

Marcos Michel de Souza

**Estudos estruturais da Purina Nucleosídeo
Fosforilase do *Mycobacterium tuberculosis*
complexada com ligantes.**

Dissertação para a obtenção do
título de mestre em Biofísica
Molecular, Área de Concentração -
Biofísica Molecular. Instituto de
Biociências, Letras e Ciências Exatas
da Universidade Estadual Paulista
"Júlio de Mesquita Filho" - UNESP.

Orientadora: Prof. Dra. Fernanda Canduri
Co-orientador: Prof. Dr. Walter Filgueira de Azevedo Jr.

São José do Rio Preto – SP

2007

Souza, Marcos Michel de.

Estudos estruturais da Purina Nucleosídeo Fosforilase do
Mycobacterium tuberculosis complexada com ligantes / Marcos Michel
de Souza. - São José do Rio Preto : [s.n], 2007.

85 f. : il. ; 30 cm.

Orientadora: Fernanda Canduri

Co-orientador: Walter Filgueira de Azevedo Júnior

Dissertação (mestrado) – Universidade Estadual Paulista, Instituto de
Biociências, Letras e Ciências Exatas

1. Biologia molecular. 2. Biofísica molecular. 3. Cristalografia. 4.
Proteínas - Estrutura. 5. Purina Nucleosídeo Fosforilase. 6. PNP
(Enzimas) 7. MtPNP (Enzima) I. Canduri, Fernanda. II. Azevedo
Júnior, Walter Filgueira de. III. Universidade Estadual Paulista, Instituto
de Biociências, Letras e Ciências Exatas. IV. Título..

CDU – 577.322

Dedico este trabalho aos meus pais, por sempre me ensinarem valores e virtudes de honradez e descência de caráter que hoje norteiam toda a minha vida e o meu trabalho acadêmico.

Agradecimentos

Gostaria de agradecer, somente a Deus.

Agradeço a Deus por todas as pessoas que ele colocou no meu caminho durante minha vida acadêmica.

Agradeço a Deus pela vida da minha orientadora de iniciação científica, Maria da Piedade Rezende da Costa, que me ensinou os primeiros passos da vida acadêmica e me apoiou a vinda para a Biofísica.

Piedade costuma dizer que o orientador é o pai acadêmico do aluno, e eu tenho muito a agradecer a Deus por essa grande família que ele me deu. Todos foram muito importantes na minha vida, minha orientadora Dra. Fernanda Canduri e meu co-orientador Dr. Walter Filgueira de Azevedo Júnior, e todos os meus irmãos acadêmicos que eu tanto amo: minha querida e sincera Alessandra Renata, o dedicado e inteligentíssimo Márcio Vinícios, meus camaradas Paulo Camilo e Rodrigo Galo, Janaína, Helen, e todos os demais.

Agradeço a Deus por ter colocado em minha vida professores muito especiais, aos quais hoje tenho o mais profundo afeto e admiração: Dr. José Ruggiero, meu colega do coral que tanto me fez sofrer na disciplina "Forças Intermoleculares", Dr. Eslo Drigo Filho, com sua gargalhada inconfundível, Dr. Jorge Chaíne, o coordenador da pós, que sempre me atendeu prontamente todas as vezes que precisei de sua ajuda, e a todos os outros professores, do departamento.

Agradeço a Deus pelos amigos que fiz no Departamento e por todas as pessoas que me ajudaram, minha amiga Cristiane, a sempre simpática Dona Rose, a prestativa Ilva, o Paulinho e o Barbosa sempre atenciosos.

Graças a Deus pelas agências de fomento à pesquisa CAPES e FAPESP sem as quais esse trabalho não seria possível.

Sou grato a Deus meus amigos de fora da Unesp, que também sempre me incentivaram nesta carreira: José Precioso Júnior, Anderson Rodrigues Silva, Fabiano Ferreira Ângelo, e ao meu grande amigo Thiago Laurêncio Rizzo.

Dou graças a Deus também pela vida do meu pai e minha mãe, que sempre lutaram para que eu pudesse estudar e sempre me deram incentivo para seguir na carreira acadêmica.

Agradeço a Deus até mesmo pelas pessoas que me fizeram algum mal, pois até as dificuldades me fizeram crescer.

Muito obrigado.

Índice Geral

Índice de Figuras	4
Índice de Tabelas	6
Lista de Abreviaturas	7
Resumo/Abstract	9/10
1. Introdução	11
1.1 A Tuberculose – uma emergência global	11
1.2 Metabolismo de Purinas – um caminho para novas drogas	14
1.3 A PNP – Purina Nucleosídeo Fosforilase	18
1.4 PNP's de diferentes origens e suas classificações	19
1.5 Análise do Mecanismo Catalítico	22
1.6 A PNP do <i>Mycobacterium tuberculosis</i> - MtPNP	25
2. Objetivos	28
3. Materiais e Métodos	29
3.1. Obtenção da proteína	29
3.2. Cristalização da MtPNP	29
3.3. Difração de Raios X	44
3.4. O Problema da Fase	47
3.5. Substituição Molecular	47
4. Resultados e Discussão	49
4.1. Cristalização e Coleta de Dados de Difração de Raios X	49
4.2. Substituição Molecular	51
4.3. Refinamento Cristalográfico	52
4.4. Descrição da Estrutura MtPNP:Adenina	53

4.5. Qualidade do Modelo	57
4.6. A Adenina no Centro do Trímero	59
4.7. O Sítio Ativo	61
4.8 O Estudo do Potencial Eletrostático	66
5. Conclusão	71
Referências	73
6. Apêndices: O problema da Fase e a Substituição Molecular	79

Índice de Figuras

Figura 1.1 - Taxas estimadas de tuberculose em 2003.	11
Figura 1.2 - (a) Reação fosforolítica reversível e (b) reação hidrolítica, catalisadas por PNPs de mamíferos	15
Figura 1.3 - Inibição ou deficiência da PNP	16
Figura 1.4 - Possíveis enovelamentos da Purina Nucleosídeo Fosforilase	22
Figura 1.5 - Possíveis ordens de ligação dos substratos	23
Figura 1.6 – Novo Mecanismo catalítico	24
Figura 1.7 - Diagrama do trímero e do monômero da MtPNP.	26
Figura 1.8 - Alinhamento das estruturas primárias da MtPNP e da HsPNP	26
Figura 1.9 - Comparação dos monômeros da HsPNP (A) e da MtPNP (B)	27
Figura 3.1 - Principais passos envolvidos na resolução da estrutura	31
Figura 3.2 - Representação esquemática de um diagrama de solubilidade	32
Figura 3.3 - Representação esquemática de gota suspensa (hanging drop), gota sentada (sitting drop) e gota sanduíche (sandwich drop).	34
Figura 3.4 - Ligantes utilizados	36
Figura 3.5 – Eixos Cristalinos	38
Figura 3.6 – Redes de Bravais	39
Figura 3.7 – Representação da Lei de Bragg	41
Figura 3.8 – Fluxograma do protocolo de substituição molecular	48
Figura 4.1 – Cristais da MtPNP	49
Figura 4.2 – Estrutura da MtPNP complexada com adenina.	53
Figura 4.3 – Densidade eletrônica da molécula de adenina	54

Figura 4.4 – <i>Omit Map</i> da adenina no centro do trímero	55
Figura 4.5 – <i>Omit Map</i> dos sítios ativos	56
Figura 4.6 – Diagrama de Ramachandran Φ - Ψ	57
Figura 4.7 – Densidade eletrônica da molécula de adenina no centro do trímero	59
Figura 4.8 – A molécula de adenina presente no centro do trímero	60
Figura 4.9 – Sítio Ativo da cadeia A	61
Figura 4.10 – Sítios ativos nas cadeias B e C.	62
Figura 4.11 – Sobreposição das estruturas de MtPNP adenina e hipoxantina	63
Figura 4.12 – Sítio ativo da MtPNP complexada com Dadme-Imucilina H	64
Figura 4.13 – Sítio ativo da MtPNP complexada com deazahipoxantina.	65
Figura 4.14 – Superfície de potencial eletrostático	66
Figura 4.15 – Superfície de potencial eletrostático (lado oposto)	67
Figura 4.16 – Superfície de potencial eletrostático (vista lateral)	68
Figura 4.17 – Superfície de potencial eletrostático do sítio ativo	69
Figura 4.18 – Superfície de potencial eletrostático do sítio ativo da HsPNP	70
Figura 4.19 – Detalhe do potencial eletrostático do sítio ativo da HsPNP	70

Índice de Tabelas

Tabela 1.1 Classificação de PNPs de diferentes origens	20
Tabela 3.1 Os tipos de redes em três dimensões	38
Tabela 4.1 Estatísticas do processamento de dados de difração de raios X para a MtPN:adenina	50
Tabela 4.2 Estatísticas do processamento de dados de difração de raios X para a MtPN:adenosina	50
Tabela 4.3 Valores obtidos na etapa de substituição molecular	51
Tabela 4.3 Estatísticas do refinamento cristalográficos da estrutura da MtPNP complexada com Adenina.	52

Lista de Abreviaturas

PNP	- Purina Nucleosídeo Fosforilase
MtPNP	- Purina Nucleosídeo Fosforilase do <i>Mycobacterium tuberculosis</i>
HsPNP	- Purina Nucleosídeo Fosforilase do <i>Homo sapiens</i> (humana)
PDB	- Protein Data Bank (Banco de dados de proteínas)
ADA	- Adenosina deaminase
AMP	- Adenosina monofosfato
GMP	- Guanosina monofosfato
IMP	- Inosina monofosfato
dGuo	- Desoxiguanosina
dGMP	- Desoxiguanosina monofosfato
dGDP	- Desoxiguanosina difosfato
dGTP	- Desoxiguanosina trifosfato
NDP	- Nucleotídeo difosfato
dNDP	- Desoxinucleotídeo difosfato
dNTP	- Desoxinucleotídeo trifosfato
Gua	- Guanina
ATP	- Adenosina trifosfato
M _r	- Massa molecular
CC	- Coeficiente de correlação
PC	- Correlação de Patterson
V _m	- Volume de Matthews
CCP4	- Collaborative Computational Project n° 4 (Projeto computacional n° 4)
MtPNP·PO ₄	- Purina Nucleosídeo Fosforilase do <i>Mycobacterium tuberculosis</i> em complexo com fosfato

- MtPNP·Acy·PO₄ - Purina Nucleosídeo Fosforilase do *Mycobacterium tuberculosis*
em complexo com Aciclovir e fosfato
- MtPNP·Hypo·PO₄ - Purina Nucleosídeo Fosforilase do *Mycobacterium tuberculosis*
em complexo com Hipoxantina e fosfato
- MtPNP·Ino·PO₄ - Purina Nucleosídeo Fosforilase do *Mycobacterium tuberculosis*
em complexo com Inosina e fosfato

Resumo

De acordo com a OMS, a tuberculose mata 5000 pessoas por dia, e se não controlada, são estimados 35 milhões de novos casos nos próximos vinte anos. A alta susceptibilidade dos infectados por HIV para a tuberculose, bem como a proliferação da tuberculose multidroga-resistente (MDR), tem criado um grande interesse mundial de expansão nos programas atuais de pesquisas sobre a tuberculose.

A Purina Nucleosídeo Fosforilase do *Mycobacterium tuberculosis* (MtPNP) é um potencial alvo para novas drogas anti-tuberculose visto que é responsável pela síntese *de novo* de ribonucleotídeos de purina. A Inibição específica da PNP poderia potencialmente levar o *M. tuberculosis* ao estado latente.

O objetivo deste trabalho é resolver a estrutura da MtPNP complexada com adenina, e analisar as interações com os ligantes. A MtPNP foi cristalizada utilizando condições experimentais descritas posteriormente, e o ligante (adenina) foi adicionado por soaking. Os dados de difração de raios X foram coletados no detector CCD utilizando fonte de radiação síncrotron (Laboratório Nacional de Luz Síncrotron, LNLS, Campinas, Brazil). O programa MOSFLM foi utilizado para o processamento, e os dados foram escalonados através do programa SCALA, apresentando o grupo espacial ortorrômico $P2_12_12$ ($a=118,96\text{\AA}$, $b=134,65\text{\AA}$, $c=44,42\text{\AA}$). O cristal foi determinado utilizando o método de substituição molecular através do programa AMoRe.

O modelo final obteve R_{free} e R_{factor} de 23.87% e 17.51% respectivamente, e resolução máxima de $1.86\text{\AA}$. Foi observada uma grande densidade eletrônica no centro do trímero, indicando a presença de adenina nesta posição. Este fato tem sido o alvo de atenção na análise da estrutura. A análise da estrutura da MtPNP complexada com adenina e o maior conhecimento sobre seu sítio ativo, fornece informações importantes sobre a base estrutural de novas drogas e novos inibidores mais específicos.

Abstract

In according to the WHO, the tuberculosis kills 5000 people every day, if does not controlled, are esteemed 35 millions of new cases in next 20 years. The high susceptibility of human immunodeficiency virus infected persons to the disease and the proliferation of multidrug-resistant (MDR) strains have created a worldwide interest in expanding current programs in tuberculosis research.

The Purine Nucleoside Phosphorylase from *Mycobacterium tuberculosis* (MtPNP) is a potential target for new drug anti-tuberculosis, whereas is responsible for *de novo* synthesis of purine ribonucleotides. The specific inhibition of *M. tuberculosis* PNP could potentially lead to the latent state of *M. tuberculosis*.

The objective of this work is to solve the structure from MtPNP complexed with adenine, and to analyze their interactions with the ligand. MtPNP was crystallized using the experimental conditions described elsewhere, and the ligand (adenine) was added by soaking. The X-ray diffraction data were collected on a CCD detector using synchrotron radiation source (Laboratório Nacional de Luz Síncrotron, LNLS, Campinas, Brazil). It was used the MOSFLM program to processing, and the data were scaled through the program SCALA, presenting orthorhombic spatial group $P2_12_12$ ($a=118,96\text{\AA}$, $b=134,65\text{\AA}$, $c=44,42\text{\AA}$). The crystal was determined by molecular replacement methods using the program AMoRe.

The final model has R_{free} and R_{factor} of 23.87% and 17.51% respectively, and maximum resolution of $1.86\text{\AA}$. It was observed a large electronic density in the center of the trimer, indicating the presence of adenine in this position. This fact was one of aims of attention in the structure analysis. The analysis of the structure of MtPNP complexed with adenine and the larger knowledge about the active site provides important information for the structure-based of new drugs and new inhibitors more specific.

1. Introdução

1.1 Tuberculose – uma emergência global.

A cada dia 5000 pessoas são mortas pela Tuberculose no mundo, sendo que 98% das mortes ocorrem nos países em desenvolvimento, atacando principalmente pessoas mais vulneráveis, como os mais pobres e desnutridos. As taxas de Tuberculose mais altas por habitantes são na África, que detêm mais de um quarto dos casos. No que se refere aos novos casos da doença, mais da metade se situam em seis países asiáticos: Bangladesh, China, Filipinas, Índia, Indonésia e Paquistão, (figura 1.1), (WHO, Global Factsheet, 2005; NUNN et al 2005). Os interesses comerciais da indústria farmacêutica e o fato da tuberculose ser uma doença típica de países ditos subdesenvolvidos, podem gerar um desinteresse na pesquisa de novos tratamentos da doença, cabendo assim às pesquisas acadêmicas gerar suporte para o desenho de novas drogas.

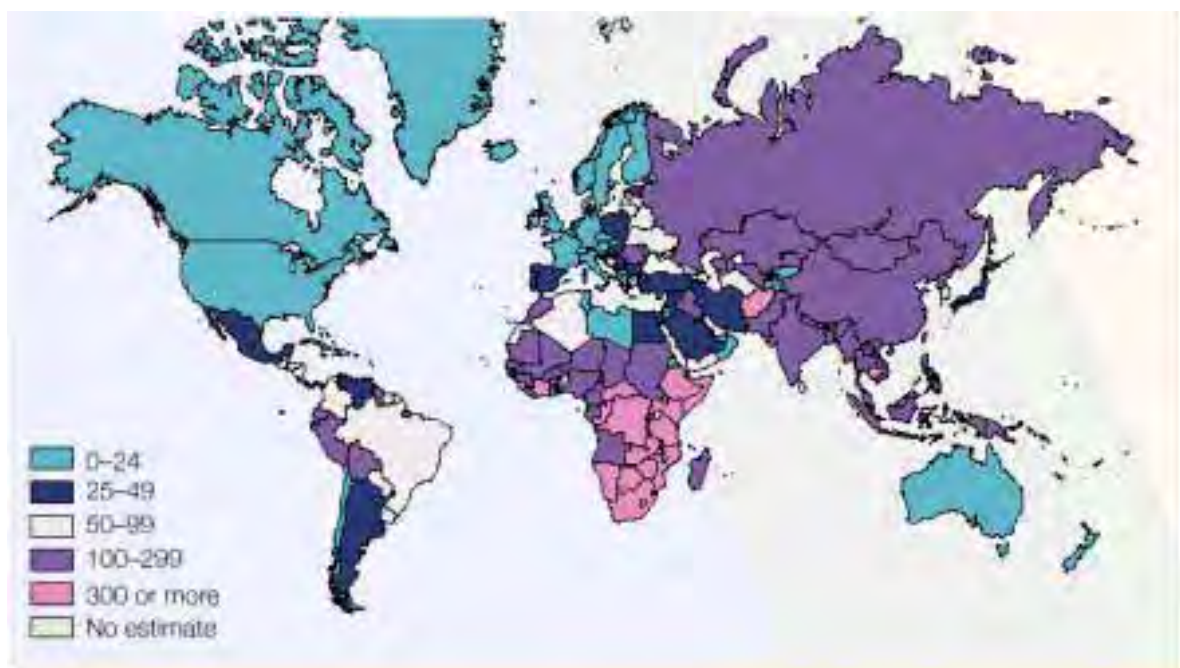


Figura 1.1 Taxas estimadas de tuberculose em 2003, (NUNN et al, 2005)

Em alguns países a epidemia da Tuberculose decresce ou se mantém estabilizada, em outras regiões, no entanto, têm uma altíssima taxa de crescimento, o que resulta num aumento global da epidemia de 1% ao ano. Em 2003 surgiram 8,8 milhões de novos casos de Tuberculose, sendo que 80% desses novos casos se

concentram em somente 22 países. O Brasil ocupa a 15ª posição entre os 22 países que configuram esta lista, tendo cerca de 90.000 casos notificados a cada ano. Em 2004 a Tuberculose foi declarada como prioridade nacional para o período de 2004-2007 (WHO, Global Factsheet, 2005, Programa de Controle de Tuberculose do Estado do Rio de Janeiro, Boletim Informativo – 2005).

A Tuberculose é uma doença infecciosa crônica, causada pelo *Mycobacterium tuberculosis* (“bacilo de Koch”), que afeta principalmente os pulmões mas pode causar lesões em qualquer parte do corpo (COTRAN, 1991).

Existem evidências de Tuberculose desde os tempos pré-históricos, encontradas em múmias do Egito antigo, e, mais recentemente, em uma múmia pré-colombiana no Peru. Mesmo sendo há muitos séculos reconhecida como uma entidade clínica grave e amplamente difundida, só no início do século passado que algumas de suas diversas manifestações foram percebidas como fazendo parte do mesmo processo. A natureza transmissível da tuberculose, no entanto, só foi reconhecida em 1868, quando Villemin produziu uma doença em coelhos semelhante à tuberculose humana, injetando material de lesões tuberculosas de origem humana nos animais (DAVIS et al., 1973; DIVISÃO DE CONTROLE DA TUBERCULOSE CVE – SES/SP www.saude.sp.gov.br/tuberculose/; SOCIEDADE BRASILEIRA DE PNEUMOLOGIA E TISIOLOGIA www.sbpt.org.br).

Durante a revolução industrial a mortalidade por tuberculose alcançou altas taxas, a doença disseminou-se na Europa, com a urbanização crescente e no século XVIII tornou-se conhecida como a peste branca. Chegou até mesmo a ser considerada uma "doença romântica", idealizada nas obras literárias e artísticas e identificada como uma doença de poetas e intelectuais nos últimos anos do século XVIII. A Tuberculose prosseguiu a sua caminhada mortífera e nos séculos 17 e 18 levava uma entre cinco

vidas. No ano de 1882, Robert Koch anuncia a descoberta do agente causador da tuberculose, o bacilo de Koch, cientificamente denominado *Mycobacterium tuberculosis*. Uma vez descoberto o bacilo vários tratamentos foram propostos e falharam, como certos climas, sanatórios, colapsoterapia (especialmente o pneumotórax terapêutico) descoberto por Forlanini, sais de ouro, de cobre, assim como o uso da tuberculina, descoberta pelo próprio Roberto Koch, que chegou a anunciar a sua descoberta como cura definitiva. (DAVIS et al.,1973; Divisão de Controle da Tuberculose CVE – SES/SP www.saude.sp.gov.br/tuberculose/; SOCIEDADE BRASILEIRA DE PNEUMOLOGIA E TISIOLOGIA www.sbpt.org.br).

Uma variante genética (cepa) do *Mycobacterium tuberculosis* resistente às drogas existentes, está presente em quase todos os 109 países estudados pela Organização Mundial de Saúde (OMS) e seus associados. A cada ano surgem 425.000 novos casos de Tuberculose Multidroga Resistente (TB-MDR), as taxas mais altas de TB-MDR correspondem aos países da antiga União Soviética e China, onde até 14% dos casos novos de Tuberculose não respondem aos tratamentos farmacológicos convencionais (WHO, Global Factsheet, 2005).

A alta susceptibilidade das pessoas infectadas com a Síndrome de Imunodeficiência Adquirida (AIDS) à Tuberculose, assim como a proliferação de cepas resistentes a multidroga, criaram um interesse mundial na expansão dos programas atuais de pesquisa sobre a Tuberculose. Novos agentes anti-micobacteriais são necessários para o tratamento e para a diminuição das complicações e dos efeitos colaterais sentidos pelos pacientes (SHI et al, 2001; NOLASCO 2005).

1.2 Metabolismo de Purinas – um caminho para novas drogas

Os nucleotídeos são ésteres fosfatados de uma pentose (ribose ou desoxirribose) nos quais uma base púrica ou pirimídica está ligada ao carbono 1 do açúcar. Os nucleosídeos trifosfatos são as unidades monoméricas que atuam como precursores de ácidos nucleicos. A importância dos nucleotídeos no metabolismo celular é indicada pela observação de que quase todas as células podem sintetizá-los “de novo”, e a partir dos produtos de degradação dos ácidos nucleicos (VOET & VOET, 2000).

Processos bioquímicos envolvendo nucleotídeos e precursores de nucleotídeos são algumas das reações mais importantes e fundamentais para a manutenção da vida (AGARWAL & PARKS 1972; DOS SANTOS, 2004). A descoberta da reação enzimática responsável pela clivagem da ligação N-glicosídica é atribuída a Levene e Medigreceanu em 1911. Kalckar, em 1945, mostrou o segundo produto da reação como sendo ribose-1-fosfato em vez de ribose como era esperado na época. Esta nova evidência forneceu argumentos para afirmar que a clivagem N-glicosídica era devido a fosforólise enzimática (Figura 1.2a), ao invés da hidrólise, e foi confirmada, pelo próprio Kalckar, em 1947 (PUGMIRE & EALICK, 2002; DOS SANTOS 2004).

A maioria dos alimentos, por serem de origem celular, contém ácidos nucleicos. Os ácidos nucleicos da dieta sobrevivem ao meio ácido do estômago e são degradados até seus nucleotídeos componentes, principalmente no intestino, pelas nucleases pancreáticas e pelas fosfodiesterases intestinais. Os nucleotídeos iônicos, que não podem atravessar as membranas celulares, são hidrolisados a nucleosídeos por uma variedade de nucleotidasas grupo-específicas e fosfatases inespecíficas. Assim, os nucleosídeos podem ser absorvidos diretamente pela mucosa intestinal e degradados até

as bases livres e ribose-1-fosfato pela ação de nucleosidases ou nucleosídeo fosforilases (VOET & VOET, 2000; DOS SANTOS 2004).

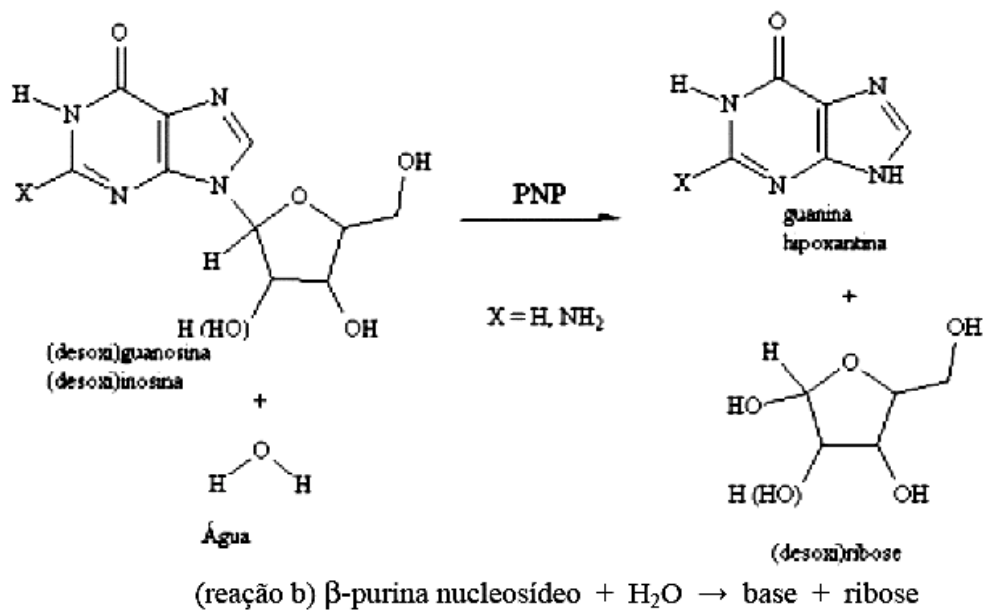
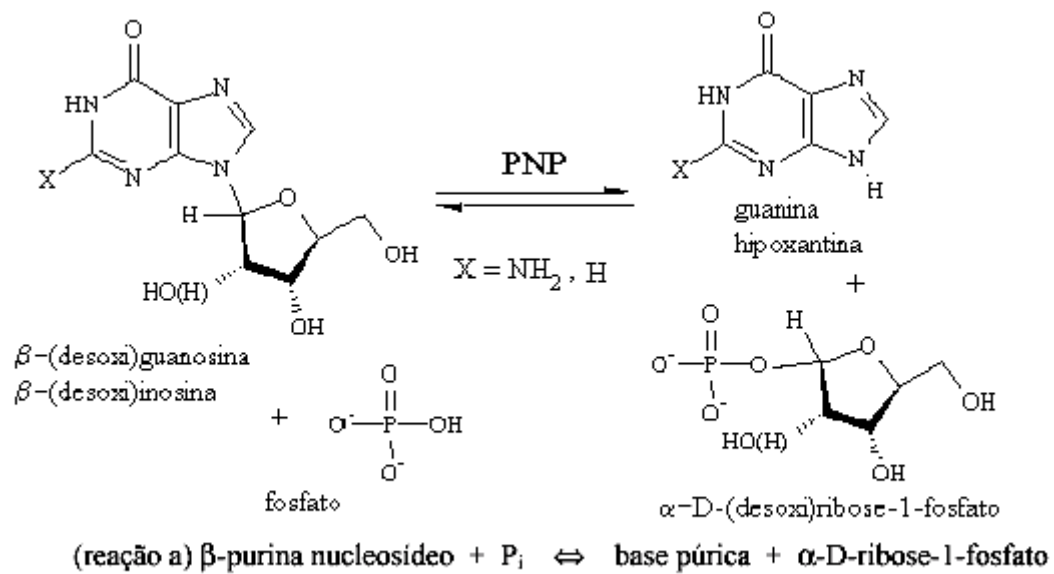
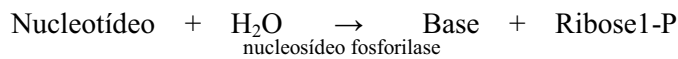
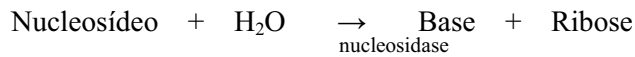


Figura 1.2 (a) Reação fosforolítica reversível, (b) reação hidrolítica, catalisadas por PNPs de mamíferos (DOS SANTOS, 2004).

Apenas uma pequena fração das bases dos ácidos nucléicos ingeridos são incorporados em ácidos nucléicos nos tecidos. Assim, as rotas de biossíntese de novo de nucleotídeos satisfazem plenamente as necessidades de nucleotídeos de um organismo. Conseqüentemente, as bases ingeridas são, em sua maioria, degradadas e excretadas. Os ácidos nucléicos celulares são também submetidos à degradação contínua como parte da reutilização contínua de quase todos os componentes celulares (VOET & VOET, 2000).

Na síntese *de novo* de ribonucleotídeos de purina a formação de adenosina monofosfato (AMP) e guanosina monofosfato (GMP), a partir de inosina monofosfato (IMP), é irreversível, mas bases púricas, nucleosídeos e nucleotídeos podem ser interconvertidos através da atividade de PNP, ADA e hipoxantina-guanina fosforribosil transferase (AVARBOCK ET AL., 1999). A deficiência ou inibição da PNP, então, levaria à não conversão de desoxiguanosina (dGuo) em guanina (Gua), acarretando no acúmulo de desoxiguanosina trifosfato (dGTP) o que inibiria alostericamente a enzima Ribonucleotídeo redutase e, conseqüentemente, a síntese de DNA celular (Figura 1.3) (BZOWSKA et al., 2000; NOLASCO 2005).

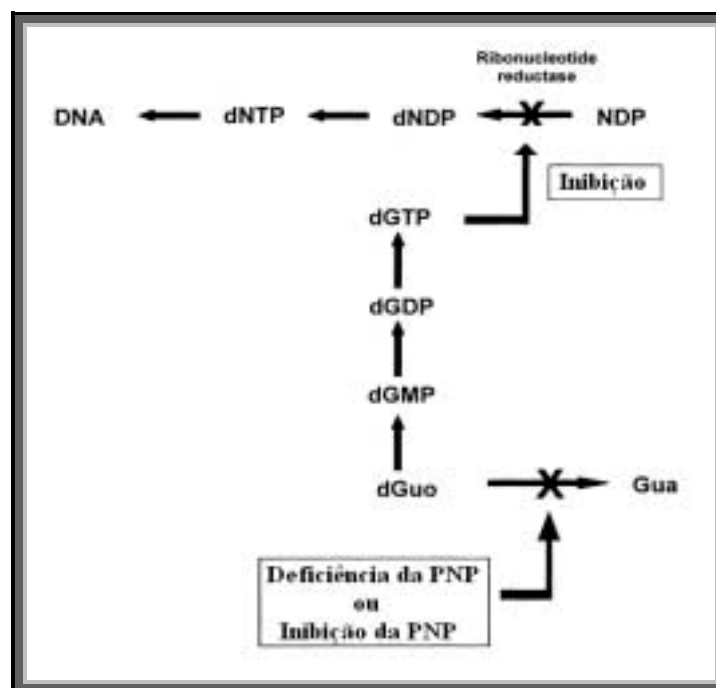


Fig 1.3 Inibição ou deficiência da PNP: levando ao acúmulo de dGTP, inibindo o Ribonucleotídeo redutase e a síntese de DNA (BZOWSKA et al., 2000; NOLASCO 2005).

No que concerne ao metabolismo humano de nucleosídeos de purina, as recorrentes deficiências hereditárias das enzimas purina nucleosídeo fosforilase (PNP) ou adenosina desaminase (ADA) causam anormalidades que são seletivamente tóxicas para linfócitos e resultam em doenças do sistema imune. Ambas as condições são herdadas em uma forma autossômica recessiva, e resultam de mutações em genes estruturais localizados nos cromossomos 14q (PNP) e 20q (ADA) (HERSHFIELD & KREDICH, 1989; DOS SANTOS 2004).

A PNP humana tem sido considerada como alvo primário para a ação de drogas desde o trabalho de Giblett onde foi relatada a ausência da atividade da HsPNP no sangue de uma criança com linfopenia e que apresentava deficiência severa das células T mas com um funcionamento normal de células B (GIBLETT ET AL 1975 *apud* DOS SANTOS 2004).

Têm-se sabido por muito tempo que a diferença metabólica mais impressionante entre parasitas e humanos, é a via de recuperação do DNA. Em contraste com as células dos mamíferos, que podem sintetizar anéis de purina de novo, todos os protozoários parasitas são classificados como auxotróficos para purinas, e podem desencadear um crescimento médio definido somente na adição de fontes de purinas (BZOWSKA et al, 2000). Além disso, cada gênero de parasita evoluiu sua própria disposição original de enzimas de recuperação de purinas, permitindo ao organismo remover purinas do hospedeiro. Este é o constituinte base para a susceptibilidade de '*Leishmania*' e '*Trypanosoma*' para análogos de purinas naturais e seus nucleosídeos, como o allopurinol (4-hidroxipirazolo[3,4-d]pirimidina), uma droga que não é tóxica para humanos e é usada para o tratamento de hiperuricemia e gota (BZOWSKA et al, 2000).

A inibição específica da Purina Nucleosídeo Fosforilase do *Mycobacterium tuberculosis* (MtPNP), poderia potencialmente conduzir a um acúmulo de nucleotídeos

de guanina. O aumento da metade da concentração de guanosina hiperfosforilada é a característica central de uma resposta pleiotrópica fisiológica chamada de resposta estrita, na qual a bactéria entra em um estado latente em resposta ao estresse nutricional, ou seja, o acúmulo de guanosina hiperfosforilada implica na latência do *Mycobacterium tuberculosis*. Foi mostrado também que a concentração de guanosina hiperfosforilada pode controlar a adaptação micobacterial para as condições de crescimento limite permitido para o tempo de sobrevivência (BASSO et al 2001).

Os estudos estruturais de enzimas, tais como a PNP, podem ser a chave para ampliar nosso conhecimento para o desenvolvimento de futuros fármacos mais seletivos e eficazes. Estes estudos estruturais potencialmente poderiam nos levar a encontrar diferenças estruturais capazes de proporcionar o desenho de inibidores seletivos que não atuariam na enzima humana.

1.3 A Purina Nucleosídeo Fosforilase – Uma enzima chave no metabolismo de purinas

A PNP tem um papel central no metabolismo de purinas, operando normalmente na via de recuperação nas células (DE AZEVEDO et al, 2003). A PNP catalisa a clivagem da ligação N-glicosídica de ribo e desoxirribonucleosídeos, na presença de ortofosfato inorgânico (Pi) como um segundo substrato, para gerar a base de purina e ribose (desoxirribose)-1-fosfato (Figura 1.2) (CANDURI et al, 2004).

O desenvolvimento do conhecimento acerca da Purina Nucleosídeo Fosforilase exige esforços principalmente no que diz respeito a estudos estruturais. Muitos grupos em todo o mundo têm se dedicado a estes estudos com destaque à PNP humana (NOLASCO, 2005; BZOWSKA et al, 2000; DE AZEVEDO et al, 2003).

Diversas estruturas cristalográficas da PNP humana (HsPNP) complexadas com substratos, análogo de substratos, ou inibidores, abriram oportunidade para o conhecimento da base estrutural da afinidade existente entre a HsPNP e diferentes ligantes (CANDURI et al, 2005).

A primeira estrutura cristalográfica da PNP humana foi determinada em 1990 a 3,2Å de resolução. Estudos cristalográficos posteriores melhoraram a resolução para 2,8Å, e estas coordenadas foram então utilizadas no embasamento estrutural de desenho de inibidores de PNP (DE AZEVEDO et al, 2003). Em 2003, foi obtida, através da difração de raios X de altíssima resolução, a estrutura da PNP recombinante a 2,3Å usando radiação síncrotron (DE AZEVEDO et al, 2003). Essa estrutura permitiu a redefinição dos resíduos envolvidos na ligação de substratos e forneceu um modelo de maior confiança para embasamentos estruturais para o desenho de inibidores, com os dados obtidos com essa estrutura, foi possível também se propor um novo mecanismo catalítico para a PNP (CANDURI et al, 2005).

1.4 PNPs de diferentes origens e suas classificações

Muitas PNPs dos mais diversos organismos tem sido identificadas e purificadas com diferentes especificidades, e diversas tentativas de classificação tem sido feita desde então. Neste trabalho mereceram maior destaque duas classificações: a classificação proposta por Bzowska et al (2000) e a classificação proposta por Pugmire & Ealick (2002).

Bzowska et al (2000) fez uso de uma classificação com base na massa molecular (M_r), composição da unidade biológica e especificidade dos substratos, separando assim em três grandes categorias as PNPs. (BZOWSKA et al, 2000; DOS SANTOS, 2004).

A classificação de Bzowska et al (2000) faz distinção entre as PNPs de baixa massa molecular homotriméricas e as de alta massa molecular homo-hexaméricas, outras PNPs que não se enquadram em nenhum desses dois grupos são classificadas como “outras”. O grupo “outras”, por sua vez, se divide em dois subgrupos de acordo com sua especificidade: o primeiro subgrupo composto de PNPs que não são homotrímeros mas tem especificidade de homotrímeros. Fazem parte do segundo subgrupo, as PNPs que não são homo-hexâmeros mas são possuidoras da mesma especificidade dos homo-hexâmeros. Esta classificação é apresentada na Tabela1.1

Tabela1. 1 Classificação de PNPs de diferentes origens proposta por BZOWSKA et al (2000).

	Alta Massa Molecular	Baixa Massa Molecular	Outras	
			Grupo1	Grupo2
Massa Molecular	110 – 160kDa	80 – 100kDa	60 – 180kDa	
Estrutura Quaternária	Homo-hexâmeros	Homotrímeros	Variada, exceto homotrímeros.	Variada, exceto homo-hexâmeros
Especificidade (substratos naturais)	Inosina, Guanosina e Adenosina	Inosina e Guanosina	Inosina e Guanosina	Inosina, Guanosina e Adenosina
Exemplos	E. Coli, Salmonella	Homo Sapiens		

No que se refere a MtPNP, não se pode afirmar com segurança sua classificação, uma vez que há escassez de estudos sobre a sua especificidade e cinética enzimática para o seu enquadramento em um dos grupos anteriormente descritos. Estudos como o do presente trabalho (o qual trata da estrutura da MtPNP complexada com adenina e adenosina) podem fornecer base para um maior entendimento e possível auxílio na sua classificação.

Outra classificação dessas enzimas foi proposta por Pugmire & Ealick (2002), que, baseados em estudos estruturais, verificaram que todas as enzimas que catalisam a quebra fosforolítica da ligação glicosídica em nucleosídeos possuem uma entre duas formas distintas de enovelamento. A descoberta de somente duas formas possíveis de enovelamento fornece a base para a definição de duas famílias distintas de nucleosídeo fosforilases.

A primeira família, Nucleosídeo Fosforilase I (NPI), apresenta enzimas com estruturas quaternárias triméricas ou hexaméricas e compartilham enovelamento comum da subunidade em estruturas α/β . A segunda família, Nucleosídeo Fosforilase II (NPPII), foi definida como sendo a família das enzimas que apresentam a estrutura quaternária dimérica onde cada subunidade característica consiste de um pequeno α -domínio separado por uma fenda de um domínio α/β maior (Figura 1.4).

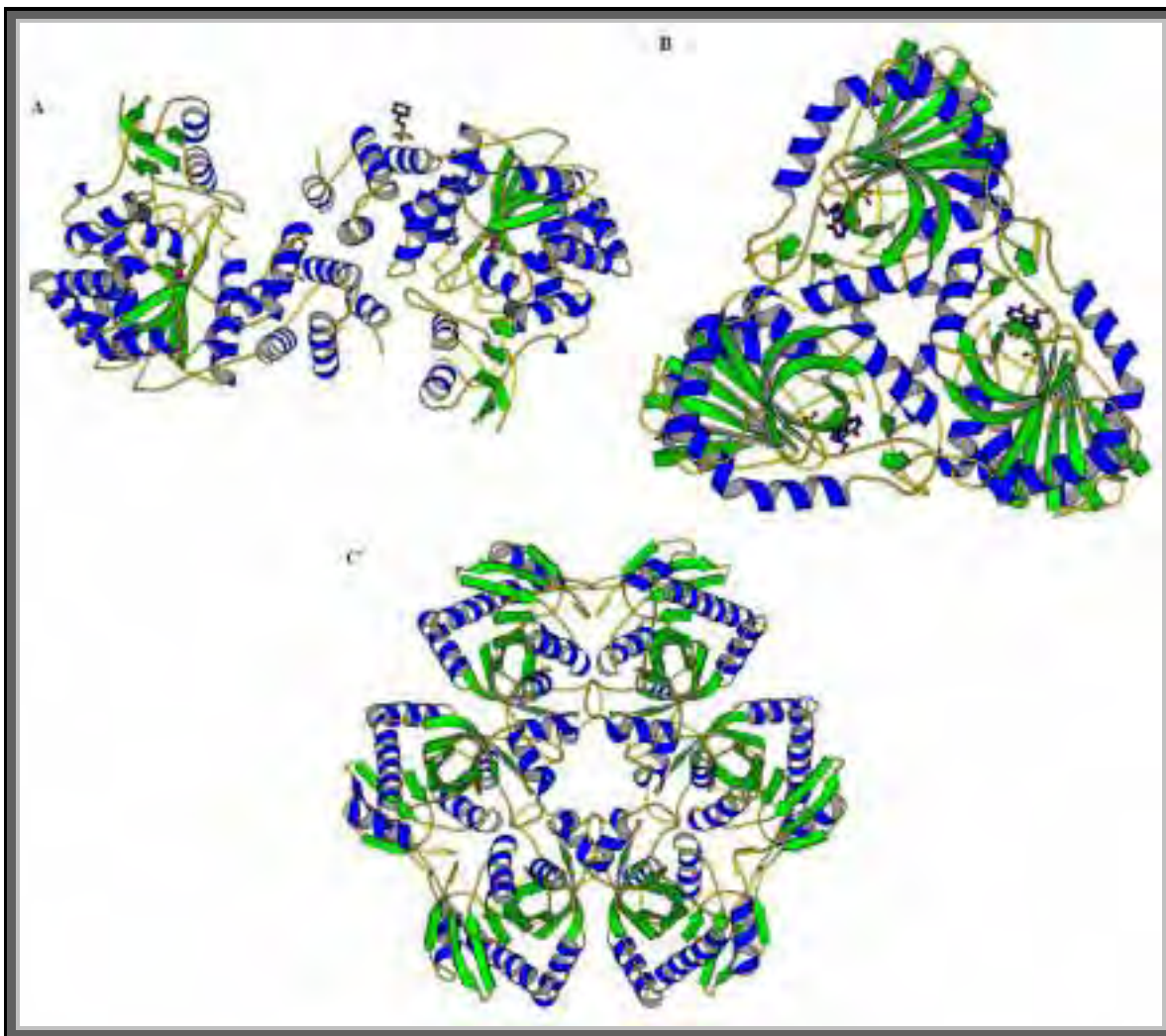


Figura 1.4 – Possíveis enovelamentos da Purina Nucleosídeo Fosforilase. A estrutura dimérica (A) é uma Pirimidina Nucleosídeo Fosforilase do *Bacillus stearothermophilus*. A estrutura trimérica (B) é uma Purina Nucleosídeo Fosforilase do *Mycobacterium bovis*. A estrutura hexamérica (C) é uma Purina Nucleosídeo Fosforilase da *Escherichia coli*.

1.5 Análise do Mecanismo Catalítico

As análises das estruturas da HsPNP demonstraram que o sítio ativo está localizado em uma cavidade profunda na proteína, na interface entre duas subunidades. O subsítio da purina localiza-se mais internamente (mais profundo no bolsão de ligação), enquanto o subsítio da ribose está mais próxima do solvente (ver figura 1.5). Erion et al (1997) propuseram a ordem de ligação dos substratos e a liberação dos produtos no qual o subsítio da purina parece ser bloqueado quando ocorre o complexo PNP:ribose-1-fosfato, que resultaria num complexo “dead-end” (sem saída). No entanto, a ligação

inicial da purina pode resultar em uma mudança conformacional que promova também interações adicionais entre a PNP e a ribose-1-fosfato (DOS SANTOS, 2004; NOLASCO, 2005; ERION et al,1997).

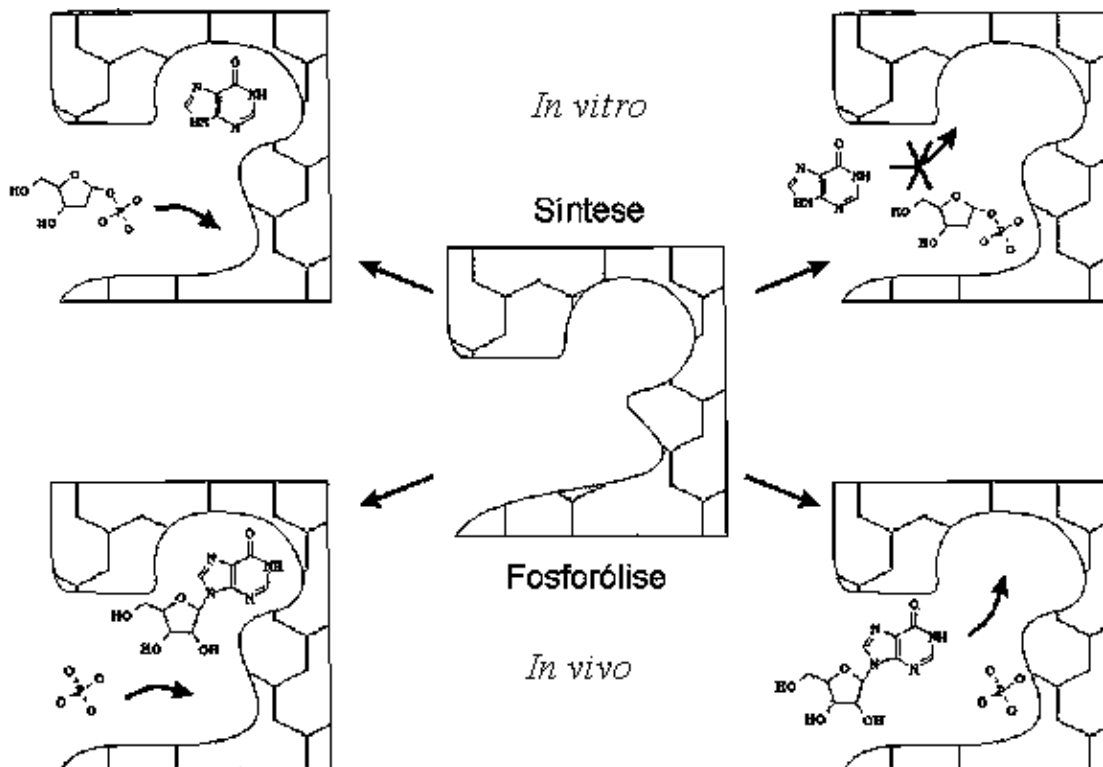


Figura 1.5– Possíveis ordens de ligação dos substratos, nas direções fosforolítica e sintética (modificado de ERION *et al.*, 1997).

Até pouco tempo atrás, o mecanismo catalítico do trímico da PNP humana ainda não havia sido corretamente elucidado. Haviam três mecanismos reportados com base em dados cristalográficos de baixa resolução, contudo todos estes modelos apresentavam várias discrepâncias com as informações funcionais. Canduri et al, (2005) propôs um novo mecanismo catalítico para PNP que explica os novos dados estruturais e análises funcionais da PNP humana. O mecanismo de Canduri et al. só foi proposto graças aos estudos estruturais cristalográficos de altíssima resolução de Azevedo et al, (2003), os quais forneceram um modelo para HsPNP a 2.3Å de resolução.

O mecanismo de Canduri et al. (2005), é ilustrado na figura 1.6, neste modelo, a His86 aceita um próton do fosfato, de modo que o último se torne um diânion (figura

1.6 A). O fosfato diânion ataca o C(1) da pentose que enfraquece a ligação glicosídica, causando a troca de elétrons entre o anel de purina, resultando em uma purina intermediária negativamente carregada (figura 1.6 B e C), a carga negativa localizada na purina O(6) faz a força eletrostática de interação com Asn243 (ND2), que arranca elétrons da ligação N-glicosídica. Esta ponte de hidrogênio intermolecular pode facilitar a redistribuição de carga negativa da ligação glicosídica enfraquecida e seu acúmulo na O(6) da base.

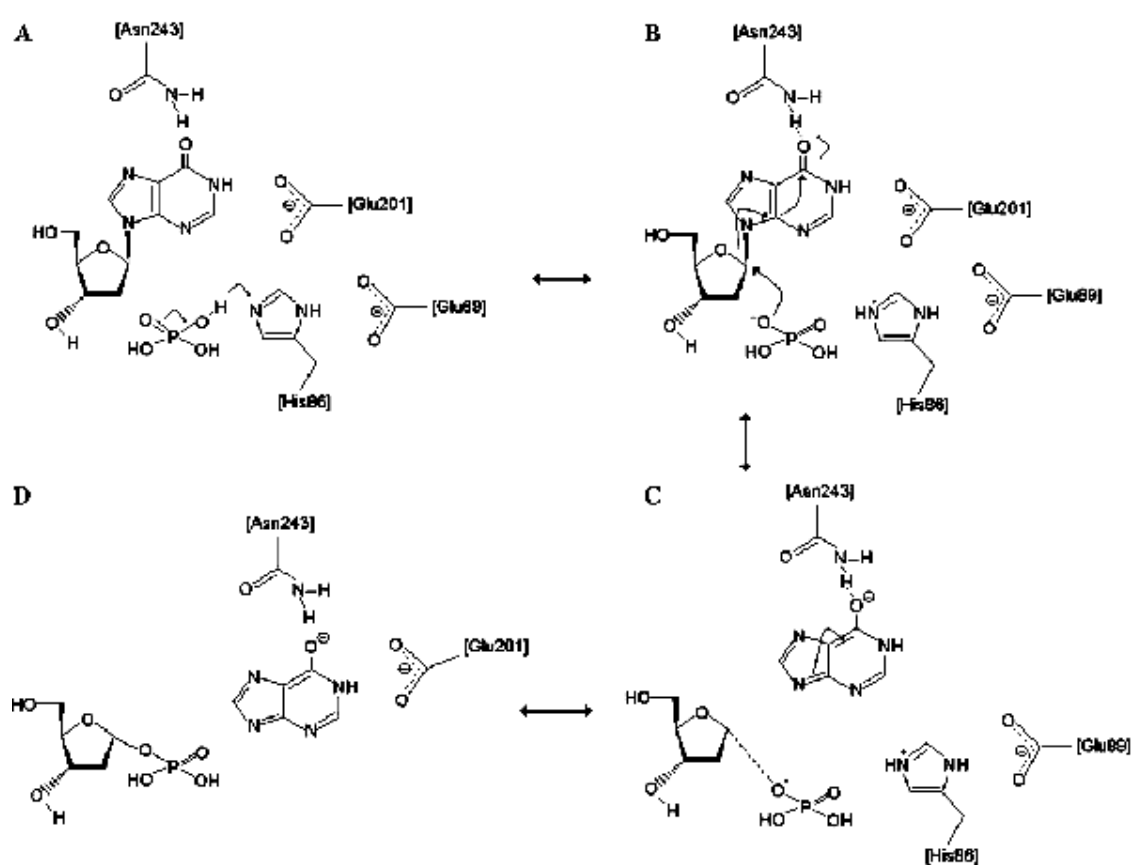


Figura 1.6 Novo mecanismo catalítico para a HsPNP proposto por Canduri et al., (2005).

1.6 A PNP do *Mycobacterium tuberculosis*

Uma proteína de sequência homóloga à da PNP humana foi identificada no genoma do *Mycobacterium tuberculosis*. O genoma do *Mycobacterium tuberculosis* é composto por 4 411 529 pares de bases e contém 3924 genes. Baseado em homologia de seqüências, funções bioquímicas foram atribuídas à aproximadamente 40% das proteínas previstas, enquanto similaridades com outras proteínas descritas foram encontradas para outros 44% e os 16% restantes não mostraram semelhança com nenhuma proteína conhecida e podem ainda guardar proteínas com funções micobacteriais específicas (NOLASCO 2005; COLE et al., 1998).

São escassos os estudos encontrados sobre a PNP do *Mycobacterium tuberculosis*, e, dentre os poucos trabalhos existentes, não se pode deixar de citar Basso et al 2001 e Nolasco et al 2004. Em 2001 em trabalho publicado no periódico Biochemistry, Basso et al. analisou a base estrutural para a inibição da MtPNP por imucilina-H e por suas partes componentes. Nolasco et al. resolveu a estrutura cristalográfica da MtPNP na forma de apoenzima e complexada com acyclovir (NOLASCO et al 2004).

A MtPNP apresenta, assim como as PNPs de mamíferos, uma estrutura simetricamente homotrimérica de arranjo triangular das subunidades. Cada monômero da proteína está enovelado numa topologia do tipo α/β , consistindo de onze fitas β circundadas por oito hélices α , e os três sítios catalíticos independentes se encontram nas interfaces das subunidades, (Figura 1.7), (NOLASCO, 2005).

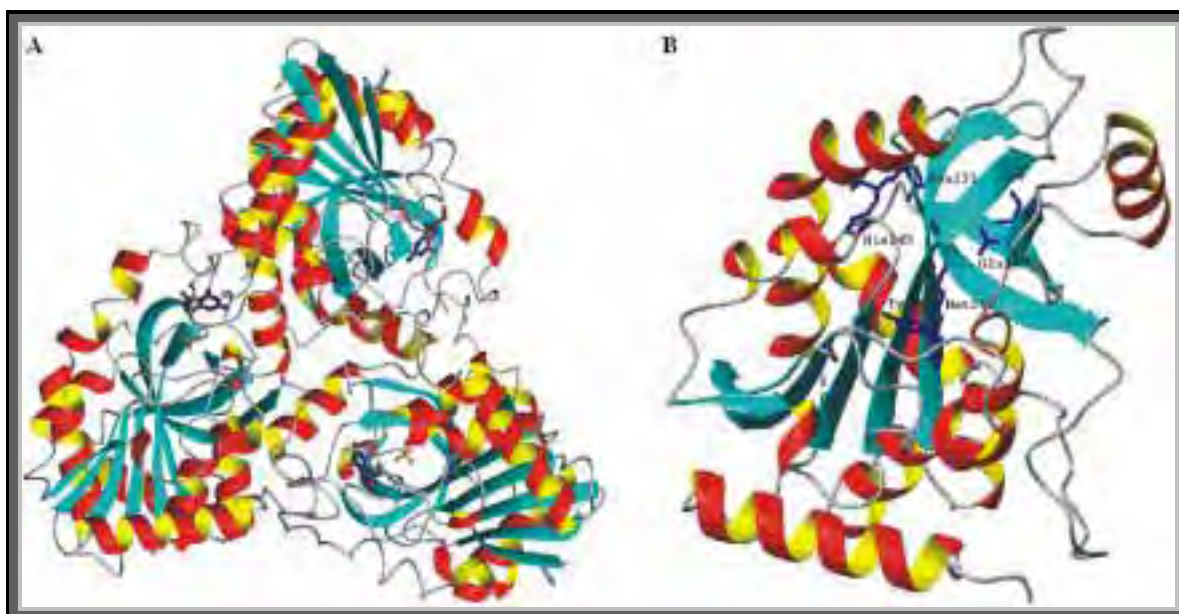


Figura 1.7 – Diagrama do trîmero (A) e do monômero (B) da PNP do *Mycobacterium tuberculosis* (NOLASCO 2005).

A estrutura primária da MtPNP tem 34% de identidade com a estrutura primária da HsPNP (Figura 1.8), porém, as estruturas secundária, terciária e quaternária são extremamente próximas (NOLASCO *et al.*, 2004). (Fig. 1.8)

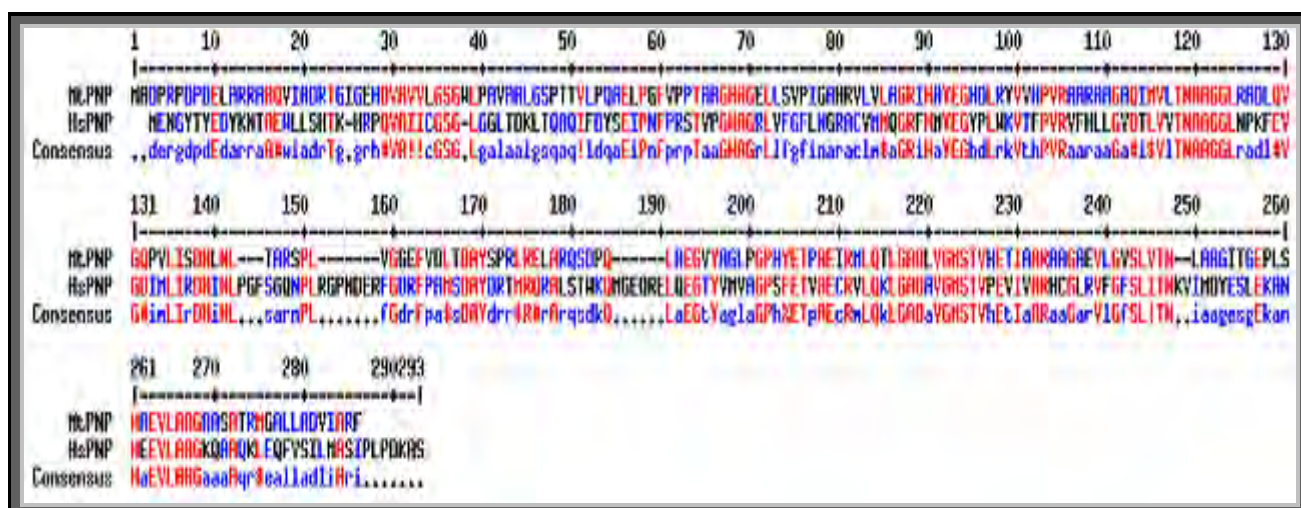


Figura 1.8 – Alinhamento das estruturas primárias da MtPNP e da HsPNP feito utilizando o programa MultAlin (CORPET, 1988). A identidade, mostrada em vermelho, é de, aproximadamente, 34%.

Mesmo sendo de diferentes origens, ambas PNPs apresentam a característica estrutural trimérica das PNPs de mamíferos. Os sítios ativos da MtPNP e da HsPNP (código de acesso no PDB: 1M73) são formados pelos mesmos resíduos e estes estão dispostos de forma idêntica (NOLASCO *et al.*, 2004) (Figura 1.9). Sobrepondo as

estruturas observou-se um RMSD de 2,5Å, em média, e, selecionando somente dos resíduos que formam o sítio catalítico do monômero A observou-se, ainda, um RMSD de 0,9Å (NOLASCO et al., 2004).

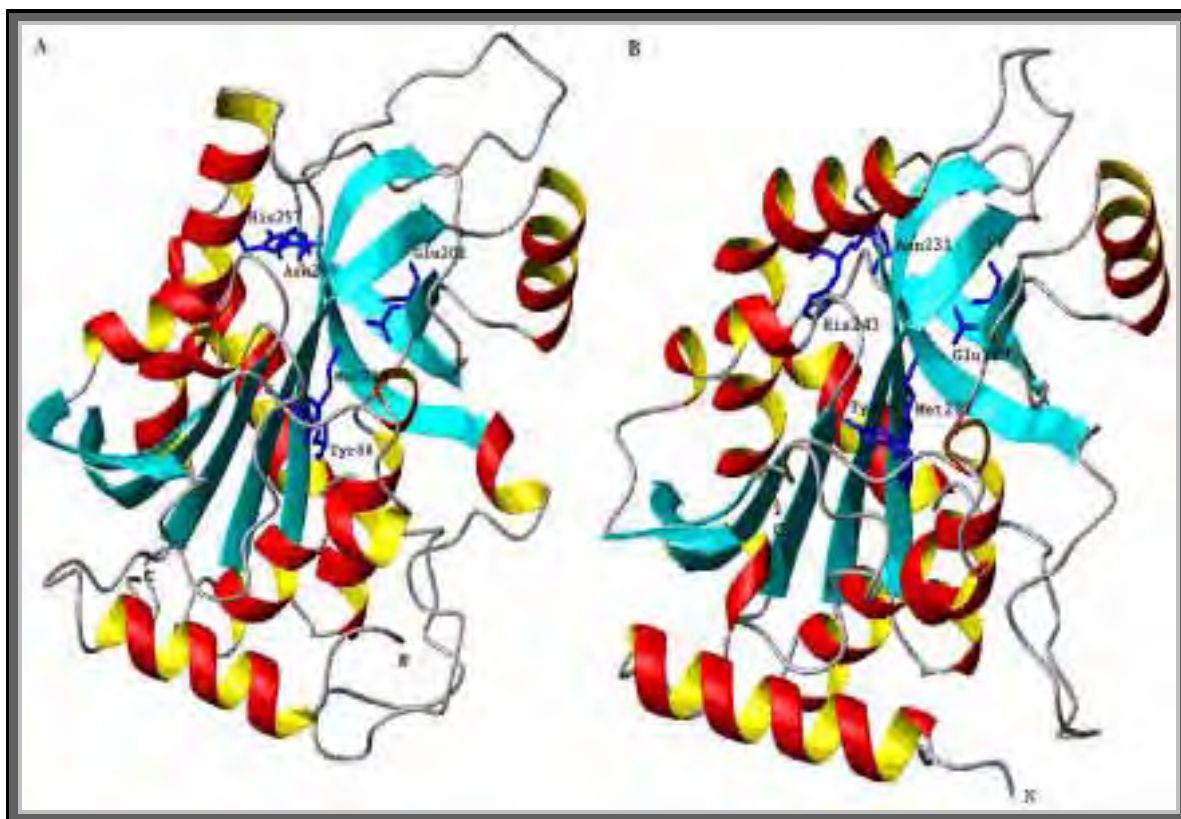


Figura 1.9 – Comparação dos monômeros da HsPNP (A) e da MtPNP (B) (NOLASCO *et al.*, 2004).

2. Objetivos

O presente trabalho tem como objetivo fazer um estudo estrutural, principalmente no que se refere ao sítio ativo da Purina Nucleosídeo Fosforilase do *Mycobacterium tuberculosis* complexado com adenosina (um possível substrato) e adenina. Uma vez que a PNP humana não cliva adenosina, a análise do sítio ativo da MtPNP complexada com adenina e adenosina pode fornecer base para o desenho de inibidores específicos para ela.

3. Materiais e Métodos

3.1 Obtenção da Proteína

A proteína foi gentilmente cedida pelo Prof. Dr. Diógenes Santiago Santos e pelo Prof. Dr. Luiz Augusto Basso, Departamento de Biologia Molecular e Biotecnologia (Pontifícia Universidade Católica do Rio Grande do Sul), Porto Alegre – Brasil, ligado à Rede Brasileira de Pesquisa em Tuberculose. A MtPNP foi obtida a partir de métodos de clonagem, expressão e purificação (BASSO *et al.*, 2001).

A massa molecular determinada por experimento de espectrometria de massa para o monômero da MtPNP foi de 27 437 Da, muito próximo da massa molecular prevista que era de 27 571 Da.

3.2 Cristalização da MtPNP e Estrutura Cristalina

A cristalização de macromoléculas

Cristalização é um processo de ordenação, no qual as moléculas assumem posições regulares no estado sólido. Um processo de cristalização pode ser considerado como um equilíbrio dinâmico entre as partículas no estado líquido e no estado sólido. Sendo assim, o cristal cresce quando o equilíbrio é deslocado no sentido de solidificação (DE AZEVEDO 1997).

O uso de cristais na biologia estrutural veio depois de 1934, quando Bernal e Crowfoot produziram o primeiro padrão de difração de raios X de uma proteína, o da pepsina. Desde então, representantes de muitas famílias de macromoléculas tem sido cristalizadas (DUCRUIX & GIEGÉ, 1992; DOS SANTOS, 2004).

Vinte anos se passaram do primeiro padrão de difração produzido até que Perutz, juntamente com seus colaboradores em Cambridge, foram capazes de mostrar que o

método de substituição isomórfica poderia ser usado para resolver o problema da fase em estruturas de proteínas. Em 1960 este método foi usado com sucesso para a solução da primeira estrutura cristalográfica de proteína, a estrutura da mioglobina, que foi seguida pela resolução da estrutura de uma enzima, a lisozima, em 1965 (BLUNDELL & JOHNSON, 1976).

O principal objetivo da cristalização de uma macromolécula biológica, tal como proteínas, ácidos nucleicos ou vírus, é o seu estudo estrutural por métodos de difração de raios X. Este estudo é fundamental para o entendimento molecular de vários mecanismos biológicos. E a cristalização é uma das partes mais críticas na resolução da estrutura de uma macromolécula biológica. Os principais passos envolvidos na resolução da estrutura de uma macromolécula biológica, a nível atômico são mostrados na figura 3.1 (CANDURI, 2001).

Entre todos estes passos, a cristalização é o menos entendido. Biocristalização, como qualquer outra cristalização, é um processo de múltiplos parâmetros que envolve os três passos clássicos da cristalização: nucleação, crescimento e término de crescimento. O que faz da biocristalização de macromoléculas biológicas diferente é o número de parâmetros envolvidos, muito maior do que aqueles envolvidos no crescimento de cristais de pequenas moléculas, além das propriedades físico-químicas peculiares destes compostos. Por exemplo, a estabilidade otimizada das macromoléculas biológicas em um meio aquoso está restrita a uma estreita faixa de pH, sendo que a principal diferença da cristalização de pequenas moléculas e de macromoléculas biológicas é a flexibilidade conformacional e a versatilidade química dessas macromoléculas e, conseqüentemente, sua maior sensibilidade às condições externas (CANDURI, 2001).

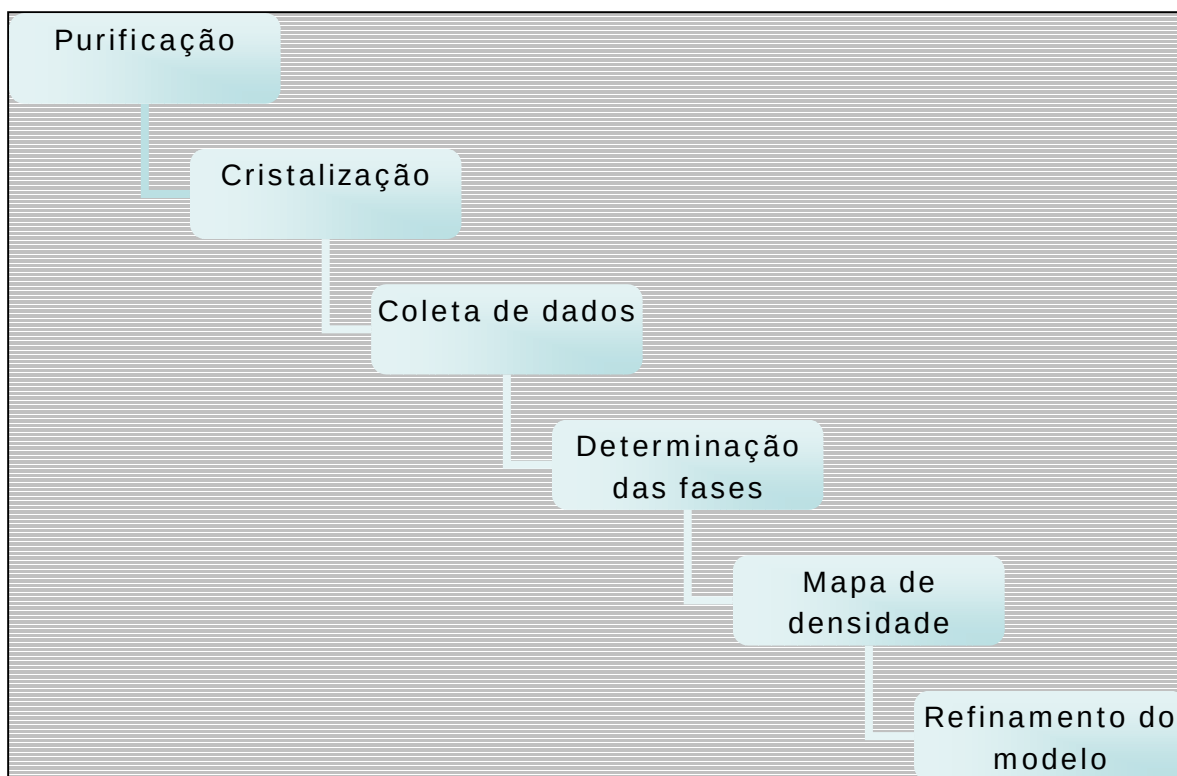


Figura 3.1 Principais passos envolvidos na resolução da estrutura de uma macromolécula biológica (CANDURI, 2001).

Devido ao fato de que proteínas e ácidos nucleicos necessitam de pH e força iônica definidos para a sua estabilidade e função, cristais de macromoléculas devem ser crescidos a partir de soluções aquosas complexas. A cristalização começa com a nucleação e as condições para se atingir a nucleação são, algumas vezes, difíceis de repetir e, desta forma, os procedimentos de semeadura com um material cristalino pré-formado se fazem necessários. Em alguns momentos este é o único método para se obter resultados reprodutíveis. A nucleação requer um estado de supersaturação maior do que aquele da fase de crescimento (DE AZEVEDO, 1997).

O término do crescimento de um cristal de macromolécula pode ter diversas causas, desconsiderando as triviais como a remoção da macromolécula do meio de cristalização, por exemplo. Entre essas causas temos os defeitos de crescimento, o envenenamento das faces e o envelhecimento da macromolécula (DE AZEVEDO, 1997).

Para a determinação da estrutura tridimensional de uma macromolécula utilizando técnicas de difração de raios X, é imprescindível se obter monocristais de tamanho apropriado para a coleta de dados de difração de raios X. Como o número de variáveis envolvidas no processo de cristalização é muito grande (como concentração de sal, temperatura, pH, agentes precipitantes e aditivos) encontrar as condições ideais pode levar dias, meses, anos ou pode ainda nunca acontecer (BLUNDELL & JOHNSON, 1976; CANDURI, 2001, NOLASCO, 2005).

Para cristalizar uma macromolécula em um determinado solvente, é necessário leva-la lentamente a um estado de supersaturação, que é termodinamicamente instável, no qual pode-se chegar a uma fase amorfa ou cristalina, quando este retornar ao equilíbrio. As fases deste processo estão ilustradas na Figura 3.2.

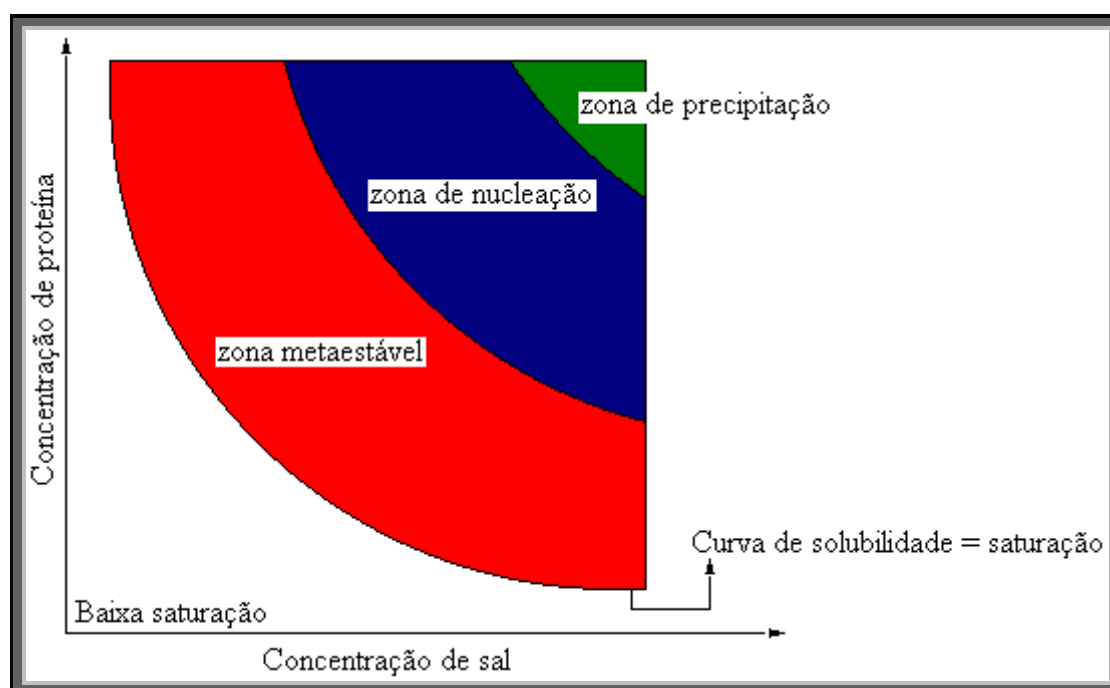


Figura 3.2 – Representação esquemática de um diagrama de solubilidade, ilustrando as regiões de variação da solubilidade de uma proteína quando a concentração do agente cristalizante é alterada (NOLASCO, 2005).

A curva de solubilidade divide as regiões de baixa saturação e de supersaturação. Na região de baixa saturação a macromolécula não será cristalizada. Logo acima se encontra a região de supersaturação, que é dividida em três zonas: zona metaestável,

zona de nucleação e zona de precipitação. Quando a supersaturação é alcançada rapidamente, a proteína em solução precipita, no entanto, quando a supersaturação é alcançada lentamente, a solução de proteína alcança a zona de nucleação, onde se formam os núcleos (primeiros agregados ordenados de proteína). Conforme a solubilidade da proteína diminui, os núcleos atingem a zona metaestável, onde ocorrerá o crescimento do cristal (GIEGÉ & DUCRUIX, 1992; NOLASCO 2005) (Figura 3.2).

Técnicas de Cristalização de macromoléculas biológicas

São utilizados quatro métodos para a cristalização de macromoléculas biológicas, (segundo GIEGÉ & DUCRUIX, 1992; CANDURI 2001): diálise, difusão em interface livre, batch e difusão de vapor.

O princípio da diálise se baseia no equilíbrio de pequenos solutos via membrana semi-permeável. A difusão de interface livre consiste em dois solventes de densidades diferentes nas quais a macromolécula não é igualmente solúvel. O batch consiste na preparação de uma solução de proteína de forma que se cristalice espontaneamente no tempo, pode-se aplicar um gradiente de temperatura, por exemplo, se a proteína for mais solúvel a temperaturas mais altas, pode-se aquecer a solução e permitir seu resfriamento lento, isso pode levar a formação de cristais.

A difusão de vapor tem por princípio o processo de equilíbrio entre duas soluções através da fase de vapor num meio fechado. A solução menos concentrada poderá perder seu solvente volátil até que potenciais químicos das duas soluções sejam iguais.

A técnica de difusão de vapor possui três variações, sendo elas: a gota suspensa (*hanging drop*), gota sentada (*sitting drop*) e a gota sanduíche (*sandwich drop*)

(Figura 3.3). A técnica da gota pendurada (*hanging drop*) é frequentemente mais utilizada, e consiste em “pendurar” uma gota contendo a macromolécula a ser cristalizada, diluída em tampão com o agente de cristalização e aditivos, em uma lamínula e equilibrá-la contra um reservatório contendo a solução do agente de cristalização a uma concentração mais alta do que na gota. O equilíbrio ocorre por meio da perda de água (da gota para o reservatório), isto leva a uma diminuição do volume da gota e, conseqüentemente, a um aumento na concentração de todos os constituintes da gota de cristalização (DUCRUIX & GIEGÉ, 1992).

Para a concentração final de proteína ser controlada, se realiza um experimento de difusão de vapor com um volume de solução precipitante – o reservatório – infinitamente grande quando comparado com o pequeno volume da solução da proteína. A gota contendo a macromolécula biológica a ser cristalizada, com o tampão, agentes precipitantes e aditivos, é equilibrada contra um reservatório contendo uma solução do agente precipitante a uma concentração mais alta que a gota (CANDURI, 2001).

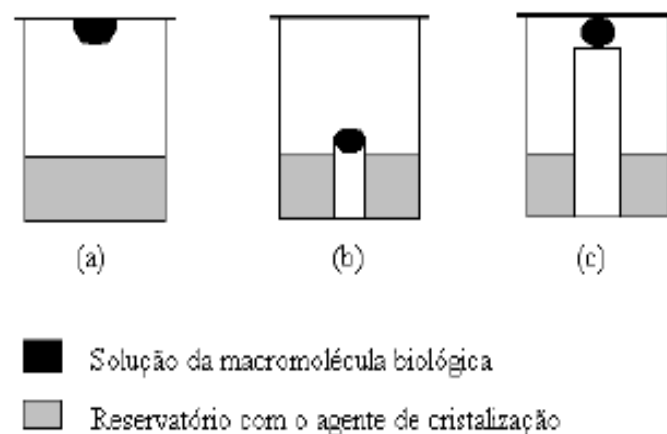


Figura 3.3 - Representação esquemática de (a) gota suspensa (*hanging drop*), (b) gota sentada (*sitting drop*) e (c) gota sanduíche (*sandwich drop*).

Se o volume da solução do reservatório é suficientemente grande comparado com o da gota, podemos assumir que quando o equilíbrio é atingido a concentração do precipitante na solução de macromolécula será igual à concentração inicial na solução do reservatório.

A cristalização da MtPNP

A MtPNP foi cristalizada segundo condições previamente estabelecidas (SHI *et al.*, 2001; Nolasco 2004, 2005). A MtPNP (1µL à uma concentração de 25mg.ml⁻¹) contendo 5 mM NaH₂PO₄ foi misturada com igual volume da solução do poço que continha 100 mM de Tris, pH 8,0, 25% de PEG 3350 e 25 mM de MgCl₂ e foi equilibrada contra 1,0 mL da solução do poço, a técnica utilizada foi a de difusão de vapor utilizando a gota suspensa.

A adição dos ligantes Adenosina e Adenina

A adição dos ligantes na enzima ocorreu através de duas técnicas distintas, para o substrato Adenosina foi utilizada a técnica de **co-cristalização**, complexada em 5 moléculas de ligante para uma molécula de proteína (5:1), na qual o cristal é feito com o ligante já complexado com a enzima. Para o produto Adenina foi utilizada a técnica de **soacking** que consiste na adição do ligante na solução que contém o cristal da apoenzima (enzima sem o ligante). No presente trabalho, o soacking foi feito adicionando diretamente o ligante na gota que continha os cristais, sendo que o ligante adicionado estava em pó, sem nenhum tipo de diluição em solventes. Isto foi necessário porque a adenina é insolúvel em todos os possíveis solventes testados. A técnica do **soacking** é limitada quando o tamanho da molécula de solvente é maior do que os canais formados dentro da rede do cristal da macromolécula (DUCRUIX & GIEGÉ, 1992). A estrutura dos dois ligantes aparece na figura 3.4.

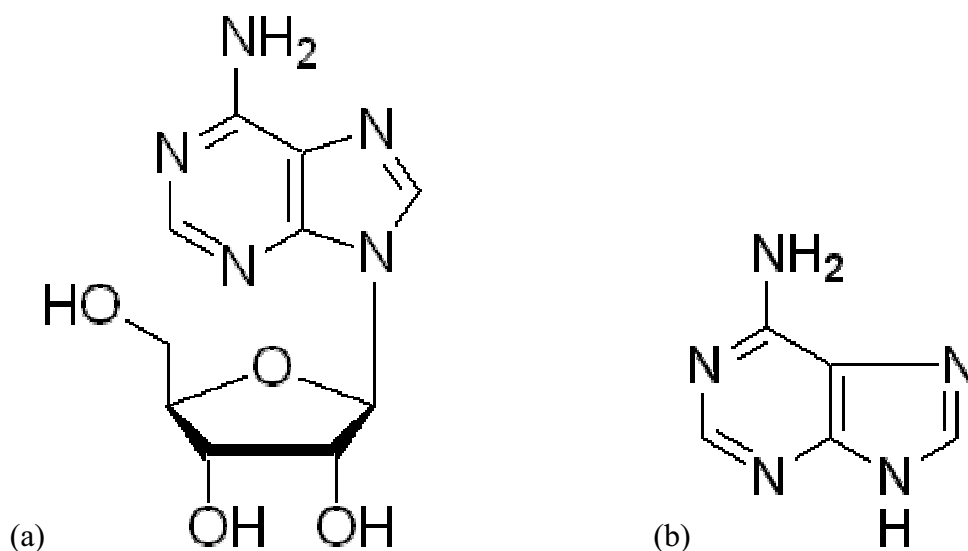


Figura 3.4 Ligantes utilizados. (a) adenosina e (b) adenina.

Estrutura Cristalina (KITTEL, 1978)

Um cristal é um arranjo periódico de átomos. Um cristal ideal é construído pela repetição infinita de uma mesma estrutura elementar. Nos cristais mais simples, a unidade estrutural é representada por um único átomo. A estrutura de todos os cristais pode ser descrita em termos de uma rede com um grupo de átomos ligados a cada ponto da rede. Este grupo é denominado **base**; ele se repete no espaço para formar a estrutura cristalina.

Um cristal ideal é composto por uma arrumação de átomos numa rede definida por três vetores fundamentais **a**, **b** e **c**, de modo que as configurações atômicas sejam exatamente iguais tanto para um observador situado em **r** quanto para um observador situado em **r'** dado por:

$$\mathbf{r}' = \mathbf{r} + u\mathbf{a} + v\mathbf{b} + w\mathbf{c} \quad (1)$$

onde u , v , w são números inteiros arbitrários. O conjunto de pontos **r'** especificados por (1) para todos os valores inteiros u, v, w define uma rede. O paralelepípedo formado pelos eixos primitivos **a**, **b** e **c** denomina-se célula primitiva (ou cela unitária), uma

célula primitiva é um tipo de (cela e não célula unitária), sendo que o volume da cela definida pelos eixos **a**, **b** e **c** é:

$$V_c = a \times b \times c \quad (2)$$

As redes cristalinas podem ser transformadas nelas mesmas por meio da rede **T** e por várias outras operações de simetria. Uma operação de simetria típica é a de rotação em torno de um eixo que passa por um ponto da rede. Existem redes que com eixos de rotação unitários, binários, ternários, quaternários e senários, que, por meio de rotações de 2π , $2\pi/2$, $2\pi/3$, $2\pi/4$ e $2\pi/6$ radianos ou múltiplos inteiros dessas rotações, transformam a rede cristalina nela mesma. Essas rotações são designadas pelos símbolos 1, 2, 3, 4 e 6. Não existem redes que se transformem nelas próprias quando submetidas a rotações de $2\pi/7$ radianos ou $2\pi/5$ radianos. O **grupo pontual da rede** é a coleção de operações de simetria que, quando aplicadas a um ponto da rede, deixam-na invariante.

Em três dimensões, os grupos de simetria pontuais estão associados com 14 tipos diferentes de redes indicadas na figura 3.6 e listados na tabela 3.1. O tipo geral de rede é o de rede triclinica. Os 14 tipos de redes estão convenientemente agrupados em sete sistemas cristalinos de acordo com os sete tipos convencionais de células unitárias: triclinico, monoclinico, ortorrômbico, tetragonal, cúbico, trigonal e hexagonal. Esses sistemas cristalinos são divididos em termos das relações axiais (entre os eixos) especiais para as células unitárias convencionais, esses eixos (**a**, **b** e **c**) e o ângulo entre eles (α , β e γ) estão definidos na figura 3.5. Os módulos a , b e c denominam-se **parâmetros de rede**.

Tabela 3.1 Os tipos de redes em três dimensões

Sistema	Número de Redes	Símbolos da Rede	Restrições para a célula convencional, para os eixos e para os ângulos.
Triclínico	1	P	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$
Monoclínico	2	P, C	$a \neq b \neq c$ $\alpha = \gamma = 90^\circ \neq \beta$
Ortorrômbico	4	P, C, I, F	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$
Tetragonal	2	P, I	$a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$
Cúbico	3	P ou sc I ou bcc F ou fcc	$a = b = c$ $\alpha = \beta = \gamma = 90^\circ$
Trigonal	1	R	$a = b = c$ $\alpha = \beta = \gamma < 120^\circ, \neq 90^\circ$
Hexagonal	1	P	$a = b \neq c$ $\alpha = \beta = 90^\circ \neq \gamma = 120^\circ$

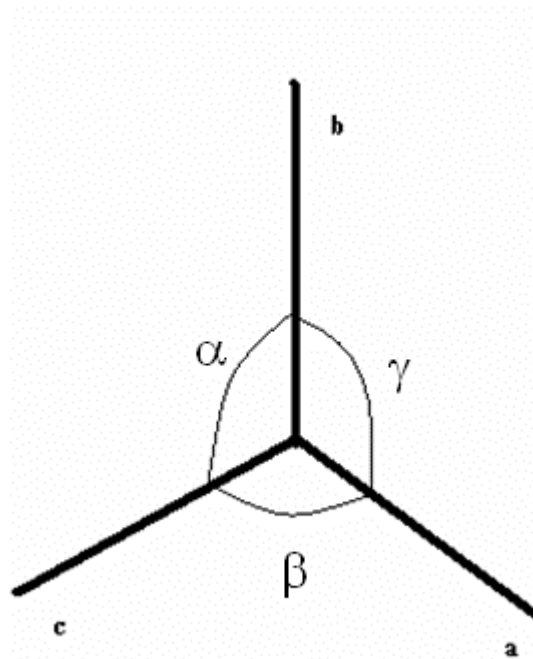


Figura 3.5 Eixos Cristalinos **a**, **b** e **c**. O ângulo α é o que está compreendido entre **b** e **c**.

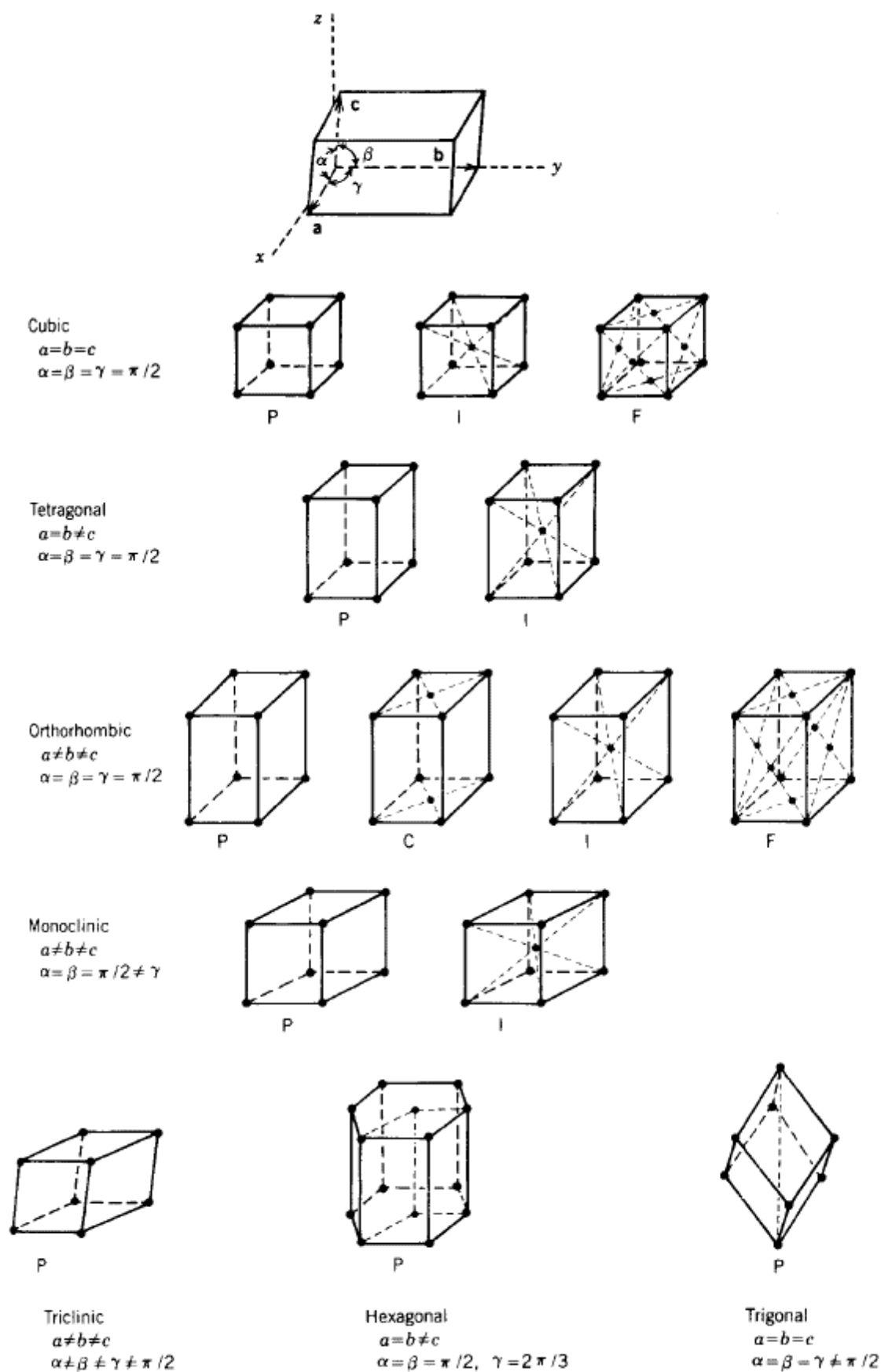


Figura 3.6. Rede de Bravais: As quatorze redes de bravais e os sete sistemas.

A posição e a orientação do plano de um cristal são determinadas por três pontos quaisquer pertencentes ao plano, desde que eles não sejam colineares. Se cada ponto estiver situado sobre um eixo cristalino, o plano pode ser especificado dando-se as posições dos pontos ao longo dos eixos em termos das constantes da rede o resultado é colocado entre parênteses (hkl).

Métodos de estudo de uma rede cristalina

Estuda-se a estrutura cristalina através da difração de fótons, nêutrons e elétrons. A difração depende da estrutura cristalina e do comprimento de onda da radiação, quando o comprimento de onda for comparável à constante da rede ou menos do que esta constante, podemos encontrar feixes difratados em direções bastante diferentes das direções originais (KITTEL, 1978).

A energia de um fóton de raios X é relacionada com seu comprimento de onda λ por $E = h\nu = hc/\lambda$, onde h é a constante de Plank, $h = 6,63 \cdot 10^{-27} \text{ erg.s} = 6,62 \cdot 10^{-34} \text{ J.s}$. Em unidades mais usadas em laboratório: $\lambda(\text{\AA}) = 12,4/E(\text{keV})$ (KITTEL, 1978).

Os raios X podem ser gerados tanto pela desaceleração dos elétrons em um alvo metálico quanto pela excitação dos elétrons dos átomos do alvo. Os núcleos, por causa de suas massas mais elevadas, efetivamente não espalham os raios X, sendo os elétrons responsáveis pelo fenômeno (KITTEL, 1978)

W. L. Bragg apresentou uma explicação para os feixes difratados por um cristal. Suponhamos que as ondas incidentes sejam refletidas especularmente (o ângulo de incidência é igual ao ângulo de reflexão) por planos de átomos paralelos no interior do cristal, sendo que cada plano reflete somente uma pequena fração da radiação. Os feixes difratados formam-se quando as reflexões provenientes de planos paralelos de átomos produzem interferência construtiva. Considerem-se somente plano paralelos da rede separados por uma distância d , como indicado na Figura 3.7, a diferença de caminho

para raios refletidos por planos adjacentes é $2d \sin\theta$, onde θ é medido a partir do plano (KITTEL, 1978).

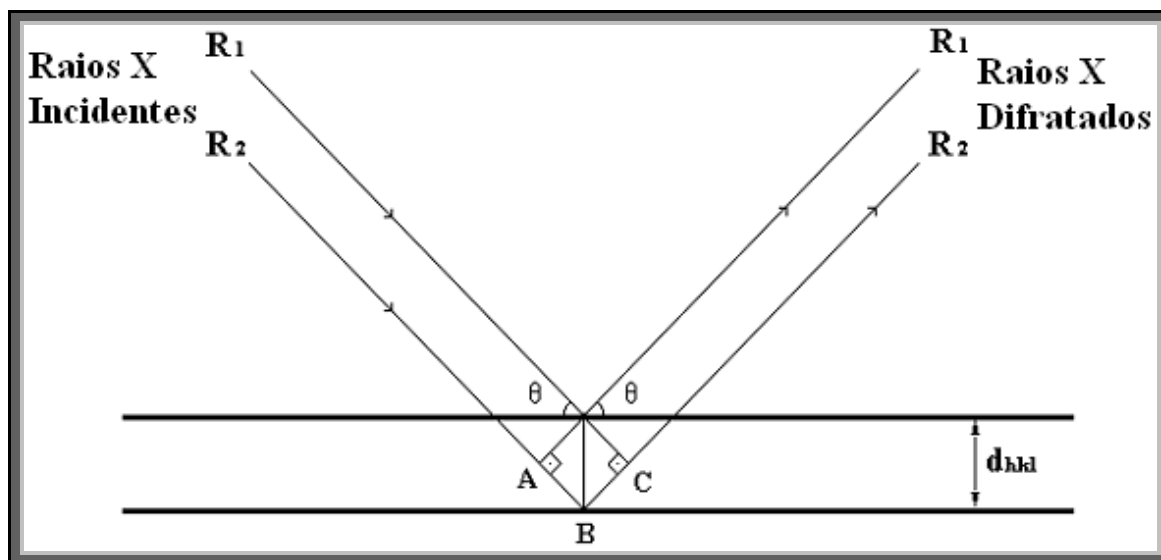


Figura 3.7 Representação esquemática da lei de Bragg (BLUNDELL & JOHNSON, 1976; NOLASCO, 2005)

A interferência construtiva da radiação proveniente de planos sucessivos ocorre quando a diferença de caminho for um número inteiro n de comprimentos de onda λ , de modo que:

$$2d \sin\theta = n\lambda \quad (3)$$

A reflexão de Bragg pode ocorrer apenas para comprimentos de onda $\lambda \leq 2d$, uma vez que a função seno varia entre zero e um. A dedução de Bragg da condição de difração fornece uma formulação clara da condição para a ocorrência da interferência construtiva das ondas espalhadas por cargas pontuais situadas nos pontos das redes. Entretanto, no que se refere a intensidade do espalhamento produzido pelos elétrons, é necessária uma análise mais profunda (Kittel, 1978; DE AZEVEDO 1997).

Qualquer propriedade física do cristal (concentração de cargas, a densidade numérica dos elétrons, a densidade de massa e a densidade dos momentos magnéticos) permanece invariante sob qualquer translação da forma $T = ua + vb + wc$, onde u, v e w são números inteiros e a, b e c são os eixos cristalinos, portanto a densidade numérica

dos elétrons $n(\mathbf{r})$ deve ser uma função periódica de $\mathbf{r}$ com períodos a , b , c nas direções dos eixos cristalinos, donde se conclui que $n(\mathbf{r} - \mathbf{T}) = n(\mathbf{r})$, (KITTEL, 1978).

Sendo a densidade numérica dos elétrons uma função periódica pode-se representá-la como uma série de Fourier, muitas propriedades do cristal podem ser relacionadas com os componentes de Fourier da densidade numérica dos elétrons (KITTEL, 1978; DE AZEVEDO 1997).

Fazendo a análise de Fourier para funções periódicas em três dimensões encontramos um conjunto de vetores $\mathbf{H}$ tais que

$$n(\mathbf{r}) = \sum_{\mathbf{H}} n_{\mathbf{H}} \exp^{i\mathbf{H} \cdot \mathbf{r}} \quad (4)$$

seja invariante em todas as translações $\mathbf{T}$ que deixam o cristal invariante. Os vetores $\mathbf{H}$ da série de Fourier são os vetores da rede recíproca, porque para eles, a série de Fourier para a densidade de elétrons possui a desejável invariância sob qualquer translação cristalina $\mathbf{T} = u\mathbf{a} + v\mathbf{b} + w\mathbf{c}$.

Cada estrutura cristalina possui duas redes: a rede cristalina e a rede recíproca. Uma figura de difração pode ser encarada como uma representação da rede recíproca do cristal, em contraste com a imagem microscópica, que é uma imagem da estrutura cristalina real. Quando giramos um cristal tanto a rede cristalina quanto a rede recíproca sofrem a rotação. Os vetores da rede cristalina possuem dimensão de [comprimento]; os vetores da rede recíproca possuem dimensão de [comprimento]⁻¹. A rede cristalina é o espaço real; a rede recíproca é uma rede no espaço de Fourier associado. As duas redes estão relacionadas através da construção dos vetores que definem os eixos da rede recíproca:

$$A = 2\pi \frac{b \times c}{a \bullet b \times c}; \quad B = 2\pi \frac{c \times a}{a \bullet b \times c}; \quad C = 2\pi \frac{a \times b}{a \bullet b \times c}. \quad (5)$$

Se $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ forem vetores primitivos da rede cristalina, então $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ serão os vetores primitivos da rede recíproca. Cada vetor é ortogonal a dois outros vetores dos eixos da rede cristalina. Portanto $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ possuem a propriedade:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{a} &= 2\pi & \mathbf{B} \cdot \mathbf{a} &= 0 & \mathbf{C} \cdot \mathbf{a} &= 0 \\ \mathbf{A} \cdot \mathbf{b} &= 0 & \mathbf{B} \cdot \mathbf{b} &= 2\pi & \mathbf{C} \cdot \mathbf{b} &= 0 \\ \mathbf{A} \cdot \mathbf{c} &= 0 & \mathbf{B} \cdot \mathbf{c} &= 0 & \mathbf{C} \cdot \mathbf{c} &= 2\pi \end{aligned} \quad (6)$$

Qualquer conjunto arbitrário de vetores primitivos $\mathbf{a}$, $\mathbf{b}$ e $\mathbf{c}$ de uma dada rede cristalina conduz ao mesmo conjunto de pontos da rede recíproca. Assim qualquer vetor $\mathbf{H}$ da seguinte forma: $\mathbf{H} = h\mathbf{A} + k\mathbf{B} + l\mathbf{C}$, ($h, k, l = \text{Inteiros}$), é denominado vetor da rede recíproca. O conjunto de vetores $\mathbf{H}$ da rede recíproca determina as possíveis reflexões dos raios X, sendo possível ter uma outra forma de expressão da lei de Bragg:

$$2\mathbf{k} \cdot \mathbf{H} + H^2 = 0 \quad (7)$$

este é o resultado central da teoria do espalhamento elástico numa rede periódica, resultado idêntico ocorrerá na teoria eletrônica sobre a estrutura das bandas de energia dos cristais (KITTEL, 1978).

Sabendo que o conjunto de vetores $\mathbf{H}$ da rede recíproca determina as possíveis reflexões dos raios X é possível introduzir os componentes de Fourier de $n(\mathbf{r})$ na equação da amplitude da onda espalhada (KITTEL, 1978, DE AZEVEDO 1997).

A amplitude da onda espalhada por um elemento de volume é proporcional à concentração local dos elétrons. Assim, (assumindo como $\mathbf{k}$ o vetor de onda do feixe incidente e como $\mathbf{k}'$ o vetor de onda do feixe de onda emergente), temos que a amplitude total de onda espalhada na direção $\mathbf{k}'$ é proporcional a integral sobre o volume do cristal do fator de fase $\exp [i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{r}]$ vezes $n(\mathbf{r})dV$ (densidade numérica dos elétrons multiplicada pelo volume). A amplitude da onda espalhada é determinada pela integral:

$$\mathbf{F} = \int dV \mathbf{n}(\mathbf{r}) \exp(-i \Delta \mathbf{k} \cdot \mathbf{r}) \quad (8)$$

Introduzindo os componentes de Fourier de $\mathbf{n}(\mathbf{r})$ temos:

$$\mathbf{F} = \sum \int dV \mathbf{n}_{\mathbf{H}} \exp[i(\mathbf{H} - \Delta \mathbf{k} \cdot \mathbf{r})] \quad (9)$$

Quando a condição $\Delta \mathbf{k} = \mathbf{H}$ for satisfeita, a amplitude de espalhamento para um cristal de N celas pode ser escrita como:

$$\mathbf{F}_{\mathbf{H}} = \int_{\text{cel}} dV \mathbf{n}(\mathbf{r}) \exp(-i \mathbf{H} \cdot \mathbf{r}) = N F(\mathbf{hkl}) \quad (10)$$

onde F é denominado **fator de estrutura**, sendo definida por uma integral sobre uma cela única, com $\mathbf{r} = 0$ em um dos vértices.

Sabendo que a onda total espalhada por todos os átomos é dada pela soma vetorial das contribuições de cada átomo, podemos reescrever o fator de estrutura, utilizando uma somatória das contribuições individuais de cada átomo, logo:

$$F(\mathbf{S}) = \sum_{j=1}^N f_j \exp 2\pi i (\mathbf{r}_j \cdot \mathbf{S}) \quad (11)$$

A partir de agora passaremos a utilizar $\mathbf{S}$ ao invés de $\Delta \mathbf{k}$.

3.3 Difração de Raios X

Os assuntos tratados nesse tópico serão explanados de forma compacta visando um entendimento qualitativo do processo de difração de raios X. Para melhor compreensão dos assuntos referidos neste tópico sugere-se uma leitura de BLUNDELL & JOHNSON (1996); DRENTH (1994); DE AZEVEDO (1997).

Feita a cristalização o próximo passo é a exposição do cristal a um feixe de raios X cujo comprimento de onda deve ser próximo ao comprimento das distâncias entre os átomos, para que seja possível a distinção dos átomos individuais. Quando um feixe de raios X atravessa um cristal ocorre uma interação entre o feixe e os elétrons dos átomos no cristal. Os elétrons passam a espalhar radiação para todas as direções, num processo

denominado espalhamento (DE AZEVEDO, 1997; CANDURI, 2001; BLUNDELL & JOHNSON 1996; DRENTH 1994).

Considerando que a luz pode se apresentar ora como onda, ora como partícula, ou seja, a sua dualidade, podemos classificar duas formas de espalhamento: o espalhamento Thomson, que considera a radiação incidente como sendo uma onda eletromagnética plana e polarizada, e o espalhamento Compton, que leva em conta a natureza corpuscular da radiação (DE AZEVEDO 1997; CANDURI, 2001).

No **espalhamento Thomson** o campo elétrico oscilante associado ao feixe de raios X que incide sobre um elétron, obriga este elétron a oscilar em torno da sua posição de equilíbrio. Sabemos que toda partícula carregada emite radiação, assim o elétron, submetido a um campo elétrico oscilante, emite uma onda eletromagnética, que possui o **mesmo comprimento de onda da radiação incidente (espalhamento elástico)**, (DE AZEVEDO, 1997).

O **espalhamento Compton**, ocorre quando raios X incidem sobre elétrons livres ou fracamente ligados e pode ser entendido a partir da teoria quântica, a partir da conservação do momento e da energia da colisão do fóton com o elétron. Assim temos que a **radiação espalhada possui um comprimento de onda maior** que o da radiação incidente (**espalhamento inelástico**) devido à transferência de energia do fóton para o elétron (DE AZEVEDO, 1997).

Experimentalmente encontra-se que a radiação espalhada pelos materiais consiste de duas partes, uma associada ao espalhamento Thomson e outra parte ao espalhamento Compton. A primeira parte (associada ao espalhamento Thomson) possui o mesmo comprimento de onda da radiação incidente; a segunda parte tem um comprimento de onda maior que a radiação incidente, com o aumento do comprimento de onda sendo dependente do ângulo de espalhamento (DE AZEVEDO, 1997).

Quando se calcula o espalhamento de raios X por um cristal se obtém uma expressão conhecida como equação do fator de estrutura. O padrão de difração (fator de estrutura) é a transformada de Fourier da densidade eletrônica da estrutura, de forma inversa, temos que a densidade eletrônica da estrutura é a transformada de Fourier do padrão de difração (DE AZEVEDO, 1997).

Reescrevendo o fator de estrutura, não mais em função da densidade numérica, mas em função da densidade eletrônica temos:

$$F(S) = \int_V \rho(r) \exp 2\pi i(r \bullet s) dv \quad (12)$$

Multiplicando ambos os lados por $(\exp -2\pi i(r' \bullet S))$ e integrando sobre o volume recíproco ($V^* = 1/V$), temos que:

$$\rho(r) = \int_{V^*} F(S) 2\pi (r_j \bullet S) dv^* \quad (13)$$

onde dv^* é o elemento de volume associado ao espaço recíproco.

A integração pode ser substituída por uma somatória, visto que, $F(S)$ não é contínuo e é diferente de zero somente os pontos do retículo recíproco. Consequentemente,

$$\rho(x, y, z) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F(hkl) \exp -2\pi i(hx + ky + lz) \quad (14)$$

Desta forma se os fatores de estrutura, $F(hkl)$, são conhecidos para todas as reflexões, h , k e l , então a densidade eletrônica, $\rho(x,y,z)$, pode ser calculada para cada ponto x , y , z , na cela unitária, (DE AZEVEDO 1997, DRENTH, 1994).

3.4 O Problema da Fase

Durante um experimento de difração de raios X só se registram as intensidades e toda informação sobre a fase é perdida. Portanto, é impossível determinar a estrutura diretamente das medidas do padrão de difração, visto que parte da informação está perdida. O problema da fase é o problema básico de qualquer determinação de estrutura (DE AZEVEDO, 1997; BLUNDELL & JOHNSON, 1976; CANDURI, 2001, DOS SANTOS 2004).

Para ser resolvido o problema da fase, se lança mão de alguns métodos indiretos. Atualmente existem quatro métodos principais para se determinar as fases dos fatores de estrutura: (i) o método de difração anômala com múltiplos comprimentos de onda, que depende da presença de átomos que espalhem forte e irregularmente na própria estrutura da proteína; (ii) métodos diretos, já usados para pequenas proteínas e em fase de desenvolvimento para proteínas maiores; (iii) a substituição isomórfica múltipla; (iv) o método de substituição molecular, que utiliza uma estrutura já resolvida, similar e/ou com enovelamento parecido já determinado, e que será descrito adiante.

3.5 Substituição Molecular

A partir da observação de que moléculas com seqüências primárias homólogas possuem enovelamentos similares, o conhecimento estrutural destas moléculas permite simular um modelo inicial do cristal a ser resolvido e, deste modelo, obter as informações necessárias sobre a fase da radiação difratada pelo cristal real (RHODES, 1999 *apud* DOS SANTOS 2004).

A substituição molecular é um dos métodos disponíveis para se resolver o problema da fase. Nesse método uma estrutura conhecida homóloga é ajustada na cela unitária da estrutura que se pretende resolver e as fases da estrutura conhecida são

usadas como uma primeira aproximação para as fases na nova estrutura. O ajuste do modelo se faz em uma procura em seis dimensões, ou graus de liberdade, que consistem basicamente em procuras de ordem rotacional e translacional (DE AZEVEDO, 1997).

Na substituição molecular deve-se usar um modelo de busca o mais homólogo possível da macromolécula a ser estudada, para proteínas uma identidade igual ou superior a 50% de homologia é suficiente. É amplamente recomendável o uso de modelos com altíssima identidade para que se possa ser evitado problemas de estimativa grosseira das fases e problemas durante a etapa de refinamento da estrutura estudada (NOLASCO, 2005).

Para se fazer a substituição molecular, foi usado o programa AMoRe do pacote de programas CCP4 (CCP4, 1994). Este programa utiliza a estrutura tridimensional do modelo escolhido, segundo alguns critérios para gerar as fases da estrutura em estudo. A figura 3.8 mostra o fluxograma deste programa, e maiores detalhes sobre ele são dados no Apêndice A, e nos trabalhos de DOS SANTOS, 2004 E DE AZEVEDO, 1997.

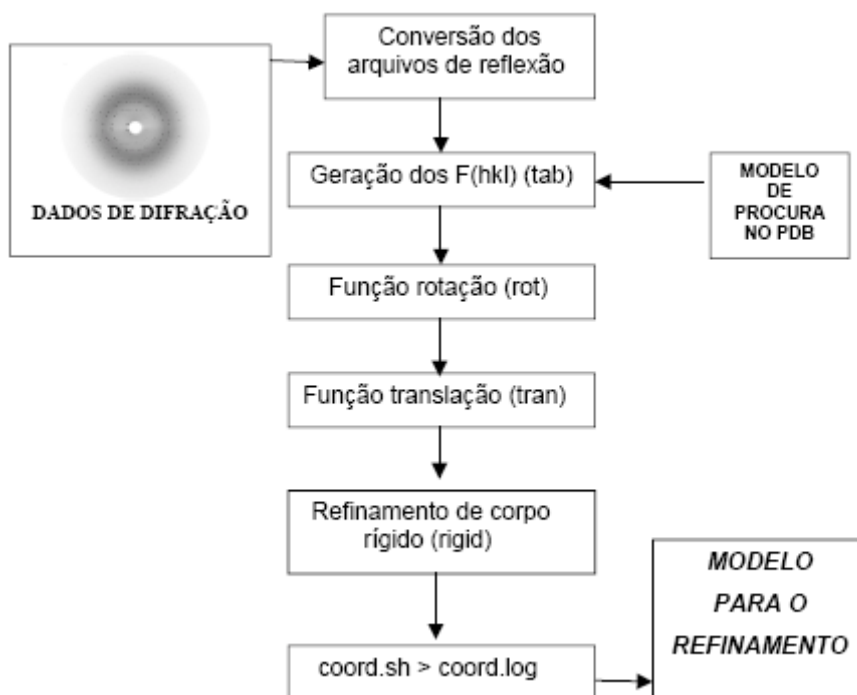


Figura 3.8 Fluxograma do protocolo de substituição molecular de acordo com o programa AMoRe do pacote CCP4 (DOS SANTOS, 2004).

4. Resultados e Discussão

4.1 Cristalização e Coleta de Dados de Difração de Raios X

Os cristais de MtPNP complexados com adenosina e aqueles complexados com adenina foram obtidos através de condições previamente estabelecidas (SHI *et al.*, 2001; NOLASCO 2004, 2005), ambos apresentaram dimensões de, aproximadamente, 0,5 x 0,6 x 0,7mm. E os dados de difração de raios X foram coletados a 1,85Å de resolução e, quando indexados, mostraram pertencer ao grupo de Laue primitivo ortorrômico $P2_12_12$, com parâmetros de cela $a=118,9\text{\AA}$, $b=134,6\text{\AA}$, $c=44,4\text{\AA}$ e os ângulos $\alpha=\beta=\gamma=90^\circ$ (valores finais obtidos após o refinamento destes parâmetros, utilizando as cento e sessenta imagens coletadas) para a MtPNP adenina e de $a=127.6\text{\AA}$, $b=144.4\text{\AA}$, $c=47.6\text{\AA}$ para a MtPNP complexada com adenosina.

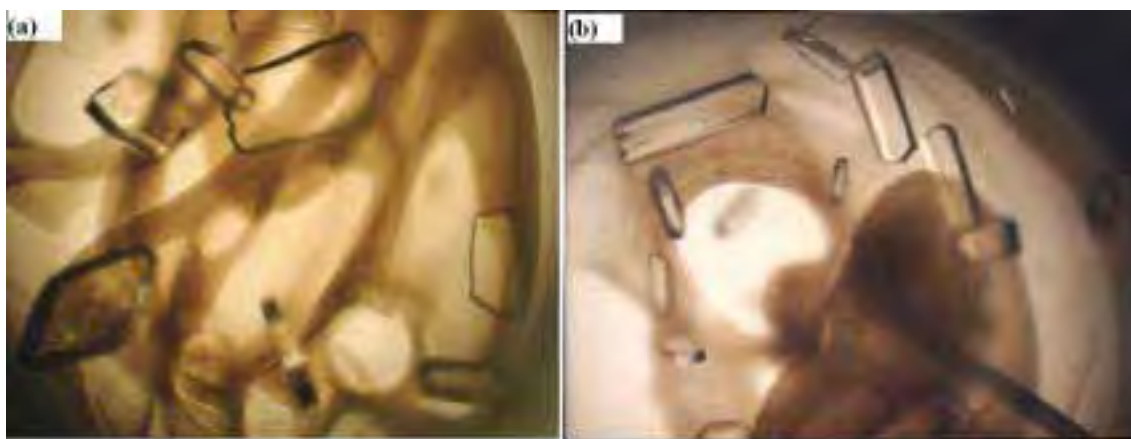


Figura 4.1 Cristais de MtPNP. Em (a) MtPNP complexada com adenosina e em (b) apoenzima.

Para a MtPNP:adenina foram observadas 288 832 reflexões, das quais 33 436 eram reflexões independentes das cento e sessenta imagens coletadas. (Tabela 4.1)

Tabela 4.1 Estatísticas do processamento de dados de difração de raios X da MtPNP: adenina

	Adenina
a (Å)	118,81
b (Å)	134,55
c (Å)	44,5
Grupo espacial	P2 ₁ 2 ₁ 2
Número de imagens utilizadas	160
Número de reflexões independentes	33.436
Completeza (%)	99,3
R [*] _{sym} (%)	7,9
Camada de mais alta resolução (Å)	1,96 a 1,86
Completeza à mais alta resolução (%)	97,6
R [*] _{sym} à mais alta resolução (%)	47

Para a MtPNP complexada com Adenosina foram observadas 439.731 reflexões, das quais 49.294 eram reflexões independentes das cento e sessenta imagens coletadas.

Tabela 4.2 Estatísticas do processamento de dados de difração de raios X da MtPNP: adenosina.

	Adenosina
a (Å)	127.6466
b (Å)	144.4295
c (Å)	47.6932
Grupo espacial	P21212
Número de imagens utilizadas	160
Número de reflexões independentes	49.294
Completeza (%)	99.1
R [*] _{sym} (%)	6,7
Camada de mais alta resolução (Å)	2.14
Completeza à mais alta resolução (%)	94.9
R [*] _{sym} à mais alta resolução (%)	22.9

4.2 Substituição Molecular

Feito o processamento e a interpretação dos dados coletados foi iniciado o procedimento de substituição molecular para solucionar o problema da fase. A estrutura utilizada como modelo de busca para os dois conjuntos de dados foi a estrutura da MtPNP complexada com hipoxantina. Para a substituição molecular foi utilizado o programa AMoRe (NAVAZA, 1994), do pacote de programas CCP4 (Collaborative Computational Project n° 4) e os valores de coeficiente de correlação (CC) e R_{factor} inicial são mostrados na tabela 4.3.

Tabela 4.3 – Valores obtidos na etapa de substituição molecular.

Substituição Molecular								
	CC (%)	R_{factor} inicial (%)	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\gamma(^{\circ})$	$T_x(\text{\AA})$	$T_y(\text{\AA})$	$T_z(\text{\AA})$
MtPNP•Adenina	69,3	37,6	68,96	14,69	202,02	0,0992	0,2833	0,4283
MtPNP•Adenosina	54,4	37,2	71,68	14,48	212,90	0,1037	0,2821	0,4239

4.3 Refinamento Cristalográfico

O refinamento cristalográfico das estruturas foi realizado com a utilização do programa REFMAC5 (MURSHUDOV *et al.*, 1997), parte integrante do pacote CCP4. O método de refinamento aplicado pelo REFMAC5 é o Método da Máxima Verossimilhança (*The Maximum Likelihood Method*) (FISHER, 1921; NOLASCO 2005).

Os valores obtidos na etapa de refinamento cristalográfico são mostrados na tabela 4.4:

Tabela 4.4 – Estatísticas do refinamento cristalográfico da estrutura da MtPNP complexada com Adenina e MtPNP complexada com Adenosina.

MtPNP·Adenina		
		37,08Å - 1,85Å
<i>Faixa de Resolução</i>		
R _{factor}	18,52%	
R _{free}		23,87%
Nº de aminoácidos (trímero)		791
Nº de moléculas de água (trímero)		560
MtPNP·Adenosina		
		44,677 – 2,142Å
<i>Faixa de Resolução</i>		
R _{factor}		19,41%
R _{free}		25,66%
Nº de aminoácidos (trímero)		791
Nº de moléculas de água (trímero)		454

4.4 Descrição da estrutura MtPNP:Adenina

A estrutura quaternária da MtPNP complexada com adenina apresenta as mesmas características das estruturas já resolvidas anteriormente, trata-se de uma estrutura simetricamente homotrimérica de arranjo triangular das subunidades, assim como as PNPs de mamíferos. Cada monômero da proteína está enovelado numa topologia do tipo α/β , consistindo de onze fitas β circundadas por oito hélices α .

Esta estrutura, no entanto caracterizou-se por uma peculiaridade: foi encontrada uma densidade eletrônica característica da adenina fora dos três sítios ativos, na região central da proteína, sem de fato pertencer a nenhuma das três subunidades, (figura 4.2).

A adenina central só se liga à proteína através de ligações com moléculas de água, não fazendo nenhuma ligação com resíduos de aminoácidos, os quais se encontram afastados desta adenina.

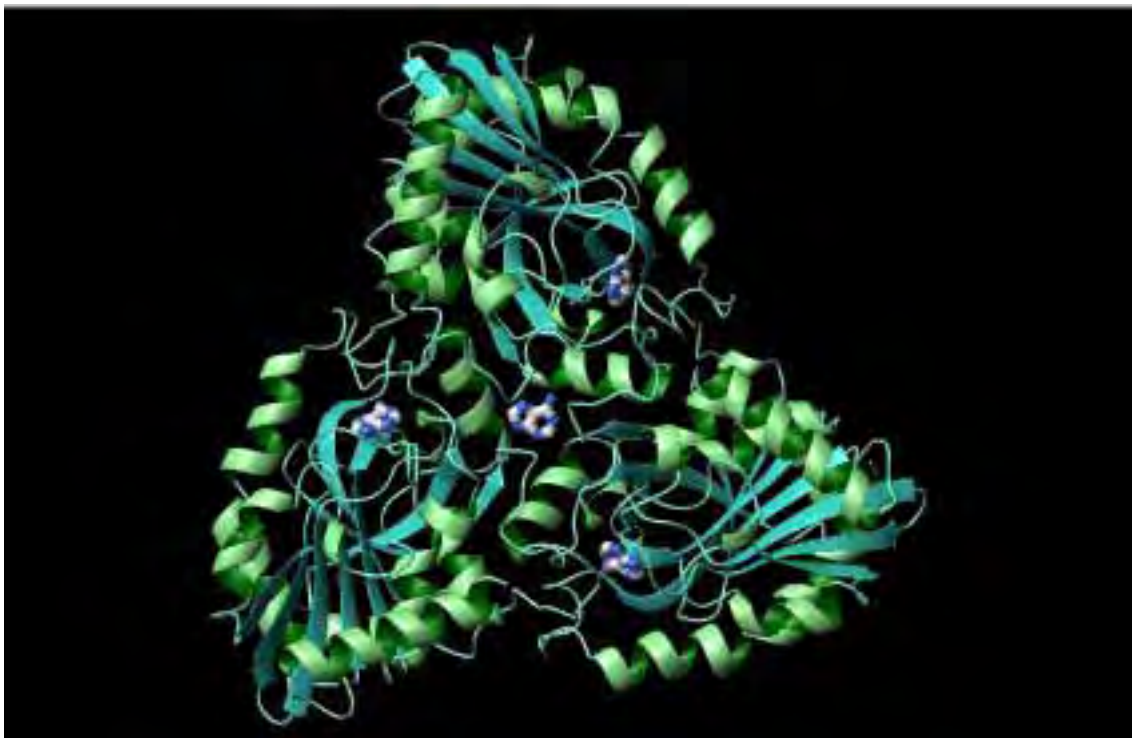


Figura 4.2 Estrutura da MtPNP complexada com adenina, em detalhe a adenina no centro do trímero. Programa utilizado Pymol (DELANO, 2004).

Ao se resolver a estrutura cristalográfica da MtPNP **complexada com adenosina**, verificou-se a presença de **adenina** nos sítios ativos bem como no centro do trímero, em nenhum dos sítios foi encontrada adenosina como era esperado. (ver figura 4.3). Este fato evidencia que: (1) a MtPNP aparentemente possui afinidade enzimática com a adenina; (2) a enzima se encontra ativa no cristal exercendo sua atividade catalítica; (3) a molécula de adenina no centro do trímero não é meramente um artifício cristalográfico, uma vez que foi observada tanto na estrutura obtida por *soaking* quanto na estrutura obtida pela enzima complexada.

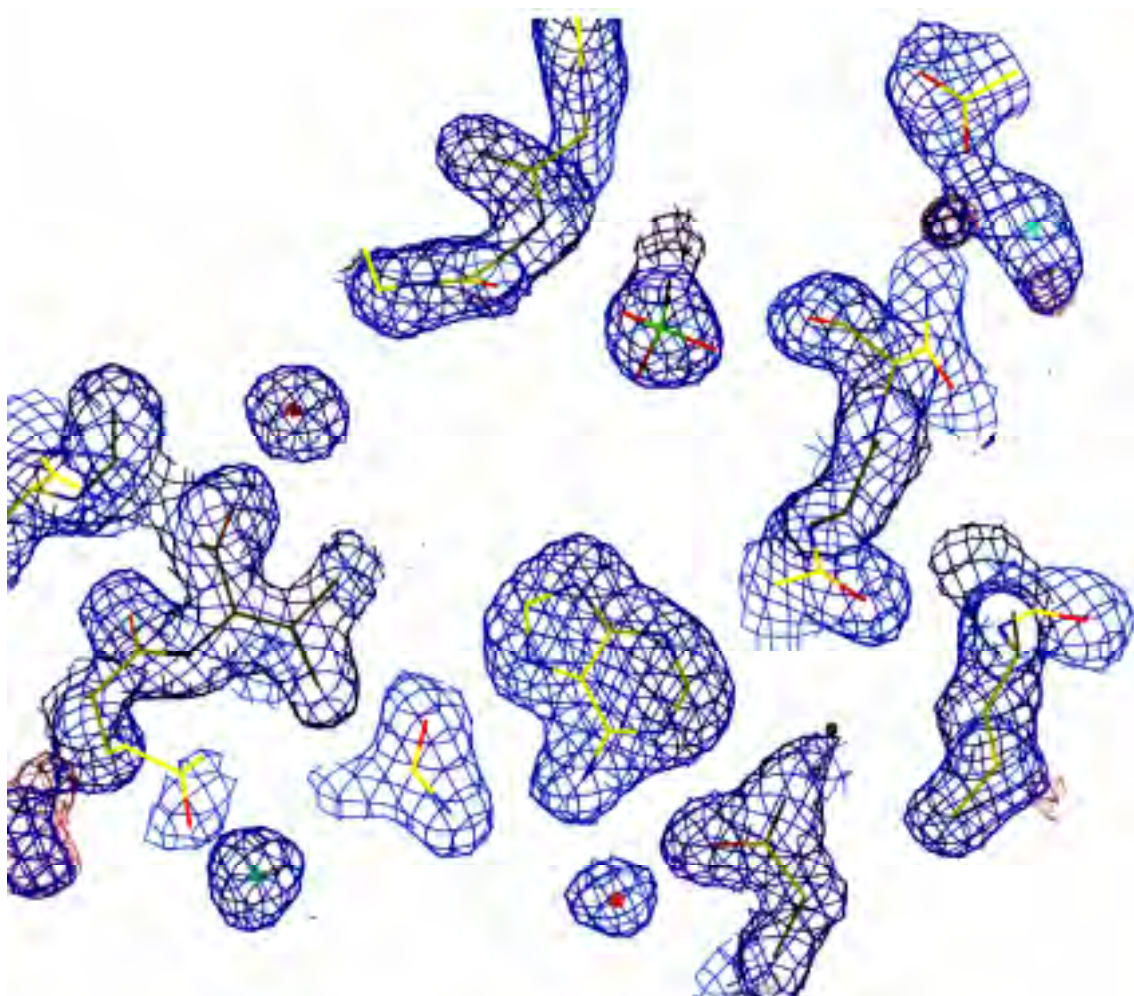


Figura 4.3 Mapa de densidade eletrônica mostrando a molécula de adenina no centro do trímero. Programa utilizado XtalView (MCREE, 1999).

Para que fosse comprovado o fato de que se tratava realmente de uma molécula de adenina no centro do trímero, foi realizado um refinamento sem os ligantes e gerado um *omit map* e foram observadas tanto a região do centro do trímero quanto os sítios ativos para efeito de comparação das densidades. O resultado mostra que realmente se trata de uma molécula de adenina no centro do trímero, com uma densidade eletrônica característica da molécula de adenina (ver Figuras 4.4 e 4.5).

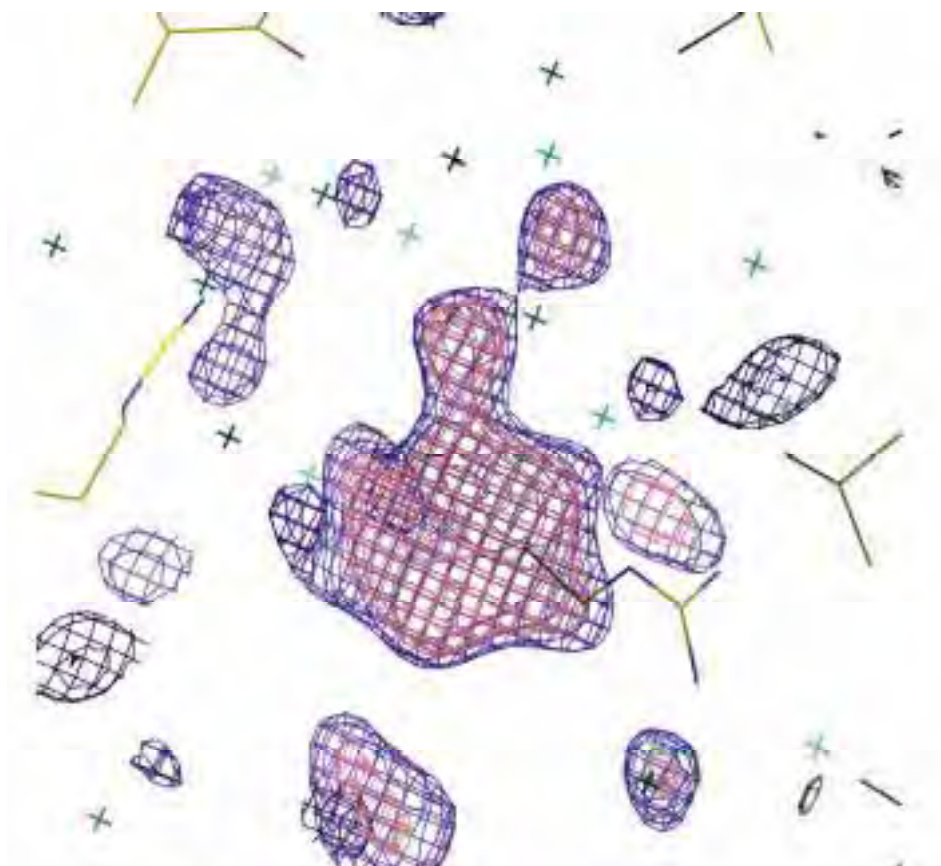


Figura 4.4 *Omit map* mostrando a densidade eletrônica da molécula de adenina no centro do trímero. Programa utilizado XtalView (MCREE, 1999).

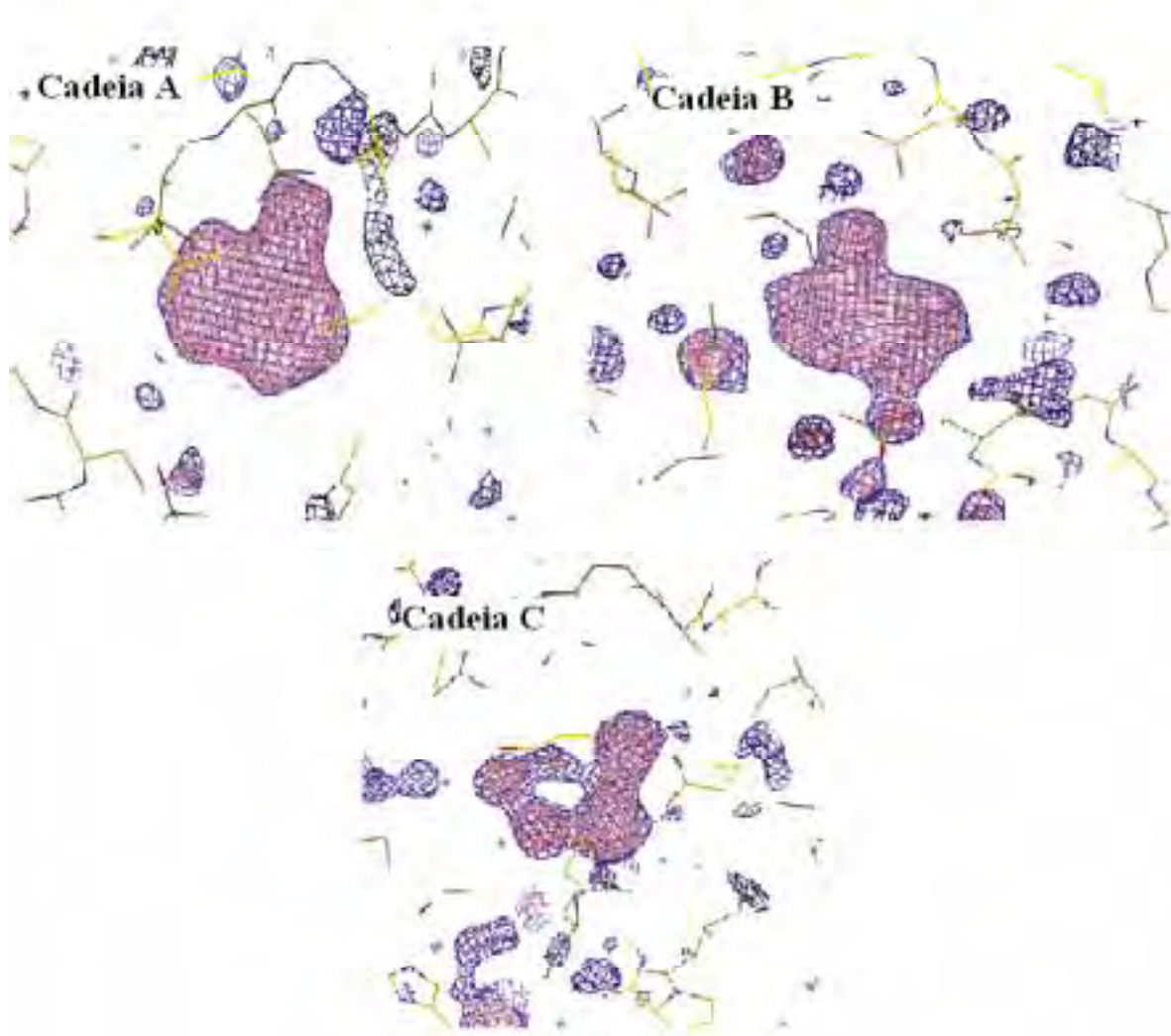


Figura 4.5 *Omit map* dos sítios ativos. Programa utilizado XtalView (MCREE, 1999).

Por se tratar de dois conjuntos com o mesmo ligante (ambos com adenina), foi optado pela análise de apenas um deles, já que ambos apresentam a mesma estrutura com o mesmo ligante. A estrutura que estará sendo analisada é a estrutura obtida através de *soaking*.

4.5 Qualidade do modelo

Para verificação da qualidade estereoquímica do modelo se fez uso do **DIAGRAMA DE RAMACHANDRAN**, o qual analisa as posições permitidas dos átomos de cada resíduo de aminoácido levando em conta os ângulos Φ e Ψ da cadeia principal.

A estrutura em questão apresenta 93,2% dos resíduos nas regiões mais favoráveis do diagrama de Ramachandran Φ - Ψ (Figura 4.6), 6,3% dos resíduos nas regiões adicionalmente permitidas, nenhum dos resíduos nas regiões generosamente permitidas e outros três resíduos (0,5%) em regiões não permitidas do mapa. As análises do mapa de densidade eletrônica ($2F_{\text{obs}} - F_{\text{calc}}$) validam a estrutura como um todo e levam a crer na sua veracidade, visto que apenas alguns aminoácidos do N- e do C-terminais não ajustam perfeitamente com suas respectivas densidades eletrônicas.

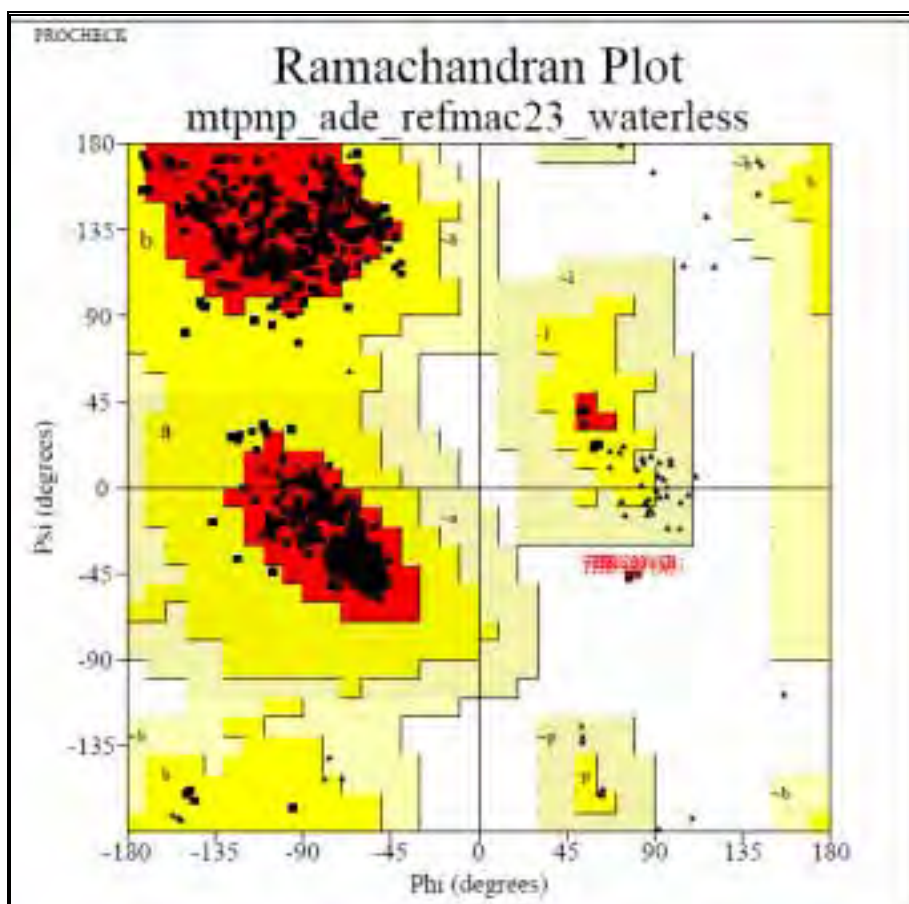


Figura 4.6 - Diagrama de Ramachandran Φ - Ψ para a estrutura da MtPNP complexada com adenina. Programa utilizado Procheck (LASKOWSKI et al 1994).

Outro parâmetro utilizado para análise da qualidade do modelo foi o **G-FACTOR**. O G-Factor fornece uma medida de uma propriedade estereoquímica da proteína, se é dada como normal ou não. Os valores do G-Factor são computados para as propriedades do ângulo de torção e geometria covalente (LASKOWSKI et al, 1993).

O G-Factor global é uma medida da normalidade da estrutura como um todo, representa uma média para o G-Factor de cada resíduo da estrutura. O valor ideal do G-Factor deve ser acima de -0,5 (MORRIS, 1986). O valor do G-Factor obtidos para o Ângulo de Torção foi de -0.25, para a Geometria Covalente foi de 0,38 e o valor Global foi de 0,02.

Foi feita análise também através do **VERIFY-3D** (BOWIE et al, 1991; LUTHY et al, 1992; KABSCH & SANDER, 1983), que mede a confiabilidade da seqüência de aminoácidos de uma proteína e o modelo tridimensional. O resultado é uma medida da compatibilidade da seqüência primária e a estrutura 3D construída, a razão entre o escore total e o escore ideal deve ser maior que 0,45 para que a estrutura seja compatível. O Escore Total obtido para a estrutura foi de 138,61, o Escore Ideal 121,29 e a razão entre eles foi de 1,14.

Além das análises anteriores foi realizada também uma análise do **RMSD da geometria ideal** para o comprimento e ângulo de ligação, através do programa X-PLOR (BRÜNGER, 1992), os valores ideais para o comprimento e ângulo de ligação devem estar abaixo de 0,1Å e 10° respectivamente. O valor obtido para o comprimento de ligação da estrutura foi de 0,016Å e o ângulo de ligação foi de 1,71°.

4.6 A adenina no centro do trímero

Como dito anteriormente, foi observada uma densidade eletrônica característica da molécula de adenina na região central do trímero, a qual foi confirmada como adenina através do *omit map* já mostrado anteriormente. A figura 4.7 mostra essa densidade eletrônica já refinada.

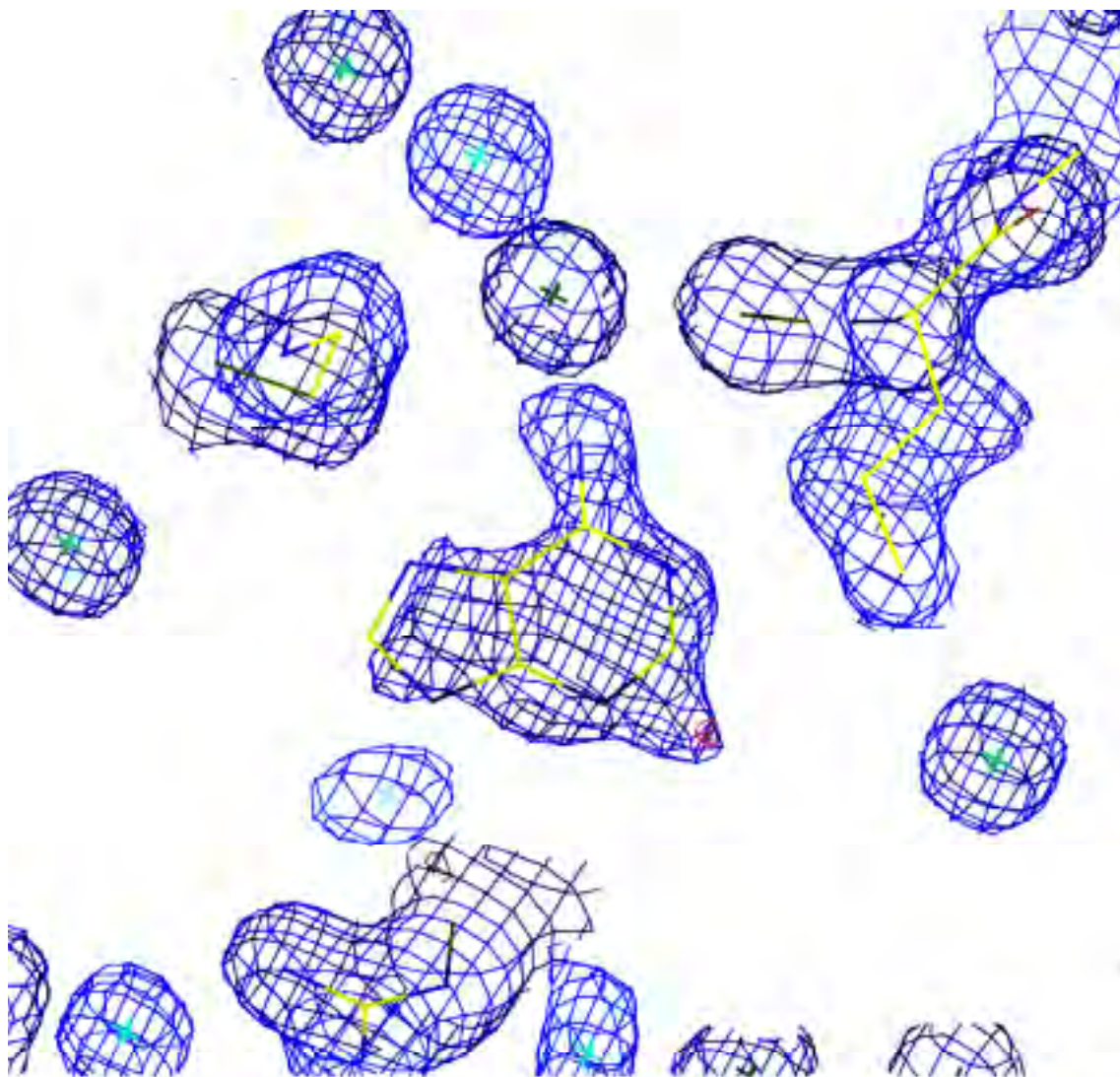


Figura 4.7 Densidade eletrônica da molécula de adenina no centro do trímero. Gerado pelo programa XtalView (MCREE, 1999).

A molécula de adenina presente no centro do trímero faz ligações de hidrogênio com quatro moléculas de água, essas moléculas de água por sua vez se ligam a outras moléculas de água e/ou resíduos de aminoácidos da proteína. A alanina 144 da cadeia C, é o resíduo mais próximo da adenina do centro do trímero distando 3,61Å do nitrogênio

da adenina. As moléculas de água, além de se ligarem a outras moléculas de água, se ligam também a alguns resíduos de aminoácidos próximos, a Alanina 144 da cadeia A, e Arginina 145 da cadeia C (ver figura 4.8).



Figura 4.8 A molécula de adenina presente no centro do trímero. Gerado pelo programa XtalView (MCREE, 1999).

Uma possível explicação para o aparecimento deste ligante fora do seu sítio ativo é o fato da proteína ter sido supersaturada de ligantes. Adiante nesse trabalho, será feita uma análise do potencial eletrostático com a qual se pretende obter mais esclarecimentos a respeito da ocupação do centro do trímero pela molécula de adenina.

4.7 O sítio ativo

Pode-se observar no sítio ativo da MtPNP, além da Asn 231 e do Glu 189, já verificados anteriormente, a presença efetiva de uma molécula de água. O Glu 189 parece estar se alternando em interações ora com essa molécula de água, ora com um átomo de nitrogênio do ligante.

Além da participação da molécula de água no sítio de ligação, outra participação ainda não evidenciada foi observada, trata-se do resíduo Ala 120, que faz interações eletrostáticas com a adenina a uma distância de 3,31 Å na cadeia A.

Foi observada também uma possível participação do ligante com o resíduo Val 205, no entanto apresentou uma distância maior que 3,5 Å nas três cadeias (ver Figuras 4.9 e 4.10).

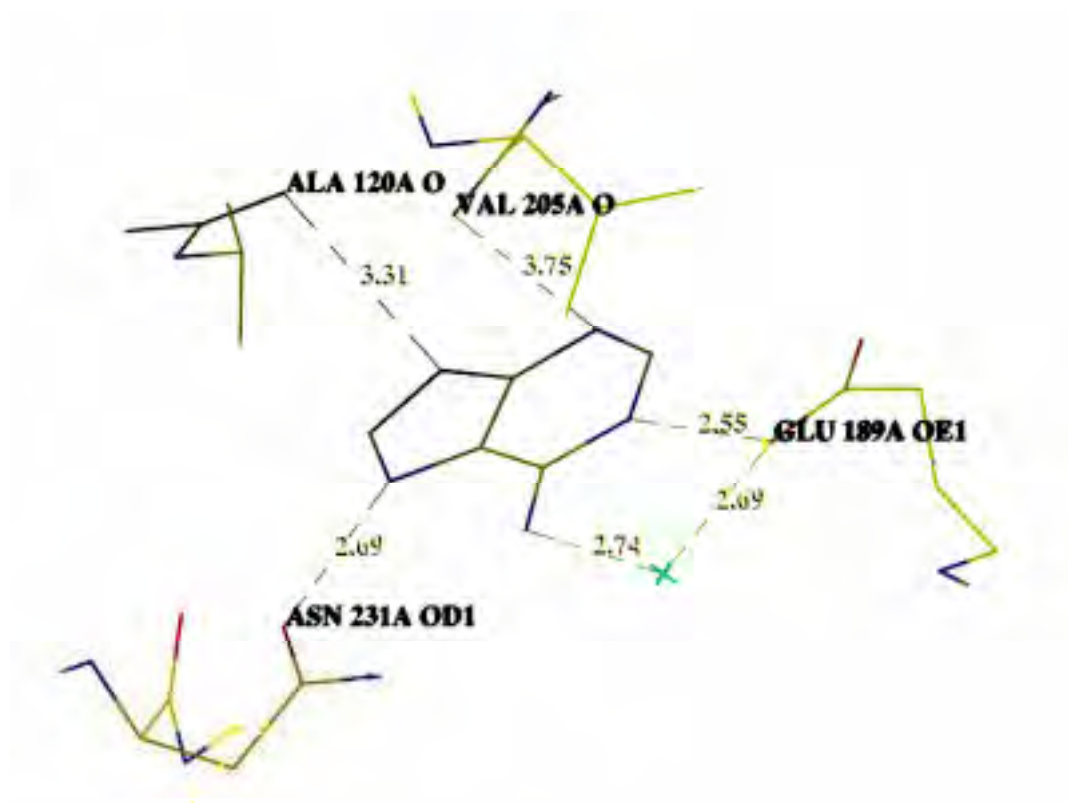
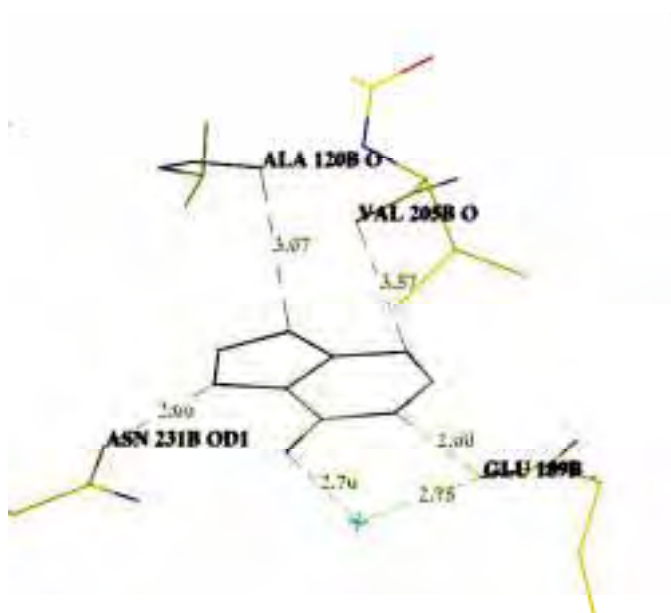


Figura 4.9 sítio ativo da cadeia A com as distâncias correspondentes de cada resíduo. Gerado pelo programa XtalView (MCREE, 1999).

Os três sítios conservam esses resíduos em distâncias praticamente iguais em cada uma das três cadeias, bem como as suas posições e ângulos de ligação como já era esperado.

(A)



(B)



Figura 4.10 em (A) o sítio ativo da cadeia B e em (B) o sítio ativo da cadeia C. Gerado pelo programa XtalView (MCREE, 1999).

Outra análise feita para atestar a participação dos resíduos Ala 120, Val 205 e da molécula de água no sítio ativo foi uma sobreposição da estrutura da MtPNP complexada com hipoxantina. Esta estrutura foi utilizada na substituição molecular para a resolução da estrutura da MtPNP complexada com adenina, por essa razão não foi utilizado nenhum programa para sobrepor as estruturas, que já se encontravam sobrepostas. Nesta análise, verificou-se, de maneira clara a concordância na presença da molécula de água no sítio ativo. A figura 4.11 mostra as duas estruturas simultaneamente.

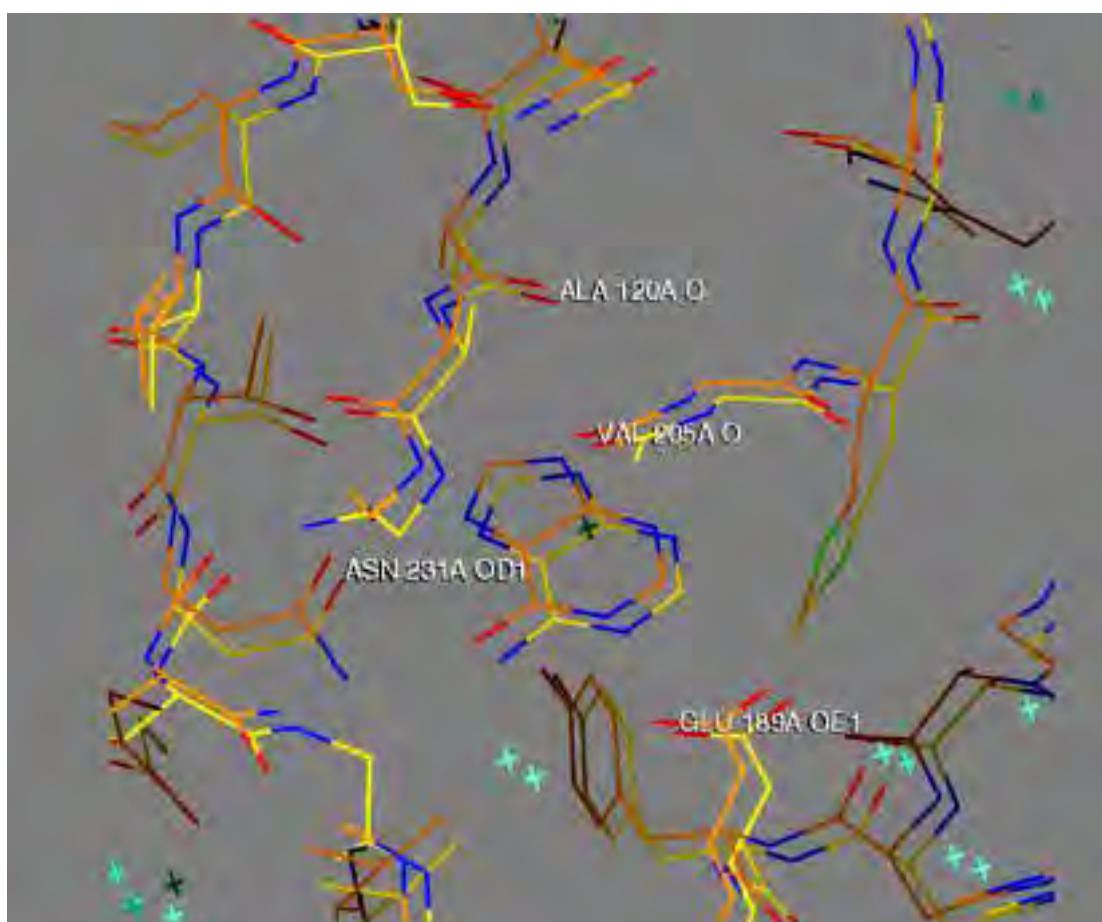


Figura 4.11 sobreposição das estruturas de MtPNP complexada com hipoxantina e com adenina, a estrutura complexada com adenina aparece em amarelo. Gerado pelo programa XtalView (MCREE, 1999).

Além da sobreposição das estruturas foram analisados também os sítios ativos de outras duas estruturas de MtPNP depositadas no banco de dados de proteínas (PDB – Protein Data Bank – www.pdb.org). Foram analisados os sítios MtPNP complexada

com deazahipoxantina (código para acesso no pdb 1I80) e com Dadme-Imucilina H (código para acesso no pdb 1G20).

Nesta análise procurou-se observar a similaridade na disposição dos resíduos participantes quando da presença da molécula de água.

Verificou-se similaridade nas duas estruturas analisadas, tanto na posição dos resíduos participantes do sítio, quanto na presença e posição da molécula de água. Essas estruturas mostraram a mesma disposição da molécula de água em relação ao ácido glutâmico e ao ligante (Ver figuras 4.12 e 4.13). Mais uma vez fazendo a confirmação do que a estrutura complexada com adenina já havia mostrado.



Figura 4.12 sítio ativo da MtPNP complexada com Dadme-Imucilina H. Programa utilizado XtalView (MCREE, 1999)..

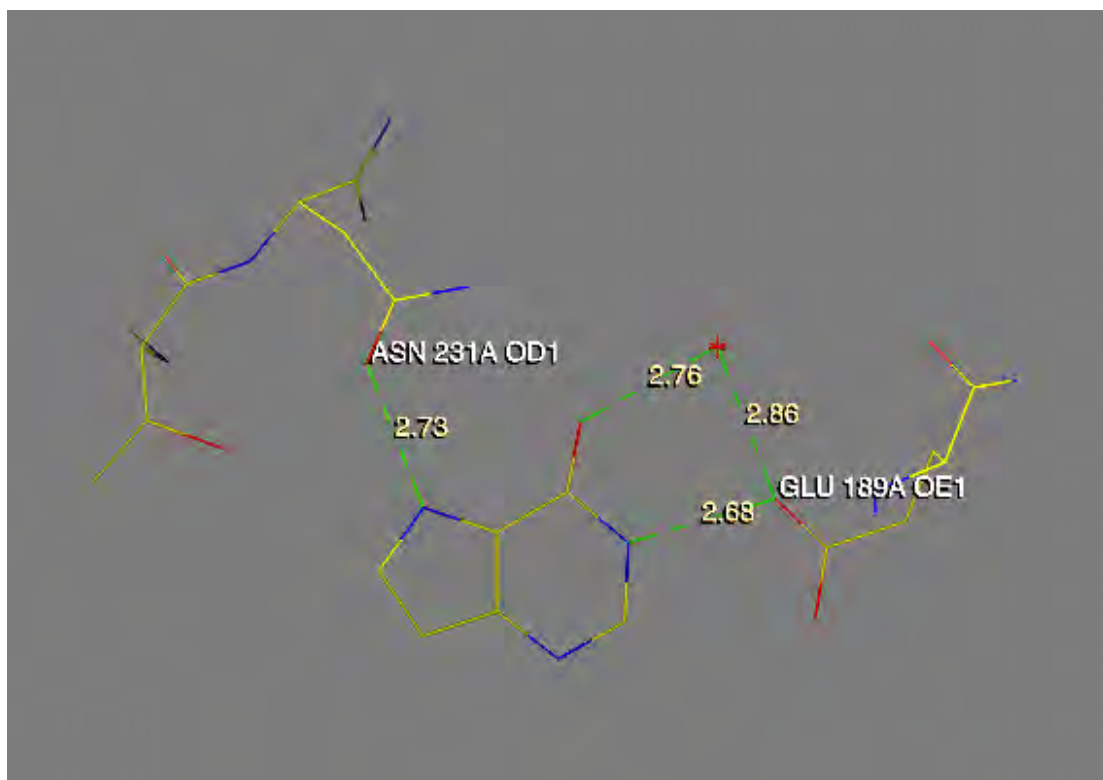


Figura 4.13 sítio ativo da MtPNP complexada com deazahipoxantina. Programa utilizado XtalView (MCREE, 1999)..

Uma vez esclarecido a presença da molécula de água e a participação dos resíduos no sítio, passou-se a uma análise comparativa com o sítio ativo da HsPNP (*Homo sapiens* Purina Nucleosídeo Fosforilase), a importância deste estudo está no fato da HsPNP não ter afinidade com os ligantes Adenosina e Adenina, a melhor compreensão dos sítios ativos poderia servir de base para se propor inibidores seletivos para a MtPNP.

Foi feita uma sobreposição das estruturas de MtPNP (complexada com Adenina) e da HsPNP complexada com guanina (estrutura depositada no PDB – Protein Data Bank, código de acesso 1V2H). Para a sobreposição foi utilizado o programa LSQKAB componente do pacote de programas CCP4 (CCP4, 1994). A sobreposição foi feita com os Carbonos α dos resíduos de aminoácidos, utilizando a HsPNP como corpo móvel e a MtPNP como rígido, as duas faixas de resíduos não equivalentes utilizadas para esta sobreposição foram do resíduo Leu 212 ao resíduo Asp 215 e do resíduo Val 217 ao

resíduo Val 222 para a HsPNP e para a MtPNP foram utilizadas as faixas de resíduos da Leu 200 ao resíduo Asp 203 e Val 205 ao resíduo Val 205, utilizando somente o monômero da cadeia A, respectivamente, e observou-se, um RMSD de 0,302Å.

4.8 Estudo do potencial eletrostático

Um estudo da superfície do potencial eletrostático feito através do programa Pymol (DELANO, 2004) mostra uma peculiaridade na MtPNP, uma predominância de um potencial positivo em um dos lados da proteína, com três grandes morros positivos e um bolsão positivo no centro da proteína. O outro lado da proteína, que fica mais longe dos ligantes apresenta um potencial mais equilibrado com predominância de cargas negativas distribuídas quase de maneira uniforme, (figura 4.14, 4.15 e 4.16).

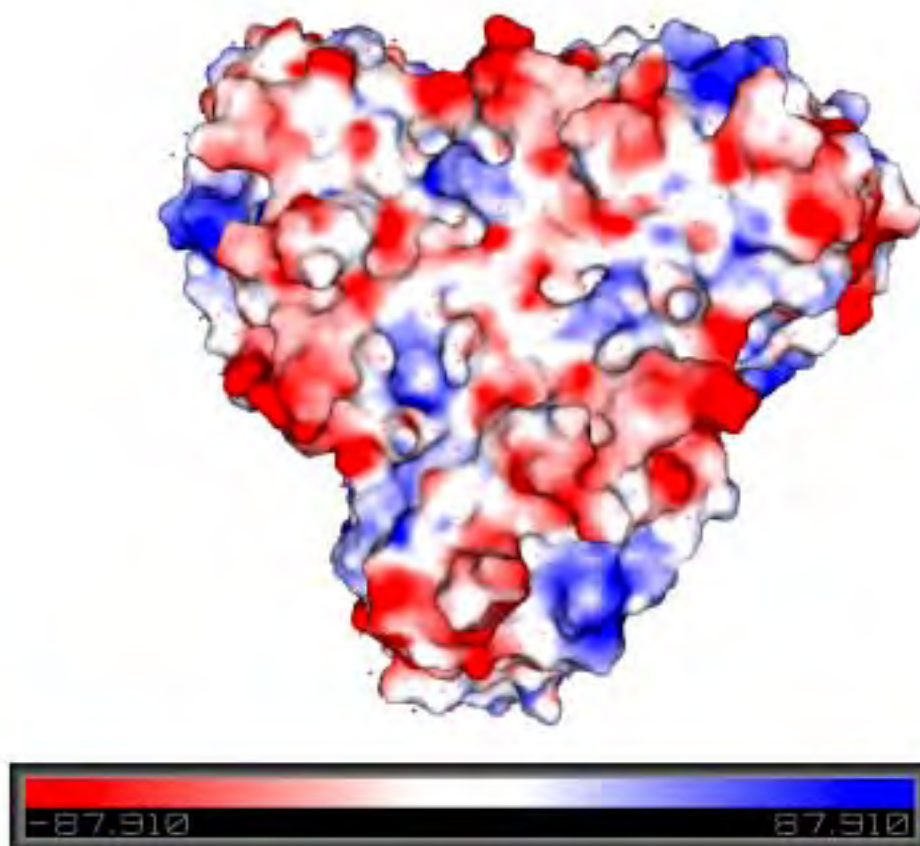


Figura 4.14 a superfície de potencial eletrostático da proteína MtPNP. Programa utilizado Pymol (DELANO, 2004).

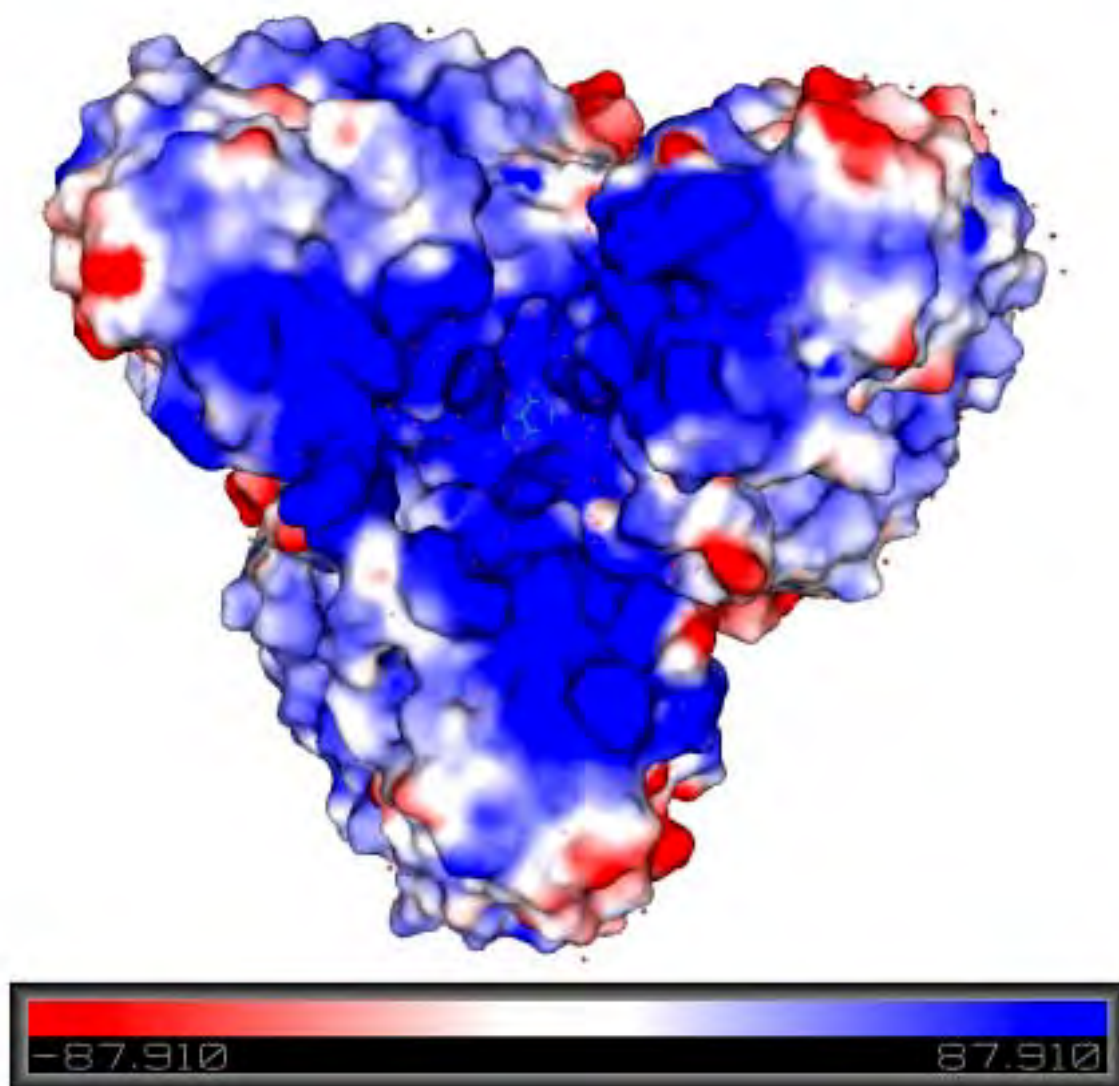


Figura 4.15 A superfície de potencial eletrostático vista do lado oposto à figura anterior, onde visualiza-se a molécula de adenina central. Programa utilizado Pymol (DELANO, 2004).

A adenina possui cinco átomos de nitrogênio e tem afinidade por cargas negativas, incluindo o oxigênio presente nas moléculas de água, o bolsão positivamente carregado no centro do trímero é um lugar energeticamente positivo, e provavelmente a molécula de adenina só se encontra nesse bolsão por causa da grande quantidade de moléculas de água, que causa uma mudança no potencial eletrostático.

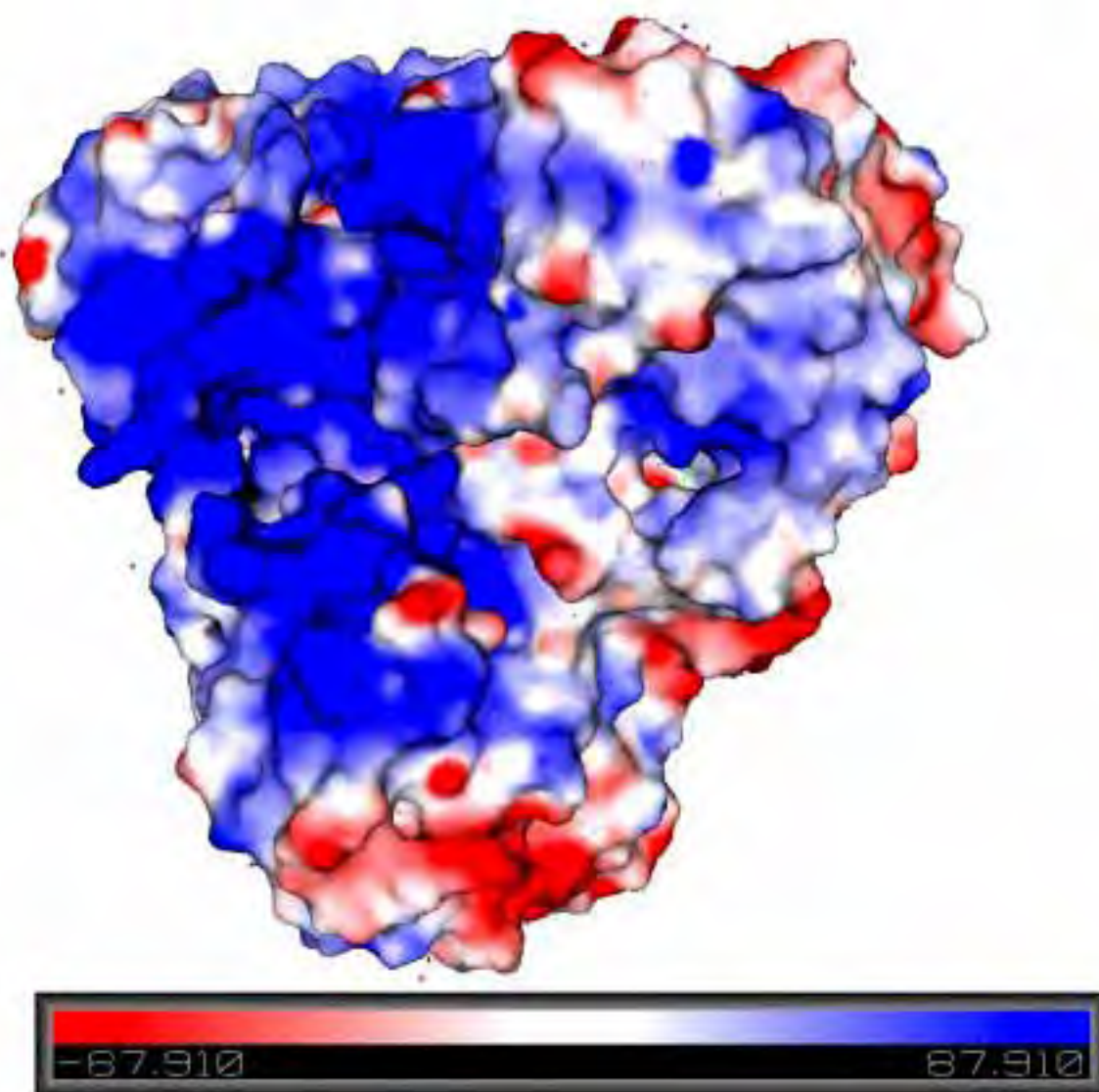


Figura 4.16 Superfície de Potencial Eletrostático (vista lateral). Programa utilizado Pymol (DELANO, 2004).

A superfície de potencial eletrostático no sítio ativo, também se mostrou carregada, com pouco espaço para moléculas de água, predominância de cargas negativas no seu interior e presença de cargas positivas mais externas (Figuras 4.17).

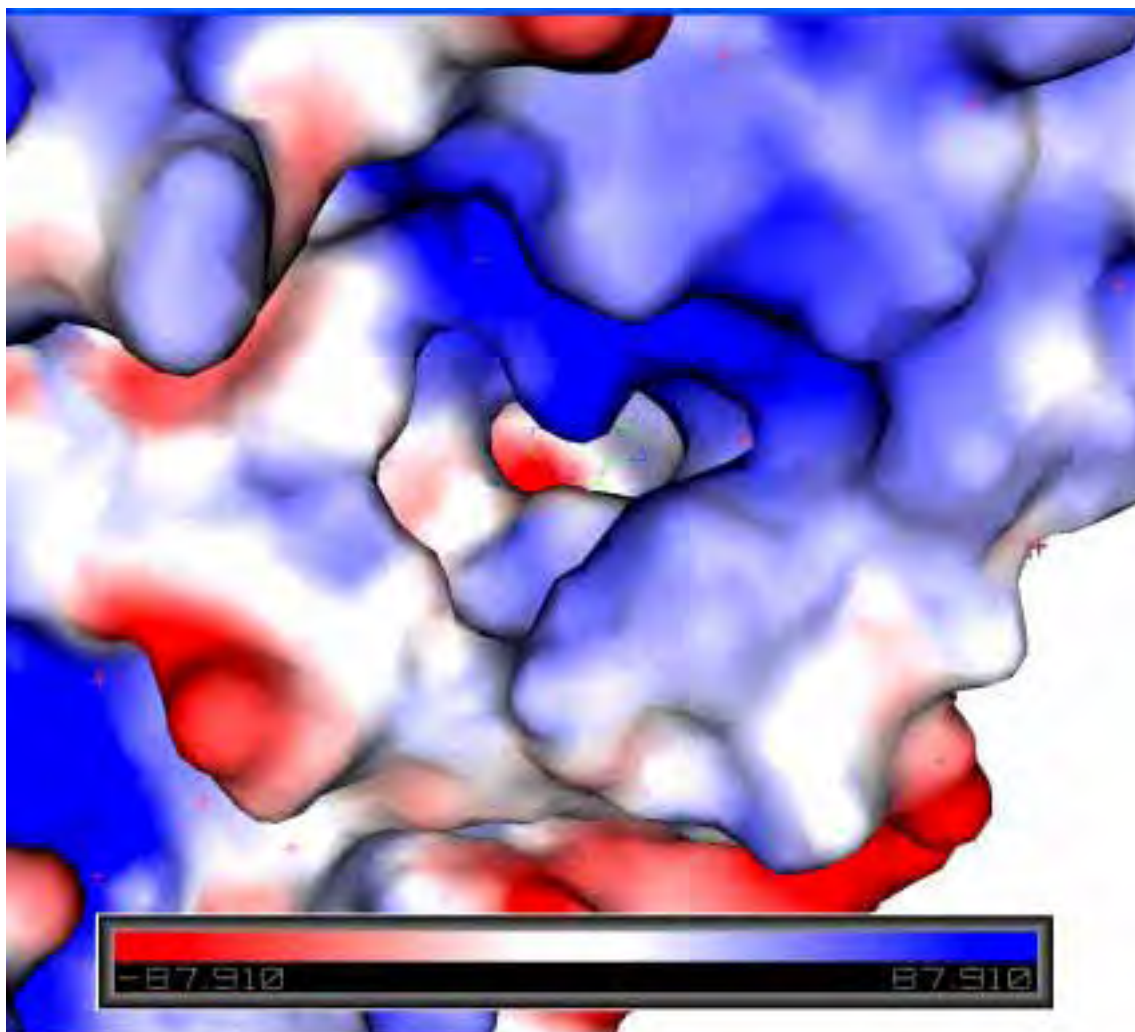


Figura 4.17 superfície de potencial eletrostático no sítio ativo. Programa utilizado Pymol (DELANO, 2004).

Uma mesma análise do sítio ativo da HsPNP mostrou grande diferença na distribuição do potencial eletrostático (Figura 4.18 e 4.19). Para a enzima humana não foi encontrada nenhuma distribuição de potencial eletrostático positivo próximo ao sítio ativo, ao contrário da MtPNP cujo exterior é todo positivo. Esta pode ser uma possível explicação pela qual existe afinidade da MtPNP pela molécula de adenosina enquanto a HsPNP não tem afinidade pela adenosina.

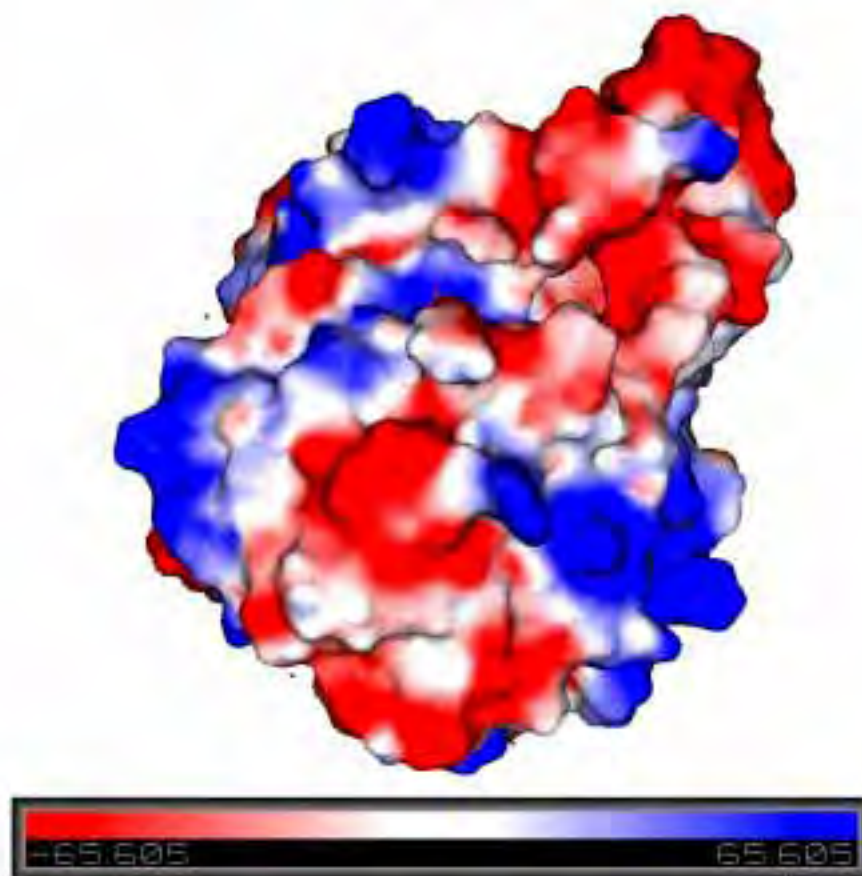


Figura 4. 18 Potencial Eletrostático do monômero da HsPNP. Programa utilizado Pymol (DELANO, 2004).

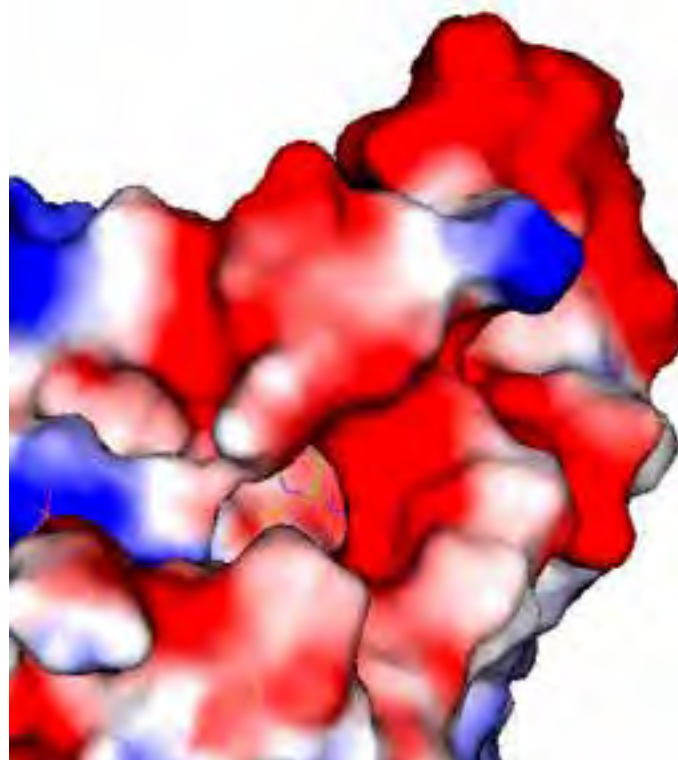


Figura 4.19 Detalhe do potencial eletrostático do sítio ativo da HsPNP. Programa utilizado Pymol (DELANO, 2004).

5. Conclusão

O estudo realizado, sendo fiel ao seu propósito inicial, fez uma análise estrutural da Purina Nucleosídeo Fosforilase para um conhecimento maior desta enzima tão importante do metabolismo do *Mycobacterium tuberculosis*, causador da tuberculose que hoje tem assolado o mundo de maneira assustadora.

O estudo mostrou ainda as peculiaridades da superfície de potencial eletrostático da MtPNP, provável causa do aparecimento de uma molécula de ligante no centro do trímero.

A MtPNP complexada com adenina tem estrutura bastante semelhante à HsPNP, e também possui o sítio ativo composto pelos mesmos resíduos, e apesar dessa grande semelhança estrutural, ambas parecem ter afinidades diferentes: MtPNP foi cristalizada com Adenosina e apresentou Adenina na sua estrutura, enquanto a enzima humana não cliva Adenosina em Adenina.

Estes estudos estruturais nos permitem fazer uma primeira classificação da MtPNP, segundo o modelo de classificação proposto por Bzowska (2000), na qual ela se classificaria como homo-trímero com especificidade de homo-hexâmero.

Uma explicação para as possíveis diferenças de afinidades da enzima do *Mycobacterium tuberculosis* e a enzima humana estaria na distribuição eletrostática absolutamente distinta nas duas. A MtPNP apresenta uma predominância de um potencial eletrostático positivo no lado em que os ligantes ficam mais expostos, e seu sítio ativo é negativo somente no interior, enquanto a HsPNP possui um sítio ativo com uma distribuição eletrostática negativa.

Os resultados desse estudo permitem uma maior compreensão do sítio ativo da MtPNP com todos os resíduos que fazem parte da catálise. Isso pode propiciar um

melhor entendimento sobre o mecanismo catalítico desta enzima ainda não completamente elucidado.

Sabe-se também que o estudo estrutural e um conhecimento mais avançado do sítio ativo desta proteína pode nos dar base para desenhos de inibidores mais potentes e específicos para o *Mycobacterium tuberculosis*. Esses inibidores, por sua vez, podem vir a ser a chave para novos medicamentos mais eficazes e menos nocivos ao paciente no combate à Tuberculose.

Referências

AGARWAL, R.P.; PARKS, R.E. JR.; Purine nucleoside phosphorylase from human erythrocytes. Volume Content and behavior of sulfhydryl groups. **J. Biol. Chem**, v. 246, n. 11, p. 3763-3768, 1971.

AVARBOCK, D.; SALEM, J.; LI, L.S.; *et al.*; Cloning and characterization of a bifunctional RelA/SpoT homologue from Mycobacterium tuberculosis. **Gene**, v. 233, p. 261-269, 1999.

BASSO, L.A.; SANTOS, D.S.; SHI, W.; *et al.*; Purine nucleoside phosphorylase from Mycobacterium tuberculosis. Analysis of inhibition by a transition-state analogue and dissection by parts. **Biochemistry**, v. 40, n. 28, p. 8196-8203, 2001.

BLUNDELL, T. L.; & JOHNSON, L. N.; **Protein Crystallography**. London: Academic Press, 1976.

BOWIE, J. U., *et al.* A method to identify protein sequences that fold into a known three-dimensional structure. **Science**, v.253, p. 347-352, 1987.

BZOWSKA, A.; KULIKOWSKA, E.; SHUGAR, D.; Purine nucleoside phosphorylases: properties, functions, and clinical aspects. **Pharmacol. Ther.** v. 88, n. 3, p. 349-425, 2000.

CANDURI, F.; **Estudo Estrutural de Proteinases Aspárticas por Cristalografia de Difração de Raios X**, UNESP, SP: 2001. Dissertação de doutorado.

CANDURI, F.; FADEL, V.; BASSO, L.A.; PALMA, M.S.; DE AZEVEDO, W.F.Jr.
Biochem. Biophys. Res. Commun. v. 327, p. 646-649, 2005.

CANDURI, F.; DOS SANTOS, D.M.; SILVA, R.G.; *et al.*; Structures of human purine nucleoside phosphorylase complexed with inosine and ddI. **Biochem. Biophys. Res. Commun**, v. 313, n. 4, p. 907-914, 2004.

Collaborative Computational Project N° 4. The CCP4 suite: programs for proteins crystallography. **Acta Crystallogr.**, v. D50, p. 760-763, 1994.

COLE, S.T.; BROSCHE, R.; PARKHILL, J.; *et al.*; Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. **Nature**, v. 393, n. 6685, p. 537-544, 1998.

CORPET, F.; Multiple sequence alignment with hierarchical clustering. **Nucleic Acids Res.**, v. 16, n. 22, p.10881-10890, 1988.

COTRAN RS, KUMAR V, ROBBINS SL. **Patologia estrutural e funcional**. 4. ed. Rio de Janeiro: Guanabara Koogan; 1991

DAVIS, B.D.; BULBECCO, R.; EISEN, H.N.; *et al.*; **Microbiologia – Infecções Bacterianas e Micóticas**, São Paulo: Edart, 1973.

DE AZEVEDO, W.F. JR.; **Introdução à biocristalografia com o estudo estruturas da quinase dependente de ciclina 2 (CDK2) complexada com inibidores**, USP, SP 1997. Dissertação de doutorado.

DE AZEVEDO, W.F. JR.; CANDURI, F.; DOS SANTOS, D.M.; *et al.*; Crystal structure of human purine nucleoside phosphorylase at 2.3Å resolution. **Biochem. Biophys. Res. Commun**, v. 308, n. 3, p. 545-552, 2003.

DELANO, W. L. **The PyMOL Molecular Graphics System**. San Carlos: CA DeLano Scientific, 2004.

DIVISÃO DE CONTROLE DA TUBERCULOSE, CVE – SES/SP. Disponível em: www.saude.sp.gov.br/tuberculose/. Acesso em junho de 2006.

DOS SANTOS, D.M.; CANDURI, F.; PEREIRA, J.H.; *et al.*; Crystal structure of human purine nucleoside phosphorylase complexed with aciclovir. **Biochem. Biophys. Res. Commun**, v. 308, n. 3, p. 553-559, 2003.

DOS SANTOS, D.M.; **Estudos estruturais da Purina Nucleosideo Fosforilase Humana**, UNESP, São Paulo, 2004. Dissertação de Mestrado.

DUCRUIX, A.; & GIEGÉ, R.; **Crystallization of nucleic acids and proteins. A practical approach**. Oxford, UK: IRL Press, 1992.

ERION, M.D.; STOECKLER, J.D.; GUIDA, W.C.; *et al.*; Purine nucleoside phosphorylase. 2. Catalytic mechanism. **Biochemistry.**, v. 36, n. 39, p. 11735-11748, 1997.

FISHER, R. A.; "On Mathematical Foundations of Theoretical Statistics", **Philosophical Transactions of the Royal Society of London**, série A 222, p. 309-368, 1921.

HERSHFIELD, M.S.; KREDICH, N.M.; S-Adenosylhomocysteine metabolism in adenosine deaminase deficient cells. **Adv. Exp. Med. Biol.**, v. 122A, p.421-425, 1980.

KABASCH, W. & SANDER, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. **Biopolymers**, v. 22, n. 12, p. 2577-2637, 1983.

KITTEL, C. **Introdução à Física do Estado Sólido**. 5. ed. Rio de Janeiro: Guanabara Dois, 1978.

LASKOWSKI, R. A., et al; PROCHECK – A program to check the stereochemical quality of protein structures, **J. Appl. Cryst.**, v. 26, p. 507, p. 283-291, 1993.

MCREE, D.E.; XtalView/Xfit - A versatile program for manipulating atomic coordinates and electron density. **J. Struct. Biol.** 125(2-3):156-65 (1999)

MORRIS, P.E. JR.; OMURA, G.A.; Inhibitors of the enzyme purine nucleoside phosphorylase as potential therapy for psoriasis. **Curr. Pharm. Des.**, v. 6, n. 9, p. 943-959, 2000. Review.

MURSHUDOV, G.N.; VAGIN, A.A.; DODSON, E.J.; Refinement of macromolecular structures by the maximum-likelihood method. **Acta Crystallogr. D Biol. Crystallogr.**, v. 53, n. Pt 3, p.240-255, 1997.

NAVAZA, J.; AMoRe: an automated package for molecular replacement, **Acta Crystallogr.**, v. 50, p. 157-163, 1994.

NOLASCO, D.O.; CANDURI, F.; PEREIRA, J.H.; *et al.*; Crystallographic structure of PNP from *Mycobacterium tuberculosis* at 1.9Å resolution. **Biochem. And Biophys. Res. Commun.**, n. 324, p. 789-794, 2004.

NOLASCO, D.O.; **Estrutura Cristalográfica da Purina Nucleosídeo Fosforilase do *Mycobacterium tuberculosis***, UNESP, SP, 2005. Tese de Mestrado.

NUNN, P.; WILLIAMS, B.; FLOYD, K.; DYE, C.; ELZINGA, G.; RAVIGLIONE, M.; Tuberculosis control in the era of HIV. **Nature Reviews – Immunology.**, n. 5, p. 819-826, 2005.

PROGRAMA DE CONTROLE DE TUBERCULOSE DO ESTADO DO RIO DE JANEIRO, boletim Informativo, Rio de Janeiro, 2005.

PUGMIRE, M.J.; EALICK, S.E.; Structural analyses reveal two distinct families of nucleoside phosphorylases. **Biochem. J.** 1;361(Pt 1):1-25 (2002) Review

RHODES, G. **Crystallography Made Crystal Clear**, p. 10-11, Academic Press, 1999.

SANTOS, D.M.; **Estudo Estrutural da Purina Nucleosídeo Fosforilase Humana**, Unesp, SP, 2004. Dissertação de mestrado.

SHI, W.; BASSO, L.A.; SANTOS, D.S.; *et al.*; Structures of Purine Nucleoside Phosphorylase from *Mycobacterium tuberculosis* in complexes with Immucillin-H and Its Pieces, **Biochemistry**, n. 40, p. 8204-8215, 2001.

SOCIEDADE BRASILEIRA DE FISIOLOGIA. Disponível em www.sbt.org.br. Acesso em junho de 2006.

VOET, D., VOET, J.G., PRAT, C. W. **Fundamentos de bioquímica**, Trad. Fett Neto, A. G. et al. Porto Alegre: Artes Médicas Sul, 2000.

WHO, (World Health Organization), Global Factsheet, 2005.

6. Apêndice

O Problema da Fase e a Substituição Molecular

Quando se deseja “resolver” uma estrutura utilizando raios X, procura-se descobrir como se encontra distribuída a densidade eletrônica na molécula. Para a obtenção da densidade eletrônica, tem-se a equação para o Fator de Estrutura que, multiplicada por $\exp(2\pi i \mathbf{r} \cdot \mathbf{S}) d\mathbf{v}$, onde $\mathbf{S}$ é igual a $\mathbf{s} - \mathbf{s}_0$ (direção da radiação espalhada menos o feixe incidente) e $\mathbf{r}$ é a posição da densidade da carga causadora do espalhamento, e integrando-se sobre todo o volume do espaço recíproco, encontra-se:

$$\rho(\mathbf{r}) = \int_{\text{Volume do espaço recíproco}} F(\mathbf{S}) \exp(-2\pi i \mathbf{r} \cdot \mathbf{S}) d\mathbf{v}_s$$

Dos experimentos de difração obtém-se um número finito de valores de $F(hkl)$ (DRENTH, 1994). A integral acima, portanto, deve ser repassada para uma somatória sobre todos os índices de Miller. Tem-se então:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} F(hkl) \exp i\alpha(hkl) \exp 2\pi i(hx + ky + lz)$$

onde $|F(hkl)|$ é a amplitude do fator de estrutura das reflexões (hkl) , x_j , y_j e z_j são as coordenadas fracionárias do jésimo átomo na cela unitária, h , k e l são os índices de reflexão, e $\alpha(hkl)$ é o ângulo da fase.

Nesta equação fica explícita a necessidade de se conhecer o módulo e a fase dos fatores de estrutura, mas pode-se obter somente o módulo a partir da difração de raios X.

Sendo $I(hkl) \propto |F(hkl)|^2$, as amplitudes $|F(hkl)|$ são facilmente encontradas. Porém, nenhuma informação é disponível sobre os ângulos das fases, tornando por isso a descoberta das fases o problema fundamental na determinação da estrutura de

qualquer. Durante um experimento de difração de raios X só se registram as intensidades e toda informação sobre a fase é perdida. Portanto, é impossível determinar a estrutura diretamente das medidas do padrão de difração, visto que parte da informação está perdida. O problema da fase é o problema básico de qualquer determinação de estrutura (DE AZEVEDO, 1997; BLUNDELL & JOHNSON, 1976; CANDURI, 2001, DOS SANTOS 2004).

Para ser resolvido o problema da fase se lança mão de alguns métodos indiretos. Atualmente existem quatro métodos principais para se determinar as fases dos fatores de estrutura: (i) o método de difração anômala com múltiplos comprimentos de onda, que depende da presença de átomos que espalhem forte e irregularmente na própria estrutura da proteína; (ii) métodos diretos, já usados para pequenas proteínas e em fase de desenvolvimento para proteínas maiores; (iii) a substituição isomorfa múltipla; (iv) o método de substituição molecular, que utiliza uma estrutura já resolvida, similar e/ou com enovelamento parecido já determinado.

Substituição Molecular

A partir da observação de que moléculas com seqüências primárias homólogas possuem enovelamentos similares, o conhecimento estrutural destas moléculas permite simular um modelo inicial do cristal a ser resolvido e, deste modelo, obter as informações necessárias sobre a fase da radiação difratada pelo cristal real (RHODES, 1999 *apud* DOS SANTOS 2004).

A substituição molecular é um dos métodos disponíveis para se resolver o problema da fase. Nesse método uma estrutura conhecida homóloga é ajustada na cela unitária da estrutura que se pretende resolver e as fases da estrutura conhecida são usadas como uma primeira aproximação para as fases na nova estrutura. O ajuste do

modelo se faz em uma procura em seis dimensões, ou graus de liberdade, que consistem basicamente em procuras de ordem rotacional e translacional (DE AZEVEDO, 1997).

Na substituição molecular deve-se usar um modelo de busca o mais homólogo possível da macromolécula a ser estudada, para proteínas uma identidade igual ao superior a 50% de homologia é suficiente. É amplamente recomendável o uso de modelos com altíssima identidade para que se possa ser evitado problemas de estimativa grosseira das fases e problemas durante a etapa de refinamento da estrutura estudada (NOLASCO, 2005).

O método de substituição molecular pode ser entendido a partir da relação da função de Patterson da proteína a ser resolvida, com a função de Patterson do modelo de procura, que é calculada a partir de dados de difração e é definida como:

$$P(u) = \int_{\text{Volume da cela unitária}} \rho(r) \rho(r+u) dv$$

sendo $u = (u \ v \ w)$ um vetor deslocamento.

A função de Patterson é um mapa vetorial de distâncias interatômicas da estrutura. A partir da equação para a densidade eletrônica para a coordenada recíproca h , obtém-se:

$$P(u) = \frac{1}{V^2} \sum_h \sum_{h'} F_h F_{h'} \exp(-2\pi i h' u) \int_{\text{Volume da cela unitária}} \exp[-2\pi i (h + h')] x dv$$

que pode ser simplificada para as três coordenadas hkl : como $P(u)$ é a transformada de Fourier de $F(hkl)^2$, esta pode ser calculada a partir dos dados de difração de raios X.

$$P(u) = \frac{2}{V} \sum_{hkl} F(hkl)^2 \cos 2\pi (hu + kv + lw)$$

O primeiro passo deve ser a colocação do modelo na mesma posição que a molécula real ocupa no seu retículo cristalino. Antes, para isto, os dados de difração

de raios X (**arq.card**) foram convertidos para um formato específico reconhecido pelo programa em um arquivo hkl (**arq.hkl**). Foi verificado também o volume da unidade assimétrica e definido uma faixa de resolução dos dados de difração de raios X a serem usados na SM. Na próxima etapa de Substituição Molecular, **tab**, o modelo de busca a ser utilizado no arquivo de entrada **tab.sh** foi organizado retirando-se o cabeçalho, ligantes e moléculas de água. Foi feito também, a partir do programa *reform* (CCP4, 1994), a mudança do fator de temperatura (fator B) dos modelos para 20 Å² e a inclusão dos parâmetros da cela unitária do modelo a ser resolvido. Os arquivos de saída gerados do procedimento acima foram search.tab (que possui os fatores de estrutura calculados (F_{calc}) do modelo de busca) e o arquivo de coordenadas atômicas (DOS SANTOS, 2004).

O **tab** calcula a distância máxima a partir do centro de massa r , com a dimensão ideal do comprimento das faces da cela de procura (considerando-a, portanto, uma cela cúbica), (DOS SANTOS, 2004).

Outro parâmetro a ser observado é a faixa de resolução a ser usada no cálculo dos mapas de Patterson. Reflexões de mais alta resolução diferem marcadamente, mesmo entre estruturas com alta homologia, visto que essas reflexões trazem informações sobre os detalhes estruturais. Reflexões de baixa resolução refletem aspectos relacionados com a estrutura secundária. Assim, pode-se inferir que nuances estruturais são dependentes da faixa de resolução que se pode utilizar tanto na substituição molecular quanto no refinamento, o que na prática implica em testar diversas faixas até que se encontre resultados satisfatórios (NOLASCO, 2005).

A partir destes e de outros parâmetros menos importantes o modelo pode passar por um passo posterior de refinamento baseado na correlação de Patterson (PC), ou coeficiente de correlação (CC). Neste momento o modelo é tratado como um corpo

rígido e sua orientação (rotação e translação) é definida de forma a minimizar a função

$F_{total}(C, C_i, t_i)$ definida como:

$$F_{total} = (1 - CC(C, C_i, t_i))$$

onde a CC é o coeficiente de correlação entre $|F_{obs}|^2$ e $|F_{model}|^2$ e é dada por:

$$CC(C, C_i, t_i) = \frac{\sum ab - (\sum a \sum b) / N}{\left\{ \left[\sum a^2 - (\sum a)^2 / N \right]^{1/2} \right\} \left\{ \left[\sum b^2 - (\sum b)^2 / N \right]^{1/2} \right\}}$$

onde $a = |F_{obs}|^2$, $b = |F_{model}(C, C_i, t_i)|^2$ e as somatórias são feitas sobre o conjunto de

reflexões observadas. E_{obs} representa os fatores de estrutura normalizados observados e

E_{model} os fatores de estrutura normalizados do modelo. Os fatores de estrutura

normalizados são dados pela seguinte equação:

$$F(H)^2 = \frac{|F(H)|^2}{\varepsilon \sum_{j=1}^N f_j^2}$$

onde H é um vetor do espaço recíproco, f_j os fatores de espalhamento atômico e ε é

um inteiro geralmente igual a 1, mas pode assumir outros valores para conjuntos

especiais de reflexões em certos grupos espaciais (STOUT & JENSEN, 1989 *apud* NOLASCO, 2005).

O coeficiente de correlação (CC) terá valores de -1 (correlação inversa perfeita),

passando por 0 (sem correlação) e com o máximo de 1 (correlação perfeita). Assim, a

minimização da função F_{total} leva à maximização da função CC , ou seja, os parâmetros

serão ajustados de forma que haja uma máxima correlação entre $|F_{obs}|^2$ e $|F_{model}|^2$.

A função rotação

Se um cristal possui um sistema de eixos ortogonais a, b e c, cada átomo de uma molécula pertencente a este cristal tem sua posição determinada por um conjunto de coordenadas fracionárias x, y e z. Pode-se calcular, então, os fatores de estrutura F (S) deste cristal, que descreve a posição dos átomos em coordenadas fracionárias, e a função de Patterson P(u).

Ao se aplicar uma rotação à molécula, de tal modo que:

$$(x, y, z).r = [C]. (x, y, z)$$

sendo (x, y, z)_r as novas coordenadas rotacionadas pela matriz de rotação C. A função de Patterson de um cristal, como um mapa vetorial das distâncias entre átomos deste cristal, rotacionará da mesma forma. Assim, aplicando-se a rotação C à função de Patterson P(u), obtém-se a função de Patterson rotacionada Pr (u).

Definindo uma função de *overlap* (sobreposição) R, para ângulos α, β e γ relacionados à matriz de rotação C, chega-se a:

$$R(\alpha, \beta, \gamma) = \int_U P(u) P_r(u_r) du$$

que terá seu valor maximizado quando houver uma correta sobreposição da molécula original e da molécula rotacionada.

A maximização da função R(α, β, γ), computada utilizando P(u), calculado dos dados de raios X, e Pr (u), e calculado do modelo rotacionado, fornece a rotação necessária para orientar o modelo a ser resolvido corretamente na unidade assimétrica do cristal.

Após encontrar o valor mais próximo do volume da unidade assimétrica da proteína a ser resolvida, utiliza-se este valor de **r** como o arquivo de entrada no próximo processo(**rot**). Coloca-se, portanto, como input no arquivo de entrada: arquivo **arq.hkl**

que possui os fatores de estrutura observados da molécula a ser resolvida (**F_{obs}**); arquivo **search.tab** que possui os fatores de estrutura calculados (**F_{calc}**) do PDB; o raio da esfera **r**; e novamente os valores dos parâmetros de cela da molécula a ser resolvida (DOS SANTOS, 2004).

Numa função de rotação só se pretende considerar os vetores intramoleculares, visto que todos os vetores no mapa de Patterson começam na origem e, por conseguinte, os vetores mais próximos da origem serão os intramoleculares. Escolhendo-se criteriosamente o raio máximo no espaço de Patterson, onde será feita a procura rotacional, as chances de se achar a rotação certa são aumentadas. O raio máximo no espaço de Patterson é um dos parâmetros de entrada para se efetuar a procura rotacional (BRÜNGER et al., 1992 *apud* NOLASCO, 2005).

A função translação

A melhor maneira de se encontrar a translação necessária ao correto posicionamento do modelo dentro da unidade assimétrica é por tentativa e erro. O modelo é transladado dentro da unidade assimétrica e os fatores de estrutura são calculados para esta nova posição (**F_{calc}**) e comparados com os fatores de estrutura observados (**F_{obs}**). A qualidade do processo de substituição molecular é monitorada através do cálculo de fatores de controle.

Os fatores usualmente utilizados são: o fator R (*Rfactor*) e o coeficiente de correlação (*CC*).

O fator R é dado pela expressão:

$$R = \frac{\sum_{hkl} \|F_{obs} - k|F_{mod} \|}{\sum_{hkl} F_{obs}}$$

sendo k um fator de escala .

Para concordâncias perfeitas, encontra-se valores de 0% para *Rfactor* e 100% para *CC*. Valores de *Rfactor* inferiores a ~40% e de *CC* superiores a ~50% indicam que a substituição molecular encontrou um modelo passível de refinamento. No procedimento para realizar a função translação, utiliza-se como arquivos de entrada novamente **arq.hkl,search.tab**, os parâmetros de cela e, desta vez, as coordenadas de rotação obtidas em *rot.log* (DOS SANTOS, 2004).

O processo acima dará as coordenadas rotacionadas e agora transladadas, sendo estas coordenadas transferidas para o arquivo de entrada de um processo de minimização por refinamento de corpo rígido (**rigid**). Esta rotina é um processo de refinamento das posições de grupos rígidos de átomos, que pode ser o de um monômero completo, contra as intensidades observadas, e dará o **arq.hkl** das coordenadas de rotação e translação e os fatores de concordância *CC* e *Rfactor* (DOS SANTOS, 2004).

No arquivo de entrada do **coord**, do último processo de SM, são incluídos as coordenadas de rotação e translação do resultado do *rigid* e também o arquivo de saída **model.pdb** do **tab**. Esta última rotina tem como saída o arquivo **imodel.pdb**, com as coordenadas do modelo a ser resolvido (DOS SANTOS, 2004).