

Plinio Delatorra

ANÁLISE CRISTALOGRÁFICA PRELIMINAR DAS
HEMOGLOBINAS DO (PEIXE) LIPOSARCUS ANISITSI



1910001018





UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO" - UNESP

Departamento de Física

PLINIO DELATORRE

**Análise Cristalográfica Preliminar das
Hemoglobinas do (Peixe) *Liposarcus anisitsi*.**

Orientador: Prof. Dr. Walter Filgueira de Azevedo Junior

Co-orientador: Prof. Dr. Gustavo O. Bonilla-Rodriguez



Dissertação apresentada para obtenção
do título de MESTRE EM BIOFÍSICA
- Área de concentração BIOFÍSICA
MOLECULAR. Instituto de Biociências,
Letras e Ciências Exatas – IBILCE da
Universidade Estadual Paulista "Júlio de
Mesquita Filho" – UNESP.

SÃO JOSÉ DO RIO PRETO – SP
Outubro - 2000

AGRADECIMENTOS

Primeiramente agradeço a Deus porque é fiel e justo. Pois nada acontece se não for da sua vontade. Pois, puros e verdadeiros são os seus caminhos.

É com carinho muito especial que agradeço ao Prof. Dr. Walter Figueira de Azevedo Jr. pela orientação dedicada, pelo apoio, pela companheirismo e amizade, que tornaram este trabalho possível. Mas, sobretudo, por ter me ensinado que sempre e tempo de acreditar que somos capazes.

Ao Prof. Dr. Gustavo Orlando Barilla-Rodriguez pela orientação.

Ao Prof. Dr. José Roberto Ruggiero pelo apoio prestado como coordenador do Programa de Pós-Graduação, por apresentar-me ao Prof. Walter e por ter me recebido prontamente quando aqui estive em visita.

Ao Prof. Dr. João Ruggiero Neto pelos ensinamentos, pela crítica deste trabalho no Exame Geral de Qualificação e por ser um grande mestre.

Ao Prof. Dr. Johnny Rizzuto pela orientação e pela crítica deste trabalho no Exame Geral de Qualificação.

**Dedico este trabalho aos meus
filhos João Victor e Mateus,
à minha esposa Daniela.**

AGRADECIMENTOS

Primeiramente agradeço a Deus porque é fiel e justo. Pois nada acontece se não for de sua vontade. Pois justos e verdadeiros são os teus caminhos.

É com carinho muito especial que agradeço ao Prof. Dr. Walter Filgueira de Azevedo Jr. pela orientação dedicada, pelo pronto apoio, pelo companheirismo e amizade, que tornaram este trabalho possível. Mas, sobretudo, por ter me mostrado que sempre é tempo de acreditar que somos capazes.

Ao Prof. Dr. Gustavo Orlando Bonilla-Rodriguez pela co-orientação.

Ao Prof. Dr. José Roberto Ruggiero pelo apoio prestado como coordenador do Programa de Pós Graduação, por apresentar-me ao Prof. Walter e por ter me recebido prontamente quando aqui estive em visita.

Ao Prof. Dr. João Ruggiero Neto pelos ensinamentos, pela análise crítica deste trabalho no Exame Geral de Qualificação e por ser um grande mestre.

Ao Prof. Dr. Johnny Rizzieri Olivieri pela colaboração prestativa na análise deste trabalho no Exame Geral de Qualificação.



Aos Profs. Dr. Mário Sérgio Palma e Dr. Luiz Fernando Delboni pela colaboração prestativa na análise deste trabalho na defesa de dissertação.

Ao Prof. Dr. Delbonni pela colaboração prestativa na análise deste trabalho na defesa de dissertação.

A todo o corpo docente do Departamento de Física, pelos ensinamentos, pela amizade e pelo ambiente harmonioso e acolhedor.

Aos colegas do grupo de biocristalografia: Carla, David, Henrique e Márcia, Willian, Valmir, Janaína e Fernanda.

Ao meu amigo Ricardo por estar sempre comigo nesta batalha, desde a graduação, pelas leituras críticas em meus trabalhos nas horas do café.

À Janaína pelas várias leituras deste trabalho, pelas críticas e sugestões.

Ao Valmir pelas coletas de dados no Síncrotron e pelos ensinamentos em muitos momentos.

À Fernanda pelas cristalizações, pela leitura crítica deste trabalho e por essa pessoa formidável que é.

Ao Flávio pela colaboração em vários momentos.

Aos colegas Adriana, Alexandre, Ângelo, Benatti, Bernadete, Ceci, Cecília, Charles, Dulcinéia, Fábio, Fátima, Gerson, Glaúcia,

Hamilton, Marcelo, Milton, Neide, Norma, Osmar, Ocimar, Pedro, Priscilla, Raquel, Renato, Rubens, Sandra, Selma pela gostosa convivência durante todo o tempo.

Aos funcionários Barbosa, Ilva e Paulo pelo auxílio técnico prestado.

Aos meus professores de graduação na UFSCar, Prof. Dr. Antonio Lima dos Santos e Prof. Dr. Miled Hassan Y. Moussa pela confiança em mim depositada.

Aos meus colegas de graduação: Rogério, Floripa, Jesus, Vermelho, Piu, Lobinho, Nerso, Tiozão, Ricardo, Teconomia, Edson, Edsão, William, Carla, Alemão e Gilmar.

Aos funcionários da biblioteca da Ufscar: Arlete e Lauro.

À CAPES pelo suporte financeiro.

RESUMO

Este trabalho apresenta um estudo cristalográfico preliminar de duas das iso-hemoglobinas do peixe cascudo (*Liposarcus anisitsi*), denominadas fração-I e fração-IV. Dados de difração de raios X, foram coletados para ambas frações com resoluções de 2.8 Å e 2.4 Å respectivamente, usando fonte de radiação síncrotron.

As frações tiveram suas estruturas preliminares resolvidas por substituição molecular, e em ambos os casos o modelo que apresentou melhor resultado foi a carbomonoxihemoglobina da truta (*Oncorhynchus mykiss*), os coeficientes de correlação obtidos foram de 67.1 % e 62.7 % e os R_{fatores} foram 36.5 % e 39 %.

Apenas a fração-IV foi submetida a todos os passos de refinamento: refinamento de corpo rígido, refinamento posicional, dinâmica molecular e refinamento do fator de vibração térmica. Após a realização de todos os passos o R_{fator} foi de 20.0 % e o R_{free} foi 31 %.

O refinamento completo da estrutura da fração IV, deverá ser feito assim que seja obtida a sequência primária.

O apêndice I foi aceito para publicação na RBEF (Revista Brasileira de Ensino de Física). Os resultados da análise cristalográfica da

fração I da Hb do cascudo foram publicados na Acta Cryst. D (Anexo I), e a análise da fração IV foi aceita para a publicação Acta Cryst. D (Anexo II).

Palavras chaves: cristalografia de proteínas, difração de raios X, hemoglobina, peixe.

Capítulo 1 - Introdução	5
1.1 - Ligação cooperativa do oxigênio à hemoglobina	5
1.2 - Efeito Bohr	7
1.3 - Mudança conformacional da molécula de hemoglobina	8
1.4 - Introdução ao estudo das hemoglobinas de peixe	10
1.5.1 - Efeito Root	12
1.6 - Cristalização	13
1.6.1 - Processos de cristalização	14
1.6.2 - Cristalização por difusão de vapor	15
1.7 - Difração de Raios X	16
1.8 - Substituição molecular	18
1.8.1 - História do Método	17
Capítulo 2 - Objetivo	19
Capítulo 3 - Materiais e Métodos	20
3.1 - Cristalização da Fração I da Hb do <i>Lipocarus</i>	20
3.1.1 - Cristalização da Fração IV da Hb do <i>Lipocarus</i>	21
3.2 - Coleta de dados da Fração I	22
3.2.1 - Coleta de dados da Fração IV	23
3.3 - Substituição molecular	23
3.3.1 - A escolha dos modelos	24
3.3.2 - A Determinação da Estrutura	25

ÍNDICE GERAL

Capítulo 1 - Introdução

1- Introdução às hemoglobinas.....	1
1.1 – O grupo heme	3
1.2 – Ligação cooperativa do oxigênio à hemoglobina.....	5
1.3 – Efeito Bohr.....	7
1.4 – Mudança conformacional da molécula de hemoglobina.....	8
1.5 – Introdução ao estudo das hemoglobinas de peixe.....	10
1.5.1 – Efeito Root.....	12
1.6 - Cristalização	13
1.6.1 – Processos de cristalização	14
1.6.2 - Cristalização por difusão de vapor.....	15
1.7 - Difração de Raios X	16
1.8 – Substituição molecular	16
1.8.1 – Descrição do Método	17
Capítulo 2 – Objetivo	19
Capítulo 3 – Materiais e Métodos	20
3.1 – Cristalização da Fração I da Hb do <i>Liposarcus anisitsi</i>	20
3.1.1 - Cristalização da Fração IV da Hb do <i>Liposarcus anisitsi</i>	21
3.2 – Coleta de dados da Fração I.....	22
3.2.1 - Coleta de dados da Fração IV.....	23
3.3 – Substituição molecular	23
3.3.1 – A escolha dos modelos.....	24
3.3.2 – A Determinação da Estrutura	25



3.4 – Alimento da seqüência do <i>Oncorhyncus mykiss</i> via Blast II	26
3.5 – Refinamento da estrutura	28
Fig. 1	28
Capítulo 4- Resultados e Discussões	33
Fig. 3	33
4.1 – Cristalização Lahb-I	33
4.1.1 – Cristalização LaHb-IV	33
4.2 – Substituição molecular	35
4.3 – Refinamento da estrutura da LaHb-IV	39
4.4 – Análise preliminar da estrutura	45
4.5 – Perspectivas futuras	47
Fig. 10	21
Apêndice I – “Simulação da densidade eletrônica para cristais unidimensionais e bidimensionais ”	53
Fig. 13	53
Apêndice II – Método da Substituição Molecular	78
Fig. 15	48
Anexo I – “Purification, crystallization and preliminary X-Ray analysis of haemoglobin I from the Armoured Catfish <i>Liposarcus anisitsi</i> ”	93
Anexo II - "Purification, crystallization, and Patterson search of hemoglobin-IV from the armored catfish <i>Liposarcus anisitsi</i> "	96
Anexo III - O programa AMoRe	110
Anexo IV Alinhamento via Blast II	113
Referências Bibliográficas.....	115

ÍNDICE DE FIGURAS

Fig. 1	2
Fig. 2	4
Fig. 3	6
Fig. 4	7
Fig. 5	8
Fig. 6	9
Fig. 7	10
Fig. 8	13
Fig. 9	15
Fig. 10	21
Fig. 11.....	22
Fig. 12	36
Fig. 13	43
Fig. 14	44
Fig. 15	48
Fig. 16	49
Fig. 17	50
Fig. 18	51
Fig. 19	52

CAPÍTULO 1

INTRODUÇÃO

1 – Introdução às hemoglobinas

Hemoglobinas (Hbs) são macromoléculas protéicas que estão contidas nos glóbulos vermelhos do sangue (eritrócitos) (fig. 1). Sua principal função é o transporte de oxigênio dos órgãos captadores para os tecidos via respiração aeróbica. Nos seres humanos cerca de 96% das hemoglobinas do sangue arterial, que passam dos pulmões para os tecidos, encontram-se saturadas de oxigênio, mas apenas 64% das Hbs do sangue venoso que retornam ao coração, apresentam-se saturadas de oxigênio. Portanto, as Hbs liberam cerca de um terço das moléculas de oxigênio que transportam no trajeto dos pulmões aos tecidos (Lenhinger *et al*, 1995).

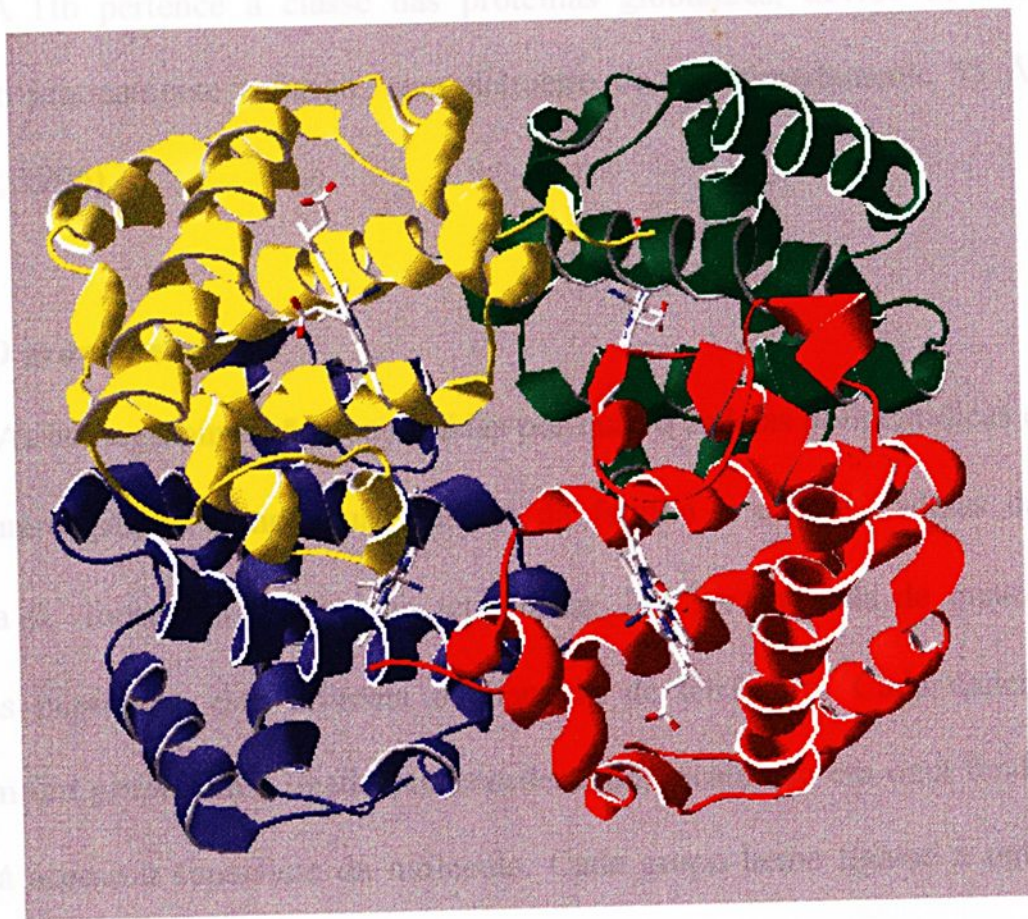


Fig. 1 – Modelo com elementos da estrutura secundária e os grupos hemes da hemoglobina da Truta, onde cada cadeia polipeptídica está colorida com uma cor distinta, e os grupos hemes são destacados.

Aproximadamente um terço da massa do sangue de mamíferos é composta de hemoglobinas, sendo esta uma proteína alostérica de papel fisiológico importantíssimo, pois um ser humano adulto consome o equivalente a 250ml de oxigênio puro por minuto. Este oxigênio é usado para prover energia para todos os tecidos e órgãos do corpo (Lenhinger *et al*, 1995).

A Hb pertence à classe das proteínas globulares, devido ao seu formato praticamente esférico, com diâmetro de aproximadamente 55 Å (Stryer, 1996).

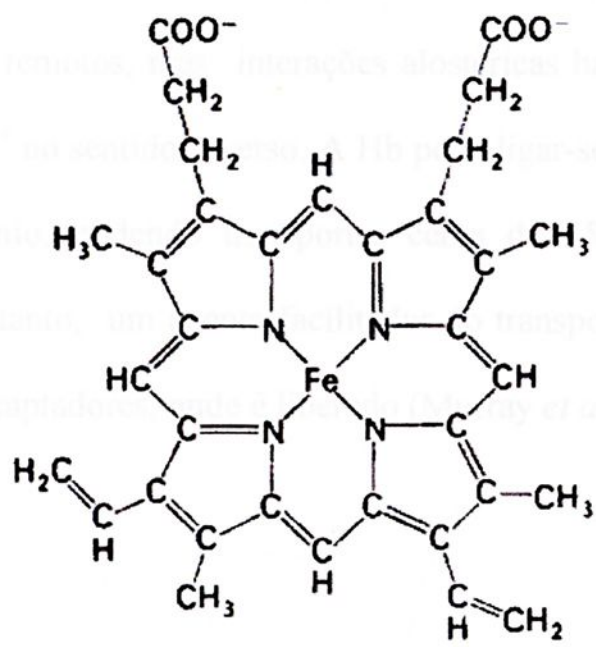
1.1 – O grupo heme

As hemoglobinas são constituídas por quatro cadeias polipeptídicas e por quatro grupos heme. Duas cadeias da família α e duas cadeias da família β , formando um arranjo tetraédrico. Por ser formada de quatro cadeias peptídicas esta proteína é chamada de tetrâmero. Cada cadeia contém um grupo heme ligado por coordenação, localizado em uma fenda que dá acesso à superfície da molécula. Cada grupo heme liga-se a uma molécula de oxigênio (Stryer, 1996).

O grupo prostético heme, responsável pela cor vermelha característica do sangue, consiste de um complexo protoporfirina IX, que é constituído de quatro anéis pirrólicos, ligados por pontes α -metênicas ($-\text{CH}=\text{}$), para formar um anel tetrapirrólico planar (fig. 2). No heme os grupos substituintes são: quatro metilas, duas vinilas e duas cadeias laterais de propionato. A protoporfirina faz ligação planar central com o átomo de ferro (Fe^{2+}) (Murray *et al*, 1988).

Fig. 2 – Grupo Prostético Heme: a) modelo esquemático b) modelo em esferas e bastões. Fonte: extraído do Site <http://archi.unl.pt/~cruzeiro/cruzeiro/04-105/08S2Mx36.html>

a)



b)

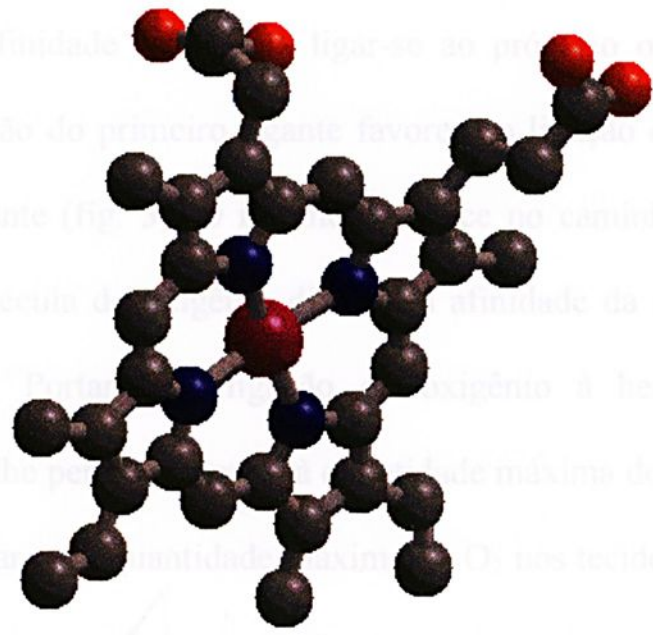


Fig. 2 – Grupo Prostético Heme: a) modelo esquelético b) modelo em esfera e bastões. Figura extraída do Site: <http://web1.andrew.cmu.edu/course/09-105/S98Slides36.html>

Como visto, a função primária da hemoglobina é ligar-se ao oxigênio, e transportá-lo na corrente sanguínea desde os órgãos captadores até os tecidos mais remotos, mas interações alostéricas habilitam a Hb a transportar CO_2 e H^+ no sentido inverso. A Hb pode ligar-se ao CO_2 após a liberação de oxigênio, podendo transportar cerca de 15% do CO_2 do sangue. Sendo, portanto, um agente facilitador do transporte de CO_2 dos tecidos aos órgãos captadores, onde é liberado (Murray *et al*, 1988).

1.2 – Ligação cooperativa do oxigênio à hemoglobina

A ligação de um oxigênio à molécula de hemoglobina causa um aumento na afinidade desta para ligar-se ao próximo oxigênio. Isso mostra que a ligação do primeiro ligante favorece a ligação do segundo e assim sucessivamente (fig. 3). O mesmo acontece no caminho inverso: a perda de uma molécula de oxigênio diminui a afinidade da proteína pelo próximo ligante. Portanto, a ligação do oxigênio à hemoglobina é cooperativa o que lhe permite ligar-se à quantidade máxima de O_2 no órgão respiratório e liberar uma quantidade máxima de O_2 nos tecidos periféricos (Stryer, 1996).

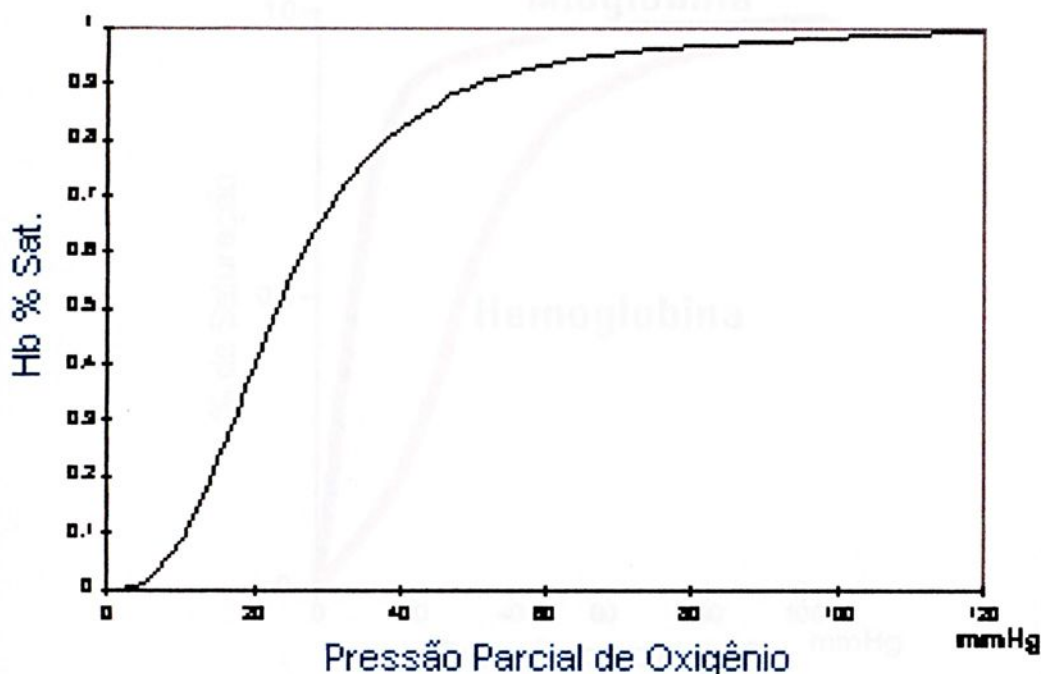


Fig. 3 - Curva de dissociação do oxigênio da molécula de hemoglobina.

Adaptada do site:

<http://www.wellesley.edu/Chemistry/chem227/bindprotns/hemoglobin.htm>

O caráter cooperativo da ligação do oxigênio a Hb, fica mais notório quando comparado à mioglobina. Suas curvas de dissociação de oxigênio (fig. 4), mostram que para uma mesma pressão parcial de oxigênio a mioglobina apresenta uma porcentagem de saturação maior do que a Hb. Isto significa que a mioglobina possui uma maior afinidade para o oxigênio do que a Hb. A afinidade de ambas pelo oxigênio, pode ser explicada pelo formato sigmoide da curva de dissociação da Hb e pelo formato hiperbólico da mioglobina (Stryer, 1996).

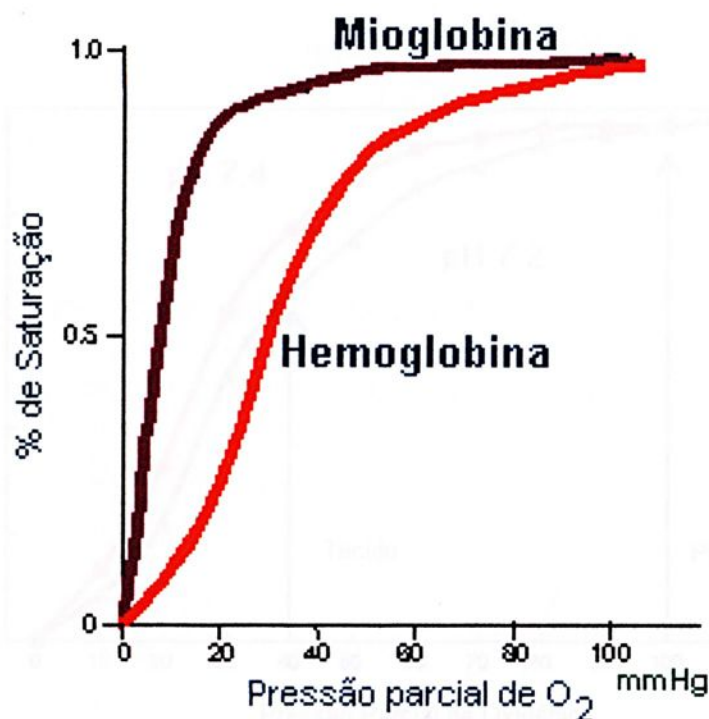


Fig. 4 – Comparação entre as curvas de dissociação de oxigênio das moléculas de hemoglobina e mioglobina.

1.3 - Efeito Bohr

Em suas observações experimentais Bohr descobriu que, a um pH abaixo do fisiológico (7,4), a hemoglobina não se ligava ao oxigênio com a mesma facilidade. Esta influência do pH na captação ou liberação do oxigênio é conhecido como efeito Bohr (Lehninger *et al*, 1995). Este efeito pode ser observado no gráfico de saturação de Hb (fig. 5).

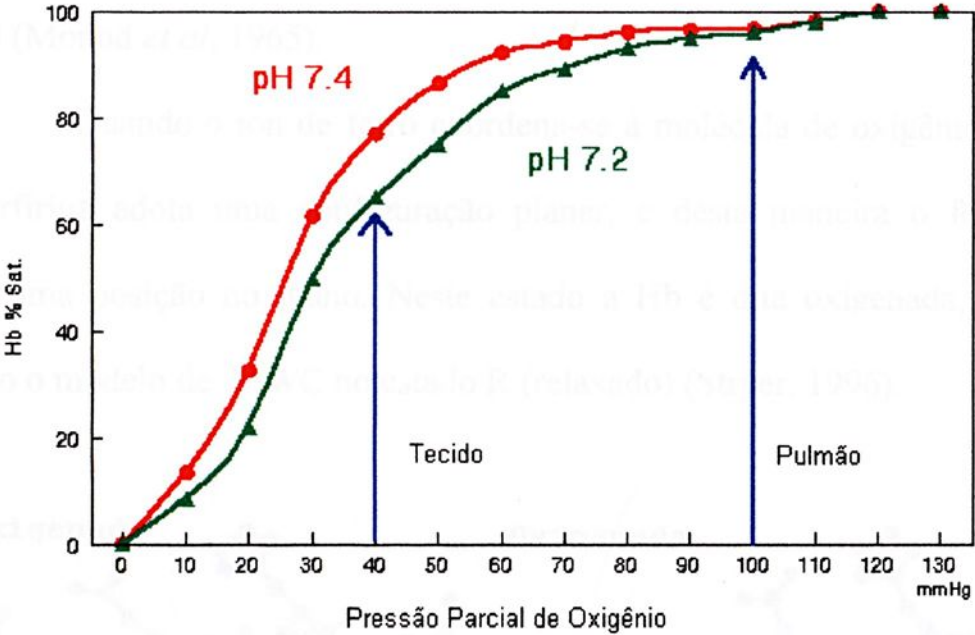


Fig. 5 – Efeito do pH na afinidade da hemoglobina pelo oxigênio. Extraída do Site: <http://ella.austin.cc.tx.us/~emeyerth/bohr.htm>

Na figura 5 observa-se que a um pH mais baixo, a uma mesma pressão de oxigênio, a Hb apresenta-se menos saturada de oxigênio.

1.4 – Mudança conformacional da molécula de hemoglobina

Quando o grupo heme não está ligado a uma molécula de oxigênio, observa-se que ele adota uma formação não planar. Isto acontece devido ao fato do íon de ferro ser empurrado para fora do plano da porfirina, cerca de 0,4 Å, em direção à histidina proximal a fim de coordena-se a ela (fig. 6). Esta ligação não planar é característica da forma

desoxigenada da hemoglobina, também conhecido como estado (T) Tenso, seguindo-se a nomenclatura do modelo de Monod, Wyman e Changeux (MWC) (Monod *et al*, 1965).

Quando o íon de ferro coordena-se à molécula de oxigênio, a protoporfirina adota uma configuração planar, e desta maneira o ferro assume uma posição no plano. Neste estado a Hb é dita oxigenada, ou seguindo o modelo de MWC no estado R (relaxado) (Stryer, 1996).

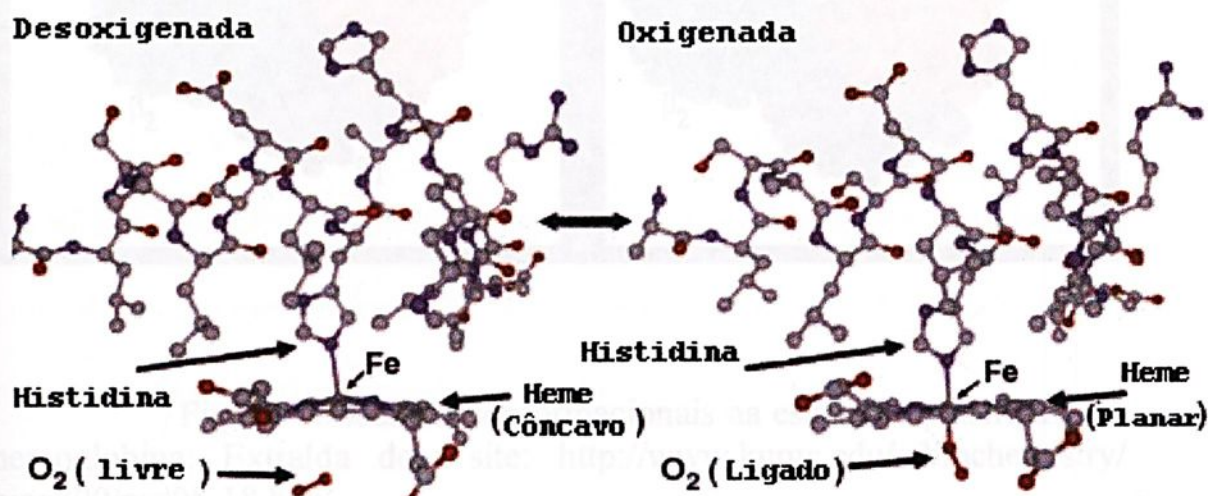


Fig. 6 – Mudanças conformacionais induzidas pelo movimento do íon de ferro.

Após ligar-se ao O₂, em pH fisiológico, a hemoglobina muda seu estado conformacional, provocando a liberação de prótons.

Há duas regiões de contato entre as cadeias α e β : $\alpha 1\text{-}\beta 1$ e $\alpha 2\text{-}\beta 2$. Após a mudança do estado tenso (estrutura desoxi) para o estado

relaxado (estrutura oxi), o dímero α_1 - β_1 rotaciona 15° em relação ao outro dímero. Fazendo com que alguns átomos na interface entre estes pares se desloquem de até 6 Å (fig. 7) (Murray *et al*, 1988).

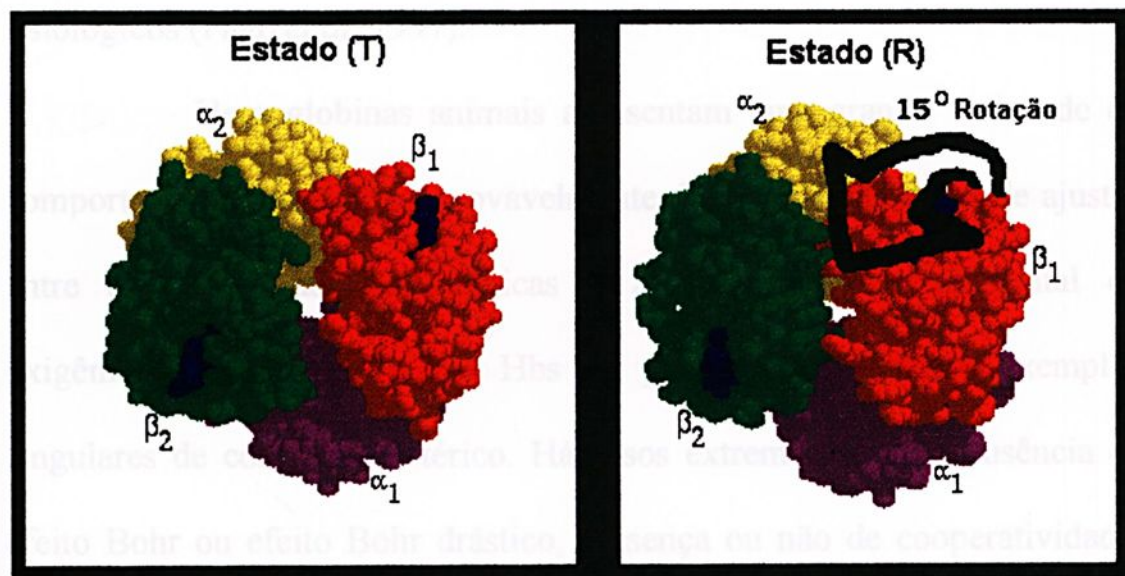


Fig. 7 - Mudanças conformacionais na estrutura quaternária da hemoglobina: Extraída do site: <http://www.kumc.edu/biochemistry/bioc800/pro05-18.htm>.

1.5 – Introdução ao estudo das hemoglobinas de peixes

Hemoglobinas de peixes têm sido estudadas nestes últimos anos, principalmente devido às suas várias propriedades funcionais e à multiplicidade de componentes presente em um único indivíduo. Em contraste com os mamíferos, os peixes mostram uma larga variação no

número de iso-hemoglobinas, apresentando também uma grande variedade de comportamentos funcionais, em termos de ligação cooperativa de oxigênio e controle alostérico por uma variedade de fatores químicos. Isto lhes confere uma grande habilidade de adaptar-se em diferentes ambientes fisiológicos (Fago *et al*, 1997).

Hemoglobinas animais apresentam uma grande variedade de comportamentos funcionais provavelmente decorrentes da gama de ajustes entre as necessidades fisiológicas e a disponibilidade ambiental de oxigênio. Nesta categoria, as Hbs de peixes têm mostrado exemplos singulares de controle alostérico. Há casos extremos como a ausência de efeito Bohr ou efeito Bohr drástico, presença ou não de cooperatividade, tetrâmeros com variada estabilidade face à dissociação e novos arranjos das cadeias de globina (De Young *et al*, 1994).

Estudos estruturais recentes de Hbs de peixes, realizados por difração de raios X, mostram que esta apresenta grande similaridade com a estrutura tridimensional da Hb humana (Camardella *et al*, 1992; Ito *et al*, 1995; Tame *et al*, 1996; Mylvaganam *et al*, 1996), com pequenas diferenças nas regiões N e C terminais, porém a identidade seqüencial entre as duas é pequena.

1.5.1 - Efeito Root

A grande variação, em diferentes organismos, das propriedades alostéricas das Hbs, especialmente as ligações com oxigênio, fazem desta proteína um paradigma para estudos adaptativos das espécies. Uma das adaptações mais interessantes é a do peixe Telost (Mylvaganam *et al*, 1996), que usa sua hemoglobina não somente para levar o oxigênio das guelras para os tecidos, mas também para oxigenar seus olhos contra uma pressão hidrostática muito grande.

Um fenômeno importante, qualitativamente similar ao efeito Bohr em Hbs humanas, encontrado em Hbs de peixes é o efeito Root, descoberto por R.W. Root (Root, 1931). A única função biológica importante do Efeito Root em Hbs é a de superar a ação do aumento do gradiente de O_2 , proveniente da sua extrema sensibilidade às variações de pH, caracterizada pela redução drástica da cooperatividade e da afinidade por ligantes em valores de pH baixo (Mylvaganam *et al*, 1996) (Ito *et al*, 1995).

O efeito Root pode ser considerado como um efeito Bohr exagerado, no qual os prótons servem como efetores alostéricos que equilibram a alta afinidade do estado conformacional T, com a baixa afinidade do estado conformacional R. Portanto, as Hbs que apresentam este efeito não podem der saturadas completamente a pHs baixos (fig. 8) (Mylvaganam *et al*, 1996).



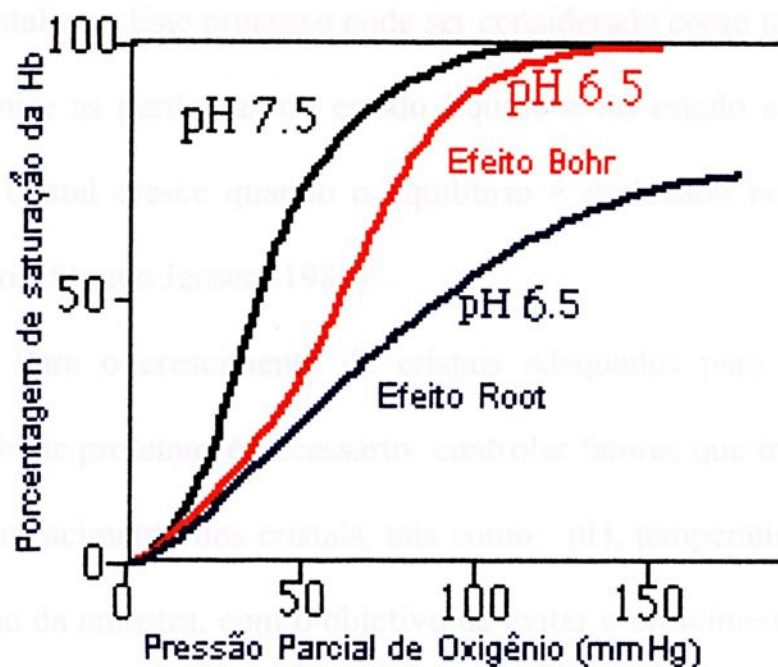


Fig. 8 – Curva de saturação para hemoglobinas que apresentam Efeito Bohr e Efeito Root.

1.6 - Cristalização

A determinação da estrutura de uma macromolécula biológica fica grandemente facilitada se um grande número de moléculas idênticas puderem ser arranjadas de maneira regular. Esta condição pode ser obtida através da cristalização. Com a obtenção de um cristal é possível usar o método da difração de raios X, para coleta de dados e realizar a análise da estrutura do cristal (Stout e Jensen, 1989).

Cristalização é um processo de ordenação, no qual os átomos ou moléculas assumem posições regulares no estado sólido. Em geral, o processo de cristalização, em seu estágio inicial, envolve o fenômeno de nucleação, onde um núcleo de crescimento é formado, através da deposição

de faces cristalinas. Este processo pode ser considerado como um equilíbrio dinâmico entre as partículas no estado líquido e no estado sólido. Desta maneira, o cristal cresce quando o equilíbrio é deslocado no sentido da solidificação (Stout e Jensen, 1989).

Para o crescimento de cristais adequados para estudos em cristalografia de proteínas é necessário controlar fatores que influenciem a cinética de crescimento dos cristais, tais como: pH, temperatura, pureza e concentração da amostra, com o objetivo de evitar o crescimento rápido de vários núcleos pequenos e indesejáveis (Drenth, 1994).

1.6.1 – Processos de Cristalização (Drenth, 1994)

Um primeiro passo na tentativa de obter um bom cristal, que sirva para a coleta de dados de difração de raios X, é certificar-se da pureza do material a ser utilizado.

O segundo passo é obter a condição de cristalização:

- (1) - Encontrar o solvente adequado para a proteína,
- (2)- O solvente é normalmente uma solução aquosa tamponante ou orgânica.
- (3)- A solução é levada à super saturação com o objetivo de provocar a nucleação.

Neste passo, para a formação do núcleo é necessário um aumento energético na tensão superficial, pois uma barreira energética tem que ser

superada para iniciar a cristalização. Esta barreira é facilmente superada a um alto grau de saturação.

- (4)- Após o início do crescimento do cristal, novas moléculas são ligadas a superfície do cristal, provocando um crescimento lento do mesmo e um decréscimo nos valores de supersaturação.

Os métodos mais utilizados para cristalização de proteínas são: diálise, interface de difusão livre, "Batch" e difusão de vapor.

1.6.2 – Cristalização por difusão de vapor

Um dos métodos usuais em cristalização por difusão de vapor é o método da gota suspensa (fig. 9). Este método consiste no preparo de gotas, com volume variando normalmente de 1 a 10 μl de solução de proteínas, misturadas em igual quantidade de solução precipitante em uma lamínula de vidro.

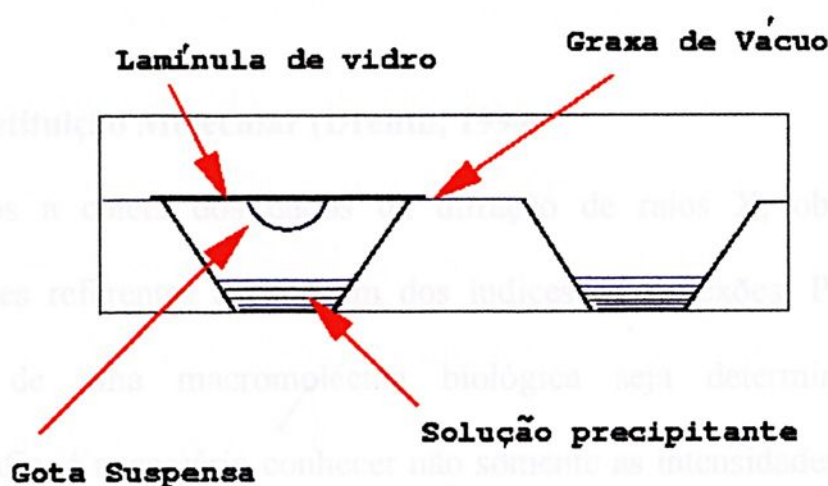


Fig. 9 – Cristalização pelo método da gota suspensa.

1.7 – Difração de Raios X (Blundell e Johnson, 1976)

A difração de raios X por cristais, ocorre como um resultado da interação dos raios X com a distribuição eletrônica de carga do retículo cristalino. A natureza ordenada da distribuição eletrônica de cargas, onde mais elétrons estão distribuídos em torno do núcleo atômico, com um arranjo regular e com uma periodicidade translacional, faz com que a superposição de espalhamentos de raios X formem regiões de interferência construtivas e destrutivas conforme a diferença de fase, produzindo um padrão de difração (apêndice I).

A única quantidade medida em um experimento de difração por um cristal é o conjunto de intensidades e suas posições relativas:

$$I_{hkl} \propto |F_{hkl}|^2$$

onde a intensidade é proporcional ao quadrado do fator de estrutura.

1.8 – Substituição Molecular (Drenth, 1994)

Após a coleta dos dados da difração de raios X, obtêm-se as intensidades referentes a cada um dos índices de reflexões. Para que a estrutura de uma macromolécula biológica seja determinada, por cristalografia, é necessário conhecer não somente as intensidades devido a cada uma das reflexões, mas também à fase relativa de cada uma delas.



Porém nenhuma informação acerca das fases pode ser obtida no experimento de difração.

Um dos métodos possíveis de resolver estruturas de macromoléculas biológicas, entre outros, é o método da substituição molecular (apêndice II), porém o uso deste método requer a existência de uma estrutura já resolvida, que tenha algum tipo de similaridade com a que deseja-se resolver. Este método tem se tornado cada vez mais usual, com o aumento do número de estruturas depositadas no banco de dados de proteínas (PDB) <http://www.rcsb.org>.

1.8.1 – Descrição do método (Drenth, 1994)

O método da substituição molecular é usado com o objetivo de encontrar uma primeira estrutura, passível de refinamento, da proteína desconhecida. Partindo do seguinte princípio: se as proteínas possuem alguma similaridade, seus padrões de difração também deverão possuir alguma concordância. Não exatamente o padrão de difração, mas sim as transformadas de Fourier dos fatores de estruturas. O método explora estas concordâncias com o objetivo de obter as fases iniciais da proteína desconhecida. Em uma primeira aproximação as fases do modelo são usadas na proteína desconhecida.

Transferir a estrutura molecular do modelo para o arranjo cristalino da proteína desconhecida é o principal problema do método. Para que isso



seja feito é necessário uma orientação adequada e um posicionamento preciso do modelo na cela unitária alvo, através de aplicações sucessivas de translações e rotações, até que um melhor índice de concordância entre os fatores de estrutura observados e calculados seja obtido.

As hemoglobinas de peixes têm sido grandemente estudadas nos últimos anos, isto se deve principalmente as variações em suas propriedades funcionais e a presença de iso-hemoglobinas. Alguns pesquisadores propuseram uma regra adaptativa para esta multiplicidade funcional e para a presença das iso formas. Por outro lado, outros pesquisadores não encontrando diferenças significativas nas iso-hemoglobinas, consideraram que estas são remanescentes evolucionários das hemoglobinas.

Independente da discordância de idéias, o estudo cristalográfico de estruturas de hemoglobinas, deverá contribuir para melhorar o entendimento das funções das hemoglobinas.

De uma forma geral, este trabalho tem por objetivo o estudo preliminar da fração IV da hemoglobina do peixe cascudo, procurando destacar algumas similaridades estruturais desta Hb com outras Hbs de peixes. Não é objetivo deste trabalho a realização de um estudo funcional da estrutura da L α Hb-IV.

Capítulo 2

Objetivos

As hemoglobinas de peixes têm sido grandemente estudadas nos últimos anos, isto se deve principalmente as variações em suas propriedades funcionais e a presença de iso-hemoglobinas. Alguns pesquisadores propuseram uma regra adaptativa para esta multiplicidade funcional e para a presença das iso formas. Por outro lado, outros pesquisadores não encontrando diferenças significantes nas iso-hemoglobinas, consideraram que estas são remanescentes evolucionário das hemoglobinas.

Independente da diversidade de idéias, o estudo cristalográfico de estruturas de hemoglobinas, deverá contribuir para melhorar o entendimento das funções das hemoglobinas.

De uma forma geral, este trabalho tem por objetivo o estudo preliminar da fração IV da hemoglobina do peixe cascudo, procurando destacar algumas similaridades estruturais desta Hb com outras Hbs de peixes. Não é objetivo deste trabalho a realização de um estudo funcional da estrutura da LaHb-IV.



CAPÍTULO 3

Materiais e métodos

3.1.1 – Cristalização da fração I da hemoglobina do *Liposarcus anisitsi*. (LaHb – I).

A forma oxigenada da LaHb – I, foi concentrada à 10 mg/ml em 50mM Tris em tampão com pH 8.5. Alguns cristais foram obtidos em diferentes condições de cristalização, utilizando-se o método da difusão de vapor em gota suspensa e a técnica da matriz esparsa (Jancarick e Kim, 1991), sendo que os melhores foram obtidos após três dias de crescimento. A fig. 10 ilustra uma caixa de cultura de células linbro, utilizada no método da matriz esparsa. As gotas foram montadas com 7 μ l de solução de hemoglobina (10 mg/ml) e com igual volume da solução do reservatório. A solução usada no reservatório foi 0.1 M HEPES pH 7.4 e 1.4 M de Acetato de Sódio. Os dois melhores cristais foram transportados para um

tubo capilar de borosilicato de 1.0 mm, para coleta de dados de difração de raios X (anexo-I).

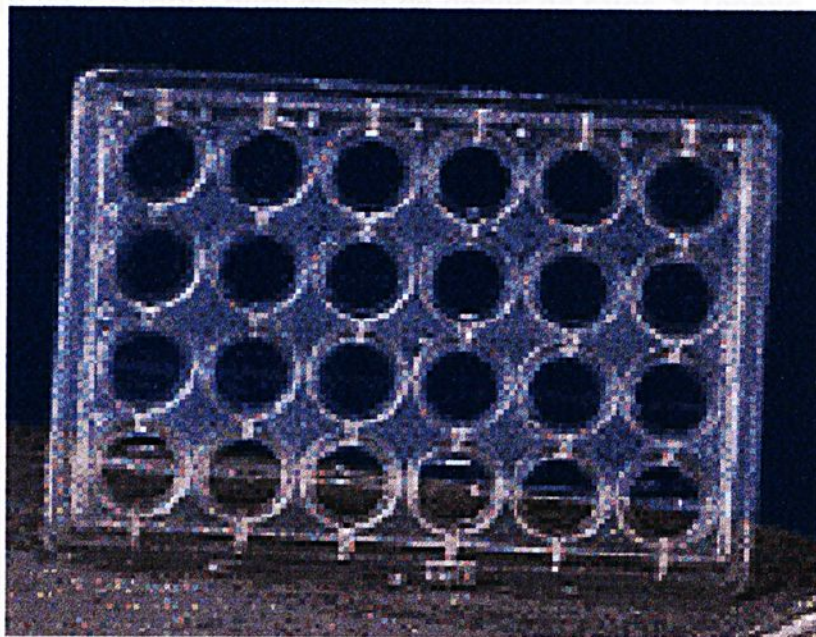


Fig. 10 – Caixa de cultura de células linbro.

3.1.2– Cristalização da fração IV da hemoglobina do *Liposarcus anisitsi* (LaHb – IV).

A forma oxigenada da LaHb – IV, foi concentrada à 10 mg/ml em 50mM Tris em tampão com pH 8.5. Alguns cristais foram obtidos em diferentes condições de cristalização, utilizando-se o método da difusão de vapor em gota suspensa e a técnica da matriz esparsa (Jancarick e Kim, 1991), sendo que os melhores foram obtidos após dois dias de crescimento. As gotas foram montadas com 7 μ l de solução de hemoglobina (10 mg/ml)

e com igual volume da solução do reservatório. A solução usada no reservatório foi de 0.1 M Tris pH 8.5, 0.2 M Acetato de Sódio e PEG 4000 30%. O melhor cristal (fig. 11) foi transportado para um tubo capilar de borosilicato de 1.0 mm, para coleta de dados de raios X (anexo-II).



Fig. 11 – Cristal da Fração IV da Hemoglobina do *Liposarcus anisitsi*

3.2.1 - Coleta de dados – fração I

Os dados de difração de raios X foram coletados para os dois cristais da LaHb – I a temperatura de 8 °C, a um comprimento de onda de 1,38 Å, usando fonte de radiação síncrotron da estação PCr do laboratório Nacional de Luz Síncrotron, LNLS, em Campinas. Foi usada uma placa

detectora de imagem MAR de 34.5 cm (MAR Research). Dados preliminares da difração de raios X para os cristais I e II foram processados a resoluções de 2.77 Å e 2.86 Å respectivamente, usando o programa Denzo (Otwinowski, 1993) e escalado com o programa SCALEPACK (Otwinowski, 1993).

3.2.2 - Coleta de dados – fração IV

Os dados de difração de raios X foram coletados para o cristal da LaHb – IV a temperatura de 8 °C, a um comprimento de onda de 1,547 Å, usando fonte de radiação síncrotron da estação PCr do laboratório Nacional de Luz Síncrotron, LNLS, em Campinas. Foi usada uma placa detectora de imagem MAR de 34.5 cm (MAR Research) com tempo de exposição de 2 minutos por imagem, com o cristal a uma distância de 150mm do detector. Com imagens foram coletadas usando uma amplitude de oscilação de 0.8°. Dados preliminares da difração de raios X foram processados a resolução de 2.4 Å, usando o programa Denzo (Otwinowski, 1993) e escalado com o programa SCALEPACK (Otwinowski, 1993)

3.3 – Substituição Molecular

Em cristalografia, existem quatro métodos que podem ser usados para a determinação dos ângulos de fase: MIR-(Substituição Isomórfica Múltipla), Substituição Molecular, Dispersão Anômala e



Métodos Diretos. Neste trabalho o método escolhido foi o da substituição molecular, devido a existência de aproximadamente 130 estruturas de hemoglobinas depositadas no banco de dados de proteínas (*Protein Data Bank-PDB*), que podem ser usadas como modelos.

3.3.1 – A escolha dos modelos

Os modelos escolhidos para a realização da substituição molecular foram os dímeros de oito hemoglobinas de peixes e uma humana, todas com coordenadas depositadas no (PDB) e com código de acesso listado na tabela 1. A utilização da Hb humana como modelo para comparação é uma prática comum no estudo cristalográfico de Hbs.

Tabela 1. Lista de hemoglobinas usadas como modelos de procura

Código de Acesso	Proteína	Resolução (Å)
1HHO	Oxihemoglobina de <i>Homo sapiens</i>	2.1
1T1N	Carbomonoxihemoglobina de peixe (<i>Trematomus akajei</i>)	2.2
1OUU	Carbomonoxihemoglobina de peixe (<i>Oncorhyncus mykiss</i>)	2.5
1CG8	Carbomonoxihemoglobina de peixe (<i>Dasyatis akajei</i>)	1.9
1HBH	Desoxihemoglobina de peixe (<i>Pagothenia bernacchii</i>)	2.2
1OUT	Desoxihemoglobina de peixe (<i>Oncorhyncus mykiss</i>)	2.3
1PBX	Carbomonoxihemoglobina de peixe (<i>Pagothenia bernacchii</i>)	2.5
1SPG	Carbomonoxihemoglobina de peixe (<i>Leiostomus xanthurus</i>)	1.95
1CG5	Desoxihemoglobina de peixe (<i>Dasyatis akajei</i>)	1.6

3.3.2 – A determinação da estrutura

Os primeiros modelos das estruturas moleculares da LaHb-I e LaHb-IV, foram obtidos usando o programa AMoRe (anexo III) (Navaza, 1994). Este programa faz parte do pacote CCP4 (*Collaborative Computing Project No. 4*) e usa o método da substituição molecular, para a determinação de estruturas moleculares. Ele é dividido basicamente em três etapas:

- 1) Avaliação dos dados obtidos através da difração de raios X e cálculo dos coeficientes de Fourier do modelo;
- 2) Rotação;
- 3) Translação;
- 4) Refinamento de corpo rígido.

Como o programa AMoRe lê somente arquivos de reflexão no formato mtz é necessário a conversão para este formato. Dentro do CCP4, o programa que realiza este trabalho é o “**f2mtz**”.

Após a devida formatação dos dados, inicia-se a primeira etapa no processo de Substituição Molecular, via AMoRe. As rotinas utilizadas nesta etapa são: o “**Sorting**” e o “**Tabling**”. A rotina “**Sorting**” ordena, empacota e avalia a qualidade dos dados medidos. As estatísticas de saída deste arquivo apresentaram, para ambos os casos LaHb-I e LaHb-IV, uma relação acima de 90% de completeza dentro da faixa de resolução analisada. A rotina “**Tabling**”, calcula os coeficientes contínuos de Fourier



do modelo, que são obtidos através da transformada de Fourier da densidade eletrônica, usando como base as coordenadas atômicas da molécula modelo.

A segunda etapa é rotacionar o modelo de maneira que ele possa ter a melhor orientação possível. A rotina do programa AMoRe, que realiza a rotação do modelo é o **“Rotting”**. Neste passo o programa procura pelos melhores picos no mapa de Patterson. Este mapa não apresenta como resultado posições atômicas individuais, por se tratar de um mapa vetorial. Seus picos apresentam as posições dos vetores entre os átomos, podendo ser tanto intramoleculares quanto intermoleculares. Sendo, o primeiro dependente da orientação da molécula e das posições dos átomos e o segundo, por se tratar de um vetor de um átomo de uma molécula para um átomo de outra molécula, depende somente da orientação das moléculas.

A saída do **“Rotting”** traz uma lista das melhores orientações, em ordem decrescente de coeficiente de correlação (CC) (apêndice II). O coeficiente de correlação é um controle estatístico que relaciona os dados calculados do modelo (fator de estrutura calculado - F_{calc}) e os provenientes do padrão de difração (fator de estrutura observado - F_{obs}), medindo a concordância entre ambos.

A terceira etapa é encontrar os parâmetros translacionais do modelo dentro da cela unitária alvo. A rotina usada nesta etapa é a **“Traning”**. Esta rotina tem como entrada a saída do **“Rotting”**, ela lê uma

rotação específica e procura a melhor translação para o modelo dentro da cela unitária. O programa faz um controle estatístico da solução, através do coeficiente de correlação e do R_{factor} . Para cada dímero $\alpha 1-\beta 1$, foi feita uma procura distinta. Encontrada a melhor translação para um dos dímeros, fixa-se este e procura-se a solução do outro. Para cada um dos modelos foi feita uma substituição molecular, sendo que em ambos os casos o melhor modelo foi o 1OUU(*Oncorhynchus mykiss*).

As melhores soluções da substituição molecular foram submetidas a um refinamento de corpo rígido (Castellano *et al*, 1991), sendo selecionada a que apresentou maior coeficiente de correlação.

Para a melhor solução foram geradas as coordenadas relativas a cada um dos átomos em formato PDB, obtendo-se assim uma estrutura preliminar das frações I e IV da hemoglobina do cascudo.

3.4 Alinhamento da seqüência do *Oncorhynchus mykiss* via Blast II

Como as hemoglobinas do cascudo não foram, ainda, seqüenciadas, foi feito um alinhamento da seqüência inicial da hemoglobina do *Oncorhynchus mykiss*, com objetivo de verificar quais aminoácidos mantinham suas posições conservadas em relação a outras hemoglobinas de peixes, e quais variações residuais eram possíveis em uma determinada posição.



O programa Blast II (Tutosowa e Madden, 1999), faz uma busca de seqüências no banco de dados de proteínas (PDB), que apresentam similaridades com a seqüência de entrada. Havendo dados que possam ser comparados o programa faz uma classificação por afinidade residual na forma de “score”. Nesta classificação é dada a seqüência de aminoácidos da proteína, bem como a sua identificação no banco de dados. Neste caso, o alinhamento da seqüência encontra-se no anexo – IV.

3.5 – Refinamento da Estrutura

Apenas a fração IV da Hb do cascudo foi submetida a um refinamento completo, ou seja, foram executados todos os passos de refinamento.

Para o refinamento da estrutura, foi usado um programa computacional de biologia estrutural, chamado *X-PLOR* (Brünger, 1992). Este programa explora o espaço conformacional das macromoléculas restringindo-as a regiões permitidas, através de combinações entre funções energias empíricas e dados experimentais. O refinamento pode ser dividido em passos (Azevedo, 1997).

O primeiro passo, antes de iniciar o refinamento, é a geração das informações moleculares da estrutura que terão suas coordenadas refinadas. Estas informações são divididas em dois arquivos, um com a topologia da macromolécula e o outro com os parâmetros químicos. As



informações topológicas incluem as massas dos átomos, as ligações (conectividade), os ângulos de ligação, os ângulos diédricos, os tipos de átomos, a carga elétrica, as ligações de hidrogênio, etc... O arquivo de parâmetros químicos armazena as constantes a serem usadas no cálculo da função energia (E_{total}), que será minimizada durante o refinamento. Estes dois arquivos são gerados a partir das coordenadas da estrutura inicial a ser refinada (Azevedo, 1997).

A próxima etapa é o refinamento de corpo rígido, que nada mais é do que um processo de refinamento das posições de grupos rígidos de átomos contra as intensidades observadas. A estrutura inicial pode ser dividida em um ou mais grupos rígidos, neste caso por se tratar de um tetrâmero de hemoglobina na unidade assimétrica, a molécula foi dividida em quatro monômeros. Para cada grupo é possível refinar até seis parâmetros, três rotacionais e três translacionais. As coordenadas de saída do refinamento de corpo rígido são usadas como coordenadas de entrada para o refinamento posicional (Azevedo, 1997).

Para o refinamento posicional o programa *X-PLOR*, usa o método do gradiente conjugado de Powell (1977). A função energia minimizada no algoritmo é dada por:

$$\begin{aligned} E_{total} = & w_A \sum E_{cristal} + w_B \sum E_{bond_distances} + w_C \sum E_{bond_angles} \\ & + w_D \sum E_{torsion_angles} + w_E \sum E_{non_bonded_contacts} + w_F \sum E_{planar_groups} \\ & + w_G \sum E_{chiral_volumes} \end{aligned}$$

os termos w s na equação são os pesos relativos para os diferentes termos de energia (apêndice V). A energia do termo cristalográfico E_{cristal} vem da diferença entre os fatores de estrutura observados (F_{obs}) e os calculados (F_{calc}). Os outros termos estereoquímicos são calculados considerando-se a diferença entre o valor real, obtido da estrutura que está sendo refinada, e os valores ideais, de forma que, quando os átomos se desviam do valor ideal ocorre um aumento no termo da energia. Estes valores considerados ideais foram tabulados para proteínas e ácidos nucleicos, através de exames de estruturas de pequenas moléculas de aminoácidos, peptídeos, nucleotídeos e outros compostos similares. O programa *X-PLOR* fornece arquivos com a geometria de todos aminoácidos, alguns nucleotídeos e moléculas de água. O programa adiciona um termo eletrostático na função energia E_{total} , para os grupos carregados presentes na macromolécula. Durante cada ciclo de refinamento o programa calcula dois índices de discordância, definidos a seguir. O R_{factor} é definido como:

$$R_{\text{factor}} = \frac{\sum_k |F_{\text{obs}}(H) - F_{\text{calc}}(H)|}{\sum_k F_{\text{obs}}(H)}$$

onde a somatória é feita sobre todas as reflexões H , dentro da faixa de resolução definida no refinamento. Este índice mostra a discordância entre os fatores de estrutura observados e calculados, sendo portanto mais baixo para estruturas mais próximas dos dados experimentais. Alguns

procedimentos de refinamento podem levar o R_{factor} a valores artificialmente bem baixos (abaixo de 20%). Entre estes procedimentos podemos citar o de aumentar o número de moléculas de água acima do necessário, que pode de forma artificial diminuir o valor do R_{factor} . Um outro índice de discordância entre os fatores de estruturas calculados e observados, bem mais robusto, é chamado de R_{free} . Este índice é calculado da mesma forma que o R_{factor} , só que usa somente uma parte das reflexões do conjunto de dados. Aproximadamente 10% das reflexões são deixadas, ao acaso, fora do refinamento, para que seja feito o controle via R_{free} . Os valores de R_{factor} e R_{free} são calculados para cada ciclo de refinamento, possibilitando um acompanhamento passo a passo do processo. Geralmente valores de R_{free} abaixo de 35% são indicativos de que a estrutura está de alguma forma correta, porém necessita de alguns refinamentos (Azevedo, 1997).

O próximo passo é fazer o refinamento de dinâmica molecular (*simulated annealing*). Se o modelo inicial não difere muito da estrutura real o refinamento posicional normalmente converge para a solução correta, entretanto, se as distâncias entre os átomos no modelo e na estrutura real são muito grandes o refinamento posicional pode ficar aprisionado em um mínimo local. A dinâmica molecular é usada para tirar o modelo deste mínimo local, permitindo que a estrutura supere barreiras energéticas da função E_{total} (Brünger, 1991). Nesta etapa o comportamento dinâmico de um sistema de partículas é simulado (Azevedo, 1997).



As coordenadas refinadas são usadas para gerar mapas de densidade eletrônica. Neste trabalho o tipo de mapa utilizado foi $2F_{\text{obs}} - F_{\text{calc}}$, com fases α_{calc} , que é usado para localizar pequenos desvios das cadeias laterais e para a observação do aspecto geral da estrutura (Azevedo, 1997).

A visualização é feita através do programa *Xtalview* (McRae, 1992), que superpõe as coordenadas do modelo com o mapa de densidade eletrônica.

4.1 – Cristalização La1b-I

O resultado da coleta de dados de difração de raios X de La1b-I, mostra que esta foi cristalizada no grupo espacial monoclinico primitivo, com extinções sistemáticas compatíveis com o grupo espacial $P2_1$. Seus parâmetros de rede são $a = 56.8 \text{ \AA}$, $b = 64.0 \text{ \AA}$, $c = 25.7 \text{ \AA}$ e $\beta = 92.2^\circ$. A unidade assimétrica contém apenas um tetramero com peso molecular de 65 kDa, com volume molecular de $2.4 \text{ \AA}^3 \text{ Da}^{-1}$ (Matthews, 1968). O conteúdo estequiométrico de solvente no cristal é de 48.8% e a densidade calculada do cristal foi de 1.184 g cm^{-3} .

4.1.1 – Cristalização La1b-IV

O resultado da coleta de dados de difração de raios X de La1b-IV, mostra que esta foi cristalizada no grupo espacial ortorrômbico primitivo, com parâmetros de rede $a = 52.6 \text{ \AA}$, $b = 104.8 \text{ \AA}$ e $c = 113.9 \text{ \AA}$.

Capítulo 4

Resultados e discussões.

4.1 – Cristalização LaHb-I

O resultado da coleta de dados de difração de raios X da LaHb-I, mostrou que esta foi cristalizada no grupo espacial monoclinico primitivo, com extinções sistemáticas compatíveis com o grupo espacial $P2_1$. Seus parâmetros de cela são $a = 56.8 \text{ \AA}$, $b = 64.0 \text{ \AA}$, $c = 86.7 \text{ \AA}$ e $\beta = 92.2^\circ$. A unidade assimétrica contém apenas um tetrâmero com peso molecular de 65 KDa, com volume molecular de $2.4 \text{ \AA}^3 \text{ Da}^{-1}$ (Matthews, 1968). O conteúdo estimado de solvente no cristal é de 48.8% e a densidade calculada do cristal foi de 1.184 g cm^{-3} .

4.1.1 – Cristalização LaHb-IV

O resultado da coleta de dados de difração de raios X da LaHb-IV, mostrou que esta foi cristalizada no grupo espacial ortorrômbico primitivo, com parâmetros de cela $a = 52.6 \text{ \AA}$, $b = 104.8 \text{ \AA}$ $c = 113.9 \text{ \AA}$.



A unidade assimétrica contém apenas um tetrâmero com peso molecular de 65 KDa, com volume molecular de $2.41 \text{ \AA}^3 \text{ Da}^{-1}$ (Matthews, 1968) e volume parcial específico de proteína de $0.74 \text{ cm}^3\text{g}^{-1}$. O conteúdo estimado de solvente no cristal é de 49 % e a densidade calculada do cristal foi de 1.18 g cm^{-3} .

Um resumo detalhado das estatísticas da difração de raios X, da fração I e IV, estão dispostas nas tabelas 4.1 e 4.2.

Tabela 4.1 - Estatística detalhada dos dados de difração de raios X síncrotron para o cristal da fração I.

INTERVALO DE RESOLUÇÃO (Å)	NÚMERO DE REFLEXÕES INDEPENDENTE	I/σ(I)	COMPLETEZA (%)	REDUNDÂNCIA	R _{MERGE} (%)
15.00-6.15	1095	18.4	74.5	2.47	6.6
6.15-4.89	1202	16.8	81.6	2.47	8.4
4.89-4.27	1180	17.2	82.6	2.39	7.7
4.27-3.88	1220	15.7	84.3	2.42	8.6
3.88-3.60	1230	12.8	86.0	2.38	9.7
3.60-3.39	1255	10.4	86.9	2.34	11.7
3.39-3.22	1255	8.1	87.4	2.35	13.8
3.22-3.08	1231	6.5	88.0	2.30	17.4
3.08-2.96	1291	5.4	88.9	2.30	19.0
2.96-2.86	1290	4.4	90.0	2.29	22.0
TODAS AS REFLEXÕES	12249	11.4	85.0	2.37	8.3

$R_{merge} = 100 \times \sum_{hkl} [\sum_i (|I_{hkl,i} - \langle I_{hkl} \rangle|)] / \sum_{hkl,i} (I_{hkl})$, onde I_{hkl} é a intensidade de uma medida individual de reflexões com índices, h , k e l , e $\langle I_{hkl} \rangle$ é a média das intensidades das reflexões.

Completeza = (nº reflexões observadas / nº reflexões possíveis)

Redundância = média das reflexões ponderada pelo nº de observações,

Tabela 4.2 - Estatística detalhada dos dados de difração de raios X síncrotron para o cristal da fração IV.

INTERVALO DE RESOLUÇÃO (Å)		NÚMERO DE REFLEXÕES INDEPENDENTE	I/σ(I)	R _{MERGE} (%)	COMPLETEZA (%)
20.00	5.87	1581	19.92	5.8	86.3
5.87	4.68	1665	20.01	6.5	95.9
4.68	4.09	1670	20.49	7.2	97.4
4.09	3.72	1652	19.34	7.8	97.8
3.72	3.46	1662	17.24	8.5	98.5
3.46	3.25	1657	15.11	9.6	98.7
3.25	3.09	1661	12.45	11.2	99.6
3.09	2.96	1666	10.59	12.5	98.7
2.96	2.84	1639	9.59	13.5	98.0
2.84	2.75	1629	8.42	15.1	98.6
2.75	2.66	1646	6.49	17.8	99.0
2.66	2.58	1620	6.10	18.9	97.1
2.58	2.52	1642	5.33	21.3	99.8
2.52	2.46	1606	5.12	22.7	96.3
2.46	2.40	1622	4.56	24.2	99.6
Todas as reflexões-		24618	15.17	8.8	97.3

$R_{merge} = 100 \times \sum_{hkl} [\sum_i (|I_{hkl,i} - \langle I_{hkl} \rangle|)] / \sum_{hkl,i} (I_{hkl})$, onde I_{hkl} é a intensidade de uma medida individual de reflexões com índices, h , k e l , e $\langle I_{hkl} \rangle$ é a média das intensidades das reflexões.

Completeza = (nº reflexões observadas / nº reflexões possíveis)

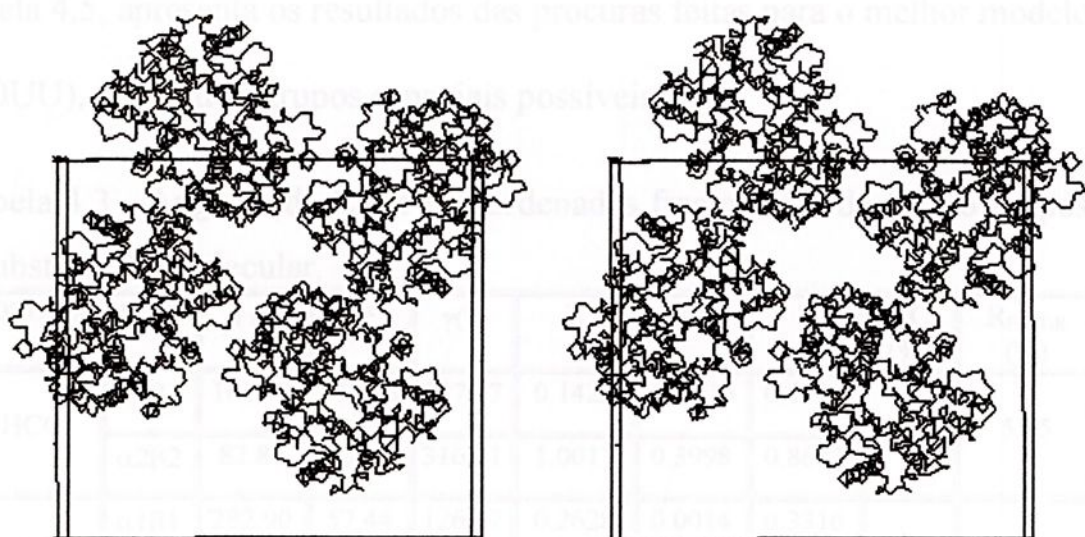
Redundância = média das reflexões ponderada pelo nº de observações.

4.2 - Substituição Molecular

As estruturas preliminares das frações I e IV da hemoglobina do peixe cascudo, foram determinadas por substituição molecular (fig. 12). Para ambas frações, o modelo que apresentou melhor resultado, entre todos os testados (tabelas 4.3 e 4.4), foi a hemoglobina da truta (1OUU) (Tame,

et al, 1996). A melhor faixa de resolução para ambos os casos foi de 10.0-4.5 Å.

a)



b)

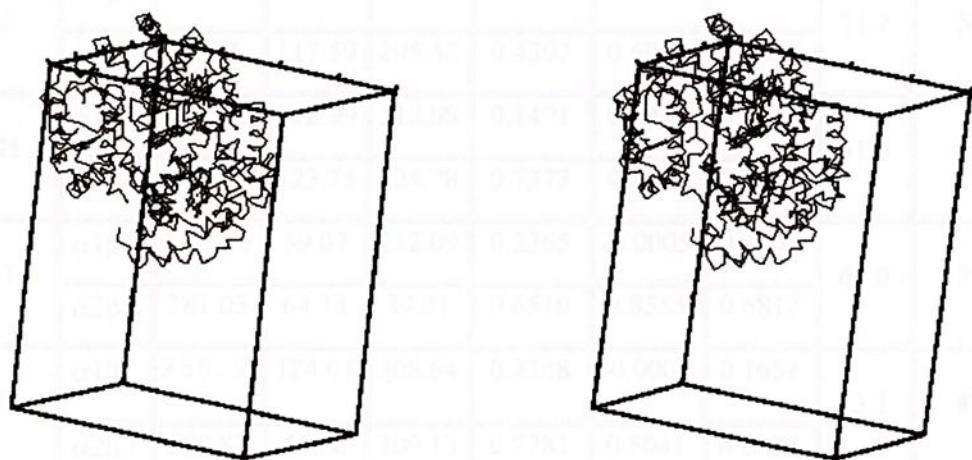


Fig. 12 – Visão estéreo das estruturas preliminares das frações IV (a) e I (b).

Para a fração IV, além da procura pelo melhor modelo, foi também necessário fazer a procura pelo grupo espacial correto. O melhores resultados foram obtidos para o grupo espacial $P2_12_12_1$ (tabela 4.4). A tabela 4.5, apresenta os resultados das procuras feitas para o melhor modelo (1OUU), nos outros grupos espaciais possíveis.

Tabela 4.3 - Ângulos de Euler e coordenadas fracionárias da fração I, após a substituição molecular.

PROTEÍNA	DIM.	α (°)	β (°)	γ (°)	X	Y	Z	CC (%)	R _{FATOR} (%)
1HCO	$\alpha 1\beta 1$	101.30	118.63	137.17	0.1422	-0.0013	0.1735	31.0	52.5
	$\alpha 2\beta 2$	82.87	65.52	316.71	1.0017	0.3998	0.8653		
1T1N	$\alpha 1\beta 1$	282.90	57.44	126.57	0.2628	0.0014	0.3316	64.3	38.2
	$\alpha 2\beta 2$	101.93	114.51	139.93	0.1514	0.3580	0.1826		
1OUU	$\alpha 1\beta 1$	258.03	121.81	127.01	0.2368	0.0008	0.1707	67.1	36.5
	$\alpha 2\beta 2$	101.45	115.82	319.59	0.6536	0.8546	0.6813		
1CG8	$\alpha 1\beta 1$	286.82	56.12	40.65	0.4888	0.0007	0.3084	31.7	52.4
	$\alpha 2\beta 2$	31.01	117.59	295.52	0.4397	0.6826	0.6530		
1HBH	$\alpha 1\beta 1$	100.58	112.99	318.98	0.1491	0.0044	0.1843	51.6	44.3
	$\alpha 2\beta 2$	256.28	123.75	124.78	0.7373	0.1434	0.6675		
1OUT	$\alpha 1\beta 1$	77.75	59.07	232.09	0.2365	-0.0005	0.1704	61.0	39.0
	$\alpha 2\beta 2$	281.03	64.14	39.81	0.6510	0.8555	0.6812		
1PBX	$\alpha 1\beta 1$	258.97	124.01	308.64	0.2348	-0.0008	0.1658	43.1	48.2
	$\alpha 2\beta 2$	278.83	56.06	309.13	0.7781	0.5041	0.8228		
1SPG	$\alpha 1\beta 1$	280.29	56.93	308.08	0.2666	0.0007	0.3286	62.9	37.8
	$\alpha 2\beta 2$	99.39	115.03	319.18	0.1480	0.3566	0.1810		
1CG5	$\alpha 1\beta 1$	44.40	86.95	10.42	0.1939	0.0016	0.1122	28.8	52.2
	$\alpha 2\beta 2$	353.41	130.16	130.94	0.7181	0.5014	0.6493		

Tabela 4.4 - Ângulos de Euler e coordenadas fracionárias da fração IV, grupo espacial $P2_12_12_1$, após a substituição molecular.

PROTEÍNA	DIM.	α (°)	β (°)	γ (°)	X	Y	Z	CC (%)	R _{FATOR} (%)
IHHO	$\alpha 1\beta 1$	85.05	92.50	309.09	0.4756	0.0743	0.0573	29.3	54.0
	$\alpha 2\beta 2$	24.30	61.13	265.74	0.1988	0.1696	0.0169		
ITIN	$\alpha 1\beta 1$	76.61	62.38	32.05	0.3706	0.1994	0.2186	24.9	54.5
	$\alpha 2\beta 2$	102.54	117.45	207.46	0.3271	0.7166	0.4570		
IOUU	$\alpha 1\beta 1$	97.53	59.75	63.02	0.3072	0.0129	0.4158	62.7	39.0
	$\alpha 2\beta 2$	79.25	58.31	24.71	0.8661	0.6949	0.7090		
ICG8	$\alpha 1\beta 1$	87.15	83.84	47.07	0.0354	0.4196	0.3626	24.3	53.9
	$\alpha 2\beta 2$	41.22	114.49	270.54	0.2430	0.1286	0.6039		
IHBH	$\alpha 1\beta 1$	75.02	58.29	26.37	0.3734	0.1950	0.2133	31.5	51.6
	$\alpha 2\beta 2$	77.70	58.34	26.45	0.3714	0.1964	0.2121		
IOUT	$\alpha 1\beta 1$	79.39	121.16	334.18	0.1326	0.3039	0.2081	42.7	47.0
	$\alpha 2\beta 2$	84.70	64.53	295.50	0.3202	0.4882	0.5812		
IPBX	$\alpha 1\beta 1$	104.74	121.58	27.26	0.1298	0.1965	0.2899	37.7	49.9
	$\alpha 2\beta 2$	78.15	61.77	29.20	0.8608	0.6946	0.2131		
ISPG	$\alpha 1\beta 1$	80.35	119.65	67.20	0.0613	0.0384	0.1216	24.3	52.7
	$\alpha 2\beta 2$	24.71	115.30	339.60	0.8311	0.1327	0.8739		
ICG5	$\alpha 1\beta 1$	73.24	62.01	30.85	0.4680	0.2456	0.2637	23.2	53.4
	$\alpha 2\beta 2$	110.85	117.22	211.13	0.4393	0.7906	0.1647		

Tabela 4.5 – Coordenadas fracionárias dos diferentes grupos espaciais possíveis para a LaHb-IV.

Grupo Espacial	Dímero	x	y	z	CC (%)	R _{factor} (%)
P222	α1β1	0.3739	0.1420	0.4332	33.0	52.6
	α2β2	0.3911	0.1471	0.4328	33.0	
P222 ₁	α1β1	0.0673	0.0119	0.3542	37.9	49.7
	α2β2	0.3845	0.6936	0.8484	37.9	
P2 ₁ 2 ₁ 2	α1β1	0.0823	0.3436	0.0926	32.1	51.6
	α2β2	0.1136	0.8662	0.9638	32.1	

4.3 - Refinamento da estrutura da LaHb-IV

O primeiro passo foi submeter a estrutura a um refinamento de corpo rígido. Quarenta ciclos deste refinamento foram realizados. O R_{factor} nesta etapa, caiu de 42.3% para 38.8% e o R_{free} de de 42.8% para 39.1%, para dados entre 10.0 e 2.4 Å. O próximo passo foi fazer um refinamento posicional da estrutura. Após a realização de 200 ciclos, o R_{factor} e o R_{free} caíram de 38.8% para 25.1% e de 39.1% para 35.8% , respectivamente. Foi usada simetria não-cristalográfica no refinamento de corpo rígido e nos refinamentos posicionais deste trabalho, com o intuito de obter-se um mapa médio, no qual os ruídos tendem a se anular, melhorando a qualidade geral do mapa de densidade eletrônica.

As coordenadas geradas no refinamento posicional foram usadas para a realização da dinâmica molecular (refinamento de simulated annealing). A temperatura inicial foi de 3000 K e os passos foram de 0.5fs. Os valores obtidos de R_{free} e R_{factor} foram 34.8% e 20.0%, respectivamente. Terminado estes primeiros ciclos de refinamentos, foi gerado um mapa de densidade eletrônica do tipo $2F_o - F_c$.

Através da análise do mapa de densidade eletrônica, foram realizados alguns ajustes posicionais nas cadeias laterais. Em algumas posições, a densidade eletrônica era claramente referente a um outro resíduo de aminoácido. Para que não houvessem dúvidas foram gerados mapas de densidade eletrônica omitindo cada um desses resíduos. Neste tipo de mapa (*omit map*) a porção do modelo que será analisada é deixada fora do cálculo da fase, portanto toda a influência da porção analisada é retirada. Nos casos em que se confirmaram a necessidade de troca, foi verificado quais resíduos poderiam ocupar a mesma posição, segundo o alinhamento e o contorno local do mapa. Havendo a possibilidade, trocavam-se os resíduos que não estavam em conformidade com a densidade eletrônica local, por outros que apresentaram melhores ajustes (figs. 13 e 14).

No total foram substituídos 13 resíduos, conforme tabelas 4.6 e 4.7. Em outros 54 resíduos houve dúvidas quanto ao posicionamento, porém ao gerar-se o mapa omitindo o modelo, não ficou claro qual resíduo deveria ocupar a posição, não sendo, portanto, realizada a troca.

Tabela 4.6 – Substituições residuais realizadas nas cadeias alfas.

Resíduo Nº	Resíduo Original	Resíduo substituído	Possíveis resíduos na posição	B_fator
3	THR (T)	SER (S)	T,S	30
90	PHE (F)	THR (T)	F,T,S,Y	35
105	ASN (N)	CYS (C)	N,C	20
116	SER (S)	ALA (A)	S,A,V,N,G,K	25
124	ILE (I)	VAL (V)	I,V,L,M,A	24
125	ALA (A)	SER (S)	A,S	25
132	ALA (A)	SER (S)	A,N,C,Q,G,S	26
133	VAL (V)	LEU (L)	V,L,F	30

Tabela 4.7 – Substituições residuais realizadas nas cadeias betas

Resíduo Nº	Resíduo Original	Resíduo substituído	Possíveis resíduos na posição	B_fator
48	VAL (V)	LEU (L)	V,L	44
105	VAL (V)	LEU (L)	V,L	20
125	GLU (E)	ASP (D)	E,D,N	44
128	ALA (A)	GLU (E)	A,E	23
129	THR (T)	ALA (A)	T,V,A	17

Após a troca dos resíduos, foram feitos novos ciclos de refinamentos posicionais. Nesta etapa o valor do R_{factor} não apresentou mudança significativa e o R_{free} caiu de 34.8 para 32.8%, indicando que o modelo foi aprimorado.

Na etapa final do refinamento é feita a inserção das moléculas de água, usando o programa Xtalview (McRee, 1992). Para isso foi gerado um mapa do tipo $F_o - F_c$, que mostra as densidades não explicadas do mapa, que podem ser provenientes de moléculas de água ou de ruídos.

Com o objetivo de verificar a veracidade das moléculas de água inseridas, foi feito um refinamento do fator térmico (b_factor). Onde todas as moléculas que apresentaram b_factor acima de 50.0 foram excluídas, pois provavelmente tratava-se de picos de ruídos. O número total de moléculas de água inseridas foi de 245.

Ao final desta etapa, os valores de R_{free} e R_{factor} foram 31.0% e 20.0%, respectivamente. Em comparação com outros refinamentos de estruturas de hemoglobinas, os valores aqui obtidos podem ser caracterizados como sendo bons, principalmente, considerando-se que o sequenciamento primário ainda não foi realizado.



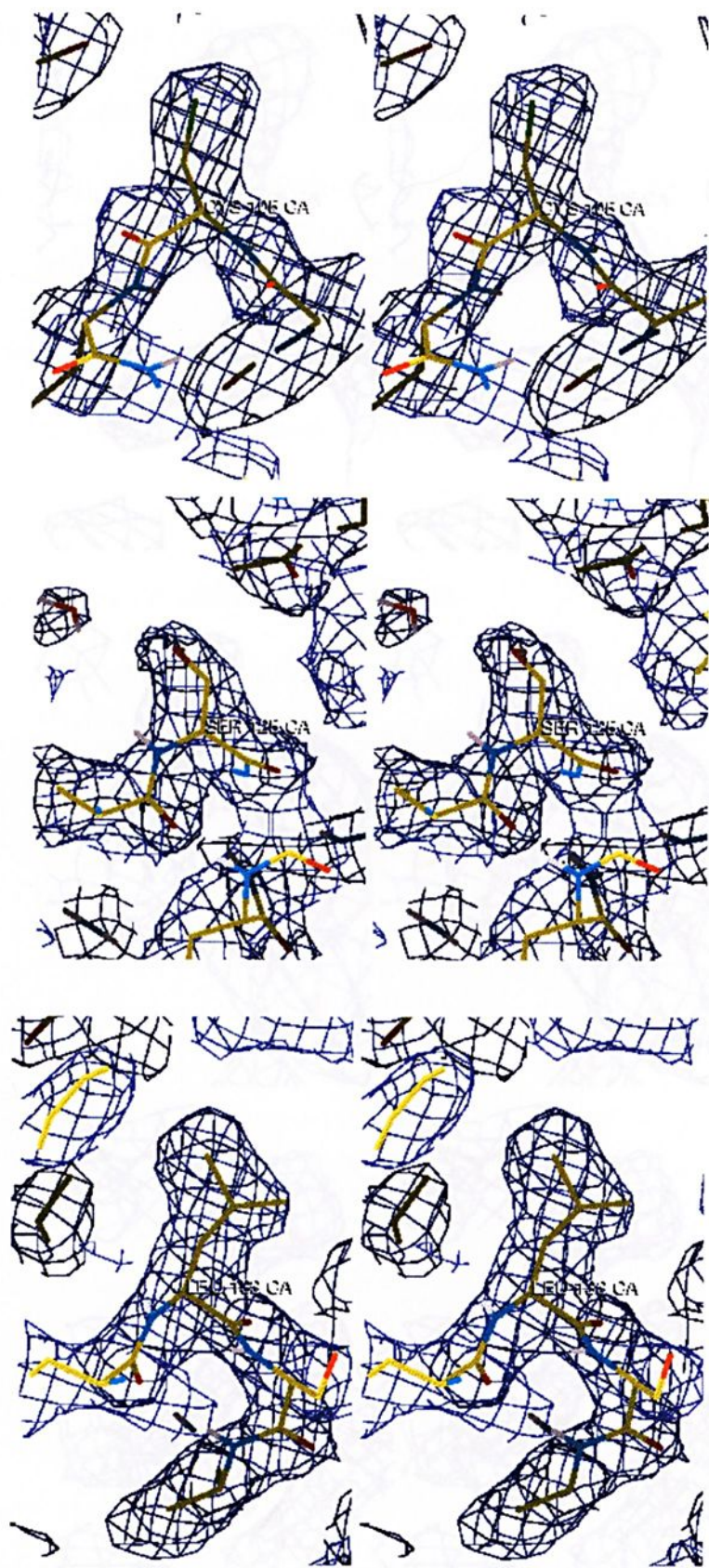


Fig. 13 - Visão estéreo dos resíduos substituídos na cadeia alfa: CYS(105), SER(125) e LEU(133). Mapa (2Fo - Fc) a 2.4 Å e corte de 1σ.

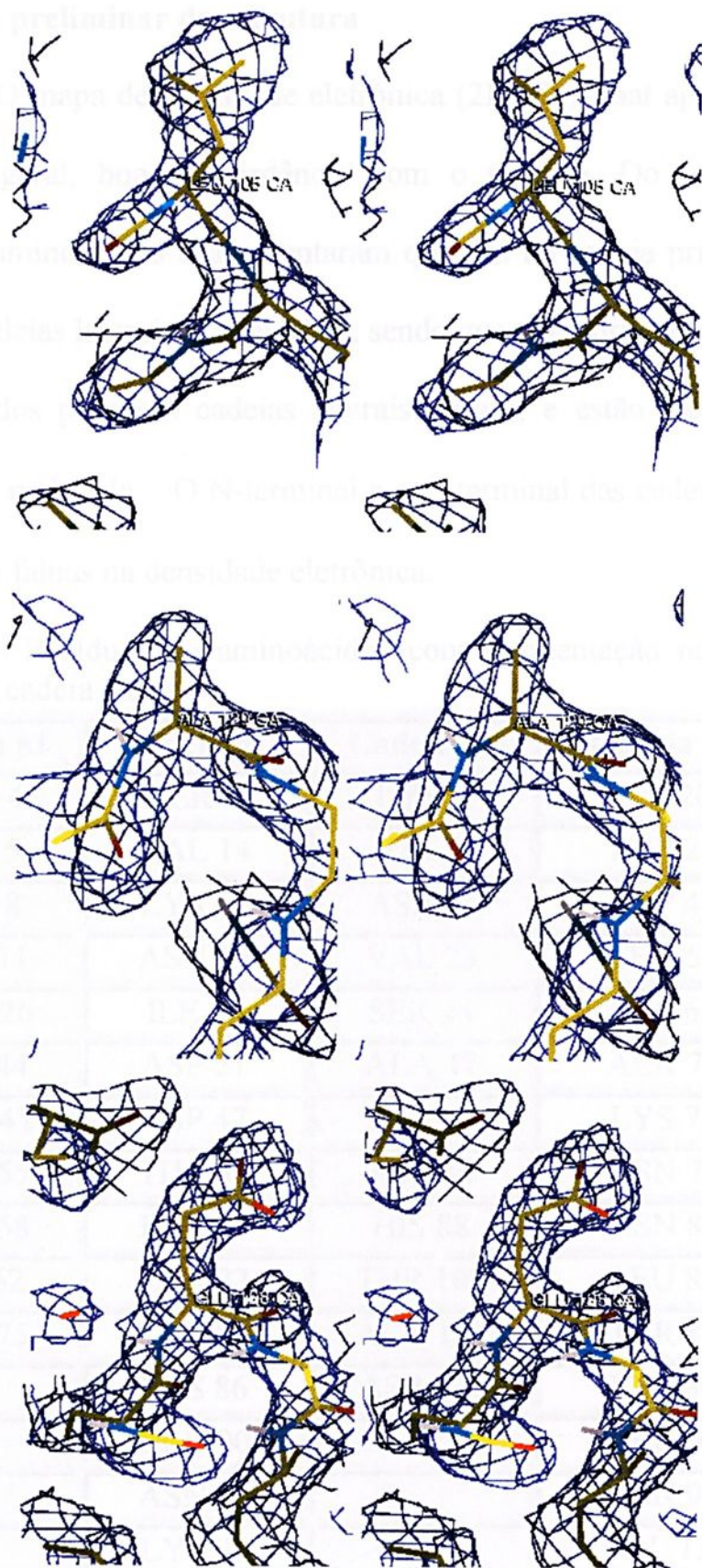


Fig. 14 - Visão estéreo dos resíduos substituídos na cadeia beta: LEU(105), GLU(128) e ALA(129). Mapa (2Fo - Fc) a 2.4 Å e corte de 1σ.

4.4 - Análise preliminar da estrutura

O mapa de densidade eletrônica (2Fo-Fc) final apresentou, de uma forma geral, boa concordância com o modelo. Do total de 576 resíduos de aminoácidos 6 apresentaram quebras na cadeia principal e 54 falhas nas cadeias laterais (Tabela 4.8), sendo que a maioria desses resíduos de aminoácidos possuem cadeias laterais longas, e estão localizados na superfície da molécula. O N-terminal e o C-terminal das cadeias β 1 e β 2, apresentaram falhas na densidade eletrônica.

Tabela 4.8 – Resíduos de aminoácidos com segmentação na densidade eletrônica da cadeia lateral.

Cadeia a1	Cadeia b1	Cadeia a2	Cadeia b2
ALA 4	SER 9	LYS 5	ILE 20
LYS 5	VAL 14	VAL 9	ASP 21
SER 8	LYS 17	ASP 22	ASP 47
LYS 11	ASN 19	VAL 23	THR 50
ALA 26	ILE 20	SER 44	LYS 65
SER 44	ASP 21	ALA 47	ALA 70
ALA 47	ASP 47	LEU 73	LYS 73
PRO 55	THR 50	ASP 86	ASN 77
LYS 58	LYS 65	HIS 88	ASN 80
ILE 62	LEU 82	THR 109	LEU 82
ASP 75	THR 84	ALA 116	THR 84
	LYS 86	ASP 139	LYS 86
	GLU 90		GLU 90
	ASN 94		ASN 94
	LYS 95		VAL 136
	PHE 97		

Na estrutura da LaHb-IV os grupos hemes apresentaram boa densidade eletrônica (fig. 15 e 16), com os quatro íons de ferro posicionados no plano do heme.

Foi feita uma análise da qualidade estereoquímica da estrutura final, usando o programa PROCHECK (Laskowski et al, 1993). Esse programa afere as distâncias ideais entre as ligações, os ângulos de ligação, os ângulos diédricos e os contatos de Van der Waals, usando para isso a comparação com proteínas modelos. Uma das saídas deste programa é o gráfico de Ramachandran, que afere as medidas dos ângulos diédricos ϕ e ψ , sendo, portanto, um bom indicador da qualidade do modelo.

O modelo final apresentou boa configuração estereoquímica com 84.5 % dos resíduos de aminoácidos presentes em regiões favoráveis e somente 1,2 % em regiões generosamente permitidas, não apresentando nenhum resíduo em regiões proibidas (fig. 17).

A caracterização preliminar da LaHb-IV, mostrou que esta não apresenta efeito Root. De posse deste resultado realizou-se uma comparação entre as estruturas secundárias desta Hb com a Hb do peixe antártico *Pagothenia bernacchii*, que apresenta efeito Root. As duas Hbs comparadas possuem estruturas secundárias próximas (fig. 18). Foram observadas pequenas variações nas estruturas, principalmente com relação ao número de resíduos em algumas hélices.



Com o objetivo de destacar a grande similaridade entre as Hbs de peixes e a Hb humana, apesar da baixa identidade seqüencial, foi feita a comparação entre suas estruturas secundárias, que também apresentaram pouca variação (fig. 18)

Foi feita a superposição da LaHb-IV com o modelo inicialmente usado para a substituição molecular (1OUU), e os cálculos dos desvios RMS foram feitos usando o programa ProFit V1.8 (Martin, 1992). Os tetrâmeros das Hbs foram superpostos (fig. 19) e os desvios RMS foram calculados para todos os átomos da cadeia principal, exceto os 5 resíduos N-terminais e C-terminais de cada monômero. O desvio RMS médio para o tetrâmero foi de 0,86 Å, com maiores picos localizados em regiões de *random coil*, mostra que a estrutura final obtida é a da fração IV da hemoglobina do cascudo.

4.5 Conclusão

O presente trabalho descreve a cristalização e a análise preliminar dos dados de difração de raios X de duas Hbs extraídas do *Liposarcus anisitsi* (LaHb-I e LaHb-IV). Foi realizado também o refinamento cristalográfico parcial das coordenadas atômicas da LaHb-IV. O resultado deste refinamento mostra que, mesmo sem a seqüência primária, foi possível obter baixos valores de R_{free} e R_{factor} , com mapas de densidade eletrônica razoáveis, o que indica que uma vez a estrutura

primária esteja disponível o refinamento deverá convergir para valores mais baixos de R_{free} e R_{factor} . O modelo final refinado será utilizado para estudos comparativos das Hbs de peixes, o que possibilitará um maior entendimento das bases moleculares do amplo espectro de atividades funcionais apresentadas por estas hemoglobinas.

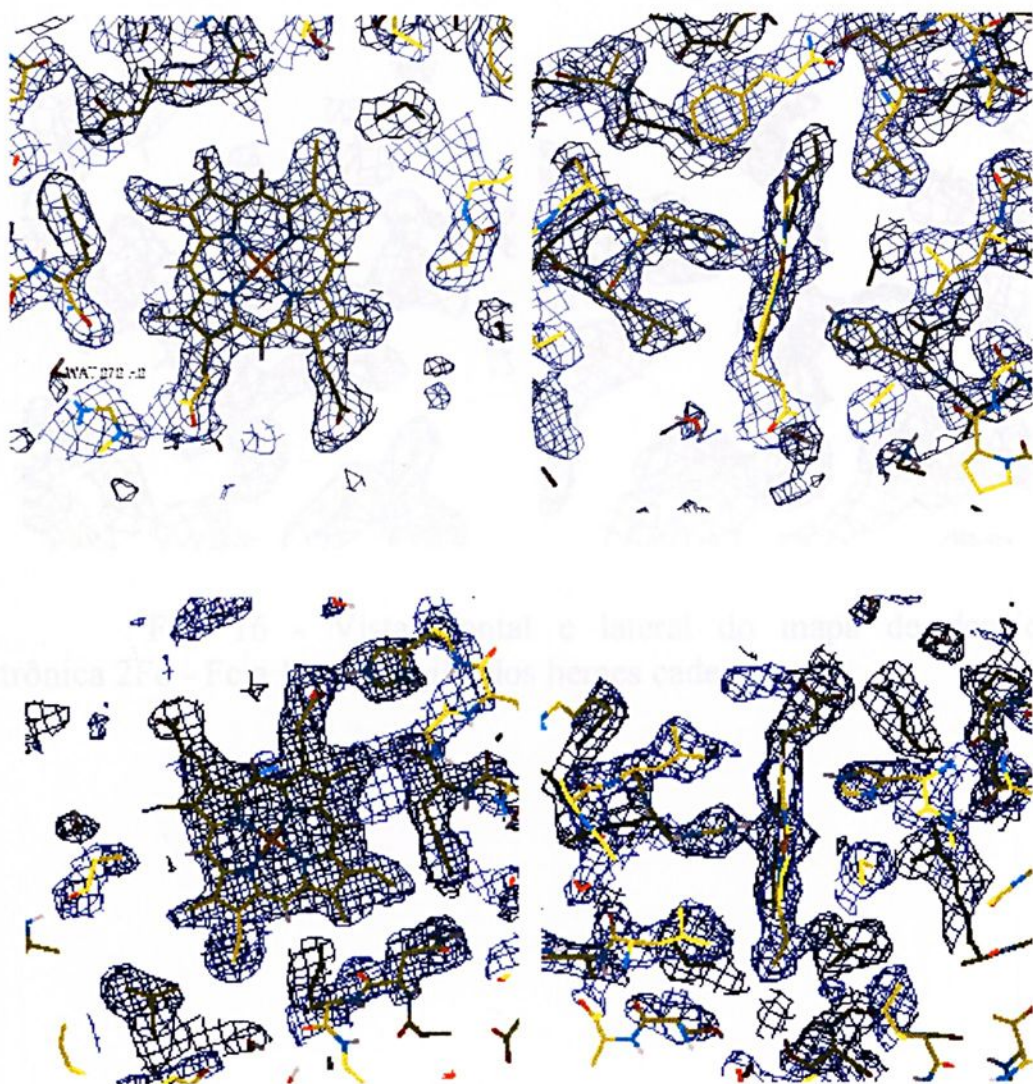


Fig. 15 - Vista frontal e lateral do mapa de densidade eletrônica $2F_o - F_c$ a 1σ , da região dos hemes cadeia alfa.

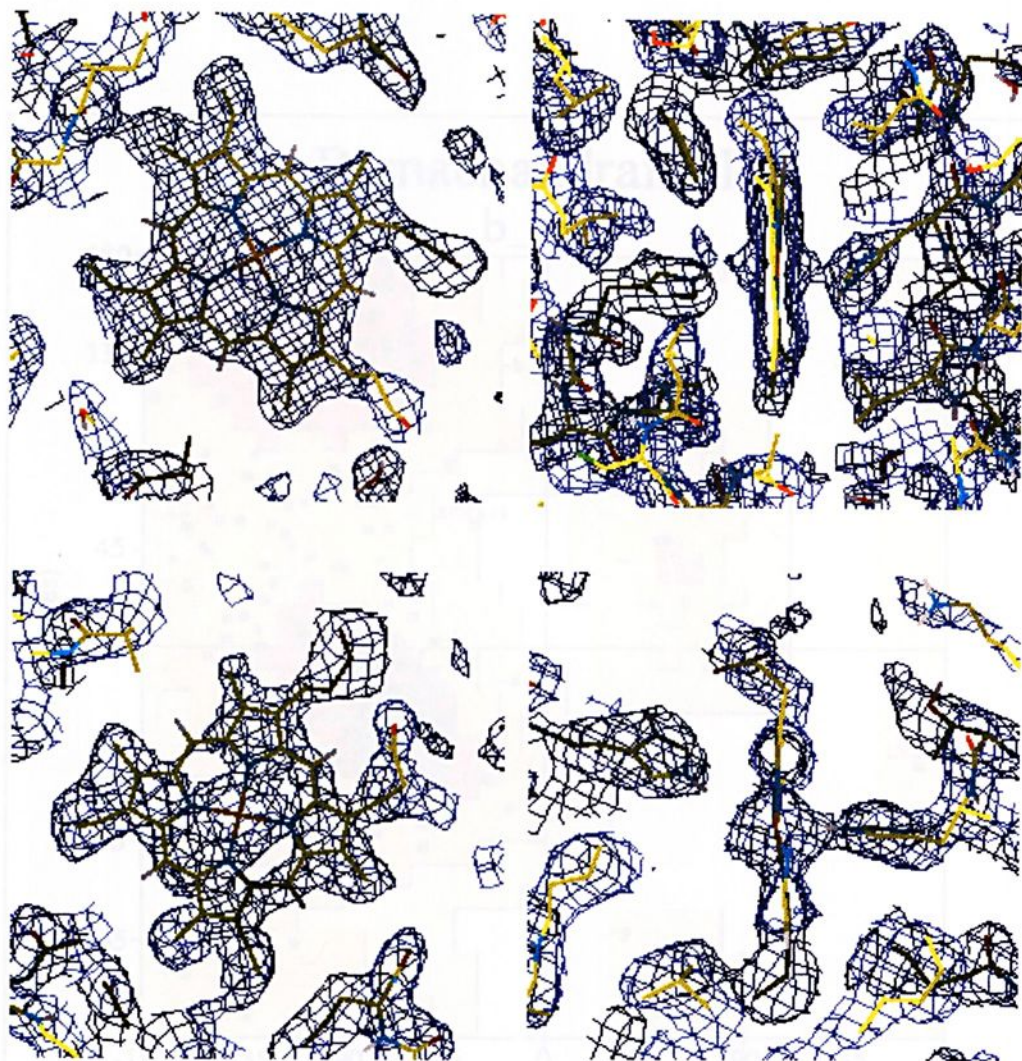


Fig. 16 - Vista frontal e lateral do mapa de densidade eletrônica $2F_o - F_c$ a 1σ , da região dos hemes cadeia beta.

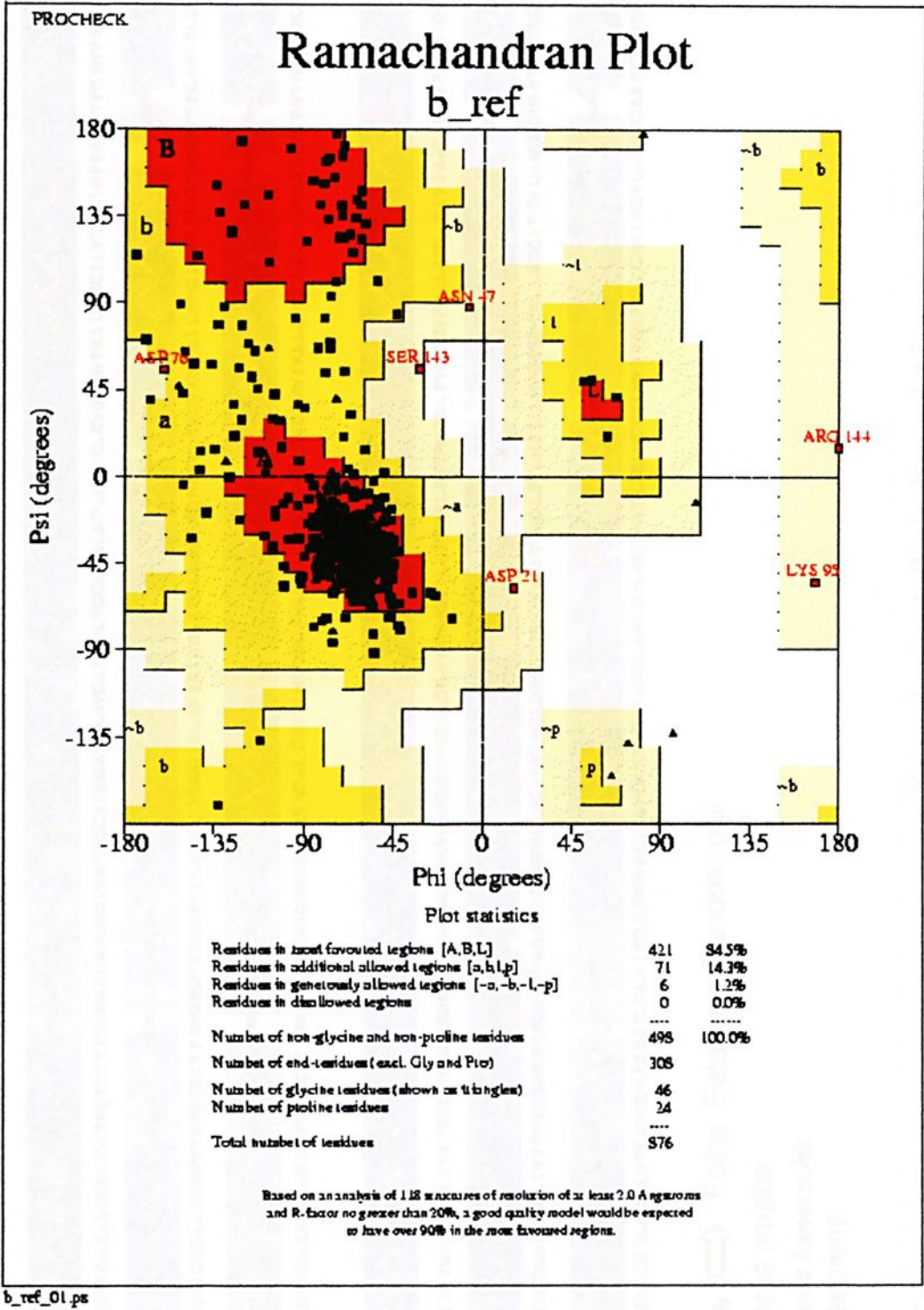


Fig. 17 - Gráfico de Ramachandran da estrutura final da LaHb_{IV} calcula através do programa PROCHECK (Laskowski *et al.*,1993)



Legenda: Hélice Folha Beta Random coil

- 1 *Liposarcus anisitsi*
- 2 *Pagothenia bernacchii*
- 3 *Homo sapiens*

Fig. 18 - Comparação entre as estruturas secundárias da fraçãoIV da hemoglobina do peixe cascudo (LaHb-IV), com a hemoglobina do peixe antártico (1PBX) e com a hemoglobina humana (1HHO) gerado pelo programa PROCHECK (Laskowski *et al.*, 1993).

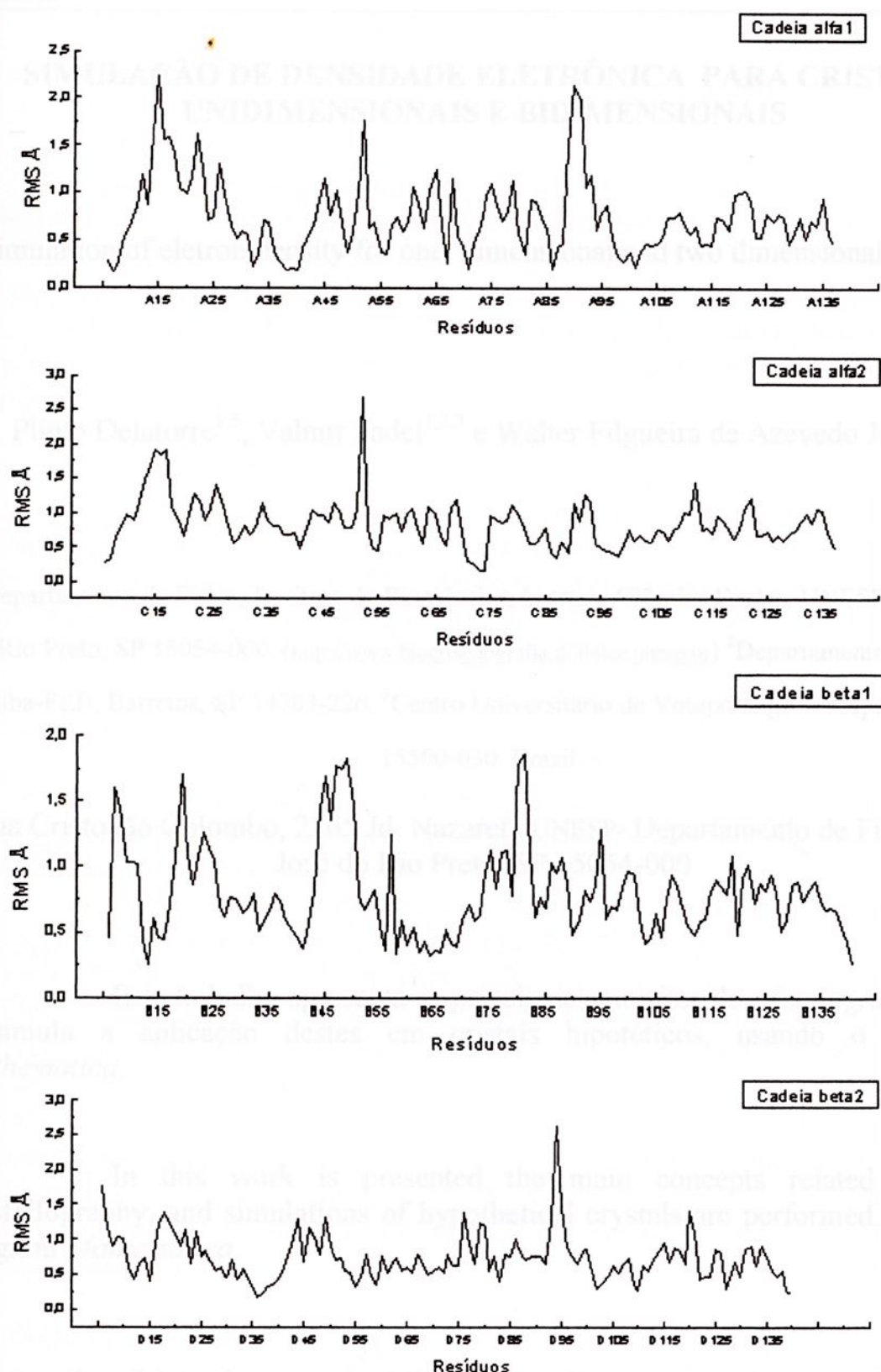


Fig. 19 - Gráficos das superposições da cadeia principal dos monômeros $\alpha 1$, $\alpha 2$, $\beta 1$ e $\beta 2$ da LaHb-IV sobre a Hb da Trout.

SIMULAÇÃO DE DENSIDADE ELETRÔNICA PARA CRISTAIS UNIDIMENSIONAIS E BIDIMENSIONAIS

(Simulation of eletron density for one dimensional and two dimensional crystals)

Plinio Delatorre^{1,3}, Valmir Fadel^{1,2,3} e Walter Filgueira de Azevedo Junior¹

¹Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas, UNESP, São José do Rio Preto, SP 15054-000. (<http://www.biocristalografia.df.ibilce.unesp.br>) ²Departamento de Física. Faciba-FEB. Barretos, SP 14783-226. ³Centro Universitário de Votuporanga. Votuporanga, SP. 15500-030. Brazil.

Rua Cristovão Colombo, 2265 Jd. Nazareth UNESP- Departamento de Física, São José do Rio Preto, SP 15054-000

Este trabalho apresenta os principais conceitos de cristalografia básica, e simula a aplicação destes em cristais hipotéticos, usando o programa *Mathematica*.

In this work is presented the main concepts related to basic crystallography, and simulations of hypothetical crystals are performed, using the program *Mathematica*.

(Este artigo foi aceito para publicação na Revista Brasileira de Ensino de Física.)



I. Introdução

Um crescente número de grupos de pesquisas no Brasil vêm se dedicando ao estudo estrutural de macromoléculas biológicas, empregando técnicas de difração de raios X. Neste cenário, um grande número de estudantes de pós-graduação em diversas instituições de ensino, com diferentes tipos de formação, têm-se iniciado ao estudo dos métodos da cristalografia, muitos destes com uma formação incompleta em física e matemática, o que dificulta a aprendizagem dos diversos conceitos cristalográficos.

Entre os principais conceitos de cristalografia, normalmente ensinados nos cursos de pós-graduação, os relacionados com o problema da fase são os que requerem maior abstração. A visualização da solução do problema da fase, aplicada à solução de estruturas simples, em muito ajuda no amadurecimento deste conceito.

O presente trabalho descreve simulações simples de cristais unidimensionais e bidimensionais, usados no curso de cristalografia I e II no DF-IBILCE-UNESP.

II. Conhecimento Teórico Básico

A análise do padrão de difração de raios X, obtido a partir de um cristal de uma substância de interesse, é de fundamental importância na obtenção de dados geométricos precisos da molécula.

A natureza eletromagnética dos raios X permite a analogia da formação da imagem deste com a formação de imagem por raios de luz. Em um microscópio óptico o objeto é iluminado por feixes de luz, que ao serem espalhados, são coletados por uma lente objetiva, que recombina todos os raios de luz para formar a imagem do objeto. Da mesma maneira, um cristal é iluminado por feixes de raios X, que após serem espalhados são coletados, em uma placa de imagem por exemplo. Mas, neste caso, para que a imagem seja reconstruída é necessário sabermos do padrão de difração, a fase e a intensidade de cada raio difratado. Porém em um experimento de difração toda a informação relativa a fase é perdida, só obtemos informações sobre as intensidades e a posição relativa dos feixes difratados [1].

II.1 Os índices de Miller

Cristais apresentam alto grau de ordenação interna, ou seja, são formados por repetições translacionais de moléculas ou átomos em todas as direções (Fig.1). Esta periodicidade interna é descrita por uma pequena unidade de



volume uniforme da qual o cristal pode ser considerado um derivado, chamada de cela unitária. Cada cela unitária é caracterizada por três vetores a , b e c que definem suas arestas e pelos ângulos γ , α e β existentes entre elas (fig.2).

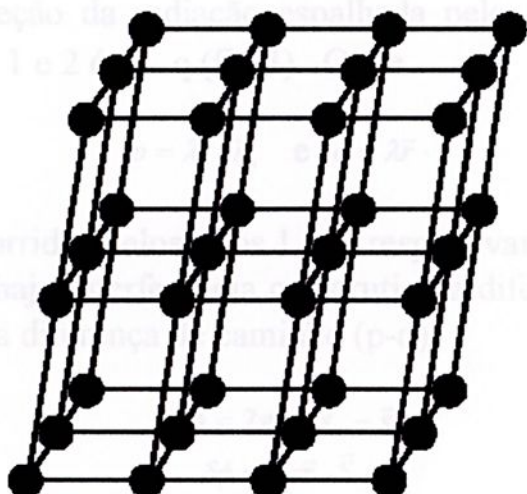


Figura 1 – Retículo Cristalino tridimensional

A lei dos índices racionais estabelece que cada uma das faces de um cristal pode ser descrita com relação a três eixos não coplanares, através de três inteiros pequenos, que são designados por h , k , l . A importância destes índices está no fato de que estes não só descrevem as faces de um cristal, mas também, descrevem grupos de planos paralelos do retículo cristalino. Portanto os feixes de difração de um cristal, podem ser descritos por eles, como feixes refletidos por estes planos. Estes inteiros são chamados de índices de Miller [2].

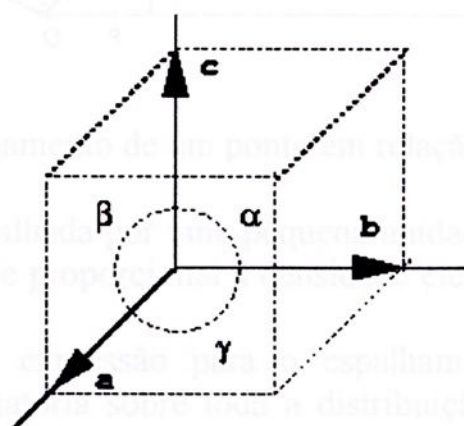


Figura 2 – Cela Unitária

II.2 Fator de Espalhamento Atômico

Ao incidir-se raios X em um átomo, cada elétron comportar-se como um centro espalhador de radiação. Definindo-se a direção da radiação incidente pelo vetor $\vec{s}_0$ e a direção da radiação espalhada pelo vetor $\vec{s}$, a diferença de caminho entre os raios 1 e 2 é $p - q$ (fig.3). Onde

$$p = \lambda \vec{r} \cdot \vec{s}_0 \quad \text{e} \quad q = \lambda \vec{r} \cdot \vec{s}$$

são os caminhos percorridos pelos raios 1 e 2, respectivamente.

Para que haja interferência construtiva a diferença de fase $\delta\phi$ tem que ser igual a $2\pi/\lambda$ vezes a diferença de caminho ($p-q$).

$$\begin{aligned} \delta\phi &= 2\pi \vec{r} \cdot (\vec{s}_0 - \vec{s}) \\ \delta\phi &= 2\pi \vec{r} \cdot \vec{S} \end{aligned} \quad (1)$$

onde $\vec{S} = (\vec{s}_0 - \vec{s})$.

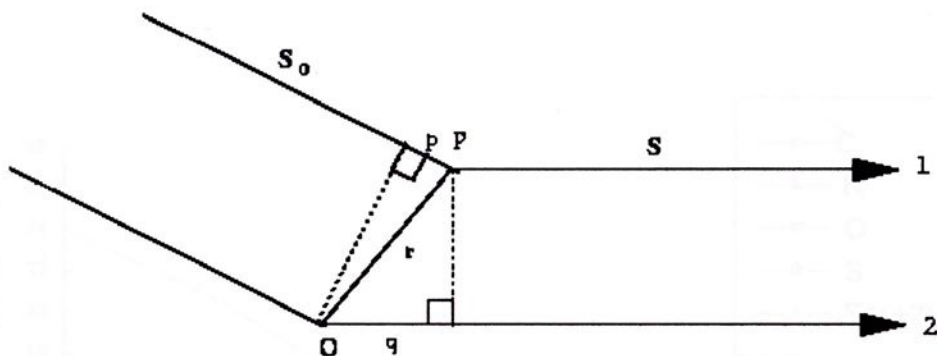


Figura 3 – Espalhamento de um ponto em relação a origem.

A onda total espalhada por uma pequena unidade de volume em uma posição $\vec{r}$, terá uma amplitude proporcional a densidade eletrônica mais um termo de fase.

Para obter-se a expressão para o espalhamento de um átomo é necessário fazer-se uma somatória sobre toda a distribuição espacial de elétrons (nuvem eletrônica).

A densidade eletrônica a uma distância $\vec{r}$ do núcleo será representada por $\rho(\vec{r})$.

O espalhamento atômico depende diretamente do número de elétrons e de suas posições relativas na nuvem eletrônica.

Dizemos que uma nuvem é centro-simétrica em torno da origem quando $\rho(\vec{r}) = \rho(-\vec{r})$. Assim, considerando um átomo com a densidade eletrônica centro-simétrica, podemos escrever a expressão para o fator de espalhamento atômico f como sendo,

$$f = \int \rho(\vec{r}) \exp(2\pi i \vec{r} \cdot \vec{S}) d\vec{r} \tag{2}$$

onde a integração é feita sobre todo o espaço $\vec{r}$.

Sendo a densidade eletrônica do átomo centro-simétrica, não faz qualquer diferença se a orientação de $\vec{S}$ em relação ao átomo muda. Desta maneira, o fator de espalhamento atômico também será centro-simétrico, isto é, ele será independente da direção de $\vec{S}$ e dependente do comprimento de $\vec{S}$:

$$|\vec{S}| = \frac{2 \sin \theta}{\lambda} \tag{3}$$

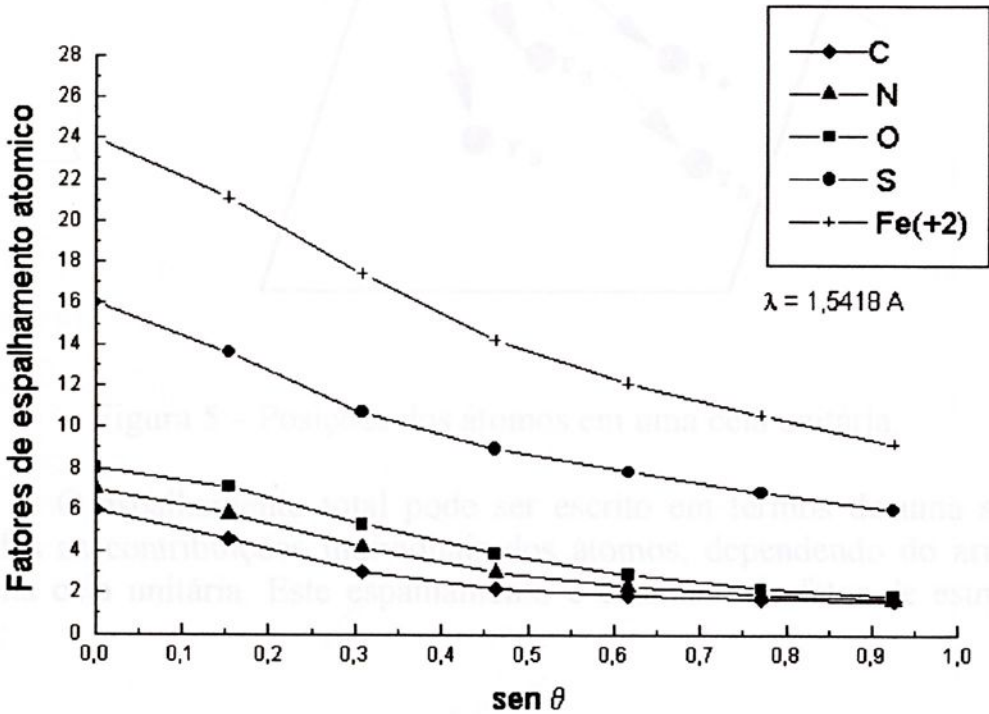


Figura 4 – Gráfico de fatores de espalhamento atômico.

Com o aumento de $\sin \theta$ há um decréscimo do fator de espalhamento atômico (fig.4), devido ao aumento da diferença de fase dos raios X espalhados por regiões distintas na nuvem eletrônica.

II.3 Fator de Estrutura

Consideremos uma cela unitária composta por n átomos, com seu próprio núcleo como origem, os átomos difratarão de acordo com seu fator de espalhamento atômico. Se a origem for transferida para a origem da cela unitária conforme figura 5, o ângulo de fase será dado por $2\pi \vec{r}_j \cdot \vec{S}$ e a onda total espalhada pelos átomos será dada pela soma de todas as contribuições individuais de cada um.

Em função da nova origem o espalhamento atômico será dado por

$$\vec{f}_j = f_j \exp(2\pi i \vec{r}_j \cdot \vec{S}) \quad (4)$$

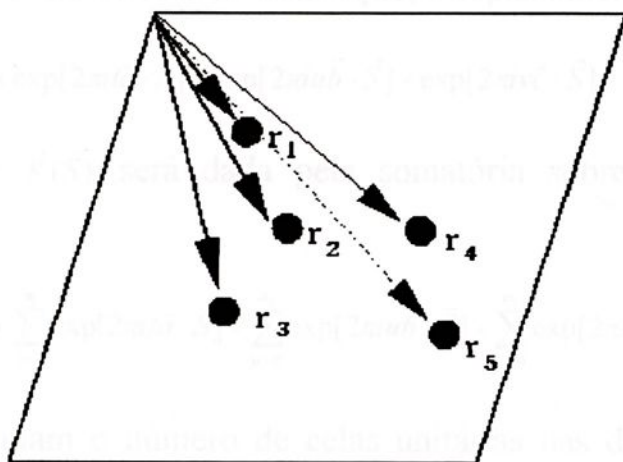


Figura 5 – Posições dos átomos em uma cela unitária.

O espalhamento total pode ser escrito em termos de uma somatória sobre todas as contribuições individuais dos átomos, dependendo do arranjo dos mesmos na cela unitária. Este espalhamento é chamado de fator de estrutura e é dado por:

$$F(S) = \sum_{j=1}^n f_j \exp(2\pi i \vec{r}_j \cdot \vec{S}) \quad (5)$$

onde $\vec{r}_j$ é usado para descrever posições no espaço vetorial direto, e $\vec{S}$ é usado para descrever posições no espaço vetorial recíproco e f_j é o fator de espalhamento atômico referente a cada átomo. Escrevendo-se $F(S)$ em termos de uma integral no volume, temos:

$$F(S) = \int_V \rho(\vec{r}) \exp(2\pi i \vec{r} \cdot \vec{S}) dV \quad (6)$$

onde V é o volume da cela unitária [2][3][4].

II.4 Espalhamento por um Cristal

Para obter o espalhamento de um cristal é necessário somar o fator de estrutura de cada cela unitária, com respeito a uma única origem. Assim, para uma cela unitária com sua própria origem na posição $t\vec{a} + u\vec{b} + v\vec{c}$, onde t , u , e v são números inteiros e $\vec{a}$, $\vec{b}$ e $\vec{c}$ são vetores de translação, o espalhamento será:

$$\vec{F}(\vec{S}) \times \exp[2\pi i t \vec{a} \cdot \vec{S}] \times \exp[2\pi i u \vec{b} \cdot \vec{S}] \times \exp[2\pi i v \vec{c} \cdot \vec{S}] \quad (7)$$

A onda total espalhada $\vec{K}(\vec{S})$ será dada pela somatória sobre todas as celas unitárias:

$$\vec{K}(\vec{S}) = \vec{F}(\vec{S}) \times \sum_{t=0}^{n_1} \exp[2\pi i t \vec{a} \cdot \vec{S}] \times \sum_{u=0}^{n_2} \exp[2\pi i u \vec{b} \cdot \vec{S}] \times \sum_{v=0}^{n_3} \exp[2\pi i v \vec{c} \cdot \vec{S}] \quad (8)$$

onde n_1 , n_2 e n_3 determinam o número de celas unitárias nas direções **a**, **b** e **c** respectivamente. Se os valores de n_1 , n_2 e n_3 forem suficientemente grandes a somatória $\sum_{t=0}^{n_1} \exp[2\pi i t \vec{a} \cdot \vec{S}]$ e as outras duas feitas sobre u e v serão sempre iguais a zero, a menos que $\vec{a} \cdot \vec{S}$ seja um inteiro h , $\vec{b} \cdot \vec{S}$ seja um inteiro k e $\vec{c} \cdot \vec{S}$ seja um inteiro l . Pode-se concluir que um cristal só espalha raios X quando as condições abaixo são satisfeitas:

$$\vec{a} \cdot \vec{S} = h$$

$$\vec{b} \cdot \vec{S} = k$$

$$\vec{c} \cdot \vec{S} = l$$

Essas condições são conhecidas como condições de Laue [5].

II.4.1 Lei de Bragg

Na tentativa de visualizar espalhamento de raios X por cristais, W.L. Bragg deduziu uma equação simples que tratava a difração como reflexões de planos paralelos em um retículo cristalino.

Feixes de raios X incidem sobre uma família de planos paralelos do cristal e sofrem reflexões de maneira semelhante a do espelho plano, portanto, o ângulo de incidência e o ângulo de reflexão são iguais (fig.6). Reflexões de máxima intensidade serão observadas em interferências construtivas. Ondas de mesma frequência e amplitude propagando-se com velocidades iguais na mesma direção e sentido, quando superpostas resultam em uma nova onda de mesma frequência e de amplitude igual a $2A\cos(\phi/2)$, onde ϕ é a diferença de fase. Se as cristas de uma corresponder as cristas da outra, a diferença de fase entre ambas será nula ($\phi = 0$), e a amplitude da onda resultante será o dobro da amplitude de cada onda componente. Diz-se nesse caso, que as ondas interferem construtivamente, ou seja, estão em fase.

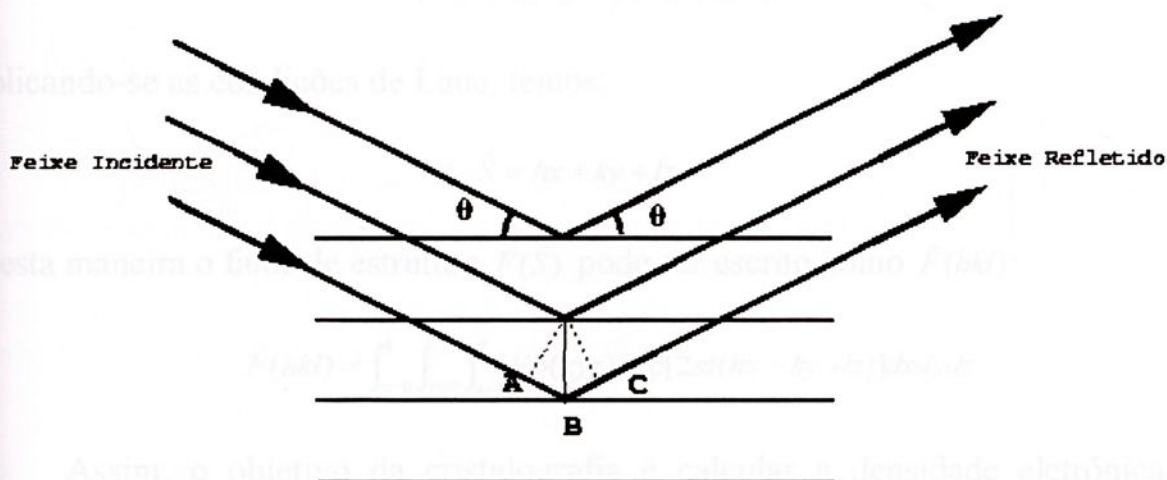


Figura 6 – Lei de Bragg

Fazendo-se com que a separação entre planos atômicos sucessivos seja igual a d , é fácil verificar que a diferença de caminho entre os dois raios incidentes 1 e 2 será dada por $AB + BC = 2d\sin\theta$. Espalhamento de planos

adjacentes do retículo devem estar em fase, ou seja, a diferença de caminhos entre os feixes deve ser um múltiplo inteiro do comprimento de onda. Desta maneira, para uma interferência construtiva teremos:

$$2d\sin\theta = n\lambda \quad (9)$$

onde λ é o comprimento de onda dos raios X e n é um número inteiro. A equação acima é conhecida como lei de Bragg.

II.5 Densidade Eletrônica

O fator de estrutura é uma função da distribuição da densidade eletrônica na cela unitária, conforme pode ser visto na equação 5, onde o vetor posição é $\vec{r} = \vec{a}x + \vec{b}y + \vec{c}z$ e a somatória é feita sobre todos os átomos da cela unitária. Transferindo-se esta somatória para uma integral sobre todos os elétrons da cela unitária obtemos a equação 6. Sabendo-se que x, y, z são as coordenadas fracionárias dos átomos na cela unitária, o produto escalar entre $\vec{r}$ e $\vec{S}$ será:

$$\vec{r} \cdot \vec{S} = x\vec{a} \cdot \vec{S} + y\vec{b} \cdot \vec{S} + z\vec{c} \cdot \vec{S}$$

aplicando-se as condições de Laue, temos:

$$\vec{r} \cdot \vec{S} = hx + ky + lz \quad (10)$$

Desta maneira o fator de estrutura $\vec{F}(\vec{S})$ pode ser escrito como $\vec{F}(hkl)$:

$$\vec{F}(hkl) = \int_{x=0}^1 \int_{y=0}^1 \int_{z=0}^1 V\rho(xyz)\exp[2\pi i(hx + ky + lz)]dx dy dz \quad (11)$$

Assim, o objetivo da cristalografia é calcular a densidade eletrônica em qualquer posição da cela unitária. Isto é possível devido ao fato de $\rho(xyz)$ ser a transformada de Fourier do fator de estrutura. Segundo as condições de Laue a difração ocorre em direções discretas. Então $\rho(xyz)$ pode ser escrito como uma função de $\vec{F}(hkl)$, por meio de uma somatória:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l \vec{F}(hkl) \exp[-2\pi i(hx + ky + lz)] \quad (12)$$

Sabendo que $\vec{F} = |F| \exp[i\alpha]$ e substituindo seu valor na somatória acima, teremos um novo valor para a densidade eletrônica:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)| \exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)] \quad (13)$$

onde $|F(hkl)|$ é a amplitude e $\alpha(hkl)$ é a fase do fator de estrutura. Em um experimento de difração de raios X por um cristal a única informação obtida são as intensidades dos feixes difratados, e conseqüentemente o módulo do fator de estrutura ao quadrado, pois a intensidade I é proporcional ao quadrado de $|F|^2$,

$$I \propto |F|^2$$

Toda informação referente à fase é perdida no experimento. Este é o problema da fase. Para a solução do problema da fase podem ser usados quatro métodos principais: MIR-(Substituição Isomórfica Múltipla), Substituição Molecular, Dispersão Anômala e Métodos Diretos (Drenth, 1993).

III. Métodos

O programa *Mathematica* foi usado para simular matematicamente a densidade eletrônica de cristais, de maneira que pudéssemos visualizar os picos e as curvas de níveis referentes a cada átomo, bem como, suas disposições geométricas no retículo cristalino. Para isso foi necessário calcular o fator de estrutura e a densidade eletrônica de cada um dos modelos cristalinos.

III.1 Os Modelos

Os modelos que descreveremos abaixo, são cristais teóricos hipotéticos, que foram idealizados com a finalidade de tornar o mais fácil possível a visualização dos conceitos cristalográficos que estarão envolvidos no processo de geração do mapa de densidade eletrônica.

Modelo I:- é um cristal unidimensional (na direção x) composto por um átomo de nitrogênio (N), dois átomos de oxigênio (O) e um átomo de ferro (Fe)

(fig.7). O comprimento de sua cela unitária é de 17 Å, com átomos posicionados conforme tabela 1.

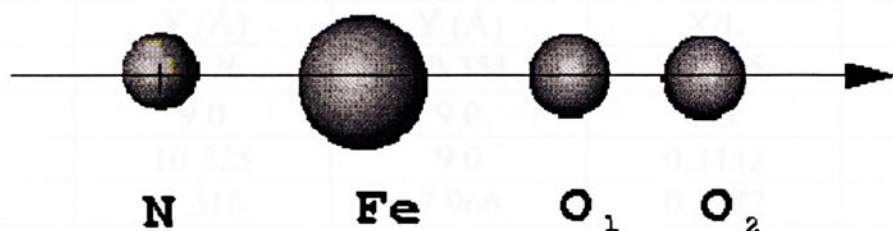


Figura 7 – Modelo I – Cristal Unidimensional

Tabela 1. - Posições dos átomos do modelo I na cela unitária com suas devidas coordenadas fracionárias.

Átomos	X (Å)	X/L
N	0	0
Fe	2,07	0.12177
O ₁	3.94	0.23177
O ₂	5.17	0.30412

Modelo II :- é um cristal bidimensional (na direção xy) de uma ligação peptídica sem a presença do átomo de hidrogênio (fig.8). O comprimento da cela unitária é de 30 Å em ambas as direções, com átomos posicionados conforme tabela 2.

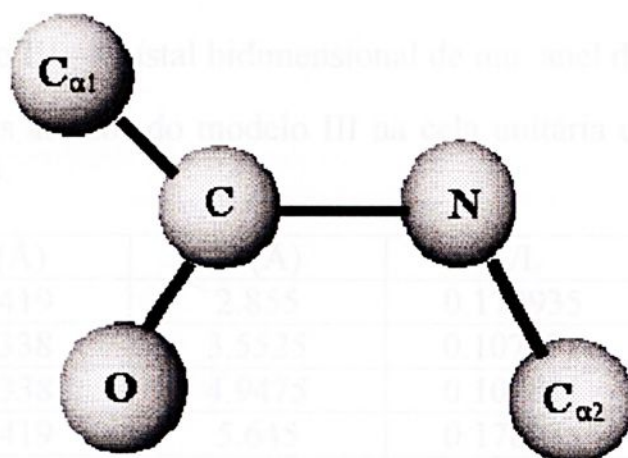


Figura 8 – Modelo II – Cristal bidimensional de uma ligação peptídica

Tabela 2. – Posições dos átomos do Modelo II na cela unitária com suas devidas coordenadas fracionárias.

Átomo	X (Å)	Y (Å)	X/L	Y/L
C _{α1}	8.326	10.351	0.2775	0.3450
C	9.0	9.0	0.3	0.3
N	10.325	9.0	0.3442	0.3
O	8.316	7.966	0.2772	0.2656
C _{α2}	11.096	7.766	0.3699	0.2589

Modelo III :- é um cristal bidimensional (na direção xy) de um anel benzeno sem a presença dos átomos de hidrogênio (fig.9). O comprimento da cela unitária é de 17 Å em ambas as direções, com átomos posicionados conforme tabela 3.

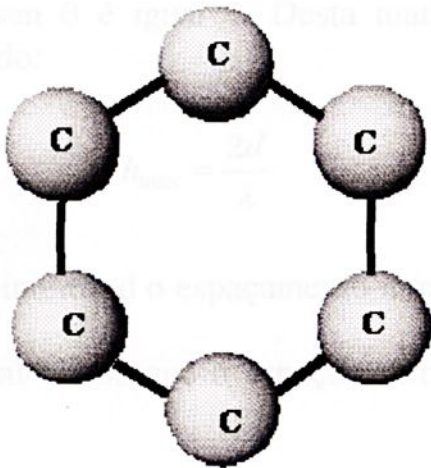


Figura 9 – Modelo III – Cristal bidimensional de um anel de benzeno

Tabela 3. - Posições dos átomos do modelo III na cela unitária com suas devidas coordenadas fracionárias.

Átomos	X (Å)	Y (Å)	X/L	Y/L
C ₁	3.0419	2.855	0.178935	0.167941
C ₂	1.8338	3.5525	0.107871	0.208971
C ₃	1.8338	4.9475	0.107871	0.291029
C ₄	3.0419	5.645	0.178935	0.332059
C ₅	4.25	4.9475	0.25	0.291029
C ₆	4.25	3.5525	0.25	0.208971

III.2 Procedimento para o cálculo do Fator de Estrutura

O fator de estrutura depende do posicionamento de cada átomo na cela unitária o que o torna dependente do fator de espalhamento atômico. Mas para efetuarmos o cálculo do fator de espalhamento atômico é necessário sabermos o número total de reflexões. Portanto, o primeiro passo para obtenção do fator de estrutura é calcular o número total de reflexões.

III.2.1 Cálculo do número de reflexões

Após cada modelo atômico estar devidamente definido podemos calcular seus índices de reflexões h, k, l , através a lei de Bragg (equação 9).

Para o caso unidimensional só teremos reflexões do tipo h e seu valor máximo é obtido quando $\sin \theta$ é igual 1. Desta maneira, podemos escrever a equação para $h_{\max}$ como sendo:

$$h_{\max} = \frac{2d}{\lambda} \tag{14}$$

onde h deve ser um número inteiro, d o espaçamento interplanar e λ o comprimento de onda do raio X.

No caso bidimensional temos que o espaçamento entre seus planos é dado por:

$$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2}$$

onde a e b são os comprimentos das arestas do cristal. Considerando-se um cristal quadrado, onde a é igual a b e o mesmo acontece com os valores dos interceptos h e k , podemos escrever:

$$\frac{1}{d^2} = \frac{2h^2}{a^2}$$

A razão $1/d$ será dada por:

$$\frac{1}{d} = \frac{h\sqrt{2}}{a}$$

Substituindo o valor da razão $1/d$ na equação 9, obtemos uma expressão para $h_{\max}$:

$$h_{\max} = \frac{2a}{\lambda\sqrt{2}} \quad (15)$$

III.2.2 Cálculo do Fator de Espalhamento Atômico

Para nosso caso usamos dados tabelados do fator de espalhamento atômico (International Table vol. III) [6], que são apresentados em função da razão $(\sin\theta)/\lambda$. A partir desses dados encontramos para cada reflexão um valor para $(\sin\theta)/\lambda$ e por interpolação sua respectiva intensidade de espalhamento, conforme implementação para o *Mathematica*.

III.2.3 Cálculo do Fator de Estrutura

O fator de estrutura para um cristal tridimensional é dado pela equação 5, que após aplicação das condições de Laue pode ser escrito da seguinte maneira:

$$F(hkl) = \sum_{j=1}^n f_j \exp(2\pi i h x_j + 2\pi i k y_j + 2\pi i l z_j)$$

onde a somatória é feita para todos os átomos do cristal.

Para o modelo 1, caso unidimensional, F_h será:

$$F(h) = f_n \exp(2\pi i x_n) + f_{Fe} \exp(2\pi i x_{Fe}) + f_{O1} \exp(2\pi i x_{O1}) + f_{O2} \exp(2\pi i x_{O2}).$$

Como o modelo é considerado centro-simétrico, teremos $\sin(-x) = -\sin x$ e $\cos(-x) = \cos x$, desta maneira a parte imaginária da exponencial é cancelada. Então $F(h)$ será uma combinação de termos reais:



$$F(h) = 2f_n \cos(2\pi h x_n) + 2f_{Fe} \cos(2\pi h x_{Fe}) + 2f_{O1} \cos(2\pi h x_{O1}) + 2f_{O2} \cos(2\pi h x_{O2})$$

onde os fatores de espalhamento atômicos possuem dependência direta das reflexões h .

Para os casos bidimensionais o fator de estrutura é dado de maneira semelhante ao caso unidimensional, levando-se em conta as reflexões do tipo h e k . Portanto, podemos escrever a expressão para $F(h,k)$:

$$F(hk) = 2 \sum_{j=1}^n f_j \cos[2\pi(x_j h + y_j k)]$$

onde f_j é uma combinação das reflexões h e k . O fator de espalhamento atômico (f_j) nos casos bidimensionais decai com valores dado pela raiz quadrada da soma de cada um dos índices da reflexão ao quadrado, conforme pode ser visto nas implementações dos modelos I e II.

III.3 Cálculo da Densidade Eletrônica

Para o caso unidimensional, conforme visto na seção II.5, a densidade eletrônica é dada por:

$$\rho(x) = \frac{1}{L} \sum_{-h}^h F(h) \exp(-2\pi i h x)$$

Expandindo em termos de coseno e seno, temos:

$$\rho(x) = \frac{1}{L} \sum_{-h}^h F(h) [\cos(2\pi h x) - i \sin(2\pi h x)]$$

Separando-se a somatória em partes, nos permite escrever a densidade eletrônica como:

$$\rho(x) = \frac{1}{L} \left\{ \sum_{-h}^{-1} F(h) [\cos(2\pi hx) - i \sin(2\pi hx)] + F_0 + \sum_1^h F(h) [\cos(2\pi hx) - i \sin(2\pi hx)] \right\}$$

Desta maneira, usando-se as propriedades de seno e cosseno, temos

$$\rho(x) = \frac{1}{L} \left\{ \sum_1^h F(h) [\cos(2\pi hx) + i \sin(2\pi hx)] + F_0 + \sum_1^h F(h) [\cos(2\pi hx) - i \sin(2\pi hx)] \right\}$$

Podemos observar na equação acima que os termos em seno se cancelam e, portanto, teremos duas somatórias idênticas em termos de cossenos. Então a densidade eletrônica será:

$$\rho(x) = \frac{1}{L} [F_0 + 2 \sum_1^h F(h) \cos(2\pi hx)]$$

No caso bidimensional a densidade eletrônica será dada por:

$$\rho(x, y) = \frac{1}{A} \sum_{-h, -k}^{h, k} F(hk) \exp[-2\pi i(hx + ky)]$$

Expandindo-se em termos de senos e cossenos, temos

$$\rho(x, y) = \frac{1}{A} \sum_{-h, -k}^{h, k} F(hk) \{ \cos[2\pi(hx + ky)] + i \sin[2\pi(hx + ky)] \}$$

Pela mesma propriedade apresentada no item anterior, os termos de seno se cancelam, então teremos:

$$\rho(x,y) = \frac{1}{A} \left\{ 2 \sum_{h,k}^{h,k} F(h,k) \cos[2\pi(hx + ky)] + 2 \sum_1^h F(h,0) \cos(hx) + 2 \sum_1^k F(0,k) \cos(ky) + 2 \sum_{h,k}^{h,k} F(h,k) \cos[2\pi(hx + ky)] \right\}$$

Somado-se os termos da dupla somatória, temos

$$\rho(x,y) = \frac{1}{A} \left\{ 4 \sum_{h,k}^{h,k} F(h,k) \cos[2\pi(hx + ky)] + 2 \sum_1^h F(h,0) \cos(hx) + 2 \sum_1^k F(0,k) \cos(ky) \right\}$$

IV. Resultados

Utilizando-se o comprimento de onda referente a radiação do CuK_α (λ = 1.5418 Å) e admitindo que a cela unitária é simétrica, de acordo com a natureza periódica da estrutura, encontramos para cada cristal uma amplitude máxima de reflexões conforme tabela 4.

Tabela 4. - Amplitude máxima de reflexões para cada modelo.

Modelos	hmax
Modelo 1	22
Modelo 2	22
Modelo 3	39

De posse do número máximo de reflexões para cada um dos cristais, foram elaboradas as implementações para o *Mathematica* do cálculo da densidade eletrônica, conforme implementação I, II e III.

Para cada modelo foi gerado um mapa de densidade eletrônica, conforme figuras 10, 11, 12. O resultado foi, como pode ser visto, mapas com picos bem definidos, que dão uma boa visão das disposições geométricas de cada átomo, bem como de suas intensidades.

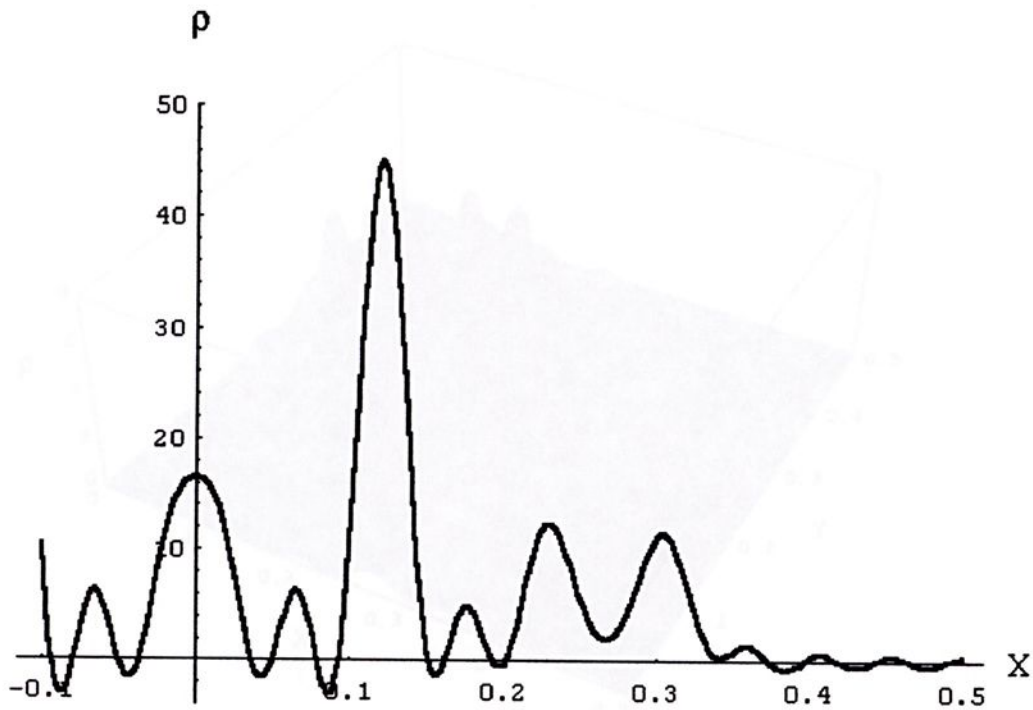


Figura 10 – Mapa de Densidade Eletrônica Modelo I.

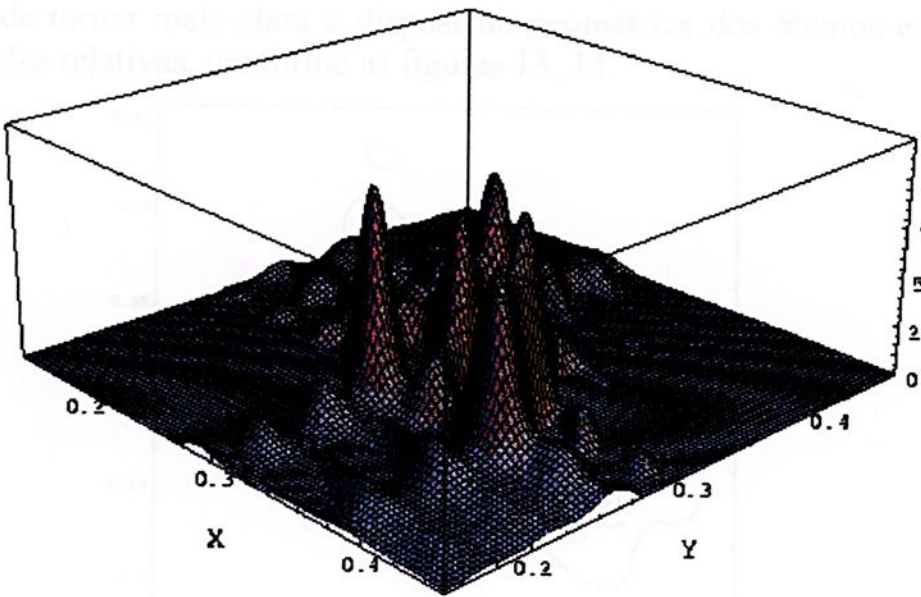


Figura 11 – Mapa de Densidade Eletrônica Modelo II.

Figura 12 – Gráfico de Contorno do Modelo II

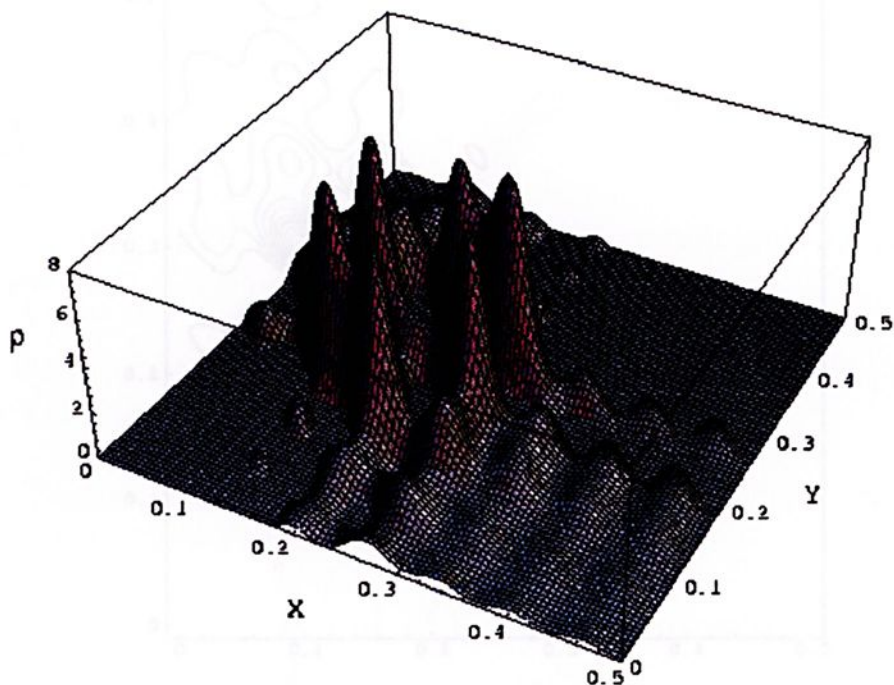


Figura 12 – Mapa de Densidade Eletrônica Modelo III.

Para os modelos bidimensionais foram geradas curvas de níveis com o objetivo de tornar mais clara a disposição geométrica dos átomos e também suas intensidades relativas, conforme as figuras 13, 14.

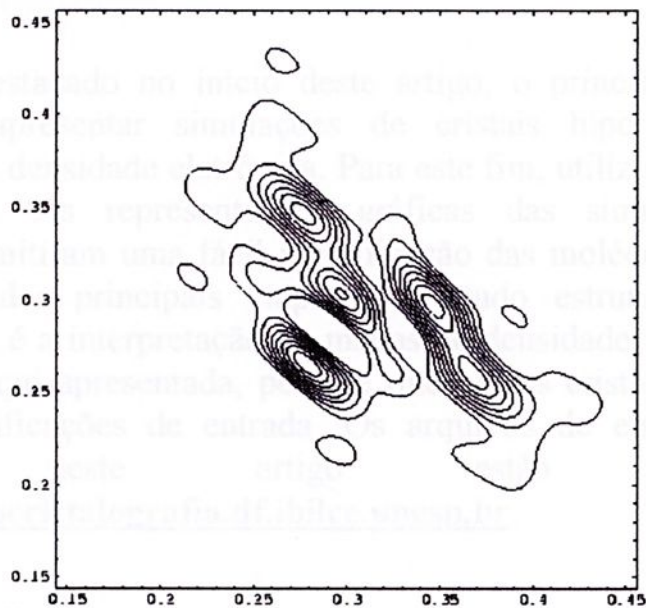


Figura 13 – Gráfico de Contorno do Modelo II.

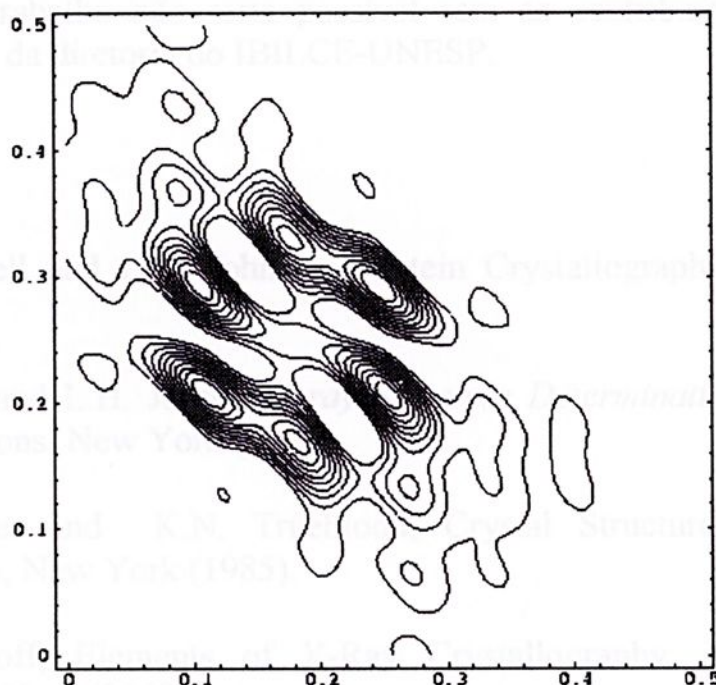


Figura 14 – Mapa de Contorno do Modelo III.

V. Conclusões

Como destacado no início deste artigo, o principal objetivo do presente trabalho era apresentar simulações de cristais hipotéticos, para facilitar a visualização da densidade eletrônica. Para este fim, utilizou-se o *Mathematica* para as simulações. As representações gráficas das simulações das densidades eletrônicas permitiram uma fácil identificação das moléculas representadas, o que enfatiza uma das principais etapas do estudo estrutural de macromoléculas biológicas, que é a interpretação de mapas de densidade eletrônica. Além disso, a metodologia, aqui apresentada, permite que outros cristais sejam simulados, com pequenas modificações de entrada. Os arquivos de entrada para os exemplos discutidos neste artigo estão disponíveis em <http://www.biocristalografia.df.ibilce.unesp.br>.

Agradecimentos:

Agradecemos aos professores Dr. Johnny Rizzieri Olivieri, Dr. Elso Drigo Filho, Dr. Vitor Barbanti Pereira Leite pela colaboração na realização deste

trabalho. Este trabalho não seria possível sem as contribuições financeiras da Capes, Fapesp e da diretoria do IBILCE-UNESP.

Referências:

[1] T.L. Blundell and L.N. Johnson, Protein Crystallography, Academic Press, London (1976).

[2] G.H. Stout and L.H. Jensen, *X-ray Structure Determination*, Second Edition, John Wiley & Sons, New York (1989).

[3] J.P. Glusker and K.N. Trueblood, Crystal Structure Analysis, Oxford University Press, New York (1985).

[4] L.V. Azároff, Elements of X-Ray Crystallography, McGraw-Hill Book Company, New York (1968).

[5] J. Drenth, Principles of protein X-ray crystallography, Editora Springer-verlag, Boston (1994).

[6] International Tables For X-Ray Crystallography (1974). Vol. III. The Kynoch Press, Birmingham, England.



Implementação I - Modelo I.

```

r = Table[((h*1.5418)/34), {h, 0, 22}]
dn = Interpolation[{{0,7.0},{0.07709,6.781},
{0.15418,6.203},{0.23127,5.42},{0.30836,4.6},
{0.38545,3.856},{0.46254,3.241},{0.53963,2.76},
{0.61672,2.397},{0.7709,2.17},{0.92508,1.944},
{1.07926,1.821},{1.23344,1.698},{1.38762,1.624},
{1.5418,1.55},{1.69598,1.497},{1.85016,1.444},
{2.00434,1.397}}]
n = dn[r]
do = Interpolation[{{0,8.0},{0.07709,7.796},
{0.15418,7.25},{0.23127,6.482},{0.30836,5.634},
{0.38545,4.814},{0.46254,4.094},{0.53963,3.492},
{0.61672,3.01},{0.7709,2.674},{0.92508,2.338},
{1.07926,2.141},{1.23344,1.944},{1.38762,1.829},
{1.5418,1.714},{1.69598,1.64},{1.85016,1.566},
{2.00434,1.514}}]
o = do[r]
dfe = Interpolation[{{0,26.0},{0.07709,25.3},
{0.15418,23.68},{0.23127,21.85},{0.30836,20.09},
{0.38545,18.40},{0.46254,16.77},{0.53963,15.31},
{0.61672,13.84},{0.7709,12.655},{0.92508,11.47},
{1.07926,10.59},{1.23344,9.71},{1.38762,9.09},
{1.5418,8.47},{1.69598,8.035},{1.85016,7.6},
{2.00434,7.295}}]
fe = dfe[r]
F = Table[ (2)n[[h]] + (2)fe[[h]]Cos[(0.7651)(h-1)] +
(2)o[[h]]Cos[(h-1)(1.4562)] + (2)o[[h]]Cos[(h-1)(1.9108)],
{h, 1, 23}]
g = Plot[(1/17)F[[1]] + (2/17)Sum[F[[h+1]]Cos[(2Pi(h))x],
{h, 1, 22}], {x, -0.1, 0.5}, PlotRange -> {-4, 50}]
Show[g, AxesLabel -> {
FontForm["X", {"Courier",14}],
FontForm["r", {"Symbol",14}]},
DefaultFont -> {"Courier", 10}]

```



Implementação II - Modelo II.

```

fc = Interpolation[{ {0,6},{1,5.95246},{2,5.78682},
{3,5.53032},{4,5.21020},{5,4.86070},{6,4.49388},
{7,4.11774},{8,3.75012},{9,3.41996},{10,3.13383},
{11,2.87802},{12,2.64951},{13,2.45143},{14,2.28443},
{15,2.14629},{16,2.03643},{17,1.94748},{18,1.86832},
{19,1.80087},{20,1.74347},{21,1.69451},{22,1.65090},
{23,1.61256},{24,1.57885},{25,1.54878},{26,1.52041},
{27,1.49319},{28,1.46747},{29,1.44282},{30,1.41848},
{31,1.39374},{32,1.36922},{33,1.34486},{34,1.32056},
{35,1.29604},{36,1.27153},{37,1.24702},{38,1.22250},
{39,1.19793}}];

fn = Interpolation[{ {0,7},{1,6.95127},{2,6.80467},{3,6.57633},
{4,6.28239},{5,5.93826},{6,5.56163},
{7,5.17288},{8,4.78496},{9,4.41048},
{10,4.05790},{11,3.73206},{12,3.43719},{13,3.17215},
{14,2.93750},{15,2.73055},{16,2.54973},{17,2.39271},
{18,2.25792},{19,2.14256},{20,2.04416},{21,1.96025},
{22,1.88600},{23,1.82146},{24,1.76587},{25,1.71780},
{26,1.67472},{27,1.63574},{28,1.60107},{29,1.56998},
{30,1.54143},{31,1.51441},{32,1.48920},{33,1.46542},
{34,1.44263},{35,1.41961},{36,1.39719},{37,1.37531},
{38,1.35389},{39,1.33298}}];

fo = Interpolation[{ {0,8},{1,7.95406},{2,7.81803},
{3,7.60447},{4,7.32595},
{5,6.99392},{6,6.62339},{7,6.23123},
{8,5.82942},{9,5.42972},
{10,5.04142},{11,4.67239},{12,4.32867},{13,4.01015},
{14,3.71871},{15,3.45398},{16,3.21672},{17,3.00407},
{18,2.81421},{19,2.64541},{20,2.49593},{21,2.36402},
{22,2.24734},{23,2.14476},{24,2.05496},{25,1.97645},
{26,1.90706},{27,1.84562},{28,1.79164},{29,1.74404},
{30,1.70122},{31,1.66164},{32,1.62596},{33,1.59370},
{34,1.56432},{35,1.53702},{36,1.51177},{37,1.48821},
{38,1.46597},{39,1.44401}}];

F = Table[(
  (2fc[Sqrt[(h)^2 + (k)^2]])Cos[(2Pi)(0.3(h) + 0.3(k))]
+ (2fc[Sqrt[(h)^2 + (k)^2]])Cos[(2Pi)(0.2775(h) + 0.3450(k))])

```



```

+ (2fc[Sqrt[ (h)^2 + (k)^2 ]])Cos[(2Pi)(0.3699(h) + 0.2589(k))]
+ (2fn[Sqrt[ (h)^2 + (k)^2 ]])Cos[(2Pi)(0.3442(h) + 0.3(k))]
+ (2fo[Sqrt[ (h)^2 + (k)^2 ]])Cos[(2Pi)(0.2772(h) + 0.2656(k))] ),
{h, 0, 27}, {k, 0, 27}];
G = Plot3D[ (4/900)Sum[F[[h+1,k+1]]Cos[(2Pi)(x(h) + y(k))],
{h, 1, 27},{k, 1, 27}] +
(2/900)Sum[F[[h+1,1]]Cos[ (2Pi) h x ], {h, 1, 27}] +
(2/900)Sum[F[[1,k+1]]Cos[ (2Pi) k y ], {k, 1, 27}] +
((1/900)*F[[1,1]]),
{x, 0.15, 0.45},{y, 0.15, 0.45}, PlotRange -> {0, 12},
PlotPoints -> {100,100}]
Show[G, ViewPoint->{1.300,-2.400,2.000},
AxesLabel -> {
FontForm["X", {"Courier",14}],
FontForm["Y", {"Courier",14}],
FontForm["r", {"Symbol",14}]}]
Show[ContourGraphics[G],
ContourShading -> False]

```



Implementação III - Modelo III.

```

fc = Interpolation[ {{0,6}, {1, 5.83498},
{2, 5.36704}, {3, 4.75428}, {4, 4.09565},
{5, 3.47407}, {6, 2.98062}, {7, 2.57599},
{8, 2.26668}, {9, 2.04816}, {10, 1.89484},
{11, 1.77611}, {12, 1.68662}, {13, 1.61467},
{14, 1.5573}, {15, 1.50738}, {16, 1.46159},
{17, 1.41848}, {18, 1.37497}, {19, 1.332},
{20, 1.28883}, {21, 1.24558}, {22, 1.20227}} ]
xc = {0.178935, 0.107871, 0.107871, 0.178935, 0.25,
0.25 }
yc = { 0.167941, 0.208971, 0.291029, 0.332059,
0.291029, 0.208971 }

```

```

F = Table[ (2(fc[Sqrt[ (h)^2 + (k)^2 ]])
(Sum[ Cos[ (2Pi)( (h)(xc[[j]]) + (k)(yc[[j]]) )],
{j, 1, 6}]),
{h, 0, 15},
{k, 0, 15} ];
g = Plot3D[ (4/289)Sum[
F[[h+1,k+1]]Cos[ (2Pi)( h x + k y )],
{h, 1, 15}, {k, 1, 15} ] +
(1/289)F[[1,1]] +
(2/289)Sum[ F[[h+1,1]]Cos[ (2Pi) h x ], {h, 1, 15}] +
(2/289)Sum[ F[[1,k+1]]Cos[ (2Pi) k y ], {k, 1, 15}],
{x, 0, 0.5}, {y, 0, 0.5}, PlotRange -> {0, 8},
PlotPoints -> {100,100}]
Show[g, ViewPoint->{1.775,-2.073,2.000},
AxesLabel -> {
FontForm[" X", {"Courier",14}],
FontForm[" Y", {"Courier",14}],
FontForm[" r", {"Symbol",14}]]
Show[ContourGraphics[g],
ContourShading -> False]

```



APÊNDICE - II

MÉTODO DA SUBSTITUIÇÃO MOLECULAR (Drenth, 1994)

1. Introdução

Para determinar a estrutura de uma proteína é necessário conhecer do seu padrão de difração: os fatores de estrutura e os ângulos de fase relacionados. Mas durante o experimento as informações sobre os ângulos de fase não são possíveis de medir, impossibilitando a determinação direta da estrutura. Essa perda de todas as informações referentes à fase é conhecida, em cristalografia como o problema da fase. Algumas técnicas foram desenvolvidas com o objetivo de resolver esse problema, entre elas o método da substituição molecular (Drenth, 1994).

O método da substituição molecular, tem-se constituído em um método rápido para determinação de estruturas de proteínas de uma mesma família. O método requer um nível de semelhança entre duas estruturas protéicas que serão correlacionadas, bem como uma identidade sequencial.

Este método pode ser dividido em três partes:



1 – Rotação do modelo:

Nesta etapa uma função rotação é usada para encontrar a orientação correta do modelo em relação ao retículo cristalino do cristal a ser determinado.

2 – O modelo obtido a partir da função rotação é posicionado corretamente no retículo cristalino:

Uma função translação encontra os parâmetros translacionais para posicionar o modelo corretamente na cela unitária do cristal.

3- Com a orientação e a posição correta é possível usar as fases do modelo, para calcular a densidade eletrônica da estrutura cristalina desconhecida.

Neste passo, a estrutura desconhecida está pronto para ser refinada.

Com o aumento do número de estruturas conhecidas, este método tende a ser usado com maior frequência, devido ao seu baixo custo e rapidez.

2. A função rotação

A função rotação é usada para encontrar a orientação correta do modelo no retículo cristalino. Isto pode ser obtido por um teste de concordância entre as funções de Patterson calculada do modelo e a obtida diretamente do experimento de difração.

Transferir a estrutura molecular da proteína resolvida para o arranjo cristalino da desconhecida é o principal problema do método da



substituição molecular. Para que isso seja feito é necessário uma orientação adequada e um posicionamento preciso do modelo na cela unitária alvo.

Dois passos são necessários para a transformação de um grupo de coordenadas para outro: **ROTAÇÃO E TRANSLAÇÃO**.

2.1 O cálculo da função rotação

Uma das maneiras possíveis de descrever uma orientação, é definir um eixo de rotação e os ângulos de rotação em torno deles, isto pode ser feito através de uma função rotação.

Assumindo-se que, em um sistema cristalino com eixos ortogonais, um átomo esteja localizado na posição:

$$\vec{X} = X_1\vec{a}_1 + X_2\vec{a}_2 + X_3\vec{a}_3$$

da rede cristalográfica com eixos $\vec{a}_1, \vec{a}_2, \vec{a}_3$. Rotacionando o sistema, em torno da origem, tem-se um novo grupo de coordenadas $\vec{a}_{r,1}, \vec{a}_{r,2}, \vec{a}_{r,3}$. A posição do átomo em relação a este novo sistema de coordenadas será:

$$\vec{X} = X_{r,1}\vec{a}_{r,1} + X_{r,2}\vec{a}_{r,2} + X_{r,3}\vec{a}_{r,3}$$

A relação entre os dois grupos de coordenadas é dada por :

$$X_r = [C]X$$

onde $[C]$ é a matriz rotação.



Rotacionar o eixo tem efeito igual ao de rotacionar a estrutura na direção oposta. O mapa de Patterson é rotacionado de maneira igual. Aplicando-se a rotação $[C]$ para a função de Patterson $P(\vec{u})$, obtém-se a função de Patterson rotacionada $P_r(\vec{u}_r)$.

Uma superposição da função $P(\vec{u})$ com sua rotação $P_r(\vec{u}_r)$, define a função rotação:

$$R(\alpha, \beta, \gamma) = \int_U P(\vec{u}) P_r(\vec{u}_r) d\vec{u}$$

onde U é o volume do mapa de Patterson.

O produto depende dos ângulos de rotação, e um valor máximo é obtido para a superposição correta (quando os dois grupos vetoriais $P(\vec{u})$ e $P_r(\vec{u}_r)$ tiverem orientações equivalentes).

Tanto $P(\vec{u})$ como $P_r(\vec{u}_r)$ podem ser expandidos em uma série de Fourier:

$$P(\vec{u}) = \frac{1}{v} \sum_h |F(\vec{h})|^2 \exp[-2\pi i \vec{h} \vec{u}]$$

$$P_r(\vec{u}_r) = \frac{1}{v} \sum_{h'} |F(\vec{h}')|^2 \exp[-2\pi i \vec{h}' \vec{u}_r]$$

mas $\vec{u}_r = [C]\vec{u}$. Substituindo em $P_r(\vec{u}_r)$, temos

$$P_r(\vec{u}_r) = \frac{1}{v} \sum_{h'} |F(\vec{h}')|^2 \exp[-2\pi i \vec{h}' [C]\vec{u}]$$

da relação $\vec{h} = [C^{-1}]\vec{h}'$ e $\vec{h}' = [C]\vec{h}$, pode-se escrever $P_r(\vec{u}_r)$ do seguinte modo:

$$P_r(\vec{u}_r) = \frac{1}{v} \sum_{\vec{h}'} |F([C]\vec{h}')|^2 \exp[-2\pi i \vec{h}' \vec{u}]$$

Substituindo os dois mapas de Patterson na função rotação;

$$R(\alpha, \beta, \gamma) = \int_U \frac{1}{v^2} \sum_{\vec{h}} |F(\vec{h})|^2 \sum_{\vec{h}'} |F([C]\vec{h}')|^2 \exp[-2\pi i \vec{h} \vec{u}] \exp[-2\pi i \vec{h}' \vec{u}] d\vec{u}$$

colocando os termos independentes de $\vec{u}$ fora da integral, tem-se:

$$R(\alpha, \beta, \gamma) = \frac{1}{v^2} \sum_{\vec{h}} |F(\vec{h})|^2 \sum_{\vec{h}'} |F([C]\vec{h}')|^2 \int_U \exp[-2\pi i (\vec{h} + \vec{h}') \vec{u}] d\vec{u}$$

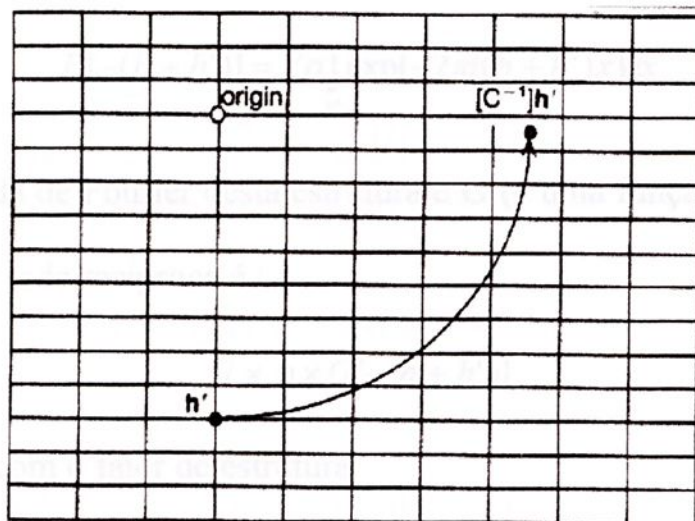


Fig. 1 - Efeito da rotação $[C^{-1}]$ no espaço recíproco.

Dois problemas aparecem aqui:

- 1 - Para o cálculo de $P_r(\vec{u}_r)$ os coeficientes devem mostrar as posições de $[C]\vec{h}'$ no espaço recíproco, que é impossível. Esta impossibilidade deve-

se ao fato de $[C]h'$ conter posições no espaço recíproco com valores não inteiros.

2 - Um enorme número de termos deve ser calculado, porque a multiplicação é requerida de cada um dos h com todos os h' .

Para resolver esses os problemas, supõe-se que exista um retículo cristalino com estrutura muito especial. Ele contém em cada cela unitária um corpo com forma e volume dado por U , onde o centro de U é a origem da cela unitária e a densidade eletrônica dentro de U é $\rho(x) = \rho$ e fora de U $\rho(x) = 0$. O Fator de estrutura desta estrutura especial, com posições na rede recíproca $-(\mathbf{h} + \mathbf{h}')$ é,

$$\bar{F}[-(\vec{h} + \vec{h}')] = V\rho \int_V \exp[-2\pi i(\vec{h} + \vec{h}')\vec{x}] d\vec{x}$$

A transformada de Fourier desta estrutura é G (é uma função interferência) e seu valor na rede recíproca é :

$$U \times \rho \times G[-(\vec{h} + \vec{h}')] =$$

Comparando com o fator de estrutura:

$$\int_V \exp[-2\pi i(\vec{h} + \vec{h}')\vec{x}] d\vec{x} = \frac{U}{V} \times G[-(\vec{h} + \vec{h}')] =$$

Substituindo o valor da integral acima na função rotação:

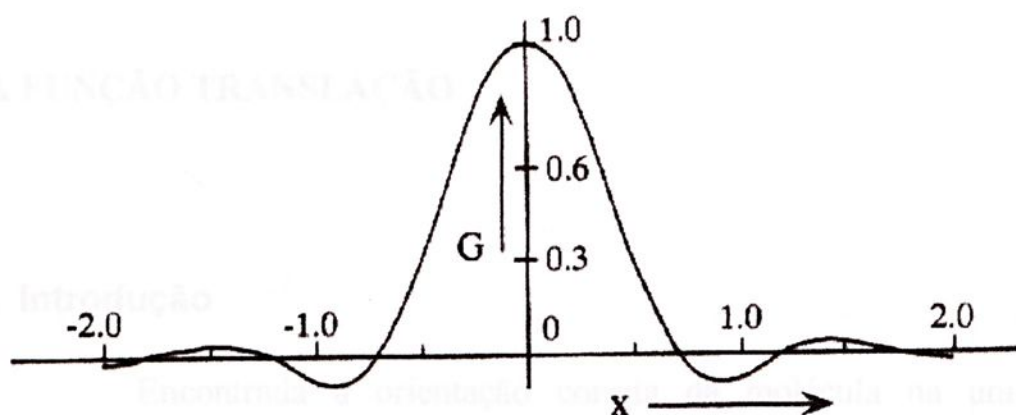
$$R(\alpha, \beta, \gamma) = \frac{U}{V^3} \sum_{\vec{h}} \sum_{\vec{h}'} |F(\vec{h})|^2 |F([C]\vec{h}')|^2 \times \bar{G}[-(\vec{h} + \vec{h}')] =$$

As propriedades de $G[-(\vec{h} + \vec{h}')]]$ permitem a solução dos dois problemas acima. Usualmente U é esférico e a transformada de uma esfera com raio r e centro na cela unitária é:

$$G = \frac{3(\sin 2\pi x - 2\pi x \cos 2\pi x)}{(2\pi x)^3}$$

x é neste caso igual a $(\vec{h} + \vec{h}') \cdot \vec{r}$.

Esta função pode ser representada graficamente dessa maneira:



No gráfico G tem seu valor máximo para $\mathbf{h}' = -\mathbf{h}$, e decresce rapidamente para valores de $\mathbf{h}'$ diferente $-\mathbf{h}$.

Desta maneira a somatória da função rotação pode ser feita por todos os $\mathbf{h}$ com um número limitado $\mathbf{h}'$. Isto resolve o segundo problema.

O primeiro problema pode ser resolvido da seguinte maneira:

- $|F([C]\vec{h}')|^2$ é calculado para um número integral de pontos na rede entorno de $([C]\vec{h}')$.

- O valor de $|F|^2$ com $([C]\vec{h}')$ é obtido por interpolação, dando um peso $G[-(\vec{h} + \vec{h}')]]$ para os valores de $|F|^2$ com pontos de rede integral em torno de $([C]\vec{h}')$.

A aplicação da rotação é importante para todas as reflexões presentes, porque o cálculo da função rotação depende basicamente da rotação do mapa de Patterson, que é determinado principalmente pelas reflexões fortes.

3. A FUNÇÃO TRANSLAÇÃO

3.1 Introdução

Encontrada a orientação correta da molécula na unidade assimétrica, temos que encontrar os parâmetros translacionais para posicionar a molécula corretamente no eixo de simetria. Para encontrar a superposição correta do modelo rotacionado é usado a função translação, que é conceitualmente relacionada a função rotação. Ela investiga a correlação entre as intensidades observadas, os vetores intermoleculares, a simetria do modelo e como este é movido sobre a cela unitária. Quando o modelo for posicionado corretamente a função deverá ter picos com valores correspondentes aos vetores translações, existentes entre as



simetrias moleculares relacionadas. Há algumas formas para a função translação, mas a função T2 é normalmente mais usada.

3.2 O CÁLCULO DA TRANSLAÇÃO

Neste exemplo, a translação será feita para uma estrutura desconhecida com cristais pertencentes ao grupo espacial P222. Esta estrutura tem uma molécula na unidade assimétrica e quatro moléculas na cela unitária. É escolhido um par de moléculas, com orientação determinada pela função rotação, porém não é conhecida sua posição na cela unitária. Através da função translação podemos determinar a posição da molécula 1 em relação a molécula 2, e subsequente para qualquer outro par de simetria que relacione moléculas.

Fig. 3 - Posição (x) na molécula 1 e (x + u) na molécula 2.

Onde O_1 e O_2 são origens locais.

A origem local da molécula 1 é em O_1 , da molécula 2 em O_2 .

(Fig. 2 e 3) e a origem da cela unitária em O . Os vetores intermoleculares de

Patterson entre duas moléculas podem ser calculados por

$$P_{12}(u) = \int \lambda_1(x) \lambda_2(x+u) dx$$

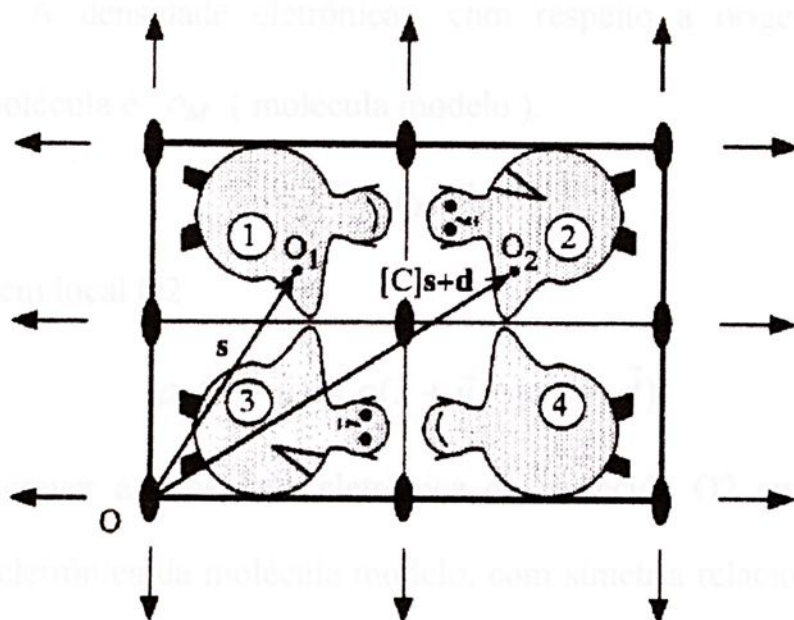


Fig. 2 - Uma cela unitária para o grupo espacial P222.

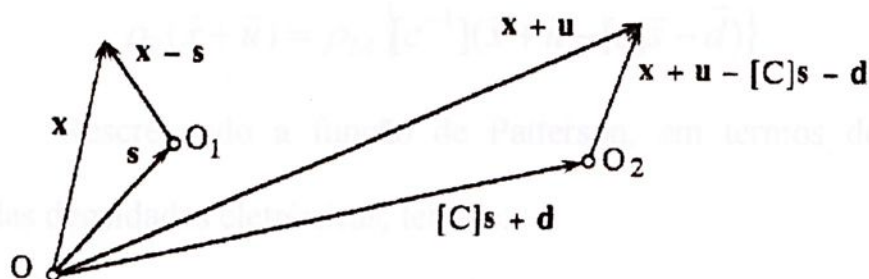


Fig . 3 - Posição ($\mathbf{x}$) na molécula 1 e ($\mathbf{x} + \mathbf{u}$) na molécula 2.

Onde O_1 e O_2 são origens locais.

A origem local da molécula 1 é em O_1 , da molécula 2 em O_2 (fig.2 e 3) e a origem da cela unitária em O . Os vetores intermoleculares de Patterson entre duas moléculas podem ser calculados por :

$$P_{1,2}(\vec{u}) = \int_V \rho_1(\vec{x}) \times \rho_2(\vec{x} + \vec{u}) d\vec{u}$$

A densidade eletrônica com respeito a origem local da primeira molécula é ρ_M (molécula modelo).

$$\rho_1(\vec{x}) = \rho_M(\vec{x} - \vec{s})$$

Para a origem local O2

$$\rho_2(\vec{x} + \vec{u}) = \rho(\vec{x} + \vec{u} - [c]\vec{s} - \vec{d})$$

Pode-se escrever a densidade eletrônica da molécula O2 em função da densidade eletrônica da molécula modelo, com simetria relacionada com a posição, dada por:

$$[c^{-1}](\vec{x} + \vec{u} - [c]\vec{s} - \vec{d})$$

$$\rho_2(\vec{x} + \vec{u}) = \rho_M\{[c^{-1}](\vec{x} + \vec{u} - [c]\vec{s} - \vec{d})\}$$

Rescrevendo a função de Patterson, em termos dos novos valores das densidades eletrônicas, tem-se:

$$P_{1,2}(\vec{u}, \vec{s}) = \int_V \rho_M(\vec{x} - \vec{s}) \times \rho_M\{[c^{-1}](\vec{x} + \vec{u} - [c]\vec{s} - \vec{d})\} d\vec{x}$$

Se ρ_M é escrito por uma serie de Fourier, em termos do fator de estrutura

$\vec{F}_M$, tem-se

$$P_{1,2}(\vec{u}, \vec{s}) = \sum_{\vec{h}} \sum_{\vec{p}} \vec{F}_M(\vec{h}) \times \vec{F}_M(\vec{p}) \exp[2\pi i(\vec{h} \cdot \vec{s} + \vec{p}[c^{-1}][c]\vec{s} + \vec{d}))] \\ \times \exp[-2\pi i\vec{p}[c^{-1}]\vec{u}] \int_V \exp[-2\pi i(\vec{h} + \vec{p}[c^{-1}])\vec{x}] d\vec{x}$$

Devido $[c]$ ser uma rotação cristalográfica, $p[c^{-1}]$ é um ponto da rede recíproca e, desta maneira, $(\vec{h} + p[c^{-1}])$ é um número inteiro e como conseqüência a integral se anula, quando é somada todas contribuições de $\mathbf{h}$ e $\mathbf{h}'$, a menos que $(\vec{h} + p[c^{-1}]) = 0$. Então o valor da integral é igual a V , com $\mathbf{p} = -\mathbf{h}[c]$. Supondo que o cálculo seja feito para um volume integral, com $V = 1$, pode-se rescrever $P_{1,2}(\vec{u}, \vec{s})$:

$$P_{1,2}(\vec{u}, \vec{s}) = \sum_{\vec{h}} \vec{F}_M(\vec{h}) \cdot \vec{F}_M^*(\vec{h}[c]) \exp[2\pi i \vec{h}(\vec{s} - [c]\vec{s} - \vec{d})] \exp[2\pi i \vec{h} \cdot \vec{u}]$$

Se $\vec{t} = -\vec{s} + [c]\vec{s} + \vec{d}$ é o vetor intermolecular (O1 e O2). Então pode-se rescrever a Função de Patterson cruzada como função do vetor $\mathbf{t}$:

$$P_{1,2}(\vec{u}, \vec{t}) = \sum_{\vec{h}} \vec{F}_M(\vec{h}) \cdot \vec{F}_M^*(\vec{h}[c]) \exp[2\pi i \vec{h} \cdot \vec{t}] \exp[2\pi i \vec{h} \cdot \vec{u}]$$

Esta é a função cruzada de Patterson do modelo estrutural, no qual duas moléculas são relacionadas por simetria cristalográfica. Ela deve ser comparada com a função de Patterson $P(\vec{u})$. Para este fim a função translação $T(\vec{t})$ é calculada:

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

Quando o vetor intermolecular $\vec{t}$ é igual ao vetor intermolecular verdadeiro $\vec{t}_0$, a função $T(\vec{t})$ deverá encontrar um valor máximo, porque a função de

Patterson $P_{1,2}(\vec{u}, \vec{t})$ calculada encaixa-se corretamente na função de

Patterson observada $P(\vec{u})$. Se $P_{1,2}(\vec{u}, \vec{t})$ é substituída em $T(\vec{t})$ e $P(\vec{u})$ por

$\sum_{\vec{p}} |F_{obs}(\vec{p})|^2 \exp[-2\pi i \vec{p} \cdot \vec{u}]$ a função será dada por:

$$T(\vec{t}) = \sum_{\vec{h}} \vec{F}_M(\vec{h}) \cdot \vec{F}_M^*(\vec{h}[c]) \exp[-2\pi i \vec{h} \cdot \vec{t}] \times \sum_{\vec{p}} |F_{obs}(\vec{p})|^2 \int_V \exp[2\pi i (\vec{h} - \vec{p}) \cdot \vec{u}] d\vec{u}$$

Devido a mesma razão antes verificada para o cálculo de $P_{1,2}(\vec{u}, \vec{t})$, aqui a integral se anulará a menos que $\vec{p}$ seja igual a $\vec{h}$. Com $\vec{p} = \vec{h}$ o valor de $T(\vec{t})$ será:

$$T(\vec{t}) = \sum_{\vec{h}} |F_{obs}(\vec{h})|^2 \vec{F}_M(\vec{h}) \cdot \vec{F}_M^*(\vec{h}[c]) \exp[-2\pi i \vec{h} \cdot \vec{t}]$$

A forma final da função translação é uma somatória de Fourier com coeficientes conhecidos. Desta maneira, ela pode ser facilmente calculada.

3.3 A Função Translação $\vec{T}_2$

A função translação $T(\vec{t})$ é bidimensional com os vetores intermoleculares $\vec{t}$ perpendiculares ao eixo de simetria. Para uma solução completa da estrutura desconhecida, as várias funções de translação bidimensionais devem ser combinadas.

Uma função translação que resulte nos picos correspondentes a todos os vetores intermoleculares da estrutura desconhecida foi desenvolvida por Crowther e Blow (1967). Esta função translação é chamada de $\vec{T}_2$.



O modelo, usado na função $\vec{T}_2$, não somente contém um par de moléculas, mas um número igual ao de moléculas da cela unitária da estrutura desconhecida.

Devido ao grande número de vetores intermoleculares, é mais conveniente o uso do vetor posição $\vec{m}$ do modelo molecular M. As posições e orientações das outras moléculas ($j = 2 \dots n$) na cela unitária, são derivadas do modelo molecular por operações $[c_j]\vec{m} + \vec{d}_j$.

Deve-se agora calcular a função de Patterson, escrevendo o fator de estrutura calculado do modelo:

$$\vec{F}_{calc}(\vec{h}, \vec{m}) = \sum_{j=1}^n \vec{F}_M(\vec{h}[c_j]) \exp[-2\pi i \vec{h}([c_j]\vec{m} + \vec{d}_j)]$$

onde $\vec{F}_M(\vec{h}[c_j])$ é a contribuição do fator de estrutura da reflexão $\vec{h}$ da j -ésima molécula em relação a origem local de orientação $[c_j]$. Atentando para o fato da origem da molécula estar em $[c_j]\vec{m} + \vec{d}_j$, pode-se escrever o fator de estrutura assim:

$$\vec{F}_{calc}(\vec{h}, \vec{m}) = \sum_{j=1}^n \vec{F}_M(\vec{h}[c_j]) \exp[-2\pi i \vec{h} \vec{d}_j] \exp[-2\pi i \vec{h}[c_j]\vec{m}]$$

O modulo do fator de estrutura ao quadrado é igual ao produto do fator de estrutura pelo seu complexo conjugado:

$$|\vec{F}_{calc}(\vec{h}, \vec{m})|^2 = \vec{F}_{calc}(\vec{h}, \vec{m}) \cdot \vec{F}_{calc}^*(\vec{h}, \vec{m})$$

Para diminuir o número de termos na função de Patterson calculada fazemos:

$$|\vec{F}_{calc}(\vec{h}, \vec{m})|^2 = Q$$

Então escrevendo-se a função de Patterson calculada:

$$\vec{P}_{calc}(\vec{m}, \vec{u}) = \sum_{\vec{h}} Q \exp[-2\pi i \vec{h} \cdot \vec{u}]$$

Calculando a nova função translação:

$$\vec{T}_2(\vec{m}) = \int_V P_{obs}(\vec{u}) \times P_{calc}(\vec{m}, \vec{u}) d\vec{u}$$

onde $P_{obs}(\vec{u})$ é a função de Patterson observada do cristal a ser determinado e $P_{calc}(\vec{m}, \vec{u})$ é a função de Patterson calculada do modelo.

Como antes, esta equação contém uma integral

$$\int_V \exp[2\pi i(\vec{h} - \vec{p}) \cdot \vec{u}] d\vec{u}$$

e de igual modo seu valor será nulo a menos que $\mathbf{h} = \mathbf{p}$, desta maneira

pode-se escrever $\vec{T}_2(\vec{m})$:

$$\begin{aligned} \vec{T}_2(\vec{m}) &= \sum_{\vec{h}} |\vec{F}_{obs}(\vec{h})|^2 \sum_j \sum_k \vec{F}_M(\vec{h}[c_j]) \vec{F}_M^*(\vec{h}[c_k]) \times \exp[-2\pi i \vec{h}(\vec{d}_j - \vec{d}_k)] \\ &\times \exp[-2\pi i \vec{h}([c_j] - [c_k])\vec{m}] \end{aligned}$$

Esta função utiliza todos os operadores de simetria simultaneamente, sendo portanto uma função tridimensional.

Purification, crystallization and preliminary X-ray analysis of haemoglobin I from the armoured catfish *Liposarcus anisitsi*

A. L. S. Smarra,^a W. F. de Azevedo Jr.,^a V. Fadel,^{a,c,d} P. Delatorre,^a M. Dellamano,^a M. F. Colombo^a and G. O. Bonilla-Rodriguez^{b*}

^aDepartamento de Física, Instituto de Biociências, Letras e Ciências Exatas, UNESP, São José do Rio Preto, SP 15054-000, Brazil,

^bDepartamento de Química e Geociências, Instituto de Biociências, Letras e Ciências Exatas, UNESP, São José do Rio Preto, SP 15054-000, Brazil, ^cDepartamento de Física, Faciba-FEB, Barretos, SP 14783-226, Brazil, and ^dCentro Universitário de Votuporanga, CEUV, Votuporanga, SP 15500-030, Brazil

Correspondence e-mail:
bonilla@qeg.ibilce.unesp.br

Liposarcus anisitsi is an armoured catfish that presents accessory air oxygenation through a modified stomach, which allows this species to survive in waters with very low oxygen content. Analysis of its haemolysate has shown the presence of four haemoglobins; this work focuses on the main component, haemoglobin I. It has been crystallized in two different forms and X-ray diffraction data have been collected to 2.77 and 2.86 Å resolution using synchrotron radiation. Crystals were determined to belong to the space groups *C2* and *P2₁* and preliminary structural analysis revealed the presence of one tetramer in the asymmetric unit in both crystal forms. The structure was determined using a standard molecular-replacement technique.

Received 31 March 1999

Accepted 1 February 2000

1. Introduction

Haemoglobin functions in aerobic organisms by dramatically increasing oxygen diffusion from the environment (aquatic or atmospheric) to its reduction site at the inner mitochondrial membrane. Aerobic metabolism in multicellular organisms would be impossible without this protein.

Owing to its abundance and availability, haemoglobin has been largely used as a model in order to gain understanding of oligomeric proteins, particularly concerning the heterotropic 'fine tuning' of oxygen binding in terms of affinity and cooperativity (Smarra *et al.*, 1999). Recent reports describing 'new' structures found for human adult haemoglobin (Smith *et al.*, 1991; Mercedes-Silva *et al.*, 1992; Seixas *et al.*, 1999) additional to those considered by the classical allosteric model of Monod *et al.* (1965) have revealed new aspects, despite the large amount of research spent on the subject.

The study of the structures of other forms of haemoglobin could contribute to providing a better understanding of haemoglobin function. The animal kingdom has plenty of isomorphous haemoglobins working under a variety of physiological requirements and subject to environmental stress, being a natural source for the study of structure-function relationships.

L. anisitsi (Siluriformes, Loricariidae) is a South American armoured catfish with accessory air oxygenation using a modified stomach. This morphological and histological adaptation allows the fish to hold swallowed air and oxygenate blood circulating on abundant capillaries close to the mucosa (Smarra *et al.*,

1997). Its blood has four haemoglobins, with isoelectric points (pI) in the range 7.2–8.3. These proteins were denominated I to IV according to their pI, starting with the highest value.

The present work was performed on the oxygenated form of Hb-I (which comprises about 50% of the haemolysate). Structural studies have also been initiated on the unliganded forms.

2. Materials and methods

2.1. Purification

Blood was collected by caudal puncture after an adaptive period of 4 d in aquaria. Subsequent procedures were carried out at low temperature (around 277 K). Red blood cells were washed by centrifugation four times with buffered saline (50 mM Tris pH 8.5, 1 mM EDTA). Haemolysis was accomplished with buffer A (50 mM Tris pH 8.5, 1 mM EDTA), followed by centrifugation (1000g for 1 h) and filtration to remove debris. An aliquot was kept for phosphate determination. For phosphate removal, haemolysate dialysis was performed against buffer A, followed by gel filtration (Sephadex G-25) on a 1.6 × 70 cm column.

Haemoglobin purification was performed on DEAE-Sephadex A-50 using a pH gradient between 50 mM Tris buffer pH 9.2 and 50 mM bis-Tris pH 6.5. The isolated components were further deionized by several passages through a mixed-bed Amberlite MB-1 (Sigma), concentrated by centrifugation on Amicon microconcentrators and stored in liquid nitrogen until use.



Table 1
Detailed X-ray diffraction statistics for synchrotron data of crystal form I.

Resolution range (Å)	Number of independent reflections	$I/\sigma(I)$	Completeness, all data (%)	Redundancy	$R_{\text{merge}}^{\dagger}$, all data (%)
15.00–6.05	1100	18.2	85.6	2.62	4.0
6.05–4.78	1208	17.6	95.0	2.61	4.4
4.78–4.16	1186	18.0	93.8	2.60	5.5
4.16–3.77	1226	16.5	93.5	2.20	6.6
3.77–3.49	1237	16.3	92.2	2.34	6.4
3.49–3.28	1263	14.2	93.1	2.50	7.0
3.28–3.12	1261	10.4	92.6	2.40	8.1
3.12–2.98	1235	9.3	90.2	2.33	12.0
2.98–2.86	1294	7.0	87.3	2.22	14.6
2.86–2.77	1304	5.8	78.9	1.90	20.1
All reflections	12314	13.3	90.2	2.37	8.9

$R_{\text{merge}}^{\dagger} = 100 \sum_{hkl} \sum_i (|I_{hkl,i} - \langle I_{hkl} \rangle|) / \sum_i I_{hkl,i}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with indices hkl and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection.

Table 2
Detailed X-ray diffraction statistics for synchrotron data of crystal form II.

Resolution range (Å)	Number of independent reflections	$I/\sigma(I)$	Completeness, all data (%)	Redundancy	$R_{\text{merge}}^{\dagger}$, all data (%)
15.00–6.15	1095	18.4	74.5	2.47	6.6
6.15–4.89	1202	16.8	81.6	2.47	8.4
4.89–4.27	1180	17.2	82.6	2.39	7.7
4.27–3.88	1220	15.7	84.3	2.42	8.6
3.88–3.60	1230	12.8	86.0	2.38	9.7
3.60–3.39	1255	10.4	86.9	2.34	11.7
3.39–3.22	1255	8.1	87.4	2.35	13.8
3.22–3.08	1231	6.5	88.0	2.30	17.4
3.08–2.96	1291	5.4	88.9	2.30	19.0
2.96–2.86	1290	4.4	90.0	2.29	22.0
All reflections	12249	11.4	85.0	2.37	8.3

$R_{\text{merge}}^{\dagger} = 100 \sum_{hkl} \sum_i (|I_{hkl,i} - \langle I_{hkl} \rangle|) / \sum_i I_{hkl,i}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with indices hkl and $\langle I_{hkl} \rangle$ is the mean intensity of that reflection.

2.2. Crystallization

The oxygenated form of haemoglobin I (LaHb-I) was concentrated to 10 mg ml⁻¹ in 50 mM Tris pH 8.5. Crystallization trials were performed using the hanging-drop vapour-diffusion method (Jancarik & Kim, 1991).

Crystals of *L. anisitsi* haemoglobin have been obtained under several different crystallization conditions using the hanging-drop vapor-diffusion and sparse-matrix methods (Smarra *et al.*, 1997). The best crystals were obtained after 3 d from drops in which 7 µl haemoglobin solution (10 mg ml⁻¹) was mixed with an equal volume of reservoir solution. 0.1 M HEPES pH 7.4, 1.4 M sodium citrate was used as the reservoir solution. The crystals were harvested in the well buffer and mounted in a 1.0 mm boro-

silicate glass capillary tube for X-ray data collection.

2.3. X-ray data collection and processing

Two crystal forms had their X-ray diffraction data sets collected at a wavelength of 1.38 Å using the Synchrotron Radiation Source [Station PCr, Laboratório Nacional de Luz Síncrotron (LNLS), Campinas, Brazil; Polikarpov, Oliva *et al.*, 1998; Polikarpov, Perles *et al.*, 1998] and a 34.5 cm MAR imaging-plate detector (MAR Research). X-ray diffraction data for crystal forms I and II were processed to 2.77 and 2.86 Å, respectively, using the program DENZO and were scaled using the program SCALEPACK (Otwinowski, 1993).

Autoindexing procedures combined with analysis of the X-ray diffraction pattern and averaging of equivalent intensities were used in the characterization of the Laue symmetry for both crystals.

2.4. Molecular replacement

The crystal structure of the oxyhaemoglobin I isolated from *L. anisitsi* was determined by standard molecular-replacement methods using the program AMoRe (Navaza, 1994). The atomic coordinates of the dimer ($\alpha\beta$) of trout haemoglobin deposited in the PDB (Tame *et al.*, 1996; PDB code 1ouu) were used as a search model. All water and phosphate molecules were removed from the search model and the temperature factors for all atoms were set to 20 Å²; however, the haem groups were retained in the model. The atomic coordinates for the search model were translated to set their centre of gravity at the origin and were then rotated so that the principal axes of inertia of the search model lie parallel to orthogonal axes.

Cross-rotation functions were calculated in the resolution range 10–4.5 Å with a sampling step of 2.5° using the program AMoRe (Navaza, 1994). These calculations were carried out with an integration radius of 25 Å. The rotation which generated the highest correlation coefficient (CC) was applied to the search model and used in the

subsequent translation-function computations based on data in the same resolution range.

3. Results and discussion

LaHb-I crystallized in two different space groups under the same crystallization conditions. Crystal form I belongs to the monoclinic C2 Bravais lattice, with unit-cell parameters $a = 57.2$ (1), $b = 62.9$ (1), $c = 183.4$ (2) Å, $\beta = 92.6$ (2)°. Assuming the asymmetric unit content to be one tetramer of molecular weight 65 kDa, the V_m value is 2.5 Å³ Da⁻¹ (Matthews, 1968). Assuming a value of 0.74 cm³ g⁻¹ for the protein partial specific volume, the calculated solvent content in the crystal is 51.5% and the calculated crystal density is 1.175 g cm⁻³. A summary of the X-ray data-collection statistics is given in Table 1.

Crystal form II also belongs to the monoclinic system; however, it has different unit-cell parameters $a = 56.8$ (1), $b = 64.0$ (1), $c = 86.7$ (2) Å, $\beta = 92.2$ °. The observed systematic absences are compatible with space group P2₁. Assuming that one tetramer is present in an asymmetric unit, the V_m value is 2.4 Å³ Da⁻¹. The solvent content in the crystal is 48.8% and the calculated crystal density is 1.184 g cm⁻³. Table 2 shows the results for data collection from crystal form II.

The X-ray diffraction data for crystal form II were used in the molecular replacement. The highest magnitude of the correlation function was obtained for the Euler angles $\alpha = 258.02$, $\beta = 121.82$, $\gamma = 127.00$ ° for the first dimer ($\alpha_1\beta_1$) and angles $\alpha = 78.53$, $\beta = 64.15$, $\gamma = 139.52$ ° for the second dimer ($\alpha_2\beta_2$). The correlation coefficient and the R factor after the translation search were 67.1 and 36.6%, respectively. The fractional coordinates are $T_x = 0.2368$, $T_y = 0.0010$, $T_z = 0.1707$ for the first dimer and $T_x = 0.3460$, $T_y = 0.3549$, $T_z = 0.3191$ for the second dimer. The X-ray diffraction data for crystal form I were also used in the molecular replacement; however, a lower correlation coefficient (52.4%) and a higher R factor (44.3%) were generated after the translation search using the same search model; therefore, X-ray data for crystal form II were used in the preliminary refinement.

Repeated cycles of refinement were carried out using the maximum-likelihood method as implemented in the program REFMAC (Murshudov *et al.*, 1997). The present values of R and R_{free} are 27.2 and 36.9%, respectively, using the sequence of trout haemoglobin (PDB code 1ouu). Amino-acid sequencing for LaHb-I using



the automated Edman technique is under way. Further refinement will be carried out as soon as the sequence become available.

We thank Dr P. Kuser and Dr I. Polikarpov (LNLS) for their help in the synchrotron data collection. The authors also thank Dr Waldener Garutti and Dr Francisco Langeani Neto, who kindly supplied specimens from the Centro de Aquicultura (Caunesp-Unesp). This work was supported by grants from FAPESP, CNPq, CAPES and FUNDUNESP (Brazil).

References

- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Mercedes Silva, M., Rogers, P. H. & Arnone, A. (1992). *J. Biol. Chem.* **276**, 17248–17256.
- Monod, J., Wyman, J. & Ghangeux, J. P. (1965). *J. Mol. Biol.* **12**, 88–118.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst. D53*, 240–255.
- Navaza, J. (1994). *Acta Cryst. A50*, 157–163.
- Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
- Polikarpov, I., Oliva, G., Castellano, E. E., Garratt, R., Arruda, P., Leite, A. & Craievich, A. (1998). *Nucl. Instrum. Methods A*, **405**, 159–164.
- Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. & Craievich, A. (1998). *J. Synchrotron Rad.* **5**, 72–76.
- Seixas, F. A., de Azevedo, W. F. Jr & Colombo, M. F. (1999). *Acta Cryst. D55*, 1914–1916.
- Smarra, A. L. S., Arni, R. K., de Azevedo, W. F. Jr, Colombo, M. F. & Bonilla-Rodriguez, G. O. (1997). *Protein Pept. Lett.* **4**(5), 349–354.
- Smarra, A. L. S., Fadel, V., Dellamano, M., Olivieri, J. R., de Azevedo, W. F. Jr & Bonilla-Rodriguez, G. O. (1999). *Acta Cryst. D55*, 1618–1619.
- Smith, F. R., Lattman, E. E. & Carter, C. W. Jr (1991). *Proteins Struct. Funct. Genet.* **10**, 81–91.
- Tame, J. R., Wilson, J. C. & Weber, R. E. (1996). *J. Mol. Biol.* **259**, 749–760.



**Purification, crystallization, and Patterson search of hemoglobin-IV from the armored
catfish *Liposarcus anisitsi***

Delatorre, P¹., Smarra, A.L.S.¹, Fadel, V^{1,3,4}, Canduri, F¹, Dellamano, M¹., Bonilla-Rodriguez,
G.O.² & de Azevedo Jr., W.F.^{1*}

¹ Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas, UNESP, São José
do Rio Preto, SP 15054-000. ² Departamento de Química e Geociências, Instituto de
Biociências, Letras e Ciências Exatas, UNESP, São José do Rio Preto, SP 15054-000.
³ Departamento de Física. Faciba-FEB. Barretos, SP 14783-226. ⁴ Centro Universitário de
Votuporanga. Votuporanga, SP. 15500-030. Brazil.

*Corresponding author: Prof. Dr. Walter Filgueira de Azevedo Jr. Departamento de Física,
IBILCE-UNESP. CP 136. CEP 15054-000. São José do Rio Preto. SP. Brazil. E-mail:
walterfa@df.ibilce.unesp.br



Synopsis

In this paper is reported the crystallization and preliminary X-ray analysis of hemoglobin-IV from the armored catfish *Liposarcus anisitsi*.

Abstract

Considerable interest is currently focused on fish hemoglobins, in order to identify the structural basis for their diversity of functional behaviors. The armored catfish *Liposarcus anisitsi* presents accessorial air breathing through a modified stomach, which allows this species to survive in waters with high oxygen debt (de Oliveira *et al.*, submitted). The analysis of its hemolysate has shown the presence of four main hemoglobins, with this work focusing on hemoglobin-IV. Hb-IV was crystallized and X-ray diffraction data collected to 2.4 Å resolution using a synchrotron radiation source. Crystals were determined to belong to the space group $P2_12_12_1$ and preliminary structural analysis revealed the presence of one tetramer in the asymmetric unit. The structure was determined using standard molecular replacement technique.

1. Introduction

Fish hemoglobins exhibit a greater variety of functional behaviors when compared to mammalian hemoglobins, in terms of cooperative ligand binding and allosteric control exerted by a variety of chemical factors, such as H^+ , Cl^- , CO_2 , and 2,3-biphosphoglycerate (BPG) (DeYoung *et al.*, 1994). This characteristic reflects a structural diversity guided by evolutionary pressure exerted on the adaptation to differing environmental and physiological needs (Perutz, 1996; DeYoung *et al.*, 1994, Fadel *et al.*, 1999).

Vertebrate hemoglobins bind cooperatively molecular oxygen, the needed electron acceptor at the respiratory chain. Several chemical agents, known as heterotropic ligands, and also temperature can influence hemoglobin oxygen equilibria (Perutz, 1996). For most vertebrate hemoglobins oxygen affinity is reduced as the pH is lowered in the range above 7.0, a property known as the alkaline Bohr effect. This property arises from preferential proton binding to the low O_2 -affinity state (T) and its stabilization (Perutz, 1996; Camardella *et al.*, 1992). An exaggerated response to proton binding can be present on several fish hemoglobins, known as the Root effect, with a presumed role for pumping oxygen into fishes' swimming bladders (Brittain, 1987).

On oxygenation human hemoglobin liberates one proton per tetramer, whereas carp hemoglobin liberates 3.6 per tetramer. Mylvaganan *et al.* (1996), based on the high resolution crystallographic structure of the carbon monoxide bound form of hemoglobin from *Leiostomus xanthurus*, proposed a molecular mechanism for the Root effect. The authors proposed that positively charged clusters across the $\beta 1$ - $\beta 2$ interface in the R-state, identified in the structure, are responsible for part of the Root effect. When pH lowers, protonation of the β N-terminus and His $\beta 147$ (HC3) within these clusters is claimed to



destabilize the R-state and promote the acid-triggered, allosteric R->T transition with concomitant O₂ release. More recently a new fish hemoglobin structure became available (Mazzarella *et al.*, 1999) and its structural analysis led to somewhat contrasting results with the above explanation of the Root effect, suggesting that the differences in the pH dependence can not be related to the modulation of the R-state as had been previously proposed.

Liposarcus anisitsi (Siluriformes, Loricariidae) is a south-american armored catfish able to survive in waters with very low oxygen content, due to its ability to gulp air at the surface. The modified stomach which acts as an acessorial gas-exchange organ permits more efficient blood oxygenation. It has been identified four hemoglobins in its hemolysate with isoelectric points ranging from 7.2 to 8.3 (Smarra *et al.*, 1997). The present work was performed on Hb-IV.

2. Materials and methods

Purification

Blood was collected after an adaptive period of 4 days in aquaria, by caudal puncture. Subsequent procedures were carried out at low temperature (around 4°C). Red blood cells were washed by centrifugation four times with buffered saline (50 mM Tris pH 8.5 containing 1 mM EDTA). Hemolysis was accomplished with buffer A: 50 mM Tris pH 8.5 containing 1 mM EDTA, followed by centrifugation (1,000 g for 1 hr) and filtration for debris removal. An aliquot was kept for phosphate determination. For phosphate removal hemolysate dialysis was performed against buffer A, followed by gel-filtration on a 1.6 x 70 cm column.

Hemoglobin purification was performed on DEAE-Sephadex A-50 using a pH gradient between buffer A and 50 mM bis-Tris pH 6.5. The isolated components were further deionized by several passages through a mixed-bed Amberlite MB-1 (Sigma), concentrated by centrifugation on Amicon microconcentrators and stored in liquid nitrogen until use. Four main hemoglobins were identified in the hemolysate: LaHb-I to IV.

Non denaturing electrophoresis was performed in 10 % polyacrylamide gel and cellulose acetate.

Analytical isoelectric focusing was performed in agarose gels. The bands containing hemoglobin were cut off and eluted in deionized water for pH measurement and isoelectric point estimation.

Crystallization and X-ray data collection

The oxygenated form of hemoglobin IV (LaHb-IV) was concentrated to 10mg/ml in 50mM Tris pH 8.5. Crystallization trials were performed using the hanging drop vapor diffusion method (Jancarik & Kim, 1991).

Crystals of LaHb-IV have been obtained in several different crystallization conditions, using the hanging drop vapor diffusion and sparse matrix methods. The best crystals were obtained after 2 day growth from drops in which 7 μ l of hemoglobin solution (10mg/ml) was mixed with equal volume of reservoir solution. 0.1M Tris pH 8.5, 0.2 M Sodium Acetate, and 30% PEG 4000 were used as the reservoir solution. The crystals were harvested in the well buffer and mounted in 1.0 mm capillary tube of borosilicate glass for X-ray data collection.



X-ray diffraction data was collected from a single LaHb-IV crystal at a temperature of 8°C, at a wavelength of 1.547 Å using the Synchrotron Radiation Source (Station PCr, Laboratório Nacional de Luz Síncrotron, LNLS, Campinas, Brazil) (Polikarpov, Oliva *et al.*, 1998; Polikarpov, Perles *et al.*, 1998) and a 34.5 cm MAR imaging plate detector (MAR Research) with an exposure time of 2 min per image. The crystal to detector distance was 150 mm. Using an oscillation range of 0.8°, 100 images were collected. The raw X-ray diffraction data were processed to 2.4 Å resolution using the program DENZO (Otwinowski, 1993) and scaled with the program SCALEPACK (Otwinowski, 1993). Auto indexing procedures, combined with analysis of the X-ray diffraction pattern and averaging of equivalent intensities was used in the characterization of the Laue symmetry.

Molecular replacement

The crystal structure of the LaHb-IV was determined by standard molecular replacement methods using the program AMoRe (Navaza, 1994). The atomic coordinates of eight different fish hemoglobins and one human hemoglobin deposited in the PDB (Brookhaven Protein Data Bank) (Abola *et al.*, 1997) were used as search models. All solvent molecules were removed from the search models, however the heme groups were kept in the search models. The search models which presented more than one dimer were modified, so that only one dimer ($\alpha\beta$) was used in the molecular replacement. The atomic coordinates for all search models were translated so that their center of gravity is at the origin, they were also rotated so that the principal axes of inertia of the search models are parallel to the orthogonal axes. The PDB accession numbers and identification of the search models are listed in Table 1. Cross-rotation functions were calculated in the resolution



range 10–4.5 Å using a sampling step of 2.5° for all 9 search models. These calculations were carried out with integration radius of 22.5 Å. The rotation which generated the highest correlation coefficient (CC) was applied to the search models and used in the subsequent translation function computations, based on data in the same resolution range.

The best solutions for each search model were selected based on the magnitude of the R_{factor} and correlation coefficient. Translation function for space groups (P222, P2₁2₁2 and P222₁) has also been computed using the same resolution range and the best search model, in order to confirm the space group.

3. Results and discussion

The Figure 1 illustrates photomicrograph of the LaHb-IV crystal suitable for X-ray diffraction experiments, with average dimensions of about 0.3 x 0.4 x 1.0 mm. The crystal has the orthorhombic space group P2₁2₁2₁ with cell parameters of $a = 52.6(1)$ Å, $b = 104.8(2)$ Å and $c = 113.9(2)$ Å. Assuming the asymmetric unit content to be one tetramer of molecular weight of 65 kDa, the V_m value is 2.41 Å³ Da⁻¹ (Matthews, 1968). Assuming a value of 0.74 cm³ g⁻¹ for the protein partial specific volume, the calculated solvent content in the crystal is 49 % and the calculated crystal density 1.18 g cm⁻³. A total of 78791 observations having $I > 1.0\sigma(I)$ of 24618 reflections were collected. The X-ray diffraction statistics are summarized in Table 2.

The results of the molecular replacement using the 9 different search models are listed in Table 3. The correlation coefficients after translation function computation range from 23.2 to 62.7 % and the R_{factor} s range from 39 to 54.5 %. The search model which

presented the best correlation coefficient and R_{factor} was for Trout hemoglobin (Accession number: 1OUU) (Tame *et al*, 1996).

Translation functions for space groups (P222, P222₁, P2₁2₁2) has been computed using the coordinates of model 1OUU as search model and the results are listed in Table 4. The correlation coefficients after translation function computation for the 3 space groups range from 32.1 to 37.9 % and the R_{factor} s range from 49.7 to 51.6 %, which strongly indicates that the correct space group is P2₁2₁2₁.

Figure 2 presents the crystal packing for the partially refined structure of the LaHb-IV. Further refinement using the slow-cooling protocols will be performed using the program X-PLOR (Brünger, 1992). The refined model of the LaHb-IV will be used for detailed comparison with other fish hemoglobins.

Acknowledgments: The authors thank Mr. J. R. Brandão Neto and Dr. I. Polikarpov (LNLS) for their help in the synchrotron data collection. We also want to thank Dr. Francisco Langeani Neto and Waldener Garutti from the Dept. of Zoology of our Institute and representatives of the Centro de Aqüicultura from the State University of São Paulo (CAUNESP-UNESP), for their help concerning specimens and technical advice. This work was supported by grants from FAPESP, CNPq, Fundo Bunka (Banco Sumitomo), and FUNDUNESP (Brazil).

References

- Abola, E. E., Prilusky, J. and Manning, N. O. (1997) Protein Data Bank archives of three-dimensional macromolecular structures. *Methods Enzymol.* **277**, 556-71.
- Brittain, T. (1987). *Comp. Biochem. Physiol.* **86B**, 473-481.
- Brünger, A. T.(1992). X-Plor Version 3.1: A system for crystallography and NMR. Yale University Press, New Haven.
- Camardella, L., Caruso, C., D'Avino, R., di Prisco, G., Rutigliano, B. and Tamburrini, M., Fermi, G., Perutz, M. F. (1992). *J. Mol. Biol.* **224**, 449-460.



- De Young, A., Kwiatkowski, L. D., Noble, R. W. (1994). Fish Hemoglobins. *Methods in Enzymol.* **231**, 124 – 150.
- De Oliveira, C., Taboga, S.R.; Smarra, A. L. S. and Bonilla-Rodriguez, G.O. (1999) Accessorial oxygenation through a modified stomach in the armored catfish *Liposarcus anisitsi* (Siluriformes, Loricariidae): morphological aspects. Submitted to *Comparative Biochemistry and Physiology A*.
- Fadel, V., Honda, R.T., Dellamano, M., Smarra, A. L. S., Delatorre, P., Olivieri, J. R., Bonilla-Rodriguez, G. O., and de Azevedo Jr., W. F. Purification, crystallization, and preliminary X-ray diffraction analysis of carbomonoxy hemoglobin-II from the fish *Piaractus mesopotamicus* (Pacu). (1999). *Acta Cryst. D*. (In Press).
- Jancarik, J. and Kim, S. H. (1991). *J. Appl. Crystallogr.* **24**, 409-411.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491-497.
- Mazzarella, L., D'Aavino, R., di Prisco, G., Savino, C., Vitagliano, L., Moody, P. C., Zagari, A. (1999). *J. Mol. Biol.* **287**, 897 – 906.
- Mylvaganam, S. E., Bonaventura, C., Bonaventura, J., Getzoff, E. D. (1996). *Nat. Struct. Biol.* **3**, 275 –283.
- Navaza, J. (1994). *Acta Cryst.* **A50**, 157-163.
- Otwinowski, Z. (1993). In *Proceedings of the CCP4 Study Weekend*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56-62. Warrington: Daresbury Laboratory.
- Perutz, M. F. (1996).. *Nat. Struct. Biol.* **3**, 211- 212.
- Polikarpov, I., Oliva, G., Castellano, E.E., Garratt, R., Arruda, P., Leite, A. & Craievich, A. (1998). *Nucl. Instrum. Methods A*, **405**, 159-164.
- Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E.E., Garratt, R. & Craievich, A. (1998). *J. Synchrotron Rad.* **5**, 72-76.



Smarra, A. L. S., Arni, R. K., de Azevedo, Jr., W. F., Colombo, M., & Bonilla-Rodriguez, G. O. (1997). *Protein and Peptide Letters*. **4**(5) 349-354.

Smarra, A. L. S., Fadel, V., Dellamano, M., Olivieri, J.R., de Azevedo, Jr., W. F. & Bonilla-Rodriguez, G. O. Crystallization, preliminary X-ray diffraction analysis and Patterson search of oxyhaemoglobin I from the wolf (*Chrysocyon brachiurus*). (1999). *Acta Cryst. D*, **55** (in press).

Tame, J.R., Wilson, J.C., and Weber, R.E. (1996). *J. Mol. Biol.* **259**, 749-760.



Table 1. List of the hemoglobins used as search models

PDB Accession number	Protein
1HHO	Oxy hemoglobin from <i>Homo sapiens</i>
1T1N	Carbonmonoxy Hemoglobin from fish (<i>Trematomus akajei</i>)
1OUU	Carbonmonoxy Hemoglobin from fish (<i>Oncorhyncus mykiss</i>)
1CG8	Carbonmonoxy Hemoglobin from fish (<i>Dasyatis akajei</i>)
1HBH	Deoxy Hemoglobin from fish (<i>Pagothenia bernacchii</i>)
1OUT	Deoxy Hemoglobin from fish (<i>Oncorhyncus mykiss</i>)
1PBX	Carbonmonoxy Hemoglobin from fish (<i>Pagothenia bernacchii</i>)
1SPG	Carbonmonoxy Hemoglobin from fish (<i>Leiostomus xanthurus</i>)
1CG5	Deoxy Hemoglobin from fish (<i>Dasyatis akajei</i>)

Table 2. Detailed X-ray diffraction statistics for synchrotron data of crystal of LaHb-IV.

Resolution Range		Number of independent reflections	I/σ(I)	R _{merge} [*] (%)	Completeness (%)
20.00	5.87	1581	19.92	5.8	86.3
5.87	4.68	1665	20.01	6.5	95.9
4.68	4.09	1670	20.49	7.2	97.4
4.09	3.72	1652	19.34	7.8	97.8
3.72	3.46	1662	17.24	8.5	98.5
3.46	3.25	1657	15.11	9.6	98.7
3.25	3.09	1661	12.45	11.2	99.6
3.09	2.96	1666	10.59	12.5	98.7
2.96	2.84	1639	9.59	13.5	98.0
2.84	2.75	1629	8.42	15.1	98.6
2.75	2.66	1646	6.49	17.8	99.0
2.66	2.58	1620	6.10	18.9	97.1
2.58	2.52	1642	5.33	21.3	99.8
2.52	2.46	1606	5.12	22.7	96.3
2.46	2.40	1622	4.56	24.2	99.6
All Reflections		24618	15.17	8.8	97.3

^{*}R_{merge} = 100 × Σ_{hkl} [Σ_i(| I_{hkl,i} - <I_{hkl}>|)]/Σ_{hkl,i} (I_{hkl}) , where I_{hkl,i} is the intensity of an individual measurement of the reflection with indices, h, k and l, and <I_{hkl}> is the mean intensity of that reflection.

Table 3. Eulerian angles and fractional coordinates after translation function computation.

Protein	Dimer	α (°)	β (°)	γ (°)	x	y	z	CC (%)	R_{factor}^* (%)
1HHO	$\alpha 1\beta 1$	85.05	92.50	309.09	0.4756	0.0743	0.0573	29.3	54.0
	$\alpha 2\beta 2$	24.30	61.13	265.74	0.1988	0.1696	0.0169	29.3	54.0
1T1N	$\alpha 1\beta 1$	76.61	62.38	32.05	0.3706	0.1994	0.2186	24.9	54.5
	$\alpha 2\beta 2$	102.54	117.45	207.46	0.3271	0.7166	0.4570	24.9	54.5
1OUU	$\alpha 1\beta 1$	97.53	59.75	63.02	0.3072	0.0129	0.4158	62.7	39.0
	$\alpha 2\beta 2$	79.25	58.31	24.71	0.8661	0.6949	0.7090	62.7	39.0
1CG8	$\alpha 1\beta 1$	87.15	83.84	47.07	0.0354	0.4196	0.3626	24.3	53.9
	$\alpha 2\beta 2$	41.22	114.49	270.54	0.2430	0.1286	0.6039	24.3	53.9
1HBH	$\alpha 1\beta 1$	75.02	58.29	26.37	0.3734	0.1950	0.2133	31.5	51.6
	$\alpha 2\beta 2$	77.70	58.34	26.45	0.3714	0.1964	0.2121	31.5	51.6
1OUT	$\alpha 1\beta 1$	79.39	121.16	334.18	0.1326	0.3039	0.2081	42.7	47.0
	$\alpha 2\beta 2$	84.70	64.53	295.50	0.3202	0.4882	0.5812	42.7	47.0
1PBX	$\alpha 1\beta 1$	104.74	121.58	27.26	0.1298	0.1965	0.2899	37.7	49.9
	$\alpha 2\beta 2$	78.15	61.77	29.20	0.8608	0.6946	0.2131	37.7	49.9
1SPG	$\alpha 1\beta 1$	80.35	119.65	67.20	0.0613	0.0384	0.1216	24.3	52.7
	$\alpha 2\beta 2$	24.71	115.30	339.60	0.8311	0.1327	0.8739	24.3	52.7
1CG5	$\alpha 1\beta 1$	73.24	62.01	30.85	0.4680	0.2456	0.2637	23.2	53.4
	$\alpha 2\beta 2$	110.85	117.22	211.13	0.4393	0.7906	0.1647	23.2	53.4

* $R_{factor} = 100 \times \sum |F_{obs} - F_{calc}| / \sum (F_{obs})$, the sums being taken over all reflections with $F / \sigma(F) > 2$ cutoff.



Table 4. Fractional coordinates after translation function computation for other possible space groups

Space Group	Dimer	X	y	z	CC (%)	R _{factor} [*] (%)
P222	α1β1	0.3739	0.1420	0.4332	33.0	52.6
	α2β2	0.3911	0.1471	0.4328	33.0	52.6
P222 ₁	α1β1	0.0673	0.0119	0.3542	37.9	49.7
	α2β2	0.3845	0.6936	0.8484	37.9	49.7
P2 ₁ 2 ₁ 2	α1β1	0.0823	0.3436	0.0926	32.1	51.6
	α2β2	0.1136	0.8662	0.9638	32.1	51.6

^{*} $R_{factor} = 100 \times \sum |F_{obs} - F_{calc}| / \sum (F_{obs})$, the sums being taken over all reflections

with $F / \sigma(F) > 2$ cutoff.

Anexo III

O programa AMoRe

O programa AMoRe procura fazer uma junção entre automação, velocidade e performance, usando para isso a análise matemática da estrutura. O sucesso deste programa, na maioria dos casos, deve-se a sua capacidade de explorar muitas soluções em potencial.

O procedimento inicial para execução do programa é obter as seguintes informações:

- Parâmetros de cela
- Resolução
- Modelo a ser testado
- Número de moléculas esperadas na unidade assimétrica
- Faixa de resolução na qual a procura será feita.



A rotina *Tabling* calcula os coeficientes contínuos de Fourier correspondente ao modelo. Estes coeficientes são obtidos através da transformada de Fourier da densidade eletrônica, com base nas coordenadas atômicas da molécula, que são colocadas em uma cela unitária com dimensões lineares aproximadamente quatro vezes maior do que o tamanho do modelo. Para evitar o uso de uma cela unitária muito grande, o que levaria a um posicionamento desfavorável da molécula, o centro de massa dos átomos é deslocado para a origem da cela unitária e as coordenadas são rotacionadas de tal maneira que o eixo principal de inércia é alinhado com os eixos da cela unitária.

A rotina *Roting* é usada para o cálculo da função rotação. O raio de integração é usualmente a distância entre o centro de massa e o átomo mais afastado. Os fatores de estrutura do modelo são calculados por interpolação. A função rotação é calculada, usando expansões em harmônicos esféricos, mas as variáveis radiais são manipuladas diferentemente, conduzindo a resultados mais apurados. O coeficiente de correlação é usado como critério de seleção.

A rotina *Traing* calcula a função translação (T_2), usando transformadas rápidas de Fourier. Os valores de T_2 são usados somente com o objetivo de selecionar os melhores picos, usando para isso o R_{factor} :



$$R_{factor} = \sum_H m_H \left| |F_H^{Obs}| - |F_H(\alpha, \beta, \gamma, x, y, z)| \right| x \left(\sum_H m_H |F_H^{Obs}| \right)^{-1}$$

e o coeficiente de correlação:

$$Cc = \sum_H m_H \Delta |F_H^{obs}| \Delta |F_H(\alpha, \beta, \gamma, x, y, z)| \times \left\{ \left[\sum_H m_H (\Delta |F_H^{obs}|) \right] \times \left[\sum_H m_H [\Delta |F_H(\alpha, \beta, \gamma, x, y, z)|]^2 \right] \right\}^{-1/2}$$

onde m_H é a multiplicidade de reflexões e $\Delta |F_H|$ corresponde aos maiores picos de T.

Anexo IV

Alinhamento via Blast II da sequência da hemoglobina do *Oncorhyncus mykiss* (10UU).

Cadeia Alfa

tmpseq_1	1	SLTAKDKSVVKAFWKGISGKADVVGAEALGRDKMLTAYPQTKTYFSHWA-DLSPGSGPVK	59
1942654	2	--.....-	58
1942654	2	--.....-	58
1942654	2	--.....-	58
1942034	2	..S.T..AR...L.D..E..SAEL.....--..VSF....I...E.GQ..G.QTPQ.R	59
1065039	2	..SD...AA.R.L.S..GKS..AI.ND..S.--.IVV.....P-.VT...PHI.	58
1065039	2	..SD...AA.R.L.S..GKS..AI.ND..S.--.IVV.....P-.VT...PHI.	58
1065039	2	..SD...AA.R.L.S..GKS..AI.ND..S.--.IVV.....P-.VT...PHI.	58
4930138	2	..SD...AA.R.L.S..GKSS.AI.ND..S.--.IVV.....I....P-.VT...PNI.	58
576142	2	..S.A..GN...A...VG.H.AEY.....--ER.FLSF.T.....P.F--...H..AQ..	56
576142	2	..S.A..GN...A...VG.H.AEY.....--ER.FLSF.T.....P.F--...H..AQ..	56
tmpseq_1	60	KHGGIIMGAIGKAVGLMDDLVGGMSSALSDLHAFKLRVDPGNFKILSHNVLTLAIHFPSD	119
1942654	59		118
1942654	59		118
1942654	59		118
1942034	60	N..AV..A.V...KSI.N...L.Q..E.....A....A...ILVISM.Y..G.	119
1065039	59	A..KKV..G.AL..SKI...KT.LME..EQ..Y.....A....N.C...VISTM..KE	118
1065039	59	A..KKV..G.AL..SKI...KT.LME..EQ..Y.....A....N.C...VISTM..KE	118
1065039	59	A..KKV..G.AL..SKI...KT.LME..EQ..Y.....A....N.C...VISTM..KE	118
4930138	59	A..KKV..G.AL..SKI...KT.LME..EQ..Y.....S....N.C...VISTM..KE	118
576142	57	G..AKVAA.LT...EHL...P.AL.E.....H.....V...L...SL.....S.L...	116
576142	57	G..AKVAA.LT...EHL...P.AL.E.....H.....V...L...SL.....S.L...	116
tmpseq_1	120	FTPEVHIAVDKFLAAVSAALADKYR	144
1942654	119		143
1942654	119		143
1942654	119		143
1942034	120	LS.....CLAL..SE...	144
1065039	119	...A.VSL....SG.AL...ER..	143
1065039	119	...A.VSL....SG.AL...ER..	143
1065039	119	...A.VSL....SG.AL...ER..	143
4930138	119	...A.VSL....SG.AL...ER..	143
576142	117	...A..ASL.....N..TV.TS...	141
576142	117	...A..ASL.....N..TV.TS...	141



Cadeia Beta

QUERY	1	VEWTDAEKSTISAVWGKVNIDEIGPLALARVLIVYPWTQRYFGSFGNVSTPAAIMGNPKV	60
P02142	1		60
P80727	1	...SAS.R...TSL...I.VA...Q.....K..DL.NA...Q..A..	60
P02140	1	R.A.IGL...L.P..L..Q....C.....AT...L.S.....	60
P02139	1	R.A.I.L...L.P..L..E....C.....F.A.Y..L.S.....	60
P56251	1	.D....RAA.K.L...IDVG....Q..S.L.....H.KG...I..N...L..A..	60
P02141	1	.D....R.A.VGL...ISV....Q....L...S.....H.ST...L.....A.	60
Q91473	2	.D....R.A.VGL...ISV....Q....L...S.....H.ST...L.....A.	61
P14524	1	.Q..AE..QL.TGL.....VA.C.GE....L.....F.A....L.S.T.VL.....	60
P07406	1	.Q..AE..QL.TGL.....VA.C.GE....L.....F.A....L.S.T.VL.....	60
P02126	1	.Q..AE..QL.TGL.....VA.C.AE....L.....F.A....L.S.T.VL.....	60
QUERY	61	AAHGKVVCGALDKAVKNMGNILATYKSLSETHANKLFVDPDNFRVLADVLTIVIAAKFGA	120
P02142	61		120
P80727	61	L...E.....DDVKG..SK..QL.NE..N.....L.G.C....L.T.L..	120
P02140	61	RT.M.G.ER.I...D..K...AP..VM.SE..H.....L...CI.VCA.M...P	120
P02139	61	RT.E.G.MR.I.D.D..K...AP..VM.SE..H.....L...CI.VCA.M...P	120
P56251	61	.E...T.M.G..R..Q..D..KNV..Q..IK.SE.IH.....L.GEII.MCVG....P	120
P02141	61	.K...T.MHG..R..Q.LDD.KN..AT..VM.SE..H.....L...CI.VCV...L..	121
Q91473	62	.K...T.MHG..R..Q.LDD.KNA.TA..VM.SE..H.....L...CI.VCV...L.P	121
P14524	61	Q....K.LTSFGE....LDS.KG.FAQ...L.CD..H...E...L.G.I.VV.L..H..K	120
P07406	61	Q....K.LTSFGE....LDS.KN.FSQ...L.CD..H...E...L.G.I.VV.L..H..K	120
P02126	61	Q....K.LTSFGD....LDS.KN.FSQ...L.CD..H...E...L.G.I.VV.L..H..K	120
QUERY	121	SFTPEIQATWQKFMKVVAAMGSRYF	146
P02142	121		146
P80727	121	G.PA....V....VA...S.LSKQ..	146
P02140	121	..NADV.EA....LS...S.LCRQ.	146
P02139	121	..S.NV.EA....LS...N.LKRQ.	146
P56251	121	HEA....LA...S.L.RQ.	146
P02141	122	V.SADT.EAF...LA...S.L.RQ.	146
Q91473	122	T.SAD..EAF...LA...S.L.RQ.	147
P14524	121	D...DC..A...LVR..AH.LARK.	145
P07406	121	D...DC..A...LVR..AH.LARK.	145
P02126	121	D...DC..A...LVR..AH.LARK.	145



Referências Bibliográficas

- Azevedo, W. F. Jr. (1997) Introdução à biocristalografia com o estudo estrutural da quinase dependente de ciclina 2 (CDK 2) complexada com inibidores. Tese de doutorado – Instituto de física – USP - São Carlos.
- Blundell, T.L., Johnson, L.N. Protein Crystallography, Academic Press, London (1976).
- Brünger, A. T.(1992). X-Plor Version 3.1: A system for crystallography and NMR. Yale University Press, New Haven.
- Camardella, L., Caruso, C., D'Avino, R., di Prisco, G., Rutigliano, B. and Tamburrini, M., Fermi, G., Perutz, M. F. (1992). *J. Mol. Biol.* **224**, 449-460.
- Castellano, E., Oliva, G., Navaza, J. (1992) *J. Appl. Cryst.* **25**, 281-284.
- Drenth, J. Principles of protein X-ray crystallography, Editora Springer-verlag, Boston (1994).
- Fago, A., Bendixen, E., Malte, H., Weber, R. E. The anodic hemoglobin of *Anguilla anguilla*. Molecular basis for allosteric effects in a root-effect hemoglobin. *J Biol Chem.* 1997 Jun 20;272(25):15628-35.
- Ito, N., Komiyama, N.H., and Fermi, G. (1995) *J. Mol. Biol.* **250**, 648-658.
- Jancarik, J. and Kim, S. H. (1991). *J. Appl. Crystallogr.* **24**, 409-411.



Lehninger, R. H., Nelson, D.L., Cox, M.M. Principles of Biochemistry. 2 ed. New York: Worth, 1995.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491-497.

Monod, J., Wyman, J.P., Changeaux, J., On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.*, **12**, 88-118, 1965.

Murray, R. K., Granner, D.K., Mayes, P. A., Rodwell, V. W. Harper's De Young, A., Kwiatkowski, L. D., Noble, R. W.(1994). Fish Hemoglobins. *Methods in Enzymol.* **231**, 124 – 150.

Navaza, J. (1994). *Acta Cryst.* **A50**, 157-163.

Otwinowski, Z. (1993). In *Proceedings of the CCP4 Study Weekend*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56-62. Warrington: Daresbury Laboratory.

Root, R. W. (1931) The respiratory function of the blood of marine fishes. *Biol. Bull.* **61**, 427-465

Stryer, L. Bioquímica. 4 ed. Rio de Janeiro. Guanabara Koogan, 1995.

Biochemistry. 23 ed. East Norwalk, Appleton & Lange, 1988.

Stout, G. H., Jensen, L. H. X-Ray Structure Determination – A practical Guide. 2 ed. John Wiley & Sons, 1989.



- Tame, J.R., Wilson, J.C., and Weber, R.E. (1996). *J. Mol. Biol.* **259**, 749-760.
- Tutosova, T. A., Maden, T. L. (1999) Blast 2 sequences – a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett.* **174**, 147-250.



