura maior correspondem às alças, normalmente expostas ao solvente, cujo coeficiente de correlação é menor que os demais resíduos.

Os fatores de temperatura são afetados pela flexibilidade da molécula de proteína, pela desordem na rede cristalina e pelos erros nas fases cristalográficas. A distinção entre flexibilidade da molécula e desordem na rede cristalina pode ser feita comparando a molécula em diferentes formas cristalinas ou em diferentes vizinhanças na mesma forma cristalina quando se tem simetria não-cristalográfica.

Como relatado no tópico 5.1, nas três formas cristalinas X, Y, Z, a região correspondente à alça CD (resíduos 61 a 66) apresentam-se com altos valores para o fator de temperatura. A posição desta alça sempre se diferencia quando as estruturas disponíveis são comparadas entre si. Na forma X, a comparação entre os monômeros também indica diferenças nas posições atômicas nesta alça². Isto evidencia que a alça CD corresponde a uma região de mobilidade da BLG.

Outras alças como α_3B (resíduos 31 a 42) e GH (resíduos 108 a 117) são mais definidas na forma X do que nas outras formas cristalinas, possivelmente, devido aos contatos intramoleculares estabelecidos com outros resíduos espacialmente adjacentes. Os resíduos da alça α_3B e GH participam, respectivamente, de interações na interface do dímero (descritas a seguir) e com os resíduos Asn88, Glu89 e Asn90 presentes na alça EF.

Na interface do dímero ocorrem as interações mostradas na figura 5.5. As interações identificadas na cadeia A→B foram: OD1 Asp 33→NH1 Arg 40, OD2 Asp 33→NH2 Arg 40,

² Como apresentado no tópico 4.2, ao final do refinamento da forma cristalina X foi liberada a restrição de simetria não-cristalográfica entre os monômeros. Nos trabalhos de Brownlow, para esta mesma forma cristalina, também há referência de remoção da restrição de simetria não-cristalográfica entre as subunidades.

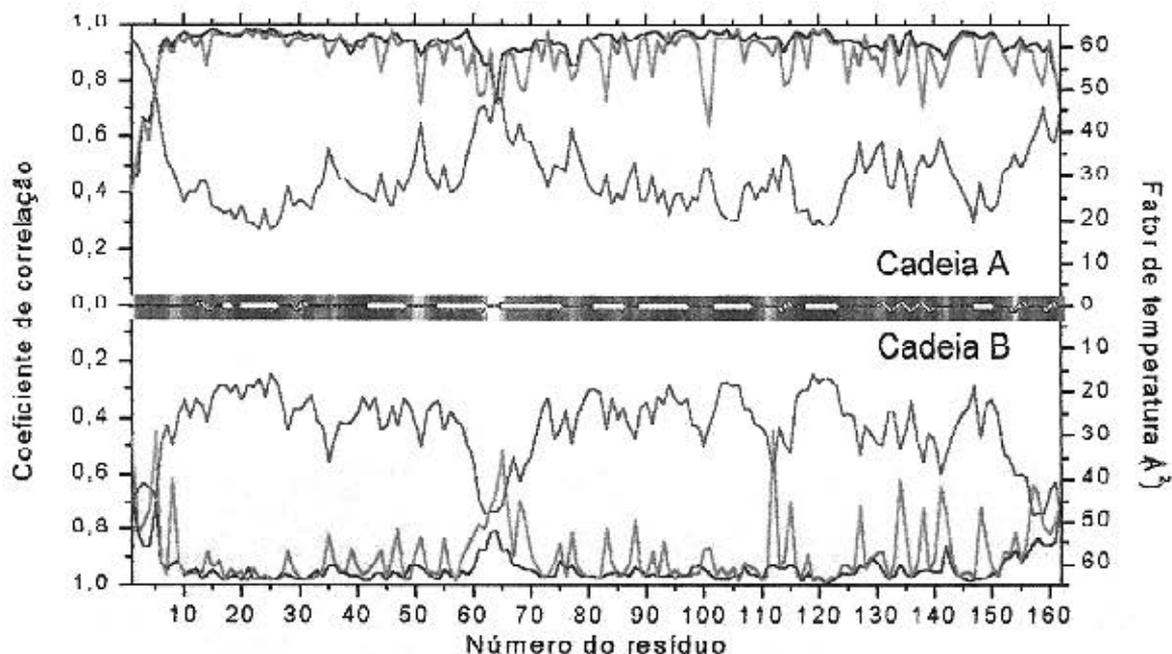


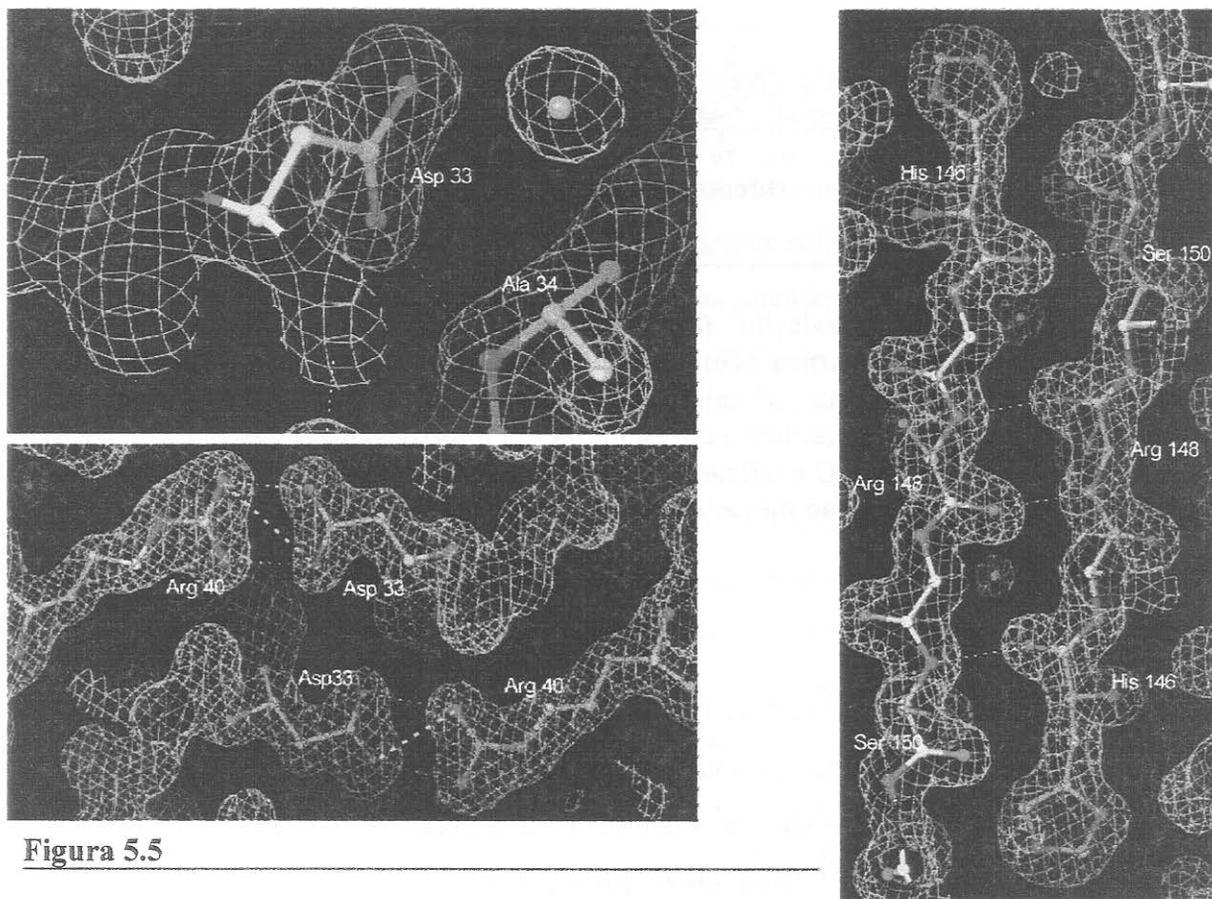
Figura 5.4

Gráfico do coeficiente de correlação* (cadeia principal em azul e cadeia lateral em verde) e do fator de temperatura (em vermelho) em função do resíduo para cada monômero da BLG na forma cristalina X. No gráfico também está presente a acessibilidade ao solvente dos resíduos representada em tons de azul. Quanto mais claro, maior a exposição do resíduo. *O coeficiente de correlação foi calculado com o programa O considerando-se a contribuição da cadeia principal e lateral

OD2 Asp 33→NH1 Arg 40, N Ala 34→OD1 Asp33, O His 146→N Ser 150, N Arg 148→O Arg 148 e as respectivas interações complementares de B→A. Nesta interface, uma área de aproximadamente 1080\AA^2 é inacessível ao solvente frente a dimerização segundo os cálculos efetuados com o programa GRASP (Nicholls & Honing, 1991).

Estudos recentes de desnaturação com Gdn-HCl indicaram que a dissociação do dímero ocorre com baixas concentrações de Gdn-HCl, cerca de 0,2 M, frente à concentração necessária para a desnaturação completa da proteína que está em torno de 3,0 M (Botelho *et al.*, comunicação pessoal). Este tipo de informação a respeito das interações na interface do dímero é importante, por exemplo, na interpretação dos espectros de fluorescência com

pressão da BLG em pH básico. Nesta faixa de pH, considera-se que a pressão promove dissociação das subunidades e desenovelamento da proteína, sendo que os espectros produzidos são similares aos induzidos por Gdn-HCl (Mesquita-Valente *et al.*, 1998). Portanto, se a dissociação ocorre facilmente, os resultados das análises dos espectros de fluorescência expressam, essencialmente, o desenovelamento da proteína. Esta hipótese foi confirmada com experimentos de fluorescência com pressão em pH 2,0 (onde a BLG é monomérica) que produziram espectros similares àqueles em pH 7,0, indicando que a dissociação teria um papel menor na análise dos dados (Mesquita-Valente *et al.*, comunicação pessoal).



Contatos na interface do dímero da beta-lactoglobulina, forma cristalina X. As interações são representadas por linhas tracejadas

Com relação à escolha da variante que este dímero representa, a modelagem da estrutura foi efetuada para a variante A, segundo a análise da densidade eletrônica na posição 118. Esta posição corresponde ao segundo ponto de mutação que distingue as variantes. O primeiro ponto de mutação, na posição 64, não foi determinante pois, não havia densidade eletrônica local suficiente para a cadeia lateral.

5.2.1 Comparação entre estruturas da BLG na forma cristalina X

Brownlow *et al.* em 1997 apresentou a primeira estrutura da beta-lactoglobulina refinada a alta resolução (código de acesso no PDB 1BEB). Esta estrutura foi resolvida a 1.8Å para um cristal também na forma cristalina X crescido em torno do pH 6,5. Comparando a presente forma cristalina X a 1.6Å com esta a 1.8Å, foram encontradas algumas diferenças. Em primeiro lugar, o modelo estrutural resolvido a mais alta resolução representa todos os 162 resíduos da beta-lactoglobulina, ao passo que o modelo 1BEB representa os resíduos 5 a 160. As estruturas foram superpostas e as diferenças entre as coordenadas calculadas, respectivamente, com os programas LSQKAB (CCP4) e COMPAR (CCP4). Algumas diferenças nas posições atômicas tanto na cadeia lateral como na cadeia principal foram observadas. Nas cadeias laterais as diferenças ocorreram em função de resíduos localizados na superfície da molécula, os quais possuem maior mobilidade. Na cadeia principal, as diferenças ocorreram no resíduo Pro 38 na cadeia A, na região dos resíduos 60 a 65 (alça CD) e na extremidade da cadeia B (figura 5.6). No resíduo Pro 38 ocorreu uma inversão na posição da ligação C=O na estrutura 1BEB como mostrado na figura 5.7. Esta inversão em 1BEB explica o porquê do resíduo Leu 39 estar uma região não permitida no diagrama de Ramachandran para esta estrutura. As demais diferenças se referem a regiões de maior flexibilidade da proteína.

Finalmente, o fator de temperatura médio para todos os átomos de proteína é de aproximadamente 30Å^2 para ambas as estruturas. Entretanto, a comparação deste parâmetro ao longo da cadeia polipeptídica entre estas estruturas (figura 5.6) revelou diferenças significativas. Neste tipo de análise, avalia-se a correlação entre o fator de temperatura e o elemento da estrutura secundária correspondente da proteína. Em 1BEB, regiões de maior flexibilidade da proteína, como extremidades e alças, possuem fatores de temperatura

semelhantes a resíduos mais internos da proteína. Ao contrário, na estrutura atual há diferenciação entre regiões da estrutura da proteína com relação aos valores do fator de temperatura. Esta diferença no comportamento do fator de temperatura indica fortes indícios de que em 1BEB não houve o refinamento deste parâmetro.

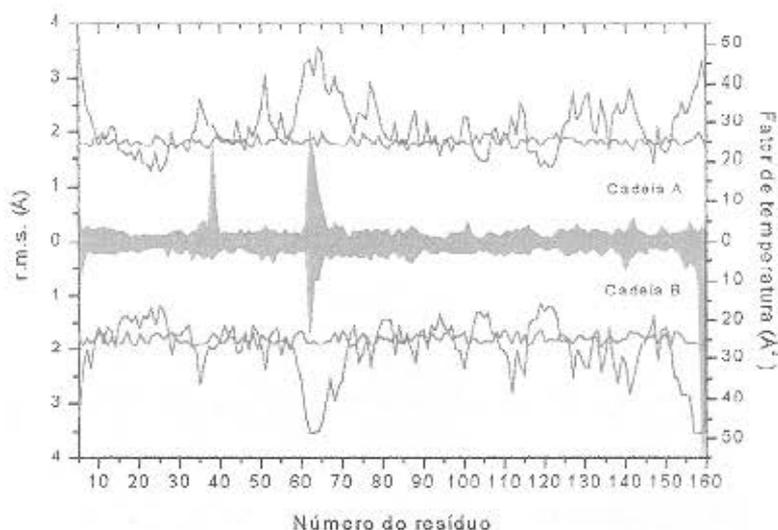


Figura 5.6

R.M.S. entre estruturas da BLG na forma cristalina X em função do resíduo (em verde). No gráfico também são mostrados o fator de temperatura em função do resíduo para a estrutural atual (em azul) e para 1BEB (em vermelho)

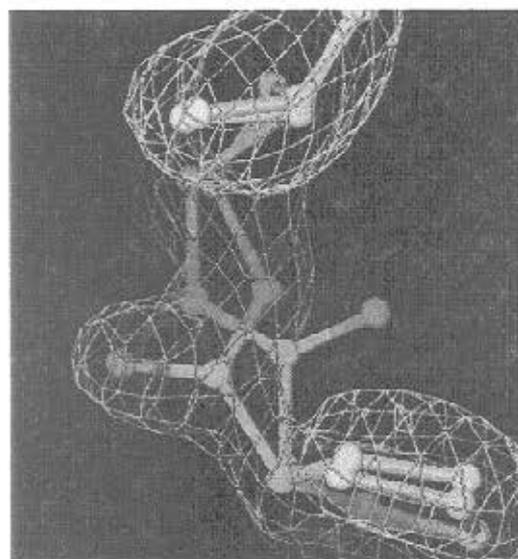


Figura 5.7

Diferenças no resíduo Pro38 nas formas cristalinas X da BLG. A ligação C=O da estrutura 1BEB não explica a densidade eletrônica local

5.3 Estruturas da BLG na forma cristalina Y

As estruturas resolvidas nesta forma cristalina representam um monômero na unidade assimétrica de cristais crescidos em pH básico. As estruturas correspondem à variante A, variante B e mistura de variantes A/B na resolução de 2,0Å, 1,95Å e 2,0Å respectivamente.

A comparação entre os fatores de temperatura das variantes A e B evidencia que a variante A possui fatores de temperatura ligeiramente maiores quando comparados a variante B (figura 5.8). Esta constatação vai de encontro aos experimentos de troca de hidrogênio ("hydrogen exchange") publicados por Dong *et al.* em 1996. Estes resultados mostraram que, embora não haja evidências de diferenças estruturais entre estas variantes, a variante A apresenta mobilidade ligeiramente maior. Nestes estudos a proteína é desnaturada com uma solução D₂O-desnaturante para trocar grupos NH por ND. O processo de enovelamento é iniciado diminuindo a concentração do desnaturante. Com isto, os grupos ND tornam-se protegidos à medida que a proteína adquire novamente sua estrutura secundária e terciária. Esta amostra é então diluída em água e as trocas entre H e D são monitoradas com NMR ou espectroscopia no infravermelho. Quanto menos rígida for a estrutura maior a taxa de troca.

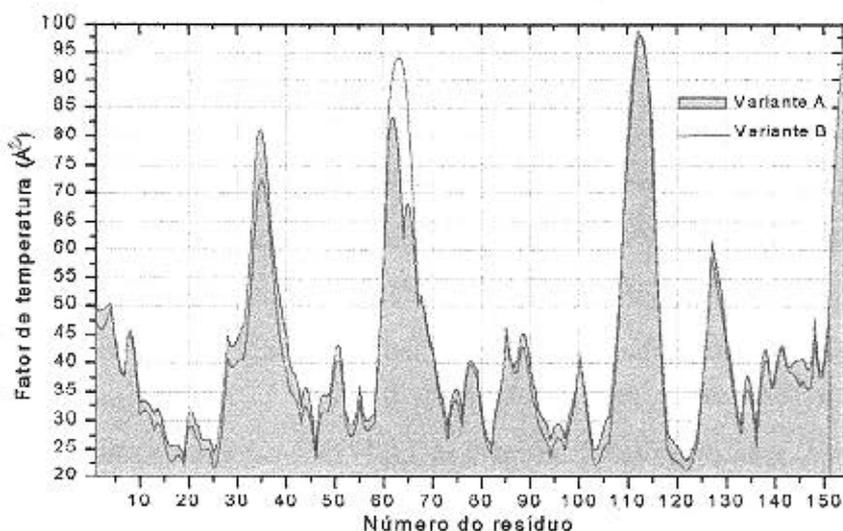


Figura 5.8

Gráfico do fator de temperatura das variantes A e B

A superposição das estruturas e o cálculo das diferenças nas coordenadas foram feitos com os programas citados no tópico 5.2.1. Os resultados estão apresentados na figura 5.9. Geralmente, regiões de alta mobilidade são caracterizadas por densidade eletrônica insuficiente e refletem grandes diferenças nas posições atômicas (figura 5.10). Nesta condição se encontram os resíduos 30 a 35, 61 a 65, 110 a 115. Como apresentado nos tópicos 5.1 e 5.2, as duas últimas zonas não são exclusivas da forma cristalina Y. Em outras formas cristalinas elas também se destacam por sua mobilidade.

As zonas 61 a 65 e 110 a 115 fazem parte das alças CD e GH respectivamente (figura 5.1) e são de especial interesse. Como dito no tópico 1.1, as mutações que distinguem as variantes A e B ocorrem nas posições 64 e 118. A primeira modificação localiza-se exatamente sobre a alça CD e acredita-se que os efeitos conformacionais desta mutação sejam encobertos em função da mobilidade desta região (Brownlow *et al.*, 1997). Já a segunda ocorre no núcleo hidrofóbico na molécula e está próxima à alça GH. As conseqüências desta mutação sobre esta alça são uma questão aberta. Entre as três formas cristalinas discutidas, a forma X é a única que exhibe uma densidade eletrônica local bem definida que pode ser explicada em função de contatos intramoleculares estabelecidos com a alça EF. Por outro lado, as estruturas disponíveis nesta forma cristalina são provenientes de cristais da mistura de isoformas A/B e a participação de cada variante na formação do cristal (e conseqüentemente, seus efeitos) é difícil de precisar.

Na figura 5.11 estão ilustrados os modelos adotados para as variantes A e B na forma cristalina Y em relação à forma X. Uma proposta para decidir se as diferenças nas coordenadas entre as variantes A e B são significantes seria introduzir uma medida de significância da variação posicional entre estes modelos, normalizando a diferença entre as coordenadas através do deslocamento térmico médio, estimado com base no fator de temperatura.

O fator de temperatura isotrópico, B, é dado por

$$B = 8\pi^2 \langle u^2 \rangle$$

onde $\langle u^2 \rangle$ é o deslocamento térmico quadrático médio de um átomo. Podemos estimar a amplitude média do deslocamento atômico, $\langle u \rangle$, assumindo que :

$$\langle u \rangle \approx \sqrt{\langle u^2 \rangle} = \sqrt{\frac{B}{8\pi^2}}$$

Na comparação entre estruturas, a média do deslocamento térmico quadrático poderia ser estimada por

$$\langle u^2 \rangle = \frac{\overline{B_A} + \overline{B_B}}{16\pi^2}$$

onde $\overline{B_A}, \overline{B_B}$ correspondem aos fatores de temperatura das estruturas comparadas.

A figura 5.12 mostra a diferença nas posições atômicas para a cadeia principal das variantes A e B da BLG normalizada pela estimativa do deslocamento atômico médio. Desde que o deslocamento térmico médio fornece uma medida da incerteza na exata localização do fragmento molecular no espaço, podemos assumir que a diferença na posição atômica é significativa se ela corresponde a pelo menos três vezes o valor correspondente de $\sqrt{\langle u^2 \rangle}$. A única parte da estrutura acima de 3σ corresponde aos resíduos da alça GH. Assim, é possível que as diferenças conformacionais na alça GH entre as variantes A e B, identificadas na forma Y, sejam efetivas.

Em termos das estruturas das variantes como um todo, a mutação Val \rightarrow Ala parece estar relacionada com a estabilidade da proteína. Experimentos de desnaturação sob pressão realizados por Mesquita *et al.* demonstraram que a isoforma B é termodinamicamente mais estável que a isoforma A com constantes de equilíbrio para a desnaturação dadas por: $(K_{eq})_B = 0,010 \pm 0,002$ e $(K_{eq})_A = 0,037 \pm 0,020$. Entretanto, para a mudança de volume foram encontrados os seguintes valores: $\Delta V_A = 51 \pm 8$ ml/mol e $\Delta V_B = 74 \pm 3$ ml/mol. A maior variação de volume para a isoforma B faz com que esta variante seja mais suscetível a desnaturação sob pressão.

A diferença em ΔV observada para as variantes A e B corresponde a 23 ml/mol (cerca de $38 \text{ \AA}^3/\text{molécula}$) e pode ser interpretada com base nos dados cristalográficos destas variantes. A substituição Val \rightarrow Ala resulta na perda de dois grupos metila ($-\text{CH}_3$)

ocasionando a formação de uma cavidade no interior da proteína. Não houve mudança estrutural significativa na vizinhança desta posição no sentido de compensar a perda destes grupos. O volume ocupado por estes dois grupos $-\text{CH}_3$ é de aproximadamente 40 \AA^3 , segundo os cálculos efetuados com o programa GRASP. Curiosamente, a diferença no volume das celas unitárias das isoformas A e B é exatamente oito vezes este valor de volume, ou seja, corresponde exatamente a soma dos volumes dos grupos $-\text{CH}_3$ perdidos para todas as oito moléculas presentes na cela unitária do cristal. Este valor também é comparável à diferença na variação de volume das variantes nos experimentos de pressão. Portanto, a formação de uma cavidade no interior da proteína poderia explicar a maior suscetibilidade frente a pressão da variante B em relação à variante A.

A comparação entre as estruturas de ambas as variantes e a mistura de variantes A/B (todas resolvidas em pH 7,9) nenhuma diferença significativa pode ser encontrada. A Mistura de variantes foi modelada como variante A também em função da densidade eletrônica do resíduo 118 pelos mesmos motivos apresentados no tópico anterior para a forma cristalina X.

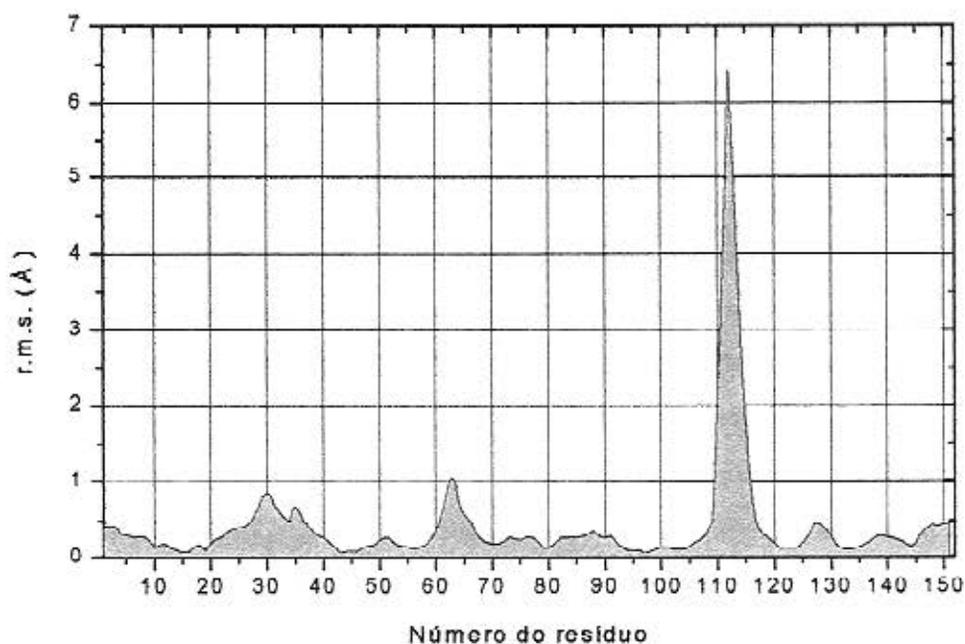


Figura 5.9

Diferenças nas coordenadas entre as variantes A e B

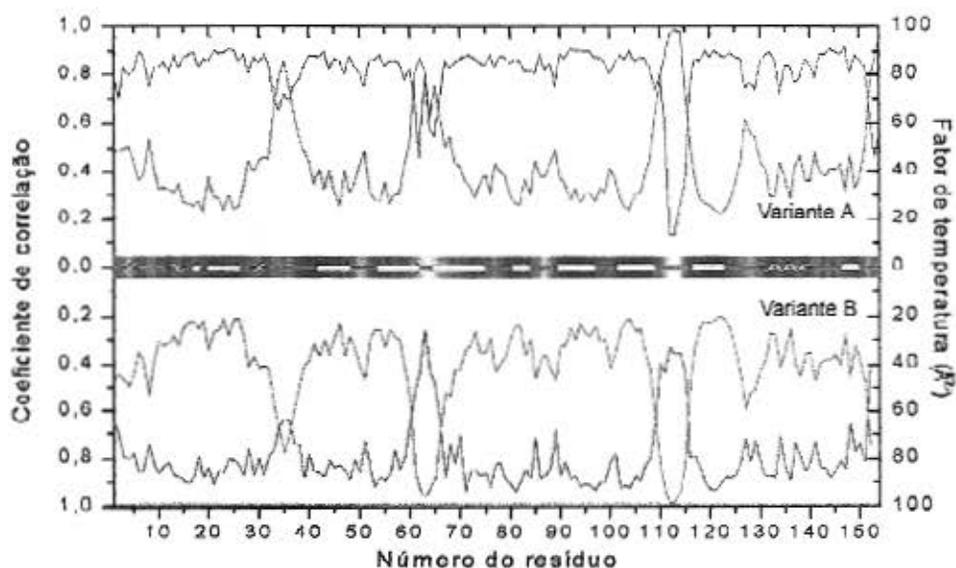


Figura 5.10

Gráfico do coeficiente de correlação (em azul) e do fator de temperatura (em vermelho) em função do resíduo para as variante A e B. Também estão apresentados os elementos da estrutura secundária e a acessibilidade dos resíduos ao solvente em tons de azul. Quanto mais escuro menos acessível e vice-versa

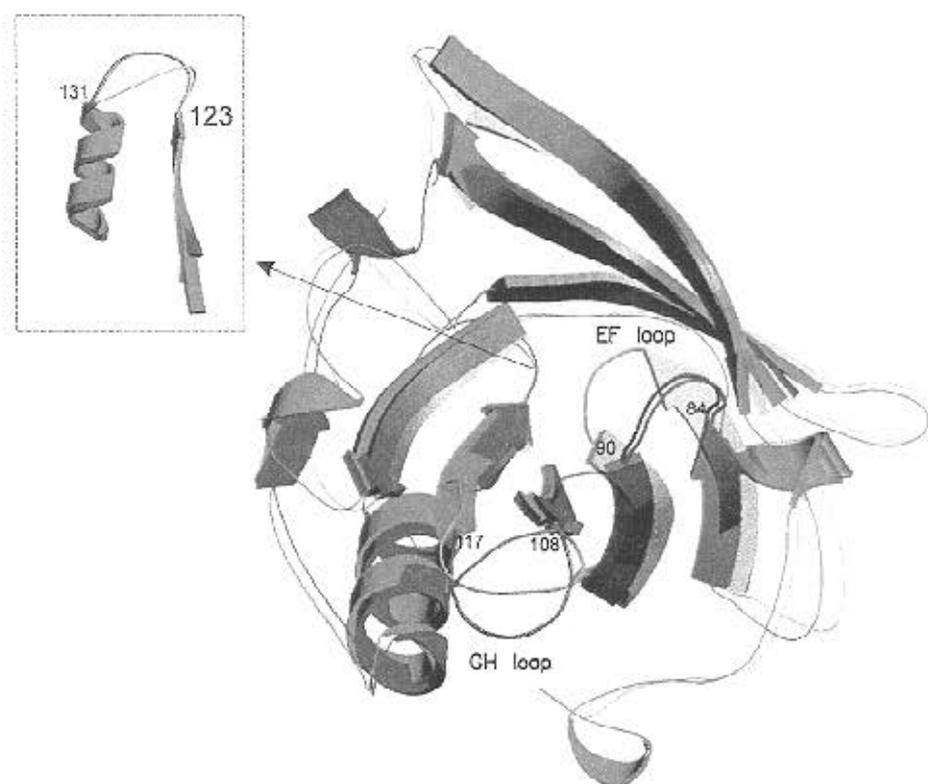


Figura 5.11

Superposição entre a forma cristalina X da mistura de isoformas e forma cristalina Y das variantes A e B. Alça em azul corresponde à variante B. Em amarelo, mistura de isoformas A e B. Em vermelho, variante A.

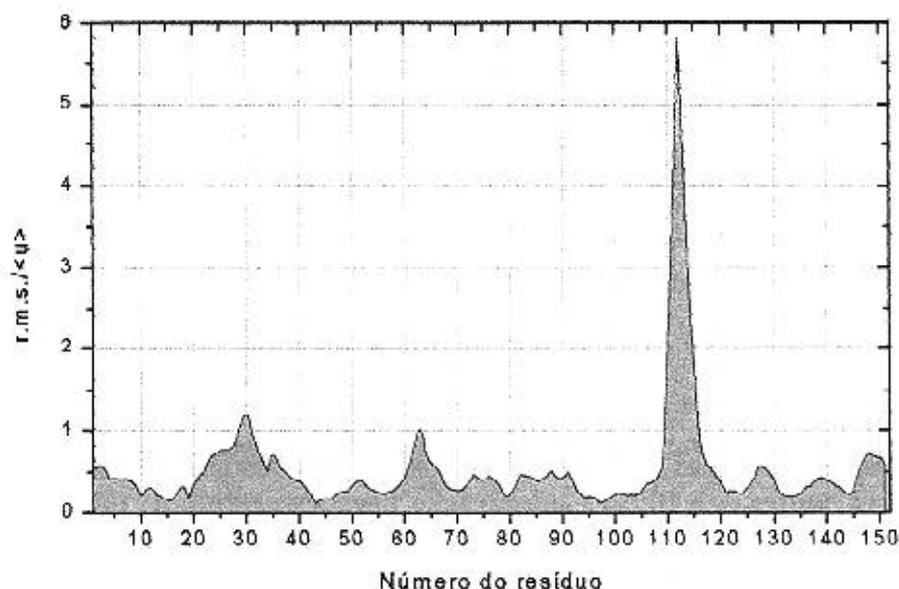


Figura 5.12

Diferenças nas coordenadas das variantes A e B da BLG normalizadas pelo deslocamento térmico médio

5.3.1 Análise da interface do dímero cristalográfico das variantes A e B

Os monômeros das isoformas A e B foram superpostos com o correspondente monômero da forma X. Na forma cristalina Y, os eixos do dímero coincidem com o eixo cristalográfico. No entanto todas as interações que ocorrem na interface do dímero não-cristalográfico (forma X) são preservadas com exceção da ponte de hidrogênio N Ala34 → OD1 Asp33* que está ausente em ambas as isoformas e a interação OD2 Asp33 → NE2 Gln35* que ocorre somente para a isoforma A. Por outro lado, os dímeros das formas X e Y não são idênticos porque os monômeros que formam o dímero estão rotacionados como corpos rígidos ao redor de um eixo entre os dois de tal maneira que, o dímero na forma Y é mais aberto. O ângulo de rotação é de aproximadamente 12° para a isoforma A e 13° para a isoforma B. A rotação resulta em uma diminuição da área inacessível ao solvente na interface do dímero em 11% para a isoforma A (área excluída 961 Å²) e 13% para isoforma B (área excluída 935 Å²) contra 1080 Å² de área excluída na forma X (figura 5.13 (a) e (b)).

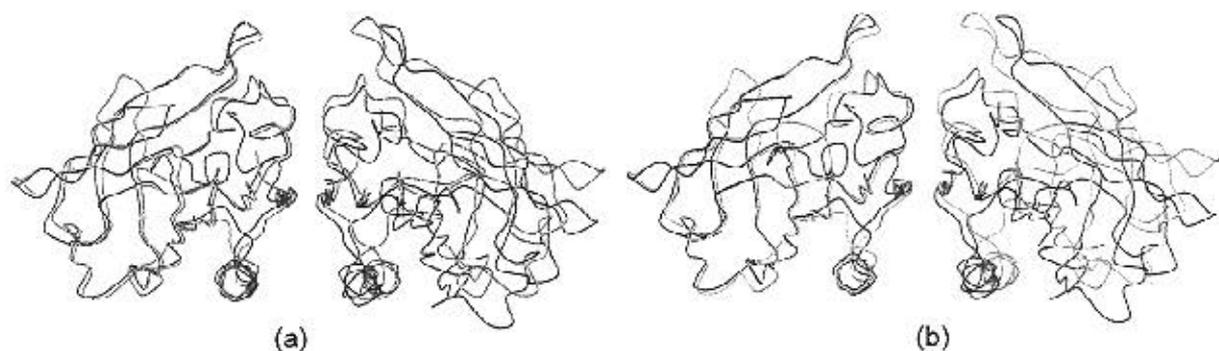


Figura 5.13

Comparação entre dímeros nas formas cristalinas X (em azul) e Y (em verde e em amarelo) para as variantes A e B. (a) Comparação para a variante A. (b) Comparação para a variante B.

5.4 A transição de Tanford

Três fatores caracterizam a transição de Tanford: exposição de um grupo carboxil, mudanças nas vizinhanças de um resíduo de tirosina e aumento da reatividade do grupo sulfidril livre (Cys 121) com reagentes que possuem Hg^{+2} . Comparando as estruturas da BLG em diferentes valores de pH foi possível identificar as mudanças estruturais que acompanham esta transição.

A área acessível ao solvente dos resíduos que possuem grupo carboxil na cadeia lateral foi inspecionada para as estruturas da BLG em pH ácido e básico. Dois resíduos tornaram-se mais expostos: Glu89 e Asp129. Estes resíduos pertencem, respectivamente, às alças EF e H α_4 que modificaram-se com a variação de pH como mostrado na figura 5.14.

Existem quatro resíduos de tirosina na seqüência da BLG. Possíveis modificações nas vizinhanças destes resíduos foram investigadas. O único resíduo que apresentou modificações em sua vizinhança foi Tyr20. Este resíduo está, especialmente, próximo da alça H α_4 que sofre alterações com o pH. Sua área acessível ao solvente aumentou de 57 Å² na mistura de isoformas A/B em pH ácido para 91 Å², na isoforma A, 112 Å², na mistura de isoforma A/B e 116 Å², na isoforma B, todas em pH básico. Chobert e colaboradores (1991) estudaram a influência do pH sobre a BLG através de estudos de hidrólise com tripsina. Nestes estudos ocorreu uma clivagem atípica na ligação Tyr20 — Ser21 que era favorecida no intervalo de

pH de 7,5 a 8,5 indicando que provavelmente o resíduo Tyr20 estivesse envolvido na transição de Tanford.

Com relação ao resíduo Cys 121 nenhuma alteração com o pH pode ser observada nas estruturas resolvidas.

Independentemente, Qin *et al.* também observaram os efeitos da transição de Tanford avaliando estruturas da BLG na forma cristalina Z para a variante A. Os resultados encontrados também apontam o envolvimento do resíduo Glu 89 na transição. Entretanto, a contribuição do resíduo Asp 129 e a mudança nas vizinhanças do resíduo Tyr20 não foram identificadas.

5.5 Deposição no PDB das estruturas cristalográficas das variantes A e B

As estruturas cristalográficas das variantes A e B foram depositadas no banco de dados PDB (<http://pdb.pdb.bnl.gov>). O código de acesso para as variantes A e B estão disponíveis e são respectivamente: 1qg5 e 1b8e.

NOTA:

As figuras 5.1, 5.2, 5.3, 5.11 e 5.13 foram feitas com os programas MOLSCRIPT (P.J. Kraulis, J. (1991) *Appl. Cryst.*, vol 24, pp 946-950) e RASTER3D (Merritt & Bacon (1997) *Meth. Enzymol.* 277, 505-524).

No presente trabalho, foram determinadas quatro estruturas da beta-lactoglobulina bovina por difração de raios-x. As amostras no estado cristalino foram obtidas na faixa de pH ácido, para a mistura de isoformas A/B, e pH básico, para a mistura de isoformas A/B, para variante A e para a variante B. Tais estruturas foram estudadas a fim de acompanhar os efeitos estruturais que ocorrem na beta-lactoglobulina com a variação de pH (transição de Tanford), bem como investigar possíveis diferenças estruturais entre as variantes A e B que viessem a justificar as diferenças nas propriedades físico-químicas destas variantes.

Comparando-se as estruturas resolvidas em pH ácido e básico foi possível identificar as mudanças estruturais sofridas pela beta-lactoglobulina com relação à transição de Tanford : duas alças, EF e H α ₄, desta proteína mudaram de conformação. Com estas mudanças, os resíduos Glu89 e Asp129 tornaram-se mais expostos e foram, assim, identificados como os resíduos carboxilados anômalos da transição de Tanford. Dos quatro resíduos de tirosina existentes na seqüência primária da proteína, Tyr20 foi o único cuja vizinhança foi afetada com as modificações na alça H α ₄ . Assim, Tyr20, juntamente com Glu89 e Asp129, estão envolvidos com a transição de Tanford.

Todas as estruturas da beta-lactoglobulina resolvidas apresentam, essencialmente, o mesmo enovelamento. Entretanto, a alça GH nas variantes A e B possui conformações distintas. A mutação Val \rightarrow Ala na posição 118 destas variantes ocasiona a formação de uma cavidade no interior da proteína que possivelmente está associada a diferença de sensibilidade frente a pressão destas variantes.

Proteínas

Neste apêndice serão tratados alguns princípios a respeito de proteínas que foram utilizados ao longo dos capítulos deste trabalho.

A.1 Constituição das proteínas

As proteínas são formadas pela união covalente de unidades fundamentais denominadas aminoácidos. Todos os aminoácidos possuem um grupo carboxila, um grupo amino, um átomo de hidrogênio e um radical (ou cadeia lateral, que distingue os aminoácidos entre si) conectados a um carbono central denominado carbono alfa (C_{α}) (figura A.1). Existem vinte tipos de aminoácidos¹ e as diferentes combinações entre eles são responsáveis pela diversidade das proteínas. Os aminoácidos estão mostrados na figura A.2 e dividem-se em: hidrofóbicos, carregados e polares. A glicina não se enquadra nestas três classes e constitui uma classe a parte.

O carbono alfa é dito assimétrico (com exceção ao C_{α} da glicina), pois existem quatro grupos substituintes diferentes entre si, e em função disto, possui duas configurações espacialmente distintas denominadas isômeros ópticos (imagens especulares não superpostas). Isômeros ópticos (ou quirais) são quimicamente idênticos nas suas reações mas possuem a propriedade de desviar o plano de uma luz polarizada para a direita ou para a esquerda.

¹ Existem outros tipos de aminoácidos que estão presentes em organismos vivos mas não em proteínas (Lehninger "Princípios de bioquímica" p. 72)

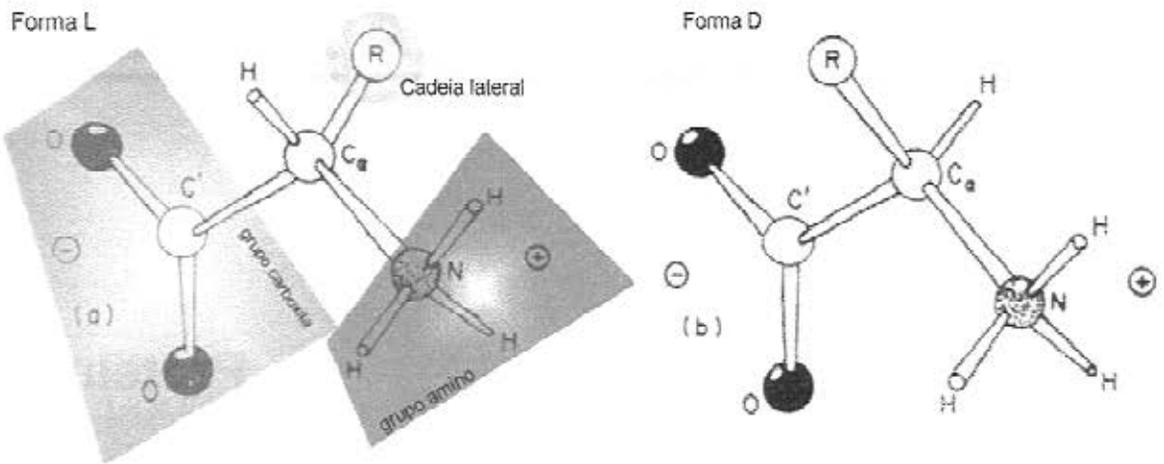
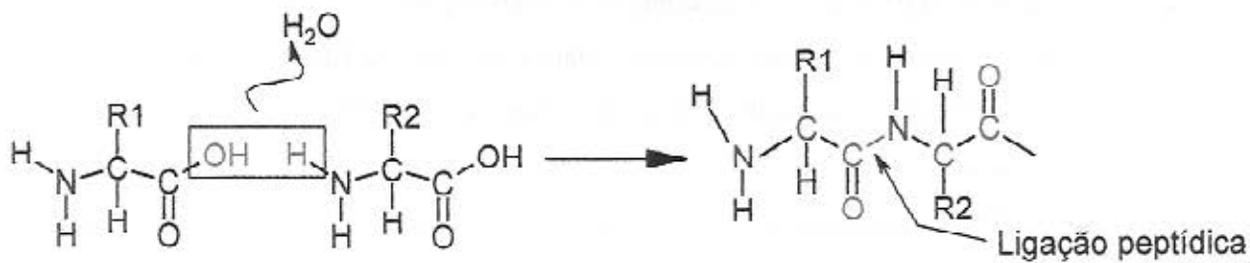


Figura A.1

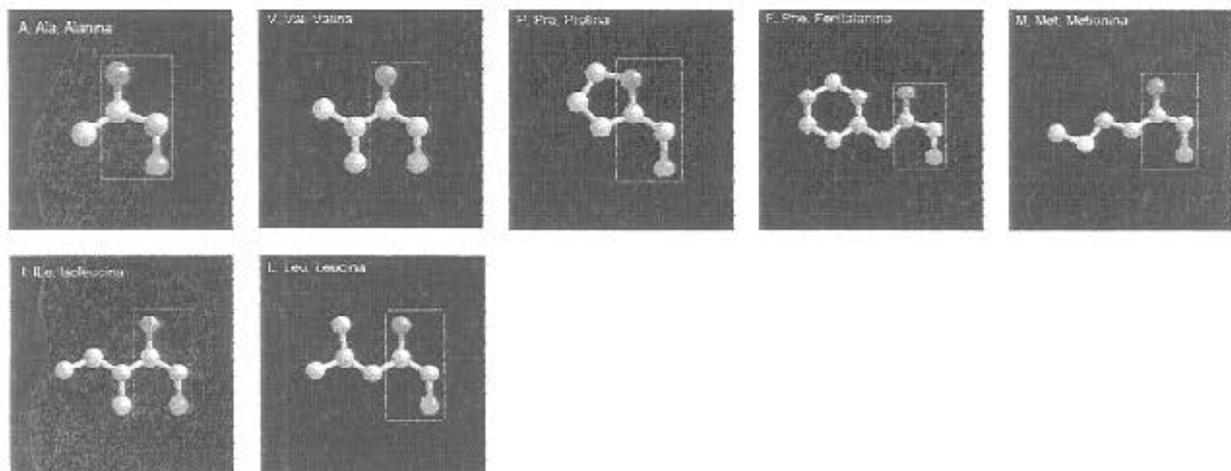
Composição dos aminoácidos e formas quirais. (a) Forma L, (b) Forma D

As formas quirais são denominadas L e D (figura A1 (a) e (b))². Entretanto, sistemas biológicos (por razões desconhecidas) utilizam-se apenas de uma das formas quirais: a forma L.

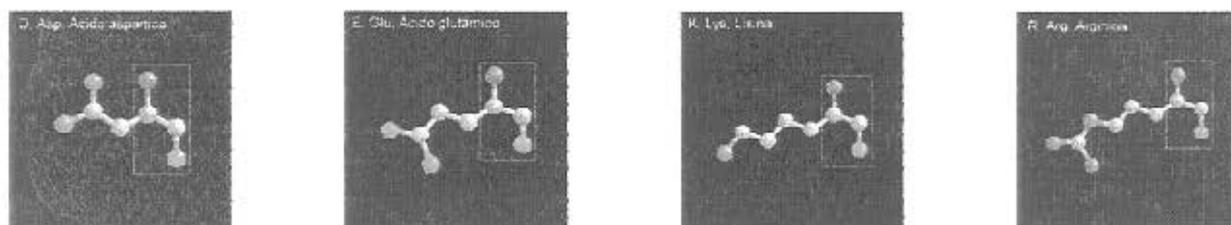
Os aminoácidos são unidos entre si através de ligações peptídicas durante a síntese protéica. Uma ligação peptídica é formada entre o grupo carboxila do primeiro aminoácido e o grupo amina do seguinte eliminando uma molécula de água como representado no esquema abaixo:



² Os símbolos L e D referem-se a configuração absoluta dos quatro substituintes ao redor do carbono assimétrico, e não à direção de rotação da luz polarizada. Os estereoisômeros que possuem configuração relacionada àquela do L-gliceraldeído são denominados L, e aqueles relacionados ao D-gliceraldeído são designados D (Lehninger "Princípios de bioquímica", p.73).



Aminoácidos carregados



Aminoácidos polares

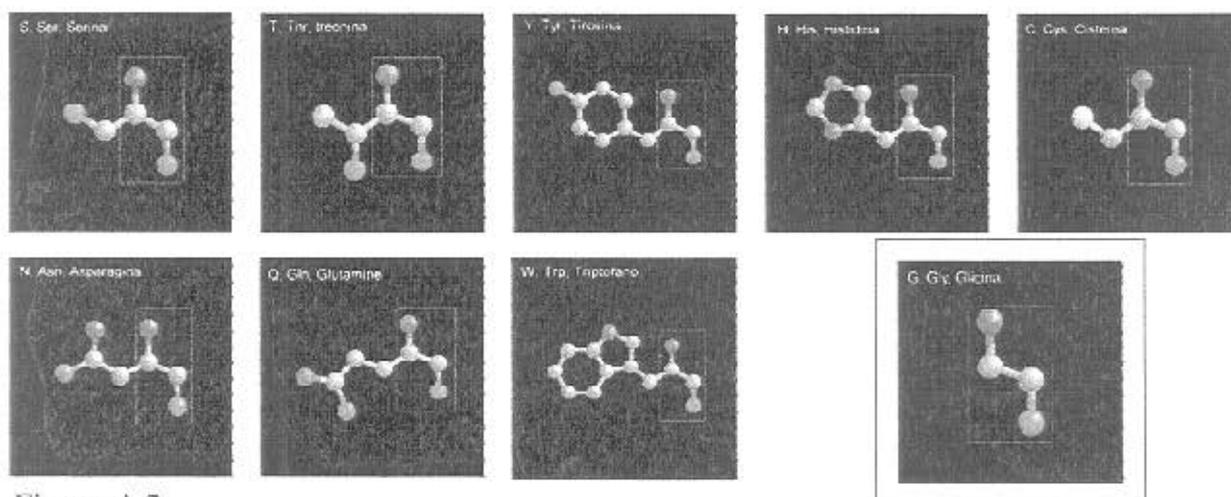


Figura A.2

Os vinte tipos de aminoácidos. A cadeia principal está destacada e os hidrogênios não estão representados

- Oxigênio
- Carbono
- Nitrogênio
- Enxofre

A cadeia assim formada é denominada cadeia polipeptídica e os aminoácidos passam a ser chamados de resíduos. Duas particularidades muito importantes da ligação peptídica são a sua planaridade e os comprimentos das ligações CN e CO que são menores que as ligações simples usuais³. Estas características restringem os graus de liberdade da cadeia polipeptídica de forma que são possíveis apenas rotações ao redor das ligações NC e CC. Rotações ao redor destas ligações são determinadas pelos ângulos ϕ e ψ e caracterizam o enovelamento (conformação) da proteína (figura A.3). A representação destes ângulos em um gráfico bidimensional de ψ contra ϕ é denominado diagrama de Ramachandran. Entretanto nem todas as combinações de ϕ e ψ são possíveis devido a impedimentos estéricos (ou combinações energeticamente desfavoráveis). Assim, um diagrama de Ramachandram é dividido em regiões permitidas e não permitidas. O aminoácido glicina é a única exceção (dos vinte existentes) podendo adotar qualquer combinação de ângulos, pois sua cadeia lateral é composta por apenas um átomo de hidrogênio (figura A.2).

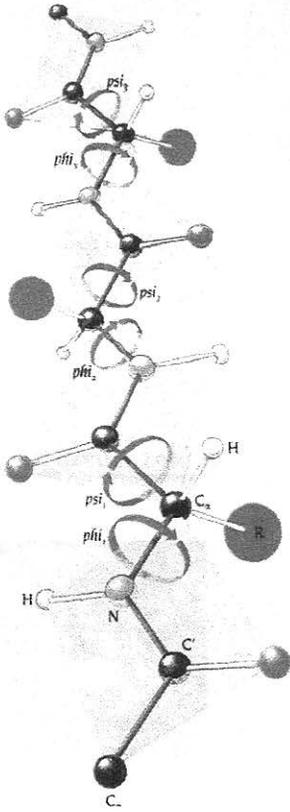


Figura A.3

Conformação da cadeia principal definida pelos ângulos ϕ e ψ ⁴

A.2 Estrutura primária, secundária, terciária e quaternária das proteínas

A seqüência de aminoácidos na cadeia polipeptídica define a estrutura primária da proteína. O enovelamento da proteína coloca resíduos distantes na estrutura primária em proximidade espacial. Interações como pontes de hidrogênio, pontes salinas e pontes dissulfeto podem ocorrer. As pontes salinas são interações que ocorrem entre grupos carregados com cargas opostas enquanto que as pontes dissulfeto são formadas entre resíduos

³ Ref. "Protein Crystallography" Blundell & Johnson.

⁴ Figura retirada do livro "Introduction to Protein Structure" Branden & Tooze e editada

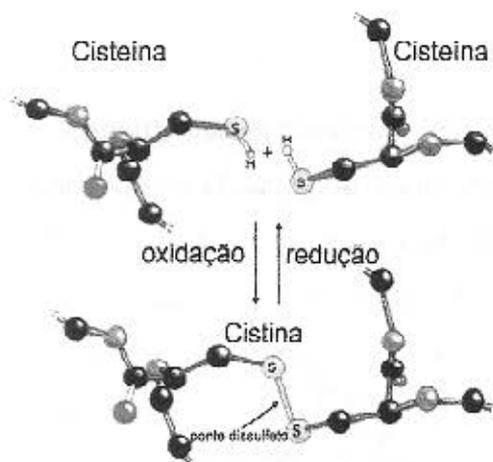


Figura A.4

Formação de uma ponte dissulfeto⁴

de cisteína que através de oxidação formam uma ligação S—S (figura A.4). A ligação entre cisteínas é denominada cistina.

Uma das razões para o enovelamento de uma proteína é manter em seu interior cadeias laterais hidrofóbicas. Como a cadeia principal é hidrofílica, acomodá-la no interior da proteína só é possível neutralizando os grupos polares N—H e C=O através de pontes de hidrogênio. As pontes de hidrogênio são obtidas através de arranjos espaciais específicos onde todos os participantes possuem os mesmos ângulos ϕ e ψ . Estes arranjos são de dois ti -

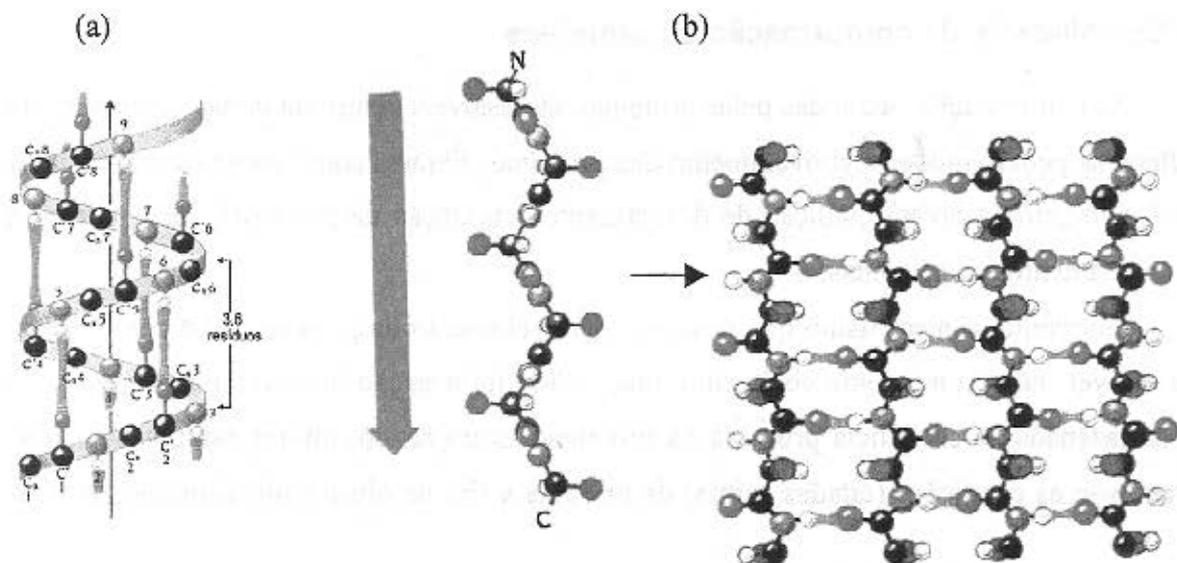


Figura A.5

Estrutura secundária das proteínas. (a) α -hélices (b) fios beta⁴

pos: α -hélices e fios beta (figura A.5 (a) e (b) respectivamente) e constituem a estrutura secundária das proteínas.

Os elementos da estrutura secundária são conectados por alças que possuem diferentes comprimentos (número de resíduos que compõem a alça) e forma irregular. O arranjo dos elementos da estrutura secundária forma a estrutura terciária da proteína que por sua vez irá compor a estrutura quaternária que contém várias cadeias polipeptídicas (figura A.6).

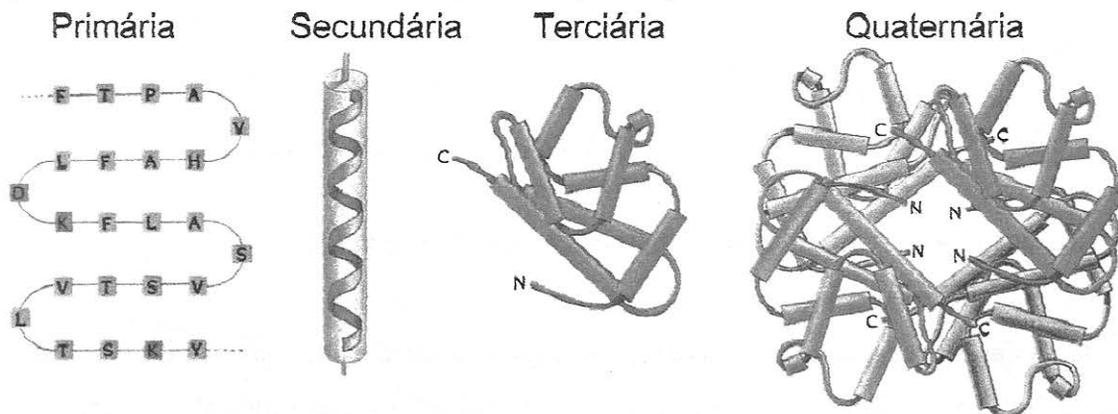


Figura A.6

Estrutura primária, secundária, terciária e quaternária das proteínas⁴

A.3 Estabilidade da conformação de proteínas

As conformações adotadas pelas proteínas são estáveis dentro de certas condições que se alteradas podem mudar o enovelamento das proteínas. Fatores como temperatura, variação de pH, aumento da pressão, adição de desnaturantes⁵, mutação de resíduos⁶, etc., interferem no enovelamento das proteínas.

Entretanto, o mecanismo que governa o enovelamento da proteína até a conformação mais estável não é totalmente conhecido, mas toda a informação necessária a este processo está armazenada na seqüência primária da proteína. Assim sendo, muitos estudos são feitos alterando-se as condições (citadas acima) da proteína a fim de obter mais informações sobre

⁵ São agentes que diminuem a estabilidade e afetam a estrutura da proteína. Exemplos de desnaturantes são guanidina-HCl (Gdn-HCl) e uréia.

⁶ Alterações na seqüência primária da proteína.

sua estabilidade. Experimentos de fluorescência são um exemplo destes estudos. Nos experimentos de fluorescência são estudados resíduos aromáticos excitando-se uma amostra de solução de proteína em um dado comprimento de onda e avalia-se a emissão destes resíduos de acordo com alguma das modificações acima citadas. Normalmente, resíduos de triptofano são estudados, pois dos aminoácidos com cadeia lateral aromática ele é o mais fluorescente e o que ocorre com menos frequência e é sensível às modificações de suas vizinhanças. A figura A.7 ilustra espectros de fluorescência com diferentes pressões do triptafano⁷.

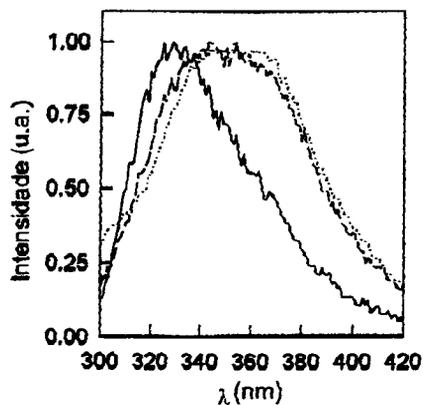


Figura A.7

Exemplo de espectros de fluorescência com diferentes pressões do triptofano

⁷ Espectro retirado do artigo da revista *Biophysical Journal* 75 (1998), 471-476 e editado.

Referências

- Aschaffenburg, R. & Drewry, J. (1955) *Nature* **176**, 218-219.
- Brownlow, S., Morais Cabral, J.H., Cooper, R., Flower, D.R., Yewdall, S.J., Polikarpov, I., North, A.C.T. & Sawyer, L. (1997) *Structure* **5**, 481-495.
- Blundell, T.L. & Johnson, L.N. (1976) *Protein Crystallography*, Academic Press, London.
- Brünger, A. T. (1992) *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT, USA.
- Brünger, A. T. (1993) *Acta Cryst.* **D49**, 24-36.
- CCP4 - Collaborative Computational Project Number 4 (1994) *Acta Cryst.* **D50**, 760-763.
- Chobert, J.-M., Dalgalarondo, M., Dufour, E., Bertrand-Harb, C. & Haertlé, T. (1991) *Biochimica et Biophysica Acta* **1077**, 31-34.
- Crowther, R.A. & Blow, D.M. (1967) *Acta Cryst.* **23**, 544-548.
- Dong, A., Matsuura, J., Allison, S.D., Chrisman, E., Manning, M.C. & Carpenter, J.F. (1996) *Biochemistry* **35**, 1450-1457.
- Dunnill, P. & Green, D.W. (1965) *J. Mol. Biol.* **15**, 147-151.
- Fällström, S.P., Ahlstedt, S. & Hanson, L.A. (1978) *Int. Arch. Allergy. Appl. Immunol.* **56**, 97-105.
- Farrell, H.M., Behe, M.J. & Enyeart, J.A. (1987) *J. Dairy Sci.* **70**, 252-258
- Green, D.W., Aschaffenburg, R., Camerman, A., Coppola, J.C., Diamand, R.D., Dunnill, P., Simmons, R.M., Komorowski, E.S., Sawyer, L., Turner, E.M.C. and Woods, K.F. (1979) *J. Mol. Biol.* **131**, 375-397.
- Hambling, S.G., McAlphine, A.S. and Sawyer, L. (1992) *Advanced Dairy Chemistry: I. Proteins* Ed. Fox, P.F. Elsevier Applied Science, New York, 141-190.
- Hendrickson, W.A. (1985) *Methods in Enzymology* **115**, 252-270.

- Hill, J.P., Boland, M.J., Creamer, L.K., Anema, S.G., Otter, D.E., Patterson, G.R., Lowe, R., Motion, R.L. and Thresher, W.C. (1996) in *Macromolecular Interactions in Food Technology, Chapter 22*, pp. 281-294, American Chemical Society, Washington, DC
- Jack, A. & Levitt, M. (1978) *Acta Cryst.* **A34**, 931-935.
- Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991) *Acta Cryst.* **A47**, 110-119.
- Katakura, Y., Totsuka, A.A., Kaminogawa, S. (1994) *Biochimica et Biophysica Acta* **1207**, 58-67.
- Kleywegt, G.J. & Jones, T.A. (1996) *Acta Cryst.* **D52**, 826-828.
- Lamzin, V.S. & Wilson, K.S. (1993) *Acta Cryst.* **D49**, 129-147.
- Laskowski, R.A. (1993) *J. Appl. Cryst.* **26**, 283-291.
- Longhi, S., Czjzek, M., Lamzin, V., Nicolas, A. (1997) *J. Mol. Biol.* **268**, 779-799.
- Malik, Z., Bottomley, R. & Austen, B. (1988) *Biochem. Society Transactions* **16**, 338-340.
- McPherson, A. (1994) *Handbook of crystal growth*, **2**, 419-463
- Monaco, H.L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L. and Eliopoulos, E.E. (1987) *J. Mol. Biol.* **197**, 695-706.
- Neurath, Robert A., Jiang, Shibo, Strick, Nathan, *et al.* (1996) *Nature Medicine* **2**, 230
- Nicholls, A., Sharp, K., & Honing, B. (1991) *Proteins: Struct., Funct., Genet.* **11**, 281-296.
- Otwinowski, Z. (1996-1997) Data processing system, version 1.9.1, *University of Virginia Patent Foundation*.
- Palmer, A.H. (1934) *J. Biol. Chem.* **104**, 359-372.
- Papiz, M.Z., Sawyer, L., Eliopoulos, E.E., North, A.C.T., Findley, J.B.C., Sirvapasadarao, R., Jones, T.A., Newcomer, M.E. and Kraulis, P.J. (1986) *Nature*, **324**, 205-231.
- Pérez, M.D. & Calvo, M. (1995) *J. Dairy Sci.* **78**, 978-988.
- Polikarpov, I. & Sawyer, L. (1995) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 5-11.
- Prez, M.D., Diaz de Villegas, M.C., Sanches, L., Aranda, P. Ena, J.M. & Calvo, M. (1989) *Biochem. J.* **106**, 1094-1097

- Qin, B.Y., Bewley, M.C., Creamer, L.K., Baker, H.M., Baker, E.N. & Jameson, G.B. (1998) *Biochemistry* **37**, 14014-14023.
- Ramachandran, G.N. & Srinivasan, R. (1961) *Nature* **190**, 159-161.
- Rossmann, M.G. & Blow, D.M. (1962) *Acta Cryst.* **15**, 24-31.
- Sheriff, S. & Hendrickson, W.A. (1987) *Acta Cryst.* **A43**, 118-121
- Spector, A.A. & Fletcher, J.E. (1970) *Lipids* **5**, 403-411.
- Tanford, C., Bunville, L.G. and Nozaki, Y. (1959) *J. Am. Chem. Soc.* **81**, 4032-4035.
- Valente-Mesquita, V.L., Botelho, M.M. & Ferreira, S.T. (1998) *Biophysical Journal* **75**, 471-476
- Weber, P.C. (1997) *Methods in enzymology*, **276**, 13-22.
- Wu, S., Pérez, M.D., Puyol, P. & Sawyer, L. (1998) manuscrito em preparação.
- Yu, H., Colucci, W.J., McLaughlin, M.L. & Barkley, M.D. (1992) *J.Am.Chem.Soc.* **114**, 8449-8454.