

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

**“DETERMINAÇÃO DA ESTRUTURA CRISTALOGRÁFICA POR DIFRAÇÃO DE RAIOS-X DA
ENZIMA GLICOSE 6-FOSFATO ISOMERASE HUMANA”**

ARTUR TORRES CORDEIRO

**Dissertação apresentada ao Instituto de Física de São Carlos, Universidade de São
Paulo, para obtenção do título de Mestre em Ciências “Física Aplicada-Opção Física
Biomolecular”**

**Orientador:
Prof. Dr. Otávio Henrique Thiemann**

**Comissão Julgadora:
Prof. Dr. Otávio Henrique Thiemann (IFSC-USP)
Prof. Dr. Luis Fernando Delboni (PUC-PC)
Dr. Igor Polikarpov (LNLS)**

OK
USP/IFSC/SBI



8-2-001370

**São Carlos - São Paulo
2001**

IFSC-USP CENTRO DE BIBLIOTECA E
INFORMAÇÃO

Cordeiro, Artur Torres

"Determinação da estrutura cristalográfica por difração de raios-x da enzima glicose 6-fosfato isomerase humana"/ Artur Torres Cordeiro – São Carlos, 2001

Dissertação (Mestrado) – Instituto de Física de São Carlos da Universidade de São Paulo, 2001
páginas: 75

Área: Física Aplicada-Opção Física Biomolecular
Prof. Dr. Otávio Henrique Thiemann

1. Estrutura Cristalográfica, 2. Estudo da enzima

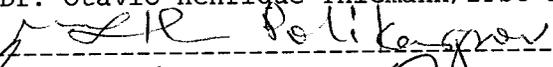
I. Título

MEMBROS DA COMISSÃO JULGADORA DA DISSERTAÇÃO DE Mestrado DE ARTUR
TORRES CORDEIRO APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO CARLOS, DA
UNIVERSIDADE DE SÃO PAULO, EM 09 DE MARÇO DE 2001

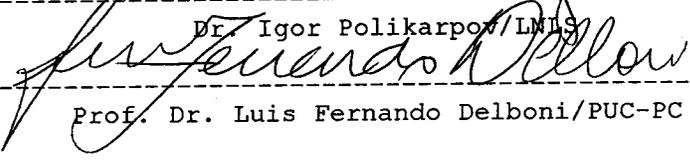
COMISSÃO JULGADORA:



Prof. Dr. Otávio Henrique Thiemann/IFSC-FFI-USP



Dr. Igor Polikarpov/LMIS



Prof. Dr. Luis Fernando Delboni/PUC-PC

DEDICATÓRIA

Aos meus pais pelo incentivo,
orientação, e exemplo de vida.

AGRADECIMENTOS

- Ao prof. Otavio H. Thiemann, meu orientador, pela incansável e valiosa orientação.
- Ao prof. Luis F. Delboni, meu ex-orientador, por acreditar na minha capacidade.
- A Leticia por sua paciência, carinho e compreensão.
- Aos prof. Glaucius Oliva e Richard Garratt pelas discussões enriquecedoras com relação a cristalografia.
- Ao Paulo H. C. Godoi pelos ensinamentos e ajuda durante as coletas de dados no RAXIS e no LNLS.
- Ao prof. Igor Poolikarpovi, seus alunos Lee e Brandão pela ajuda e tempo cedido para a coleta de dados no LNLS
- A todos os colegas de pós-graduação pelos bons momentos que compartilhamos.
- A Wanda e Bianca pela ajuda nos Laboratórios de bioquímica e biologia molecular.
- A Wladerez, Jaciara e Maria Helena pela ajuda com os assuntos acadêmicos e burocráticos.
- Ao amigos Cristiano, Alexandre e Magno por me receberem em sua casa e por todos os bons momentos que compartilhamos diariamente.
- A FAPESP pelo apoio financeiro.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

RESUMO

Capítulo 1	Sobre a enzima PGI, seu estudo por cristalografia e difração	1
Capítulo 2	Sub-clonagem do gene da PGI-hum	6
Capítulo 3	Expressão e purificação da PGI-hum	15
	3.1. Expressão da PGI-hum recombinante	15
	3.2. Purificação da PGI-hum	18
Capítulo 4	Determinação de parâmetros cinéticos da PGI-hum	23
Capítulo 5	Cristalização da PGI-hum	30
Capítulo 6	Coleta de dados de difração de raios-X dos cristais de PGI-hum	33
Capítulo 7	Análise qualitativa dos conjuntos de dados de difração	42
Capítulo 8	Substituição molecular	47
Capítulo 9	Refinamento e análise estrutural	55
Capítulo 10	Conclusões gerais e perspectivas futuras	67
Anexo		68
Referências bibliográficas		73

LISTA DE FIGURAS

Figura 1: Metabolismo da glicólise, indicando-se nas reações de	2
Figura 2: Via de formação da pentoses mostrando as enzimas	3
Figura 3: Representação esquemática dos primeiros ciclos da	7
Figura 4: Representação esquemática das etapas necessárias	9
Figura 5: Representação esquemática da transformação de	10
Figura 6: Mapa dos primer obtidos da <i>Life technologies</i>	11
Figura 7: Gel de agarose 1% com brometo de etídeo	14
Figura 8: Representação esquemática de um operon	15
Figura 9: representação esquemática de uma <i>E. coli</i>	16
Figura 10: Gel SDS-PAGE a 15% de poliacrilamida	18
Figura 11: Perfil cromatográfico da troca catiônica (A),	22
Figura 12: Gel SDS-PAGE 15% de poliacrilamida mostrando	22
Figura 13: Gráfico de velocidade por concentração de	24
Figura 14: Gráfico de Lineweaver-Burk para uma enzima	25
Figura 15 : Representação esquemática do sistema de reações	26
Figura 16: Inibidores testados contra a PGI-hum	27
Figura 17: Gráfico da relação entre [F6P] e a velocidade de	28
Figura18: Composição química do intermediário de alta	29
Figura 19: Caixinha de cristalização resumindo as melhores	32
Figura 20: cristal da PGI-humana crescido	32
Figura 21: Diagrama de Argand. O vetor A representa uma	34
Figura 22: (A) feixe espalhado , representado por s , pode ser	34
Figura 23: Diagrama de Argand para oito vetors da direção	36
Figura 24: (A) diferença de fase entre feixes refletidos por	37
Figura 25: Construção de Ewald mostrando pontos da rede	38
Figura 26: Construção de Ewald ilustrando o método de oscilação.	39
Figura 27: Representação de uma célula unitária (a esquerda)	42
Figura 28: Representação das células: primitiva (P), face centrada	43
Figura 29: (a) representa uma célula unitária bi-dimensional	49
Figura 30: Alinhamento das sequências de PGI-coelho (rab PGI)	50
Figura 31: visualização da unidade assimétrica de PGI-humana	52

Figura 32: A esquerda está representada a seção $\chi=180.^{\circ}$ da projeção	53
Figura 33: Mapa de patterson visualizado pelo programa Xtalview	53
Figura 34: Representação da célula unitária de PGI-hum com	54
Figura 35: Gráfico mostrando a variação dos valore de R	60
Figura 36: Representação do traçado de carbonos- α da PGI-hum	61
Figura 37: gráfico de ramachandram para todos os resíduos do	61
Figura 38: Mapa de densidade eletrônica $2\sigma_{AF_0} - D.F_c$	62
Figura 39: Representação esquemática da estrutura secundária	63
Figura 40: Representação esquemática do mecanismo de reação	63
Figura 41: Visão geral do dímero de PGI-hum, mostrando	64
Figura 42: Representação planar obtida do programa <i>LIGPLOT</i>	65
Figura 43: Mapa de densidade eletrônica $2\sigma_{AF_0} - D.F_c$	66

LISTA DE TABELAS

Tabela I: síntese do programa de PCR executado para amplificação	11
Tabela II: Reagentes utilizados para cada medida de atividade	26
Tabela III: variáveis definidas durante os ensaios de eficiência	27
Tabela IV: Porcentagem de atividade restante devido a presença	28
Tabela V: resumo da análise qualitativa para os conjuntos de dados	45

LISTA DE ABREVIATURAS

5 PAA	D-arabinoato-5-fosfato
5 PAAm	D-arabinoamida-5-fosfato
5 PAH	D-arabinohidroximato-5-fosfato
5 PAHz	D-arabinohidroxiazida-5-fosfato
6PGH	D-gluconato-6-fosfato
cDNA	DNA codificante
C _f	Concentração final
D.O.	densidade ótica
dNTP,s	deoxinucleotídeos trifosfato
<i>E. coli</i>	<i>Escherichia coli</i>
F6P	frutose-6-fosfato
FM	fator de maturação
FMA	fator de motilidade autócrina
FPLC	<i>Fast flow liquid chromatography</i>
G6P	glicose-6-fosfato
G6PDH	glicose-6-fosfato desidrogenase
IFSC-USP	Instituto de Física de São Carlos - Universidade de São Paulo.
IPTG	isopropil- β -D-tiogalactosídeo
LNLS	Laboratório Nacional de Luz Síncrotron
MPD	metil pentano díol
mRNA	RNA mensageiro
NADP	Nicotinamina Dinucleotídeo Adenina fosfato
NKL	neuroleucinas
PCR	Reação em Cadeia da Polimerase (<i>Polimerase Chain Reaction</i>)
PDB	Protein Data Bank
PEG	Poli etileno glicol
PGI	glicose 6-fosfato isomerase
PGI-hum	glicose 6-fosfato isomerase humana
pI	ponto isoelétrico
SDS- PAGE	<i>sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>

T.C. Tampão de Corrida (BICINA 50 mM, pH 7,8)
TAE Tampão 40 mM TRIS acetato , 2 mM EDTA e pH 8,5

RESUMO

O trabalho realizado como parte do programa de mestrado em física aplicada sub-área biomolecular, teve como objeto de estudo a enzima glicose-6-fosfato isomerase de humanos (PGI-hum). Este trabalho envolveu principalmente três áreas de estudos: biologia molecular, bioquímica e cristalografia.

A parte de biologia molecular refere-se a sub-clonagem do gene da PGI-hum a partir de uma biblioteca de cDNA de cérebro de feto humano - capítulo 2 - e a expressão deste gene em bactérias *Escherichia coli* - capítulo 3.

A parte de bioquímica envolve a purificação e caracterização de parâmetros cinéticos da enzima PGI-hum recombinante. Além dos parâmetros cinéticos, foram realizados ensaios de eficiência inibitória para quatro compostos similares ao substrado, cedidos em colaboração com o pesquisador Dr. Laurent Salmon (Lab. de Química Biorgânica e Bioinorgânica de Universidade de Paris-XI). Esta etapa encontra-se descrita nos capítulos 3 e 4.

Uma vez definido o protocolo de purificação e confirmada a atividade enzimática para a PGI-hum recombinante, iniciou-se a terceira etapa do projeto: cristalização e determinação da estrutura por difração de raios-X. O primeiro passo, nesta última fase do trabalho, foi determinar as condições de cristalização da PGI-hum - capítulo 5. Depois de obtidos os cristais, foram coletados dois conjuntos de dados de difração de raios-X, sendo o primeiro coletado em uma fonte convencional de raios-X e o segundo com um feixe proveniente de luz síncrotron - capítulo 6. A análise da qualidade e comparação dos conjuntos de dados - capítulo 7 - indicou qual conjunto seria utilizado nas etapas seguintes da determinação da estrutura da PGI-hum.

Para resolução da estrutura cristalográfica da PGI-hum foi utilizado o método de substituição molecular com base na estrutura homóloga da PGI de coelho - capítulo 8. O refinamento estrutural da PGI resultou em uma estrutura com 2.1Å de resolução e fatores R e R_{free} satisfatórios - capítulo 9. Uma análise estrutural preliminar é apresentada indicando uma geometria adequada para a proteína e descrevendo as principais características estruturais desta enzima em comparação com a PGI de coelho.

ABSTRACT

This work is presented as part of the Master degree requirements of the Applied Physics program, Biomolecular Physics area. The purpose of this work is the structural study by of the human glucose-6-phosphate isomerase (PGI-hum). This work has involved mainly three areas: Molecular Biology, Biochemistry and Crystallography.

The Molecular Biology work was intended for the cloning of the human open reading frame of PGI-hum from a fetal human brain cDNA library – Chapter 2 – and its expression in *Escherichia coli* – Chapter 3.

The biochemistry work has involved the PGI-hum purification and the determination of its kinetic parameters of the recombinant protein. Inhibitory efficiency measurements were made with four compounds kindly provided by Dr. Laurent Salmon (Laboratoire de Chimie Bioorganique et Bioinorganique - Université Paris-XI – France). This work is described in Chapters 3 and 4.

Once defined the expression and purification protocols and confirmed its enzymatic activity for the recombinant PGI-hum, a third phase was initiated in the project: The crystallization and structure determination by X-ray diffraction. The first step in this last phase of the project was determining the conditions for crystallization of the PGI-hum – Chapter 5. Once obtained the crystals, two data sets were collected. One data-set was collected “in house” X-ray source and a second data set was collected at the Synchrotron beam line (Laboratório Nacional de Luz Sincrotron – LNLS – Campinas) – Chapter 6. The analysis of the quality and the comparison of the two data sets, presented in Chapter 7, indicated that second data set was the best to be used at the following steps .

For the resolution of the atomic structure of the PGI-hum the method of molecular substitution based on the structure of the rabbit homologue enzyme was employed – Chapter 8. The refinement of the PGI-hum structure at 2.1 Å resolution and satisfactory R and R_{free} factors is presented in Chapter 9 of this dissertation. A preliminary structural analysis is presented indicating an adequate geometry of the protein and describing the most important structural features of the PGI-hum compared to its homologue, the rabbit PGI.

CAPÍTULO 1

Sobre a enzima PGI, seu estudo por cristalografia e difração de raios-X.

Metabolismo

O metabolismo celular de todos os seres vivos utiliza praticamente as mesmas enzimas distribuídas em algumas seqüências de reações bem conhecidas denominadas vias metabólicas. Essas vias metabólicas são responsáveis pela manutenção da homeostase do organismo. A via glicolítica, por exemplo, é responsável pela transformação de glicose em compostos mais simples e a concomitante formação de moléculas armazenadoras de energia, conhecidas como ATP (adenosina-trifosfato). Tais moléculas funcionam como verdadeiras “moedas de energia” e são consumidas por enzimas em outras vias metabólicas e demais processos fisiológicos, como por exemplo a contração muscular. A via glicolítica pode ser dividida em duas etapas (fig.1). A primeira, conhecida como via de Embden-Meyerhof, é a fase anaeróbica na qual a glicose é capturada do meio extra-celular e degradada até piruvato gerando como resultado líquido apenas duas moléculas de ATP; o piruvato (ou ácido pirúvico) pode ser convertido em ácido láctico ou a alguns derivados alcóolicos. A segunda fase da glicólise ocorre nas mitocôndrias, depende de oxigênio (processo aeróbico) e é responsável pela produção de 36 moléculas de ATP (por molécula de glicose).

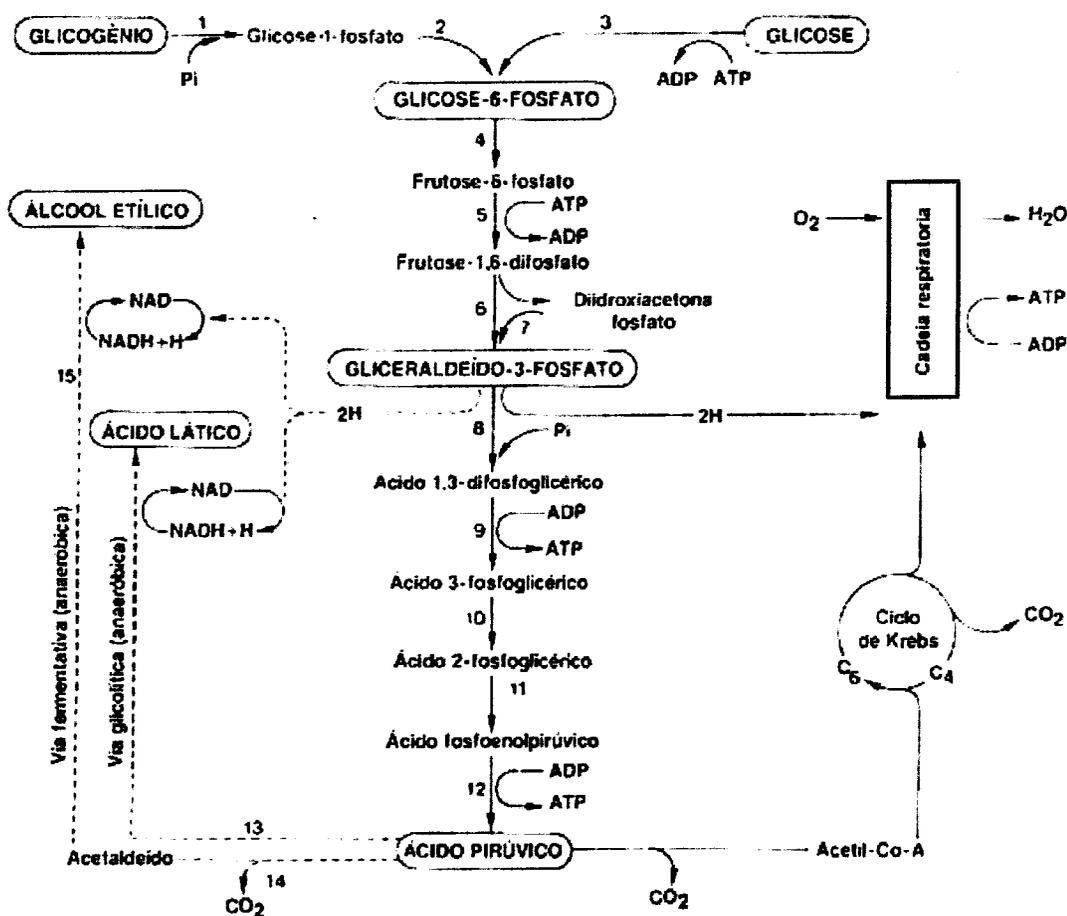


Figura 1: Metabolismo da glicólise, indicando-se nas reações de 1 a 12 a via de Embden-Meyerhof e suas relações com: a glicólise, seguindo pelo ciclo de Krebs e cadeia transportadora de elétrons, e com a fermentação (metabolismo anaeróbico) onde o ácido pirúvico é convertido a ácido láctico ou a álcool etílico. As enzimas de que dependem as reações numeradas são: (1) fosforilase, (2) fosfoglicomutase, (3) hexokinase, (4) *glucose-6-fosfato isomerase*, (5) fosfofrutokinase, (6) aldolase, (7) triosefosfato-isomerase, (8) desidrogenase da triosefosfato, (9) fosfoglicerokinase, (10) fosfogliceromutase, (11) enolase, (12) piruvatokinase, (13) desidrogenase láctica, (14) carboxilase pirúvica e (15) desidrogenase alcoólica.

Outra via crucial no metabolismo é a via das pentoses (fig2). Esta é responsável pela produção da ribose-5-fosfato presente nos ácidos nucleicos e concomitante produção de NADPH₂, coenzima necessária como agente redutor em muitas reações metabólicas. A via tem como ponto de partida a glicose-6-fosfato e produz, por descarboxilação, as pentoses, açúcares de 5 carbonos. O ciclo retorna a via Embden-Meyerhof pela produção de frutose-6-fosfato e gliceraldeído-3-fosfato.

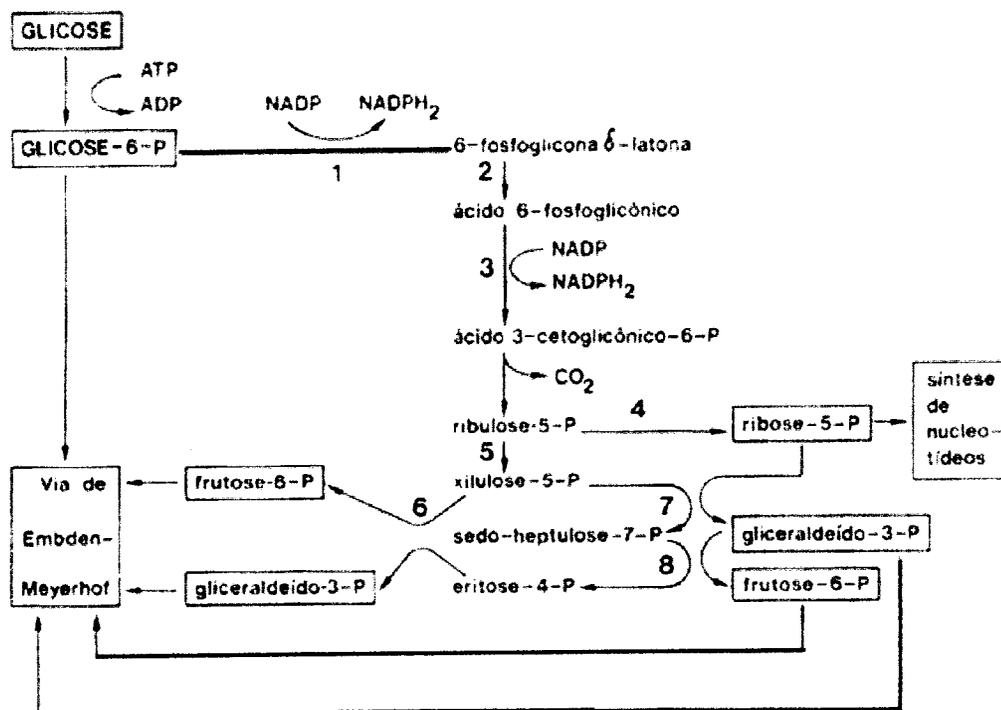


Figura 2: Via de formação das pentoses mostrando as enzimas participantes: (1) glicose-6-fosfato desidrogenase, (2) lactonase, (3) fosfogluconato desidrogenase, (4) fosforribose-isomerase, (5) fosfopentose isomerase, (6 e 7) transcetolase e (8) transaldolase.

A PGI é a segunda enzima da glicólise sendo responsável pela conversão da glicose-6-fosfato (G6P) para frutose-6-fosfato (F6P) e também participa da via das pentose, processo acoplado a primeira fase da glicólise [1]. Dessa forma a PGI é considerada um alvo potencial para o desenho racional de drogas.

Desenho racional de drogas

Uma das aplicações do desenho racional de drogas consiste em projetar compostos que inibam a enzima de parasitas, sem afetar o metabolismo do hospedeiro. Para isso é preciso estudar paralelamente tanto a enzima alvo do parasita, quanto a sua homólogo presente no metabolismo do ser humano. Isso é necessário por que a glicólise faz parte do metabolismo da maioria dos seres vivos, desde uma bactéria até um mamífero. Dessa forma todas as enzimas pertencentes a glicólise de parasitas também são encontradas no metabolismo dos seus hospedeiros e poderiam ser bloqueadas por um inibidor pouco seletivo causando efeitos colaterais indesejáveis.

Apesar de serem bastante parecidas, enzimas homólogas entre parasitas e hospedeiros apresentam diferenças estruturais a nível atômico. No caso de um

protozoário e do ser humano, essas diferenças são o resultado de milhões de anos de evolução; e é com base nessas pequenas diferenças que novas drogas poderão ser projetadas.

Outros papeis importantes da PGI-hum

Além de ser um potencial alvo para o desenvolvimento de novos fármacos, a PGI também desempenha importantes funções no metabolismo humano. A deficiência desta enzima em humanos resulta num caso particular de anemia descrito inicialmente em 1968 [2] no qual os eritrócitos assumem formas não esféricas. Desde então foram reportados, em mais de quarenta famílias, casos ligados a deficiência de PGI o que torna este quadro patológico, a terceira enzimo-patologia mais comum das células vermelhas do sangue, ficando atrás apenas da deficiência de fosfato-desidrogenase e piruvato-quinase. [3,4]. Além disso, foi determinado que a deficiência de PGI é uma doença hereditária, tipo autossômica recessiva e em alguns casos está associada a disfunções neurológicas [5].

Recentemente, foi mostrado que três proteínas importantes, neuroleucinas (NKL), fator de motilidade autócrina (FMA) e fator de maturação (FM) são todas bem parecidas, senão idênticas, a PGI-humana [6].

NKL é um fator de crescimento neurotrófico, isto é, ele promove a regeneração *in vivo* de neurônios motores. Esta ação está associada a forma monomérica da NKL, enquanto que a forma dimérica apresenta capacidade de catalisar a conversão de G6P a F6P. Ainda mais interessante foi a comprovação de que a PGI-hum reduzida a monômero apresenta atividade similar a NKL. Isso tudo leva a crer que a PGI e a NKL ou são a mesma enzima, ou estão evolutivamente bem próximas [7].

FMA representa uma classe de proteínas denominadas citocinas. Este grupo de proteínas controla o crescimento e a diferenciação celular em processo como a embriogênese, a resposta imune e processos inflamatórios [8]. FMA foi inicialmente identificado pela sua capacidade de estimular diretamente a migração celular, sendo então considerado um agente indutor de metástase.

FMA de fibrosarcoma de ratos apresentou atividade de isomerase similar a PGI, enquanto que a PGI de coração de coelho foi capaz de estimular a motilidade em fibrosarcoma de ratos da mesma forma que FMA's endógenas. Além do que,

inibidores de PGI demonstraram capacidade de inibir a motilidade em células induzidas por FMA. Estes resultados evidenciam que a FMA é uma fósfoglicose isomerase [9].

FM é um mediador da diferenciação de células de mieloma (tipo leucemia) denominadas HL-60 em monócitos terminais. A presença de FM reduz a proliferação de células do mieloma enquanto que os monócitos imaturos adquirem receptores do sistema complemento e aumentam a capacidade fagocitária. O excesso de FM não induz nenhuma diferenciação. A sequência de aminoácidos do FM apresentou 100% de homologia com a PGI e o FMA. Níveis elevados de atividade da PGI foram detectados no plasma de pacientes com leucemia aguda [10].

Cristalografia e difração de raios-X

O estudo de cristais de macromoléculas por difração de raios-X é atualmente a principal ferramenta empregada para determinação de estruturas de proteínas. A maioria das estruturas atualmente depositadas no PDB[11], foram resolvidas por difração de raios-X. A partir de uma estrutura com resolução maior do que 2.5 Å, é possível caracterizar a forma e distribuição de cargas no sítio ativo das proteínas. Com este tipo de informação, pode-se iniciar o desenho racional de um fármaco que se ligue ao sítio ativo e iniba a atividade catalítica desta proteína. Apesar de ser uma técnica aplicável a compostos que podem variar em tamanho desde um pequeno fármaco até capsídeos virais compostos pelo arranjo ordenado de centenas de subunidades proteicas, esta ferramenta esbarra na dificuldade de obter os cristais das proteínas de interesse. A variedade de formas e distribuição não homogênea de cargas nas superfícies das proteínas dificulta a ordenação das subunidades de um cristal em um arranjo ordenado e periódico. Em geral os cristais de materiais biológicos apresentam pequenas dimensões e baixa dureza [12]. Assim a obtenção de cristais de proteína requer ensaios de cristalização que varrem diversas condições de cristalização variando principalmente parâmetros como pH, força iônica, temperatura, natureza do solvente e etc ... sem que se possa garantir a obtenção de bons cristais.

CAPÍTULO 2

Sub-clonagem do gene da PGI-hum

2.1) Introdução

Sub-clonagem é o conceito empregado para descrever o experimento que transfere um gene previamente clonado de plasmídeo para outro vetor. Clonar, por definição, significa isolar um gene de um organismo, determinar sua sequência de nucleotídeos e transferir cópias deste gene para um organismo diferente. Em organismos eucariotos a clonagem requer uma etapa adicional para retirada de exons – regiões não codantes –, de forma que o gene clonado apresente apenas introns – regiões necessárias para transcrição do gene na proteína de interesse. Dessa forma, a sub-clonagem do gene da PGI-hum foi realizada para transferir o gene de uma biblioteca de cDNA para um plasmídeo com características especiais para expressão heteróloga.

As principais etapas na sub-clonagem de um gene de biblioteca de cDNA para um plasmídeo de interesse são : a construção dos oligodesoxirribonucleotídeos (oligos), a amplificação do gene de interesse por PCR (*polymerase chain reaction* - reação em cadeia da polimerase) e a ligação específica das extremidades do gene com as extremidades do plasmídeo linearizado. O plasmídeo servirá como vetor para o gene durante a transformação do organismo a ser geneticamente modificado, no caso bactérias *E. coli*.

Uma biblioteca de cDNA representa um conjunto de genes expressos simultaneamente em um determinado tecido de um organismo. Para construção de uma biblioteca de cDNA, inicialmente isola-se o conjunto de mRNA do tecido desejado, em seguida utilizando as enzima transcriptase reversa e DNA polimerase,

copia-se o mRNA em cDNA e por último cada fragmento de cDNA é inserido em um plasmídeo. O resultado disto é uma população de plasmídeos cada qual contendo uma cópia de um gene – constituído apenas pelos introns – expressos no tecido estudado. Bibliotecas de cDNA são atualmente comercializadas por empresas que fornecem materiais genéticos como reagentes para uso em laboratórios.

Um oligo é um fragmento pequeno de DNA fita simples (entre 15 a 100 bases) essencial para a início da replicação de uma fita de DNA. O oligo se liga por complementariedade de bases a uma fita simples ou desemparelhada do DNA, em seguida ele permite que a DNA polimerase adicione bases na extremidade 3' do oligo respeitando a complementariedade imposta pela fita molde. Conhecendo-se a seqüência do gene de interesse é possível construir oligos complementares as extremidades 5' e 3' do mesmo e utilizá-los em uma reação de PCR para amplificá-lo (fig.3).

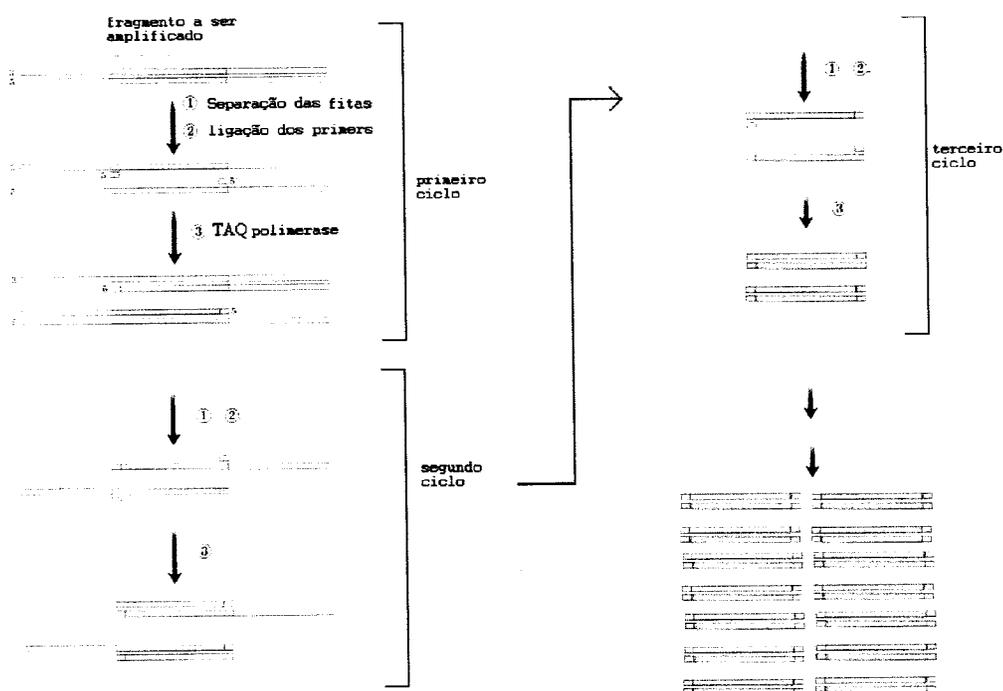


Figura 3: Representação esquemática dos primeiros ciclos da reação em cadeia da DNA polimerase (PCR). Cada ciclo passa por três etapas com diferentes temperaturas o que determina a realização de eventos diferentes. A separação das fitas (1) ocorre a temperaturas próximas a 94°C, em seguida a uma redução na temperatura para em torno de 50°C onde ocorre a hibridização dos oligos e por último, eleva-se a temperatura do sistema para o valor ótima da enzima TAQ polimerase (aprox. 72 °C) onde ocorre a polimerização da dupla fita. Após o quarto ciclo os fragmentos nascentes apresentarão a seqüência dos oligos acopladas nas suas extremidades.

O sistema de amplificação por PCR consiste basicamente de uma mistura dos primers, mais o gene de interesse presente na biblioteca de cDNA, a DNA-polimerase e dNTP's. Os dNTP's correspondem a uma mistura das quatro bases nitrogenadas que serão adicionadas pela DNA polimerase durante a complementação das fitas simples do gene de interesse. Dessa forma os dNTP's podem ser considerados como um dos substratos para a reação.

O PCR é uma reação em cadeia promovida por ciclos de temperatura onde ocorre sucessivamente, abertura da fita dupla de DNA do gene de interesse, pareamento dos oligos as regiões complementares em fita simples de DNA, ancoragem da enzima DNA-polimerase e adição dos dNTP's complementares a fita simples de DNA.

Uma vez que o gene tenha sido amplificado, ele carrega nas extremidades seqüências correspondentes aos oligos. Assim os oligos são construídos de forma a apresentarem seqüências para o reconhecimento e clivagem por enzimas de restrição previamente escolhidas. A escolha das enzimas de restrição para as quais se desenham os oligos é determinada pela seqüência do gene amplificado e pelo poli-linker do plasmídeo no qual o gene será inserido. O poli-linker é uma região especial do plasmídeo, nele estão presentes os sítios de clivagem de diversas enzimas de restrição. Portanto ao construir um oligo deve-se escolher de preferência duas enzimas diferentes e que reconheçam também o poli-linker do plasmídeo. Assim pode-se garantir que as extremidades do gene serão complementares às do plasmídeo e que a ligação ocorrerá com o gene em apenas uma orientação (Fig.4). A seqüência do gene amplificado também é determinante das enzimas de restrição a serem utilizadas, uma vez que não se pode escolher enzimas que tenham sítios de clivagem no interior do gene.

A ligação do gene amplificado ao vetor é efetivada por uma enzima denominada DNA ligase que reconhece extremidades complementares e catalisa o pareamento das bases livres desta região. A ligase não faz distinção quanto a origem das extremidades do vetor ou do inserto, basta que as bases sejam complementares para que ocorra a reação.

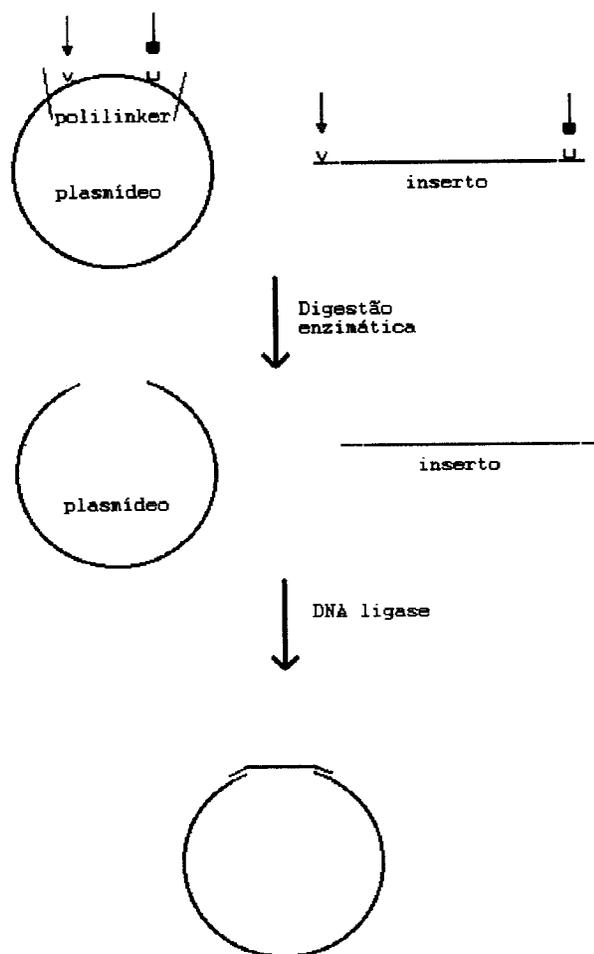


Figura 4: Representação esquemática das etapas necessárias para a montagem de um plasmídeo recombinante contendo um gene de interesse inserido na região do poli-linker. Os oligos acoplados nas extremidades do gene amplificado apresentam sítios específicos para duas diferentes enzimas de restrição (setas vermelhas e azuis).

Para que o produto de digestão enzimática tenha extremidades complementares é necessário que os mesmos sítios de restrição estejam presentes também no poli-linker do plasmídeo. Após a digestão pelas enzimas de restrição, a ligação entre o plasmídeo linearizado e o inserto é catalizada pela enzima DNA ligase. Pelo fato de utilizar duas enzimas diferentes na digestão enzimática, a ligação entre os fragmentos de DNA ocorrerão preferencialmente em uma única direção.

A etapa final na sub-clonagem de um gene consiste na introdução do produto de ligação em um organismo apto a propagar e expressar o plasmídeo assimilado, este processo é denominado transformação gênica. Durante a transformação as bactérias, são submetidas a um choque térmico que aumenta a permeabilidade da membrana facilitando a assimilação de DNA livre em solução (fig.5).

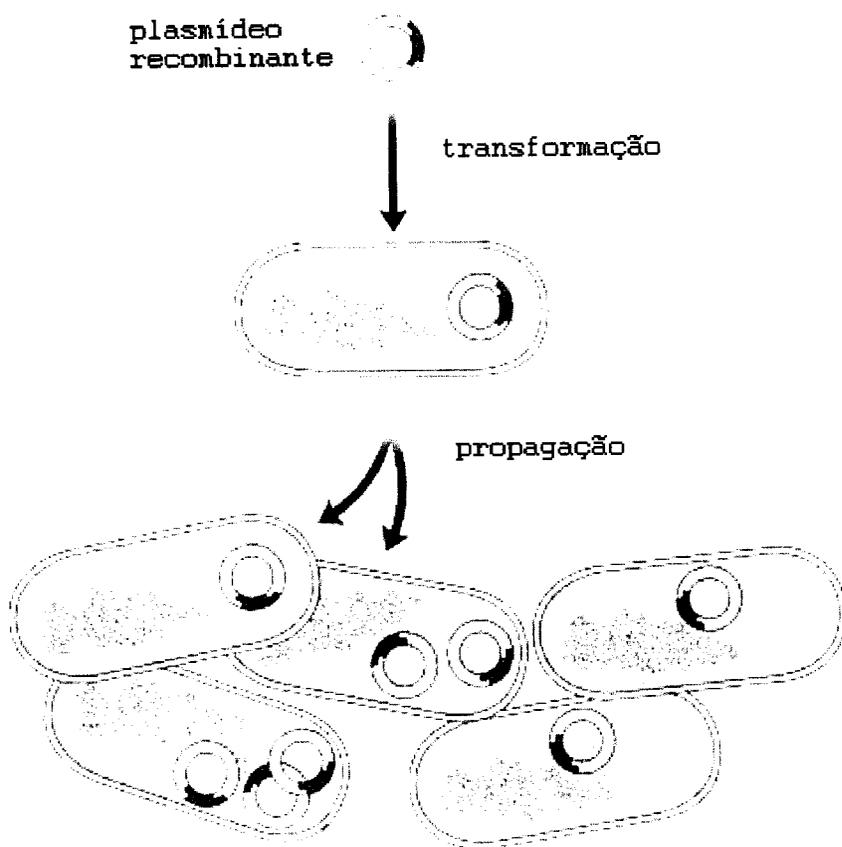


Figura 5: Representação esquemática da transformação de células competentes de *E. coli* por um plasmídeo recombinante contendo um gene de interesse. Após a transformação, o organismo geneticamente transformado passa a propagar e replicar o plasmídeo assimilado.

Para aumentar a capacidade das células assimilarem fragmentos exógenos de DNA, tais bactérias passam por um tratamento com cloreto de cálcio que aumenta a permeabilidade da membrana celular [13]. As bactérias submetidas a tal tratamento passam a ser denominadas células competentes.

2.2) Métodos

Dois oligos A e B foram encomendados para *Life Technologies* de forma que o oligo A inserisse um sítio de clivagem no gene para a enzima de restrição *Nde* I enquanto o oligo B apresentasse um codon de terminação para transcrição e um sítio de clivagem para a enzima de restrição *Hind* III (fig.6)

agarose correspondente ao gene da PGI-hum foi eluída em 30 µl de H₂O utilizando o kit para eluição de DNA *Gen Clean Kit II* da *BIO 101 Inc.*

Para a digestão do vetor e inserto com as enzimas apropriadas foram montados os seguintes sistemas de digestão:

Digestão do plasmídeo

3,5 µl de Tampão React2	(10x)
29,5 µl de pET29-a(+)	(100 ng/µl)
1 µl de HindIII	(20 U/µl)
1 µl de NdeI	(20 U/µl)

Digestão do gene da PGI-hum

3,5 µl de Tampão React2	(10x)
29,5 µl de e do produto de amplificação por PCR	
1 µl de HindIII	(20 U/µl)
1 µl de NdeI	(20 U/µl)

As enzimas e tampões foram obtidos as *Life Technologies*. Cada sistema foi incubado a 37 °C por 3 horas. Em seguida foram aplicados em gel de agarose (1% em TAE com Brometo de etídeo a 0,4µg/ml) e as bandas foram eluídas utilizando o kit para eluição de DNA *Gen Clean Kit II* da *BIO 101 Inc.* Cada banda foi eluída em 20 µl em de H₂O.

A montagem do plasmídeo recombinante do vetor pTE29-a(+) e o gene da PGI-hum, previamente digeridos pelas enzimas de restrição *Nde I* e *Hind III*, foi realizada na seguinte mistura de ligação:

1,5 µl de tampão ligase	(10x)
1,0 µl de T4 DNA ligase	(1 U/µl)
1,5 µl de pET29-a(+)	(100 ng/µl)
11 µl de PGI-hum eluído do gel de agarose	

Este sistema foi mantido a temperatura ambiente (23 a 25 °C) por 6 horas. O controle deste experimento foi realizado substituindo-se o gene de PGI-hum por H₂O. A enzima T4 DNA ligase e o tampão ligase foram obtidos da *PROMEGA*.

Para realizar a transformação de células competentes de *E. coli* capazes de regular a expressão do plasmídeo pET29-a(+), o produto de ligação foi inicialmente

transformado em *E. coli*, cepa DH5 α . Após propagação pela cepa DH5 α , grandes quantidades do plasmídeo recombinante foram extraídas por lise alcalina no sistema MINIPREP (QIAGEN) e novamente transformadas em *E. coli* de cepas BL21(DE3) e ER2566, específicas para regulação da expressão de genes inseridos no plasmídeo pET29a(+).

A transformação de células competentes de *E. coli* através de choque térmico foi realizada misturando-se 10 μ l do produto de ligação com 100 μ l de células competentes de *E. coli* tipo DH5 α . Para aplicação do choque térmico a mistura foi mantida envolta em gelo por quinze minutos; em seguida transferida por dois minutos para um banho-maria a 42 °C, retornando para o gelo por mais cinco minutos.

Após o choque térmico, foi adicionado 1 ml de meio LB [13] à mistura. Este sistema foi incubado a 37°C com rotação de 250 rpm por uma hora. Após o período de incubação centrifugou-se as células, resuspendeu-as em 100 μ l do próprio sobrenadante e este volume foi espalhado sobre placas de petri contendo meio sólido tipo LB com canamicina na concentração final de 50 μ g/ml. As placas foram mantidas em estufa a 37°C por 24 horas. O gene de resistência a canamicina, presente no plasmídeo pET29-a(+), permitiu que o crescimento de colônias a partir de células que assimilaram o plasmídeo durante o choque térmico.

Após o período de 24 horas a 37 °C, as colônias obtidas foram transferidas para 5 ml de meio LB líquido e incubadas a 37 °C com agitação de 250 rpm. Em seguida os plasmídeos destas colônias foram extraído por lise alcalina com o kit QUANTUM PREP (BioRad).

O resultado da transformação *E. coli* das cepas DH5 α , BL21(DE3) e ER2566 foi confirmado por eletroforese em gel de agarose (1% em TAE com Brometo de etídeo a 0,4 μ g/ml) do produto da digestão destes plasmídeos com as enzimas de restrição *Nde* I e *Hind* III.

2.3) Resultados e discussão

Os plasmídeos extraídos de *E. coli* das cepas DH5 α , BL21(DE3) e ER2566, foram submetidos a dois sistemas diferentes de digestões: uma por apenas *Hind* III e outra concomitantemente pelas duas enzimas *Nde* I e *Hind* III (fig.7).

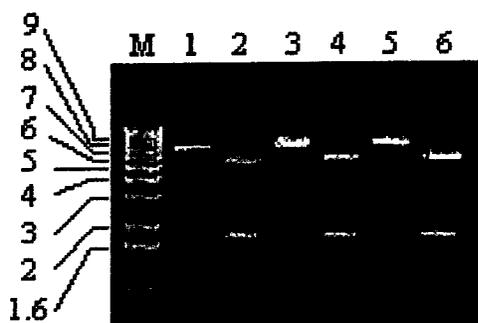


Figura 7: Gel de agarose (1% em TAE com Brometo de etídeo a 0,4 μ g/ml) mostrando o produto da digestão dos plasmídeos extraídos das três cepas de *E. coli* utilizadas no montagem do sistema de expressão. Da esquerda para direita, mostrando a cepa da qual foi extraído o plasmídeo e as enzimas usadas na digestão: marcador de peso molecular (M); DH5 α digerido por *Hind* III (1); DH5 α digerido por *Hind* III e *Nde* I (2); BL21(DE3) digerido por *Hind* III (3); BL21(DE3) digerido por *Hind* III e *Nde* I (4); ER2566 digerido por *Hind* III (5); e ER2566 digerido por *Hind* III e *Nde* I (6).

Apesar do resultado positivo para a transformação de *E. coli* da cepa DH5 α esta etapa serviu apenas para produzir e purificar plasmídeo recombinante em grandes quantidades. A cepa de *E. coli* DH5 α não apresenta o promotor necessário para regulação do plasmídeo pET29-a(+), porém é uma cepa de alta capacidade de transformação.

Pelo resultado da digestão do DNA plasmidial, extraído das três cepas de *E. coli* utilizadas para a montagem do sistema de expressão, pode-se verificar que todas as cepas foram devidamente transformadas com o plasmídeo contendo o gene da PGI-hum. Isto fica claro pelo padrão de bandas no gel (fig. 7). Onde o plasmídeo foi digerido apenas com *Hind* III observa-se uma única banda de alto peso molecular (entre 7 e 8 Kb) correspondendo ao plasmídeo conectado ao inserto pelo sítio de *Nde* I. Onde houve a digestão com as duas enzimas (*Hind* III e *Nde* I) observa-se dois fragmentos, o de maior peso molecular (~ 5.5 Kb) equivale ao plasmídeo enquanto o de menor peso molecular (1.8 Kb) equivale ao gene da PGI-hum.

CAPÍTULO 3

Expressão e purificação da enzima PGI-hum

3.1) Expressão da PGI-hum recombinante

3.1.1) Introdução

A expressão de genes hexógenos em bactérias é regulada por sistemas constituídos de operadores e reguladores. O modelo *operon* é constituído de um conjunto de genes que incluem, um gene regulador, uma região operadora e um conjunto de genes estruturais entre os quais podem ser inseridos o gene de interesse e um gene seletivo (Fig. 8).

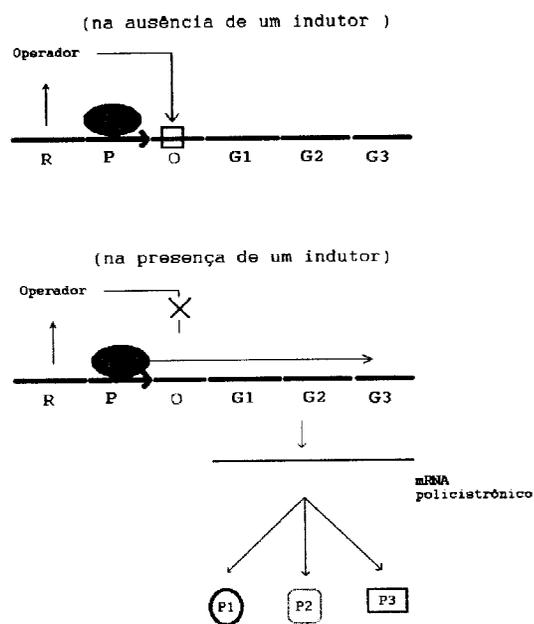


Figura 8: Representação esquemática de um operon *Lac* contendo uma região reguladora (R), um promotor (P), uma região operadora (O) e três genes (G1, G2 e G3) que quando transcritos resultam em um único mRNA policistrônico que será traduzido nas três proteínas (P1, P2 e P3) codificadas pelos respectivos genes. Na ausência de um agente indutor, o operador pode agir como uma proteína repressora da expressão gênica, ligando-se a região operadora e bloqueando a passagem da RNA polimerase. Ao adicionar o indutor, o operador perde afinidade pelo sítio operador permitindo assim a passagem da RNA polimerase e consequentemente a expressão dos

O gene regulador codifica uma proteína que apresenta afinidade pela região operadora localizada na fita de DNA. Entre o gene regulador e a região operadora existe a região promotora (ou simplesmente promotor) que serve como ponto de ancoragem da RNA-polimerase. Se a região operadora estiver bloqueada, a transcrição dos genes posteriores não será realizada porque a enzima RNA-polimerase não será capaz de transpor a região operadora.

O plasmídeo pET29-a(+) apresenta uma região operadora *lac*, um promotor específico para a T7 RNA polimerase e um gene de resistência ao antibiótico canamicina. A proteína repressora é um tetrâmero de subunidades idênticas de 37 KDa. Cada subunidade apresenta um sítio de ligação para isopropil-tiogalactosídeo (IPTG) que se liga fortemente ao repressor e impede a ligação deste com a região operadora. Dessa forma o complexo IPTG/operador permite a expressão dos genes associados ao operon.

O IPTG é portanto um indutor não metabólico da expressão de genes associados ao operon *lac*.

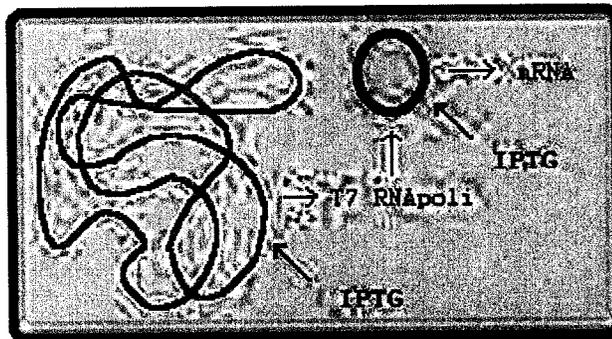


Figura 9: Representação esquemática de uma *E. coli* transformada com o plasmídeo pET29-a(+) (em preto). Esta cepa de *E. coli* apresenta em seu DNA circular (em cinza) o gene para expressão da T7 RNA polimerase de bacteriófago λ sobre a regulação de um operon *lac*. Na presença de isopropil-tiogalactosídeo (IPTG), ocorre a expressão de T7 RNA polimerase que irá promover a transcrição do genes associados ao operon *lac* do plasmídeo pET29-a(+). O IPTG também libera o sítio operador presente no plasmídeo.

Uma característica importante deste plasmídeo e a especificidade do seu promotor pela T7 RNA polimerase do bacteriófago λ . Mesmo que a região operadora presente-se desbloqueada, a transcrição só ocorrerá na presença de T7 RNA polimerase, pois a RNA polimerase da *E. coli* apresenta pouca afinidade por este promotor. Dessa forma uma alternativa para expressar genes do plasmídeo pET29-a(+) em *E. coli* é utilizar cepas que apresentem o DNA genômico do bacteriófago

sobre a regulação de um outro operon *lac*. O DNA do fago pode estar inserido no DNA da bactéria ou associado a outro plasmídeo constitutivo (fig.9).

Um dos principais problemas da expressão de proteínas hexógenas a *E. coli* é que a bactéria pode não ser capaz de enovelar corretamente a proteína de interesse ou forma corpúsculos de inclusão. Corpúsculos de inclusão são formados geralmente por proteínas que foram mal enoveladas, podendo conter também proteínas corretamente enoveladas que por motivos ainda não conhecidos formem uma massa protéica insolúvel.

3.1.2) Métodos

A expressão foi realizada com ambas as cepas *E. coli* BL21(DE3) e ER2566 para verificar se o produto da transcrição do gene inserido no vetor pET29-a(+) corresponde a uma proteína de peso molecular semelhante ao descrito na literatura para a PGI-hum [14]. Também foi analisado o efeito da concentração de IPTG sobre o nível de expressão gênica do plasmídeo recombinante.

O ensaio de expressão seguiu os passos descritos abaixo e teve seu resultado analisado por eletroforese em gel SDS-PAGE.

1. Preparar um pré-inóculo a 37 °C e 250 rpm, com 5 ml de meio LB contendo kanamicina ($C_f = 50 \mu\text{g/ml}$) para colônias BL21(DE3) e ER2566.
2. Inocular 5 ml de meio LB/kanamicina com 0,5 ml de cada pré-inóculo.
3. Incubar a 37°C com agitação de 250 rpm até atingir a D.O. de aproximadamente 0.6.
4. Adicionar IPTG para $C_f = 0,5$ e 1,0 mM
5. Incubar por mais duas horas a 37°C com agitação de 250 rpm.
6. Retirar alíquotas de 500 μl , centrifugar, ressuspender o pellet em 40 μl de H_2O e adicionar 20 μl de tampão de amostra (TA) c/ β -mercaptoetanol 3x.
7. Ferver as alíquotas e aplicar em SDS-PAGE (15%).

3.1.3) Resultados e discussão

Pelo SDS-PAGE (fig.10) pode-se observar que ambos os sistemas de expressão montados estão expressando uma proteína entre 55 e 60 KDa o que está de acordo com o valor descrito por Carter [14] em 1969. Este resultado sugere que a proteína expressa seja a PGI-hum recombinante clonada a partir da biblioteca cDNA de cérebro de feto humano.

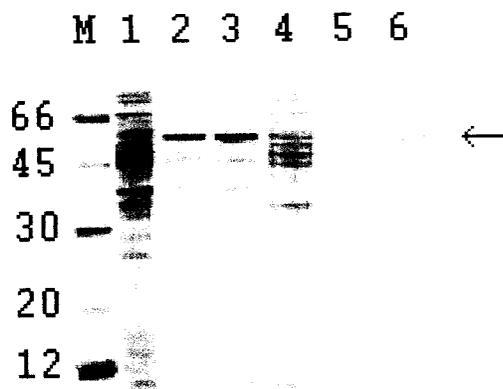


Figura 10: SDS-PAGE a 15% de poliacrilamida (método Laemmili). A coluna de números a esquerda refere-se ao peso em KDa das proteínas usadas como marcadores (M); em seguida, extrato protéico total antes da adição de IPTG (1) , após duas horas de indução com 0.5 mM (2) e 1,0 mM (3) de IPTG, todos referentes a expressão da cepa *E. coli* BL21(DE3). As colunas (4), (5) e (6) referem-se respectivamente ao extrato protéico total antes da adição de IPTG, após duas horas de indução com 0,5 mM e após duas horas de indução 1.0 mM de IPTG, todas utilizando a cepa *E. coli* ER2566. A seta indica a banda correspondente a PGI-hum.

Comparando o nível de expressão entre as duas cepas de *E. coli* observa-se que a BL21(DE3) apresenta uma expressão mais acentuada. Também pode-se concluir que 0.5 mM de IPTG para a BL21(DE3) é suficiente para obter uma indução efetiva.

Com base nestes resultados foi escolhido o sistema clonado em *E. coli* BL21(DE3) para prosseguir na tentativa de montar um protocolo de purificação da PGI-hum recombinante.

3.2) Purificação da PGI-hum recombinante

3.2.1) Introdução

O protocolo estabelecido para purificação da PGI-hum recombinante seguiu três etapas básicas utilizadas em purificação de proteínas. Primeiramente realizou-se uma precipitação por sulfato de amônia $(\text{NH}_4)_2\text{SO}_4$ seguida por duas etapas de

cromatografia, uma de troca iônica e finalizando com uma de exclusão por tamanho em gel filtração.

A precipitação de proteínas por adição de sais é conhecida do inglês como *salting out* [15]. O aumento da concentração de sais na solução requer mais moléculas d'água para coordenação e solvatação dos íons livres, isto resulta em uma competição entre íons e proteínas por moléculas d'água para solvatação. Com o aumento crescente da quantidade de íons, as proteínas passam a se aglomerar para reduzir a superfície acessível ao solvente e terminam por precipitarem. O *salting out* geralmente é aplicado a extratos protéicos totais devido a sua baixa resolução e a possibilidade de se trabalhar com grandes volumes. Um vez precipitada, muitas proteínas são facilmente resolubilizadas sem danos conformacionais.

A cromatografia de troca iônica separa as proteínas em função de sua carga líquida [15]. Utilizam-se resinas catiônicas para separar proteínas apresentando carga líquida positiva, enquanto resinas aniônica são usadas para separar proteínas de carga negativa. Conhecendo-se o ponto isoelétrico (pI) e ajustando o pH da solução, pode-se definir a carga líquida de uma proteína. Em pH inferior ao pI, prevalecem as cargas positivas e em pH superior ao pI a carga líquida passa a ser negativa. No caso da PGI-hum, foi utilizada uma resina catiônica empacotada em coluna fechada e acoplada a um sistema a baixa pressão, tipo FPLC.

A cromatografia de gel filtração [15] separa as proteínas em função do tamanho das partículas presentes na fase móvel. A resina de gel filtração (fase estacionária) é composta por um polímero insolúvel mas altamente hidratado, como dextran, agarose ou poliacrilamida, capaz de reter partículas de acordo com o tamanho dos poros formados. Partículas menores são retardadas pela malha de poros, enquanto as maiores migram com maior velocidade eluindo primeiro da coluna.

3.2.2) Métodos

Apartir do ensaio de expressão descrito anteriormente, foi estabelecido um protocolo para expressão do gene da PGI-hum transformado em cepas de *E. coli* BL21(DE3) utilizando um volume maior de meio de cultura. O meio de cultura foi alterado para 2xYT [13] por apresentar uma maior concentração de nutrientes em sua composição.

O protocolo final para expressão e purificação da PGI-hum pode ser separado em quatro etapas principais que seguem os passos abaixo:

Expressão e lise celular:

1. Inocular 500 ml de 2xYT/ kanamicina ($C_f = 50 \mu\text{g/ml}$) com 5 ml de pré-inóculo; crescer a 37°C e 250 rpm por 2,5 horas.
2. Adicionar IPTG ($C_f = 0,5 \text{ mM}$) e incubar por mais 4 horas a 37°C e 250 rpm.
3. Centrifugar a 6000g por 10 minutos a 4°C ; resuspender o pellet em 20 ml de tampão de corrida (T.C.) BICINA 50 mM, pH 7,8.
4. Lise celular por choque térmico: adicionar 2 ml de solução de lisozima (5 mg/ml), 4 μl de DNase/RNase (10 U/ μl) e aplicar oito ciclos de congelamento em N_2 líquido e descongelamento a 37°C em banho maria.

Precipitação com $(\text{NH}_4)_2\text{SO}_4$:

5. Centrifugar a 10000g por 20 minutos a 4°C e recuperar o sobrenadante (proteína total).
6. Realizar uma precipitação com 50 % de $(\text{NH}_4)_2\text{SO}_4$ do extrato proteico total, descartar o precipitado e recuperar o sobrenadante (proteína 50%).
7. Levantar o sobrenadante recuperado a 80 % de $(\text{NH}_4)_2\text{SO}_4$, descartar o sobrenadante e resuspender o pellet em 9 ml de T.C.
8. Dialisar o produto da precipitação por $(\text{NH}_4)_2\text{SO}_4$ duas vezes contra 250 ml de T.C.

Troca Catiônica, coluna de 8 ml (SP-sepharose-HP):

9. Equilibrar a coluna com 100 ml de T.C.
10. Aplicar 10 ml de amostra (obtidas da precipitação com $(\text{NH}_4)_2\text{SO}_4$)
11. Lavar a coluna com 80 ml de T.C.
12. Aplicar um gradiente salino até 2 M de NaCl em T.C. durante 64 ml de corrida (6 vezes o volume da coluna).
13. Concentrar por ultra centrifugação (em CENTRIPREP30) as frações referentes ao pico coletado até um volume final de 4 ml.

Gel filtração, coluna de 115ml (superdex-200)

14. Equilibrar a coluna com 240 ml de T.C. (mais 150 mM de NaCl no tampão)
15. Aplicar a amostra concentrada (no passo 13) com fluxo de 1 ml/min
16. Coletar frações de 2 ml cada.

3.2.3) *Resultados e discussão*

A primeira etapa da purificação foi definida em experimento prévio (não mostrado) variando-se a concentração de $(\text{NH}_4)_2\text{SO}_4$ de 30 a 80 % em intervalos de 10 %. Verificou-se que a grande parte das proteínas presentes no lisado celular precipitaram com 50 % de $(\text{NH}_4)_2\text{SO}_4$, enquanto a PGI-hum começou a precipitar a partir de 60 % (dados não mostrados). Dessa forma, iniciando-se a precipitação com 50 % de $(\text{NH}_4)_2\text{SO}_4$, conseguiu-se separar uma grande massa de contaminantes que poderiam dificultar as etapas cromatográficas seguintes. Outra vantagem de iniciar a purificação por um *salting out* é que após eliminar os contaminantes com 50 % de $(\text{NH}_4)_2\text{SO}_4$ (passo 6), a precipitação de toda PGI-hum a 80 % mostrou-se uma forma eficiente para concentrar a proteína. Uma vez precipitada, a massa proteica pode ser ressuspensa no volume adequado (passo 7) para ser aplicado do coluna cromatográfica de troca catiônica.

As etapas cromatográficas foram realizadas num sistema de FPLC (Pharmacia) no qual foi monitorada a absorbância a 280 nm. A troca catiônica, realizada na coluna de 8 ml tipo SP-sepharoseHP (fig.11a), resultou em um pico eluído com aproximadamente 1,6 M de NaCl. Isso indica que a interação da proteína com a resina é bastante forte, pois em geral um gradiente de até 1 M de NaCl é suficiente para eluir a maioria das proteínas em interações com colunas de troca iônica.

Na etapa de gel filtração em coluna de 115 ml tipo superdex-200, foi aplicado o conteúdo referente ao pico eluído da troca catiônica, mediante prévia concentração até 4 ml. Esta concentração é um pre-requisito imposto pela coluna de gel filtração, uma vez que a amostra a ser aplicada não deve ultrapassar 5% do volume da coluna.

O perfil cromatográfico da gel filtração (fig.11b) evidencia que, apesar de eluído como um pico único da coluna de troca catiônica, este pico não é composto por uma única proteína. A PGI-humana elui da gel filtração como um pico único com volume de eluição de aproximadamente 75 ml, depois de 100 ml de corrida surgem contaminantes de baixo peso molecular mas que contribuíam para a absorbância a 280 nm do pico eluído da troca catiônica.

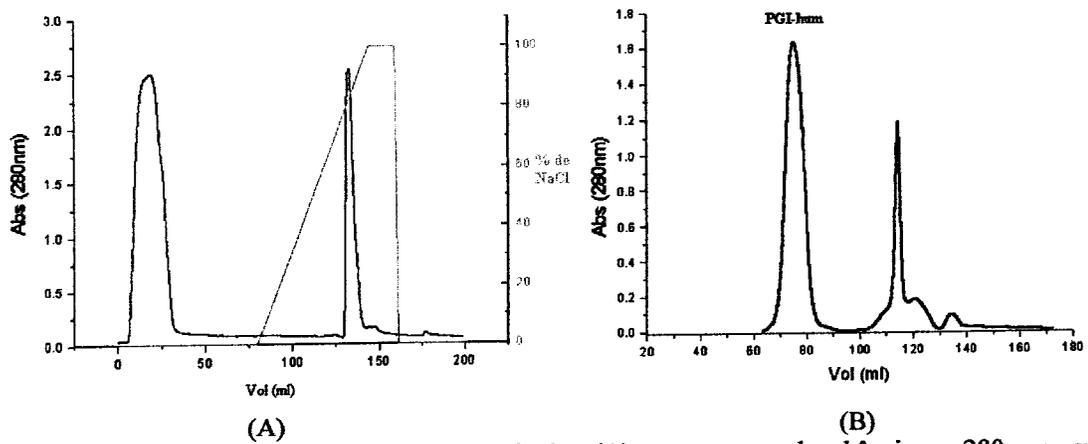


Figura 11: Perfil cromatográfico da troca catiônica (A), em preto a absorbância a 280 nm, em vermelho e o gradiente salino em porcentagem até 2M de NaCl. Perfil cromatográfico da gel filtração (B) relacionando a absorbância a 280 nm com o volume de eluição em ml.

O acompanhamento das três etapas de purificação por SDS-PAGE (15%, método Laemmli) (fig.12) confirma a presença de contaminantes de baixo peso molecular após a troca catiônica e revela a presença de uma pequena impureza caracterizada como uma fraca banda logo acima da PGI-hum.

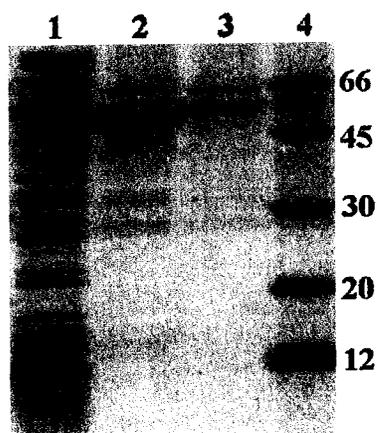


Figura 12: SDS-PAGE 15% de poliacríamida mostrando o conteúdo proteico das três etapas de purificação da PGI-humana: extrato proteico após a precipitação com $(\text{NH}_4)_2\text{SO}_4$ (1), extrato proteico referente ao pico coletado da troca catiônica (2), pico coletado a 75 ml após injeção da amostra na gel filtração (3) e marcadores de peso molecular (4).

CAPÍTULO 4

Determinação de parâmetros cinéticos da PGI-hum recombinante

4.1.) Introdução

O estudo de cinética enzimática tem como objetivo fornecer informações sobre o mecanismo da reação catalisada por uma enzima. A afinidade de uma enzima pelo seu substrato, a taxa de conversão entre o substrato e o produto, a interação da enzima com cofatores e o tipo de interação com inibidores, são todos exemplos de informações obtidas a partir de ensaios enzimáticos.

Para a determinação de parâmetros cinéticos, uma condição determinante é que a enzima trabalhe sempre com excesso de substrato. Dessa forma, as taxas de conversão do substrato em produto não serão afetadas pela variação da concentração do substrato, podendo ser monitoradas apenas pela formação do produto. Outra consequência importante é que a concentração do complexo enzima-substrato pode ser considerada constante, esta condição caracteriza o estado estacionário da reação [16].

Uma maneira simples de determinar parâmetros cinéticos tipo V_{max} e K_m é monitorar a formação de produto para diferentes concentrações iniciais de substrato, mantendo a concentração de enzima constante. Este tipo de estudo pode ser realizado com um espectrofotômetro desde que o produto apresente um máximo (ou pico) de absorvância característico e diferente dos demais reagentes.

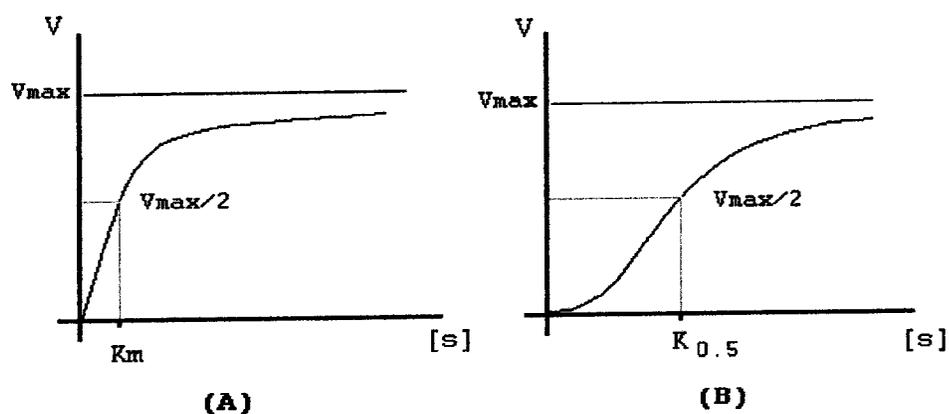


Figura 13: Gráfico de velocidade de reação por concentração de substrato ($V \times [S]$). Em (A) a curva descrita é característica de uma enzima não cooperativa que segue a cinética de Michaelis-Menten, onde K_m equivale a metade da velocidade máxima (V_{max}). Em (B) está representada a curva sigmoidal de uma enzima cooperativa. O aumento inicial da concentração do substrato resulta numa mudança não linear da velocidade da reação, isso indica um aumento da afinidade da enzima pelo substrato.

O resultado desse tipo de experimento pode ser visualizado em um gráfico de velocidade da reação por concentração de substrato (fig. 13) e pela forma da curva pode-se determinar se a enzima interage de forma cooperativa ou não com o substrato. Uma enzima não cooperativa apresenta uma cinética tipo Michaelis-Menten (fig. 13a) até atingir um platô para o qual tende a velocidade máxima da reação. No caso de uma enzima cooperativa, este crescimento é descrito por uma curva sigmoidal (fig. 13b), resultado do aumento de afinidade da enzima pelo substrato causada pelo próprio substrato.

K_m é um parâmetro que pode ser interpretado como a afinidade de uma enzima michaeliana pelo seu substrato e é caracterizado como a concentração de substrato na qual metade dos sítios ativos estão ocupados. Dessa forma o K_m é calculado como sendo a concentração de substrato na qual a enzima apresenta metade de sua $V_{máx}$. A forma mais comum de se calcular o K_m e V_{max} para uma enzima é através do gráfico duplo recíproco de Lineweaver-Burk. Neste gráfico o eixo x representa o inverso da concentração do substrato e o eixo y representa o inverso da velocidade. O ponto onde a reta corta o eixo x equivale a $-1/K_m$ enquanto que o ponto onde ela corta o eixo y equivale a $1/V_{max}$.

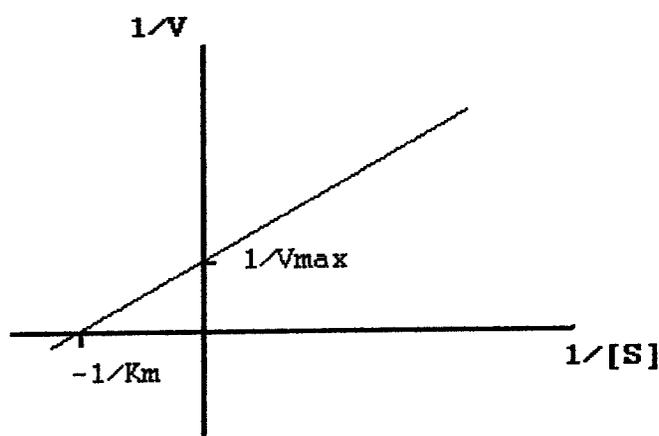


Figura 14: Gráfico de Lineweaver-Burk para uma enzima michaeliana. Este gráfico é construído pela relação do inverso da velocidade pelo inverso da concentração de substrato ($1/V \times 1/[S]$) e fornece valores precisos para V_{max} e K_m . A inclinação da reta é dada pela relação K_m / V_{max} .

Para testar a eficiência de diferentes inibidores contra uma enzima, pode-se medir a velocidade de reação com uma concentração de substrato definida e adicionando-se a mesma quantidade de cada inibidor verificar o quanto a presença de cada inibidor reduz a velocidade da reação medida apenas na presença do substrato. Dessa forma pode-se escolher o melhor inibidor para, em seguida, definir a constante de inibição (K_i) em relação a determinado substrato. A medida de K_i é um pouco mais complexa pois requer a medida da constante de afinidade (similar ao K_m) para diferentes concentrações de inibidor.

4.2) Métodos

Os parâmetros cinéticos K_m e V_{max} da PGI-hum recombinante foram determinados medindo a velocidade (atividade por unidade de tempo) de reação para diferentes concentrações de substrato através de uma variação do método de Gracy e Tilley [17].

A medida da atividade da PGI-hum é feita indiretamente medindo a formação de NADPH durante a conversão de glicose-6-fosfato (G6P) a 6-fosfoglicona-d-latona (Fig 15). O sistema para teste de atividade consiste em fornecer frutose-6-fosfato (F6P) e NADP para as enzimas PGI-hum e glicose-6-fosfato desidrogenase (G6PDH). A conversão de G6P em F6P, pela PGI-hum, fornece o substrato - F6P - para a enzima G6PDH que realiza a síntese de 6-fosfoglicona- δ -lactona mediante a conversão de NADP a NADPH.

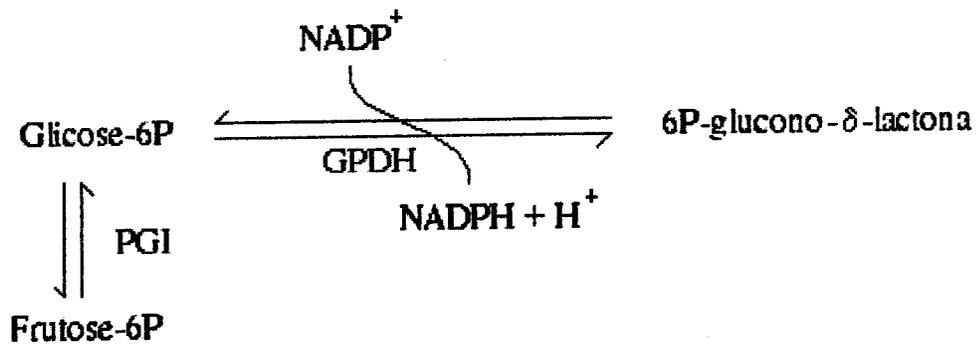


Figura 15 : Representação esquemática do sistema de reações utilizados para medir a atividade da glicose-6-fosfato isomerase (PGI). A conversão de glicose-6P para frutose-6P faz parte da via glicolítica, enquanto a conversão de glicose-6P para 6P-glucono- δ -lactona pela enzima glicose 6-fosfato desidrogenase (GPDH) faz parte da via das pentoses.

A formação de NADPH é lida por um espectrofotômetro no comprimento de onda igual a 340 nm. A diferença de absorbância a 340 nm entra no cálculo da velocidade da PGI-hum pela seguinte fórmula:

$$V = [\Delta Abs (UA/min) \cdot 1000] / 6,22$$

onde, ΔAbs é a variação de absorbância (em UA) a 340 nm por minuto de reação, 6,22 é o valor do coeficiente de extinção molar de NADPH (em UA/mM) [17], 1000 é um fator de correção para as unidades. A unidade de velocidade é definida como μmol de substrato consumido por minuto de reação. O ensaio enzimático da PGI-hum foi padronizado de acordo com a tabela abaixo, variando a concentração de F6P. Para cada medida de velocidade o ensaio foi realizado em triplicata

Tabela II: Reagentes utilizados para cada medida de atividade da PGI-hum. Os reagentes e volumes aplicados aparecem na ordem em que são adicionados a cubeta de reação.

Ordem reacional	Reagente	Volume (ml)	Conc. Final (mM)
1	trietanolamina	200	100
2	H ₂ O	639	
3	F6P	50	Variação de 0.1 a 1.3
4	NADP	50	0.4
5	MgCl ₂ . 6 H ₂ O	50	7
6	G6PDH	1	(~ 1 unidade)
7	PGI-hum	10	~ 15 unidades

O volume total para cada teste de atividade é de 1 ml e a cubeta utilizada para as medidas de absorbância possuía um caminho ótico de 1 cm.

Para o teste de eficiência inibitória foram utilizados quatro diferentes inibidores, gentilmente fornecidos pelo Dr. Laurent Salmon (Lab. de Química Biorgânica e Bioinorgânica de Universidade de Paris-XI) e por ele denominados: 5PAH, 5PAHz, 5PAAm e 5PAA (Fig.13) [18].

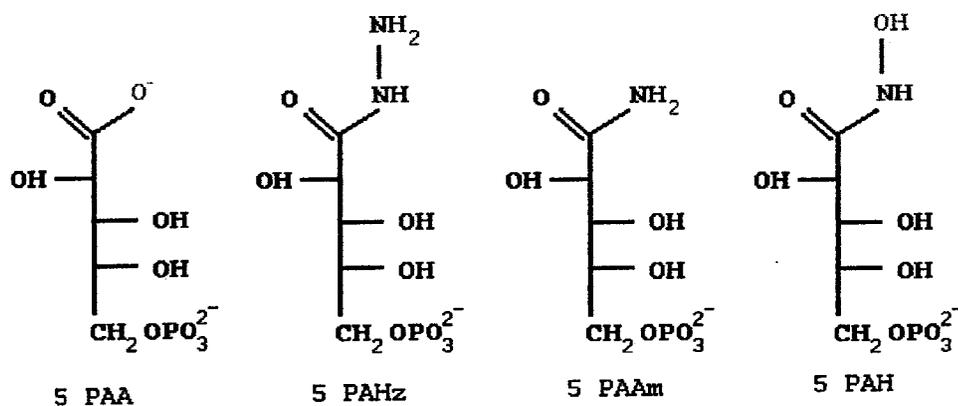


Figura 16: Inibidores testados contra a PGI-hum: D-arabinoato-5-fosfato (5 PAA), D-arabinohidroxiazida-5-fosfato (5 PAHz), D-arabinoamida-5-fosfato (5 PAAm) e D-arabinohidroximato-5-fosfato (5 PAH).

Os inibidores foram acrescentados com reagentes para teste de atividade e a absorbância medida foi comparada com controles realizados sem inibidores. As principais variáveis definidas para o teste de inibição estão descritas a seguir (tab. III).

Tabela III: variáveis definidas durante os ensaios de eficiência inibitória realizados para os inibidores : 5PAH, 5PAHz, 5PAAm e 5PAA

Conc. da PGI-hum	0.7 nM
Conc. de F6P	0.8 mM
Conc. de Inibidor	20 μ M
Relação F6P/Inibidor	40x
Var. máxima de Abs (280nm)	0,123 (100%)
Conc. de PGI	19,8 μ M

Todas as concentrações referem-se as calculadas para um ensaios realizados com um volume final de 1 ml.

4.3) Resultados e discussão

Valores de K_m igual a 96 μ M e V_{max} igual a 21 μ M/min foram calculados a partir do gráfico de Lineweaver-Burk (fig.17b).

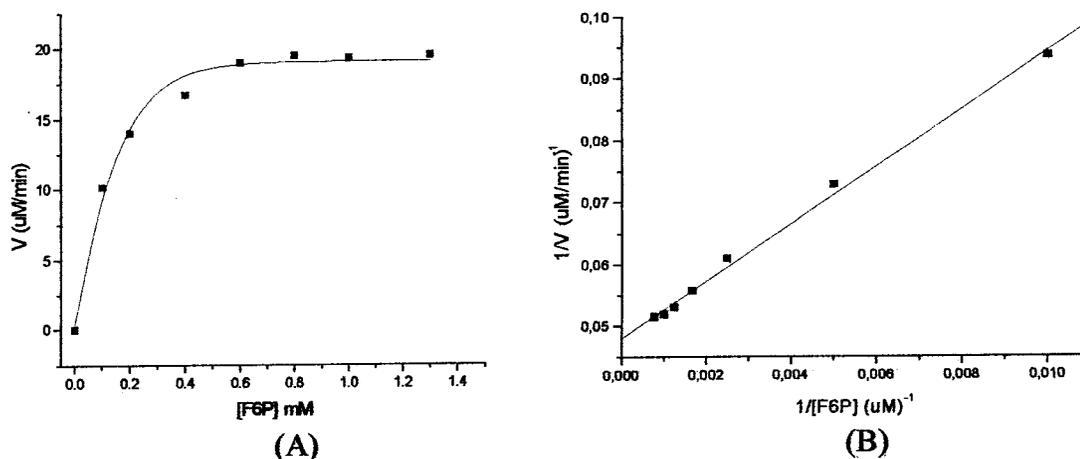


Figura 17: Gráfico da relação entre [F6P] e a velocidade de reação (A). Gráfico de Lineweaver-Burk ou também conhecido como duplo recíproco (B).

O teste de eficiência inibitória para os quatro compostos disponibilizados indicou uma maior capacidade de inibição para o composto 5PAA seguido do 5PAH, enquanto que os compostos 5 PAHz e 5PAAm apresentaram uma ação inibitória insignificante (tab.4).

Tabela IV: Porcentagem de atividade restante devido a presença dos inibidores

Inibidor	% de atividade restante
5PAH	71,5
5PAHz	95
5PAAm	97,5
5PAA	65

O próximo passo no estudo dos inibidores será escolher entre o 5PAH e o 5PAA para determinar o K_i correspondente e em seguida iniciar tentativas de co-cristalização da PGI-hum na presença do inibidor.

Um detalhe importante ao comparar a eficiência inibitória e a composição química dos quatro compostos refere-se ao grupamento associado ao carbono-1. A princípio pode-se concluir que a presença de um oxigênio em 5 PAA e de uma hidroxila em 5 PAH seriam as responsáveis pela maior capacidade inibitória quando comparados aos inibidores 5 PAAm e 5 PAHz (fig. 16).

A única estrutura de um complexo entre uma PGI de mamífero (coelho) com um inibidor fosfatado [19] indica que duas pontes de hidrogênio são formadas entre o oxigênio de uma carboxila em C-1 com um o grupo amida de cadeia lateral de uma arginina. O inibidor presente no complexo é o D-gluconato-6-fosfato (6PGH), que difere do inibidor 5 PAA apenas por apresentar um carbono a mais (fig. 18)

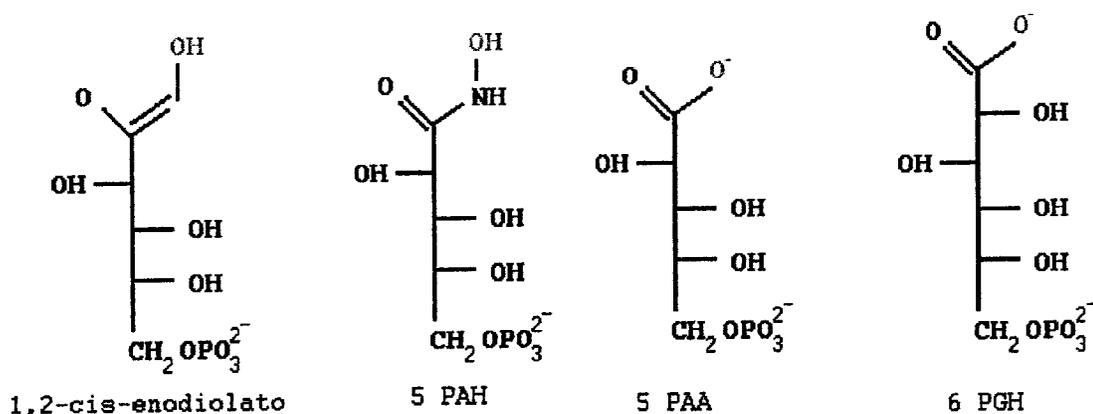


Figura18: Composição química do intermediário de alta energia da reação de isomerização catalisada pela PGI-hum (1,2-cis-enodiolato), dos inibidores D-arabinohidroximato-5-fosfato (5 PAH), D-arabinoato-5-fosfato (5 PAA) e D-gluconato-6-fosfato. Figura adaptada de [20].

Valores de K_i dos inibidores 5 PAH, 5 PAA e 6 PGH, foram calculados por Hardré, R e coautores [20] para as enzimas PGI de levedura, músculo de coelho e para *Bacillus stearothermophilus*. Seus resultados indicam que para as três enzimas o melhor inibidor foi o 5 PAH, seguido pelo 5 PAA e por último o 6 PGH. O fato do 5 PAH apresentar uma composição química mais semelhante ao intermediário da reação deve ser uma das justificativas para seu alto efeito inibitório.

Em seguida ao analisar o desempenho dos inibidores 5 PAA e 6 PGH surge a hipótese de que a redução de seis para cinco carbonos resulte numa mudança no padrão de interação, principalmente com relação as possíveis pontes de hidrogênios estabelecidas entre a enzima e o carboxila na posição C-1 destes inibidores. Para esclarecer esta dúvida seria necessária a resolução da estrutura de um complexo entre o inibidor 5PAH e 5 PAA com a PGI de humano ou de coelho, para que se pudesse comparar os contatos realizados pelos inibidores.

CAPÍTULO 5

Cristalização da PGI-hum

5.1) Introdução

Cristalizar uma proteína significa passar de um estado desordenado em solução, onde as moléculas de proteína apresentam-se com um alto grau de liberdade, para um estado ordenado onde existe um padrão de interação de uma molécula com suas vizinhas, e esse padrão é mantido ao longo de todo cristal. Para uma proteína passar de um estado em solução para uma forma cristalina é necessário induzir a aproximação dessas moléculas preservando uma ordem de agregação, ou seja, sem que a proteína forme uma massa precipitada amorfa.

Na prática, para cristalizar uma proteína é necessário purificar a proteína até uma solução o mais homogênea possível, concentrá-la até valores entre 5 e 20 mg/ml e colocar pequenas amostras dessa solução na presença de agentes precipitantes. Esses agentes em geral sequestram a água que solvata a proteína induzindo-a a formar agregados cristalinos ou precipitados amorfos.

O processo de cristalização em geral é bastante sensível a variações de temperatura, pH, pressão e umidade. Portanto ao realizar ensaios de cristalização de uma proteína é ideal estabilizar o maior número desses fatores. Isso normalmente é obtido mantendo a proteína e o agente precipitante na presença de uma tampão e realizando os experimentos em câmaras de temperatura e humidade controladas.

Os experimentos iniciais geralmente são realizados combinando diferentes valores de pH, temperatura, tipos e concentrações de agentes precipitantes. A HAMPTON comercializa conjuntos de soluções - fatorial incompleto - que combinam diversos tampões e agentes precipitantes nos quais foram obtidos cristais

de proteínas já caracterizadas e outras condições mais que ajudam a representar uma distribuição no espaço de parâmetros a ser explorado. Esses conjuntos ficaram conhecidos como “fatoriais HAMPTON”.

Um dos métodos mais comuns para montar os ensaios de cristalização é denominado difusão de vapor em gota suspensa. Em geral, neste método uma amostra da solução de proteína, entre 2 a 6 μl , é misturada sobre uma lamínula siliconizada com um mesmo volume da solução contendo o agente precipitante e o tampão (solução precipitante). Em seguida esta lamínula servirá de tampa para um poço com aproximadamente 1 ml da solução precipitante. A mistura de proteína e solução precipitante fica suspensa sobre o poço. Graça de vácuo é usada para vedar a lamínula na borda do poço.

A gota de proteína mais solução precipitante apresenta-se diluída em relação a solução do poço. Uma vez que o sistema encontra-se vedado, as duas soluções irão entrar em equilíbrio pela transferência de água da gota suspensa para a solução no poço. A perda de água e conseqüente aumento da concentração de proteína e precipitante na gota induz a cristalização ou a formação de precipitado.

5.2) Métodos

O método de cristalização em todos os ensaios com a PGI-hum foi o difusão de vapor por gota suspensa, com uma única exceção onde foi utilizado o difusão de vapor em gota sentada, mas que não deu resultados promissores. A princípio foram testadas todas as condições dos fatoriais I e II da Hampton nas temperaturas de 4 e 18 °C. Em seguida as condições destes fatoriais que apresentaram indícios de cristalização foram refinadas realizando-se pequenas variações de pH, concentração do agente precipitante, concentração de proteína e volume da gota suspensa.

5.3) Resultados e discussão

Para o primeiro ensaio de cristalização foram montados os fatoriais I e II da HAMPTON a 4 e a 18 °C, com gotas de 3 μl de proteína a 6,2 mg/ml (mais 3 μl de sol. precipitante). Indícios de cristalização foram obtidos para a condição 46 do fatorial I (0,2M de acetato de Ca^{++} , 0,1M de cacodilato de Na^+ pH=6,5 e 18%

PEG8000) ; 30 do fatorial I (0,1M de HEPES pH=7,5 , 10% de PEG 6.000 e 5% de MPD) e 38 do fatorial II(0,1M de HEPES pH=7,5 e 20% de PEG 10.000), todas aparecendo em ambas as temperaturas.

As condições encontradas nos fatoriais I e II apresentavam polietilenoglicol (PEG) com diferentes massas molares médias – 6000, 8000 e 10000 Kda – como principal agente precipitante. O refinamento de tais condições resultou em um fatorial que pode ser montados em uma única caixinha de cristalização de matrix 4 por 6 (fig 19).

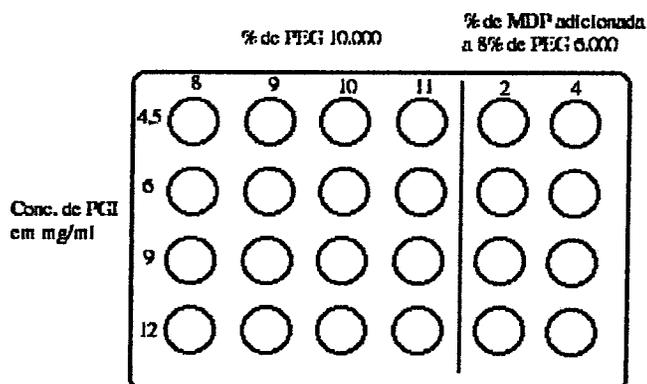


Figura 19: Caixinha de cristalização resumindo as melhores condições de cristalização obtidas para a PGI-hum. Cada círculo representa um poço de cristalização no qual combinava-se diferentes concentrações de agente precipitante (PEG 10.000 e MPD mais PEG 6.000) com diferentes concentrações de proteína (4.5 , 6, 9, 12 mg/ml).

Além do agente precipitante, cada condição do fatorial refinado (fig 19) continha o tampão 0.1 M HEPES em pH 7.5. Os melhores cristais foram obtidos com 7 dias de isolamento a temperatura de 18 °C.

Os cristais apresentam morfologia semelhante a paralelepípedo com crescimento acentuado em uma das direções e geralmente acumulam defeitos nas extremidades (fig.20).

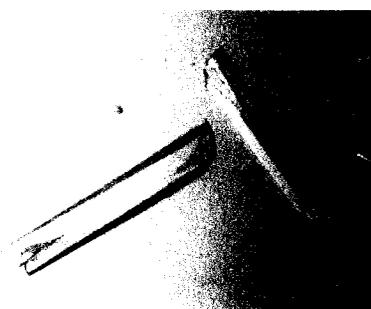


Figura 20: cristal da PGI-humana crescido em 0,1M de HEPES pH=7,5 , 10% de PEG10.000. Cristais de dimensões 0,1 x 0,1 x 0,6 mm (cristal do canto esquerdo).

CAPÍTULO 6

Coleta de dados de difração de raios-X dos cristais de PGI-hum

6.1) Introdução

O raio-X é uma onda eletromagnética que ao interagir com os elétrons de um sistema cristalino sofre um espalhamento cuja a resultante da interferência dos raios-X espalhados é conhecido como padrão de difração. Parte da energia do feixe eletromagnético que incide no cristal é absorvida pelos elétrons e depois emitida de forma elástica na mesma frequência da onda incidente, ou seja, os elétrons tornam-se fontes espalhadoras de raios-X de mesma frequência e comprimento de onda da fonte primária [21].

Um cristal é considerado um arranjo periódico no qual a *cela unitária* é a menor unidade que se repete ao longo de três direções **a**, **b** e **c**. Qualquer átomo dentro de uma cela unitária pode ser descrito por um vetor $\mathbf{r} = x.\mathbf{a} + y.\mathbf{b} + z.\mathbf{c}$, onde **a**, **b** e **c** são vetores cujo módulo equivale ao tamanho de uma aresta da cela unitária (ou parâmetro de rede). O espaço descrito por este conjunto de vetores é denominado espaço direto e x, y e z são coordenadas fracionárias.

O padrão de difração é o resultado da interferência das várias ondas espalhadas por todos os átomos do cristal [22] e pode ser descrito como o resultado da soma destas ondas. Para facilitar esta soma, as ondas de mesma frequência (ou comprimento) são representadas por um vetor num sistema de dois eixos (fig 21).

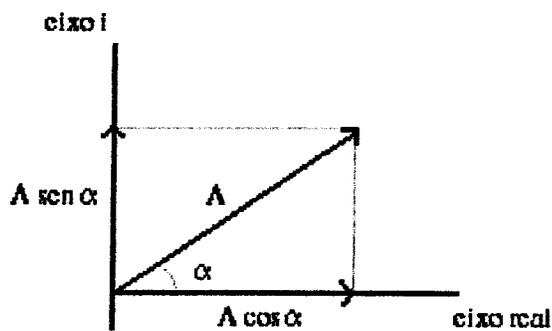
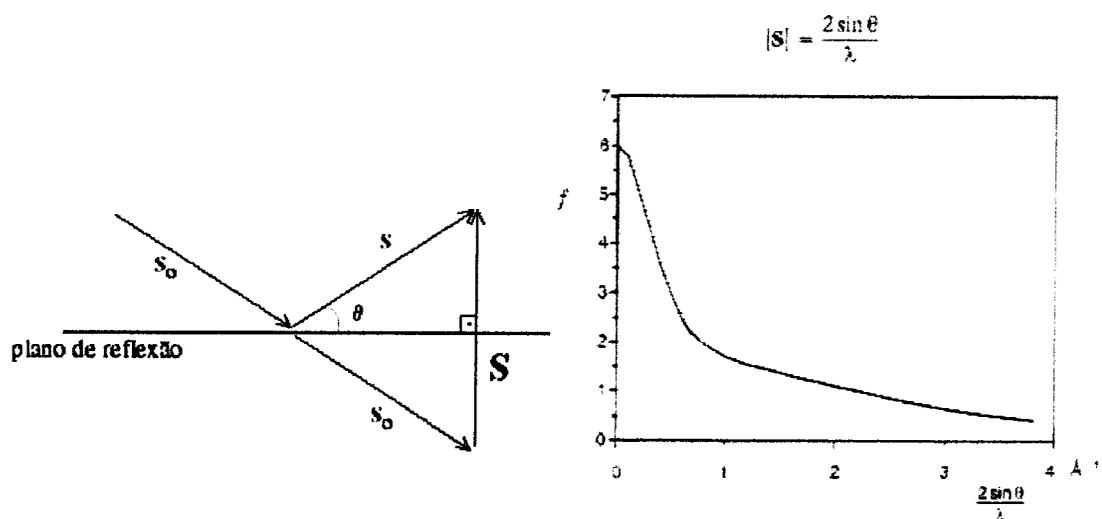


Figura 21 : Diagrama de Argand. O vetor A representa uma onda com componentes $A \cos \alpha$ no eixo real e $A \sin \alpha$ no eixo imaginário. O comprimento do vetor (ou amplitude da onda) é dado pelo $|A|$.

Dessa forma o vetor estabelece um ângulo α com o eixo real equivalente a sua fase em relação a onda (ou feixe) incidente. O somatório de várias ondas pode ser realizado somando-se as componentes no eixo real e as do eixo imaginário separadamente. Ainda assim seria muito trabalhoso somar todas as ondas provenientes do espalhamento dos átomos de um cristal. Este problema é simplificado por uma conveniência matemática do tipo $A \cos \alpha + A i \sin \alpha = A \exp[i\alpha]$.

Durante o espalhamento de raios-X por um cristal, um átomo cuja a posição é descrita por um vetor r em relação a uma origem arbitrária, absorve o feixe incidente (s_0 de comprimento $1/\lambda$) e passa a emitir raios-X em todas as direções. Existe uma direção em particular representada por um vetor s , cujo vetor diferença S entre o feixe espalhado s e o feixe incidente s_0 é perpendicular a um plano de reflexão



imaginário (fig. 22a).

(A)

(B)

Figura 22: (A) feixe espalhado por um cristal, representado por s , pode ser descrito como uma reflexão do feixe s_0 com relação ao plano de reflexão perpendicular a S . (B) fator de espalhamento

atômico para o átomo de carbono em função $(2\text{sen}\theta) / \lambda$. f é expressado com o número de elétrons quando o feixe apresenta $\theta = 0$.

Este plano “reflete” o feixe incidente com um ângulo θ igual ao ângulo de incidência obedecendo a lei de Bragg (descrita adiante).

O espalhamento de um único átomo dentro de uma cela unitária é dado por

$$f = f \exp[2\pi i \mathbf{r} \cdot \mathbf{S}]$$

onde f é o fator de espalhamento atômico, e \mathbf{r} representa a posição do átomo em relação a origem. A intensidade do espalhamento de um único átomo depende do módulo de \mathbf{S} que é igual a $(2\text{sen}\theta) / \lambda$ (fig.22b). Considerando que em uma cela unitária existem vários átomos, o espalhamento total do conteúdo de uma cela unitária corresponde ao somatório do espalhamento de cada átomo presente. Este espalhamento é dado por

$$F(\mathbf{S}) = \sum_j f_j \exp[i2\pi \mathbf{r}_j \cdot \mathbf{S}]$$

onde $F(\mathbf{S})$ é denominado *fator de estrutura* e j são os átomos da cela unitária.

Empregando os vetores \mathbf{a} , \mathbf{b} e \mathbf{c} para descrever a posição de uma cela unitária dentro do cristal, qualquer cela passa a apresentar um posição definida por $\mathbf{R} = u \cdot \mathbf{a} + v \cdot \mathbf{b} + w \cdot \mathbf{c}$, onde u , v e w são números inteiros. A onda total espalhada pelo cristal $K(\mathbf{S})$ pode ser obtida pelo somatório do espalhamento de todas as celas unitárias descritas por este sistema de coordenadas.

Considere o espalhamento de uma única cela unitária $k(\mathbf{S})$,

$$k(\mathbf{S}) = F(\mathbf{S}) \exp(i2\pi \mathbf{R} \cdot \mathbf{S}),$$

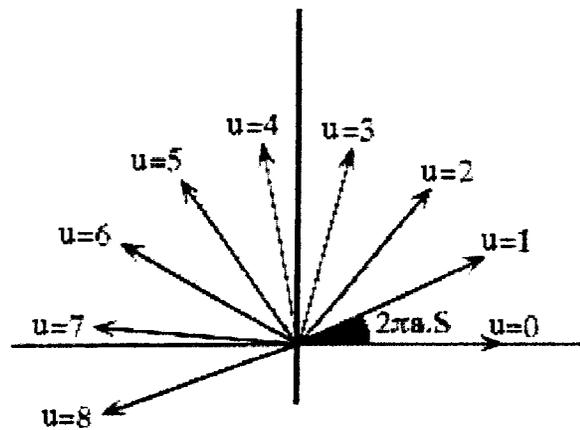
onde o termo $F(\mathbf{S})$ é o fator de estrutura, ou seja, o espalhamento de todos os átomos da cela unitária, e o termo $2\pi \mathbf{R} \cdot \mathbf{S}$ corresponda o ângulo de fase, ou seja, a posição desta cela em relação o origem do cristal e ao feixe incidente. O espalhamento total do cristal $K(\mathbf{S})$, seria obtido por

$$\begin{aligned} K(\mathbf{S}) &= F(\mathbf{S}) \exp i2\pi [(u \cdot \mathbf{a} + v \cdot \mathbf{b} + w \cdot \mathbf{c}) \cdot \mathbf{S}] \\ &= F(\mathbf{S}) \times \exp i2\pi u \cdot \mathbf{a} \cdot \mathbf{S} \times \exp i2\pi v \cdot \mathbf{b} \cdot \mathbf{S} \times \exp i2\pi w \cdot \mathbf{c} \cdot \mathbf{S} \end{aligned}$$

$$K(\mathbf{S}) = F(\mathbf{S}) \times \sum_{u=1}^{n_1} \exp(i2\pi u \vec{a} \cdot \vec{S}) \times \sum_{v=1}^{n_2} \exp(i2\pi v \vec{b} \cdot \vec{S}) \times \sum_{w=1}^{n_3} \exp(i2\pi w \vec{c} \cdot \vec{S})$$

onde n_1 , n_2 e n_3 equivalem ao número de celas unitárias na direção \mathbf{a} , \mathbf{b} e \mathbf{c} , respectivamente. Pela representação de Argand do espalhamento na direção \mathbf{a} , (fig

23) fica claro que a soma de todos esses vetores é igual a zero, uma vez que u é um número muito grande, todo vetor espalhado numa dada direção apresenta um igual



porém com direção oposta. Dessa forma ao somar todos estes vetores, o resultado seria nulo. Isto vale para as outras duas direções.

Figura 23: Diagrama de Argand para oito vetores da direção \mathbf{a} , cada um representando o espalhamento de uma cela unitária. Cada vetor apresenta módulo 1 e fase $2\pi u \mathbf{a} \cdot \mathbf{S}$.

A única maneira de $\mathbf{K}(\mathbf{S})$ ser diferente de zero é quando o produto escalar $(\mathbf{a} \cdot \mathbf{S})$ for um número inteiro qualquer. $(\mathbf{a} \cdot \mathbf{S})$ sendo inteiro significa que o espalhamento de todas as celas unitárias em \mathbf{a} apresentam a mesma direção. Isso é facilmente percebido tratando $\mathbf{K}(\mathbf{S})$ como uma função de sen e cos. Os termos contendo $\text{sen}(i2\pi \mathbf{a} \cdot \mathbf{S})$ igualam-se a zero e saem da expressão, ficando apenas os termos de $\text{cos}(2\pi \mathbf{a} \cdot \mathbf{S})$. Estes por sua vez são múltiplos de $\text{cos}(2\pi)$ e ao realizar o somatório do n_i de u , o resultado será o próprio u , ou seja, o número de celas unitárias presentes na direção de \mathbf{a} . Esta condição também é imposta as outras duas direções de forma que para haver difração por um cristal é preciso que: $\mathbf{a} \cdot \mathbf{S} = h$, $\mathbf{b} \cdot \mathbf{S} = k$ e $\mathbf{c} \cdot \mathbf{S} = l$, onde h , k e l são números inteiros. Estas condições são denominadas como *condições de Laue*.

$$K(S) = F(S) \times \left[\left(\sum_u^{n_1} \cos(2\pi u \vec{a} \vec{S}) + \sum_u^{n_1} \text{sen}(i2\pi u \vec{a} \vec{S}) \right) \times \left(\sum_v^{n_2} \cos(2\pi v \vec{b} \vec{S}) + \sum_v^{n_2} \text{sen}(i2\pi v \vec{b} \vec{S}) \right) \right. \\ \left. \times \left(\sum_w^{n_3} \cos(2\pi w \vec{c} \vec{S}) + \sum_w^{n_3} \text{sen}(i2\pi w \vec{c} \vec{S}) \right) \right]$$

$$K(S) = F(S) \times \left[\left(\sum_u^{n_1} \cos(2\pi u \vec{a} \vec{S}) \right) \times \left(\sum_v^{n_2} \cos(2\pi v \vec{b} \vec{S}) \right) \times \left(\sum_w^{n_3} \cos(2\pi w \vec{c} \vec{S}) \right) \right]$$

$$K(S) = F(S) \times u \times v \times w$$

Seguindo este raciocínio, conclui-se que a amplitude de onda espalhada por um cristal depende do fator de estrutura $F(S)$ e do número de células unitárias do cristal (u.v.z).

As condições de Laue impõem que o produto escalar entre os vetores \mathbf{a}/h e \mathbf{S} é igual a 1, ou seja, a projeção de \mathbf{a}/h sobre \mathbf{S} , multiplicada por $|\mathbf{S}|$ é igual a 1. Como isso é verdade para as outras duas direções, a extremidade dos vetores \mathbf{a}/h , \mathbf{b}/k e \mathbf{c}/l definem um plano perpendicular ao vetor \mathbf{S} . Portanto, o feixe espalhado pode ser visualizado como um feixe refletido por um plano de reflexão $(h k l)$. A diferença de fase de feixes refletidos por sucessivos planos corresponde a $2d \text{sen} \theta / \lambda$. Pela condição de Laue, a projeção \mathbf{a}/h sobre o vetor \mathbf{S} é igual a $1/|\mathbf{S}|$ e corresponde a distância entre dois planos (fig 24), ou seja $d = 1/|\mathbf{S}|$. Sabe-se também que $|\mathbf{S}| = 2 \text{sen} \theta / \lambda$, portanto combinando estas duas relações obtém-se a relação conhecida como Lei de Bragg: $\lambda = 2d \text{sen} \theta$.

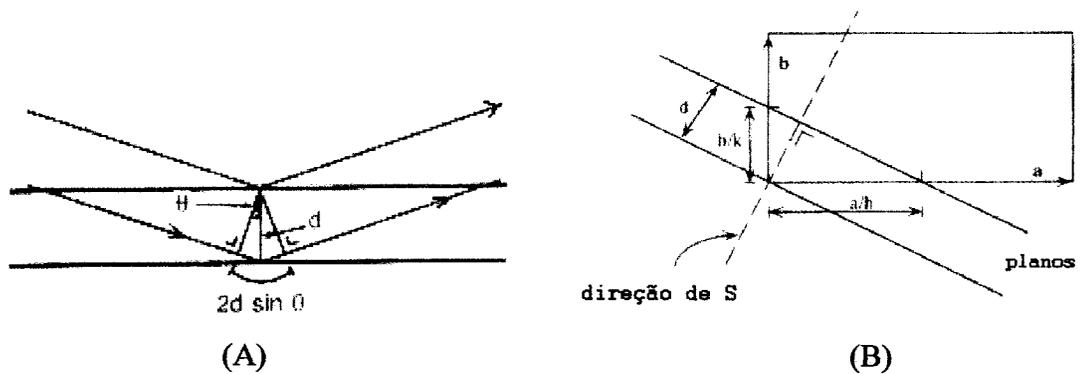


Figura 24: (A) diferença de fase entre feixes refletidos por planos paralelos. (B) Cella unitária bi-dimensional, mostrando os vetores \mathbf{a} , \mathbf{b} e a direção do feixe espalhado perpendicular ao plano formado pelos vetores \mathbf{a}/h e \mathbf{b}/k .

O plano $(h k l)$ de índices $h = 1, k = 0$ e $l = 0$ apresenta o vetor espalhamento $S(100)$ perpendicular a este plano e apresenta um módulo igual a $1/d(100)$. Este vetor é denominado a^* e apresenta correspondentes b^* e c^* definidos pelos vetores $S(010)$ e $S(001)$, respectivamente. Estes três vetores podem ser combinados independentemente para descrever qualquer reflexão $S(h k l)$ com módulo equivalente a $1/d(h k l)$. Portanto as reflexões de uma rede cristalina definidas pelos vetores a, b e c , são descritas por uma rede com vetores base a^*, b^* e c^* , denominada rede recíproca. A rede direta e a rede recíproca estão inversamente relacionadas, isto implica que quanto maior os vetores de base da rede direta, menor o tamanho dos vetores de base que descrevem a rede recíproca.

Um conjunto de dados de difração ou espalhamento de raios-X por um cristal de proteína consiste em obter uma amostragem do espaço recíproco, ou seja, o conjunto de dados é composto pela intensidade I , um desvio associado $\sigma(I)$ e pelos índices $(h k l)$ de cada reflexão medida.

Para facilitar a compreensão da rede recíproca e de como ocorre o evento de difração é comum utilizar uma representação geométrica denominada construção de Ewald (fig 25). Esta consiste de uma esfera de raio $1/\lambda$ construída de forma que a origem da rede recíproca esteja centrada na direção do feixe incidente e localizada sobre a circunferência da esfera. O centro da esfera é o ponto onde o feixe incidente é espalhado e a reflexão ocorre em direção a qualquer ponto da rede recíproca que esteja localizado sobre a superfície da esfera, esta condição de espalhamento está de acordo com a Lei de Bragg.

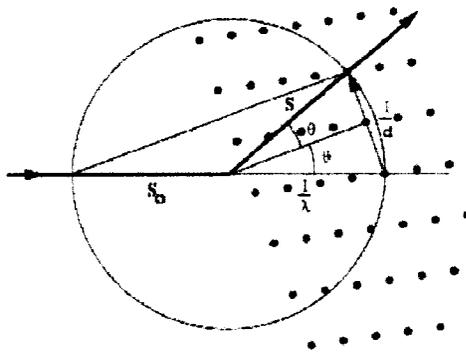


Figura 25: Construção de Ewald mostrando pontos da rede recíproca, a esfera de raio $1/\lambda$ e o espalhamento do feixe incidente s_0 em direção a um ponto sobre a circunferência da esfera. O vetor espalhamento $S(h k l)$ tem origem no ponto $(0 0 0)$ da rede recíproca, módulo $1/d$ e índices $h k l$ dados pelo ponto $(h k l)$ localizado sobre a esfera de Ewald.

Dessa forma a reflexão $S(h\ k\ l)$ é o vetor que sai da origem da rede recíproca até o ponto de índice $(h\ k\ l)$ localizado sobre a esfera. A rede recíproca gira exatamente como o cristal, assim para os demais pontos da rede passarem pela circunferência da esfera, ou seja, entrarem em condição de difração, basta rotacionar o cristal em relação a fonte de raios-X.

Para a coleta de dados de um cristal de proteína pode-se utilizar uma estratégia conhecida como método de oscilação[23]. Este método consiste em rotacionar o cristal em relação ao feixe de raios-X incidente com um ângulo de oscilação φ definido (fig 26).

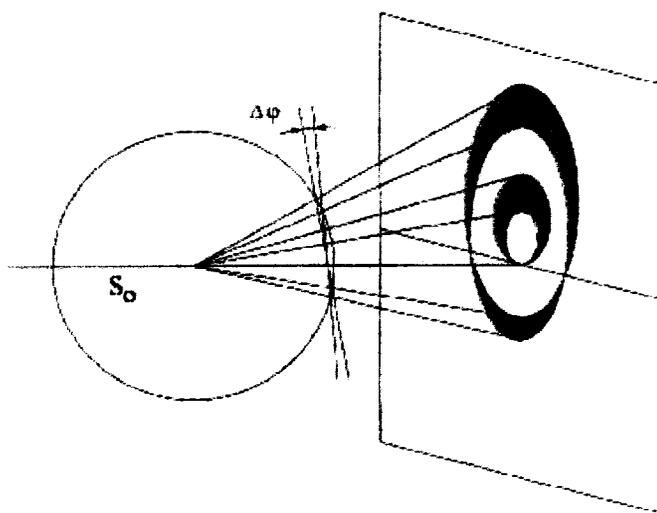


Figura 26: Construção de Ewald ilustrando o método de oscilação. Quando é rotacionado de um ângulo φ , pontos da rede recíproca pertencentes a um mesmo plano entram em condição de difração gerando um padrão de difração em faixas semelhantes a semi-luas.

A partir de um ângulo inicial, o cristal permanecerá oscilando um certo grau φ de oscilação, durante o qual serão coletados os pontos de difração correspondentes aos planos cristalinos que estiverem em condições de difração. Para a próxima imagem, o ângulo inicial de oscilação será o equivalente ao ângulo inicial da imagem anterior acrescido de φ graus.

Para registrar os padrões de difração de raio-X de uma proteína costuma-se utilizar como anteparo detectores do tipo placas de imagens (“image plates”). Este tipo de detector consiste de uma placa com uma fina camada de um composto inorgânico dotado de centros de cor, que são capazes de armazenar energia provenientes de fótons. Após capturar os fótons espalhados pelo cristal, esta placa é varrida por uma feixe de laser vermelho que resulta na emissão de luz azul pelos

centros de cor energizados. Dessa forma se tem a posição e intensidade do feixe difratado pelo cristal.

6.2) Métodos

Para a coleta de dados de difração dos cristais obtidos da PGI-hum foram utilizadas fontes de raios-X provenientes de um anodo rotatório de cobre, montada no Lab. de cristalografia do IFSC-USP, e outra proveniente de uma fonte de luz sincrotron, montada na linha de Cristalografia de Proteínas do Lab. Nacional de Luz Sincrotron (LNLS) em Campinas. As principais vantagens de realizar uma coleta de dados no LNLS está no aumento da intensidade do feixe incidente e numa melhor colimação e menor divergência do mesmo. Esses aspectos irão refletir num padrão de difração com maior resolução e melhor definição dos pontos e menor tempo de coleta por imagem.

A coleta realizada no IFSC-USP utilizou um anodo rotatório de cobre ($\lambda=1,5406 \text{ \AA}$) (RIGAKU) como fonte de raios-X, operando a 50 KV e 100 mA e um detector de imagens tipo RAXIS-IIC, a uma distância de 130 mm do cristal. A distância do detector ao cristal limitou a coleta de dados até uma resolução máxima de 2.45 \AA . O ângulo de oscilação foi de 1° e cada imagem coletada foi exposta por 30 minutos.

Para a coleta no LNLS, foi selecionado um comprimento de onda (λ) de 1.5450 \AA , e o detector de imagem do tipo MAR345 foi posicionado a 170mm do cristal, o que permitia coletar pontos até 2.0 \AA de resolução. Cada imagem foi obtida com $0,75^\circ$ de oscilação e 4 minutos de exposição.

Tanto para coletas no LNLS, quanto para coletas no Lab. de cristalografia do IFSC-USP, os cristais foram obtidos na condição 0,1 M de HEPES pH=7,5 com 10% de PEG10.000 e solução de proteína a 8 mg/ml. Os cristais foram montados em "loops" de menos de 1 mm de diâmetro máximo e mantidos constantemente em fluxo de nitrogênio gasoso. Este fluxo é soprado sobre o cristal com uma temperatura controlada de 100 K. Para impedir a formação de gelo (cristais de água) durante o resfriamento do cristal, a solução que solvata o cristal de proteína deve ser substituída por uma solução crio protetora. Esta solução consiste de uma mistura de água com glicerol, etileno-glicol, MPD ou outros compostos orgânicos, geralmente

uma concentração de 20 a 40%. No caso da PGI-hum, a solução crioprotetora era composta pela mesma solução do poço com 20 % de MPD.

6.3) Resultados e discussão

Para o cristal montado no IFSC-USP, foram coletadas um total de 79 imagens de oscilação, o que corresponde a 79° da rede recíproca. O conjunto de dados coletados no LNLS é constituído de 187 imagens de oscilação, totalizando 140° da rede recíproca.

Uma vez que o cristal é mantido dentro de um fluxo de nitrogênio gasoso a temperatura próxima de 100K, os danos causados pela radiação ionizante incidente são minimizados. Isso permite que o cristal seja exposto ao raio-X por um longo período de tempo. A coleta de dados no IFSC-USP só não varreu uma espaço maior da rede recíproca devido a interrupção do fluxo de nitrogênio durante o reabastecimento dos reservatórios de nitrogênio.

No LNLS, apesar do conjunto de dados bem amostrado, o fator limitante é o tempo de feixe disponível para o usuário realizar seus experimentos. Gasta-se parte deste tempo, para escolha e montagem do melhor cristal dentre vários daqueles que são levados no dia do experimento. A realização de uma análise preliminar da qualidade dos cristal em uma fonte externa, por exemplo no próprio IFSC-USP, permite uma economia de tempo de feixe que passa a ser usada para uma coleta maior de dados. Esta análise preliminar só foi possível após a aquisição de um *kit* (HAMPTON) para congelamento e transporte de cristais em reservatório de nitrogênio líquido. Dessa forma, os melhores cristais obtidos em um ensaio de cristalização eram analisados quanto ao limite de resolução a que difratavam e a qualidade dos pontos de difração; em seguida eram congelados para serem transportados ao LNLS no dia marcado para o experimento.

Uma vez de posse dos dados de difração, a próxima etapa consiste em analisar estatisticamente a qualidade destes dados e determinar parâmetros cristalográficos como grupo espacial, parâmetros de rede e o conteúdo da cela unitária.

CAPÍTULO 7

Análise qualitativa dos conjuntos de dados de difração de raios-X

7.1) Introdução

Para interpretar as informações do espaço recíproco é necessário ter conhecimento dos aspectos de simetria intrínsecos a um cristal. As relações de simetria entre as moléculas de um cristal tem uma importância prática na determinação da estrutura cristalina uma vez que, reduz o número de parâmetros únicos a serem determinados.

São sete os sistemas cristalinos [22] e são necessários seis parâmetros para descrever a cela unitária: três eixos a , b e c e três ângulos interaxiais α , β e γ (fig.27).

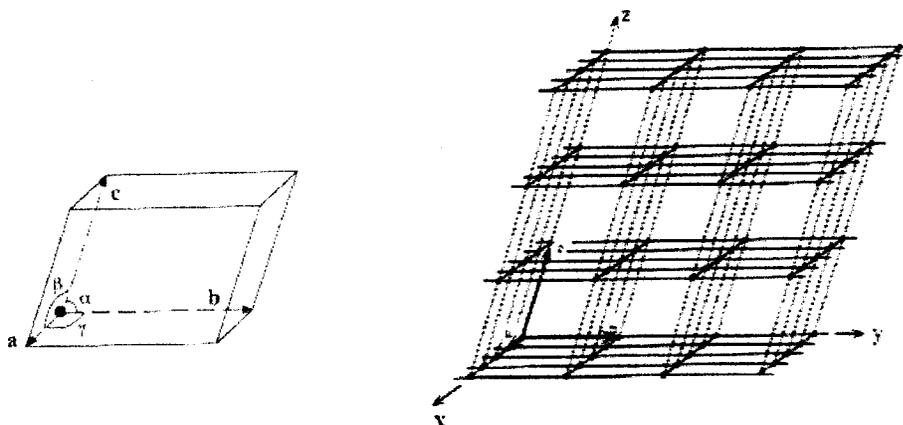


Figura 27: Representação de uma cela unitária (a esquerda) mostrando os três parâmetros de rede a , b e c e três ângulos interaxiais α , β e γ . A direita, o empacotamento cristalino.

Uma vez determinado o sistema cristalino, deve-se caracterizar o tipo de *rede de Bravais* em função da posição de pontos no interior da malha cristalina. Uma rede primitiva (P), apresenta pontos apenas nos vértices das celas; a presença de um ponto

adicional ao centro de uma das faces da cela cria a rede centrada que é representada pelo nome do eixo que intercepta a face; se este ponto no entanto estiver no centro da cela a rede é denominada corpo centrado (I) e caso existam pontos em todas as faces, a rede passa a ser representada pela letra F (fig. 28)

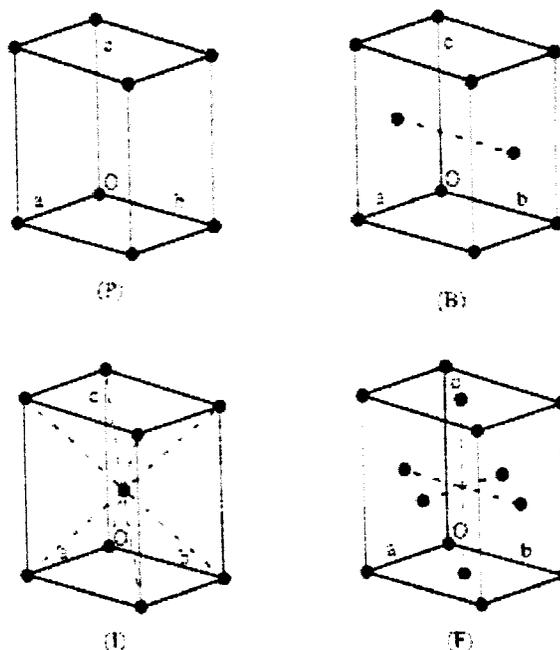


Figura 28: Representação das celas: primitiva (P), face centrada em na direção de eixo b (B), corpo centrado (I) e todas as faces centradas (F).

Por último defini-se o grupo espacial descrevendo as operações de simetria que relacionam as moléculas dentro da rede cristalográfica. Estas operações são de denominadas *operações de simetria cristalográfica*, enquanto as operações de simetria não cristalográfica descrevem a relação entre moléculas dentro da unidade assimétrica. A unidade assimétrica é a menor fração da cela unitária capaz de reproduzir todo o cristal apenas pela a aplicação das operações de simetria cristalográfica.

Para cristais de proteína, além das translações simples, são permitidos apenas operações de simetria descritas por eixos de rotação e roto-translação; pontos de inversão e planos de reflexão não são possíveis uma vez que o enantiômero de uma proteína deveria ser formados por D-aminoácidos. Isso não é encontrado na natureza.

Para se distinguir entre um eixo de rotação e um eixo de roto-translação deve-se analisar as *extinções sistemática* no padrão de difração. Por exemplo, os eixos de

roto-translação com operações de translação de meia unidade de cela unitária (2_1) geram a ausência de reflexões de índices $(h\ 0\ 0)$, $(0\ k\ 0)$ e $(0\ 0\ l)$ quando os índices h , k e l são ímpares, respectivamente [24].

A determinação dos parâmetros de rede e do grupo espacial fornece informações sobre a estrutura do cristal, mas não da proteína que o constitui. A estrutura da proteína será determinada a partir do conjunto de dados usados para definir estes parâmetros. A qualidade do conjunto de dados irá influenciar na qualidade final da estrutura determinada para a proteína. Portanto, a análise da qualidade dos dados coletados é uma etapa que indicará se deve-se prosseguir com a determinação da estrutura da proteína, ou deve-se retornar alguns passos e obter um conjunto de dados melhor a partir da difração de um novo cristal.

7.2) Métodos

A análise das imagens coletadas para os conjuntos de dados da PGI-hum foi realizada com o pacote de programas *HKL* [25]. Este pacote inclui três módulos, *xdisplay* para visualizar as imagens, *denzo* para indexar e integrar cada imagem e *scalepack* para escalonar e refinar os parâmetros do cristal.

Durante a indexação e integração de cada imagem o programa *denzo* prediz os parâmetros de cela unitária e define para cada ponto do padrão os índices h, k, l , I e $\sigma(I)$; onde I e $\sigma(I)$ referem-se a intensidade e o desvio para cada reflexão h, k, l . Essa predição é realizada a partir da distribuição dos pontos de uma imagem, da distância do detector ao cristal, do λ e da posição do feixe. De posse dos parâmetros de cela unitária, o programa lista os possíveis sistemas cristalinos que melhor se ajustam aos parâmetros observados.

Inicialmente, o *scalepack* definirá fatores de escala que serão aplicados às I 's e $\sigma(I)$'s de cada filme de forma que reflexões coletadas em diferentes filmes possam ser comparadas entre si. Em seguida o programa rejeitará todas as reflexões que falharem em um teste de $I/\sigma(I)$. Este teste tem como objetivo eliminar reflexões cuja intensidade medida I , não seja significativamente maior do que a intensidade de fundo ("background"), isto é, $\sigma(I)$.

A análise qualitativa de um conjunto de dados realizada pelo *scalepack* pode ser avaliada principalmente por parâmetros como: completudeza, redundância, I/σ , R_{merge} , χ^2 e o limite de resolução.

Para obter detalhes da estrutura de uma proteína é necessário trabalhar com resoluções maior do que 2.5 Å. Porém, para se determinar até que limite de resolução um conjunto de dados pode ser utilizado deve-se avaliar, em ordem de importância, os seguintes parâmetros: a relação sinal ruído (I/σ), em média quantas vezes estas reflexões foram coletadas (redundância) e a soma dos desvios da média das intensidades dentro de uma faixa de resolução (R_{merge}). Valores aceitáveis para esses parâmetros seriam: $I/\sigma > 0$, redundância > 2 e $R_{\text{merge}} < 0.400$.

O χ^2 é a razão média quadrada entre o erro posicional do ajuste feita para cada reflexão e o erro esperado, desta forma ele avalia a qualidade dos ajustes realizados durante a integração das imagens. χ^2 deve ser mantido entre 2 e 1 para que não interfira em cálculos subsequentes. A completudeza indica quantas reflexões foram coletadas em relação ao total de reflexões que poderiam ser coletadas de um cristal. Quanto maior a completudeza mais dados estarão disponível para o refinamento do modelo.

7.3) Resultados e discussão

Apesar de ter coletado reflexões até 2.45 Å, conjunto de dados coletado no IFSC-USP apresenta bons valores de χ^2 e um estatística boa para um limite de resolução de 2.6 Å.

Tabela V: Resumo da análise qualitativa para os conjuntos de dados coletados com uma fonte de raios-X tipo anodo rotatório de cobre montado no IFSC-USP e coletados com feixe proveniente de luz sincrotron na linha de Cristalografia de Proteínas do LNLS, Campinas-SP.

Data-set	IFSC-USP	LNLS
Limite de resolução (Å)	2.45	2.1
Reflexões totais	194074	672976
Reflexões únicas	83143	138444
Completudeza(última faixa) (%)	93.7 (84.0)	99.6 (99.4)
R_{merge} * (última faixa) (%)	10.8 (36.7)	11.1 (39.9)
Redundância	2.3	4.8
I/σ	6.3	6.7

* $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$

Apesar de ter um conjunto de dados aceitáveis até 2.6 Å de resolução, o limite de difração máximo indicava que seria possível obter um conjunto de dados com uma melhor estatística e maior limite de resolução se a fonte de raios-X fosse proveniente de luz SINCROTRON. Portanto novos cristais foram crescidos repetindo-se as melhores condições de cristalização. Doze cristais foram previamente analisados no IFSC-USP e congelados em nitrogênio líquido para serem transportados ao LNLS, Campinas-SP. No dia do experimento no LNLS, o melhor cristal foi montado na linha de Cristalografia de Proteínas e um novo conjunto de dados foi coletado otimizando assim o tempo disponível para a coleta.

A otimização do tempo de coleta resultou num conjunto de dados com maior redundância e completeza do que o dados coletados anteriormente. E a facilidade de escolher previamente e transportar apenas os melhores cristais, resultou na escolha de um cristal que difratou com qualidade até 2.1 Å de resolução.

Após integrar e escalonar todas as 187 imagem do conjunto de dados coletados no LNLS, o pacote de programas *HKL* apontou as distorções de rede levando a concluir que tratava-se de um cristal ortorrômbico primitivo, com parâmetros da cela unitária $a=80.650$, $b=107.960$ e $c=270.888$ Å e ângulos $\alpha=\beta=\gamma=90^\circ$. Analisando o padrão de extinções sistemáticas dos dois conjuntos de dados foi possível definir o grupo espacial ortorrômbico primitivo ($P2_12_12_1$) com eixos de roto-translação 2_1 nas três direções **a**, **b** e **c**.

CAPÍTULO 8

Substituição Molecular

8.1) introdução

A intensidade do feixe difratado $I(h\ k\ l)$ registrada no conjunto de dados, é proporcional ao quadrado do módulo do fator de estrutura, ou seja, aplicando-se os fatores de correção apropriados, tem-se que,

$$I(h\ k\ l) \cong |F(h\ k\ l)|^2 = |F(S)|^2$$

Ao invés de considera o fator de estrutura como um somatório discreto sobre todos os átomos de uma cela unitária, pode-se representá-lo pela soma integral de toda densidade eletrônica contínua da cela unitária. Dessa forma,

$$F(S) = \sum_j f_j \exp[2\pi i \mathbf{r}_j \cdot \mathbf{S}]$$

pode ser representado por

$$F(S) = \int_{\substack{\text{célula} \\ \text{unitária}}} \rho(\vec{r}) \cdot \exp[i2\pi \vec{r} \cdot \vec{S}] \cdot dv$$

onde $dv = dx \cdot dy \cdot dz$, x , y e z são coordenadas fracionárias, V é o volume da cela unitária e $\rho(\mathbf{r})$ é a densidade eletrônica na posição \mathbf{r} . Considerando $\mathbf{r} \cdot \mathbf{S} = x \cdot h + y \cdot k + z \cdot l$, pode-se escrever

$$F(hkl) = \int_{x=0}^1 \int_{y=0}^1 \int_{z=0}^1 \rho(xyz) \cdot \exp[i2\pi(x \cdot h + y \cdot k + z \cdot l)] \cdot dx \cdot dy \cdot dz$$

Porém, o interessante para a cristalografia de proteínas é calcular a densidade eletrônica para toda posição x , y e z da cela unitária. Partindo da equação anterior de

$F(h k l)$, é possível obter a $\rho(x y z)$ utilizando a *Transformada de Fourier*. Esta é uma importante ferramenta utilizada pelos cristalógrafos para relacionar o espaço recíproco com a rede cristalina no espaço direto.

$$F(hkl) = \int_{x=0}^1 \int_{y=0}^1 \int_{z=0}^1 \rho(xyz) \cdot \exp[i2\pi(x.h + y.k + z.l)] \cdot dx dy dz$$

Espaço recíproco



Transformada de Fourier

$$\rho(xyz) = \frac{1}{V} \cdot \sum_h \sum_k \sum_l F(hkl) \cdot \exp[-i2\pi(x.h + y.k + z.l)]$$

Espaço direto

Substituindo, $F(h k l)$ por $|F(h k l)| \cdot \exp[i\alpha]$, obtém-se,

$$\rho(xyz) = \frac{1}{V} \cdot \sum_h \sum_k \sum_l |F(hkl)| \cdot \exp[-i2\pi(x.h + y.k + z.l) + \alpha(hkl)]$$

onde $\alpha(h k l)$ corresponde ao conjunto de fases iniciais. Apesar de bem estabelecida as relações entre o espaço direto e o recíproco, o grande problema dos cristalógrafos é que para obter a densidade eletrônica em qualquer ponto da cela unitária $\rho(x y z)$ eles contam apenas com as intensidades das reflexões e não as fases $\alpha(h k l)$. Este problema é conhecido como o *problema das fases*.

Uma forma de encontrar as fases iniciais $\alpha(h k l)$ consiste em usar as coordenadas estruturais de uma proteína homóloga a aquela que se pretende resolver com o conjunto reflexões coletado. Esta técnica é conhecida como método da *substituição molecular*[26]. A ferramenta que permite usar um conjunto de fases iniciais emprestadas de outra proteína é o cálculo e operação dos mapas de Patterson.

A função de Patterson $P(\mathbf{u})$ ou $P(u v w)$ é uma soma de Fourier definida no espaço direto, com coeficiente igual a intensidade das reflexões mas desconsiderando os ângulos de fase

$$P(\mathbf{u}) = \frac{1}{V} \cdot \sum_s |F(S)|^2 \cdot \cos[2\pi\vec{u}\vec{S}]$$

Escrita desta forma, o mapa de patterson pode ser obtido para uma proteína de estrutura desconhecida a partir dos valores de $|F(S)|^2$ obtidas do conjunto de dados. Outra forma de escrever a função de Patterson é

$$P(u) = \int_{\vec{r}_1} \rho(\vec{r}_1) \times \rho(\vec{r}_1 + \vec{u}) . dv$$

O significado físico desta expressão é : para $P(u)$ ser diferente de zero, deve existir densidade eletrônica num ponto r_1 e também em outro ponto com deslocamento igual a u a partir de r_1 . Assim o produto das densidades $\rho(r_1)$ e $\rho(r_1 + u)$ será diferente de zero. Caso não exista um átomo que confira densidade eletrônica ao ponto $(r_1 + u)$, a função neste ponto será nula e , conseqüentemente $P(u) = 0$. Isso significa que u representa o vetor entre dois átomos qualquer da cela unitária (fig. 29).

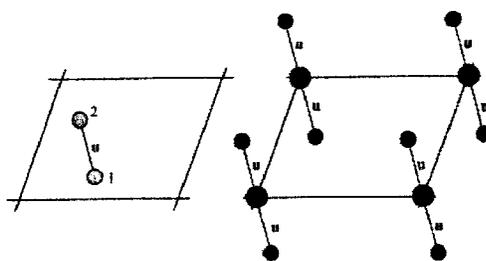


Figura 29: (a) representa uma cela unitária bi-dimensional com dois átomos separados por u . (b) Os vetores de patterson da cela unitária.

Calculando um mapa de Patterson a partir da intensidade das reflexões obtidas para uma proteína e outro mapa com base apenas na posição dos átomos de uma proteína homóloga, é possível sobrepor os dois mapas e obter a orientação da proteína homóloga com relação a cela unitária da proteína de interesse. Realizando esta operação, está se posicionando a proteína homóloga dentro da cela unitária da proteína alvo do estudo cristalográfico, dessa obtém-se um conjunto de fases iniciais para esta proteína.

Para realizar o posicionamento do modelo na cela unitária da nova molécula são aplicadas uma matriz rotacional, parametrizada pelos ângulos eulerianos α , β e γ ; seguida de uma translacional em coordenadas fracionárias x , y e z sobre o mapa de Patterson do modelo, de forma que este sobreponha o mapa da nova molécula. Para validar a qualidade da solução encontrada utiliza-se um coeficiente de correlação (C) entre as funções de patterson do modelo de busca e da estrutura alvo.

8.2) Métodos

O programa utilizado para realizar a substituição molecular da PGI-humana foi o *AMoRe (A-utomatic Mo-lecular Re-placement)*[26]. Foram utilizadas as coordenadas cartesianas do dímero de PGI de coelho, com 92% de identidade (fig.30) com a PGI-hum, cristalizada no grupo espacial ortorrômbico C222₁ e resolvida até 2.5 Å de resolução (código PDB: 1dqr).

```

rab PGI 1 AALTRNPQFQ KLOQWHREHG SELNLRHLEFD TDKERFNHFS LTLNTNHGHI
hum PGI 1 AALTRDPQFQ KLQQWYREHR SELNLRRLFD ANKDRFNHFS LTLNTNHGHI
*****
rab PGI 51 LLDYSKNLVT EEVMHMLLDL AKSRGVEAAR ESMFNGEKIN STEDRAVLHV
hum PGI 51 LVDYSKNLVT EDVMRMLVDL AKSRGVEAAR ERMFNGEKIN YTEGRAVLHV
*.....*
rab PGI 101 ALRNRSTPI VVDGKDVME VNKVLDKMKK FCQVRVSGDW KGYTGKTITD
hum PGI 101 ALRNRSTPI LVDGKDVME VNKVLDKMKK FCQVRVSGDW KGYTGKTITD
*****
rab PGI 151 VINIGIGGSD LGPLMVTEAL KPYSSGGPRV WFVSNIDGTH IAKTLAELNP
hum PGI 151 VINIGIVGSD LGPLMVTEAL KPYSSGGPRV WYVSNIDGTH IAKTLAELNP
*****
rab PGI 201 ESSLFIIASK TFFTQETITN AKTAKDWFLS SAKDPSTVAK HFVALSTNTA
hum PGI 201 ESSLFIIASK TFFTQETITN AETAKEWFLQ AAKDPSAVAK HFVALSTNTT
*****
rab PGI 251 KVKEFGIDPQ NMFEEFDWVG GRYSLWSAIG LSIALHVGFD NFEQLLSGAH
hum PGI 251 KVKEFGIDPQ NMFEEFDWVG GRYSLWSAIG LSIALHVGFD NFEQLLSGAH
*****
rab PGI 301 WMDQHFRTPP LEKNAPVLLA MLGIWYINCF GCETQAVLPY DQYLHRFAAY
hum PGI 301 WMDQHFRTPP LEKNAPVLLA LLGIWYINCF GCETHAMLPY DQYLHRFAAY
*****
rab PGI 351 FQQGMESNG KYITKSGARV DHQTGPIVWG EPGTNGQHAF YQLIHQGTKM
hum PGI 351 FQQGMESNG KYITKSGTRV DHQTGPIVWG EPGTNGQHAF YQLIHQGTKM
*****
rab PGI 401 IPCDFLIPVQ TQHPPIRKLH HKILLANFLA QTEALMKGKS TEEARKELQA
hum PGI 401 IPCDFLIPVQ TQHPPIRKLH HKILLANFLA QTEALMRGKS TEEARKELQA
*****
rab PGI 451 AGKSPEDLMK LLPHKVFEGN RPTNSIVFTK LTPFILGALI AMYEHKIFVQ
hum PGI 451 AGKSPEDLER LLPHKVFEGN RPTNSIVFTK LTPFMLGALV AMYEHKIFVQ
*****
rab PGI 501 GVVWDINSFD QWGVELGKQL AKKIEPELDG SSPVTSHDSS TNGLINFIKQ
hum PGI 501 GIWDINSFD QWGVELGKQL AKKIEPELDG SAQVTSHDAS TNGLINFIKQ
*.....*
rab PGI 551 QREAK
hum PGI 551 QREAR
*****

```

Figura 30: Alinhamento das sequências de PGI-coelho (rab PGI) e da PGI-hum (hum PGI). Os resíduos conservados são indicados por um asterisco, as mutações estão coloridas de vermelho e os resíduos descritos como catalíticos estão coloridos de azul.

O conjunto de reflexões utilizadas na substituição foi o coletados no IFSC-USP. Os cálculos rotacionais e translacionais foram realizados numa faixa de resolução entre 15 e 3,5 Å. A mesma rotina foi utilizada posteriormente com o conjunto de dados obtidos no LNLS.

O programa CNS [27] foi utilizado para calcular um mapa de patterson com o conjunto de dados do IFSC-USP e listar a posição dos maiores picos encontrados na cela unitária. Esse cálculo teve como objetivo verificar a presença de translações relacionando subunidades da PGI-hum dentro da unidade assimétrica. O programa MATTEWS_COEFF foi usado para calcular o conteúdo de solvente e coeficiente de Matthew do cristal [28]. A presença de simetria rotacional não cristalografia dentro da unidade assimétrica foi obtida pelo cálculo da função de auto-rotação do mapa de patterson, este cálculo foi realizado com o programa AMoRe [26] e GLRF[29].

8.3) Resultados e discussão

O motivo de não ter utilizado diretamente o conjunto de dados obtido no LNLS para o cálculo de substituição molecular, é que este só foi coletado aproximadamente um mês após a coleta e processamento dos dados obtidos no IFSC-USP. Além disso, para otimizar o cálculo de substituição molecular e facilitar a sobreposição das moléculas são usados apenas as reflexões de baixa resolução. Após a coleta dos dados no LNLS, as mesmas etapas de substituição molecular foram recalculadas e confirmaram as informações já obtidas para com os dados do IFSC-USP.

O AMoRe encontrou uma solução única com C igual a 68.8 % para a substituição molecular da PGI-humana, apresentando uma unidade assimétrica composta por dois dímeros. A solução é indicada por α , β , γ , x , y e z :

	α	β	γ	x	y	z
#	209.9	5.6	237.9	19.9	3.3	49.6
#	336.6	3.9	24.7	15.6	107.8	117.5

sendo os três primeiros parâmetros correspondentes aos ângulos de rotação e os três seguintes correspondentes ao vetor de translação. Cada linha corresponde a posição de um dímeros de PGI-hum na unidade assimétrica.

Para a PGI de coelho, a unidade assimétrica é composta por um dímero apenas. Sendo as cela unitárias das duas proteínas quase idênticas, essa diferença quanto a composição da unidade assimétrica é compensada pela diferença de grupo espacial. A ausência de uma centragem no cristal de PGI-humana deixa espaço na cela unitária para o posicionamento de mais um dímero na unidade assimétrica.

A visualização do empacotamento cristalino (fig. 31), o coeficiente de Matthews ($2,41 \text{ \AA}^3/\text{Da}$) e o conteúdo de solvente no cristal (48,64%) indicam que a solução proposta pelo AMoRe está correta. Estas conclusões fazem sentido uma vez que não se observam sobreposições entre moléculas vizinhas e os valores dos coeficientes são próximos aos observados para a PGI de coelho.

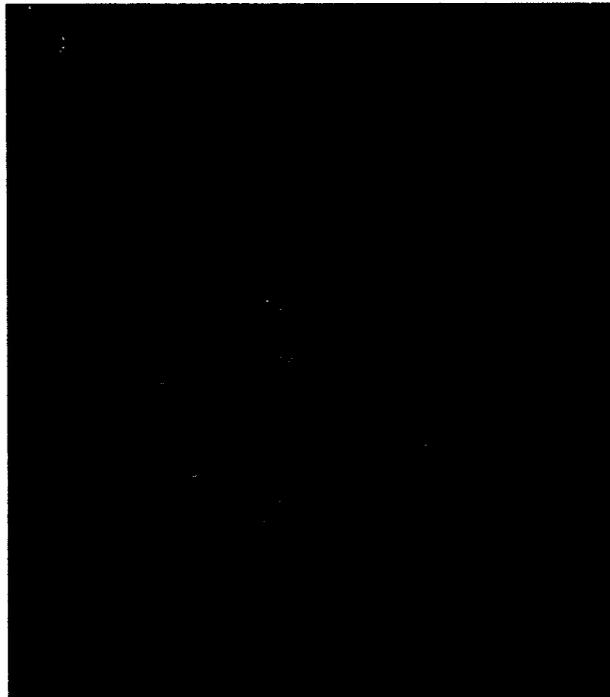


Figura 31: visualização da unidade assimétrica de PGI-humana em branco (apenas segmentos de $C\alpha$), juntamente com algumas moléculas relacionadas por simetria dentro da cela unitária.

A função de auto-rotação calculada pelo AMoRe na faixa de resolução de 15 a 3.5 \AA e raio de integração de 25 \AA indicou a presença de um eixo de ordem 2 não cristalográfico com coeficiente de correlação de 30% relacionando um monômero com o outro pertencente ao mesmo dímero. O pico correspondente ao eixo não cristalográfico pode ser visualizado na seção $\chi=180^\circ$ da projeção estereográfica da função de auto-rotação (fig32a).

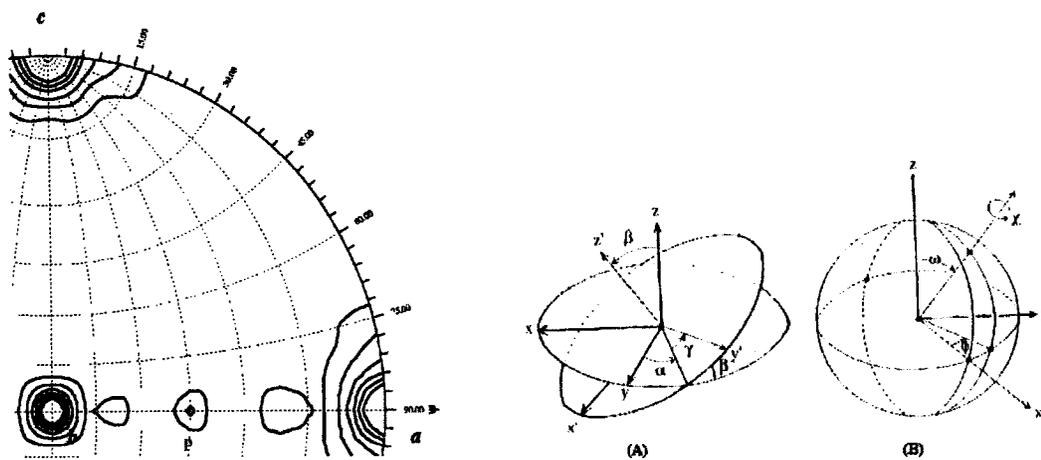


Figura 32: A esquerda está representada a seção $\chi=180^\circ$ da projeção estereográfica e o pico correspondente ao eixo não cristalográfico de ordem dois aparece indicado pela letra P. Este eixo é perpendicular ao eixo cristalográfico c e forma um ângulo de aproximadamente 45° com o eixo a. A direita (em A) estão representados os ângulos eurlianos α , β e γ usados para descrever a posição do eixo em relação aos eixos da rede cristalina; (em B) os eixos de coordenadas polares usados na busca do eixo de simetria não cristalográfica. ω e ϕ determinam do eixo de rotação, enquanto χ define a rotação sobre o próprio eixo.

Este eixo apresenta ângulos eurlianos de $\alpha=42.5^\circ$, $\beta=172.5^\circ$ and $\gamma=42.5^\circ$, correspondendo a um eixo de ordem dois quase paralelo ao plano cristalino ab e orientado próximo a diagonal deste plano.

O mapa de patterson, calculado pelo CNS, localizou um pico com 75σ de

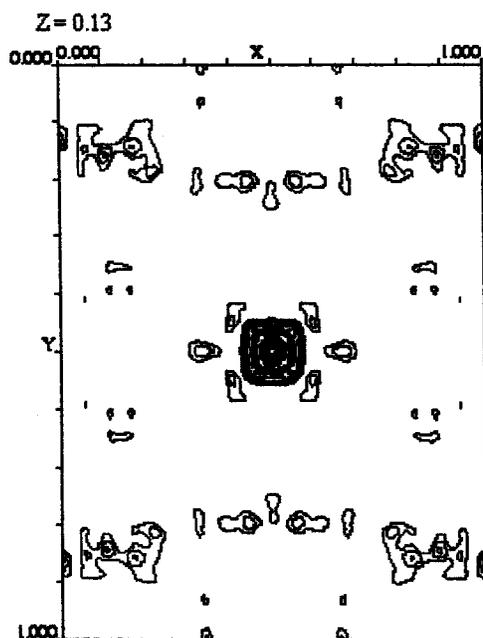


Figura 33: Mapa de patterson visualizado pelo programa Xtalview. Seção mostrando um plano paralelo ab com altura 0.13 em c. O pico no centro da figura correspondente a translação de 0.5 em a, 0.5 em b e 0.13 em c.

altura com coordenadas fracionárias de 0.5 em a, 0.5 em b e 0.13 em c (fig 33). Este

pico equivale a uma translação de um dos dímero sobre o outro dentro da unidade assimétrica.

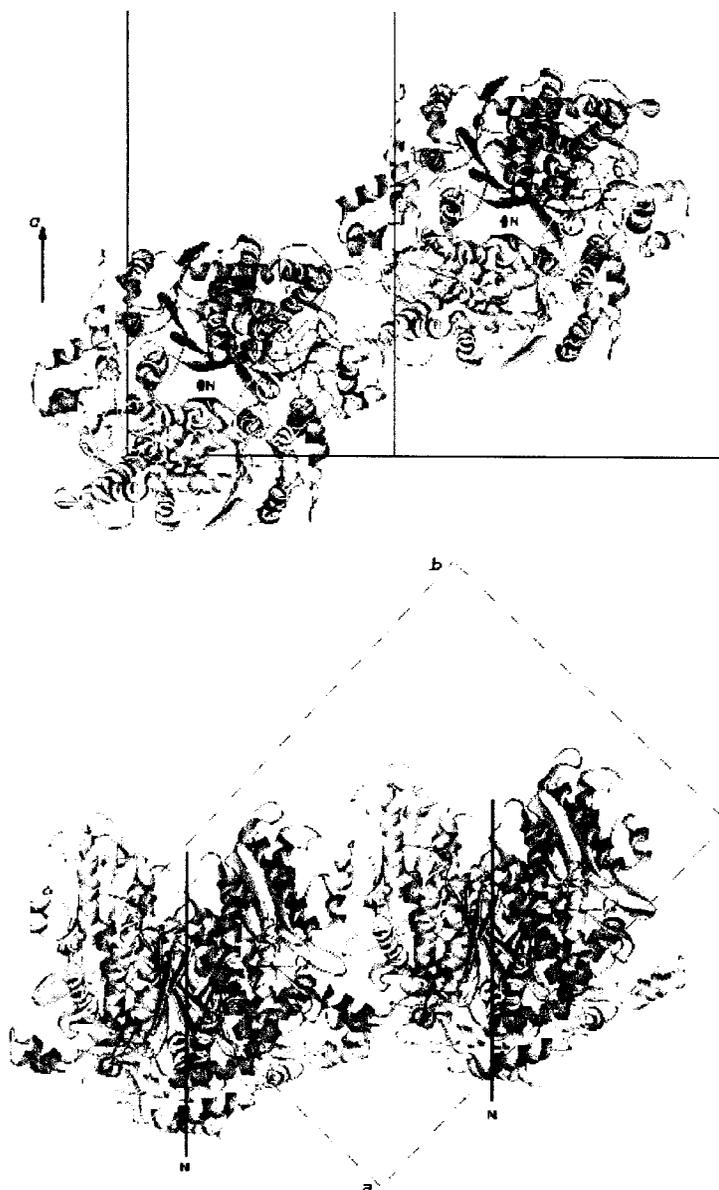


Figura 34: Representação da cela unitária de PGI-hum com moléculas de PGI de coelho correspondente a unidade assimétrica. O eixo não cristalográfico (em azul) está orientado paralelo ao plano ab e perpendicular a c (A). O eixo apresenta uma orientação próxima a diagonal entre a e b (B).

Após a visualização das moléculas de PGI de coelho dentro da cela unitária do cristal de PGI-hum, a localização do eixo não cristalográfico e a translação observada no mapa de Patterson tornam-se mais evidentes (fig.34). A unidade assimétrica apresenta dois homo dímeros.

CAPÍTULO 9

Refinamento e análise estrutural

9.1) Introdução

O refinamento é o ajuste do modelo atômico ao conjunto de dados de difração, respeitando as restrições estereoquímicas inerentes a uma estrutura proteica [30]. O conjunto de dados corresponde as reflexões medidas a partir de um cristal. As restrições estereoquímicas são definidas como tamanho de ligação, ângulos, planaridade e outros parâmetros geométricos, que podem ser considerados constantes ou limitados a um pequeno intervalo de valores para qualquer proteína.

O refinamento é realizado de maneira a minimizar a função do tipo

$$E = E_{\text{geom.}} + E_{\text{raios-X}},$$

onde $E_{\text{geom.}}$ é uma função que descreve a geometria da proteína e impõe restrições

$$R = \frac{\sum_{hkl} \| |F_o| - k |F_c| \|}{\sum_{hkl} |F_o|} \times 100\% \qquad R_{\text{free}} = \frac{\sum_{hkl \in \text{test set}} \| |F_o| - k |F_c| \|}{\sum_{hkl} |F_o|}$$

estereoquímicas; $E_{\text{raios-X}}$ é a função que descreve o ajuste dos fatores de estrutura calculados (F_c) a partir do modelo atual com os fatores de estrutura (F_o) obtidos do conjunto de dados. Para acompanhar o ajuste do modelo aos dados de difração é comum utilizar duas figuras de mérito, R e R_{free} :

A diferença de R e R_{free} é que o R_{free} é calculado a partir de uma fração do conjunto de dados (test set) que não foi usada durante o refinamento [31]. De forma que o modelo usado para obter F_c não tenha influência dessas reflexões. A função

comumente aplicada para o refinamento de macromoléculas utiliza o método dos mínimos quadrados (do inglês, *least-square residual function*) LSQ:

$$E^{LSQ} = E_{geom.} + \sum_{hkl} w_a (|F_o| - k |F_c|)^2$$

onde $|F_o|$ e $|F_c|$ são as amplitudes dos fatores de estrutura observado e calculado; k é um fator de escala e w_a é um peso estatístico que define a contribuição geométrica e cristalográfica para a função E^{LSQ} .

De acordo com essa função, as coordenadas x, y, z e fator de temperatura B são ajustados para que as reflexões tenham um $|F_c|$ mais próximo possível de $|F_o|$, reduzindo assim o resíduo cristalográfico de E^{LSQ} . O fator de temperatura é um parâmetro que mede a distribuição de átomos equivalentes em relação a uma posição de equilíbrio.

Ao ajustar os parâmetros x, y, z e B , de forma a minimizar a diferença $|F_o| - k \cdot |F_c|$ para todas as reflexões, está sendo minimizado conseqüentemente o fator R , ou seja, o R não pode ser considerado uma medida independente do ajuste feito aos parâmetros, uma vez que estes são tendenciados por todas as reflexões do conjunto de dados. Daí surge a necessidade de separar uma pequena fração (de 5 a 10 %) das reflexões que irão compor o “test-set” e não entram no cálculo de $|F_c|$, permitindo assim uma validação cruzada monitorada por R_{free} .

Atualmente um novo método, conhecido como “maximum likelihood” (ML), tem sido usado como validação cruzada para melhorar o refinamento por substituição molecular [32]. σ_A^{cv} é um valor obtido por validação cruzada e usado para quantificar erros associados ao modelo durante a obtenção de $|F_c|$. σ_A^{cv} é então usado para computar o valor esperado para $\langle |F_o| \rangle^{cv}$ e a variância associada σ_{ML}^{2cv} . Estes dois termos são incorporados a função *maximum likelihood*:

$$E^{ML} = E_{geom.} + \sum_{hkl_working-set} \left(\frac{1}{\sigma_{ML}^{2cv}} \right) \cdot (|F_o| - \langle |F_o| \rangle^{cv})^2$$

onde *hkl_working-set* é o conjunto de dados menos a fração excluída (*test-set*) para a validação cruzada, e a variância σ_{ML}^{2cv} serve como peso para a contribuição em particular de cada reflexão no termo $E_{raios-X}$. Isso permite a consideração de todas as

reflexões medidas com intensidade positiva $I(\sigma) > 0$, uma vez que as reflexões menos intensas terão uma contribuição menor devido ao peso $1/(\sigma_{ML}^2)$ associado.

Existem basicamente três formas de variar E^{ML} em função dos parâmetros atômicos: minimização de energia, dinâmica molecular e resfriamento simulado (*simulated annealing*). A função E^{ML} pode ser representada por uma superfície irregular composta de picos e vales que correspondem respectivamente a barreiras energéticas e mínimos de energia.

Os métodos de minimização tendem a buscar mínimos de energia em uma vizinhança restringida por pequenas variações das coordenadas atômicas. Estes métodos não são capazes de sobrepor barreiras energéticas e procurar mínimos de energia mais adequados. Postula-se que exista apenas um mínimo de energia, denominado *mínimo global*, que representa a melhor solução para a equação de E^{ML} em função da variação dos parâmetros atômicos.

Os métodos de dinâmica molecular fornecem a cada átomo uma massa (m_i) e uma velocidade de acordo com a energia cinética determinada pela “temperatura” associada ao sistema

$$E_{kin} = \sum_i^n \frac{1}{2} \cdot m_i (\partial \vec{r}_i / \partial t)^2 \quad T_{atual} = \frac{2 \cdot E_{kin}}{3 \cdot n \cdot k_b}$$

onde \vec{r}_i é o vetor coordenada, n é o número de átomos e k_b é a constante de Boltzmann. A energia cinética pode ser usada pelo sistema para sobrepor barreiras de energia potencial e assim realizar uma busca maior no espaço de coordenadas por possíveis candidatos a mínimo global.

O controle da quantidade de energia fornecida ao sistema é proporcional a velocidade atribuída aos átomos e pode ser definida por uma temperatura alvo

$$v_i^{nova} = \frac{\partial \vec{r}_i}{\partial t} \cdot \left(\frac{T}{T_{atual}} \right)^{1/2}$$

Os métodos de “simulated annealing” consistem em resfriar gradativamente o sistema para que a busca possa varrer inicialmente um grande espaço conformacional e em seguida passe a se restringir, a medida que a temperatura é reduzida, entorno do mínimo global.

Para o refinamento de fases iniciais obtidas por substituição molecular, costuma-se substituir os resíduos que diferem entre o modelo de busca e a proteína

no cristal por alaninas. Outra alternativa, é atribuir ocupação zero para os átomos de cadeia lateral posteriores ao C β . Realiza-se então uma etapa de refinamento de corpo-rígido para melhorar a posição da nova molécula dentro da cela unitária. Nesta etapa de refinamento, toda a molécula é tratada como um corpo rígido e a faixa de resolução empregada vai até próximo de 3.0 Å de resolução, para facilitar o ajuste da molécula dentro da densidade eletrônica.

O passo seguinte consiste em uma etapa de “simulated annealing” seguida por uma minimização de coordenadas e fator de temperatura[33]. Ao final deste protocolo, calcula-se mapas de densidade eletrônica que servirão para manipular o modelo de forma que os átomos se encaixem melhor dentro da densidade observada. A atribuição de ocupação para a cadeia lateral dos resíduos mutados devem ser realizada a medida que a densidade eletrônica correspondente aos átomos passa a ser observada nos mapas.

O processo de refinamento é bastante interativo. Ele exige a análise dos mapas e a interferência manual no posicionamento de regiões divergentes entre a proteína cristalizada e a utilizada como modelo de busca para substituição molecular. Cada passo de refinamento deve ser monitorado pelos valores de R e R_{free}. O decaimento equivalente de ambos indica um refinamento correto em direção a melhor solução para a estrutura cristalina.

Quando os valores de R e R_{free} deixam de decair, isto indica que as coordenadas atômicas atuais são as que melhor representam o padrão de difração coletado. Ainda é possível reduzir os valores de R e R_{free} pela adição de moléculas d'água com função estrutural na estabilização da proteína. Essas moléculas podem ser visualizadas em mapas diferença (|F_o| - |F_c|) como picos esféricos de intensidade superior a 2 σ .

9.2) Métodos

O modelo da PGI-hum foi refinado usando o programa CNS. Uma etapa de refinamento tipo corpo-rígido foi realizada com todas as reflexões entre 30 e 3.0 Å pertencentes ao conjunto de dados coletado no LNLS. Nesta etapa o modelo era constituído pela PGI-hum considerando os resíduos mutados em relação a PGI de coelho com ocupação zero para os átomos de cadeia lateral após o C β . Após este

primeiro ciclo, todas as etapas seguintes foram realizadas com todas as reflexões coletadas até 2.1 Å de resolução.

Inicialmente foram realizados três ciclos de simulated annealing a 5000 K, seguidos por 100 passos de refinamento de coordenadas e 10 passos de refinamento de fator B. Mais oito ciclos com temperatura de 2500 K e refinamentos semelhantes foram realizados. A cada ciclo eram calculados mapas $2\sigma A|F_o| - D.|F_c|$ e $m.|F_o| - D.|F_c|$. Após inspeção visual dos mapas utilizando o programa *O*, resíduos que se apresentavam fora da densidade eletrônica eram manipulados de maneira a melhorar o ajuste do modelo aos mapas. Durante estes ciclos iniciais não foi utilizado a simetria não cristalográfica, ou seja os quatros monômeros da unidade assimétrica estavam sendo refinados independentemente.

Para os ciclos seguintes, foi escolhido o monômero que apresentava melhor ajuste a densidade eletrônica e este serviu como base para o refinamento usando a simetria não cristalográfica restrita, implementada pelo arquivo *ncs.def* do CNS. O próximo ciclo foi realizado com temperatura de 2500 K, 100 passos minimização de coordenadas, 10 passos de minimização de fator B e uma constante de 500 Kcal/mol.Å² associada a simetria não cristalográfica. Com o decorrer do refinamento, a temperatura foi reduzida para 500 K e os ciclos de minimização foram elevados para 500 passos de coordenadas e 50 passos de fator B.

Quando os valores de R e R_{free} pararam de decair pela simples manipulação do modelo, foram realizadas buscas de moléculas d'água pela rotina *water_pick* do programa CNS. As moléculas d'água foram inicialmente selecionadas como picos superior a 3σ no mapa $m.F_o - D.F_c$. Estes picos deveriam estar entre 2.6 e 4.0 Å de qualquer átomo da proteína ou entre 2.0 e 3.2 Å de átomos de oxigênio ou nitrogênio, deveriam respeitar também uma geometria adequada para formação de pontes de hidrogênio. Com o objetivo de aumentar o número de moléculas d'água, a altura para seleção de picos no mapa $m.F_o - D.F_c$ foi reduzida para 2.5 e 2.0 sem alterar os demais critérios. Em seguida, picos com 1.5σ de altura e a qualquer distância entre 2.5 e 4.2 Å eram selecionados. A distância passou a ser o único critério de seleção, não mais importando a geometria de coordenação dessas moléculas. Os parâmetros de distância foram então relaxados para picos entre 2.2 e 4.3 Å, sendo a última seleção realizada na faixa entre 2.1 e 4.5 Å. Quando a adição

de moléculas deixou de reduzir os valores de R e R_{free} , o refinamento foi dado por encerrado.

9.3) Resultados e discussão

O refinamento de corpo rígido resultou em valores iniciais de R e R_{free} iguais a 34,24 e 32,51 %. Após o primeiro ciclo de simulated annealing, R decaiu para 28,26 e o R_{free} foi para 31,52. Estes valores decaíram constantemente com exceção de uma pequena variação devido ao início da adição de moléculas d'água (fig. 35).

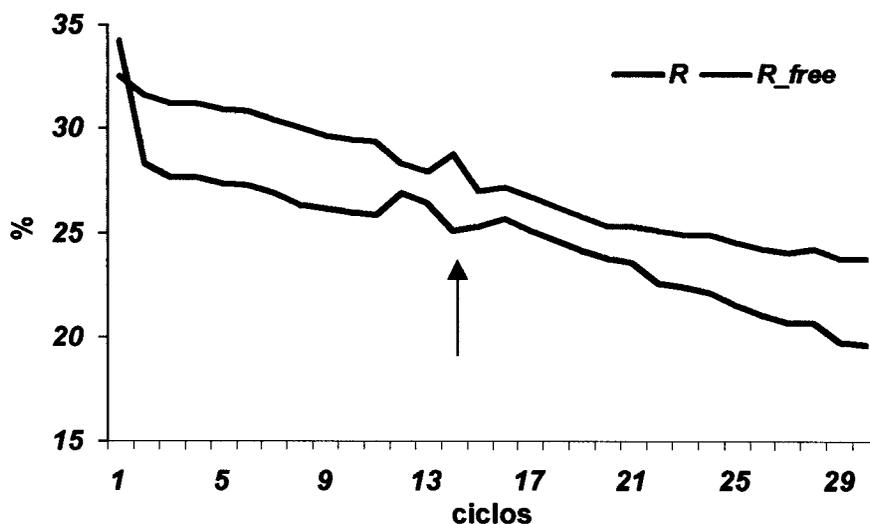


Figura 35: Gráfico mostrando a variação dos valores de R (azul) e R_{free} (roxo) durante os ciclos de refinamento. Cada ciclo era constituído por uma etapa de simulated annealing, seguida por um refinamento de coordenadas e fatores de temperatura. A seta indica o ciclo no qual iniciou-se a adição de moléculas d'água.

Os valores finais de R e R_{free} correspondem a 19,7 e 23,7 %, sendo que na faixa de máxima resolução, entre 2,23 e 2,10, estes valores sobem para 23,3 e 28,0 % respectivamente. Atualmente a melhor estrutura de PGI disponível no banco de dados PDB (PGI de coelho, código 1dqr), apresenta resolução máxima de 2,5 Å e valores de R e R_{free} iguais a 26,5 e 21,2 %. Outra diferença importante entre as duas estruturas é que a de coelho apresenta o inibidor D-gluconato-6-fosfato complexado ao sítio ativo. Como a identidade entre a PGI de coelho e a PGI-hum é superior a 90%, pode-se comparar as mudanças estruturais relacionadas a ligação do inibidor no sítio ativo da proteína.

A principal diferença entre a PGI-hum e a de coelho está localizada em dois loops denominados loop_A e loop_B. O primeiro estende-se dos resíduos 208 a 216

e inclui o resíduo catalítico GLU_216. O loop_B está logo adjacente ao loop_A e é constituído pelos resíduos 256 a 262 (fig. 36)

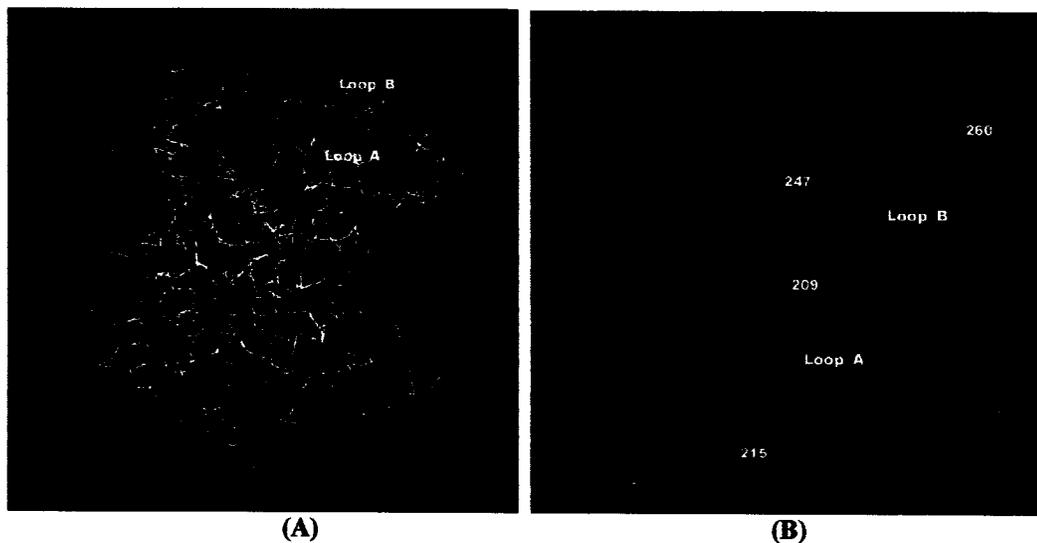


Figura 36: Representação do traçado de carbonos- α da PGI-hum (amarelo) e da PGI de coelho (roxo), indicando a posição do loop's A e B . Em (A) tem-se uma visão geral da sobreposição das duas proteínas, enquanto (B) apresenta um visão mais próxima com a presença do inibidor D-gluconato-6-fosfato e posição dos loop's na sequência de ambas as proteínas.

Aparentemente pode-se inferir que a ligação do inibidor faz com que o loop_A se feche sobre o ligante trazendo consigo o loop_B.

A geometria da PGI-hum foi analisada pelo gráfico de Ramachandran (fig.37) para todos os resíduos e individualmente para cada tipo de resíduo. Pelo gráfico para todos os resíduos pode-se observar que apenas três resíduos estão em áreas desfavoráveis.

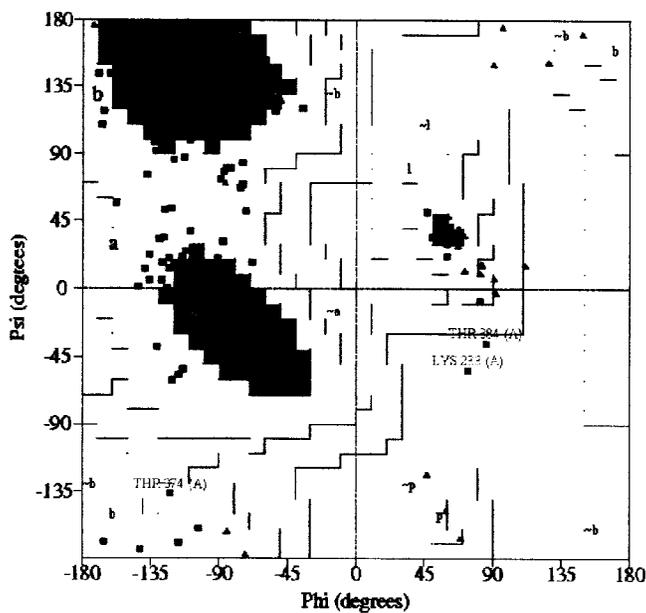


Figura 37; gráfico de ramachandran para todos os resíduos do monômero A da PGI-hum. Apenas dois resíduos em região não permitida, Lys_233 e Thr_384. E um resíduo em região desfavorável, Thr_374.

Região	Nº de res.	% de res.
Favorável	436	88,8
Permitida	52	10,6
Desfavorável	1	0,2
Imprópria	2	0,4

A análise da densidade eletrônica e das interações com a vizinhança para os três resíduos que se encontram fora das regiões permitidas indica que apesar da geometria desfavorável estes resíduos estão bem definidos em termos de densidade eletrônica e realizando contatos com resíduos vizinhos que lhes confere estabilidade (fig. 38).

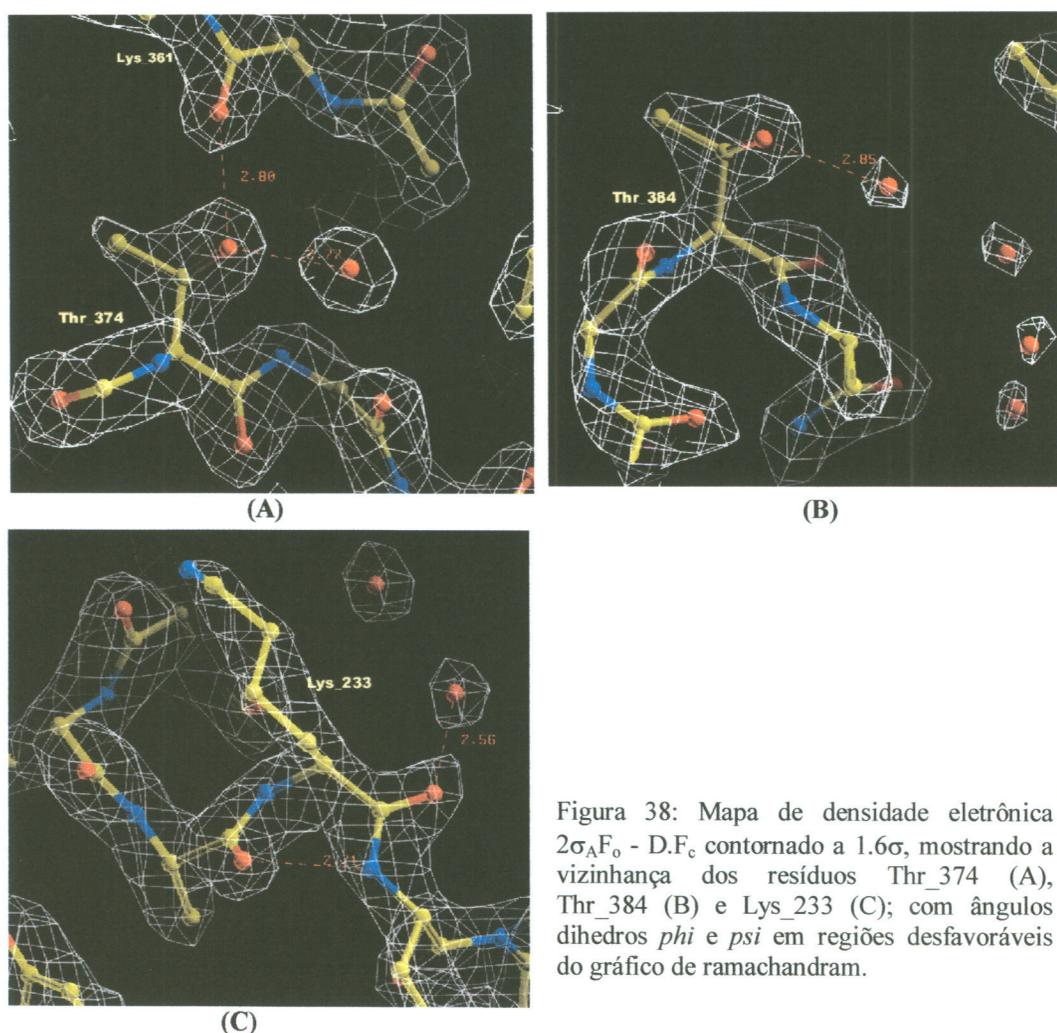


Figura 38: Mapa de densidade eletrônica $2\sigma_{A}F_{o} - D.F_{c}$ contornado a 1.6σ , mostrando a vizinhança dos resíduos Thr_374 (A), Thr_384 (B) e Lys_233 (C); com ângulos dihedros *phi* e *psi* em regiões desfavoráveis do gráfico de ramachandram.

A treonina 374 é estabilizada por pontes de hidrogênio com uma molécula d'água e com o oxigênio da cadeia principal da Lys_361, já a treonina 384 é estabilizada apenas por uma ponte de hidrogênio com uma molécula d'água. A lisina 233, com ângulos *phi* e *psi* impróprios, está localizada em uma região bastante acessível ao solvente (fig.39) e apresenta densidade eletrônica suficiente para sustentar esta conformação (fig.38c).

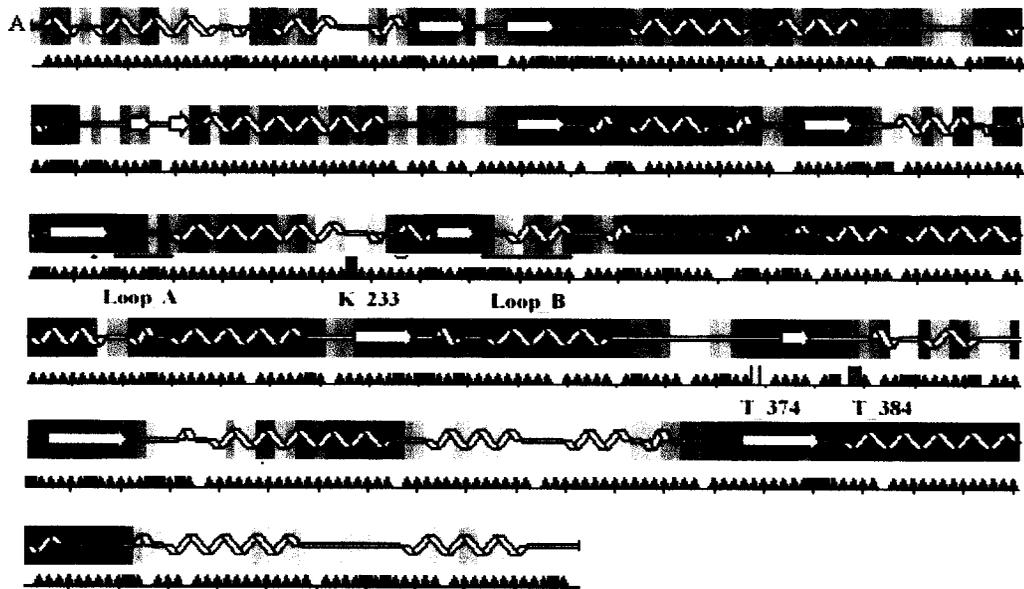


Figura 39: Representação esquemática da estrutura secundária da PGI-hum adaptada do programa *PROCHECK*. A fita amarela estreita representa regiões sem estrutura secundária definida, as setas representam fitas β e as ondulações representam α -hélices. Os tons de azul mais fortes indicam que o resíduo apresenta-se pouco exposto ao solvente e os tons mais claros indicam alta exposição ao solvente. Cada resíduo pode -ser representado por um triângulo azul, um quadrado verde, uma barra roxa ou uma barra vermelha, estes símbolos representam a região do gráfico de ramachandram em que cada resíduo se encontra. Dessa forma, o triângulo azul indica uma região favorável; o quadrado verde, uma região permitida; a barra roxa, uma região desfavorável e uma barra vermelha, uma região imprópria.

A unidade de PGI-hum que apresenta atividade de isomerase corresponde a um dímero (fig. 41a). A tríade catalítica é formada pela His_388 de um monômero, Glu_216 e Lys_518 do monômero adjacente(Fig.41c). O mecanismo de ação enzimática descrito para a PGI de coelho baseia-se no mecanismo proposto por Rose e colaboradores[34]. O primeiro passo é a ligação do açúcar fosfatado ao sítio ativo devido a coordenação do grupo fosfato principalmente pelas treoninas do loop_A. Em seguida, um próton é transferido da enzima para o oxigênio do anel do açúcar resultando na abertura do anel .

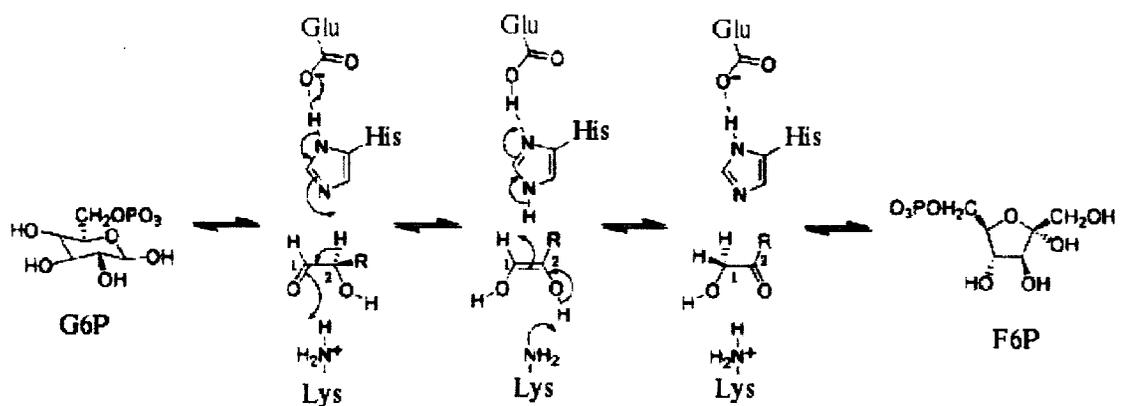


Figura 40: Representação esquemática do mecanismo de reação catalisado pela PGI. Para a PGI-hum os resíduos envolvidos são Glu_216, Lys_518 e His_388.

O próton em C-2 do substrato está agora relativamente ácido porque encontra-se numa posição α em relação a um grupo carbonil (fig. 40). Uma base enzimática retira o próton em C-2 e ocorre uma transferência de um par de elétrons para formação de uma dupla ligação entre C-1 e C-2. Esta conformação do substrato é equivalente ao intermediário de alta energia 1,2-cis-enodiolato. O mesmo próton que é removido de C-2 é agora transferido de volta ao carbono C-1. A última etapa envolve a transferência de um par de elétrons do oxigênio que formará o anel para o carbono C-2. Após a formação dessa ligação o oxigênio da ketose aceita um próton do solvente para completar o anel correspondente a frutose 6-fosfato.

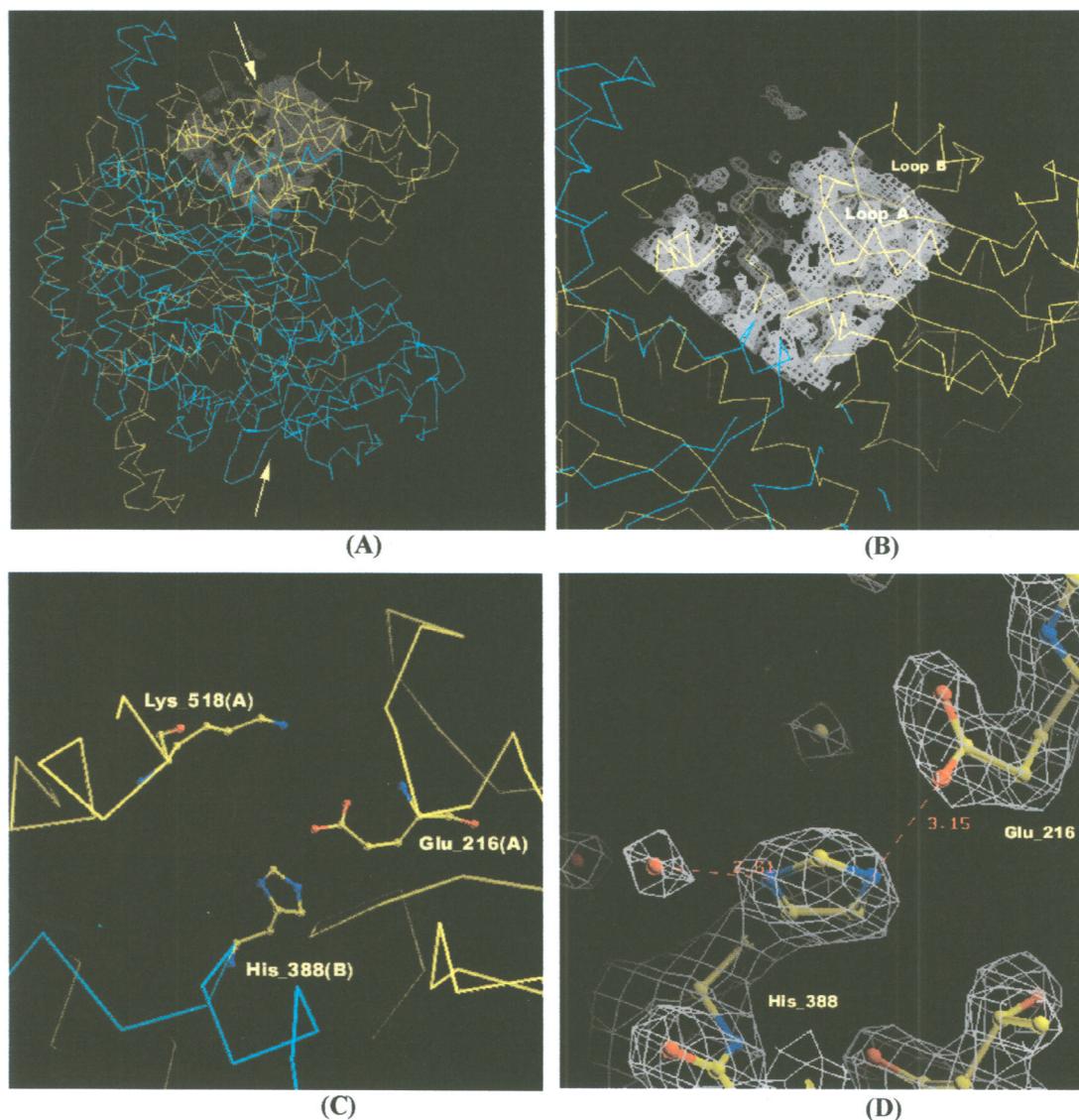


Figura 41: Visão geral do dímero de PGI-hum, mostrando o mapa $2\sigma_A F_0 - D.F_c$ contornado a 1.3σ para a região do sítio ativo(A). Uma secção mais estreita (B) que permite ver no mesmo mapa a densidade eletrônica da cavidade onde localiza-se o sítio ativo. Em (C) detalhe do sítio ativo mostrando a His_388 de um monômero (em azul), Glu_216 e Lys_518 do monômero adjacente (em amarelo). Em (D) densidade eletrônica do mapa $2\sigma_A F_0 - D.F_c$ contornado a 1.3σ , mostrando a histidina, o glutâmico e duas moléculas d'água localizadas próximo ao sítio ativo.

A especificidade por açúcares fosfatados é resultante da coordenação do fosfato pelos resíduos do Loop_A. Em um exercício de modelagem molecular, o inibidor D-gluconato-6-fosfato da PGI de coelho foi posicionado na região correspondente à PGI-hum através de uma simples sobreposição das duas proteínas. A análise dessa transferência do inibidor para a PGI-hum indica quais os principais resíduos responsáveis pela coordenação do fosfato (fig.42).

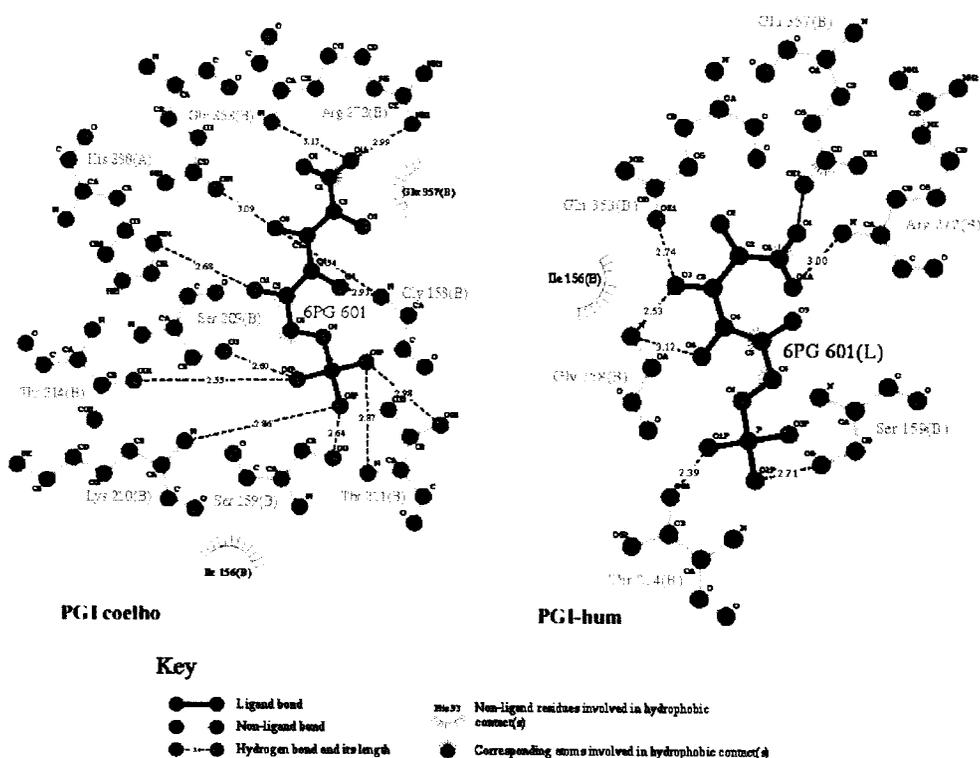


Figura 42: Representação planar obtida do programa *LIGPLOT* para os resíduos que interagem com o ligante D-gluconato-6-fosfato (6PG) na estrutura da PGI de coelho e no modelo da PGI-hum.

Dos resíduos contidos no Loop_A (209 a 216) da PGI de coelho, apenas a treonina 214 mantém a interação com o oxigênio do fosfato, para o modelo da PGI-hum; ou seja, como o Loop_A da PGI-hum apresenta-se mais aberto, os demais resíduos (Ser_209, Lys_210 e Thr_211) que coordenam os oxigênios do fosfato só irão formar pontes de hidrogênio com o inibidor se o Loop_A se fechar sobre o fosfato (fig.42). O único resíduo que não está presente no Loop_A e que coordena o fosfato nas duas PGI's é a serina 159. Este resíduo talvez seja o primeiro a ancorar o fosfato, possibilitando que os demais resíduos da Loop_A se acomodem em condições de completar a geometria de coordenação do fosfato.

A análise da densidade eletrônica ao longo da sequência de aminoácidos da PGI-hum recombinante permitiu a identificação de duas mutações. A primeira delas corresponde a substituição de um valina por uma glicina na posição 157 e a segunda é a substituição de uma lisina por um ácido glutâmico na posição 375. Ambas as mutações podem ser consideradas silenciosas quanto a atividade da proteína, são plenamente confirmadas pela densidade eletrônica e pelos contatos com os resíduos vizinhos (fig.43).

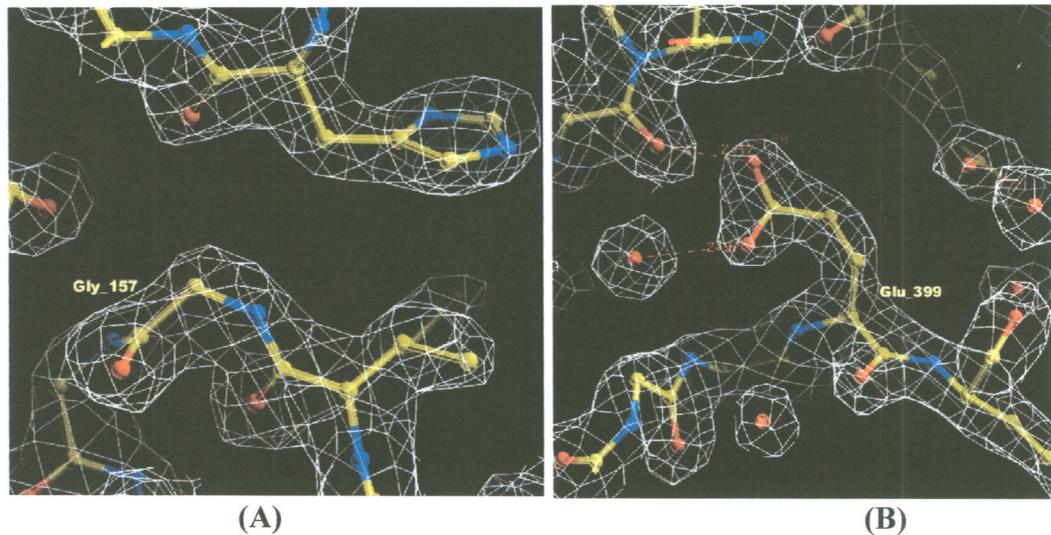


Figura 43: Mapa de densidade eletrônica $2\sigma_A F_o - D.F_c$ contornado a 1.6σ , mostrando as mutações Val > Gly_157 (A) e Lys > Glu_399 (B).

Não existe densidade eletrônica que justifique a presença de um carbono- β para o resíduo 157, dessa forma, ao invés de uma valina o resíduo 157 corresponde a uma glicina (fig. 43a). A substituição da lisina de caráter positivo por um glutâmico de carga negativa é estabilizada por pontes de hidrogênio com a Gly_375 e uma molécula d'água (fig. 43b).

CAPÍTULO 10

Conclusões gerais e perspectivas futuras

De posse de um protocolo bem estabelecido para expressão, purificação e cristalização da PGI-hum, é possível planejar novos experimentos para enriquecer o conhecimento a respeito do mecanismo de isomerização e demais funções associadas a esta enzima.

Com a estrutura apresentada neste trabalho, pode-se realizar um estudo de modelagem molecular para ajudar a compreender as modificações estruturais associadas a ligação de inibidores e substratos. De imediato, o ideal seria obter co-cristais de PGI-hum complexada ao inibidor comercial 6PG e determinar sua estrutura. Uma comparação entre as PGI's de coelho e humana, complexadas com o mesmo inibidor, iria determinar se o estudo estrutural da PGI de coelho representa de forma adequada o que ocorre com a PGI-hum.

A partir da estrutura de PGI-hum apresentada neste trabalho e da mesma enzima co-cristalizada com um inibidor o estudo das mudanças conformacionais entre o estado livre e o inibido, ganharia um maior embasamento experimental. A modelagem molecular passaria a ser empregada na proposta de novos inibidores da PGI-hum.

Outro aspecto que pode ser explorado com técnicas de biologia molecular é o efeito de mutações dirigidas a resíduos que são considerados essenciais para a atividade de isomerização e para o enovelamento correto da enzima.

O estudos cinéticos devem ser continuados com o objetivo de definir constantes de inibição K_i para os demais compostos que apresentaram alguma atividade inibitória durante os teste de eficiência de inibição realizados neste trabalho.

ANEXO I

Human phosphoglucose isomerase: expression,
purification, crystallization and preliminary
crystallographic analysisArtur T. Cordeiro,^a Paulo H. C.
Godoi,^{a,b} Luis F. Delboni,^{a†}
Glaucius Oliva^{a,b} and Otavio H.
Thiemann^{a*}^aLaboratory of Protein Crystallography and
Structural Biology, Physics Institute of São
Carlos, University of São Paulo - USP,
Av. Trabalhador São-carlense 400, PO Box 369,
13566-590 São Carlos - SP, Brazil, and^bChemistry Institute of São Carlos - USP,
Av. Trabalhador São-carlense 400, PO Box 369,
13566-590 São Carlos - SP, Brazil† Present address: Pontifícia Universidade
Católica de Poços de Caldas, MG, Brazil.

Correspondence e-mail: thiemann@ifsc.usp.br

Phosphoglucose isomerase (PGI) is the second enzyme in the glycolytic pathway and catalyzes an aldose–ketose isomerization. Outside the cell, PGI has been found to function as both a cytokine and as a growth factor. The human *pgi* gene was cloned and the expressed enzyme was purified to homogeneity. Isomorphous crystals were obtained under two conditions and belong to the $P2_12_12_1$ space group, with unit-cell parameters $a = 80.37$, $b = 107.54$, $c = 270.33$ Å. A 94.7% complete data set was obtained and processed to a limiting resolution of 2.6 Å. The asymmetric unit contains two hPGI dimers according to density calculations, a self-rotation function map and molecular-replacement solution.

Received 2 October 2000
Accepted 16 January 2001

1. Introduction

Phosphoglucose isomerase (PGI; E.C. 5.3.1.9) is a multifunctional enzyme composed of homodimers with molecular masses ranging from 100 to 120 kDa. With the exception of some intracellular parasites, PGIs are involved in the glycolytic and gluconeogenesis pathways in most known organisms (Andersson *et al.*, 1998). PGIs catalyze the intracellular isomerization reaction of D-glucose-6-phosphate (G6P) to D-fructose-6-phosphate (F6P). The proposed catalytic mechanism involves several steps via the formation of a *cis*-enediol intermediate, similar to triose phosphate isomerase.

The sequence similarity between PGIs and neuroleukin (NL; Faik *et al.*, 1988), autocrine motility factor (AMF; Li & Chirgwin, 2000) and differentiation and maturation mediator (DMM; Xu *et al.*, 1996) indicates a broader role for this enzyme in cell function. Recent investigation has shown the stimulation of mouse tumour cells' motility and the enhancement of neurite outgrowth in progenitors of neuronal cells by the addition of *Bacillus stearothermophilus* PGI to the culture medium (Sun *et al.*, 1999).

The human homologue hPGI and its iso-enzymes have been purified from several tissues with no particular tissue specificity. hPGI has been characterized by the analysis of human genetic defects affecting the *pgi* gene (Xu & Beutler, 1994). Defects in the *hpgi* loci are the cause of many autosomal recessive disorders such as haemolytic anaemia. Depending on the severity of the PGI deficiency, neonatal death can occur from neurological disorders.

The crystallographic structures of the rabbit, *Bacillus* and pig PGI enzymes have been

previously reported at resolutions ranging from 2.3 to 6.0 Å (Jeffery *et al.*, 2000; Hsiao *et al.*, 1997; Muirhead & Shaw, 1974). Although the sequence similarity of hPGI, cytokines and growth factors reveal a broader enzyme function, no structural data has been obtained for the human enzyme. The availability of a crystal structure of hPGI could help elucidate the enzyme catalytic mechanism and its substrate specificity. We present here the first successful expression, crystallization and preliminary X-ray characterization of the human PGI enzyme. The data obtained shows that the hPGI enzyme crystallizes in conditions similar to those of the rabbit homologue (Jeffery *et al.*, 2000).

2. Materials and methods

2.1. Human PGI expression, purification and characterization

The *hpgi* gene was amplified from a human brain cDNA library (GIBCO-BRL) by the polymerase chain reaction (PCR) based on the available human PGI sequence (accession No. NM_000175). Oligodeoxynucleotide primers for PCR amplification (GIBCO-BRL) introduce both *Nde*I and *Hind*III restriction sites at the 5' and 3' end, respectively (5'-AGAGCT-CCCATATGGCCGCTCTCACCCGGGAC-3', *Nde*I; 5'-agctaaagctttattggactctggcctcgcg-3', *Hind*III) for cloning into the pET29a(+) expression vector (Novagen). The PCR reaction, containing 2 pmol of each primer and approximately 50 ng of the human brain cDNA, was carried out in a GeneAmp 2400 thermocycler (Perkin-Elmer CETUS) with 2.5 U of AmpliTaq DNA polymerase (Promega) according to the manufacturer's

instructions. The sample was subjected to 2 min denaturation at 367 K followed by 30 cycles of denaturation at 367 K for 0.5 min, annealing at 313 K for 0.5 min and extension at 345 K for 1 min. A DNA band of approximately 1.8 kbp in length was gel purified by the NaI glass powder method (Vogelstein & Gillespie, 1979). The purified DNA was digested with the restriction enzymes *NdeI* and *HindIII*, subcloned into the pET29a(+) vector and transformed into BL21(DE3) competent cells. The cloned *hpgi* was sequenced in an ABI377 DNA sequencer to confirm its sequence identity. Cells from a single colony were grown overnight at 310 K, 250 rev min⁻¹ in 5 ml LB medium containing 50 µg ml⁻¹ kanamycin. A larger cell culture was grown in 2×YT with 25 µg ml⁻¹ kanamycin at 310 K, 250 rev min⁻¹ until OD₆₀₀ = 0.6. The culture was then induced for 4 h with 0.5 mM IPTG.

The *E. coli* BL21(DE3) cells were harvested by centrifugation at 4000g for 30 min. The cell pellet was suspended in 20 ml of 50 mM Bicine pH 7.8 (buffer A) and cell lysis was obtained by adding lysozyme (0.5 mg ml⁻¹) and repeated cycles of freeze-thawing. The crude extract was clarified by centrifugation (6000g, 10 min) and brought to 50%(w/v) ammonium sulfate for 20 min with slow shaking. The suspension was separated by centrifugation for 15 min at 20 000g, 277 K and subjected to a second precipitation step in 80%(w/v) ammonium sulfate as described above. The protein pellet was solubilized in 9 ml buffer A and dialyzed against 250 ml of the same buffer with two buffer exchanges. The hPGI from the 50–80% fraction was loaded onto a 16 ml SP Sepharose-HP column equilibrated with buffer A. The hPGI eluted at approximately 1.3 M NaCl in a linear gradient of buffer B (buffer A plus 2 M NaCl). hPGI was dialyzed against 50 mM Bicine, 150 mM NaCl pH 7.8 and applied to a 115 ml Superdex-200 column equilibrated with 50 mM Bicine, 150 mM NaCl pH 7.8. The human PGI solution was concentrated by ultrafiltration using Centriprep 30 and Centricon 10 (Millipore) to 8 mg ml⁻¹ in 12.5 mM Bicine pH 7.5, 37.5 mM NaCl. The isoelectric point for the recombinant hPGI was determined by gel isoelectric focusing in the pH range 3–9 (PHAST system, Pharmacia). The kinetic parameters K_m and V_{max} were measured at room temperature following the formation of NADPH at 340 nm for 1 min (Gracy & Tilley, 1975).

2.2. Crystallization, diffraction data collection, processing and initial phasing

The crystallization conditions for the hPGI protein were initially screened by the hanging-drop vapour-diffusion method using the sparse-matrix kits Crystal Screen 1 and 2 from Hampton Research at 277 and 291 K. A solution of 3 µl hPGI protein containing 8 mg ml⁻¹ enzyme in 12.5 mM Bicine pH 7.5, 37.5 mM NaCl was mixed with an equal volume of the Crystal Screen 1 and 2 well solutions (500 µl per well) to form the drop. Screening different concentrations of precipitating agent and pH further optimized the initial crystallization conditions.

The best crystallization condition for the hPGI was obtained at 291 K in droplets consisting of 3 µl of protein solution at 8 mg ml⁻¹ in the buffer described above and 3 µl of reservoir solution. The reservoir solution (500 µl) contained 10% PEG 10000, 100 mM HEPES pH 7.5.

Diffraction data were collected on an R-AXIS IIC image-plate detector mounted on a Rigaku RU200B rotating-anode generator operating with a copper target at 50 kV and 100 mA. Crystals were mounted in nylon loops (Hampton Research) after quick soaking in a cryoprotectant solution consisting of the same reservoir solution and 20% MPD. A total of 79 frames were collected using the oscillation method. Individual frames consisted of 1° oscillation and 30 min exposure at a crystal-to-detector distance of 130 mm. The data were processed using *DENZO* and *SCALEPACK* from the *HKL* suite (Otwinowski & Minor, 1997). The Matthews coefficient and solvent content were calculated with the program *MATTHEWS_COEFF*. The self-rotation function was calculated using *AMoRe* (Navaza, 1997) and *GLRF* (Tong, 1997) in the resolution range 15–3.5 Å with an integration radius of 25 Å. The Patterson map was calculated using *CNS* (Brunger *et al.*, 1998) with 35–2.6 Å resolution data.

Initial phasing was obtained by molecular replacement using *AMoRe* (Navaza, 1997) and the dimeric rabbit PGI structure as model (PDB code 1dqr) at 2.5 Å. The search for the rotation and translation functions was performed using data in the resolution range 15–3.5 Å from the hPGI data set, equivalent to 28 025 reflections.

3. Results and discussion

The availability of large quantities of pure protein is of paramount importance for biophysical studies and especially for crystallographic structure determination. The

structures of PGI from *B. stearrowthermophilus* (Hsiao *et al.*, 1997) and more recently the rabbit phosphoglucose isomerases (rPGI) have been determined (Jeffery *et al.*, 2000). In spite of the intense investigation of this multifunctional enzyme, no structural data has been reported for the human homologue. Starting from a human brain cDNA library, the procedure leading to the expression, purification, crystallization and preliminary crystallographic analysis of human PGI is described.

The recombinant hPGI protein was overexpressed and purified to homogeneity by a three-step purification protocol. Spectrophotometric assays and electrophoresis followed the purity of the sample. On a 15% SDS-PAGE, the purified hPGI protein migrates as a single band of about 60 kDa (Fig. 1a) corresponding to the expected hPGI monomer (Carter & Yoshida, 1969). The relative molecular size determined from size-exclusion chromatography showed that the recombinant hPGI elutes as a homodimer of 120 kDa, as previously reported for other PGI enzymes (Carter & Yoshida, 1969) (data not shown). The expression and purification protocol described here allowed the recovery of approximately 5 mg of pure recombinant hPGI enzyme, with an enzymatic specific activity of 440 U mg⁻¹. The kinetic parameters of hPGI were analysed and a K_m value of 96 mM for F6P and a V_{max} of 21 mM min⁻¹ were obtained, indicating that the recombinant hPGI enzyme is appropriate for crystallographic analysis and represents well the wild-type enzyme.

The hPGI protein was crystallized as described in §2. The best condition (10% PEG 10 000, 100 mM HEPES pH 7.5) resulted in the formation of rectangular-shaped crystals of approximately 0.1 × 0.1 × 0.6 mm dimensions after one week at 291 K. The crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 80.37$, $b = 107.54$, $c = 270.33$ Å (Fig. 1b). Statistical analysis of the merged reflections indicated that the collected data set is better described as having a limiting resolution of 2.6 Å, with a total of 147 487 recorded reflections merged to 69 187 unique reflections. The overall completeness and R_{merge} correspond to 94.7 and 9.9%, respectively, and in the last shell 95.6 and 28.5%, respectively. Assuming two dimers in the asymmetric unit, the calculated Matthews coefficient is 2.41 Å³ Da⁻¹ (Matthews, 1968). A solvent content of 48.64% was obtained considering a protein density of 1.34 g cm⁻³.

The self-rotation function analysis showed the presence of a single non-crystallographic twofold axis with a peak correlation coefficient of 30%. The peak is at Eulerian angles $\alpha = 42.5^\circ$, $\beta = 172.5^\circ$, $\gamma = 42.5^\circ$, corresponding to a twofold rotation axis approximately parallel to the unit-cell ab plane and at an angle of 45° to the a axis. Other peaks were observed with a lower height and no symmetry axis could be associated with them; therefore, we considered these to be spurious peaks. The twofold rotation axis relates one molecule of the dimer to the other, representing an intradimer twofold axis. These results were confirmed by *GLRF* (Tong, 1997). The stereographic projection of section $\chi = 180^\circ$, represented in Fig. 2, shows the three crystallographic axes along a , b and c and a single non-crystallographic axis between axes a and b .

The molecular-replacement studies using the dimeric rabbit PGI (rPGI) as the search model, having 92% sequence identity to the hPGI, indicated a solution for the rotation and translation functions with two dimers in the asymmetric unit. A final correlation coefficient of 68.8% was found, which was well above the remaining peaks. The observed positions for the four independent hPGI molecules in the asymmetric unit are in agreement with the expected twofold non-crystallographic axis as seen by *GLRF* and *AMoRe*.

The molecules are organized as two independent dimers identically oriented and related by a translation of (0.5, 0.5, 0.134) in fractional coordinates along the cell axes. The calculated Patterson map shows a strong peak associated with those coordinates. This peak is 17.6% of the height of the origin peak, corresponding to a 75σ level, and is the only significant peak in the whole map apart from the origin.

Comparing the rPGI and hPGI proteins, many similarities can be observed with respect to their crystallization conditions, unit-cell parameters, Matthews coefficient and solvent content (Table 1). As a consequence, the same number of molecules in the unit cell are expected, though in a

somewhat different arrangement since the proteins crystallized in different space groups: $C222_1$ for rPGI and $P2_12_12_1$ for hPGI. Indeed, in the rabbit crystal, two dimers are related by the exact crystallographic C -centring operation, whereas in the human PGI crystal two dimers are related by a non-crystallographic translation (Figs. 2 and 3).

Finally, the successful crystallization of human PGI enzyme suitable for structure determination should allow us to answer many of the fundamental questions that remain unclear about the mechanisms of catalysis and of its multifunctional activity. We are currently refining the molecular-replacement solution.

Figure 1

Purification and crystallization of human PGI. (a) SDS-PAGE analysis of hPGI purification results. Lane 1, protein pellet after 50–80% ammonium sulfate fraction; lane 2, eluted fraction from SP Sepharose-HP column; lane 3, eluted fraction after Superdex-200; lane 4, molecular-weight markers (kDa). (b) Human PGI crystals obtained in 10% PEG 10 000, 100 mM HEPES pH 7.5 at 291 K by the hanging-drop vapour-diffusion technique. The lower left crystal has approximate dimensions of $0.1 \times 0.1 \times 0.6$ mm.

Figure 2

Preliminary crystallographic analysis. Stereographic projection of $\kappa = 180^\circ$ section of the self-rotation function map of hPGI limited to polar angles $\phi = 0-105^\circ$, $\psi = -105^\circ$ contoured at 0.6σ as calculated with the program *GLRF*.

Figure 3

Ribbon representation of the two independent hPGI dimers in the crystal unit cell. (a) View down the non-crystallographic twofold axis relating the monomers in each dimer; (b) view down the crystal c axis. 'N' (in blue) indicates a single non-crystallographic axis between axes a and b .

Table 1

Comparison between hPGI and rPGI crystallization conditions and crystal characteristics.

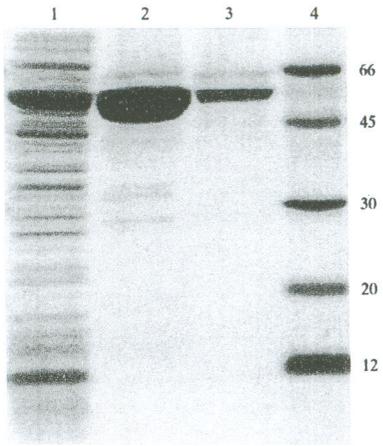
	hPGI	rPGI
Crystallization conditions	10% PEG 10 000, 100 mM HEPES pH 7.5	13% PEG 8000, 250 mM Mg acetate, 100 mM Na cacodylate pH 6.5
Space group	$P2_12_12_1$	$C222_1$
Unit-cell parameters (Å)		
a	80.37	82.69
b	107.54	115.27
c	270.33	271.84
Matthews coefficient	2.68	2.41
Solvent content (%)	53.70	48.64

This work was supported in part by a research grant 98/14979-7 to OHT (FAPESP). GO is an International Scholar of the Howard Hughes Medical Institute. ATC and PHCG are FAPESP student fellowship awardees. We would like to thank the members of the Protein Crystallography and Structural Biology Group (IFSC-USP) for helpful discussions in the course of this work and to express our gratitude to Dr Flávio H. Silva (DGE-UFSCar) for making available the Human Brain cDNA library used in this study.

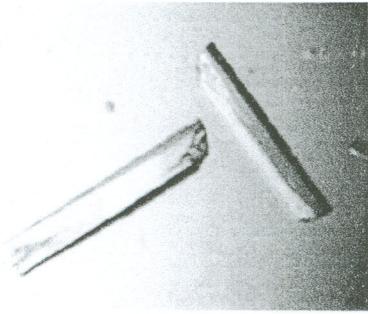
References

- Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Ponten, T., Alsmark, U. C., Podowski, R. M., Naslund, A. K., Eriksson, A. S., Winkler, H. H. & Kurland, C. G. (1998). *Nature (London)*, **396**, 133–140.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst. D54*, 905–921.
- Carter, N. D. & Yoshida, A. (1969). *Biochim. Biophys. Acta*, **181**, 12–19.
- Falk, P., Walker, J. I., Redmill, A. A. & Morgan, M. J. (1988). *Nature (London)*, **332**, 455–457.
- Gracy, R. W. & Tilley, B. E. (1975). *Methods Enzymol.* **41**, 392–400.
- Hsiao, C. D., Chou, C. C., Hsiao, Y. Y., Sun, Y. J. & Meng, M. (1997). *J. Struct. Biol.* **120**, 196–200.
- Jeffery, C. J., Bahson, B. J., Chien, W., Ringe, D. & Petsko, G. A. (2000). *Biochemistry*, **39**, 955–964.
- Li, X. & Chirgwin, J. M. (2000). *Biochim. Biophys. Acta*, **1476**, 363–367.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Muirhead, H. & Shaw, P. J. (1974). *J. Mol. Biol.* **89**, 195–203.
- Navaza, J. (1997). *Methods Enzymol.* **276**, 581–594.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Sun, Y. J., Chou, C. C., Chen, W. S., Wu, R. T., Meng, M. & Hsiao, C. D. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 5412–5417.
- Tong, L. (1997). *Methods Enzymol.* **276**, 594–611.
- Vogelstein, B. & Gillespie, D. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 615–619.
- Xu, W. & Beutler, E. (1994). *J. Clin. Invest.* **94**, 2326–2329.
- Xu, W., Seiter, K., Feldman, E., Ahmed, T. & Chiao, J. W. (1996). *Blood*, **87**, 4502–4506.

GR2113.Fig. 1

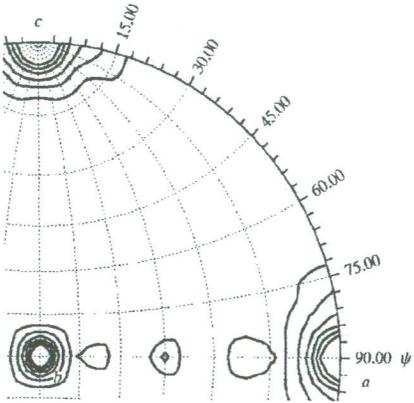


(a)

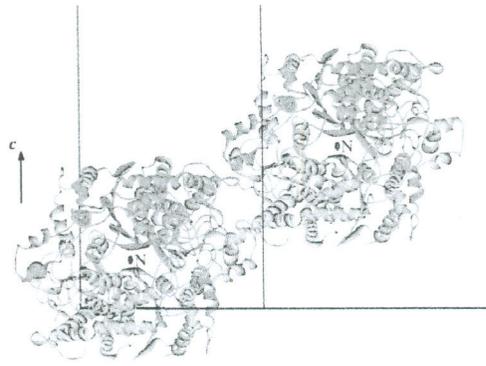


(b)

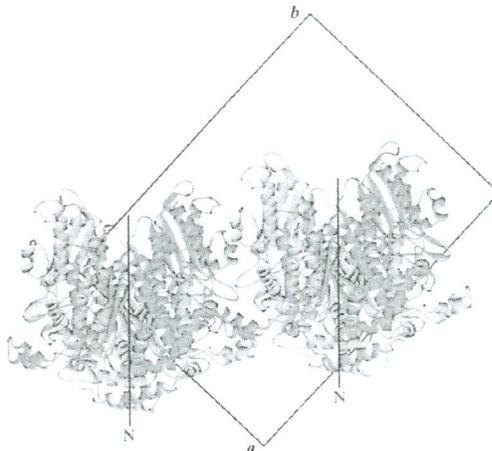
GR2113.Fig. 2



GR2113.Fig. 3



(a)



(b)

REFERÊNCIAS

- [1] STRYER, L. (1992). Bioquímica. 3^o ed. Editora Guanabara Koogan S.A. Rio de Janeiro p 881.
- [2] BAUGHAN, M.A. et al (1968). Hereditary hemolytic anemia associated with glucosephosphate isomerase (GPI) deficiency - A new enzyme defect of human erythrocytes. *Blood* 32:236.
- [3] FUJII, H. et al (1990). Recent progress in the molecular genetic analysis of erythroenzymopathy. *Am. j. Hematol.* 34:301.
- [4] ARYA, R. et al (1995). Hereditary red cell enzymopathies. *Blood Rev.* 9:165.
- [5] KUGLER, W. et al (1998). Molecular basis of neurological dysfunction coupled with haemolytic anaemia in human glucose-6-phosphate isomerase (GPI) deficiency. *Hum. Genet.* 103(4):450-454.
- [6] NIINAKA, Y. et al (1998). Expression and secretion of neuroleukin/phosphohexose isomerase/maturation factor as autocrine motility factor by tumor cells. *Cancer Rev.* 58(12):2667-2674.
- [7] MIZRACHI, Y. (1989). Neurotrophic activity of monomeric glucosephosphoisomerase was blocked by human immunodeficiency virus (HIV-1) and peptides from HIV-1 envelope glycoprotein. *J. Neurosci. Res.* 23(2):217-224.
- [8] THOMAS, H. et al (1991). Effects of interferons and other cytokines on tumors in animals: a review. *Pharmacol Ther* 52(3):307-330.
- [9] WATANABE, H. et al (1996). Tumor cell autocrine motility factor is the neuroleukin/phosphohexose isomerase polypeptide. *Cancer Rev.* 56(13):2960-2963.
- [10] XU, W. et al (1996). The differentiation and maturation mediator for human myeloid leukemia cells shares homology with neuroleukin or phosphoglucose isomerase. *Blood* 87(11):4502-4506.
- [11] <http://www.rcsb.org/pdb>
- [12] DUCRIX, A. e CIEGÉ, R. (1992) *Crystallization of Nucleic Acids and Proteins / A Practical Approach.* IRL PRESS, Oxford, p331.

- [13] AUSUBEL, F. et al (1995). Short Protocols in Molecular Biology. 3° ed. , Wiley & Sons.
- [14] CARTER, N. D. et al (1969). Purification and characterization of human phosphoglucose isomerase. *Biochim. Biophys* 181:12-19.
- [15] DOONAN, S. (1996) Protein Purification Protocols. 1° ed . Human Press Inc. p. 405
- [16] NELSON, D.L. E COX, M.M.(2000). Lehninger Principles of Biochemistry. 3° ed. Worth Publishers, New York. p1152
- [17] GRACY, R. W. & TILLEY, B. E. (1975). Phosphoglucose isomerase of human erythrocytes and cardiac tissue. *Methods Enzymol.*;41:392-400.
- [18] HARDRÉ, R. and SALMON, L. (1999). Competitive inhibitors of yeast phosphoglucose isomerase: synthesis and evaluation of new types of phosphorylated sugars from the synthon D-arabinonolactone-5-phosphate. *Carbohydrate Research* 318: 110-115.
- [19] JEFFERY, C. J., BAHNSON, B. J., CHIEN, W., RINGE, D. & PETSKO, G. A. (2000). Crystal structure of rabbit phosphoglucose isomerase, a glycolytic enzyme that moonlights as neuroleukin, autocrine motility factor, and differentiation mediator. *Biochemistry*. 39(5):955-964
- [20] HARDRÉ, R. et. al. (1998). Competitive inhibitors of yeast phosphoglucose isomerase: synthesis and evaluation of new types of phosphorylated sugars from the synthon D-arabinolactone-5-phosphate. *Bioorg. & Med. Chem. Letters*, 8: 3435-3438
- [21] SHERWOOD, D. (1976). Crystals, X-rays and Protein. 1ª edição, Longman Group Limited . London, p702.
- [22] DRENTH, J. (1994). Principles of Protein X-ray crystallography. Springer-Verlag. New York, p311.
- [23] DAUTER, Z.(1999). Data-collection strategies. *Acta Cryst.* D55:1703-1717.
- [24] STOUT, G.H. e JENSEN, L.H. (1989). X-Ray Structure Determination-A practical Guide. 2ª edição, John Wiley & Sons. New York, p453.
- [25] OTWINOWSKI, Z. e MINOR, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode *Methods Enzymol.* 276, 307-326.
- [26] NAVAZA, J. (1997). *Methods Enzymol.* 276, 581-594.

- [27] BRUNGER, A.T. et al (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D54*: 905-921.
- [28] MATTHEWS, B. W. (1968). Solvent content of protein crystals. *J.Mol.Biol.* 33, 491-497.
- [29] TONG L. (1997). Rotation function calculations with GLRF program. *Methods Enzymol.* 276, 594-611.
- [30] BRUNGER, A.T. e RICE, L.M. (1997). Crystallographic Refinement by Simulated Annealing: Methods and Applications *Methods Enzymol.* 277:243-269.
- [31] BRUNGER, A.T. (1997). Free R Value: Cross-Validation in Crystallography *Methods Enzymol.* 277:367-396.
- [32] BRUNGER, A.T *et al* (1997). Cross-Validated Maximum Likelihood Enhances Crystallographic Simulated Annealing Refinement. *Proc. Natl. Acad. Sci. USA*, 94:5018-5023.
- [33] BRUNGER, A.T *et al* (1999). Extending The Limits of Molecular Replacement Through Combined Simulated Annealing and Maximum-Likelihood Refinement. *Acta Cryst. D55*:181-190.
- [34] ROSE, I.A. (1962) *Brookhaven Symp. Biol.* 15:293-309.