

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

Efeito das mutações I 16T, I 21V, I 47T e S94A na
afinidade da enzima 2- *trans*-Enoil-ACP (CoA) Redutase de
Mycobacterium tuberculosis pelo cofator NADH: Estudos
por simulação pela dinâmica molecular e *docking* molecular

Evelyn Koeche Schroeder

Orientadores: Dr. Diógenes Santiago Santos

Dr. Luiz Augusto Basso

Co-orientador: Dr. Osmar Norberto de Souza

Tese submetida ao Programa de Pós-Graduação
em Biologia Celular e Molecular do Centro
Biotecnologia da UFRGS como requisito parcial
para a obtenção do grau de Doutor em Ciências.

Porto Alegre, Novembro de 2004

O presente trabalho foi realizado entre junho/1999 e junho/2004, no Centro de Biotecnologia da UFRGS e no Instituto de Informática da PUCRS, sob orientação dos Professores Dr. Diógenes Santiago Santos e Dr. Luiz Augusto Basso e co-orientação do Professor Dr. Osmar Norberto de Souza.

Evelyn Koeche Schroeder

Esta tese foi julgada adequada para a obtenção do título de DOUTOR EM CIÊNCIAS e aprovada em sua forma final pelos orientadores e pela Comissão Examinadora no Programa de Pós-Graduação em Biologia Celular e Molecular do Centro de Biotecnologia da UFRGS.

Orientadores:

Prof. Dr. Diógenes Santiago Santos
Centro de Pesquisa em Biologia Molecular e Funcional
Pontifícia Universidade Católica do Rio Grande do Sul

Prof. Dr. Luiz Augusto Basso
Instituto de Biociências
Universidade Federal do Rio Grande do Sul

Co-orientador:

Prof. Dr. Osmar Norberto de Souza
Instituto de Informática
Pontifícia Universidade Católica do Rio Grande do Sul

Comissão Examinadora:

Prof. Dr. Paulo Mascarello Bisch
Instituto de Biofísica Carlos Chagas Filho
Universidade Federal do Rio de Janeiro

Prof. Dr. Walter Filgueira de Azevedo Junior
Instituto de Biociências Letras e Ciências Exatas
Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP

Prof. Dr. Giancarlo Pasquali
Centro de Biotecnologia
Universidade Federal do Rio Grande do Sul

A meus pais
Edni e Edilza
e ao *Rodrigo*

“Discovery is not a result of logical thought,
even though the end result
is intimately bound to
the rules of logic.”

Albert Einstein

Agradecimentos

- Aos professores Dr. Diógenes Santiago Santos, Dr. Luiz Augusto Basso e Dr. Osmar Norberto de Souza pela orientação, incentivo e oportunidades de trabalho.
- Ao Prof. Diógenes, que desde o início apostou em mim e no projeto, disponibilizando a infra-estrutura necessária para a realização deste trabalho. Pelo entusiasmo e exemplo de gestão do grupo de pesquisa.
- Ao Prof. Basso, que sempre me estimulou, incentivou e principalmente compartilhou experiências em conversas e discussões sobre o trabalho e o meio científico.
- Ao Prof. Osmar, que inicialmente me acolheu no Laboratório Nacional de Luz Síncrotron para os primeiros passos na área de modelagem e simulação de biomoléculas, e depois me recebeu no Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas do PPGCC/PUCRS.
- À Prof. Dra. Fabiana Horn, que disponibilizou seu gabinete no Depto de Biofísica para o início deste trabalho.
- Aos professores Dr. Giancarlo Pasquali e Dr. Jorge Guimarães, que participaram de minha comissão de acompanhamento no PPGBCM - UFRGS.
- À colega Claudia Fernandes pelo trabalho em equipe e discussões no dia a dia do laboratório, mas principalmente pela amizade.
- Ao colega M.Sc. Marco Stelmar, pela importante ajuda na utilização dos clusters e na construção de programas para leitura dos dados gerados.

- Ao Dr. Luis Augusto Colembergue Kluszo, pela ajuda nas análises estatísticas dos resultados dos experimentos de *docking* molecular.
- Ao colega Hermes L. Neubauer de Amorim, que me incentivou a trabalhar na área de modelagem e simulação de biomoléculas e sempre esteve disposto a dividir experiências.
- Às famílias Koeche Schroeder e Belloc Soares pelo apoio incondicional.
- A todos os amigos que fiz durante esta jornada, agradeço através do Jaim, Luiz Pedro, Simone, Procópio, Fernanda, Lurdinha, Isabel, Claudia, Lala e Hugo.
- Aos colegas e amigos do Laboratório de Microbiologia Molecular do Centro de Biotecnologia da UFRGS, em especial à Clotilde Amorim Pinto.
- Aos amigos que fiz no Laboratório Nacional de Luz Sincrotron – LNLS, em Campinas, e à Valéria e Maurício, grandes amigos, que me receberam e apoiaram nos 3 meses que passei em Campinas.
- Aos colegas do Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas do PPGCC da PUCRS: Claudia, Lala, Marco, Andrei, Jorge, Camila, Bruna e Mariana. E aos amigos que fiz na PUCRS, entre eles o Marcon, a Ana Paula e o Gonzáles, e, em especial, ao Paulo Fernandes pelas conversas e desabafos.
- À direção e equipe do Programa de Pós-Graduação em Ciência da Computação e ao Centro de Processamento de Alto Desempenho do Instituto de Informática da PUCRS que me acolheram e permitiram a realização de grande parte deste trabalho.
- Ao Programa de Pós-Graduação em Biologia Celular e Molecular e ao Centro de Biotecnologia, em especial à Sílvia e ao Luciano.
- À CAPES, CNPq/MCT e FAPERGS, pela bolsa e auxílio financeiro a projetos dos laboratórios.

Resumo

O aumento do número de casos de tuberculose e o surgimento de cepas de *Mycobacterium tuberculosis* (MTB) resistentes a múltiplas drogas, entre elas a isoniazida (INH) representa um sério problema de saúde pública. A enzima InhA, ou 2-*trans*-Enoil-ACP (CoA) redutase de MTB, catalisa a redução NADH-dependente de ácidos graxos α,β -insaturados, precursores dos ácidos micólicos (importantes componentes do envelope celular do MTB). Mutações no gene estrutural da InhA estão associadas à resistência *in vivo* à INH devido a uma menor afinidade pela molécula de NADH, sugerindo que o mecanismo de resistência deva estar relacionado com interações específicas entre a enzima e o cofator. Para verificar os eventos moleculares associados à afinidade da enzima pelo ligante e identificar os aspectos moleculares relacionados com a resistência, foram realizados estudos de simulação por dinâmica molecular (DM) dos sistemas InhA-NADH (espécie selvagem *wt* e mutantes) completamente solvatados.

Todos os sistemas enzimáticos apresentaram grande flexibilidade durante as trajetórias por DM. Apesar da flexibilidade, no complexo *wt* InhA-NADH, a molécula de NADH permanece firmemente ligada ao sítio da enzima numa conformação estendida. Nos complexos mutantes I21V e I16T, onde as mutações ocorrem na alça rica em glicina, as interações entre enzima e cofator são menos efetivas, permitindo que a porção pirofosfato do NADH experimente importantes mudanças conformacionais e se afaste de sua região de ligação, indicando, provavelmente, um estágio inicial da dissociação do ligante. No mutante I47T, a substituição da Ile por Thr causa uma contração no sítio de ligação do NADH, que é refletida no rearranjo conformacional do NADH e na expulsão de moléculas de água importantes para a associação do cofator. O mutante S94A apresenta

comportamento muito semelhante à enzima selvagem, como esperado a partir dos experimentos cristalográficos.

As afinidades das enzimas pelo NADH foram avaliadas por experimentos de *docking* molecular, onde as estruturas instantâneas geradas durante as trajetórias da DM foram utilizadas como forma de considerar explicitamente a flexibilidade da macromolécula. Os resultados do *docking* molecular mostraram que todos os mutantes apresentam menor afinidade pelo NADH, exceto o mutante S94A, cuja energia livre de ligação do NADH foi estatisticamente igual à observada para a enzima selvagem.

Os resultados apresentados neste trabalho devem contribuir para o entendimento do mecanismo molecular específico de resistência à INH, que é crucial para o desenvolvimento de novos fármacos para o controle da tuberculose.

Abstract

The increasing prevalence of tuberculosis in many areas of the world, coupled with the rise in drug-resistant *Mycobacterium tuberculosis* (MTB) strains presents a major threat to global health. InhA, the enoyl-ACP reductase from MTB, catalyzes the NADH-dependent reduction of long chain *trans*-2-enoyl-ACP fatty acids, an intermediate in mycolic acids biosynthesis. Mutations in the structural gene for InhA are associated with isoniazid resistance *in vivo* due to a reduced affinity for NADH, suggesting that the mechanism of drug resistance may be related to specific interactions between enzyme and cofactor within the NADH binding site. In order to compare the molecular events underlying this ligand affinity in the wild type and S94A, I21V, I16T and I47T clinical mutant enzymes, and to identify the molecular aspects related to resistance, molecular dynamics (MD) simulations of fully solvated NADH-InhA (*wt* and mutants) systems were performed.

All InhA-NADH systems showed a large flexibility during the MD trajectories. Although very flexible, in the *wt* InhA-NADH complex, the NADH molecule keeps its extended conformation firmly bounded to the enzyme's binding site. In the I21V and I16T mutant complexes, where mutated residues were located in the glycine rich loop, interactions between enzyme and cofactor became more labile, and the NADH pyrophosphate moiety undergoes to considerable conformational changes, becoming more hydrated and moving apart from its binding site, probably indicating the initial phase of ligand expulsion. In the I47T mutant, the substitution of an Ile residue for Thr causes a binding site contraction with conformational changes of the NADH molecule and expulsion of water molecules important for cofactor binding to the enzyme. The S94A mutant showed to be very similar to the *wt* enzyme, in agreement to crystallographic experiments observations.

The enzyme-ligand affinities were evaluated by molecular docking experiments which were performed in the trajectory ensembles of MD snapshots as

a way to explicit consider the macromolecular flexibility. All mutant enzymes had lower affinities for the NADH molecule, except the S94A mutant, whose free energy of NADH binding was statistically similar to that of the wild type enzyme.

This results presented here should contribute to our understanding of specific molecular mechanisms of drug resistance, which is crucial for designing more potent antimycobacterial agents for controlling tuberculosis.

Lista de Figuras

Figura 1.1. Representação esquemática da parede celular de <i>Mycobacterium</i>	07
Figura 1.2. Componente peptidoglicano da parede celular de <i>Mycobacterium</i>	08
Figura 1.3. Estrutura da porção terminal do arabinogalactano.....	09
Figura 1.4. Esquema da pirólise de um ácido β -hidroxicarboxílico (ácido micólico), levando à separação da cadeia α e do meroaldeído.....	10
Figura 1.5. Estrutura dos ácidos micólicos identificados em <i>M. tuberculosis</i>	12
Figura 1.6. Biossíntese de ácidos graxos em <i>M. tuberculosis</i>	16
Figura 1.7. Algumas drogas que reconhecidamente inibem o sistema FAS II de <i>M. tuberculosis</i>	18
Figura 1.8. Ativação da pró-droga isoniazida e formação do aduto covalente com o NAD(H).....	21
Figura 1.9. Reação de redução de um tioéster α,β -insaturado catalisada pela InhA.....	24
Figura 1.10. Mecanismo químico da reação de redução catalisada pela enzima InhA.....	25
Figura 1.11. Estrutura da subunidade C da enzima InhA (PDB id: 1BVR).....	27
Figura 1.12. Homotetrâmero da InhA (PDB id: 1BVR).....	28
Figura 1.13. Relação entre as quatro subunidades do homotetrâmero da InhA.....	29
Figura 1.14. A cavidade de ligação da enzima InhA com seus ligantes NADH e C ₁₆	30
Figura 1.15. O substrato C ₁₆ e cofator NADH na conformação que assumem dentro do sítio ativo da enzima InhA (subunidade C de 1BVR).....	32
Figura 1.16. A predominância de aminoácidos não carregados no sítio de ligação do substrato C ₁₆ e cofator NADH.....	33
Figura 1.17. Posição da cadeia lateral da Phe149 dentro do sítio da InhA e o canal de água bloqueado por este resíduo.....	34
Figura 1.18. Moléculas de água no sítio ativo, próximas à ligação dupla do substrato.....	35
Figura 1.19. Sobreposição dos sítios de ligação do NADH da <i>wt</i> InhA e da S94A InhA.....	38
Figura 1.20. Comparação das interações da porção central do NADH com seu sítio de ligação nas estruturas da <i>wt</i> InhA 1ENY.....	39
Figura 1.21. Sobreposição dos sítios ativos da <i>wt</i> InhA associada ao NADH e associada ao complexo inibidor NAD-INH.....	40

Figura 1.22. Inibição da InhA por TCL e GEQ.....	41
Figura 3.1. Representação da variação da energia potencial para as interações de van der Waals e de Coulomb em função da distância interatômica r_{ij}	54
Figura 3.2. Representação das condições periódicas de contorno.....	57
Figura 3.3. Distâncias em Å entre os oxigênios centrais do grupo pirofosfato do NADH nos sítios ativos do homotetrâmero da InhA 1BVR.pdb.....	60
Figura 3.4. Seqüência da enzima InhA espécie selvagem (268 aminoácidos) utilizada na simulação por dinâmica molecular.....	63
Figura 3.5. Definição do sistema de simulação: complexo <i>wt</i> InhA-NADH e contra-íons Na^+ solvatados por moléculas de água em uma caixa retangular.....	64
Figura 3.6. Mutação do aminoácido serina pelo aminoácido alanina.....	67
Figura 3.7. Mutação do aminoácido isoleucina pelos aminoácidos valina e treonina para a construção dos mutantes I21V, I16T e I47T.....	68
Figura 3.8. Átomo teste (<i>probe atom</i>) passando pela malha tridimensional e registrando a energia potencial em cada ponto desta malha.....	79
Figura 4.1. Visualização aproximada do sítio de ligação do NADH na enzima InhA.....	92
Figura 4.2. Gráficos de energia da simulação do sistema <i>wt</i> InhA-NADH.....	94
Figura 4.3. Gráfico do RMSD do esqueleto polipeptídico em função do tempo para as trajetórias dos sistemas <i>wt</i> InhA-NADH e apo-InhA.....	97
Figura 4.4. Identificação das estruturas secundárias na estrutura da enzima InhA.....	98
Figura 4.5. Fatores de temperatura (<i>B-factor</i>) da cadeia principal dos resíduos da enzima InhA nas simulações <i>wt</i> InhA-NADH e apo InhA.....	99
Figura 4.6. Flexibilidade do sistema <i>wt</i> InhA-NADH ao longo da simulação por dinâmica molecular.....	100
Figura 4.7. Gráfico do raio de giro (Rg) para os últimos 2 ns das simulações <i>wt</i> InhA-NADH e apo-InhA.....	101
Figura 4.8. Flexibilidade da cadeia lateral da Lys165 nos últimos 2ns das simulações apo-InhA e <i>wt</i> InhA-NADH.....	103
Figura 4.9. Flexibilidade das cadeias laterais da Tyr158 e Phe149 nos últimos 2 ns das simulações apo-InhA e <i>wt</i> InhA-NADH.....	104
Figura 4.10. Movimentação da cadeia lateral da Tyr158 durante a simulação do complexo ternário formado entre a <i>wt</i> InhA, o NAD^+ e o TCL, no estudo do modo de ligação do TCL na Enoil Redutase de <i>M. tuberculosis</i>	105
Figura 4.11. Estrutura média do sítio de ligação do NADH nos últimos 2 ns de simulação do complexo <i>wt</i> InhA-NADH.....	108
Figura 4.12. Moléculas de água conservadas durante a trajetória do sistema <i>wt</i> InhA-NADH no sítio de ligação do NADH.....	112

Figura 4.13. Gráficos do RMSD do esqueleto polipeptídico em função do tempo para as trajetórias dos sistemas mutantes S94A-NADH, I21V-NADH, I16T-NADH e I47T-NADH, com a identificação da mutação <i>on the fly</i> em 1,9 ns, e comparação com o sistema <i>wt</i> InhA-NADH.....	116
Figura 4.14. Fatores de temperatura da cadeia principal dos resíduos da enzima InhA nas simulações do sistema <i>wt</i> InhA-NADH e mutantes S94A-NADH, I21V-NADH, I16T-NADH e I47T-NADH	117
Figura 4.15. Comparação da abertura da alça de ligação do substrato (resíduos 196-219) em relação às alças A e B nas simulações do sistema <i>wt</i> InhA-NADH e mutantes S94A-NADH, I21V-NADH, I16T-NADH e I47T-NADH.....	118
Figura 4.16. Sítio de ligação do NADH no mutante I21V InhA. Identificação do aminoácido mutado Val21 e dos aminoácidos e moléculas de água importantes para a associação do cofator	120
Figura 4.17. Sítio de ligação do NADH no mutante I16T InhA. Identificação do aminoácido mutado Thr16 e dos aminoácidos e moléculas de água importantes para a associação do cofator	123
Figura 4.18. Comparação da conformação do NADH na estrutura média da fase de produção do sistema <i>wt</i> InhA-NADH com suas estruturas médias nos mutantes I21V-NADH e I16T-NADH	124
Figura 4.19. Comparação dos ângulos diedros α_A e γ_A do NADH na fase de produção das trajetórias do mutante I21V-NADH e do sistema <i>wt</i> InhA-NADH.....	124
Figura 4.20. Comparação da distribuição das moléculas de água a 5 Å do NADH nos sistemas <i>wt</i> InhA-NADH e mutante I21V-NADH.....	125
Figura 4.21. Sítio de ligação do NADH no mutante I47T InhA e comparação da conformação do NADH na estrutura média da fase de produção do sistema <i>wt</i> InhA-NADH e do mutante I47T-NADH.....	128
Figura 4.22. Comparação da conformação do NADH e distribuição das moléculas de água a 5 Å do cofator nos sistemas <i>wt</i> InhA-NADH e mutante I47T-NADH.....	129
Figura 4.23. Tamanho do NADH nas simulações do sistema <i>wt</i> InhA-NADH e mutantes S94A-NADH, I21V-NADH, I16T-NADH e I47T-NADH.....	130
Figura 4.24. Comparação do número de pontes de hidrogênio formadas entre o NADH e seu sítio de ligação dos sistemas <i>wt</i> InhA-NADH e mutantes S94A-NADH, I21V-NADH, I16T-NADH e I47T-NADH	131
Figura 4.25. Sítio de ligação do NADH no mutante S94A InhA e comparação da conformação do NADH na estrutura média da fase de produção do sistema <i>wt</i> InhA-NADH e do mutante S94A-NADH.....	133
Figura 4.26. Moléculas de água cristalográficas não reproduzidas durante a simulação por DM do sistema <i>wt</i> InhA-NADH.....	137
Figura 4.27. Gráfico do RMSD do esqueleto polipeptídico em função do tempo para as trajetórias do sistema cWAT-NADH em relação à estrutura cristalográfica, e comparação com o sistema <i>wt</i> InhA-NADH.....	139

Figura 4.28. Gráficos dos raios de giro e distâncias entre centros de massa da enzima e do NADH nos sistemas <i>wt</i> InhA-NADH (azul escuro) e cWAT-NADH.....	140
Figura 4.29. Distância entre a C=O da Gly14 e o OH da Ser94 e a molécula de água wat276 durante a simulação por DM do sistema cWAT-NADH.....	143
Figura 4.30. Moléculas de água cristalográficas reproduzidas durante a simulação por DM do sistema cWAT-NADH.....	144
Figura 4.31. Comparação do <i>docking</i> molecular do NADH na estrutura cristalina da InhA espécie selvagem e do mutante S94A.....	147
Figura 4.32. <i>Docking</i> do NADH na estrutura cristalina da InhA espécie selvagem com consideração das moléculas de águas cristalográficas.....	149
Figura 4.33. Avaliação de energia livre de ligação (FEB) calculadas por experimentos de <i>docking</i> molecular via <i>simulated annealing</i> do NADH na trajetória dos sistemas da InhA.....	152
Figura 4.34. Distribuições da energia livre de ligação (FEB) calculadas por experimentos de <i>docking</i> molecular via <i>simulated annealing</i> do NADH na trajetória dos sistemas <i>wt</i> InhA e cWAT-InhA.....	154
Figura 4.35. Avaliação da energia livre de ligação (FEB) do NADH calculados por experimentos de <i>docking</i> molecular nos sistemas da InhA via algoritmos genéticos.....	156

Lista de Tabelas

Tabela 1.1. Processos celulares e alvos moleculares inibidos pelos principais fármacos utilizados no tratamento da tuberculose.....	06
Tabela 1.2. Parâmetros de cinética enzimática da enzima InhA de <i>Mycobacterium tuberculosis</i> espécie selvagem e mutantes.....	22
Tabela 1.3. Estruturas tridimensionais da enzima InhA de <i>M. tuberculosis</i> depositadas no Swiss Protein Data Bank.....	26
Tabela 1.4. Pontes de hidrogênio do NADH em seu sítio de ligação na InhA.....	36
Tabela 1.5. Pontes de hidrogênio entre moléculas de água bem ordenadas e o NADH em seu sítio de ligação na InhA.....	37
Tabela 3.1. Composição do sistema de simulação do complexo da enzima InhA espécie selvagem com o NADH (<i>wt</i> InhA-NADH).....	64
Tabela 3.2. Composição do sistema de simulação do complexo do mutante S94A da enzima InhA espécie selvagem com o NADH (S94A-NADH).....	67
Tabela 3.3. Composição dos sistemas InhA-NADH das simulações por dinâmica molecular.....	69
Tabela 3.4. Torções possíveis identificadas pelo programa AutoTors na molécula de NADH.....	82
Tabela 3.5. Comparação dos algoritmos de busca e da flexibilidade do ligante nos experimentos de <i>docking</i> do NADH no sítio de ligação da enzima InhA.....	84
Tabela 3.6. Resumo dos experimentos de <i>docking</i> do NADH nas trajetórias das simulações por DM via <i>simulated annealing</i>	89
Tabela 3.7. Resumo dos experimentos de <i>docking</i> do NADH nas trajetórias das simulações por DM via algoritmos genéticos.....	90
Tabela 4.1. Resumo geral dos parâmetros estruturais da InhA observados nos últimos 2 ns das simulações por dinâmica molecular.....	96
Tabela 4.2. Resíduos de aminoácidos implicados no mecanismo catalítico da enzima InhA de MTB.....	102
Tabela 4.3. Pontes de hidrogênio formadas entre o NADH e aminoácidos do sítio de ligação da <i>wt</i> InhA.....	108
Tabela 4.4. Pontes de hidrogênio formadas entre o NADH e moléculas de água que medeiam a interação com o sítio de ligação da <i>wt</i> InhA com ocupância maior do que 50% durante os últimos 2 ns de simulação.....	111

Tabela 4.5. Número de moléculas de água na primeira esfera de solvatação da molécula do NADH calculado sobre os últimos 2 ns de simulação dos sistemas InhA-NADH.	118
Tabela 4.6. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante I21V-InhA.	119
Tabela 4.7. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante I16T-InhA.	122
Tabela 4.8. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante I47T-InhA.	127
Tabela 4.9. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante S94A-InhA.	132
Tabela 4.10. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do sistema cWAT-InhA.	141
Tabela 4.11. Comparação das energias livre de ligação (FEB) do NADH nas estruturas cristalinas da InhA espécie selvagem, do mutante S94A e dos modelos dos mutantes I21V, I16T e I47T calculadas a partir de experimentos de <i>docking</i> com o protocolo de <i>summulated annealing</i> .	148
Tabela 4.12. Valores médios e variância da energia livre de ligação (FEB) do NADH na trajetória dos sistemas da InhA, calculados por experimentos de <i>docking</i> molecular via <i>simulated annealing</i> considerando o ligante rígido na forma estendida.	153
Tabela 4.13. Valores de energia livre de ligação (FEB) do NADH na trajetória dos sistemas da InhA, calculados por experimentos de <i>docking</i> molecular via algoritmos genéticos, considerando a flexibilidade do ligante nas ligações do pirofosfato.	156

Lista de Abreviações

ACP	<i>acyl carrier protein</i> , ou proteína carreadora de acilas
AcpM	ACP característico de <i>Mycobacterium tuberculosis</i>
ADPR	<i>adenosine diphosphate ribosyl group</i> , ou ribosil-adenosina difosfato
anel A	anel adenina do NADH
anel N	anel nicotinamida do NADH
1BVR-C	subunidade C da estrutura tridimensional da enzima InhA, depositada no PDB sob o código: 1BVR
BAAR	bacilo álcool-ácido resistente
C₁₆	2- <i>trans</i> -hexadodecenoil tioéster de N-acetilcisteamina: tioéster de cadeia longa utilizado como um análogo ao substrato natural na cristalização da enzima InhA
CER	cerulenina
CoA	coenzima A
CMN	<i>Corynebacterium-Mycobacterium-Nocardia</i> , subgrupo de eubactérias que tem como característica a produção de estruturas únicas de parede celular de quimiotipo IV
C_n	cadeia de n carbonos
Cn	átomo de carbono n
3D	tridimensional
DE	β-hidroxiacil desidratase
DM	dinâmica molecular
DSSP	<i>dictionary of protein secondary structure</i>
EMB	etambutol
ENR	enoil redutase
EP	<i>evolutionary algorithm</i> , ou algoritmos evolucionários
ETH	etionamida
FabI	enoil redutase de <i>E. coli</i>
FAS	<i>fatty acid synthase</i> , ou sistema de síntese de ácidos graxos
FAS I	<i>fatty acid synthase type I system</i> , ou sistema de síntese de ácidos graxos do tipo I
FAS II	<i>fatty acid synthase type II system</i> , ou sistema de síntese de ácidos graxos do tipo II
GA	algoritmos genéticos

GEQ	5-[[4-(9H-fluoren-9-il)piperazin-1-il]carbonil]-1H-indol, um inibidor aril-indol-piperazina
I16T InhA	mutante Ile16Thr da enzima InhA
I21V InhA	mutante Ile21Val da enzima InhA
I47T InhA	mutante Ile47Thr da enzima InhA
INH	isoniazida
InhA	enzima 2- <i>trans</i> -Enoil-ACP (CoA) Redutase de <i>Mycobacterium tuberculosis</i>
inhA	gene que codifica a enzima InhA
KAR	β -cetoacil redutase
KAS	β -cetoacil sintase
KasA/B	β -cetoacil-ACP sintases de <i>Mycobacterium tuberculosis</i>
KatG	enzima catalase-peroxidase de <i>Mycobacterium tuberculosis</i> responsável pela ativação da pró-droga isoniazida em um intermediário eletrofílico (droga ativada) que se liga ao NADH no sítio ativo da InhA para inativá-la.
LAM	lipoarabinomanano, um dos componentes da parede celular do <i>M. tuberculosis</i>
LGA	<i>Lamarckian genetic algorithm</i>
MabA	β -cetoacil-ACP redutase de <i>Mycobacterium tuberculosis</i>
mAGP	<i>mycolyl-arabinogalactan-peptidoglycan complex</i> , ou complexo micolil-arabinogalactano-peptideoglicano
MDR	<i>multi-drug resistant</i> , resistente a múltiplas drogas
MM	mecânica molecular
MTB	<i>Mycobacterium tuberculosis</i>
ns	nanossegundos: 10^{-9} s
NAD⁺	nicotinamida adenina dinucleotídeo, forma oxidada
NADH	nicotinamida adenina dinucleotídeo, forma reduzida
NAD-INH	complexo covalente formado entre a isoniazida ativada e o NAD(H), que se liga de forma não covalente à enzima InhA inibindo sua atividade
ORF	<i>open reading frame</i> ou região de leitura aberta
ps	picossegundos: 10^{-12} s
PAS	ácido <i>p</i> -aminosalicílico
PDB	<i>Protein Data Bank</i> ou banco de dados de estrutura tridimensional de proteínas
PDB id	Código de identificação das coordenadas tridimensionais no PDB
PZA	pirazinamida
ribose A	anel ribose próximo à adenina do NADH
ribose N	anel ribose próximo à nicotinamida do NADH
RIF	rifampicina
RMSD	<i>root mean square deviation</i> , ou desvio médio quadrático
SA	<i>simulated annealing</i>

S94A InhA	mutante Ser94Ala da enzima InhA
SDR	<i>short chain dehydrogenase/reductase</i> , família de enzimas desidrogenase/redutase
SM	estreptomicina
TCL	triclosan, ou 2,4,4'-triclora-2'hidroxifenil éter
TB	tuberculose (doença)
TLM	tiolactomicina
wt	<i>wild type</i> , ou espécie selvagem
ΔG_{bind}	energia livre de ligação

Índice

Resumo.....	iii
Abstract.....	v
Lista de Figuras.....	vii
Lista de Tabelas.....	xi
Lista de Abreviações.....	xiii
1 - Introdução.....	01
1.1 - A tuberculose e seu agente etiológico, o <i>Mycobacterium tuberculosis</i>	01
1.2 – A parede celular do <i>M. tuberculosis</i> e os ácidos micólicos.....	07
1.3 – A biossíntese de ácidos graxos em <i>M. tuberculosis</i> : Os sistemas FAS I e FAS II	14
1.4 – Inibição do sistema FAS II pela Isoniazida (INH) e identificação de seu alvo molecular.....	19
1.5 – A enzima InhA.....	24
2 - Objetivos.....	43
3 - Materiais e Métodos.....	44
3.1 - Simulação por Dinâmica Molecular de sistemas biológicos.....	44
3.1.1 – Fundamentos da Dinâmica Molecular.....	45
3.1.2 – O campo de força.....	50
3.1.3 – Condições de contorno.....	56
3.1.4 – Organização de uma simulação por Dinâmica Molecular.....	58
3.2 - Simulação por Dinâmica Molecular da enzima InhA de <i>M. tuberculosis</i>	60
3.2.1 – Simulação do complexo <i>wt</i> InhA-NADH.....	63
3.2.2 – Preparação e simulação dos mutantes S94A, I16T, I21V e I47T.....	66
3.2.2.1 - Mutante S94A InhA.....	66
3.2.2.2 - Mutantes I16T, I21V e I47T InhA.....	67
3.3 - O <i>Docking</i> Molecular.....	70
3.3.1 - O programa Autodock3.0.....	76

3.3.1.1 – O módulo AutoGrid.....	78
3.3.1.2 – O módulo AutoDock.....	80
3.4 – <i>Docking</i> do NADH nas estruturas da enzima InhA espécie selvagem e mutantes S94A, I16T, I21V e I47T.....	82
3.4.1 - <i>Docking</i> do NADH na estrutura cristalina da enzima InhA (1ENY).....	85
3.4.2 - <i>Docking</i> do NADH na estrutura cristalina no mutante S94A-InhA (1ENZ).....	86
3.4.3 - Consideração das águas cristalográficas no <i>docking</i> do NADH na estrutura cristalina da enzima InhA (1ENY).....	87
3.4.4 - <i>Docking</i> do NADH nas estruturas dos mutantes S94A, I16T, I21V e I47T da enzima InhA.....	87
3.5 – <i>Docking</i> do NADH nas trajetórias da enzima InhA espécie selvagem e mutantes S94A, I16T, I21V e I47T.....	88
3.5.1 - <i>Docking</i> do NADH nas trajetórias das simulações por DM via <i>simulated annealing</i>	89
3.5.2 - <i>Docking</i> do NADH nas trajetórias das simulações por DM via algoritmos genéticos.....	90
4 - Resultados e Discussão.....	91
4.1 – Avaliação da convergência e estabilidade das simulações por dinâmica molecular dos sistemas da InhA (espécie selvagem e mutantes).....	94
4.2 – Simulação do sistema <i>wt</i> InhA-NADH e da apo-InhA.....	97
4.2.1 - Comportamento dinâmico dos resíduos conservados Lys165, Tyr158 e Phe149 no sítio de ligação do cofator.....	102
4.2.2 – Ligação do NADH no sítio ativo da enzima <i>wt</i> InhA.....	106
4.3 – Simulação por DM dos mutantes S94A, I21V, I16T e I47T da enzima InhA complexados com o NADH.....	115
4.4 – Análise da importância da consideração das águas cristalográficas na simulação do sistema <i>wt</i> InhA-NADH.....	137
4.5 – Estudos de <i>docking</i> molecular para a avaliação da afinidade da enzima InhA pelo NADH.....	145
4.6 – Utilização explícita da flexibilidade da enzima InhA nos experimentos de <i>docking</i>	150

4.6.1 – Consideração da flexibilidade da enzima InhA e do ligante em experimentos de <i>docking</i>	155
5 - Considerações Finais e Perspectivas	158
6- Referências Bibliográficas.....	162
7 - Apêndices.....	175

Anexos

- 1 - **Schroeder, E.K.; de Souza, O.N.; Santos, D.S.; Blanchard, J.S. & Basso, L.A., 2002.** Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. *Curr. Pharm. Biotech.*, 3, 197-225.
- 2 - **Schroeder, E.K.; Basso, L.A.; Santos, D.S. & de Souza, O.N., 2005.** Molecular Dynamics Simulation Studies of the Wild-Type, I21V, and I16T Mutants of Isoniazid-Resistant *Mycobacterium tuberculosis* Enoyl Reductase (InhA) in Complex with NADH: Toward the Understanding of NADH-InhA Different Affinities. *Biophys. J.* , 89, 876-884.

Introdução

1.1- A tuberculose e seu agente etiológico, o *Mycobacterium tuberculosis*

As doenças causadas por infecções por microrganismos, dentre elas a tuberculose (TB), têm sido alvo de preocupações nos últimos 25 anos. Apesar do desenvolvimento de agentes quimioterápicos eficientes, esta doença continua levando à morte pessoas em diversas partes do mundo. Em adição à quimioterapia, a vacinação profilática com BCG (*Bacillus Calmette-Guerin*) contra infecções micobacterianas resulta na proteção de algumas comunidades, mas não apresenta eficiência comparável em outras. De acordo com a Organização Mundial da Saúde, aproximadamente um terço da população mundial está infectada, e a estimativa é que, desta fração, 10% desenvolva a doença ao longo de sua vida (WHO, 2004a; Bloom & Murray, 1992). Anualmente, são mais de 8 milhões de pessoas que desenvolvem a tuberculose ativa, resultando numa taxa de 2 milhões de mortes/ano (WHO, 2004a; WHO, 2004b). A cada dia, mais de 25 mil pessoas ficam doentes e 5 mil pessoas acabam morrendo (WHO, 2003). No Brasil, de acordo com o relatório da Funasa envolvendo os anos de 1980 a 2000, a tuberculose gira em torno de 80 a 100 mil casos/ano, resultando em uma taxa de óbitos em torno de 6 mil mortes/ano (FUNASA, 2002).

Apesar da TB ser considerada por muitos como uma doença do século XIX, na verdade ela nunca foi erradicada. A epidemia de TB é muito mais preocupante hoje em dia do que em toda a história da humanidade. Nos últimos anos, tem-se verificado um aumento em sua incidência: entre 1997 e 2000, o número total de novos casos de TB aumentou 1,8% ao ano (Corbett *et al.*, 2003). Em termos mundiais, a TB é atualmente responsável pelo segundo maior número de mortes de jovens e adultos entre as doenças infecciosas. Estima-se que, entre 2002 e 2020,

se não houver um esforço para seu controle, aproximadamente 1 bilhão de pessoas serão infectadas, mais de 150 milhões ficarão doentes, e a tuberculose levará à morte mais de 38 milhões de pessoas (WHO, 2001; WHO 2002). Dada a gravidade da situação global da doença, em 1993, a Organização Mundial da Saúde declarou a TB um sério problema de saúde pública mundial, como forma de alertar os setores de saúde pública e os governos para a necessidade de contenção desta doença.

A TB é uma doença infecciosa crônica causada por micobactérias do complexo *Mycobacterium tuberculosis*, que compreende o *Mycobacterium bovis*, o *Mycobacterium africanum*, o *Mycobacterium microti*, mas, principalmente, o *Mycobacterium tuberculosis* (MTB).

O *Mycobacterium tuberculosis* é um bacilo pequeno (1-4 µm de comprimento e 0,3-0,6 µm de diâmetro), sem flagelos, com uma parede celular rica em lipídeos, que facilmente retém a coloração de carbol-fucsina, até mesmo na presença de álcool acidificado (bacilo álcool-ácido resistente - BAAR). É um bacilo estritamente aeróbio, capaz de sobreviver em micro-aerofilia, com tempo de crescimento muito lento (de aproximadamente 24 horas), que coloniza preferencialmente as cavidades das paredes pulmonares, onde a pressão de oxigênio é alta (pO₂ 100 – 140 mm Hg) e o meio levemente alcalino (pH 7,4) (Grosset, 1980). Como é capaz de resistir ao ataque, sobreviver e se multiplicar no interior de células fagocitárias, como macrófagos, o MTB é considerado um parasita intracelular (Clemmens, 1996).

A infecção é normalmente transmitida de pessoa para pessoa pela inalação de pequenas partículas de aerossóis (diâmetro de 1-5 µm, contendo de 1-10 bacilos) expelidas por pacientes com a forma pulmonar da doença através da fala, espirro e, principalmente, pela tosse. Estas partículas são inaladas pelo indivíduo sadio, atingindo o espaço alveolar, onde iniciam sua multiplicação e são fagocitadas pelos macrófagos do hospedeiro (Hiriyanna & Ramakrishnan, 1986). A infecção primária por MTB envolve a replicação do organismo no foco pulmonar inicial de infecção, mas pode se espalhar para outros locais dentro do pulmão e, eventualmente, disseminar-se para outras partes do corpo do hospedeiro (Russell, 2001).

Apesar do sucesso inicial do bacilo em parasitar o hospedeiro, a infecção primária é quase sempre assintomática em adultos. Com o ataque do sistema

imunológico do hospedeiro, a replicação do bacilo ativo e sua disseminação são controladas. Entretanto, dificilmente ele é erradicado (Bloom & Murray, 1992; Glickman & Jacobs, 2001). O acúmulo de monócitos, neutrófilos e macrófagos no local da infecção gera uma estrutura granulomatosa compacta, conhecida como tubérculo.

A estrutura e a evolução das lesões causadas pelo bacilo do tubérculo dependem de múltiplos fatores relacionados com a ocorrência da infecção (primoinfecção ou reinfecção), virulência do bacilo, sistema de defesa específico do hospedeiro, resposta imune e com fatores genéticos e ambientais (Rhee *et al.*, 1999; FUNASA, 2001).

Aproximadamente 30% dos indivíduos expostos são infectados. Destes, 60-90% possuem uma resposta imune efetiva, permitindo a devida contenção da infecção. À medida que células T, monócitos e macrófagos são recrutados para a área em torno do bacilo, ele diminui sua replicação e espera pacientemente, não sendo capaz de transmitir a infecção (Russell, 2001). O MTB normalmente sobrevive dentro dos tecidos durante anos, neste estado de dormência. Eventualmente, aproveitando um momento de deficiência do sistema imunológico, o bacilo dormente é reativado, causando o ressurgimento da doença, normalmente muitas décadas após a infecção inicial (Manabe & Bishai, 2000). A existência destas formas dormentes dificulta o controle da doença, provavelmente devido à baixa atividade bactericida das drogas anti-TB contra estes bacilos em estado dormente ou semi-dormente.

Por causa dessa persistência característica, na era pré-antibióticos, a evolução clínica da infecção era extremamente variável entre os indivíduos. Enquanto alguns sucumbiam rapidamente à progressão da infecção primária, outros eram capazes de conter a doença (embora temporariamente). Mas a reativação da TB após um grande período de incubação era relativamente comum.

A introdução da terapia com antibióticos reduziu substancialmente o risco de reativação, confirmando a incapacidade de vários pacientes em eliminar o MTB dos tecidos sem uma intervenção terapêutica. A descoberta da estreptomicina (SM) em 1944, primeira droga utilizada no tratamento da TB, baixou a taxa de reativação para 20% - um avanço substancial, visto que antes 50-80% dos casos de TB resultavam em óbito alguns anos após o diagnóstico, mas ainda longe do ideal

(Gomez & McKinney, 2004). Entretanto, a monoterapia com estreptomicina muitas vezes levava à resistência bacteriana e ao fracasso do tratamento. Antevendo essas dificuldades, Paul Ehrlich já havia sugerido, em 1913, que a quimioterapia contra um patógeno deveria ser feita não com uma, mas duas ou três drogas diferentes (McKinney, 2000). Isto se tornou possível a partir do fim dos anos 50, com a utilização de isoniazida (INH), da rifampicina (RIF) e da pirazinamida (PZA) em uma terapia combinada, reduzindo, assim, a taxa de reincidência para valores mais aceitáveis.

Uma quimioterapia efetiva contra a TB deve incluir, inicialmente, uma ação bactericida contra o bacilo ativo, seguida da esterilização da população de bacilos dormentes ou semi-dormentes (Mitchison, 1985). Os agentes terapêuticos de primeira linha utilizados em seu tratamento são: INH, RIF, SM, etambutol (EMB) e PZA. Os três primeiros agentes apresentam grande atividade bactericida contra bacilos metabolicamente ativos, enquanto o último é ativo contra bacilos semi-dormentes em ambientes intracelulares ácidos. Os fármacos de segunda linha etionamida (ETH), cicloserina, e o ácido *p*-aminosalicílico (PAS), todos bacteriostáticos, só são utilizados no tratamento da TB resistente aos fármacos de primeira linha, ou quando estes são desaconselháveis por causa de seus efeitos colaterais desagradáveis (Scior *et al.*, 2002). As fluorquinolonas, apesar de serem bacteriostáticas contra o MTB, não são utilizadas como agente terapêutico de primeira linha, mas são importantes no tratamento dos organismos resistentes.

O uso racional destes fármacos depende da sensibilidade do bacilo (Ginsburg *et al.*, 2003), mas o esquema de tratamento mais prescrito, e recomendado pela Organização Mundial da Saúde, deve ser de, no mínimo, seis meses e envolve a combinação de quatro antibióticos: INH, RIF, PZA, e EMB ou SM, nos dois primeiros meses, seguido da combinação de INH e RIF por mais quatro meses (ou de INH combinada com ETH ou EMB por mais seis meses).

Até o início dos anos 1980, o tratamento com esses quimioterápicos facilitou o controle e o declínio dos números da doença. Entretanto, o número de casos começou a crescer a partir da segunda metade da década de 80. Com o surgimento da infecção por HIV, o número de pessoas co-infectadas com HIV e TB tornou-se um grande problema, à medida que houve um aumento do número de hospedeiros com infecção ativa capazes de transmitir a infecção. Além da AIDS,

vários outros fatores inter-relacionados têm contribuído para a re-emergência da TB como: (i) a concentração de pessoas nos grandes centros urbanos; (ii) o aumento da pobreza da população; (iii) o declínio das condições de saúde pública; e (iv) o abandono do tratamento.

Outro fator que tem contribuído para o aumento do número de casos de TB, e também para o aumento da taxa de mortalidade, é a emergência de cepas de MTB resistentes a múltiplas drogas (MDR-MTB)¹. Apesar da resistência a drogas ocorrer naturalmente numa frequência de 1×10^{-6} dentro de uma população de bacilos, regimes de tratamento inapropriados e abandono do tratamento têm uma grande influência na emergência de cepas multi-resistentes (Basso & Blanchard, 1998). Isso leva à seleção de bacilos de MTB com alterações em genes cromossomais que codificam proteínas-alvo dos fármacos de primeira linha (Tabela 1.1).

Portanto, torna-se urgente o desenvolvimento de novos fármacos anti-TB mais efetivos - tanto contra os bacilos sensíveis, quanto contra os resistentes - que tenham seus efeitos colaterais diminuídos e resultem em uma redução no tempo necessário para o tratamento.

Uma abordagem é a investigação de processos celulares particulares do *Mycobacterium* como alvo para intervenção, a exemplo da maioria dos fármacos utilizados no tratamento da TB, que interferem na síntese e na formação de componentes únicos da parede celular do MTB.

¹ Cepas MDR (*multi drug resistant*) são definidas como cepas resistentes a no mínimo isoniazida e rifampicina, os dois fármacos de primeira linha mais utilizados nos esquemas de tratamento da tuberculose.

Tabela 1.1. Processos celulares e alvos moleculares inibidos pelos principais fármacos utilizados no tratamento da tuberculose (McKinney, 2000).

Fármaco	Processo celular alvo	Alvo molecular (<i>gene</i>)
Isoniazida	Síntese de ácidos micólicos – parede celular	Enoil-ACP redutase (<i>inhA</i>)
Etionamida	Síntese de ácidos micólicos – parede celular	Enoil-ACP redutase (<i>inhA</i>)
Pirazinamida ^a	Transporte através da membrana celular	Indefinido
Etambutol	Síntese de arabinogalactano – parede celular	Arabinosil transferase (<i>embAB</i>)
Cicloserina	Síntese de peptídeoglicano – parede celular	D-alanil, D-alanina ligase (<i>ddlA</i>) e Alanina racemase (<i>alr</i>)
Rifampicina	Síntese de RNA	RNA polimerase – subunidade β (<i>rpoB</i>)
Estreptomicina	Síntese de polipeptídeos	16S rRNA (<i>rrs</i>) e proteína ribossomal 12S (<i>rpsL</i>)
Fluorquinolonas	Síntese de DNA	DNA girase (<i>gyrAB</i>)
PAS (ácido <i>p</i> -aminosalicílico)	Síntese de ácido fólico	Diidropteroato sintetase (<i>folP</i>)

^a Inicialmente a ação da pirazinamida era atribuída à biossíntese de ácidos graxos (sistema FAS-I) mas, recentemente, observou-se que esta droga está relacionada com alterações no potencial de membrana, afetando o transporte através da membrana (Zhang *et al.*, 2003).

1.2- A parede celular do *M. tuberculosis* e os ácidos micólicos

O envelope celular do MTB possui uma arquitetura e uma composição diferente das demais bactérias, e é reconhecido por sua baixa permeabilidade. O gênero *Mycobacterium* pertence ao subgrupo *Corynebacterium-Mycobacterium-Nocardia* (CMN) de eubactérias actinomicetos (bactérias Gram-positivas com DNA rico em C e G), que tem como característica a produção de estruturas únicas de parede celular, chamada de parede celular de quimiotipo IV.

A parede celular do MTB é predominantemente formada por três macromoléculas covalentemente ligadas: peptideoglicano, arabinogalactano e ácidos micólicos, formando o chamado complexo micolil-arabinogalactano-peptideoglicano (mAGP) (Figura 1.1).

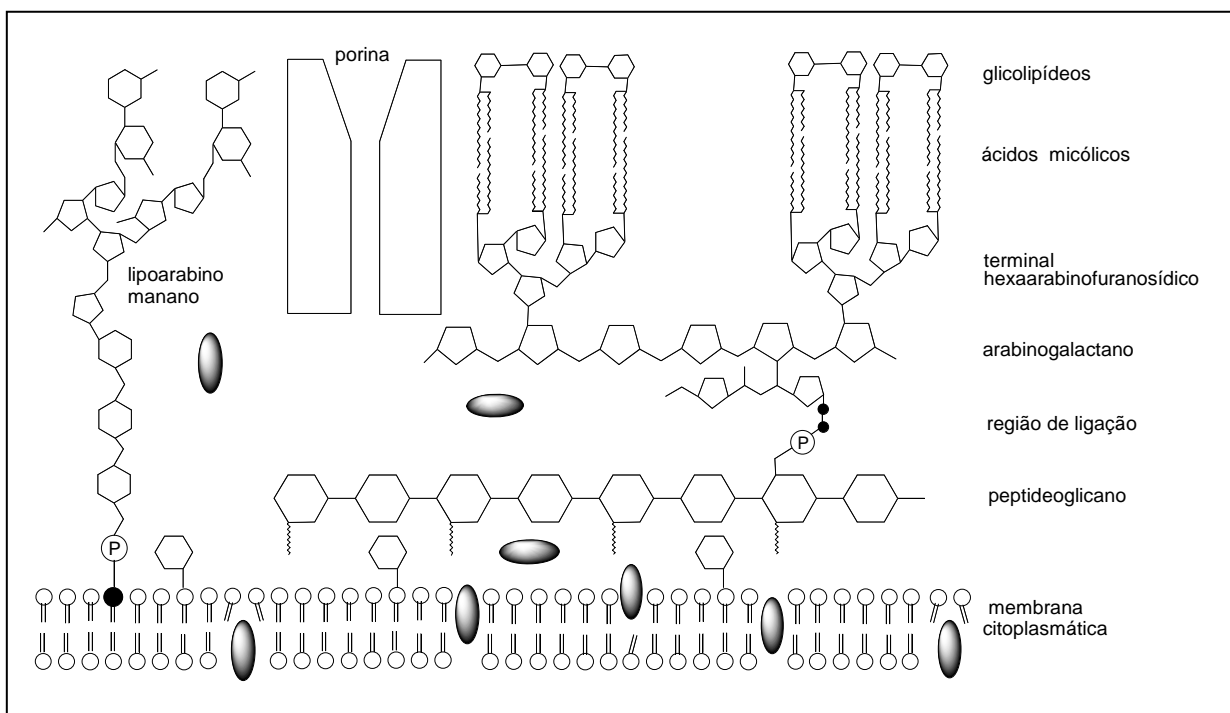


Figura 1.1. Representação esquemática da parede celular de *Mycobacterium*. A membrana citoplasmática é encapsulada por uma camada de peptideoglicano, que se liga à cadeia arabinogalactano através de uma região de ligação dissacarídeo-fosfato. A porção terminal do arabinogalactano é composta por uma estrutura hexaarabinofuranosídica ramificada esterificada por ácidos micólicos. As cadeias expostas de ácidos micólicos interagem com outras cadeias de ácidos micólicos dos glicolípideos dimicoliltrealose. Outros importantes componentes da parede celular de *Mycobacterium* são o lipoarabinomanano, que se liga diretamente à membrana citoplasmática (Schroeder *et al.*, 2002), e as porinas, que atravessam a camada hidrofóbica externa e permitem a entrada de nutrientes hidrofílicos (Niederweis, 2003).

A membrana citoplasmática de *Mycobacterium* é encapsulada por uma camada de peptideoglicano, um polissacarídeo formado de unidades repetidas de ácido *N*-acetil- β -D-glucosaminil-(1 \rightarrow 4)-*N*-glicomurâmico ligadas a cadeias peptídicas L-alanil-D-isoglutamil-*meso*-diaminopimelil-D-alanina (Figura 1.2), que utiliza o ácido *meso*-diaminopimélico como o ácido diamínico, e possui o resíduo de ácido murâmico *N*-glicosilado (ao invés de *N*-acetilado como em outras bactérias), permitindo assim a formação de pontes de hidrogênio adicionais e aumentando a rigidez da estrutura do peptideoglicano (Brennan & Nikaido, 1995).

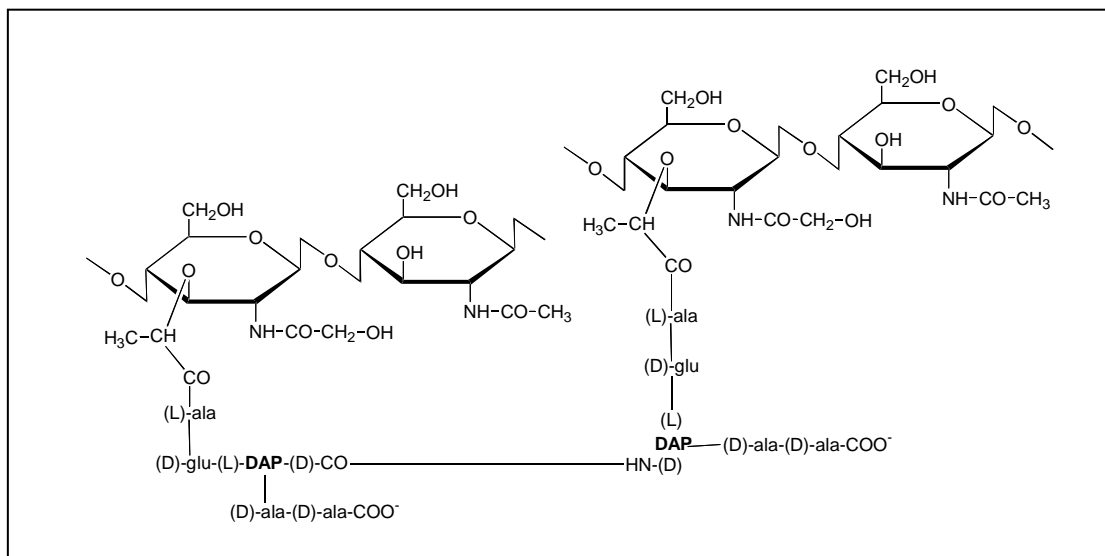


Figura 1.2. Componente peptideoglicano da parede celular de *Mycobacterium*, formado por unidades de ácido *N*-acetil- β -D-glucosaminil-(1 \rightarrow 4)-*N*-glicomurâmico ligadas a cadeias peptídicas L-alanil-D-isoglutamil-*meso*-diaminopimelil-D-alanina (Basso & Blanchard, 1998).

A estrutura do peptideoglicano é ligada à porção arabinogalactano por meio de uma ligação fosfodiéster entre alguns dos resíduos de ácido murâmico e o dissacarídeo de ligação do arabinogalactano.

O arabinogalactano é um polissacarídeo ramificado cuja primeira porção (proximal) é formada de cadeias de galactose que se ligam à porção distal de cadeias de arabinoses. A porção terminal do arabinogalactano é composta por uma estrutura hexaarabinofuranosídica ramificada, que apresenta 2/3 de seus anéis esterificados por ácidos micólicos (Figura 1.3), posicionados perpendicularmente à

membrana citoplasmática. Além de serem encontrados covalentemente ligados ao arabinogalactano da parede celular, os ácidos micólicos também são encontrados esterificando trealoses (dimicoliltrealoses) e gliceróis que se ligam não covalentemente à estrutura mAGP. Os ácidos micólicos ligados ao mAGP fazem parte de uma estrutura ordenada e formam uma barreira hidrofóbica externa de baixa fluidez que confere à célula de *Mycobacterium* uma permeabilidade extremamente baixa (Dinadayala *et al.*, 2003).

Recentemente, foi verificada a existência de canais (porinas) que atravessam esta camada hidrofóbica do envelope celular e permitem a entrada de nutrientes hidrofílicos ao interior da célula (Niederweis, 2003; Faller *et al.*, 2004).

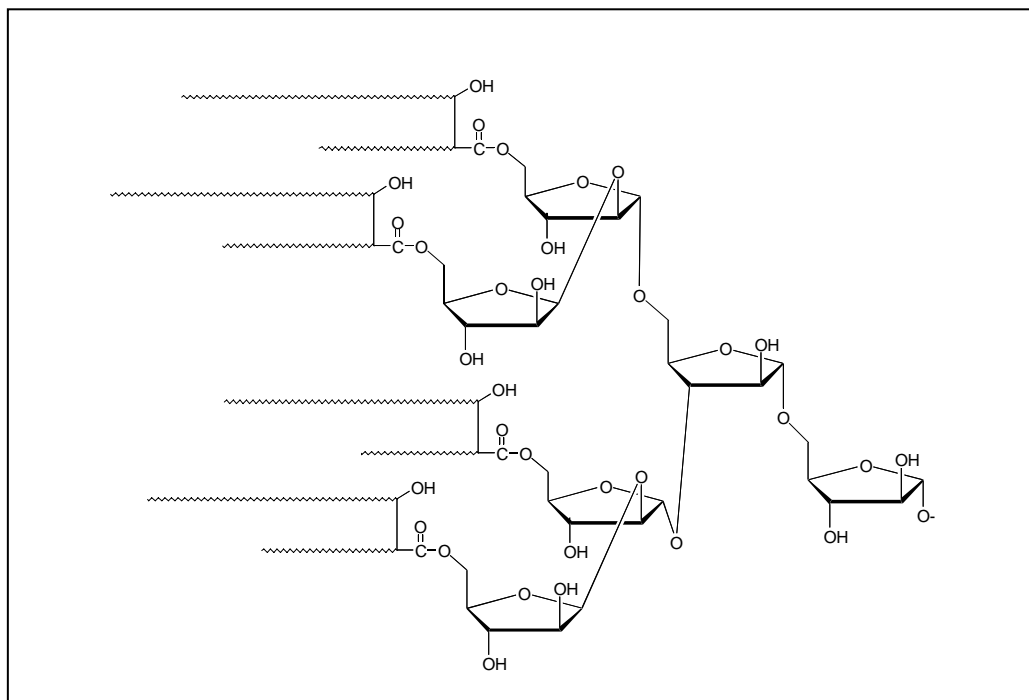


Figura 1.3. Estrutura da porção terminal do arabinogalactano. Estrutura hexaarabinofuranosídica ramificada, com 2/3 de seus anéis esterificados por ácidos micólicos (Schroeder *et al.*, 2002).

Em adição aos ácidos micólicos, o envelope celular contém uma grande variedade de lipídeos e glicolipídeos que se associam não covalentemente ao envelope celular e contribuem para a hidrofobicidade da superfície externa. Entre os glicolipídeos não covalentemente associados à estrutura mAGP, o que mais se destaca em MTB é a cadeia lipoarabinomanano (LAM). O LAM é diretamente

ancorado na membrana citoplasmática e apresenta diversas funções biológicas relevantes para a patogenicidade do MTB.

O conteúdo total de lipídeos na parede celular de *Mycobacterium* é superior à 60% em peso, dos quais mais de 40% apresenta-se covalentemente ligados à estrutura da parede celular (Draper, 1982).

Apesar de várias moléculas serem importantes para a formação da estrutura da parede celular, os ácidos micólicos são os componentes majoritários, e são característicos de *Mycobacterium* e outros gêneros relacionados. Os ácidos micólicos são ácidos graxos α -alquil β -hidroxilados de alto peso molecular, compostos de duas cadeias: uma menor, com 20-24 carbonos (C_{20-24}), e uma maior, com 50-60 carbonos (C_{50-60}). A clivagem térmica (pirólise) de ácidos micólicos leva à separação destas duas cadeias: a menor na forma de ácido graxo, chamada de cadeia α por ocupar a posição α ao grupo carboxílico, e a maior, chamada de cadeia meromicólica, na forma de aldeído (Figura 1.4). A cadeia α é saturada e a cadeia meromicólica é funcionalizada em intervalos regulares com grupos ciclopropila (α -micolatos), α -metilcetona (cetomicolatos) ou α -metilmetoxicetona (metoximicolatos) (Figura 1.5).

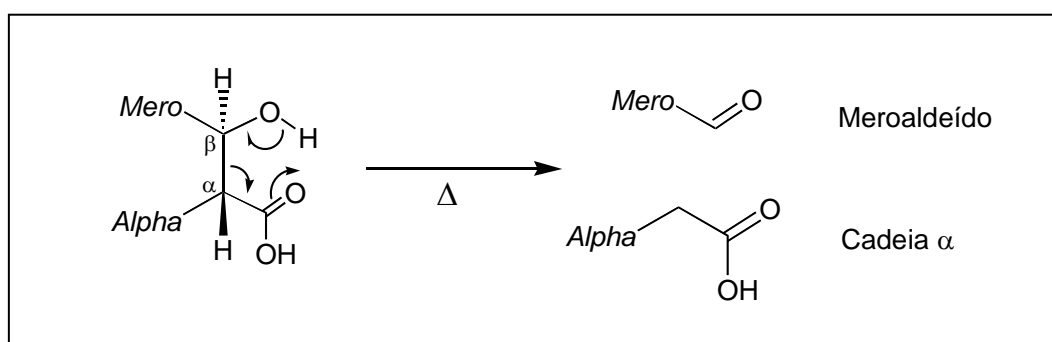


Figura 1.4. Esquema da pirólise de um ácido β -hidroxicarboxílico (ácido micólico), levando à separação da cadeia α e do meroaldeído (Barry III *et al.*, 1998).

Os ácidos micólicos de *Mycobacterium* são diferentes daqueles de outros gêneros como *Corynebacterium*, *Nocardia* e outros, pois estes últimos: (i) são maiores (C_{70-90}); (ii) possuem cadeias α mais longas (C_{20-26}); (iii) suas cadeias

meromicólicas possuem um ou dois grupos (duplas ligações ou anéis ciclopropano) capazes de introduzir um desvio da cadeia da molécula; (iv) podem conter funções oxigenadas além do grupo β -hidroxílico; e (v) podem apresentar grupos metila na cadeia meromicólica (Brennan & Nikaido, 1995). Nos outros gêneros, a única funcionalidade encontrada é a insaturação de cadeia.

As modificações presentes nos ácidos micólicos de *Mycobacterium* podem ocorrer em duas regiões da cadeia meromicólica, chamadas de distal (próxima à terminação ω da cadeia) e proximal (próximo do β -hidroxiácido). As modificações polares (grupos metiléter, cetonas, ésteres e epóxidos) tendem a ser restritas às posições distais, enquanto que as modificações não polares ocorrem em ambas posições, e incluem insaturações *cis* ou *trans* e ciclopropanos *cis* ou *trans* (Barry III *et al.*, 1998).

O MTB produz três classes de ácidos micólicos, que diferem entre si pela presença e/ou natureza dos substituintes na porção distal da cadeia meromicólica. Na série dos α -micolatos, a cadeia meromicólica apresenta dois anéis ciclopropano-*cis*. Na série dos metoximicolatos, esta cadeia é formada por um grupo metoxila com uma metila em α , na posição distal, e um anel *cis*- ou α -metil-*trans*-ciclopropano na posição proximal. Na série dos cetomicolatos, ela apresenta uma cetona com um grupo α -metila na posição distal, e o anel *cis*- ou α -metil-*trans*-ciclopropano na posição proximal (Figura 1.5) (Yuan *et al.*, 1997). Cada uma destas classes de cadeias meromicólicas de *Mycobacterium* ainda podem ser subdivididas em três tipos, em relação à ausência ou presença de insaturações *cis* ou *trans* nas cadeias, sendo predominante as cadeias do tipo 1 (sem insaturação) (Watanabe *et al.*, 2001; Watanabe *et al.*, 2002).

No MTB, as cadeias mais curtas de α -micolatos são predominantes (~57%), seguidas dos metoximicolatos (~37%), e dos cetomicolatos (~11%). É importante destacar que dentro de cada classe de ácidos micólicos de MTB existem consideráveis heterogeneidades em relação ao tamanho das cadeias α e meromicólicas, assim como posição das funcionalidades. A estimativa é que, dentre as classes majoritárias de ácidos micólicos de MTB, devam existir aproximadamente 500 isômeros estruturais (Barry III *et al.*, 1998).

Apesar da longa cadeia saturada da cadeia α , assim como a ainda mais longa cadeia meromíclica favorecerem um empacotamento regular e paralelo das cadeias carbonadas dos ácidos micólicos, a presença de funções oxigenadas, insaturações *cis* e ciclopropanos *cis* introduzem um desvio de cadeia. Estas características modulam o ordenamento dos ácidos micólicos, e desempenham uma função fundamental na estrutura e função do envelope celular de *Mycobacterium*.

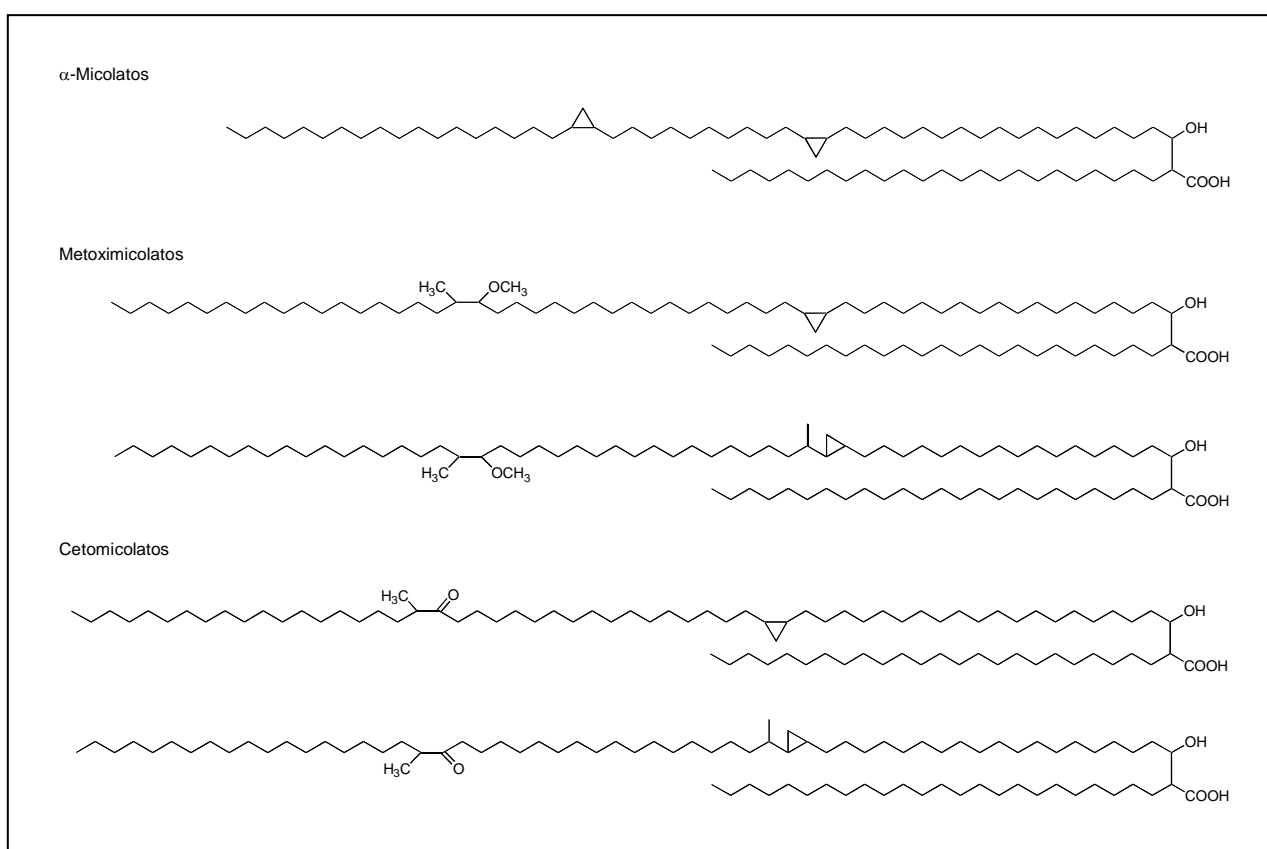


Figura 1.5. Estrutura dos ácidos micólicos identificados em *Mycobacterium tuberculosis*. (i) α -Micolatos com dois ciclopropanos *cis* em sua cadeia meromíclica; (ii) Metoximicolatos com grupos α -metil metiléter na posição distal e ciclopropanos *cis* ou α -metil *trans*-ciclopropano na posição proximal; e (iii) Cetomicolatos com um grupo α -metilcetona na posição distal e grupos iguais aos dos metoximicolatos na posição proximal (Schroeder *et al.*, 2002). Somente as cadeias do tipo 1 (sem insaturação) de cada uma das classes de ácidos micólicos estão representadas.

A natureza dos ácidos micólicos tem um papel fundamental na determinação da fluidez e permeabilidade do envelope celular do MTB, e modificações em sua

estrutura/composição levam a alterações das características da parede celular, podendo resultar até na inviabilidade da bactéria.

O alto conteúdo de ácidos micólicos e a sua importância na formação da parede micobacteriana fazem da biossíntese de ácidos micólicos uma via essencial para o MTB.

1.3- A biossíntese de ácidos graxos em *M. tuberculosis*

Os sistemas FAS I e FAS II

Dada a necessidade do MTB de sintetizar uma série de lipídeos únicos, não foi surpreendente o grande número de enzimas dedicadas ao metabolismo destas moléculas revelado pelo seqüenciamento de seu genoma em 1998 (Cole *et al.*, 1998; Camus *et al.*, 2002): mais de 250 em MTB, comparada com apenas 50 em *E.coli*, concordando com a presença de uma grande variedade de moléculas lipofílicas simples e complexas no envelope celular.

Dentre os genes envolvidos na síntese de ácidos micólicos, observou-se que o MTB possui as enzimas que fazem parte do chamado sistema de síntese de ácidos graxos, ou sistema FAS (*fatty acid synthase*), um dos processos bioquímicos mais importantes para o crescimento celular, diferenciação e homeostase.

No sistema FAS, a síntese de ácidos graxos é realizada a partir de acetil-CoA e malonil-CoA através de reações seqüenciais e intermediários acila (Wakil *et al.*, 1983). Inicialmente o acetil-CoA e o malonil-CoA são ativados por uma ligação tioéster com o grupo fosfopanteteína do cofator ACP ou CoA, ou com o complexo β -cetoacil sintase (por funções malonil-CoA-ACP transferase e acetil-CoA-ACP transacetilase), onde são sucessivamente adicionados resíduos malonato com concomitante eliminação de CO₂, resultando na adição de mais dois carbonos à cadeia de ácido graxo. O alongamento da cadeia de ácidos graxos no sistema FAS ocorre via repetidos ciclos de condensação, β -ceto redução, desidratação, e redução da ligação dupla resultante, catalisados, respectivamente, por funções β -cetoacil sintase (KAS), β -cetoacil redutase (KAR), β -hidroxiacil desidratase (DE), e enoil redutase (ENR), onde o agente redutor utilizado é o NADPH (Wakil *et al.*, 1983; Voet & Voet, 1995; Nelson & Cox, 2000).

Em bactérias e plantas, as sete funções enzimáticas para a síntese de ácidos graxos (6 enzimas e o ACP) estão localizadas em sete cadeias polipeptídicas independentes (codificadas por genes diferentes), que formam um complexo multi-enzimático onde as sucessivas enzimas ocupam sítios adjacentes (chamado de sistema FAS II).

Os sistemas FAS de leveduras e vertebrados também são complexos multi-enzimáticos, entretanto, em levedura as sete funções enzimáticas residem em duas cadeias polipeptídicas multi-funcionais, com três atividades na subunidade α e quatro na subunidade β . Nos vertebrados, uma única cadeia polipeptídica com múltiplos domínios, e arranjada na forma de homodímero, catalisa os sucessivos passos da síntese. Estes sistemas de síntese de ácidos graxos são chamados de sistema FAS I (Cronan & Rock, 1996).

Em contraste com os demais organismos, o *Mycobacterium* possui os dois sistemas de síntese de ácidos graxos, o FAS I e o FAS II. O sistema FAS I produz não somente ácidos graxos de cadeias mais curtas (C_{16} e C_{18}), os produtos usuais da síntese *de novo* de ácidos graxos, mas também de cadeias mais longas (C_{24} e C_{26}). Essa distribuição bimodal de produtos é resultante da habilidade da sintase multifuncional do FAS I de atuar tanto como uma *de novo* sintase, utilizando pequenos *primers*, quanto como uma enlongase, utilizando *primers* mais longos, que normalmente são os produtos da síntese *de novo* de ácidos graxos (Kolattukudy *et al.*, 1997). O sistema multienzimático FAS II, ACP-dependente e incapaz de realizar síntese *de novo*, é responsável pelo alongamento do palmitoil-ACP (um C_{16}) a ácidos graxos de cadeias longas entre C_{24} e C_{56} . Tanto a enzima polifuncional de FAS I, quanto o sistema multienzimático de FAS II são constituídos de proteínas citoplasmáticas (Salman *et al.*, 1999).

Na síntese de ácidos graxos em MTB, o sistema FAS I sintetiza ácidos graxos (pela condensação sucessiva de resíduos malonato ao acetil-CoA inicial) numa distribuição bimodal de C_{16-18} e C_{24-28} , que são removidos do complexo enzimático por uma transacilase na forma de acil-CoA de cadeia longa. A enzima *mtFabH*, que tem especificidade pelos produtos de cadeia menor do sistema FAS I, catalisa a condensação destes precursores acil-CoA com malonil-AcpM (um éster de ácido malônico ligado ao ACP característico de MTB, o AcpM). O produto é então alongado pelo sistema FAS II, resultando em cadeias carbônicas longas precursoras da cadeia meromíclica (C_{50-60}) dos ácidos micólicos. Os produtos mais longos do sistema FAS I (C_{24-26}) também são retirados do ciclo de alongamento para serem utilizados (provavelmente ligado à CoA) como substrato na formação da cadeia α dos ácidos micólicos (Figura 1.6).

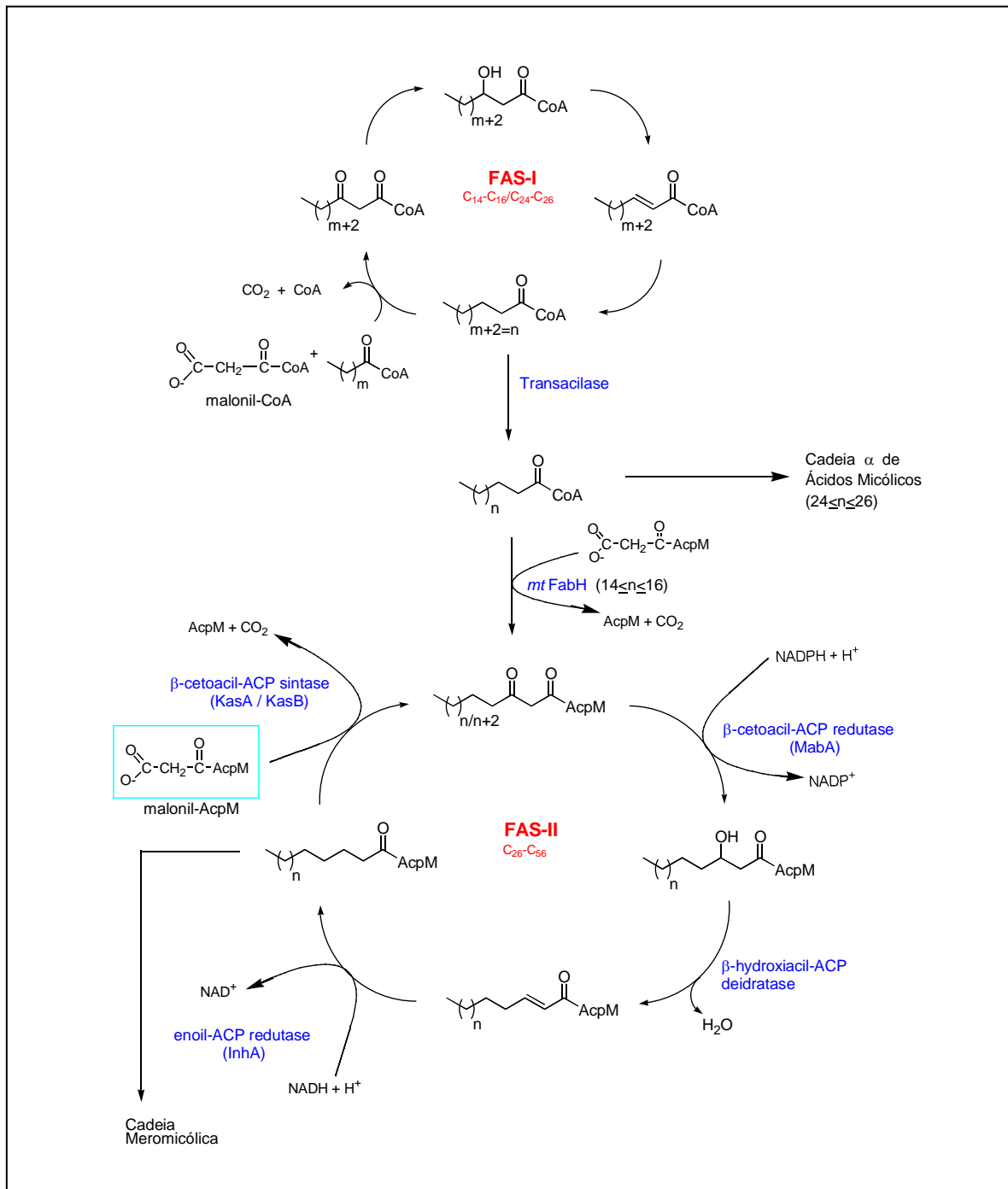


Figura 1.6. Biossíntese de ácidos graxos em *Mycobacterium tuberculosis*. O sistema FAS I sintetiza ácidos graxos numa distribuição bimodal de C₁₆₋₁₈ e C₂₄₋₂₈, que são removidos do complexo enzimático por uma transacilase na forma de acil-CoA de cadeia longa. A enzima mtFabH, que tem especificidade pelos produtos de cadeia menor do sistema FAS I, catalisa a condensação desses precursores acil-CoA com o malonil-ACP. O produto é então alongado pelo sistema FAS II, resultando em cadeias carbônicas longas precursoras da cadeia meromícica (C₅₀₋₆₀) dos ácidos micólicos. Os produtos mais longos do sistema FAS I (C₂₄₋₂₆) são utilizados (provavelmente ligado a CoA) como substrato na formação da cadeia α dos ácidos micólicos (Schroeder *et al.*, 2002).

Com o seqüenciamento completo do genoma do *Mycobacterium tuberculosis* H37Rv (Cole *et al.*, 1998; Camus *et al.*, 2002), os genes que codificam as enzimas do sistema FAS II foram identificados. O alongamento de cadeia é feito por quatro diferentes enzimas. Inicialmente, a etapa de condensação de uma unidade de malonil-AcpM à cadeia crescente de ácido graxo (na forma de acil-AcpM) pode ser catalisada por duas β -cetoacil-ACP sintases, KasA e KasB (Schaeffer *et al.*, 2001; Slayden & Barry III, 2002), liberando o AcpM e CO₂. O grupo ceto do β -cetoéster resultante é então reduzido pela enzima MabA, uma β -cetoacil-ACP redutase NADPH-dependente (Banerjee *et al.*, 1998; Marrakchi *et al.*, 2002), e a eliminação do grupo β -OH por uma β -hidroxiacil-ACP desidratase forma um derivado 2-*trans*-enoil-AcpM. A etapa final do ciclo é a redução da ligação dupla do enoil-AcpM pela enzima InhA, uma 2-*trans*- enoil-ACP redutase, que ao contrário das demais redutases dos sistemas FAS, é uma enzima NADH-dependente, e leva ao acil-AcpM correspondente que novamente é condensado ao malonil-AcpM, até que o ácido graxo de tamanho desejado seja sintetizado (Figura 1.6).

As atividades correspondentes nos dois sistemas FAS estão relacionadas em função e estrutura, mas geralmente não se observa uma homologia total de seqüência.

A grande importância da biossíntese de ácidos graxos para a síntese de ácidos micólicos em MTB, assim como a existência de um sistema exclusivo para o alongamento da cadeia, fazem do sistema FAS II um alvo atraente para intervenção (Payne *et al.*, 2001). Compostos como triclosan (TCL), hexaclorofeno, tiolactomicina (TLM), cerulenina (CER), etionamida, e o tuberculostático de primeira linha, isoniazida reconhecidamente atuam na inibição das enzimas desse sistema (Figura 1.7) (Schroeder *et al.*, 2002).

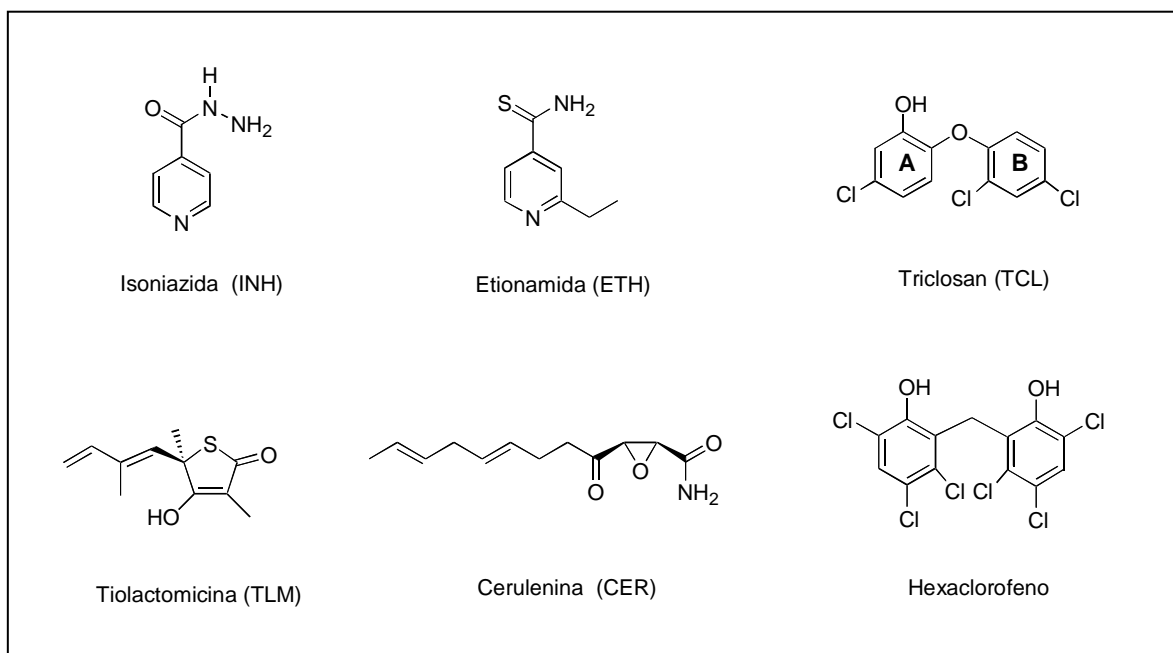


Figura 1.7. Algumas drogas que reconhecidamente inibem o sistema FAS II de *M. tuberculosis*.

1.4- Inibição do sistema FAS II pela Isoniazida (INH) e identificação de seu alvo molecular

A isoniazida (INH), ou hidrazida do ácido isonicotínico (Figura 1.7), tem sido a base da quimioterapia anti-TB por mais de 50 anos, desde sua descoberta como um agente efetivo no tratamento de infecções por *M. tuberculosis* e *M. bovis* em 1952 (Middlebrook, 1952).

Enquanto outras espécies de micobactérias são pouco suscetíveis à INH, o *M. tuberculosis* e *M. bovis* são extremamente sensíveis a essa droga em concentrações que variam de 0,02 a 0,2 µg/mL (Bernstein *et al.*, 1952). Mas o surgimento de cepas resistentes foi relatado logo após o início de sua utilização (Middlebrook, 1952).

Uma das primeiras conseqüências observadas no tratamento com INH foi a perda da característica álcool-ácido resistente do *Mycobacterium*, indicando que esta droga é ativa contra componentes da parede celular.

Embora a INH tenha sido objeto de investigações bioquímicas por várias décadas, o alvo de ação e resistência a essa droga só foi esclarecido nos últimos anos. O tratamento do MTB com INH inibe a síntese de ácidos micólicos (Winder & Collins, 1970) e causa o acúmulo de ácidos graxos saturados de 16 carbonos (Takayama *et al.*, 1975). A inibição desta via biossintética está relacionada com a morte do bacilo (Takayama *et al.*, 1972): o MTB passa por uma seqüência de alterações moleculares que resulta em lise após 24 horas, ou seja, após aproximadamente uma geração de bacilos (Takayama *et al.*, 1973).

Utilizando uma abordagem genética para localizar o alvo da isoniazida, Jacobs e colaboradores identificaram uma ORF (*open reading frame*), chamada de *inhA*, onde um único evento mutagênico era suficiente para conferir resistência à esta droga (Banerjee *et al.*, 1994). Dessa forma ficou demonstrado que o fenótipo de resistência à droga INH devia-se a uma mutação no gene estrutural *inhA*, que levava à substituição de um aminoácido serina na posição 94, por uma alanina. Essa mutação resultava na incapacidade da INH de inibir a síntese de ácidos micólicos. A super-expressão do produto do gene *inhA* espécie selvagem em *M.*

smegmatis e *M. bovis* também conferia resistência à INH e ETH (Banerjee *et al.*, 1994). Assim, concluiu-se que o produto desse gene seria o alvo primário da ação da INH (e também da ETH, um análogo estrutural da INH).

Análises bioquímicas e a determinação da estrutura tridimensional do produto do gene *inhA* identificaram a enzima InhA como uma enoil-ACP redutase NADH-dependente, com especificidade por substratos de cadeia longa (consistente com seu envolvimento na síntese de ácidos micólicos) (Quémard *et al.*, 1995; Dessen *et al.*, 1995).

Apesar de sua estrutura simples, a ação da INH na enoil redutase não é direta. Experimentos de cinética enzimática em estado estacionário mostraram que este composto é um inibidor fraco da atividade enzimática, sugerindo que a INH *per se* não é a forma ativa da droga, mas deve ser ativada para interagir com a enzima (Quémard *et al.*, 1995).

Estudos de cepas resistentes confirmaram a INH como uma pró-droga que é convertida à sua forma biologicamente ativa por uma enzima com atividade catalase/peroxidase, codificada pelo gene *katG* (Zhang *et al.*, 1992; Zhang *et al.*, 1993; Heym *et al.*, 1993). Isolados clínicos que se tornam resistentes à INH o fazem primariamente por mutação na KatG, tornando o bacilo incapaz de ativar a pró-droga (Musser *et al.*, 1996). A deleção ou mutação deste gene tem sido associada com a diminuição da sensibilidade à INH em aproximadamente 50% dos isolados clínicos de MTB (Heym *et al.*, 1995).

A natureza da espécie ativa da INH não é certa, mas sabe-se que os produtos estáveis da reação enzimática com KatG (ácido isonicotínico, isonicotinamida e 4-carboxialdeído piridina) (Johnsson & Schultz, 1994; Johnsson *et al.*, 1995) não são tóxicos ao bacilo em concentrações fisiológicas. Acredita-se que a toxicidade *in vivo* deva-se à formação de espécies eletrofílicas (radical e/ou ânion) capazes de interagir com componentes celulares.

Em 1996, Quémard *et al.* demonstraram que a inativação da enzima InhA por INH (na presença de KatG) requer a formação inicial de um complexo InhA-NADH (Quémard *et al.*, 1996). A estrutura cristalina da enzima complexada com INH e NADH mostrou que o anel acil-piridínio da INH está covalentemente ligado ao anel nicotinamida do NAD⁺ dentro do sítio ativo da enzima (Rozwarski *et al.*, 1998). Assim, postulou-se que, no MTB, a INH ativada se liga ao NAD(H) dentro do

sítio da enzima, e este bi-substrato formado é responsável pela inibição (Figura 1.8). A síntese e isolamento recente desse bi-substrato na ausência da enzima levaram ao questionamento quanto a seu local de formação: se ele realmente se dá no sítio, ou antes de sua associação com a InhA (Wilming & Johnsson, 1999; Lei *et al.*, 2000; Rawat *et al.*, 2003). Independente de onde é formado, a inibição da InhA pela INH só ocorre pela associação do composto covalente NAD-INH. A possibilidade de existência de outras formas ativas do bi-substrato que inibam a enzima, como por exemplo, isômeros cíclicos, também já foi sugerida (Broussy *et al.*, 2003), mas aguarda confirmação.

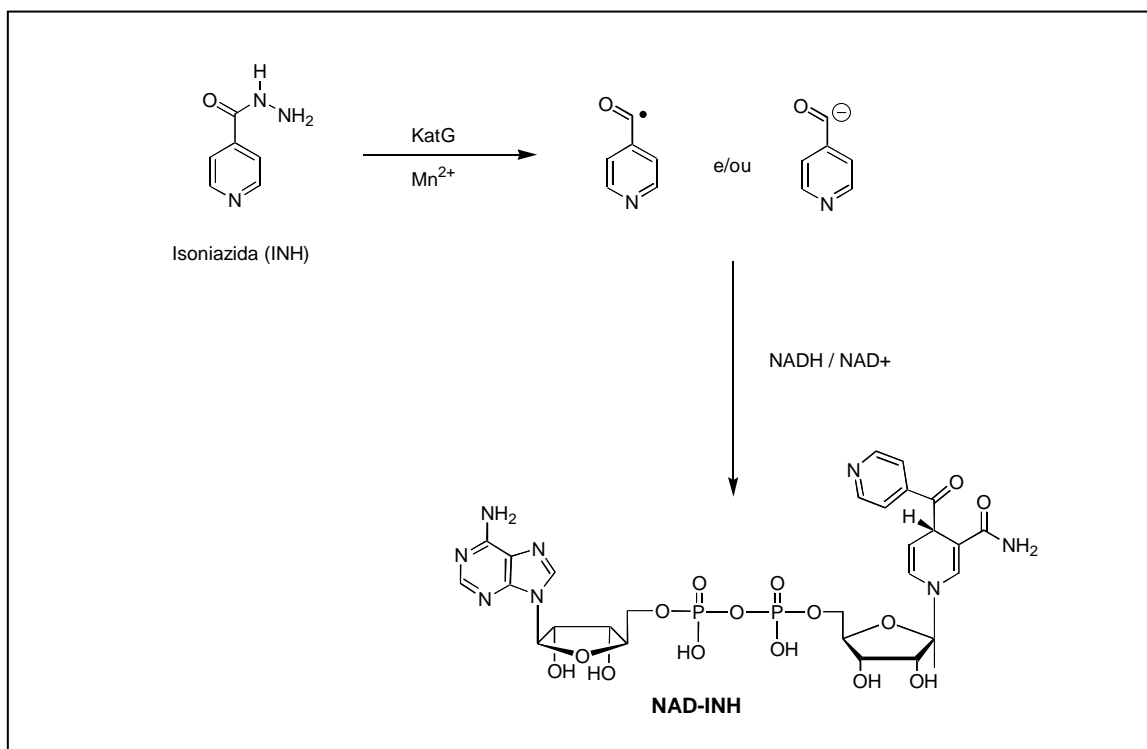


Figura 1.8. Ativação da pró-droga isoniazida e formação do aduto covalente com o NAD(H).

O certo é que a resistência parece estar associada à interação da enzima com a coenzima. Isolados clínicos resistentes à INH que apresentam mutações no gene *inhA* (mas não em *katG*), e representam aproximadamente 20% dos casos de resistência, resultam em substituições de aminoácidos do sítio de ligação do NADH. Estudos cinéticos com os mutantes I16T, I21V e I47T mostraram que, nesses

mutantes, o NADH apresenta valores maiores de K_m e uma constante de dissociação (K_d) de 10 a 150 vezes maior do que o observado para a enzima espécie selvagem (Tabela 1.2), enquanto que a $V_{m\acute{a}x}$ e o K_m do substrato variam muito pouco (Basso *et al.*, 1998a). O mesmo foi observado para o mutante S94A² (Dessen *et al.*, 1995; Quémard *et al.*, 1995). Isto indica que a resistência à INH deve-se a uma menor afinidade da enzima InhA pelo NADH e, conseqüentemente, pelo bi-substrato NAD-INH.

Tabela 1.2. Parâmetros de cinética enzimática da enzima InhA de *Mycobacterium tuberculosis* espécie selvagem (*wt*) e mutantes (Basso *et al.*, 1998a).

Enzima InhA	$V_{m\acute{a}x}$ (U/mg)	K_m substrato ^a (μ M)	K_m NADH (μ M)	K_d NADH (μ M)
<i>wt</i>	33 ± 1	75 ± 5	56 ± 4	0,57 ± 0,04
I16T	39 ± 2	85 ± 8	149 ± 10	5,95 ± 0,33
I21V	25 ± 1	102 ± 13	104 ± 7	13,9 ± 1,7
I47T	23 ± 1	42 ± 4	243 ± 25	85,1 ± 6,3

^a Substrato: 2-*trans*-dodecenoil-CoA (C₁₂)

Apesar da InhA estar reconhecidamente envolvida na resistência à INH, o acúmulo de ácido saturado C₂₆ (Takayama *et al.*, 1975), que não seria esperado pela inibição de uma enoil redutase, e a identificação de um complexo covalente formado pela enzima KasA, AcpM e INH quando o MTB era tratado com esta droga levaram à proposição de que o alvo principal da INH seria a β -cetoacil sintase KasA, e não a enoil redutase InhA (Mdluli *et al.*, 1998; Slayden *et al.*, 2000). Isto levaria a uma situação bioquimicamente incomum, onde duas enzimas consecutivas do sistema FAS II seriam alvo de um mesmo inibidor, por mecanismos completamente diferentes.

Estudos posteriores mostraram que a super-expressão de *inhA*, e não de *kasA*, é que confere resistência a INH e ETH em *M. smegmatis*, *M. bovis*, e *M. tuberculosis*, portanto a KasA não seria alvo da INH (Larsen *et al.*, 2002). Recentemente, outro estudo para identificar o papel da InhA e da KasA em relação à atividade da droga evidenciou bioquímica e geneticamente que somente a

² Parâmetros cinéticos medidos para o mutante S94A, em relação à enzima espécie selvagem (*wt*) com o substrato 2-*trans*-octenoil-ACP: *wt* InhA - K_m substrato = 2 ± 1 μ M e K_m NADH = 8 ± 0 μ M, S94A InhA - K_m substrato = 3 ± 1 μ M e K_m NADH = 38 ± 0 μ M (Dessen *et al.*, 1995).

inibição de InhA, mas não de KasA, é que induz o acúmulo do complexo contendo a KasA, e que a KasA não é alvo da INH, provando assim que a InhA é definitivamente o alvo principal da INH (Kremer *et al.*, 2003).

Além dos genes *inhA*, *katG* e *kasA*, outros genes também foram relacionados com a sensibilidade e/ou resistência do bacilo à INH, entre eles *ahpC* e *ndh*. O gene *ahpC* codifica uma alquil hidroperoxidase que atua como uma redutase antioxidante. O aumento da expressão da AhpC diminuiria a concentração intracelular de peróxido, evitando assim a formação de espécies eletrofílicas da INH mediada pela KatG (Sherman *et al.*, 1996). O gene *ndh* codifica uma NADH desidrogenase envolvida na cadeia de respiração celular. Defeitos nesta desidrogenase aumentam a razão NADH/NAD⁺ e conferem resistência à INH, possivelmente por dois processos: promovendo a redução de peróxidos, o que interferiria no processo de ativação da droga pela KatG, ou deslocando o aduto NAD-INH do sítio ativo da InhA (Miesel *et al.*, 1998).

Enquanto a maioria dos mecanismos de resistência propostos até o momento está de alguma forma relacionada com a ativação da pró-droga INH, a resistência causada diretamente pelo alvo molecular da droga ativada deve-se a mutações em resíduos envolvidos no sítio de ligação do cofator (NADH) da InhA.

Por se tratar de uma enzima específica do MTB, e por ser comprovadamente o alvo primário da principal droga utilizada no combate à TB, a InhA apresenta-se como um alvo interessante para o desenvolvimento de novas drogas anti-TB. O entendimento de seu mecanismo de ação enzimática, assim como a compreensão dos fatores que influenciam na ligação do cofator no sítio ativo da enzima e em seu mecanismo de resistência podem contribuir para o desenvolvimento de novos agentes terapêuticos anti-TB que atuem diretamente nesta enzima sem necessidade de ativação prévia.

1.5- A enzima InhA

A enzima InhA, ou 2-*trans*- Enoil – ACP (CoA) Redutase (EC 1.3.1.9, 269 aa; ~29 kD), catalisa a última reação da seqüência de alongamento da cadeia de ácidos graxos do sistema FAS II de MTB (Figura 1.6). Esta enzima é responsável pela redução NADH-dependente de derivados 2-*trans*-enoil-ACP de ácidos graxos de cadeias longas em seu correspondente saturado, resultando na redução estereoespecífica da dupla ligação do tioéster α,β -insaturado (Figura 1.9), cuja porção não lipídica pode ser tanto uma proteína carreadora de acilas (ACP) quanto a coenzima A (CoA) (Quémard *et al.*, 1995). Esta enzima tem preferência por substratos de cadeia carbônica longa (12-24 carbonos), característica consistente com a sua função na síntese de ácidos graxos saturados, precursores de ácidos micólicos (Blanchard, 1996).

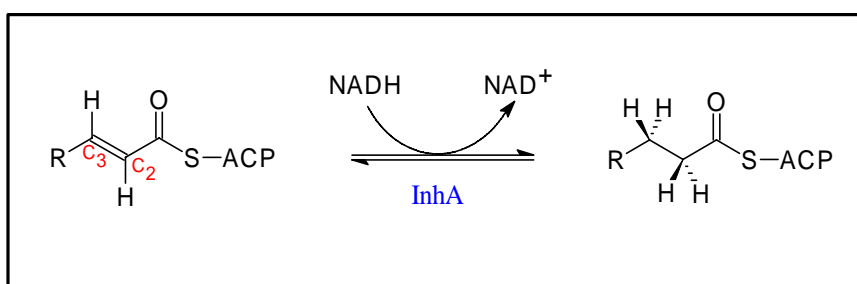


Figura 1.9. Reação de redução de um tioéster α,β -insaturado catalisada pela InhA.

Neste processo, a enzima InhA catalisa a transferência de hidreto (H^-) do NADH para o substrato, resultando na redução da dupla ligação α -tioéster (entre C2 e C3) da porção lipídica do substrato.

Ensaio de efeito isotópico determinaram a estereoquímica da transferência do hidreto. A reação se processa pela transferência direta do hidrogênio 4S do anel nicotinamida do NADH, na forma de hidreto, ao carbono C3 do substrato 2-*trans*-enoil-tioéster, com a formação de um enolato intermediário. O colapso deste enolato, via protonação do C2, leva ao produto saturado (Quémard *et al.*, 1995 - Figura 1.10). A investigação do curso estereoquímico da reação mostrou que a

redução se processa pela adição *syn* de hidrogênio à ligação dupla (Fillgrove & Anderson, 2000).

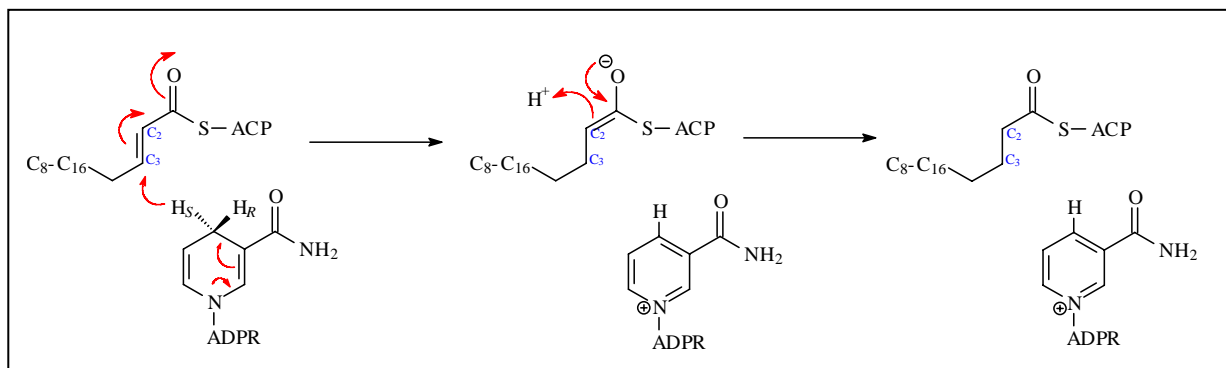


Figura 1.10. Mecanismo químico da reação de redução catalisada pela enzima InhA. O hidrogênio H_S é transferido, na forma de hidreto (H^-), do anel nicotinâmico da coenzima NADH para o carbono C_3 do substrato. Um enolato intermediário é formado, e o NADH é oxidado a NAD^+ . A adição de um próton (H^+) ao carbono C_2 leva à redução do substrato (Quémard *et al.*, 1995).

Estudos do mecanismo enzimático em estado-estacionário da enoil redutase de MTB indicaram que, apesar do mecanismo cinético não ser estritamente ordenado, o NADH se liga preferencialmente à enzima e é seguido pelo substrato (Quémard *et al.*, 1995).

Um estudo mais aprofundado do mecanismo de ação e de inibição tornou-se possível com o conhecimento das características estruturais da InhA, pela resolução de suas estruturas cristalinas. A primeira estrutura resolvida corresponde ao complexo bimolecular entre a enzima espécie selvagem (*wt* InhA) e seu cofator, o NADH (PDB id: 1ENY - Dessen *et al.*, 1995). A estrutura do mutante S94A (S94A InhA) resistente à INH também foi determinada para estudo dos fatores estruturais envolvidos na resistência (PDB id: 1ENZ - Dessen *et al.*, 1995). Em 1998, o complexo inibidor formado pela ligação covalente da INH ativada com o NAD(H) (NAD-INH) foi cristalizado no sítio ativo da enzima (PDB id: 1ZID - Rozwarski *et al.*, 1998). Apesar destas estruturas já indicarem a região de ligação do derivado de ácido graxo, o modo de ligação do substrato só foi determinado em 1999, com a cristalização da InhA com o cofator e um análogo ao substrato (tioéster formado por um derivado de ácido graxo α,β -insaturado de 16 carbonos ligado ao grupo N-

acetilcisteamina - chamado de C₁₆) (PDB id: 1BVR - Rozwarski *et al.*, 1999). Recentemente foi publicada a estrutura tridimensional da InhA inibida pelo triclosan (TCL) e por um derivado aril-indol-piperazina (GEQ) (PDB id: 1P45 e 1P44 - Kuo *et al.*, 2003) (Tabela 1.3).

Tabela 1.3. Estruturas tridimensionais da enzima InhA de *Mycobacterium tuberculosis* depositadas no Swiss Protein Data Bank (em janeiro/2004).

Enzima	Ligante	Resolução	PDB id	Referência
<i>wt</i> InhA	NADH	2,2 Å	1ENY	Dessen <i>et al.</i> , 1995
S94A InhA	NADH	2,7 Å	1ENZ	Dessen <i>et al.</i> , 1995
<i>wt</i> InhA	NAD-INH ^a	2,7 Å	1ZID	Rozwarski <i>et al.</i> , 1998
<i>wt</i> InhA	NAD ⁺ e C ₁₆ ^b	2,8 Å	1BVR	Rozwarski <i>et al.</i> , 1999
<i>wt</i> InhA	NAD ⁺ e TCL ^c	2,6 Å	1P45	Kuo <i>et al.</i> , 2003
<i>wt</i> InhA	NAD ⁺ e GEQ ^d	2,7 Å	1P44	Kuo <i>et al.</i> , 2003

^a NAD-INH: complexo inibidor da InhA, formado pelo anel acil-isonicotínico da isoniazida (INH) ligado covalentemente ao NADH; ^b C₁₆: tioéster 2-*trans*-hexadodecenoil-(N-acetilcisteamina); ^c TCL: triclosan; ^d GEQ: 5-[[4-(9H-fluoren-9-il)piperazin-1-il]carbonil]-1H-indol

A InhA pertence a uma família de enzimas com função desidrogenase/redutase, as SDR (*short chain dehydrogenase/reductase*), que utilizam como cofator uma molécula de NAD(H) ou NADP(H), e apresenta similaridade na seqüência de aminoácidos com duas enzimas enoil-ACP-redutases previamente caracterizadas: a FabI de *E. coli* (36% de identidade) e a ENR de *Brassica napus* (28% de identidade) (Parikh *et al.*, 2000). A principal característica desta família é a topologia do esqueleto peptídico, onde cada subunidade é formada por um único domínio com um núcleo central do tipo *Rossmann fold* onde está localizado o sítio de ligação do cofator (Rossmann *et al.*, 1975; Jörnvall *et al.*, 1995; Oppermann *et al.*, 2003).

Na InhA, várias α -hélices e fitas β do *Rossmann fold* se estendem em torno do sítio de ligação do NADH e formam uma profunda cavidade para o substrato (Sacchettini & Blanchard, 1996). Sua estrutura geral lembra uma cadeira onde o sítio de ligação do cofator é uma cavidade entre o “encosto” e o “assento” da estrutura e o sítio de ligação do substrato localiza-se no “encosto” (Figura 1.11) (Dessen *et al.*, 1995).

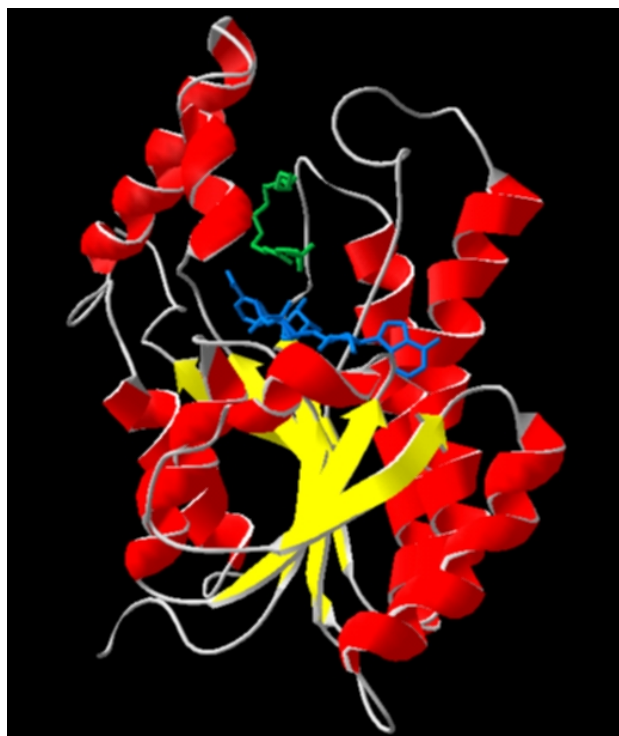


Figura 1.11. Estrutura da subunidade C da enzima InhA (PDB id: 1BVR) com α -hélices em vermelho e fitas β em amarelo. O NADH (azul) está posicionado sobre o leque de fitas β do *Rossmann fold*, com um análogo ao substrato (verde) na cavidade de ligação, logo acima do NADH. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A InhA é um homotetrâmero em solução aquosa, formado pelas subunidades C, D, E e F (Figura 1.12). A análise independente de cada subunidade revela que a região da cadeia no lado oposto ao sítio ativo tem um fator de temperatura (*B-factor*) muito baixo (região de baixa mobilidade), e deve estar envolvido no contato com as outras subunidades, na parte interna do homotetrâmero. De fato, uma considerável porção de cada subunidade do homotetrâmero está envolvida em contatos intermoleculares. A superfície de contato estimada pelo programa DSSP (Kabsch & Sander, 1983) entre a subunidade C e D é de 1550 \AA^2 (~12% da subunidade), entre C e E é de 700 \AA^2 (~5% da subunidade), e entre C e F é de 1500 \AA^2 (~12% da subunidade). Logo, a superfície total de contato de qualquer uma das subunidades do homotetrâmero é de 3750 \AA^2 (aproximadamente 29% da superfície total) (Rozwarski *et al.*, 1999).

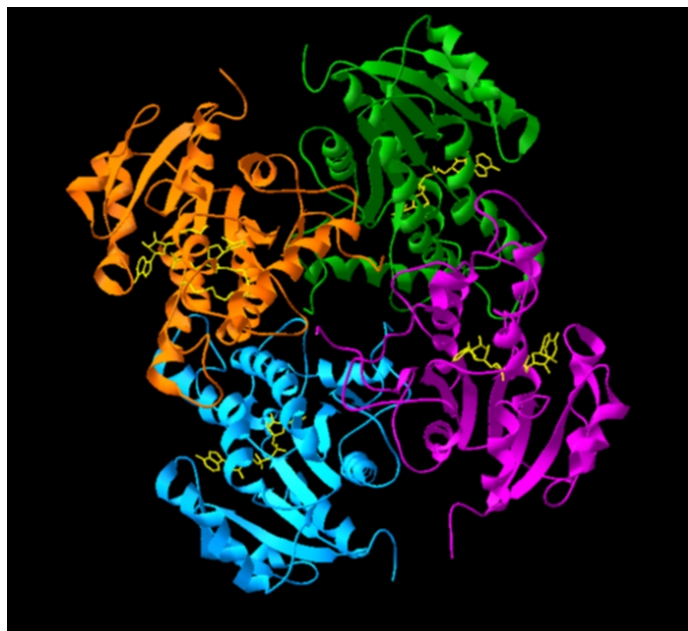


Figura 1.12. Homotetrâmero da InhA (PDB id: 1BVR). Cada subunidade é mostrada de uma cor diferente: 1BVR-C (laranja), 1BVR-D (azul), 1BVR-E (magenta) e 1BVR-F (verde), com os ligantes NAD⁺ e o substrato (somente na subunidade 1BVR-C) em amarelo. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Os aminoácidos que delimitam o lado mais externo (em relação ao tetrâmero) da cavidade de ligação, à esquerda do substrato, formam duas α -hélices perpendiculares sustentadas por alças, e são chamados de “alça de ligação do substrato” (resíduos 196-219)³. Os aminoácidos à direita do substrato formam duas grandes alças. Enquanto as alças à direita do substrato encontram-se do lado da enzima que faz contato com as demais subunidades do homotetrâmero, as duas hélices da alça de ligação do substrato estão voltadas para fora do tetrâmero, em direção ao solvente (Figura 1.13A). Os sítios de ligação das quatro subunidades, distantes entre si (mais de 40 Å entre subunidades vizinhas), estão voltados para lados opostos na estrutura do homotetrâmero (Figura 1.13B).

³ A alça de ligação do substrato na InhA é maior do que os das demais ENRs, de acordo com a especificidade da ENR de *M. tuberculosis* de reduzir substratos maiores.

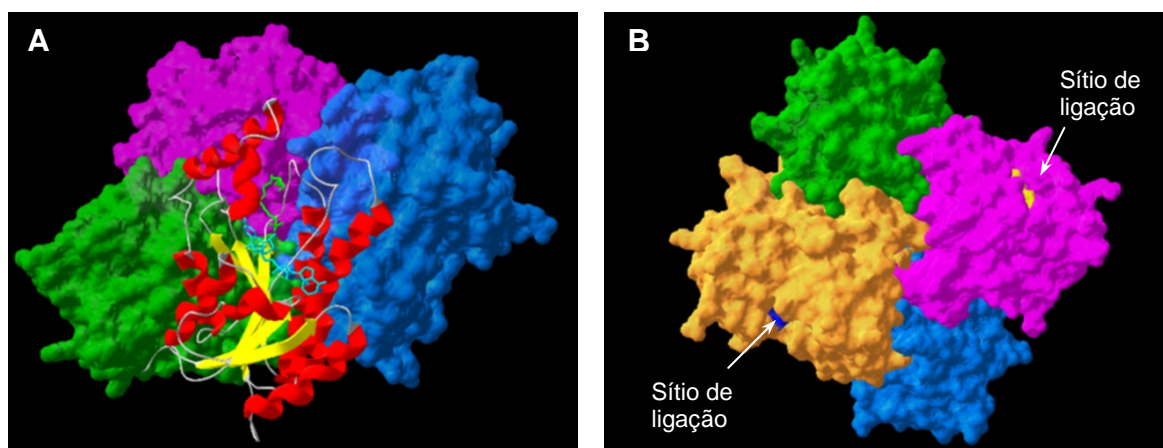


Figura 1.13. Relação entre as quatro subunidades do homotetrâmero da InhA. (A) Representação da subunidade 1BVR-C em relação a suas estruturas secundárias, com os ligantes NADH (ciano) e C₁₆ (verde) e superfícies moleculares das demais subunidades (azul, magenta e verde). As alças à direita do substrato fazem contato com a subunidade vizinha (em azul), enquanto as duas hélices da alça de ligação do substrato estão voltadas para fora do tetrâmero; (B) Os sítios de ligação das subunidades estão voltados para lados opostos na estrutura do homotetrâmero. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Em cada uma das subunidades, o NADH assume uma conformação estendida dentro de sua cavidade de ligação, ao longo das porções carbóxi-terminal do núcleo de fitas β . O anel nicotinamida do NADH se liga ao fundo, enquanto a porção adenina se direciona para fora da cavidade de ligação. Acima do NADH, o substrato C₁₆ assume uma conformação em forma de “U” (Figura 1.14), e é fixado pela alça de ligação do substrato.

A cavidade de ligação do substrato tem um formato oval, com dimensões de aproximadamente 16 Å x 13 Å x 7 Å (Rozwarski *et al.*, 1999). Um lado fica completamente aberto e exposto ao solvente, enquanto que o outro lado contém somente uma pequena abertura. A porção terminal da longa cadeia lipídica se orienta para fora do sítio de ligação, possibilitando assim a redução de substratos de cadeias mais longas. A porção *N*-acetilcisteamina do C₁₆ também se dirige em direção ao solvente, concordando com a necessidade desta porção do substrato natural, isto é, a porção ACP, de não entrar no sítio catalítico da enzima por ser um grupo volumoso (Rozwarski *et al.*, 1999).

Observando a cavidade de ligação dos ligantes NADH e C₁₆ na InhA (Figura 1.14), pode-se supor que a entrada de um substrato grande do tipo do C₁₆ não seja fácil. Conforme proposto por Rozwarski *et al.* (1999), as duas hélices da esquerda

(que são sustentadas por alças) devem se deslocar, abrindo espaço para a entrada do substrato. Esta movimentação deve ser possível, pois a alça de ligação do substrato (formado por estas duas hélices) é externa e não faz contato com outras subunidades do tetrâmero, e a abertura em direção ao solvente está desimpedida (Rozwarski *et al.*,1999).

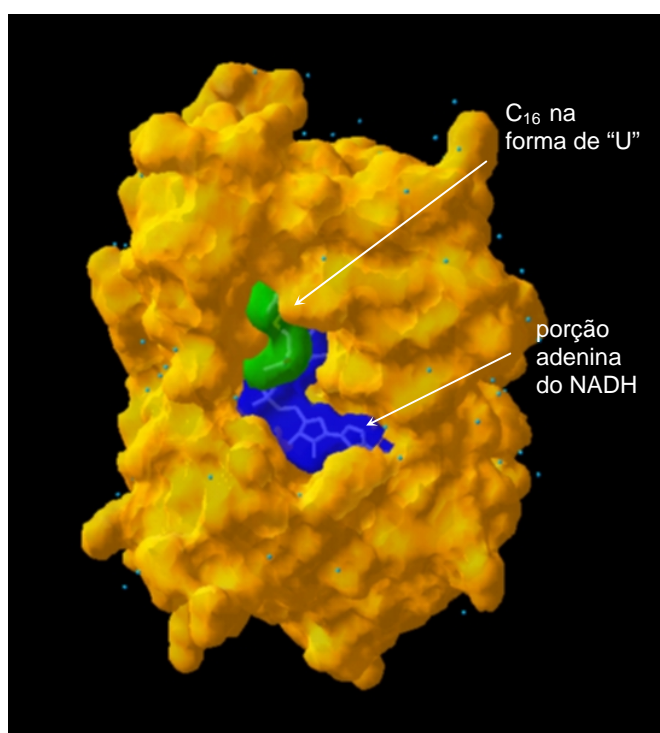


Figura 1.14. A cavidade de ligação da enzima InhA (superfície molecular da 1BVR subunidade C, em laranja) com seus ligantes NADH (azul) e C₁₆ (verde), e oxigênios das moléculas de água (solvente, esferas azuis). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Além disso, na parte superior da InhA, acima do sítio de ligação do C₁₆, na região de contato entre as duas hélices à esquerda e as alças à direita do substrato, há o predomínio de aminoácidos não carregados, e as interações entre estes dois segmentos da molécula são majoritariamente do tipo hidrofóbicas. Assim, o movimento de abertura da alça de ligação para a entrada do substrato não causa o rompimento de qualquer interação mais forte como, por exemplo, de uma ponte de hidrogênio, tornando possível o movimento de abrir e fechar sugerido para a entrada do substrato no sítio ativo.

De fato, na sobreposição da estrutura do complexo binário da InhA (1ENY), com o complexo ternário (1BVR) podemos identificar duas importantes mudanças conformacionais na proteína quando ela se liga também ao substrato. A alça de ligação do substrato (resíduos 196-219) se desloca⁴ mais de 2,0 Å, e a cadeia lateral da Tyr158, que no complexo binário está direcionada para o centro da cavidade de ligação do substrato, gira aproximadamente 60°. Ambas modificações ocorrem na cavidade de ligação de modo a acomodar melhor o substrato. Como as alças de ligação do substrato não estão envolvidas em contatos entre as subunidades, sua movimentação também não rompe com a organização do homotetrâmero. Esta flexibilidade e a independência das alças de ligação do substrato permitem ao sítio ativo uma certa adaptabilidade para acomodar substratos de vários tamanhos (Rozwarski *et al.*, 1999).

Apesar de experimentalmente não ter sido observada uma ordem certa de ligação do NADH e do substrato no sítio ativo da enzima, a pequena preferência pela ligação inicial da molécula de NADH seria facilmente compreendida. A ligação do substrato antes da entrada do NADH dificultaria a entrada subsequente deste NADH no sítio ativo da enzima.

Como o substrato assume uma conformação na forma de “U” dentro do sítio da InhA, a ligação dupla entre C2 e C3 a ser reduzida fica adjacente à posição C4 do anel nicotinamida do NADH. O carbono C3 do tioéster insaturado encontra-se mais próximo (4,8 Å) ao carbono C4 do NADH do que o carbono C2 (5,1 Å) (Figura 1.15). O anel nicotinamida do NADH está orientado de forma a permitir a transferência de seu H_{4S} ao carbono eletrofílico do substrato. Este arranjo espacial reforça o mecanismo proposto para a reação de redução, envolvendo a transferência do H_{4S} (na forma de hidreto) do NADH para o C3 do substrato (Figura 1.10).

De acordo com Rozwarski *et al.* (1999), considerando que a ligação dupla do substrato está muito próxima à volta da cadeia lipídica (C4-C8) pode-se especular que, durante a catálise do substrato, o sítio ativo deve impor uma conformação muito limitada ao substrato, podendo deformar o plano da ligação dupla e facilitar

⁴ RMSD calculado com o programa SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997) sobre os C α dos resíduos da alça de ligação do substrato (resíduos 196-219).

assim a conversão dos carbonos sp^2 planares em carbonos sp^3 tetraédricos (pelo tensionamento da ligação dupla *trans*) (Rozwarski *et al.*, 1999).

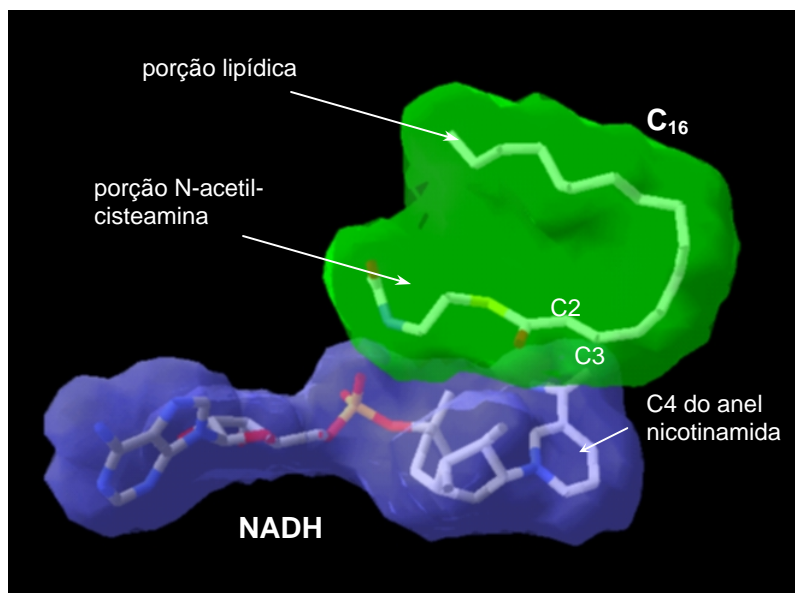


Figura 1.15. O substrato C₁₆ e cofator NADH na conformação que assumem dentro do sítio ativo da enzima InhA (subunidade C de 1BVR). Superfície molecular do NADH em azul do C₁₆ em verde. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A superfície do sítio ativo em torno do substrato é composta por aminoácidos apolares (representando 72% da área da cavidade de ligação do substrato), e por aminoácidos polares (representando 28% desta área – Figura 1.16).

A porção lipídica do substrato é completamente cercada de resíduos hidrofóbicos, a maioria deles pertencentes aos aminoácidos da alça de ligação do substrato (Ala198, Met199, Ala201, Ile202, Leu207, Ile215, Leu218). Os outros resíduos que também envolvem o substrato, mas não fazem parte da alça de ligação do substrato incluem Met103, Phe149, Met155, Tyr158 e Met161. A estrutura cristalina da InhA com o C₁₆ não apresenta moléculas de água próximas ao substrato. As poucas moléculas de água no sítio ativo estão envolvidas em interações com o NADH e aminoácidos de seu sítio de ligação.

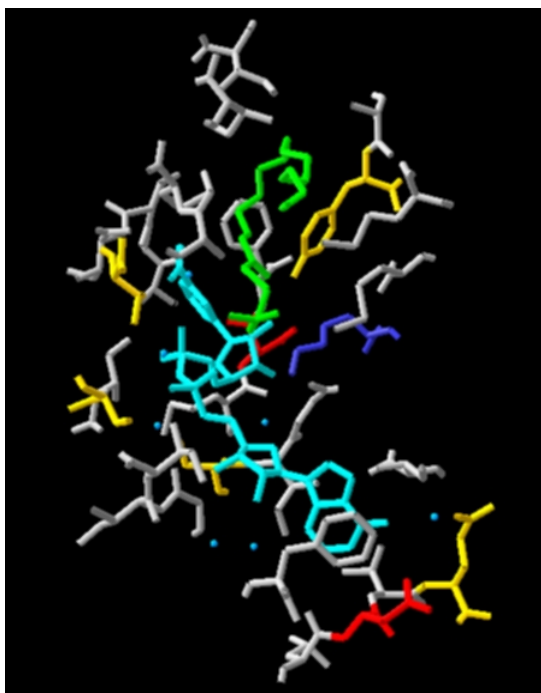


Figura 1.16. A predominância de aminoácidos não carregados no sítio de ligação do substrato C₁₆ (verde) e cofator NADH (ciano). Os aminoácidos coloridos de acordo com suas características: apolares em branco, polares em amarelo, básicos em azul e ácidos em vermelho. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A cadeia lateral da Phe 149, que se posiciona próximo ao anel nicotinamida do NADH, no fundo da cavidade de ligação, também ajuda na fixação da conformação em “U” do substrato. Mas o resíduo Phe149 parece ter um papel no mecanismo de ação da enzima mais importante do que simplesmente auxiliar nesta fixação da conformação do substrato no sítio ativo. A Phe149 esconde um canal que se estende até o outro lado da enzima. Na estrutura tridimensional da InhA com o C₁₆ e o NADH (1BVR) e também no complexo binário (1ENY), este canal possui moléculas de água bem ordenadas. Através dele seria possível a aproximação de uma molécula de água ao sítio ativo da enzima (Rozwarski *et al.*, 1999). Essa molécula de água poderia servir como a fonte de próton (H⁺) necessária para a redução do C2 da ligação dupla do tioéster (Figura 1.17).

A seqüência de fotos da Figura 1.17 mostra que o canal de água é perpendicular ao plano do anel fenila da Phe149, seguindo em direção à região central do homotetrâmero. Neste canal, de aproximadamente 12 Å de extensão, pode-se observar a presença de pelo menos duas moléculas de água, próximas ao sítio ativo (Figura 1.18).

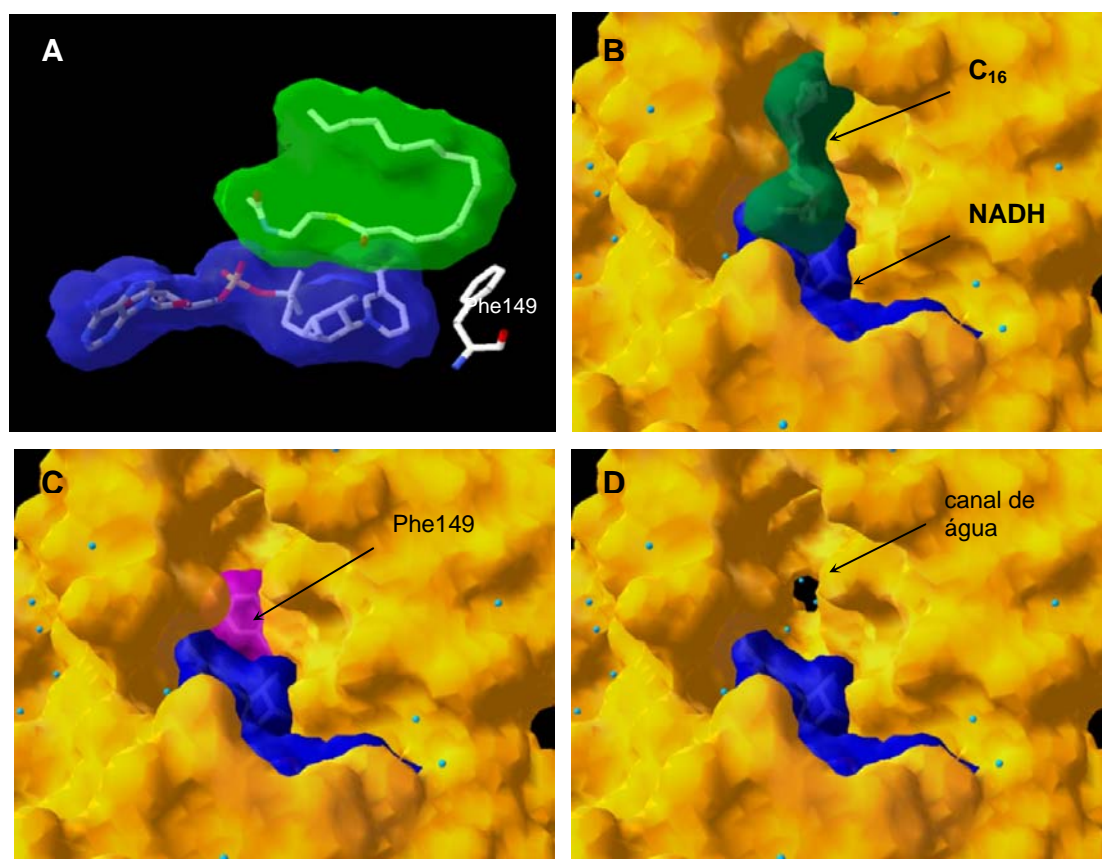


Figura 1.17. Posição da cadeia lateral da Phe149 dentro do sítio da InhA e o canal de água bloqueado por este resíduo. (A) Anel fenila da Phe149, que se posiciona próximo ao anel nicotinamida do NADH (superfície molecular em azul), no fundo da cavidade de ligação, auxiliando na fixação da conformação do substrato C₁₆ (superfície molecular em verde); (B) Vista aproximada do sítio ativo da InhA (superfície molecular em amarelo), com o NADH, o C₁₆ e moléculas de água (esferas em ciano); (C) Phe 149 (magenta) no fundo do sítio ativo da InhA (o C₁₆ foi deletado); (D) Canal de água que é aberto com a deleção da Phe149. Figuras geradas com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

O acesso destas águas ao substrato é bloqueado somente pelo anel fenila da Phe149, e a desobstrução deste canal pela rotação da cadeia lateral da Phe149 permitiria o acesso das moléculas de água ao substrato. A região central do homotetrâmero possui diversos resíduos arginina (resíduos carregados), que poderiam funcionar na regulação do acesso de moléculas de água ao interior do homotetrâmero, e ao canal de água próximo ao sítio ativo de cada subunidade. De fato, na estrutura cristalina do homotetrâmero (1BVR) esta cavidade central abriga uma grande quantidade de moléculas de água.

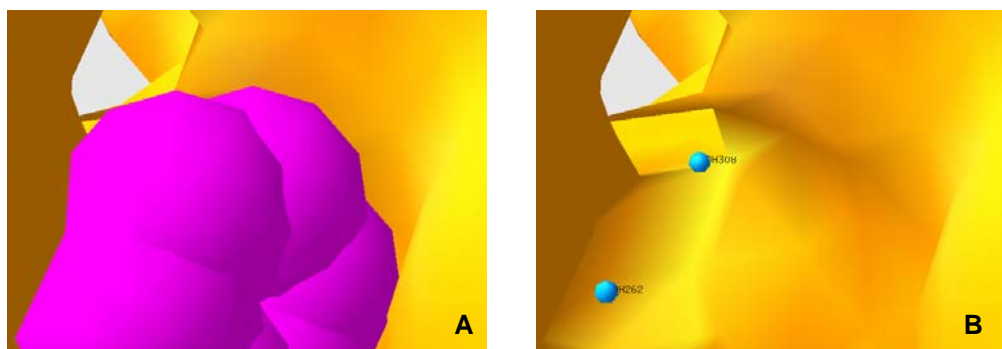


Figura 1.18. Moléculas de água no sítio ativo, próximas à ligação dupla do substrato. (A) Superfície de van der Waals da Phe149 (magenta) bloqueando o acesso das moléculas de água ao centro catalítico; (B) A retirada da Phe149 permite o acesso das moléculas de água (ciano) ao sítio catalítico. Figuras geradas com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Os resíduos Tyr e Lys do sítio ativo da enzima são altamente conservados e também parecem desempenhar um papel fundamental no mecanismo de ação enzimática. A família das enoil redutases possui um padrão de seqüência Tyr-(X)_n-Lys, onde n pode variar de 3 a 7 resíduos. Na InhA, esses resíduos: Tyr158 e Lys 165 são separados por 6 aminoácidos e ambos se encontram no sítio ativo da enzima. Baseado na primeira estrutura 3D da InhA (cristalizada com o NADH, 1ENY), propôs-se que a cadeia lateral da Lys165 seria responsável pela estabilização do enolato intermediário formado na redução do substrato (Quémard *et al.*, 1995). Entretanto, a investigação do mecanismo catalítico desta enzima por mutação sítio-direcionada de resíduos importantes para a catálise, mostrou que a substituição da Lys165 por aminoácidos apolares levava à inabilidade da enzima em se ligar ao NADH, indicando que o papel principal da Lys165 é na ligação do cofator. A substituição da Tyr158 por Phe ou Ala (dois aminoácidos apolares) resulta na diminuição da constante catalítica da reação (k_{cat}), não alterando o K_m do substrato, enquanto que a substituição por Ser não altera a atividade da enzima. Assim, a Tyr158 é que atua na estabilização do estado de transição da transferência de hidreto, fazendo uma ponte de hidrogênio com a carbonila do substrato (Parikh *et al.*, 1999). A estrutura do composto ternário 1BVR confirmou a participação da Tyr158 na associação com a carbonila do substrato, e da Lys165 na associação do cofator.

O modo de ligação do NADH é um fator importante, tanto para a reação quanto para a resistência. Apesar de seu sítio de ligação na InhA ser composto basicamente de aminoácidos hidrofóbicos, seu reconhecimento é mediado por interações com as cadeias laterais de alguns aminoácidos polares, e com átomos de oxigênio da carbonila e nitrogênio da amida das cadeias principais dos aminoácidos apolares. Estas interações são importantes para a fixação da conformação estendida do NADH dentro de seu sítio de ligação (Tabela 1.4).

Tabela 1.4. Pontes de hidrogênio do NADH em seu sítio de ligação na InhA.

Resíduo	Grupo do aminoácido	Átomo do NADH ^a	Distância (Å) ^b
Gly14	C=O	O'A3	2,75
Ser20	OH	OPA2	2,42
Ile21	N	OPN2	2,87
Asp64	COO ⁻	N6A	2,84
Val65	N	N1A	3,00
Ile95	C=O	O'N3	3,16
Gly96	N	O'A4	3,36
Lys165	NZ	O'N2	2,80
Lys165	NZ	O'N3	2,86
Ile194	C=O	N7N	3,38
Ile194	N	O7N	2,75
Thr196	C=O	OPN2	3,73

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1. ^bDistância entre os átomos pesados.

Os aminoácidos que também fazem parte da cavidade de ligação e estão até 4 Å do NADH são: Ile15, Ile16, Phe41, Leu63, Gln66, Ser94, Phe97, Ile122, Met147, Asp148, Phe149, Ala 191, Gly192, Pro193, Met199. Os resíduos Ser13, Ser19, Ala22, Ile47, His93, Val189, Ala190, Arg195 e Ala198 estão entre 4 e 5 Å do NADH e também delimitam o sítio de ligação.

Outras interações na região de ligação do cofator são importantes para a manutenção da conformação do sítio e ligação do NADH, entre elas uma série de pontes de hidrogênio mediadas por moléculas de água bem ordenadas. Algumas

destas moléculas de água fazem pontes de hidrogênio diretamente com o NADH (Tabela 1.5), enquanto outras interagem apenas com resíduos do sítio de ligação.

Tabela 1.5. Pontes de hidrogênio entre moléculas de água bem ordenadas e o NADH em seu sítio de ligação na InhA.

Molécula de água ^a	Átomo do NADH ^b	Distância (Å) ^c
WAT276	OPN2	2,87
WAT277	O'A2	2,62
WAT280	O7N	2,99
WAT295	O'N2	3,28
WAT298	N3A	3,20
WAT302	OPA2	3,13
WAT306	O'N3	3,26

^aNumeração das moléculas de água correspondente à numeração das águas utilizadas na simulação por dinâmica molecular da enzima *wt* InhA com as águas cristalográficas. ^bNome dos átomos da molécula de NADH de acordo com o Apêndice 1. ^cDistância entre os átomos pesados.

Essas interações entre enzima e cofator são importantes para a afinidade da InhA com o NADH. Conforme citado anteriormente, a substituição do resíduo Ser94 por Ala no sítio de ligação do NADH leva a um aumento de 5 vezes na constante de Michaelis do NADH (K_m), indicando uma menor afinidade da enzima pelo NADH (Dessen *et al.*, 1995).

A sobreposição das cadeias principais das estruturas cristalinas da proteína espécie selvagem (1ENY) e do mutante S94A resistente à INH (1ENZ - RMSD=0,34Å) mostra que a geometria dos sítios ativos da *wt* InhA e da InhA S94A, assim como a posição e conformação do NADH nestes sítios são semelhantes (Figura 1.19) e a única diferença significativa observada é na orientação da cadeia principal da Gly14.

Na estrutura da espécie selvagem, um oxigênio do grupo pirofosfato do NADH forma uma ponte de hidrogênio com o nitrogênio da cadeia principal da Ile21 (2,87 Å), e com uma molécula de água [WAT276] (2,87 Å). Essa molécula de água é bastante ordenada, pois também interage com a hidroxila da Ser94 (3,0 Å), com a carbonila da Gly14 (2,93 Å) e com os nitrogênios das cadeias principais da Ala22 (2,36 Å) e Ile21 (2,82 Å).

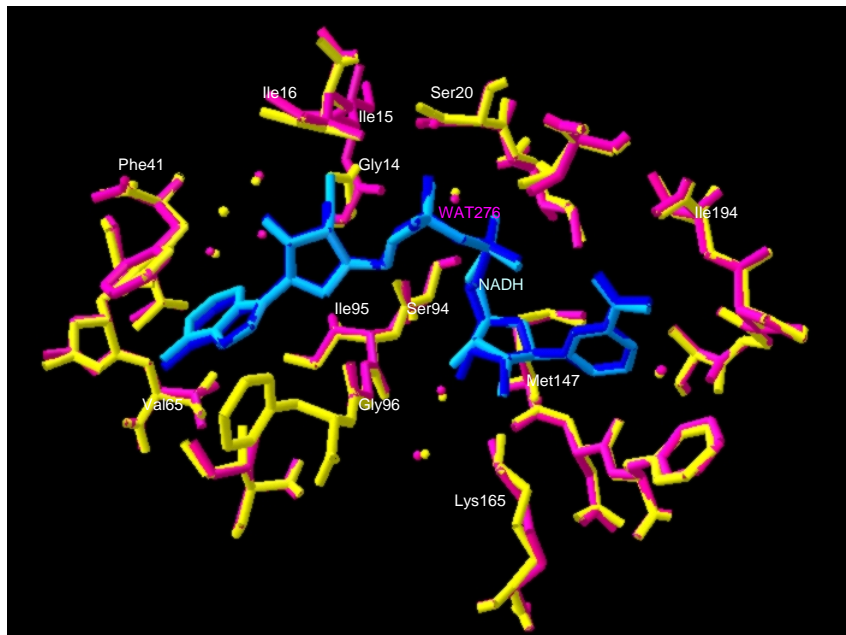


Figura 1.19. Sobreposição dos sítios de ligação do NADH na *wt* InhA (1ENY) e na S94A InhA (1ENZ). Resíduos da 1ENY em magenta, e o NADH em azul. Resíduos da 1ENZ em amarelo e NADH em ciano. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

No mutante S94A, o grupo carbonila da Gly14 gira em relação ao nucleotídeo, rompendo a ponte de hidrogênio com a molécula de água feita na estrutura da espécie selvagem. Na estrutura cristalina do mutante S94A, esta molécula de água não está tão ordenada, e forma somente uma ponte de hidrogênio com o oxigênio do pirofosfato do NADH (3,05 Å) (Pantano *et al.*, 2002) (Figura 1.20)

A mutação S94A causa uma pequena aproximação dos OH's da ribose A com o sítio ativo da proteína, levando a uma maior interação com a carbonila da Gly14. Como a Gly é um aminoácido não ramificado, ela é bastante flexível e pode arranjar-se de forma a possibilitar a maior interação possível entre o OH da ribose A com a C=O da Gly14. De fato, a carbonila da cadeia principal da Gly14 sofre uma rotação de aproximadamente 80° no ângulo ψ (psi), aproximando seu oxigênio carbonílico do OH da ribose.

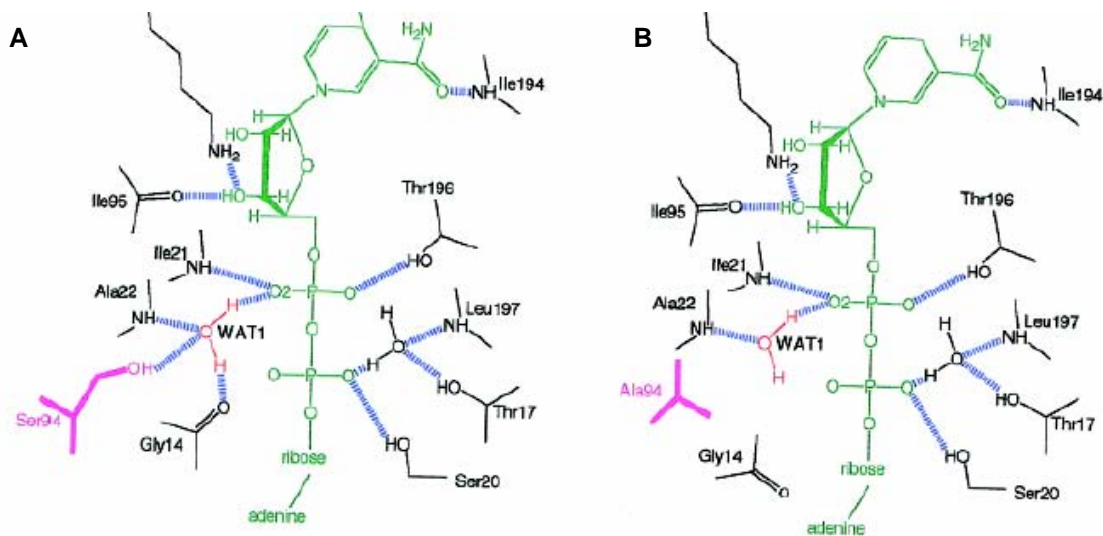


Figura 1.20. Comparação das interações da porção central do NADH com seu sítio de ligação nas estruturas da *wt* InhA 1ENY. (A) e na S94A InhA 1ENZ (B). Resíduos do sítio de ligação em preto, NADH em verde, pontes de hidrogênio em azul. Os resíduos mutados são destacados em magenta. A molécula de água WAT1 corresponde à molécula de água WAT276 no desenho acima. Retirado de Pantano *et al.*, 2002.

De acordo com Dessen *et al.*, esta mudança conformacional e a perturbação no padrão da rede de pontes de hidrogênio no sítio do NADH seria responsável pela menor afinidade do mutante S94A pelo NADH, o que afetaria diretamente a subsequente ligação do complexo inibidor NAD-INH (Dessen *et al.*, 1995).

A determinação da estrutura cristalina do complexo inibidor NAD-INH com a InhA (1ZID) mostrou que sua estrutura é isomorfa àquela cristalizada somente com o NADH (1ENY). Nela o anel acil-isonicotínico da INH está covalentemente ligado ao carbono C4 do anel nicotinamida do NADH, dentro do sítio ativo. Esse grupo substitui estereoespecificamente o hidrogênio 4S do C4 do NADH, que seria o H⁺ a ser transferido no processo de redução do enoil-ACP, e ocupa parte da região de ligação do substrato, causando um impedimento estérico e evitando a aproximação e ligação do mesmo (Figura 1.21). Estudos de modelagem molecular demonstraram que o aduto enantiomérico 4*R* é estereoquimicamente proibido (Scior *et al.*, 2002).

Na comparação do sítio ativo da 1ENY com a 1ZID não são observadas grandes diferenças conformacionais no sítio ativo da enzima (RMS = 0,45 Å). Entretanto, na enzima nativa, o NADH liga-se à base da cavidade, e a cadeia lateral da Phe149 posiciona-se sobre o anel nicotinamida, parecendo proteger a porção

reativa do NADH do solvente. Na presença do complexo inibidor, a cadeia lateral da Phe149 se afasta do anel, girando aproximadamente 90° e criando espaço para o anel acil-isonicotínico. Com isso, a cadeia lateral da Phe 149 se torna adjacente ao anel acil-isonicotínico, permitindo uma interação do tipo π -stacking do anel aromático do aminoácido com o anel piridina do grupo acil-isonicotínico (Figura 1.21). Este bisubstrato interage com grande afinidade e de forma não covalente com a enzima espécie selvagem. A interação com a Phe149 deve aumentar a afinidade da enzima pelo substrato inibidor, em comparação com somente o NADH. Estudos recentes demonstraram que este complexo NAD-INH é um inibidor competitivo que se associa fortemente à enzima, causando mudanças conformacionais (Rawat *et al.*, 2003).

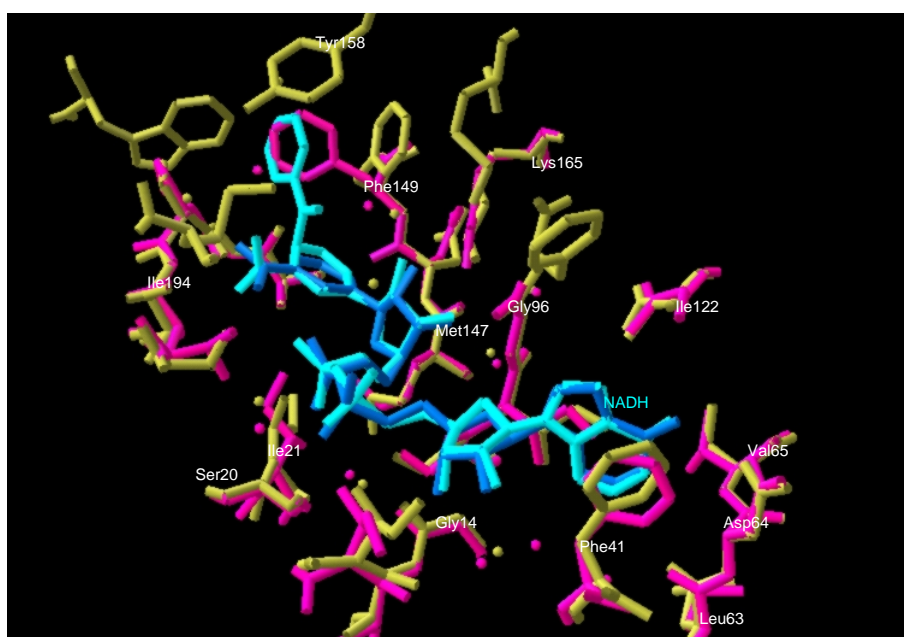


Figura 1.21. Sobreposição dos sítios ativos da *wt* InhA associada ao NADH (1ENY) e associada ao complexo inibidor NAD-INH (1ZID). 1ENY: aminoácido em magenta e NADH em azul; 1ZID: aminoácidos em dourado e NAD-IHN em ciano. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

O triclosan (TCL – Figura 1.7), um antibacteriano de espectro amplo, é um inibidor da InhA na escala de submicromolar (Parikh *et al.*, 2000), conhecido também por inibir as enoil redutases de outros organismos como *E. coli*, *B. napus* e *P. falciparum*. Seu modo de ligação com a ENR de *E. coli* e *B. napus* é conhecido e esperava-se que na InhA, que possui identidade de seqüência na região de ligação

do TCL com estas duas ENR, o TCL também se ligasse da mesma forma (Schroeder *et al.*, 2002). A determinação recente da estrutura tridimensional do complexo ternário formado pela InhA, NAD⁺ e TCL confirmou esta expectativa (PDB id: 1P45 - Kuo *et al.*, 2003) (Figura 1.22A).

A estrutura de uma das subunidades de 1P45 mostrou que a inibição da InhA por TCL ocorre de forma diferente da inibição por INH. Na inibição por TCL, este composto interage com a enzima e com o NAD⁺ de forma não covalente. O anel do TCL que contém o grupo hidroxila (anel A - Figura 1.7) faz uma interação do tipo π -stacking com o anel nicotinamida aromático do NAD⁺, e o grupo OH faz ponte de hidrogênio com uma das hidroxilas da ribose N do cofator e com a Tyr158 do sítio ativo da enzima, como anteriormente previsto por estudos de modelagem e simulação por dinâmica molecular (Schroeder *et al.*, 2002). O anel diclorofenila do TCL (anel B - Figura 1.7) fica ortogonal ao anel A (ângulo diedro entre os dois anéis fenila de aproximadamente 90°), e os átomos de cloro se projetam em direção ao solvente (Figura 1.22A).

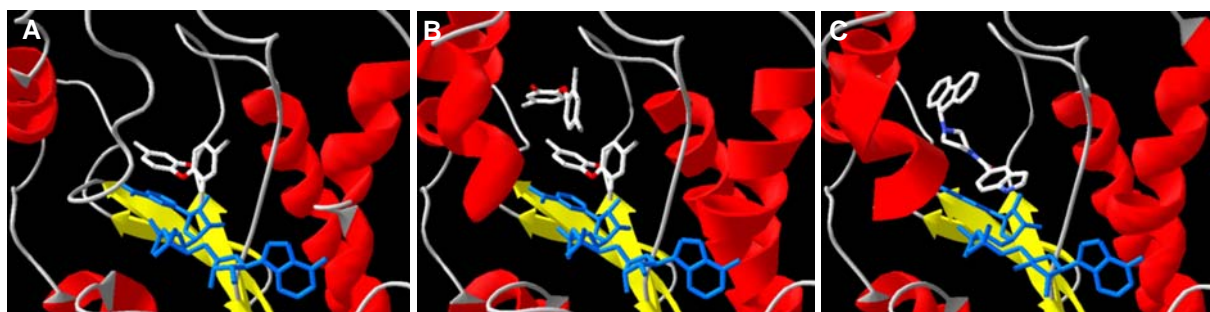


Figura 1.22. Inibição da InhA por TCL e GEQ. (A) Sítio da *wt* InhA associada ao NAD⁺ (azul) e a uma molécula de TCL (1P45); (B) Sítio da *wt* InhA associada ao NAD⁺ (azul) e a duas moléculas de TCL (1P45); (C) Sítio da *wt* InhA associada ao NAD⁺ (azul) e ao inibidor GEQ, um aril-indol-piperazina (1P44). Enzima InhA representada em relação a sua estrutura secundária: α -hélices em vermelho, fitas β em amarelo e alças em cinza. Figuras geradas com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A segunda subunidade da estrutura 1P45 possui duas moléculas de TCL ligadas à InhA. Nesta subunidade, o primeiro TCL se liga de forma idêntica a quando está sozinho no sítio, e a segunda molécula de TCL encontra-se invertida

em relação à primeira, ocupando uma área quase completamente hidrofóbica próxima à alça de ligação do substrato (Figura 1.22B).

A procura por inibidores da InhA análogos ao TCL identificou como composto líder uma aril-indol-piperazina (GEQ) que também foi cristalizada dentro do sítio da enzima. A grande molécula deste inibidor ocupa a mesma região das duas moléculas de TCL, e seu grupo carbonila também faz ponte de hidrogênio com uma das hidroxilas da ribose N do NAD⁺, e com a Tyr158 (Figura 1.22C).

Tanto as moléculas de TCL quanto o GEQ não se ligam covalentemente ao NAD⁺, mas interagem com ele dentro do sítio da enzima. A substituição de aminoácidos que perturbem o sítio de ligação do cofator (como, por exemplo, I194T e A239T) também diminui a sensibilidade da InhA por estes compostos.

Portanto, independente da formação ou não de ligação covalente entre inibidor e NAD(H) dentro do sítio ativo da enzima, a inibição da InhA tanto por INH, quanto por TCL e GEQ, não depende somente das interações da enzima alvo com estas moléculas, mas também com o cofator. Assim, mutações no sítio de ligação do NAD(H) efetivamente diminuem a afinidade da enzima pelo complexo inibidor.

Objetivos

Este trabalho tem como objetivo o estudo computacional da enzima 2-*trans*-Enoil-ACP (CoA) Redutase, produto do gene *inhA* de MTB, para o entendimento dos eventos moleculares responsáveis por sua associação com o cofator NADH e por seu mecanismo de resistência à droga INH.

Para o desenvolvimento deste trabalho, foram propostas as seguintes etapas:

- Simulação da dinâmica molecular da InhA espécie selvagem (*wt* InhA) na sua forma livre e na forma ligada ao NADH;
- Construção dos mutantes clínicos resistentes à INH S94A, I16T, I21V e I47T, via modelagem molecular, e simulação por dinâmica molecular destes mutantes associados ao NADH.
- Avaliação teórica da afinidade enzimas - NADH por experimentos de *docking* molecular.

A comparação entre a enzima InhA espécie selvagem e seus mutantes em relação à associação com o NADH nos possibilita uma melhor compreensão dos fatores relacionados com a afinidade com o NADH e a resistência à INH. A utilização do conjunto de estruturas geradas nas simulações pela dinâmica molecular, para a avaliação teórica da afinidade das enzimas pelo NADH, nos permitem a consideração da grande flexibilidade dos sistemas moleculares em estudo.

A enzima InhA desempenha uma função única e faz parte da via biossintética de ácidos micólicos, essenciais para o MTB. Portanto, ela representa um bom alvo para o desenvolvimento de novos inibidores. O entendimento dos eventos moleculares responsáveis por sua associação com o cofator e por seu mecanismo de resistência deverá contribuir para o desenho racional de novos fármacos anti-tuberculose.

Materiais e Métodos

3.1- Simulação por dinâmica molecular de sistemas biológicos

O estudo de sistemas biológicos inicialmente esteve limitado à observação e interpretação de dados obtidos por técnicas experimentais que analisam amostras macroscópicas. Essas informações retratam o comportamento médio de conjuntos de átomos e moléculas. Com o avanço contínuo de técnicas experimentais como difração de raios X, ressonância magnética nuclear e outros métodos espectroscópicos, tornou-se possível uma visão mais detalhada de diversos processos biológicos pelo acesso a propriedades atômicas de macromoléculas biológicas, como proteínas e ácidos nucleicos (Šali, 1998; Wüthrich, 2000).

Informações estruturais, energéticas e dinâmicas, em nível atômico, de biomoléculas de grande mobilidade geralmente não são acessíveis através de técnicas experimentais. Nesses sistemas pode-se utilizar métodos teóricos para a investigação de seus aspectos moleculares. Se os métodos teóricos forem suficientemente adequados para prever as propriedades de sistemas biológicos, eles podem ser utilizados no estudo de propriedades moleculares com uma resolução espacial, energética e temporal que não poderia ser alcançada experimentalmente (van Gunsteren & Mark, 1992).

Em termos quântico-mecânicos, as leis que governam o comportamento atômico são conhecidas. Entretanto, sistemas biomoleculares são muito complexos, e a resolução da equação de Schrödinger para sistemas de múltiplos átomos é impraticável. Assim, utiliza-se a mecânica clássica associada a funções semi-empíricas que descrevem as interações entre os átomos do sistema molecular. A mecânica clássica descreve adequadamente a maioria das propriedades dos sistemas biológicos, e o método da simulação por dinâmica

molecular é a abordagem teórica mais apropriada para esses estudos. (Karplus, 2002).

Na simulação por dinâmica molecular (DM), as equações de movimento de Newton que governam o movimento de sistemas de corpos múltiplos em função do tempo são numericamente resolvidas, de acordo com as condições de contorno apropriadas para a geometria e simetria do sistema em estudo. Fundamentada nos princípios básicos da mecânica clássica, a DM permite o entendimento do comportamento dinâmico microscópico dos átomos que compõem um sistema. Além de permitir esse panorama dinâmico microscópico, a DM também pode ser empregada como forma de representar e determinar as propriedades de equilíbrio do sistema, como: grandezas termodinâmicas (pressão, temperatura, volume, etc), estrutura, e energia livre.

Para descrever o comportamento microscópico de um sistema a partir das leis da mecânica clássica, a DM necessita, como entrada, uma descrição das interações interatômicas. A qualidade dos resultados obtidos na DM depende da qualidade dessa descrição. A abordagem mais utilizada envolve a utilização de um modelo, ou campo de força. A simplificação resultante da utilização de um campo de força padrão reduz consideravelmente o custo computacional e permite que estes cálculos sejam feitos para proteínas, membranas e sistemas biológicos maiores (Tuckerman & Martyna, 2000).

3.1.1 – Fundamentos da dinâmica molecular

Na simulação por DM, uma trajetória (configurações moleculares em função do tempo) de um sistema molecular é gerada pela integração simultânea da equação do movimento de Newton para todos os átomos do sistema molecular:

$$\frac{dv_i(t)}{dt} = m_i^{-1} \times F_i(t) \quad (1)$$

$$\frac{d\mathbf{r}_i(t)}{dt} = \mathbf{v}_i(t) \quad (2)$$

Onde m_i é a massa da partícula i , \mathbf{F}_i é a força exercida sobre esta partícula, \mathbf{v}_i é a velocidade, e \mathbf{r}_i a posição da partícula i .

Assim, o comportamento dinâmico de um sistema de N partículas pode ser calculado pela solução da equação diferencial de segunda ordem:

$$\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}_i(r_1, r_2, \dots, r_N) \times m_i^{-1} \quad (3)$$

A força \mathbf{F}_i exercida pelas outras partículas do sistema sobre uma partícula i também é dada pelo negativo do gradiente da função de interação atômica (V), que depende das coordenadas de todas as N partículas do sistema (van Gunsteren & Berendsen, 1990; van Gunsteren & Mark, 1992):

$$\mathbf{F}_i(t) = -\frac{\partial V[\mathbf{r}_1(t), \mathbf{r}_2(t), \dots, \mathbf{r}_N(t)]}{\partial \mathbf{r}} \quad (4)$$

Como sistemas biomoleculares são formados por inúmeros átomos (na ordem de centenas ou milhares), estas equações não podem ser resolvidas analiticamente. Sob estas circunstâncias, as equações do movimento devem ser resolvidas numericamente. Assim, a integração é feita em pequenos passos⁵, cada um separado no tempo por um Δt . A força \mathbf{F}_i sobre cada átomo na configuração no tempo t é calculada como a soma dos vetores de suas interações com os outros átomos do sistema, e é considerada constante no intervalo de tempo entre os dois passos. A partir da força, podemos calcular a aceleração dos átomos, que são combinados com suas posições e velocidades no tempo t , para então calcular as posições e velocidades no tempo $t+\Delta t$. As forças que atuam nos átomos em suas novas posições são então determinadas, levando a novas posições e velocidades

⁵ Para uma resolução numérica mais eficiente e precisa das equações da DM, a variação de tempo Δt deve ser bem menor do que a menor frequência das vibrações do sistema (normalmente a frequência de vibração dos átomos de hidrogênio, o movimento molecular mais rápido). Geralmente, aplica-se um passo de tempo de 0,5 a 2 femtosegundos (1 femtosegundo = 10^{-15} segundos).

no tempo $t+2\Delta t$, e assim por diante. A estabilidade da dinâmica é assim favorecida: os átomos seguem mais precisamente suas trajetórias clássicas, e a energia total do sistema tende a ser conservada.

Quando a energia potencial é uma função contínua das posições e o passo de tempo é suficientemente pequeno para considerar-se que as posições variem suavemente com o tempo, dado um conjunto de posições atômicas num instante t , as posições no passo seguinte e anterior podem ser obtidas por uma expansão de Taylor de $\mathbf{r}(t)$:

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \mathbf{a}(t)\Delta t^2/2 + \dots \quad (5)$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \mathbf{a}(t)\Delta t^2/2 - \dots \quad (6)$$

Diferentes algoritmos de integração podem ser utilizados para a resolução desta equação diferencial, mas o mais utilizado é o algoritmo de Verlet (Verlet, 1967). O algoritmo de Verlet para a propagação das posições das partículas do sistema, que corresponde ao somatório das equações 5 e 6, desprezando os termos de ordem superiores é dado por:

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \mathbf{a}(t)\Delta t^2 \quad (7)$$

No algoritmo de Verlet para a propagação das posições, a velocidade é eliminada⁶, e a cada passo de tempo a aceleração $\mathbf{a}(t)$ é avaliada a partir da força F_i (de acordo com as equações 3 e 4). As equações de movimento de Newton são calculadas para cada partícula do sistema, a cada incremento de tempo. Assim, as posições atômicas se propagam no tempo, gerando uma trajetória.

É importante notar que durante o intervalo Δt , a aceleração é considerada constante, e, a menos que passos infinitesimais sejam considerados, isso pode representar uma fonte de erro. Quanto menor o passo de tempo Δt , mais a trajetória se aproxima da trajetória “real”. Entretanto, um passo menor significa um maior número de etapas para cobrir um certo tempo total, elevando muito o custo

⁶ Neste algoritmo, a velocidade não é determinante da trajetória das partículas, entretanto, ela pode ser estimada pela equação $\mathbf{v}(t) = (\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)) / 2\Delta t$ que corresponde à diferença entre as equações 5 e 6.

computacional (que é inversamente proporcional ao tamanho do passo de tempo). Uma alternativa é a utilização de algoritmos específicos que eliminem os movimentos de alta frequência, tornando possível a utilização de passos de tempo um pouco maiores⁷.

Uma das desvantagens do algoritmo de Verlet é que um incremento pequeno Δt leva a diferenças grandes nas posições, podendo levar a imprecisões na trajetória. Além disso, neste algoritmo, a velocidade não é calculada explicitamente, o que é um problema para o cálculo da temperatura e energia cinética do sistema (Jensen, 1999). Estes problemas podem ser resolvidos pela utilização do algoritmo *leap-frog* (ou meio-passo, uma variação do algoritmo de Verlet), desenvolvido para melhorar a avaliação das velocidades. Nele, as velocidades são atualizadas em tempos intermediários àqueles da atualização das posições. As velocidades nos meio-passos são então definidas como:

$$\mathbf{v}(t+\Delta t/2) = [\mathbf{r}(t+\Delta t) - \mathbf{r}(t)]/\Delta t \quad (8)$$

e

$$\mathbf{v}(t-\Delta t/2) = [\mathbf{r}(t) - \mathbf{r}(t-\Delta t)]/\Delta t \quad (9)$$

Desta forma, a propagação das posições pode ser obtida diretamente pelo rearranjo da equação 8:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t+\Delta t/2)\Delta t \quad (10)$$

Onde a atualização das posições agora depende das velocidades $\mathbf{v}_i(t+\Delta t/2)$. O cálculo das velocidades pode ser feito pela aproximação da equação 1 para pequenos intervalos de tempo (van Gunsteren & Mark, 1992):

$$\mathbf{v}_i(t+\Delta t/2) = \mathbf{v}_i(t-\Delta t/2) + m_i^{-1}\mathbf{F}_i[\mathbf{r}_1(t),\mathbf{r}_2(t),\dots,\mathbf{r}_N(t)]\Delta t \quad (11)$$

⁷ Como algumas coordenadas do sistema variam muito rapidamente em função do tempo (como vibrações de ligações, vibrações angulares, etc.), é necessário utilizar pequenos passos de tempo para observá-las. Se Δt for superior a poucos femtosegundos, a solução pode acumular erros exponenciais. Para a simplificação dos cálculos computacionais são utilizados algoritmos que restringem as coordenadas internas (ligações, ângulos, etc.) a seus valores ideais, como o algoritmo SHAKE (Ryckaert *et al.*, 1977; van Gunsteren & Berendsen, 1977). Essa restrição elimina os movimentos de alta frequência do sistema a ser integrado e geram trajetórias aproximadas. Esta aproximação não é prejudicial, uma vez que os movimentos de alta frequência são pouco relevantes em trajetórias longas (Elber *et al.*, 2002).

Desta forma, as atualizações das posições e velocidades são feitas fora de fase, numa diferença de $\Delta t/2$ ⁸. Evitando o cálculo da velocidade e da aceleração no mesmo momento, os erros introduzidos pelo truncamento da expressão são minimizados. Além de sua menor suscetibilidade a erros numéricos, o *leap-frog* ocupa menos espaço de armazenamento na memória, e é um dos algoritmos mais simples, estáveis e eficientes aplicados a sistemas biomoleculares. Uma de suas maiores vantagens é a utilização das velocidades para o cálculo das novas posições atômicas, o que permite o acoplamento do sistema a um banho térmico (pela correção das velocidades).

Como a energia potencial é uma função das posições, e a variação destas posições depende das velocidades, o controle das velocidades representa, além do controle da energia cinética, também o controle da energia potencial – e conseqüentemente, o controle da energia total do sistema.

Como as equações do movimento consideram que não há forças externas atuando sobre o sistema, a simulação padrão por DM gera um *ensemble* microcanônico⁹ (NVE), isto é, na simulação de um sistema isolado de N átomos e com um volume V , sua energia total se conserva, e sua trajetória se propaga em uma superfície de energia constante, enquanto sua temperatura e pressão podem flutuar. Entretanto, muitas vezes é desejável que a simulação de um sistema biológico seja feita a temperatura constante (*ensemble* canônico¹⁰) ou a temperatura e pressão constantes (*ensemble* isobárico-isotérmico¹¹). Para isso, deve ser feito um controle de temperatura e pressão do sistema (van Gunsteren & Berendsen, 1990).

O valor da temperatura ($T(t)$) a cada instante da simulação por DM é dado pela média das energias cinéticas sobre todos os átomos do sistema:

⁸ O intervalo de tempo Δt utilizado em sistemas moleculares normalmente está entre 1 e 10 fs.

⁹ **Ensemble microcanônico**: estado termodinâmico que corresponde a um sistema isolado, e é caracterizado por número de átomos N , volume V e energia total E constantes (NVE).

¹⁰ **Ensemble canônico**: estado termodinâmico caracterizado por número de átomos N , volume V e temperatura T constantes (NVT).

¹¹ **Ensemble isobárico-isotérmico**: estado termodinâmico caracterizado por número de átomos N , pressão P e temperatura T constantes (NPT).

$$\frac{1}{2} \sum_{i=1}^N m_i \langle v_i^2 \rangle = \frac{3}{2} N k_B T \quad (12)$$

onde N é o número de átomos do sistema, $\langle v_i^2 \rangle$ é a média do quadrado das velocidades do átomo i , e k_B é a constante de Boltzmann.

Se esta temperatura for diferente da temperatura desejada (T_o), todas as velocidades podem ser re-escaladas por um fator $(T_o / T(t))^{1/2}$ em cada uma das etapas, para a obtenção da temperatura desejada. Outra alternativa é o acoplamento do sistema ao chamado “banho térmico” (Berendsen *et al.*, 1984), que gradualmente adiciona ou remove energia do sistema. A energia cinética do sistema é modificada pela multiplicação da velocidade por um fator λ que depende de T_o , $T(t)$ e de uma taxa de transferência de calor τ .

$$\lambda = \left[1 + \frac{\Delta t}{\tau} \left(\frac{T_o}{T(t)} - 1 \right) \right]^{1/2} \quad (13)$$

Procedimento análogo pode ser feito para a simulação de um sistema à pressão constante (Berendsen *et al.*, 1984; Jensen, 1999).

3.1.2 – O campo de força

A simulação por DM (ou simplesmente uma minimização de energia) requer a escolha de uma forma adequada para se tratar a função de energia potencial $V[\mathbf{r}_1(t), \mathbf{r}_2(t), \dots, \mathbf{r}_N(t)]$ na equação 4, e de parâmetros adequados para utilizá-la. Ela deve ser simples o suficiente para permitir a simulação de propriedades de moléculas complexas e precisa o suficiente para estimar adequadamente as configurações e energias do sistema (Cornell *et al.*, 1991).

A função que descreve a energia potencial – e, portanto, a interação entre os diversos átomos de um sistema molecular – juntamente com seus parâmetros é chamada de campo de força.

Um campo de força é composto de expressões que representam as forças intra e intermoleculares do sistema em questão. A associação de funções simples da mecânica molecular que descrevem as interações entre partículas de um sistema, de dados empíricos e de cálculos de mecânica quântica de pequenos sistemas modelo permite o cálculo da energia potencial de grandes sistemas moleculares em função de suas coordenadas atômicas (Leach, 1996; Weiner *et al.*, 1984).

Diferentes campos de força foram desenvolvidos e implementados em programas¹² para as mais diversas aplicações (Jensen, 1999). Como regra, todos possuem a mesma estrutura básica para a função $V[\mathbf{r}_1(t), \mathbf{r}_2(t), \dots, \mathbf{r}_N(t)]$, mas diferem entre si pelos valores de seus parâmetros. As funções que descrevem os campos de força são compostas por termos independentes, que representam a vibração linear através de uma ligação, a vibração angular e as interações de ângulos diedros, e as interações não-ligantes entre os pares de átomos, como as interações de van der Waals e interações eletrostáticas.

As simulações por dinâmica molecular realizadas neste trabalho utilizam a função de energia potencial do pacote AMBER (Pearlman *et al.*, 1995), que possui a seguinte forma (Cornell *et al.*, 1995):

$$\begin{aligned}
 V(r_1, r_2, \dots, r_N) = & \sum_{bonds} K_r (r - r_{eq})^2 & (14) \\
 & + \sum_{angle} K_\theta (\theta - \theta_{eq})^2 \\
 & + \sum_{dihedrals} \frac{V_m}{2} [1 + \cos(n\phi - \gamma)]
 \end{aligned}$$

¹² Diversos programas de simulação por dinâmica molecular foram desenvolvidos para a simulação de biomoléculas, como: AMBER (Pearlman *et al.*, 1995), CHARMM (Brooks *et al.*, 1983), GROMOS (van Gunsteren & Berendsen, 1987), OPLS (Jorgensen & Tirado-Rives, 1988; Jorgensen *et al.*, 1990), e outros.

$$+ \sum_{i < j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon r_{ij}} \right)$$

+ termos adicionais

onde r_1, r_2, \dots, r_N representam as coordenadas cartesianas de todos os átomos do sistema. Nesta equação, os dois primeiros termos representam as contribuições dos modos de vibração ao longo dos eixos das ligações covalentes e deformações angulares do sistema. Eles descrevem o desvio em relação ao comprimento de ligação ideal (de equilíbrio) r_{eq} e ao ângulo ideal θ_{eq} . Como em temperaturas ordinárias os comprimentos das ligações químicas e oscilações angulares variam próximos a seus valores de equilíbrio, eles obedecem a funções potenciais que podem ser descritas pela lei de Hooke para um sistema de massas unidas por molas, onde K_r é a constante de Hooke associada a uma ligação química específica, e K_θ ao ângulo formado entre duas ligações consecutivas. Os valores de r_{eq} e θ_{eq} dependem do tipo dos átomos e de suas hibridizações, e podem ser obtidos a partir de dados de espectroscopia de raios X e microondas, enquanto que as constantes de força K_r e K_θ , que determinam a flexibilidade da ligação e do ângulo, podem ser avaliados a partir das frequências no infravermelho e cálculos de mecânica quântica, ou pela comparação de cálculos dos modos normais de vibração com as frequências observadas (Karplus & Petsko, 1990).

O terceiro termo da equação 14 acrescenta a contribuição das forças relacionadas aos ângulos de torção (ângulos diedros), e representa a energia de deformação da rotação em relação a um eixo formado entre dois átomos covalentemente ligados, com uma constante torsional K_ϕ . Neste termo, n é o número de mínimos para a torção de uma ligação química específica (periodicidade), ϕ é o ângulo diedro, e γ é a fase deste potencial torsional. Este potencial rotacional é considerado periódico e é modelado pelo cosseno, ou o somatório sobre todas as funções cosseno (Leach, 1996).

O quarto termo da equação 14 representa a contribuição das interações entre átomos não ligados, e é composto de três partes: (i) o termo de repulsão das nuvens eletrônicas, que evita a interpenetração de átomos muito próximos; (ii) um

termo de atração relativo às forças de dispersão de London entre átomos; e (iii) um termo eletrostático de Coulomb, que pode ser atrativo ou repulsivo, dependendo se as cargas atômicas parciais q_i e q_j têm sinais opostos ou iguais. A combinação dos dois primeiros termos das interações não-ligantes corresponde ao potencial de Lennard-Jones 12-6, que tem seu mínimo a uma distância interatômica igual à soma dos raios de van der Waals dos átomos. Os parâmetros A e B dependem da natureza dos átomos envolvidos, e são determinados por métodos experimentais e cálculos teóricos (Karplus & Petsko, 1990).

A expressão do campo de força utilizado em versões anteriores do pacote AMBER ainda apresentava explicitamente um termo que modelava a função de interação do tipo pontes de hidrogênio (Weiner *et al.*, 1984) e interações não-ligantes para átomos distantes entre si por menos de 4 ligações. Entretanto, na versão atual, estes tipos de interações estão implicitamente contempladas pela atribuição de parâmetros mais exatos e da utilização de um fator de escalonamento (Cornell *et al.*, 1995).

Em uma simulação por DM, a maior parte do tempo da avaliação da energia potencial é utilizado para o cálculo das interações entre pares de átomos não ligados covalentemente, e, portanto, cresce exponencialmente com o número de átomos do sistema. Em princípio, este tipo de interação deveria ser avaliado para todos os pares de átomos, mas o custo computacional seria impraticável, principalmente devido ao grande tamanho dos sistemas biomoleculares e, recentemente, à consideração explícita das moléculas de solvente. Por isso, considerações práticas geralmente devem ser aplicadas para a simplificação deste problema, principalmente em simulações muito longas.

A grandes distâncias, as interações entre átomos não-ligantes se aproximam assintoticamente de zero. Portanto é possível aumentar a eficiência computacional ignorando estas interações em distâncias relativamente grandes, utilizando um raio de corte para as interações não-ligantes. Desta forma, as interações entre dois átomos separados por uma distância maior do que o raio de corte são desprezadas (van Gunsteren & Berendsen, 1990).

Para as interações de van der Waals, a desconsideração de interações a mais de 9 Å é razoável, uma vez que nestas distâncias o termo de van der Waals é apenas alguns décimos de seu valor máximo (Figura 3.1A). Entretanto, as

interações eletrostáticas, que diminuem numa proporção de $1/r$, são ainda significativas até em distâncias mais longas (Figura 3.1B), e o simples truncamento do cálculo destas interações além do raio de corte pode resultar em distorções graves de várias propriedades do sistema molecular¹³ (Smith & Pettit, 1991; Schreiber & Steinhauser, 1992a e 1992b).

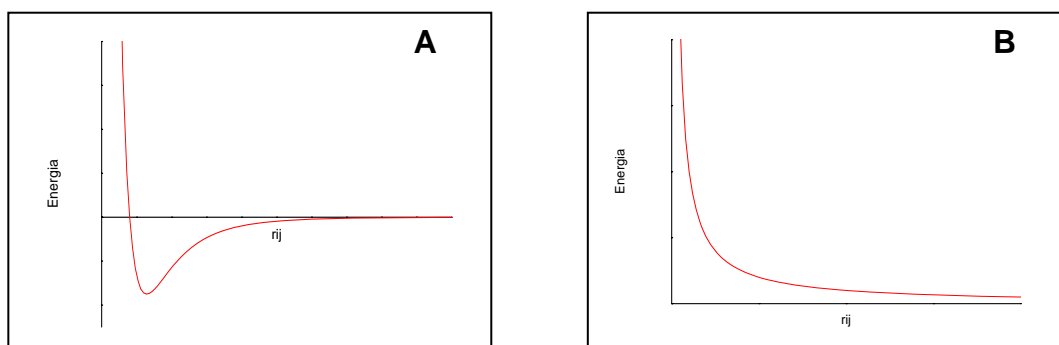


Figura 3.1. Representação da variação da energia potencial para as interações de van der Waals (A) e de Coulomb (B) em função da distância interatômica r_{ij} .

Uma alternativa para a redução desses problemas pode ser a utilização de funções que induzam as interações não-ligantes a se aproximarem de zero de uma forma lenta e contínua. Exemplos destas funções são as chamadas *shifting function* (que lentamente leva a zero essas interações no raio de corte) e *switching function* (onde estas interações são lentamente levadas à zero após uma certa distância r_{on} , até o raio de corte r_{off}) (Brooks *et al.*, 1983; Tasaki *et al.*, 1993).

O efeito do truncamento também pode ser diminuído pelo acoplamento de um segundo raio de corte mais externo (*twin-range cutoff*). Neste método, as interações dentro do raio de corte menor são calculadas a cada etapa da simulação, enquanto as interações dentro da faixa entre o raio de corte menor e um raio de corte maior são calculadas a cada 5-10 etapas, e são consideradas constantes durante este período. Neste caso uma região maior é considerada, sem um custo adicional muito grande (van Gunsteren & Berendsen, 1990). Outra alternativa é a utilização de um método chamado campo de reação (*reaction field*)

¹³ Quando se aplica um raio de corte, a descontinuidade do sistema a um raio maior do que o raio de corte causa uma perturbação na simulação por DM. Isso aumenta artificialmente a energia cinética dos átomos, aumentando assim a temperatura do sistema.

(Tironi *et al.*, 1995), onde o meio fora da esfera do raio de corte é modelado com um potencial contínuo específico (Gargallo *et al.*, 2000). Estas técnicas podem amenizar alguns dos problemas da utilização do raio de corte, mas nem sempre os elimina completamente.

Um progresso considerável no tratamento das forças eletrostáticas em sistema biológicos foi feito com a utilização do somatório de Ewald, como o método PME (*particle-mesh Ewald*) (Darden *et al.*, 1993; Essmann *et al.*, 1995). Este método impõe uma periodicidade infinita ao sistema, e permite o cálculo de todas as interações eletrostáticas sem truncamento, levando, em princípio, a uma avaliação mais exata destas interações. Transformadas de Fourier são utilizadas para acelerar o cálculo das interações no espaço recíproco. Este método resulta em trajetórias estáveis e foi logo adotado na simulação de proteínas (Fox & Kollman, 1996) e ácidos nucléicos (Cheatham & Kollman, 1996; Cheatham *et al.*, 1997) em solução.

Sob algumas condições, a periodicidade do sistema pode levar a distorções na simulação de líquidos e soluções (Hünenberger & McCammon, 1999a e 1999b; Weber *et al.*, 2000), portanto alguns cuidados devem ser tomados para a construção destes sistemas, entre eles a adição de contra-íons em regiões próximas a grupos carregados da macromolécula (Ibragimova & Wade, 1998) e utilização de uma camada de solvatação suficientemente grande para não causar distorções nas interações (consideração de um sistema diluído) (Norberto de Souza & Ornstein, 1997).

Na verdade, a escolha do método de tratamento das interações eletrostáticas depende das características dos sistemas e do balanço entre o grau de exatidão necessário e o custo computacional da simulação (van Gunsteren & Berendsen, 1990).

3.1.3 - Condições de contorno

Na simulação de sistemas de tamanho finito ($N \ll$ número de Avogadro = $6,02 \times 10^{23}$ moléculas), alguns cuidados devem ser observados em relação ao meio de simulação (van Gunsteren *et al.*, 1995).

As primeiras simulações de biomoléculas foram realizadas no vácuo, que mimetiza a fase gasosa. Entretanto, a utilização desta condição de contorno na simulação de sistemas macromoleculares em solução pode levar a distorções na superfície do sistema, que tende a se deformar para minimizar a área superficial (van Gunsteren & Karplus, 1982; van Gunsteren & Berendsen, 1990; van Gunsteren *et al.*, 1992).

O efeito causado pelo vácuo pode ser reduzido pela solvatação da macromolécula por uma esfera de moléculas de água. Mas isso somente transfere os efeitos para a superfície água-vácuo. Este efeito da distorção do sistema pode ser diminuído pela utilização de uma "camada protetora" (*extended wall region*), onde o movimento dos átomos é restrito para minimizar a influência do vácuo. Assim, esta região protege a parte do sistema que pode se mover livremente do vácuo que o rodeia (van Gunsteren & Mark, 1992).

Um método mais adequado, e adotado para a minimização dos efeitos de um sistema finito, é a utilização de condições periódicas de contorno. Nestas condições, os átomos do sistema macromolecular são colocados em uma caixa cúbica preenchida com o meio de simulação¹⁴. Esta caixa é tratada como se estivesse rodeada por caixas idênticas por todos os lados (Figura 3.2).

No cálculo das interações, somente os vizinhos (ou imagens) mais próximos são considerados. Assim, um átomo i da caixa central só interage com um átomo j (ou imagem do átomo j) dentro de um certo raio de interação. Desta forma, o átomo i não interage simultaneamente com o átomo j e a imagem periódica deste átomo. Este método é computacionalmente mais caro, mas tem sido normalmente empregado devido a uma representação mais real dos sistemas biológicos.

¹⁴ O meio em que os sistemas biológicos são solvatados tem grande influência em sua estrutura, dinâmica e termodinâmica. Na simulação por DM, uma representação adequada do solvente é importante para a reprodução das forças que atuam no sistema em estudo. A aproximação mais simples no tratamento do solvente é a sua representação como um meio contínuo com determinada constante dielétrica. Entretanto, esta abordagem não permite um estudo das interações soluto-solvente. Uma descrição mais real destas interações se tornou possível com a utilização de moléculas explícitas de solvente ao redor da macromolécula biológica (Rossy & Karplus, 1979; Smith & Pettit, 1994).

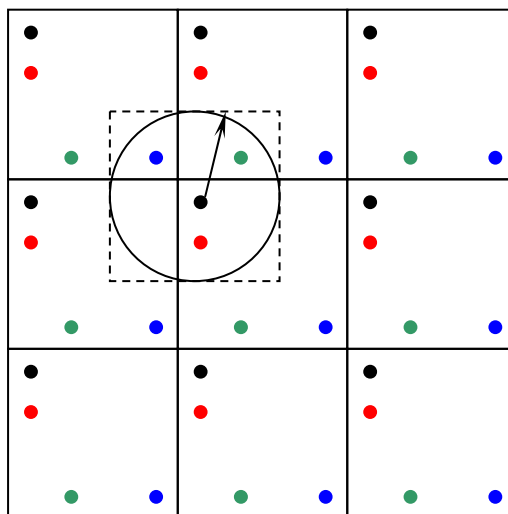


Figura 3.2. Representação das condições periódicas de contorno. Um átomo i (preto) da caixa central só interage com um átomo j (vermelho, verde ou azul), ou imagem do átomo j , dentro de um certo raio de interação. Desta forma, o átomo i não interage simultaneamente com o átomo j e sua imagem periódica. Retirado de van Gunsteren & Berendsen, 1990.

Na simulação sob condições periódicas de contorno e consideração explícita do solvente (geralmente água), a biomolécula é colocada em uma caixa retangular cheia de moléculas de solvente. Um número suficiente de moléculas de solvente deve ser incluído para garantir que a molécula do soluto esteja fora de alcance de suas próprias imagens (Smith *et al.*, 1993). Esta caixa é replicada em todas as direções, mas os cálculos são feitos apenas para a caixa central. Nestas condições, se uma molécula de solvente deixa a caixa por um lado, ela imediatamente entra novamente pelo lado oposto (com a mesma velocidade), garantindo assim que o número de átomos do sistema permaneça constante durante a simulação (van Gunsteren & Berendsen, 1990).

Como a complexidade de uma simulação por DM aumenta rapidamente com o número de átomos, uma simulação com um modelo atomístico de moléculas de água (TIPS, SPC, etc), pode ser muitas vezes (10^1 - 10^3) mais caro do que uma simulação sem solvente explícito (Smith *et al.*, 1993), mas avanços computacionais já permitem que a utilização destas condições sejam comuns.

3.1.4 – Organização de uma simulação por Dinâmica Molecular

Uma simulação por DM é composta de diversas etapas. Inicialmente é necessário o conhecimento das coordenadas e velocidades dos átomos que compõem o sistema, para que as novas posições e velocidades sejam calculadas, e a trajetória gerada. Se as condições de simulação não forem microcanônicas (NVE), então é necessário estabelecer uma temperatura (e/ou pressão) desejada e levar o sistema a esta temperatura (e/ou pressão). Só depois que o sistema estiver equilibrado nas condições desejadas é que as propriedades dinâmicas do sistema podem ser calculadas. Assim, uma simulação por DM é normalmente dividida em três partes: iniciação, equilíbrio e produção (Karplus & Petsko, 1990; Norberto de Souza, 1994).

Iniciação

O ponto de partida de uma simulação por DM é a atribuição das condições iniciais do sistema, como coordenadas e, opcionalmente, as velocidades dos átomos (temperatura) da estrutura inicial. A estrutura inicial da biomolécula a ser simulada por DM pode ser obtida a partir da estrutura tridimensional determinada através da cristalografia por difração de raios X ou por ressonância magnética nuclear (ambas técnicas experimentais), ou por um modelo gerado por estudos de modelagem molecular. Esta molécula é então colocada em um meio de simulação (que pode ser vácuo, meio contínuo ou de solvente explícito) e, se necessário, são adicionados contra-íons para garantir a neutralidade do sistema. O sistema é então relaxado com uma minimização de energia para corrigir possíveis tensões e interações indesejáveis que existam na conformação inicial. Uma vez estabelecidas as coordenadas iniciais de simulação, as velocidades podem ser calculadas diretamente a partir da equação 4, ou pela atribuição das velocidades através de uma distribuição de Maxwell-Boltzmann à temperatura desejada.

Aquecimento e equilíbrio

A atribuição das velocidades para uma temperatura desejada pode introduzir algumas perturbações no sistema, por isso é recomendado que o sistema seja aquecido lentamente. Inicialmente as velocidades são atribuídas para uma baixa

temperatura (geralmente 10K), e através de sucessivas e crescentes termalizações (como 10→50 K, 50→100 K, 100→150 K, etc) esta temperatura é elevada até a temperatura desejada, pela reatribuição das velocidades a cada nova temperatura (Norberto de Souza & Ornstein, 1999).

Após alcançada a temperatura desejada, é ainda provável que o sistema não esteja em equilíbrio, e variações em algumas propriedades podem ser observadas. O sistema é então deixado equilibrar a esta temperatura até que não se observe mais grandes variações em suas energias cinéticas e potenciais, estrutura, temperatura, pressão, e parâmetros que se deseja observar¹⁵. A fase de equilíbrio é alcançada quando estas propriedades tornam-se estáveis, isto é, não se alteram sistematicamente com o tempo.

Produção

Depois que o sistema entra em equilíbrio, ele passa para a chamada fase de produção. Somente nesta fase é que as propriedade médias dos sistemas são calculadas.

¹⁵ A determinação da fase de equilíbrio depende dos parâmetros que se quer observar. Portanto não existe um tempo exato para que o equilíbrio se estabeleça, já que diferentes parâmetros estabilizam em diferentes tempos.

3.2 - Simulações por dinâmica molecular da enzima InhA de *M. tuberculosis*

Neste trabalho foram realizadas simulações por dinâmica molecular do complexo formado pela enzima InhA espécie selvagem e mutantes S94A, I16T, I21V e I47T com o cofator NADH¹⁶, para se avaliar o efeito destas mutações na associação da enzima com o cofator.

Visto que se deseja avaliar a interação das enzimas com o cofator, para a simplificação dos sistemas de simulação foi utilizada somente uma das subunidades da InhA, pois considerou-se que a grande distância entre elementos do sítio ativo de uma subunidade com a subunidade vizinha (maior do que 40 Å) (Figura 3.3) indica que os sítios ativos possam ser independentes entre si.

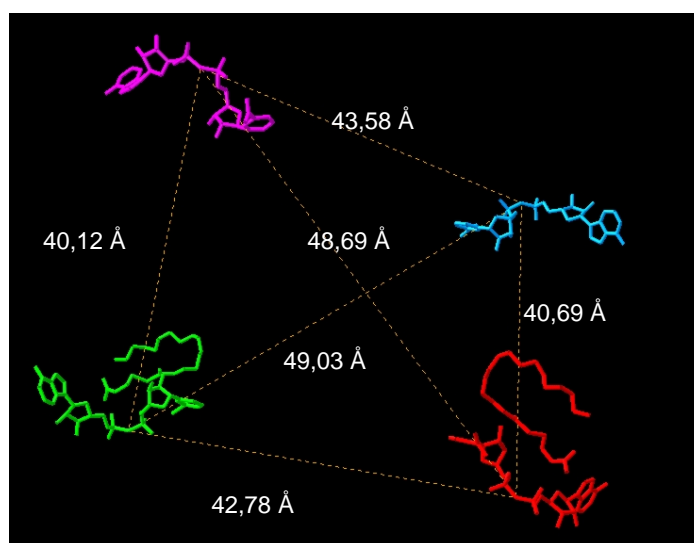


Figura 3.3. Distâncias em Å entre os oxigênios centrais do grupo pirofosfato do NADH nos sítios ativos do homotetrâmero da InhA 1BVR.pdb. Ligantes NADH e C₁₆ da subunidade C (verde), NADH da subunidade D (magenta), NADH da subunidade E (ciano), e NADH e C₁₆ da subunidade F (vermelho). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

¹⁶ As simulações receberam os seguintes códigos: *wt* InhA-NADH (complexo formado entre a InhA espécie selvagem e o NADH), S94A-NADH (complexo formado entre o mutante S94A da InhA e o NADH), I16T-NADH (complexo formado entre o mutante I16T da InhA e o NADH), I21V-NADH (complexo formado entre o mutante I21V da InhA e o NADH), I47T-NADH (complexo formado entre o mutante I47T da InhA e o NADH), cWAT-NADH (complexo formado entre a InhA espécie selvagem e o NADH, considerando as águas cristalográficas), apo-InhA (enzima espécie selvagem sem o cofator).

As coordenadas iniciais da macromolécula espécie selvagem e cofator foram obtidas a partir das coordenadas da estrutura cristalográfica do complexo binário InhA-NADH (Dessen *et al.*, 1995), que apresenta a melhor resolução (2,2 Å), e está identificada no *Protein Data Bank* (PDB) com o código 1ENY. Os mutantes foram modelados *in silico* a partir da estrutura instantânea da InhA em 1,9 ns.

A simulação do complexo *wt* InhA-NADH também foi realizada considerando-se as 41 moléculas de águas cristalográficas a 5 Å de qualquer átomo na estrutura 3D 1ENY.pdb, mas somente para a espécie selvagem (cWAT-NADH). A enzima InhA espécie selvagem também foi simulada sem a presença do cofator (apo-InhA), na sua forma livre em solução.

Em todos os sistemas, o complexo InhA-NADH foi tratado sob condições periódicas de contorno, e o solvente (água) foi considerado explicitamente. A macromolécula, de dimensões aproximadas 44 Å x 27 Å x 35 Å, foi colocada em uma caixa retangular com moléculas de água do tipo TIP3P (Jorgensen *et al.*, 1983), de forma que a macromolécula estivesse cercada por pelo menos uma camada de 10 Å de água por todos os lados. De acordo com Norberto de Souza & Ornstein (1999), essa camada de 10 Å de água em volta da proteína é suficiente para a consideração do sistema de simulação como uma solução diluída. Íons de sódio (Na⁺) foram adicionados para garantir a neutralidade explícita dos sistemas de simulação.

Todas as simulações foram realizadas num *ensemble* isobárico-isotérmico (NPT), numa temperatura de 298,16K (25 °C) e pressão de 1 atm, com o programa SANDER do pacote AMBER 6.0¹⁷ (Case *et al.*, 1999).

Os parâmetros da macromolécula, com todos os átomos explicitamente considerados (modelo atomístico, ou *all-atom model*), contra-íons e moléculas de

¹⁷ O AMBER (*Assisted Model Building with Energy Refinement*) (Pearlman *et al.*, 1995) é um conjunto de programas para a simulação de biomoléculas por DM. O termo *Amber* as vezes é utilizado para se referir ao campo de força (Cornell *et al.*, 1995) implementado no pacote AMBER. Seu código fonte é em sua maioria escrito em linguagem FORTRAN, e opera em sistema UNIX. Este conjunto de programas pode ser utilizado em estações de trabalho, super computadores, *clusters* de PCs, ou em PCs. Seu principal módulo é o programa SANDER (*Simulated Annealing with NMR-Derived Energy Restraints*) que realiza as minimizações de energia, simulações por DM e refinamentos de RMN. SANDER usa o processo de gradientes conjugados para a minimização de energia por mecânica molecular, geralmente após alguns passos do método de *steepest-descent* para corrigir energias iniciais muito desfavoráveis. Para a simulação por DM, ele utiliza a versão *leap-frog* do algoritmo de Verlet para o cálculo da trajetória. O algoritmo SHAKE pode ser usado para a restrição das vibrações de ligações, e a regulação da temperatura pode ser feita tanto por reescalonamento periódico das velocidades, quanto por banho térmico, onde a cada etapa as velocidades são levemente modificadas por um fator que depende da temperatura instantânea e da temperatura desejada (Berendsen *et al.*, 1984). O pacote AMBER ainda possui uma série de programas desenvolvidos para a análise das trajetórias geradas pelas simulações.

água TIP3P¹⁸ foram obtidos da biblioteca parm99.dat do AMBER 6.0.

As cargas parciais e parâmetros para o NADH (Apêndice 1) foram gentilmente cedidos pelo Prof. Dr. Ulf Ryde (Depto de Química Teórica, *Lund University*, Suécia) (Ryde, 1995). Eles foram gerados a partir de cálculos *ab initio* com RHF/6-31G* e ajustados com o programa RESP (Bayly *et al.*, 1993), e são totalmente compatíveis com o campo de força do AMBER (Cornell *et al.*, 1995).

A preparação dos sistemas de simulação, assim como as análises das trajetórias foram feitas em um computador Dell Pentium IV, de 1 GHz com 512 MB de memória RAM e 80 GB de disco rígido. As simulações foram rodadas com o pacote AMBER 6.0 nos clusters AMAZÔNIA (cluster com 16 computadores HP NetServer E60, cada um com dois processadores Pentium III de 550MHz, 256 MB de memória principal, e HD de 9.1GB (SCSI)) e OMBRÓFILA (cluster com 16 computadores HP e-pc, cada um com um processador Pentium III de 1GHz, 256MB de memória principal, e HD de 20GB). O cluster OMBRÓFILA foi construído com o projeto da FAPERGS intitulado “Engenharia Computacional de Proteínas: Análise de Mutantes da Proteína InhA de *Mycobacterium tuberculosis*”, coordenado pelo Prof. Osmar Norberto de Souza. Ambos clusters são administrados pelo CPAD/PUCRS¹⁹.

De acordo com avaliações de tempo-desempenho, melhores resultados foram obtidos utilizando-se 8 nós do *cluster* OMBRÓFILA para cada uma das simulações. Em média, cada pacote de 50 ps de simulação dos sistemas InhA-NADH foram realizados em 15 horas, sendo portanto necessário aproximadamente 45 dias para as simulações de 3,5 ns destes sistemas.

Os mesmos pacotes de 50 ps de simulação dos sistemas, quando rodados em 16 nós do cluster AMAZÔNIA foram finalizados em 8 horas, entretanto este cluster pertence ao Centro de Processamento de Alto Desempenho da Faculdade de Informática da PUCRS e deve ser compartilhado com outros grupos de pesquisa.

¹⁸ O modelo de água TIP3P é um modelo rígido, com uma distância O-H e ângulo H-O-H fixos em 0,9572 Å e 104,52°, respectivamente, com carga -0,834e para o Oxigênio e 0,417 para cada um dos Hs, e parâmetros de Lennard-Jones $\epsilon=76,54$ e $\sigma=3,15$.

¹⁹ O CPAD (Centro de Processamento de Alto Desempenho) da PUCRS é coordenado pelo Prof. Dr. César A. De Rose, e o apoio na manutenção dos clusters é realizada pelo M.Sc. Marco Aurélio Stelmar Netto.

3.2.1 – Simulação do complexo *wt* InhA-NADH

A preparação do sistema *wt* InhA para simulação foi realizada com os programas Prep, Link, Edit e Parm do pacote AMBER 4.1. A molécula de InhA foi preparada a partir das coordenadas da estrutura cristalográfica 1ENY.pdb, com 268 aminoácidos (Figura 3.4). Esta estrutura não apresenta coordenadas para o primeiro resíduo Met1, e a Thr2 original foi substituída por um resíduo Ala.

As coordenadas do NADH dentro de seu sítio de ligação da enzima também foram obtidas da estrutura 1ENY.pdb, e foram utilizadas em conjunto com os parâmetros do campo de força gerados pelo Prof. Dr. Ulf Ryde (Apêndice 1).

Todos os átomos do sistema foram tratados explicitamente (modelo atomístico), e seus resíduos de aminoácidos foram considerados em pH fisiológico. Neste pH, a proteína apresenta uma carga líquida – 4, e o NADH apresenta carga – 2, por isso o sistema a ser simulado foi neutralizado pela adição de 6 íons de sódio (Na^+) próximos a resíduos Asp na superfície da proteína, a 5 Å de distância.

ALA	GLY	LEU	LEU	ASP	GLY	LYS	ARG	ILE	LEU	VAL	SER	GLY	ILE	ILE	THR
ASP	SER	SER	ILE	ALA	PHE	HIE	ILE	ALA	ARG	VAL	ALA	GLN	GLU	GLN	GLY
ALA	GLN	LEU	VAL	LEU	THR	GLY	PHE	ASP	ARG	LEU	ARG	LEU	ILE	GLN	ARG
ILE	THR	ASP	ARG	LEU	PRO	ALA	LYS	ALA	PRO	LEU	LEU	GLU	LEU	ASP	VAL
GLN	ASN	GLU	GLU	HID	LEU	ALA	SER	LEU	ALA	GLY	ARG	VAL	THR	GLU	ALA
ILE	GLY	ALA	GLY	ASN	LYS	LEU	ASP	GLY	VAL	VAL	HIE	SER	ILE	GLY	PHE
MET	PRO	GLN	THR	GLY	MET	GLY	ILE	ASN	PRO	PHE	PHE	ASP	ALA	PRO	TYR
ALA	ASP	VAL	SER	LYS	GLY	ILE	HID	ILE	SER	ALA	TYR	SER	TYR	ALA	SER
MET	ALA	LYS	ALA	LEU	LEU	PRO	ILE	MET	ASN	PRO	GLY	GLY	SER	ILE	VAL
GLY	MET	ASP	PHE	ASP	PRO	SER	ARG	ALA	MET	PRO	ALA	TYR	ASN	TRP	MET
THR	VAL	ALA	LYS	SER	ALA	LEU	GLU	SER	VAL	ASN	ARG	PHE	VAL	ALA	ARG
GLU	ALA	GLY	LYS	TYR	GLY	VAL	ARG	SER	ASN	LEU	VAL	ALA	ALA	GLY	PRO
ILE	ARG	THR	LEU	ALA	MET	SER	ALA	ILE	VAL	GLY	GLY	ALA	LEU	GLY	GLU
GLU	ALA	GLY	ALA	GLN	ILE	GLN	LEU	LEU	GLU	GLU	GLY	TRP	ASP	GLN	ARG
ALA	PRO	ILE	GLY	TRP	ASN	MET	LYS	ASP	ALA	THR	PRO	VAL	ALA	LYS	THR
VAL	CYS	ALA	LEU	LEU	SER	ASP	TRP	LEU	PRO	ALA	THR	THR	GLY	ASP	ILE
ILE	TYR	ALA	ASP	GLY	GLY	ALA	HID	THR	GLN	LEU	LEU				

Figura 3.4. Seqüência da enzima InhA espécie selvagem (268 aminoácidos) utilizada na simulação por dinâmica molecular. Os resíduos HIS foram individualmente avaliados em relação a sua protonação, e substituídos por HID (posição δ protonada) ou HIE (posição ϵ protonada). Resíduos coloridos de acordo com sua polaridade: positivos (azul), negativos (vermelho), polares (laranja) e apolares (cinza).

Este complexo foi inicialmente solvatado em um sistema cúbico, adicionando-se caixas de 216 moléculas de água equilibradas pelo método de

Monte Carlo (Jorgensen *et al.*, 1993), desprezando-se aquelas moléculas muito próximas aos átomos da proteína (distância de até 2,5 Å entre átomos pesados), e também aquelas moléculas a mais de 10 Å da proteína, resultando em uma caixa retangular de dimensões iniciais 77,5 Å x 64,3 Å x 70,0 Å (Figura 3.5) com a seguinte composição (Tabela 3.1):

Tabela 3.1. Composição do sistema de simulação do complexo da enzima InhA espécie selvagem com o NADH (*wt* InhA-NADH).

wt INHA-NADH	Resíduos	Átomos
<i>wt</i> InhA	1-268	1-4008
NADH	269	4.009-4.079
6 Na+	270-275	4.080-4.085
TIP3P (9.132 moléculas)	276-9.407	4.086-31.481
Total	9.407 resíduos	31.481 átomos

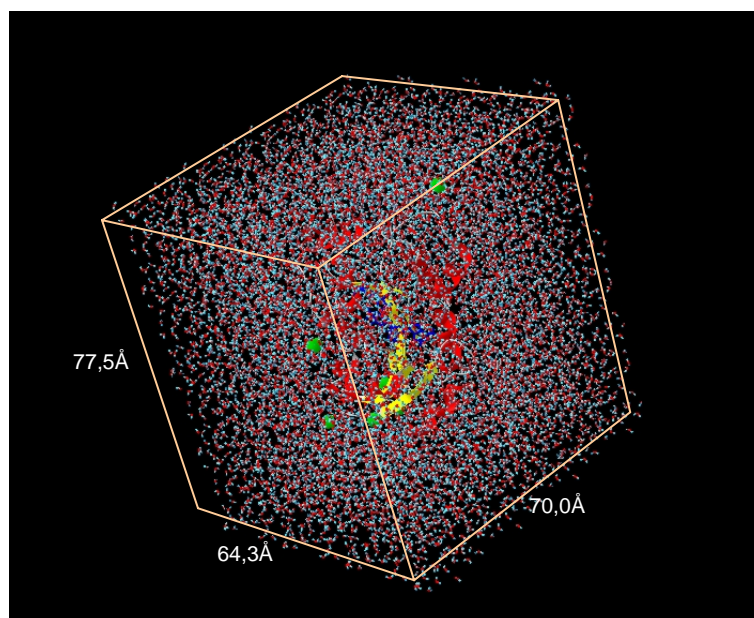


Figura 3.5. Definição do sistema de simulação: complexo *wt* InhA-NADH e contra-íons Na⁺ solvatados por moléculas de água em uma caixa retangular de dimensões 77,5 Å x 64,3 Å x 70,0 Å. InhA (estruturas secundárias em amarelo e vermelho), NADH (azul), íons Na⁺ (superfície molecular verde), e moléculas de água do tipo TIP3P (ciano e vermelho). Este sistema foi replicado periodicamente em todas as direções. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A relaxação das interações desfavoráveis entre as moléculas de solvente²⁰ e entre estas e o soluto foi realizada por 100 passos de minimização de energia (com o algoritmo *steepest descent*), seguido de 20 ps de dinâmica molecular (com passos de 2 fs), para a elevação da temperatura do solvente de 10 a 298,16 K e equilíbrio da densidade da caixa nesta última temperatura²¹. Ambos procedimentos foram realizados com restrição (*constrains*) dos movimentos da proteína e do ligante, utilizando a opção BELLVY do SANDER. Os contra-íons foram tratados como parte do solvente.

O soluto foi então relaxado, com etapas sucessivas de minimização de energia de 500 passos (100 primeiros passos com *steepest descent* e os 400 passos seguintes com o método gradiente conjugado), onde os movimentos da macromolécula e do ligante foram gradualmente liberados (pela diminuição da constante de força de restrição), até que seu movimento fosse completamente liberado²².

A seguir, o aquecimento gradual do sistema até 298,16K foi realizado em pequenas dinâmicas de 2 ps cada, totalizando 12 ps de aquecimento (10→50 K, 50→100 K, 100→150 K, 150→200 K, 200→250 K, 250→298,16 K). As velocidades iniciais do sistema foram atribuídas de acordo com a distribuição de Maxwell-Boltzman a 10 K. A cada etapa de aquecimento as velocidades foram reatribuídas na nova temperatura, e os valores das dimensões da caixa de simulação foram atualizados²³.

Depois de alcançada a temperatura desejada, a simulação do sistema foi realizada à temperatura e pressão constantes de 298,16 K e 1 atm pelo acoplamento do sistema a um banho térmico (Berendsen *et al.*, 1984) e regulação da pressão por reescalonamento isotrópico (em relação a cada molécula). O algoritmo SHAKE (Ryckaert *et al.*, 1977), com uma tolerância de 10^{-5} Å, foi utilizado para a correção das ligações contendo átomos de hidrogênio, permitindo assim o uso de um passo de 2 fs para a integração das equações de movimento. Nenhuma outra medida restritiva foi aplicada ao sistema após a fase de equilíbrio. As

²⁰ Nas simulações por DM, as moléculas de água do tipo TIP3P e os contra-íons, chamados de ClO, foram considerados como solvente. A proteína e o ligante NADH foram considerados como solutos.

²¹ Arquivo de entrada para minimização do solvente: mini_00.in; e para termalização/equilíbrio: mdcp_00.in (Apêndice 2).

²² Arquivos de entrada para a minimização do soluto: mini_01.in a mini_07.in (Apêndice 2).

²³ Arquivos de entrada para a termalização/equilíbrio do soluto: mdcp_01.in a mdcp_06.in (Apêndice 2).

interações eletrostáticas entre átomos não-ligados foram avaliadas pelo método PME, que utilizou uma malha de aproximadamente 1 Å para a interpolação do somatório no espaço direto. As interações de Lennard-Jones foram avaliadas utilizando uma distância máxima de 9 Å²⁴.

Os dados da simulação por dinâmica molecular do sistema *wt* InhA-NADH foram coletados por 3,1 ns, e as estruturas para análise foram salvas a cada 0,5 ps.

Este mesmo protocolo foi utilizado para a simulação da enzima *wt* InhA na ausência do ligante NADH (apo-InhA) (exceto que somente 4 contra-íons foram necessários para a neutralização deste sistema) e para o complexo *wt* InhA-NADH com a consideração das águas cristalográficas (cWAT-NADH) (Apêndice 3). Para a preparação deste último sistema, as moléculas das águas cristalográficas foram inicialmente consideradas como parte do soluto e depois foram ajustadas para serem devidamente tratadas como solvente durante a simulação por DM.

3.2.2 – Preparação e simulação dos mutantes S94A, I16T, I21V e I47T

3.2.2.1 - Mutante S94A InhA

O mutante S94A InhA foi gerado *in silico* durante a simulação do complexo da enzima espécie selvagem e NADH. Ao invés da mutação ser modelada diretamente a partir da estrutura cristalográfica 1ENY.pdb, ela foi realizada na estrutura instantânea da enzima em 1,9 ns de simulação do sistema *wt* InhA-NADH.

Na mutação de um resíduo serina (aminoácido polar) por alanina (um aminoácido menor e apolar), o grupo OH da serina é substituído por um hidrogênio (Figura 3.6). A substituição do aminoácido Ser94 por Ala na estrutura da *wt* InhA em 1,9 ns de simulação se deu pela deleção das coordenadas do hidrogênio da hidroxila, e a substituição do átomo de oxigênio deste grupo por um hidrogênio (mantendo suas coordenadas) no arquivo de coordenadas (Apêndice 4).

²⁴ Arquivo de entrada para a simulação da DM: mdcp_prod_02.in (Apêndice 2)

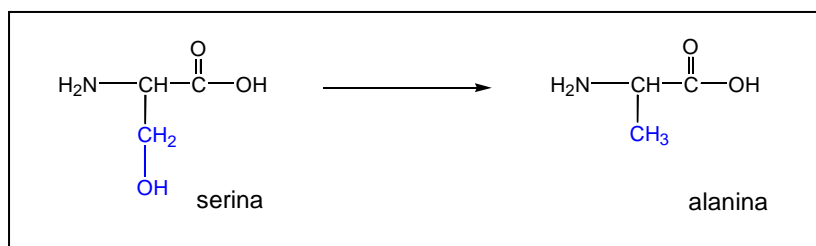


Figura 3.6. Mutaç o do amino cido serina pelo amino cido alanina.

As velocidades do sistema *wt* InhA-NADH (inclusive das  guas e contra- ions) foram transferidas para o novo sistema S94A-NADH, simulando assim o que foi chamado de muta o *on-the-fly*. Na transfer ncia/atribui o das velocidades para o novo sistema foi necess ria a manipula o do arquivo de velocidades para a dele o da velocidade do  tomo de hidrog nio da hidroxila da Ser94. Como foram utilizadas as velocidades do sistema *wt* InhA-NADH   1,9 ns, j  termalizado e com trajet ria est vel, as etapas de aquecimento e equilibra o do sistema mutante foram evitadas. Assim, a din mica prosseguiu como se a muta o tivesse ocorrido ao longo da trajet ria da din mica molecular, simulando uma muta o espont nea. A composi o do novo sistema de simula o est  apresentada na tabela abaixo (Tabela 3.2):

Tabela 3.2. Composi o do sistema de simula o do complexo do mutante S94A da enzima InhA esp cie selvagem com o NADH (S94A-NADH).

S94A-NADH	Res�duos	�tomos
S94A InhA	1-268	1-4007
NADH	269	4.008-4.078
6 Na+	270-275	4.079-4.084
TIP3P (9.132 mol�culas)	276-9.407	4.085-31.480
Total	9.407 res�duos	31.480 �tomos

A simula o do sistema mutante S94A-NADH foi continuada   temperatura constante de 298,16 K e press o de 1 atm pelo acoplamento do sistema a um banho t rmico (Berendsen *et al.*, 1984) e regula o da press o por reescalonamento isotr pico. O algoritmo SHAKE (Ryckaert *et al.*, 1977) foi utilizado

para a correção das ligações contendo átomos de hidrogênio, e a integração das equações do movimento foram feitas a cada 2 fs. As interações eletrostáticas entre átomos não-ligados foram avaliadas pelo método PME, com um raio de corte de 9 Å para as interações de Lennard-Jones.

Os dados da dinâmica do sistema S94A-NADH foram coletados por 3,1 ns, e as estruturas para análise foram salvas a cada 0,5 ps.

3.2.2.2 - Mutantes I16T, I21V e I47T InhA

A construção dos mutantes I16T, I21V e I47T da enzima InhA foi realizada por procedimentos semelhantes à preparação do mutante S94A. Como estas mutações também envolveram a substituição de aminoácidos de cadeias laterais maiores por aminoácidos de cadeias laterais menores (Figura 3.7), as construções dos mutantes envolveram somente a deleção de coordenadas de alguns dos átomos dos resíduos mutados. Uma atenção especial foi tomada para a observação da ordem de ligação dos centros assimétricos dos novos resíduos (Apêndice 4).

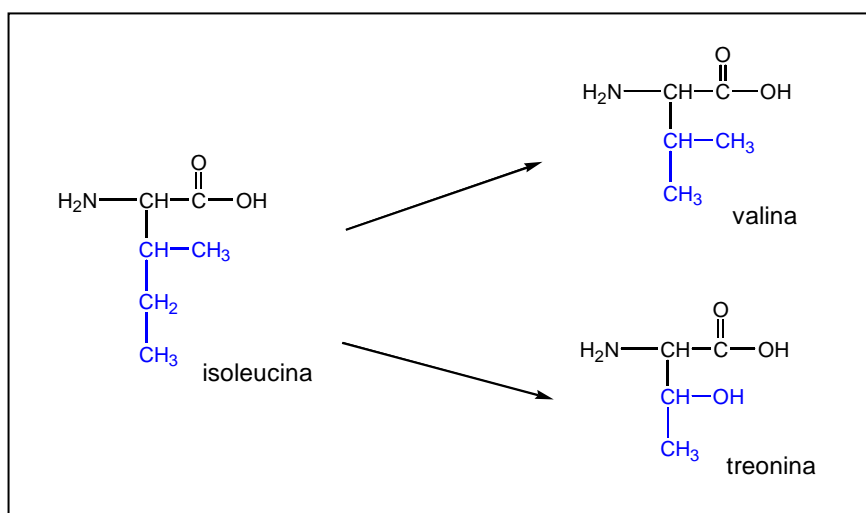


Figura 3.7. Mutações do aminoácido isoleucina pelos aminoácidos valina e treonina para a construção dos mutantes I21V, I16T e I47T.

Após as mutações *on the fly*, a simulação dos mutantes seguiu o mesmo protocolo das simulações anteriores.

A composição dos sistemas de simulações²⁵ encontra-se resumida na Tabela 3.3, a seguir:

Tabela 3.3. Composição dos sistemas InhA-NADH das simulações por dinâmica molecular.

Componentes	<i>apo</i> -InhA	<i>wf</i> InhA-NADH	<i>cWAT</i> -NADH	S94A-NADH	I16T-NADH	I21V-NADH	I47T-NADH
Proteína	268 resid.	268 resid.	268 resid.	268 resid.	268 resid.	268 resid.	268 resid.
NADH	----	1 moléc.	1 moléc.	1 moléc.	1 moléc.	1 moléc.	1 moléc.
Na ⁺	4 íons	6 íons	6 íons	6 íons	6 íons	6 íons	6 íons
Águas cristal.	----	----	41 molec.	----	----	----	----
Águas TIP3P	8.229 moléc.	9.132 moléc.	9.033 moléc.	9.132 moléc.	9.132 moléc.	9.132 moléc.	9.132 moléc.
Nº total resid.	8.501	9.407	9.349	9.407	9.407	9.407	9.407
Nº total átom.	28.699	31.481	31.307	31.480	31.476	31.478	31.476

²⁵ As composições detalhadas dos sistemas de simulações encontram-se no Apêndice 3.

3.3 – O *Docking* Molecular

O entendimento detalhado das interações entre receptores biológicos e seus ligantes é essencial para a interpretação de inúmeros fenômenos bioquímicos, e é de enorme importância prática para aplicações médicas e industriais, sendo a base da estratégia do desenho racional de drogas (Kuntz *et al.*, 1994; Lybrand, 1995; Lengauer & Rarey, 1996).

A determinação de como um ligante se associa aos bio-receptores é, tanto para o desenho de drogas quanto para o estudo da ligação de substratos, uma questão importante da biologia molecular estrutural.

Nos últimos anos, técnicas computacionais como o *docking* molecular foram desenvolvidas para a predição da estrutura (ou estruturas) de complexos intermoleculares formados entre duas ou mais moléculas, e têm sido amplamente empregadas no estudo das interações proteína-ligante. As informações estruturais obtidas do complexo modelado teoricamente podem auxiliar no esclarecimento do mecanismo molecular de reconhecimento, e podem até indicar como a estrutura de um ligante ou receptor²⁶ pode ser modificada para melhorar sua função biológica, ou quais as características necessárias para o desenho de novos compostos que aumentem ou inibam suas funções (Hou *et al.*, 1999).

O *docking* molecular é realizado com a utilização dos chamados algoritmos de *docking*, que são capazes de gerar um grande número de complexos ligante-receptor, classificando-os de acordo com sua energia de interação, e identificando aqueles cuja interação ligante-receptor sejam mais efetivas. Estes algoritmos utilizam os dados da estrutura tridimensional de um determinado receptor para: (i) caracterizar o sítio de ligação; (ii) colocar o possível ligante neste sítio, e (iii) avaliar a qualidade do acoplamento (Kuntz, 1992; Krumrine *et al.*, 2003).

Entretanto, a solução do problema de *docking* não é trivial, principalmente se for considerada a natureza macromolecular das proteínas (o que implica em um conjunto de coordenadas dimensionalmente muito grande) e a complexidade dos

²⁶ No *docking* molecular, o receptor é a macromolécula biológica (geralmente uma proteína ou DNA) à qual um ligante se associa. O ligante pode ser tanto uma pequena molécula (substratos ou inibidores), quanto outra macromolécula (normalmente uma proteína).

fatores energéticos que governam estas interações. Se o ligante também for uma proteína, estas dificuldades tornam-se ainda maiores.

Independentemente da natureza dos complexos receptor-ligante e do método utilizado para o *docking*, algumas questões devem ser contempladas pelos métodos de *docking*, e podem ser resumidas como a combinação de dois componentes: estratégia de busca e uma função de avaliação/classificação. Qualquer método de *docking* deve explorar o espaço conformacional disponível para a interação entre as duas moléculas, e avaliar/classificar as conformações geradas pelo processo de busca²⁷ (Sotriffer *et al.*, 2000).

Assim, o processo de *docking* pode ser descrito como a tentativa para encontrar-se o mínimo global de uma função de avaliação de múltiplas variáveis que sejam capazes de discriminar entre aquelas estruturas nativas e as não-nativas (Trosset & Scheraga, 1998).

Um bom algoritmo de busca deve explorar exhaustivamente todas as formas de ligação entre ligante e receptor. Isto inclui tanto a exploração de todos os seis graus de liberdade translacionais e rotacionais do ligante, quanto os graus de liberdade conformacionais do ligante e do receptor. Entretanto, isso é impraticável devido ao tamanho do espaço conformacional que deve ser analisado. Como consequência, somente uma pequena amostra deste espaço total pode ser examinado, e um balanço entre o custo computacional e a amostragem do espaço conformacional deve ser considerado. Portanto, para que esta amostragem seja computacionalmente possível, uma série de restrições e aproximações devem ser utilizadas para a redução da dimensão do problema, num esforço para a localização do mínimo global da forma mais eficiente possível (Taylor *et al.*, 2002).

A aproximação mais comum nos algoritmos de *docking* é o tratamento do ligante e receptor como corpos rígidos baseado no modelo de reconhecimento molecular do tipo “chave e fechadura”, proposto por Emil Fisher em 1894. Neles, somente os seis graus de liberdade translacionais e rotacionais são explorados. Um exemplo deste tipo de algoritmo é a versão inicial do programa DOCK (Kuntz *et al.*, 1982), que foi criado para a procura de moléculas com um alto grau de complementaridade com o sítio de ligação do receptor. Neste programa, toda a

²⁷ O mais natural seria a avaliação em relação à energia livre de ligação ΔG_{bind} , mas este valor não é facilmente acessível. Portanto, são utilizadas funções alternativas que modelem a ΔG_{bind} da forma mais realística possível.

superfície da macromolécula é analisada, na busca de reentrâncias, que poderiam formar o sítio ativo. Estes sítios são preenchidos com conjuntos de esferas sobrepostas que tocam a superfície molecular em pelo menos dois pontos, derivando uma imagem negativa do sítio de ligação. O ligante é então colocado dentro deste conjunto de esferas e orientado dentro do sítio de ligação. A orientação é conferida para garantir que não há nenhuma interação estérica desfavorável entre ligante e receptor. Se a orientação for aceita, o programa calcula a energia desta interação, que corresponde a um escore. Novas orientações são geradas pela reorientação da molécula no conjunto de esferas. A melhor orientação, isto é, aquela de melhor escore, é registrada para análises subsequentes (Leach, 1996).

Entretanto, o tratamento de ligante e receptor como moléculas rígidas é uma simplificação muito grande, e está limitado àquelas conformações próximas à conformação experimentalmente observada para o complexo ligante-receptor.

Como a geometria molecular pode mudar bastante na associação ligante-receptor (*induced fit*), a flexibilidade conformacional deve ser considerada nos algoritmos de busca. Os programas mais modernos de *docking* molecular se preocupam em incluir também estes graus de liberdade conformacionais, mas levam em conta somente o espaço conformacional do ligante, assumindo que o receptor é rígido. Nestes programas a busca por diferentes conformações é incorporada em alguma etapa do algoritmo (Taylor *et al.*, 2002).

A consideração da flexibilidade do ligante leva ao aumento de seu número de graus de liberdade, e o tamanho do espaço conformacional de investigação se torna bem maior. Por isso, um processo de busca mais efetivo se faz necessário. Os algoritmos que consideram a flexibilidade do ligante podem ser divididos em três tipos de busca: sistemática, determinística e estocástica (Taylor *et al.*, 2002).

Os algoritmos de busca sistemática são baseados em um conjunto de valores para cada grau de liberdade formal, e cada um destes valores do conjunto é explorado de uma forma combinada durante a busca. À medida que mais graus de liberdade são considerados, o número de avaliações necessárias aumenta rapidamente. Assim, são utilizados limites para evitar que o algoritmo explore desnecessariamente posições que levem a soluções erradas. Um exemplo deste tipo de algoritmo é o algoritmo de construção progressiva (*incremental*

construction), que geralmente divide o ligante em pedaços e os testam independentemente. Programas como o DOCK 4.0 (Ewing *et al.*, 2001) e FlexX (Rarey *et al.*, 1996) utilizam esta abordagem.

Nos algoritmos de busca determinísticos, um estado inicial precisa ser definido. Este estado inicial determina o movimento que é feito para gerar o próximo estado. Uma seqüência de estados é assim criada. Desta forma, a utilização das mesmas condições iniciais e dos mesmos parâmetros gera exatamente o mesmo estado final. Exemplos destes métodos são: minimização de energia e dinâmica molecular. Um dos problemas dos métodos determinísticos é que eles podem ficar confinados a uma região de mínimo local, mas ajustes nas condições de simulação podem minimizar este problema (Brooijmans & Kuntz, 2003).

Os algoritmos de busca estocásticos fazem modificações aleatórias no sistema inicial, geralmente mudando um grau de liberdade por vez. Um dos maiores problemas das buscas estocásticas é a incerteza de convergência. Para resolver este problema, várias buscas independentes devem ser feitas. Exemplos de buscas estocásticas são os métodos de Monte Carlo e os algoritmos evolutivos. Algoritmos de busca estocásticos, como *simulated annealing* (SA), programação evolutiva (*evolutionary programming* - EP) e algoritmos genéticos (*genetic algorithm* - GA), têm sido amplamente aplicados aos problemas de *docking* (Baxter *et al.*, 2002). Um exemplo de utilização deste tipos de algoritmos é o programa AutoDock que utiliza o método de Monte Carlo com *simulated annealing* (Goodsell & Olson, 1990). Em cada iteração do processo de Monte Carlo, a conformação interna do ligante é alterada (pela rotação em torno de uma ligação), ou a molécula toda é aleatoriamente rotada ou transladada. A energia do ligante dentro do sítio de ligação é calculada utilizando mecânica molecular, e o movimento pode ser aceito ou rejeitado de acordo com esta energia.

No *docking* molecular com algoritmos genéticos ou programação evolutiva, aquelas condições que apresentam melhor performance dentro de um conjunto de condições são selecionadas para servir de ponto de partida para a geração de novas condições, num processo de evolução convergente. O escore do acoplamento de cada estrutura (com determinada orientação e conformação)

dentro do sítio de ligação é utilizado na seleção daquelas coordenadas iniciais para a avaliação da próxima interação (Morris *et al.*, 1998).

Um método de *docking* molecular ideal seria aquele que permitisse que ambos, ligante e receptor, explorassem seus graus de liberdade conformacionais. Alguns métodos que tratam da flexibilidade do ligante também conseguem amostrar a flexibilidade do receptor, mas a maioria somente considera a flexibilidade do sítio ativo do receptor. Entre eles está a utilização de bibliotecas de confôrmeros (Leach, 1994; Jackson *et al.*, 1998), o uso de características médias de um conjunto de estruturas (Knegtel *et al.*, 1997) e o movimento de cadeias laterais do sítio de ligação do receptor (Nakajima *et al.*, 1997; Pak & Wang, 2000).

A geração de uma grande quantidade de modos de ligação não significa muito se não houver uma forma de avaliar/classificar cada uma destas conformações de uma forma precisa e eficiente. A função de avaliação/classificação deve ser capaz de distinguir o modo de ligação real (observado experimentalmente) dos outros modos propostos pelos algoritmos de busca.

Existem vários tipos de interação entre ligante e receptor que devem ser modeladas²⁸. Tais interações incluem complementaridade geométrica, interações específicas (como interações entre cargas), interações de solvatação/dessolvatação, interações hidrofóbicas e pontes de hidrogênio (Vieth *et al.*, 1998a e 1998b). Quanto mais tipos de interações forem contemplados, melhor o grau de predição da formação do complexo ligante-receptor.

Mas uma função de avaliação rigorosa é computacionalmente muito cara, por isso ela normalmente deve ser simplificada, levando a uma perda de precisão. Os métodos de avaliação podem variar desde os campos de força da mecânica

²⁸ Antes da associação, ambas moléculas interagem com o solvente. A dessolvatação das porções polares ou carregadas destas moléculas é energeticamente desfavorável. Esta energia desfavorável é apenas parcialmente compensada pelas interações eletrostáticas e/ou pontes de hidrogênio formadas no complexo. A dessolvatação das partes não polares das superfícies moleculares libera moléculas de água, o que aumenta a entropia do sistema. Este ganho entrópico é favorável à formação do complexo. Além do aumento da entropia do solvente, também ocorrem mudanças na entropia dos solutos: (i) há a perda da entropia translacional e rotacional quando as duas moléculas se associam, (ii) há variações na entropia vibracional com a formação do complexo, e (iii) há mudanças na entropia conformacional devido à restrição dos ângulos diedros com a formação do complexo. A variação da energia interna das moléculas também deve ser considerada, tanto porque ela geralmente é maior no complexo do que no estado não-complexado, quanto porque esta energia de restrição pode desestabilizar o complexo (Brooijmans & Kuntz, 2003).

molecular²⁹, como AMBER (Cornell *et al.*, 1995), CHARMM (Brooks *et al.*, 1983) ou OPLS (Jorgensen & Tirado-Rives, 1988), até funções empíricas de avaliação de energia livre³⁰ (Böhm, 1994; Eldridge *et al.*, 1997) ou funções derivadas da análise de um grande conjunto de complexos conhecidos³¹ (*knowledge based functions*) (Muegge & Martin, 1999).

Os métodos de *docking* atualmente disponíveis utilizam estas funções de avaliação de duas maneiras. A primeira faz uso da função de avaliação completa para classificar a conformação proteína-ligante. O sistema é então modificado pelo algoritmo de busca e a mesma função de avaliação é usada para classificar a nova estrutura.

O método alternativo é a utilização de funções de avaliação em dois estágios. Assim, uma função simplificada é utilizada para direcionar a busca, e uma função de avaliação mais rigorosa é utilizada para classificar as estruturas selecionadas. Estes métodos mais direcionados fazem simplificações sobre a energia da hipersuperfície, geralmente omitindo termos computacionalmente caros, como interações eletrostáticas, e considerando somente alguns tipos de interações como pontes de hidrogênio. Uma séria limitação da maioria das funções de avaliação/classificação é a tendência de negligenciar os efeitos da solvatação, ou usar modelos de solventes de forma estática (Taylor *et al.*, 2002).

Neste trabalho, a avaliação da interação do NADH com a enzima InhA e mutantes via *docking* molecular foi realizada utilizando-se o programa AutoDock 3.0, que utiliza o campo de força do AMBER (o mesmo utilizado nas simulações por dinâmica molecular) para avaliação das energias de interação. O programa AutoDock 3.0 é um programa já bem estabelecido e é distribuído gratuitamente (mediante registro) para uso acadêmico pelo site <http://www.scripps.edu/pub/olson-web/doc/autodock>.

²⁹ Os campos de força da mecânica molecular já estão bem caracterizados e estabelecidos. Assume-se que a energia livre de ligação em solução pode ser substituída pela entalpia de ligação na fase gasosa: $\Delta G_{bind} \approx \Delta H_{bind}$.

³⁰ A energia livre é considerada como o somatório de todas as interações localizadas: $\Delta G_{bind} \approx \sum \Delta G_i f_i(r)$, onde f_i é uma função das coordenadas e ΔG_i é um fator de importância. É definido um conjunto de termos, como pontes de hidrogênio, interações iônicas, interações hidrofóbicas, entropia de ligação, etc. O valor de ΔG_i é calibrado com base na afinidade de ligação de um conjunto de complexos conhecidos.

³¹ Utiliza informações de dados estruturais já conhecidos de complexos ligantes-receptores para detectar as geometrias favoráveis de ligações. É feita uma análise estatística do banco de dados estruturais, e a energia é calculada em função da densidade de probabilidade.

3.3.1 - O programa AutoDock 3.0

O AutoDock 3.0 é um conjunto de programas (AutoTors, AutoGrid e AutoDock³²) desenvolvidos por Olson e colaboradores (Goodsell & Olson, 1990; Goodsell *et al.*, 1996; Morris *et al.*, 1996; Morris *et al.*, 1998) para a predição das interações de ligantes com receptores macromoleculares, que auxilia no desenho de compostos bioativos (mais especificamente no desenho de drogas assistido por computador).

Como em qualquer procedimento de *docking*, dois quesitos conflitantes tiveram que ser balanceados: a necessidade de um procedimento confiável e preciso, e o desejo de manter a demanda computacional em níveis razoáveis.

O AutoDock 3.0 pode utilizar várias técnicas para a exploração conformacional do ligante (com uma rápida avaliação de energia utilizando potenciais de afinidade molecular), combinando as vantagens de uma busca espacial completa e uma avaliação energética precisa (Morris *et al.*, 1998).

Na exploração conformacional utilizando a técnica de Monte Carlo com *simulated annealing*, a proteína permanece estática durante a simulação e a molécula do substrato faz um movimento aleatório no espaço em torno da proteína, dentro do sítio de ligação. A cada passo da simulação é aplicada uma modificação pequena e aleatória em cada grau de liberdade do substrato: translação em relação ao seu centro geométrico, orientação, e rotação nos ângulos diedros flexíveis. Esta modificação resulta numa nova configuração, cuja energia é avaliada utilizando uma malha de afinidade previamente calculada pelo módulo AutoGrid. Esta nova energia é comparada com a energia da etapa anterior. Se a nova energia for menor, a nova configuração é imediatamente aceita. Se a nova energia for maior, o resultado é tratado probabilisticamente em função da temperatura, de acordo com a expressão:

$$P(\Delta E) = \exp(-\Delta E/k_B T) \quad (15)$$

³² AutoTors é o primeiro programa a ser utilizado, e também é o mais simples. Ele define quais as ligações do ligante podem sofrer rotações, o que diretamente afeta o número de graus de liberdade do ligante, e, portanto, a complexidade computacional (Goodsell *et al.*, 1996). Cada ângulo rotacional adiciona graus de liberdade extra ao sistema, assim ligantes com muitos ângulos rotacionais rapidamente tornam-se muito complexos para cálculo. O AutoGrid calcula malhas tridimensionais de energias de interação baseado no receptor, utilizando o campo de força do AMBER. O AutoDock é propriamente o programa que faz o *docking* molecular.

onde ΔE é a diferença de energia, k_B é a constante de Boltzmann, e T a temperatura em K.

A altas temperaturas, quase todas as etapas são aceitas, enquanto que a baixas temperaturas poucas estruturas de maior energia o são, resultando em uma seleção conformacional maior.

A conformação do ligante no complexo é sucessivamente refinada a medida que a simulação se processa com o controle de temperatura.

A simulação é feita em uma série de ciclos a uma dada temperatura. Cada ciclo comporta um grande número de etapas individuais, que são aceitas ou rejeitadas de acordo com a temperatura corrente. Após um certo número de aceitações ou rejeições das etapas, o próximo ciclo se inicia com uma diminuição da temperatura calculada pela expressão:

$$T_i = g * T_{i-1} \quad (16)$$

onde T_i é a temperatura na etapa i , e g é uma constante entre 0 e 1 (Goodsell & Olson, 1990).

Cada ciclo inicia-se com o ligante na conformação de menor energia encontrada no ciclo anterior, garantindo assim a convergência do método.

O AutoDock 3.0 também pode utilizar algoritmos genéticos para exploração conformacional do ligante no sítio do receptor. O algoritmo genético inicia pela criação aleatória de uma população com um número de indivíduos definidos pelo usuário. Para cada um dos indivíduos são atribuídos “genes” (ou valores) de translação, orientação, rotação e torção variados (dentro de uma faixa determinada pelo usuário). A criação da população inicial é seguida de ciclos sucessivos de gerações até que o número máximo de gerações ou o número máximo de avaliações de energia seja alcançado.

Uma geração consiste em cinco estágios: mapeamento e avaliação, seleção, cruzamento, mutação e seleção do “indivíduo” mais adaptado (melhor conformação) (Morris *et al.*, 1998). O mapeamento transforma o “genótipo” de cada indivíduo em um “fenótipo” correspondente, que permite que a avaliação em relação à energia de interação seja feita. Aqueles indivíduos com melhor avaliação

são selecionados para cruzamento. O cruzamento e mutações são feitos aleatoriamente entre os membros da população, de acordo com os parâmetros determinados para estes estágios. Estas etapas geram novos indivíduos que substituem aqueles que lhes deram origem (mantendo o número da população constante). O parâmetro determinado pela seleção determina quantos indivíduos automaticamente sobreviverão na próxima geração.

Este método faz diversas iterações através das gerações até que o critério de terminação seja alcançado.

A grande maioria dos algoritmos genéticos está fundamentada nas características da evolução Darwiniana, seguindo a genética Mendeliana. Neles, as características do genótipo (dado pelo cruzamento e mutações) determinam o fenótipo (determinado pela avaliação de energia).

O programa AutoDock 3.0, contudo, permite a utilização do chamado algoritmo genético Lamarckiano (*Lamarckian genetic algorithm* - LGA), que está baseado na afirmação de que as características fenotípicas adquiridas durante a vida de um indivíduo podem se tornar hereditárias.

Quando o LGA é usado, fenótipos próximos ao fenótipo escolhido são avaliados, e um novo fenótipo (com melhor avaliação de energia) é selecionado. Fazendo o caminho inverso, este fenótipo é associado a um genótipo, que caracteriza o indivíduo selecionado (Morris *et al.*, 1998). Assim, no LGA, os valores das variáveis são ajustados de uma forma determinística por uma otimização local, e estes “genes adaptados” são passados para as próximas gerações se apresentarem uma melhor avaliação de energia³³ (Thormann & Pons, 2001).

3.3.1.1 – O módulo AutoGrid

Independentemente do método de busca, a energia de interação do substrato com a proteína é avaliada a cada passo da simulação, utilizando-se potenciais de afinidade molecular determinados pelo módulo AutoGrid. Com este programa, a proteína é envolvida numa malha tridimensional, e um átomo teste

³³ No LGA, após cada geração é feita uma avaliação das conformações próximas à conformação selecionada (*local search*), e se houver uma conformação de melhor avaliação de energia, ela substitui a anterior.

visita cada ponto desta malha, registrando sua energia de interação com a proteína (Goodford, 1985) (Figura 3.8). Para cada ponto da malha, a energia de interação daquele ponto é avaliada para todos os átomos do receptor a menos de 8 Å de distância. A malha de afinidades representa uma tabela de consulta para a rápida avaliação das energias de interação. Diferentes malhas de afinidade são calculadas, uma para cada tipo de átomo do ligante (normalmente C, N, O, e H), incluindo energias de dispersão/repulsão e de pontes de hidrogênio. Além destas malhas de afinidade, o AutoDock necessita de uma malha de potencial eletrostático, que é gerada pelo AutoGrid, e representam as interações coulômbicas entre a macromolécula e uma partícula teste de carga e ($1,60219 \times 10^{-19}$ C).

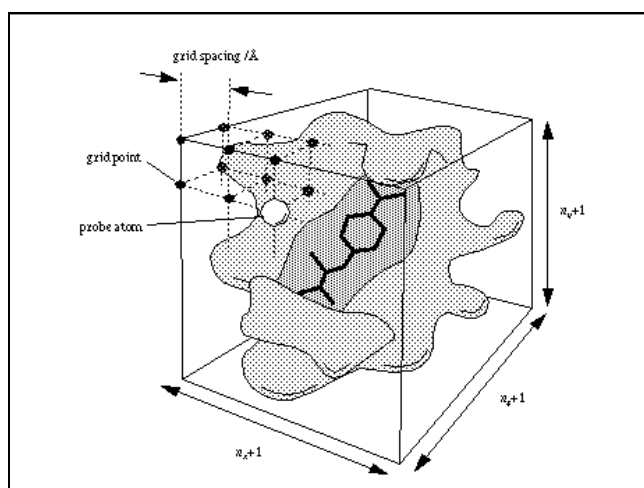


Figura 3.8. Átomo teste (*probe atom*) passando pela malha tridimensional e registrando a energia potencial em cada ponto desta malha. Retirado de: *AutoDock User Guide* – Morris *et al.*, 2001.

Estas avaliações de afinidade são feitas baseadas no campo de força do AMBER (Weiner *et al.*, 1984). As energias de dispersão/repulsão são calculadas com o tradicional potencial de Lennard-Jones 12-6, e as pontes de hidrogênio são tratadas com o potencial 12-10. As energias de interação eletrostática são calculadas pela expressão do potencial de Coulomb (Goodsell *et al.*, 1996).

Durante a simulação, estas malhas são usadas como referência para a avaliação rápida da energia de interação. Para cada átomo do substrato numa dada conformação, a energia de interação é calculada pela interpolação dos valores de afinidade dos oito pontos da malha em torno da posição deste átomo. A energia de

interação eletrostática de cada átomo é calculada pela multiplicação do valor daquele ponto da malha de potencial eletrostático pela carga parcial do átomo. Estes cálculos são proporcionais somente ao número de átomos do ligante, e não ao número de átomos da proteína. Como estes potenciais de afinidade são pré-calculados, uma rápida avaliação da energia no momento do *docking* é facilmente obtida.

3.3.1.2 – O módulo AutoDock

A versão AutoDock 3.0 apresenta um avanço em relação às versões anteriores. Nesta versão, o módulo AutoDock utiliza um campo de força empírico para a predição da energia livre de ligação, e, portanto, para a predição da constante de ligação para o ligante testado.

No módulo AutoDock, a avaliação da energia livre de interação é calculada utilizando um campo de força que inclui cinco termos:

$$\begin{aligned}
 \Delta G_{bind} = & \Delta G_{vdW} \sum_{ij} \left(\frac{A_{ij}}{r_{ij}^{12}} + \frac{B_{ij}}{r_{ij}^6} \right) \\
 & + \Delta G_{Hbond} \sum_{ij} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} + \frac{D_{ij}}{r_{ij}^{10}} \right) \\
 & + \Delta G_{elec} \sum_{ij} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \\
 & + \Delta G_{sol} \sum_{ij} (S_i V_j + S_j V_i) e^{(-r_{ij}/2\sigma^2)} \\
 & + \Delta G_{tor} N_{tor}
 \end{aligned} \tag{17}$$

onde os cinco termos ΔG da direita correspondem a coeficientes otimizados empiricamente baseados em um conjunto de complexos com constantes de ligação conhecidas (Morris *et al.*, 1998).

O somatório é realizado sobre todos os pares do átomo teste i , e todos os átomos j da proteína. Na equação 17, os três primeiros termos, também presentes nas versões anteriores do AutoDock, correspondem às interações não-ligantes dos campos de força da mecânica molecular: o primeiro termo modela as forças de dispersão/repulsão, e utiliza o potencial padrão de Lennard-Jonnes 12-6, o segundo termo modela as pontes de hidrogênio utilizando o potencial de Lennard-Jonnes 12-10, onde $E(t)$ corresponde a um coeficiente baseado no ângulo t formado entre o átomo teste, o hidrogênio e o átomo alvo (doador) na proteína, e o terceiro termo corresponde ao potencial de Coulomb para as interações eletrostáticas (Morris *et al.*, 1996). O quarto termo modela o efeito da dessolvatação e os efeitos hidrofóbicos, utilizando parâmetros de solvatação S e volumes atômicos V da proteína e dos átomos teste. O termo final estima a contribuição das rotações internas, rotação global e translação do ligante para a entropia de ligação (Zhao *et al.*, 2001).

A constante de inibição k_i pode então ser calculada em relação à variação de energia livre de ligação, de acordo com a equação 18:

$$\Delta G_{bind} = RT \ln k_i \quad (18)$$

onde R é a constante geral dos gases ($1,987 \text{ cal K}^{-1} \text{ mol}^{-1}$), e T é a temperatura em K (considerada $298,15 \text{ K}$ – temperatura ambiente)³⁴ (Morris *et al.*, 1998).

³⁴ É importante ressaltar que esta equação não apresenta o sinal negativo porque a constante de inibição k_i é definida para a reação de dissociação $EI \leftrightarrow E + I$, enquanto o ΔG_{bind} se refere ao processo inverso de ligação $E + I \leftrightarrow EI$, onde E é a enzima e I o inibidor.

3.4 - *Docking* do NADH nas estruturas da enzima InhA espécie selvagem e mutantes S94A, I16T, I21V e I47T

A avaliação da energia de associação do NADH à enzima InhA espécie selvagem e mutantes por experimentos de *docking* molecular foram conduzidos com o programa AutoDock 3.0 (Morris *et al.*, 1998).

Inicialmente foram realizados diversos experimentos de *docking* com a estrutura cristalográfica da InhA 1ENY.pdb para a determinação do protocolo de *docking* mais apropriado para o sistema. Nestes experimentos foram testados os métodos de busca por *simulated annealing* (SA) e algoritmos genéticos Lamarckianos (LGA), assim como diferentes protocolos para cada um destes métodos.

Em ambos os métodos, o ligante foi testado como molécula rígida ou como molécula flexível. A flexibilidade do ligante foi considerada nas 16 ligações que podem sofrer torções (identificadas pelo módulo AutoTors), ou somente nas 6 ligações centrais da molécula do ligante (porção pirofosfato), indicadas na Tabela 3.4. Quanto maior o número de ligações consideradas flexíveis, maior o tempo necessário para o experimento.

Tabela 3.4. Torções possíveis identificadas pelo programa Autotors na molécula de NADH.

Torção	Entre os átomos do NADH	Torção	Entre os átomos do NADH
1	N9A e C1'A	9	C5'N e C4'N
2	C4'A e C5'A	10	C1'N e N1N
3*	C5'A e O5'A	11	C3N e C7N
4*	O5'A e PA	12	C2'N e O2'N
5*	PA e O3P	13	C3'N e O3'N
6*	O3P e PN	14	C3'A e O3'A
7*	PN e O5'N	15	C2'A e O2'A
8*	O5'N e C5'N	16	C6A e N6A

* 6 ligações centrais da molécula do ligante (porção pirofosfato).

A conformação do ligante na estrutura cristalográfica 1ENY.pdb foi utilizada como condição inicial para os experimentos de *docking*. Em alguns deles, a posição (coordenadas) cristalográfica do NADH foi considerada como posição inicial, mas na maioria dos experimentos sua posição em relação à proteína foi levemente transladada dentro do sítio de ligação.

Os hidrogênios do ligante foram adicionados com o programa MOE (MOE, 1998), e as cargas atômicas parciais utilizadas foram determinadas a partir de cálculos quântico-mecânicos com HF/6-31C* (cedidas pelo Prof. Dr. Ulf Ryde – Apêndice 1) (Ryde, 1995) e ajustadas com o programa RESP (Bayly *et al.*, 1993).

A definição do grau de flexibilidade do ligante foi realizada com o módulo AutoTors. Neste programa, as ligações do ligante que podem sofrer torções são definidas pelo usuário, e as cargas parciais dos hidrogênios apolares são adicionadas às cargas parciais dos átomos pesados aos quais estão ligados, pois no *docking*, somente os hidrogênios polares do ligante são tratados explicitamente.

Para o *docking* com o programa AutoDock3.0, o arquivo de entrada da macromolécula deve conter as coordenadas de todos os átomos do sistema (no formato .pdb), com uma coluna extra contendo as cargas parciais de cada um destes átomos. As cargas parciais da InhA foram adicionadas com os módulos Link e Edit do pacote AMBER4.1, gerando um arquivo de extensão .pdbq. As cargas parciais foram determinadas no modelo átomos unidos (Weiner *et al.*, 1984), mas o modelo atomístico (Cornell *et al.*, 1995) também pode ser utilizado.

Os parâmetros de solvatação atômica da macromolécula foram assinalados com o utilitário *addsol* do AutoDock 3.0, gerando um arquivo de extensão .pdbqs, com as informações necessárias da macromolécula para o cálculo das malhas de afinidades.

As malhas de afinidades, utilizadas para a rápida avaliação da energia de interação entre o ligante e a macromolécula, foram calculadas utilizando o módulo AutoGrid. Malhas de afinidades independentes foram calculadas para cada tipo de átomo do ligante NADH (C, A³⁵, N, O, P e H), juntamente com a malha de potencial eletrostático.

³⁵ Corresponde aos carbonos aromáticos da molécula do ligante NADH.

Os experimentos de *docking* propriamente ditos (via SA ou LGA) foram realizados com o módulo AutoDock, de acordo com os parâmetros determinados no arquivo de entrada de *docking* (exemplo no Apêndice 5)

Nos experimentos via SA, temperaturas iniciais de 37, 100, 300, 500 e 1000 °C foram testadas, tanto com redução geométrica, como com redução linear da temperatura. O número de corridas e o número de ciclos, assim como o número total de passos aceitos e/ou rejeitados, foram variados para a determinação dos parâmetros mais apropriados para os experimentos de SA.

No método de SA, o estado inicial do ligante é determinado aleatoriamente pelo AutoDock, a partir de um gerador de números aleatórios baseados na contagem de tempo ou no número do processo na máquina. Desta forma, diferentes experimentos com protocolos idênticos podem não reproduzir o valor de energia livre de ligação. Para a padronização dos experimentos de SA, utilizou-se um número fixo para a geração da conformação inicial, fixando a variável correspondente na fonte do programa.

Para a determinação das melhores condições experimentais para o *docking* via LGA, variou-se o tamanho da população, o número de indivíduos que sobrevivem para a próxima etapa, a razão de mutação e cruzamento e número de iterações na busca local.

Resultados muito próximos foram obtidos com os dois métodos de busca, entretanto o método SA apresentou melhores avaliações de energia de interação e melhor reprodução da geometria de ligação, e por isso foi escolhido como o método de *docking* quando o ligante foi considerado rígido. O LGA se mostrou um método mais apropriado quando a flexibilidade do ligante foi considerada (Tabela 3.5).

Tabela 3.5. Comparação dos algoritmos de busca e da flexibilidade do ligante nos experimentos de *docking* do NADH no sítio de ligação da enzima InhA (1ENY).

Macromolécula	Posição inicial do NADH	Flexibilidade do NADH	Algoritmo de busca	ΔG_{bind} (kcal/mol)
1ENY	cristalográfica	rígido	SA	-15,37 ± 0,16
1ENY	cristalográfica	flexível	SA	- 9,37 ± 1,92
1ENY	deslocada	rígido	SA	-15,27 ± 0,26
1ENY	deslocada	flexível	SA	- 4,50 ± 2,25
1ENY	deslocada	rígido	LGA	-10,56 ± 0,81
1ENY	deslocada	flexível	LGA	-13,20 ± 0,05

3.4.1 - *Docking* do NADH na estrutura cristalina da enzima InhA (1ENY)

O *docking* do NADH na estrutura cristalina InhA espécie selvagem foi realizado com a estrutura 1ENY, sem as águas cristalográficas e sem o ligante. A adição dos átomos de hidrogênio e o assinalamento das cargas atômicas parciais foram feitos com o programa AMBER. Os parâmetros de solvatação da macromolécula foram adicionados com o utilitário *addsol*.

A conformação inicial do NADH correspondeu àquela da estrutura cristalina 1ENY, e as cargas calculadas pelo Prof. Dr. Ulf Ryde foram utilizadas para o assinalamento das cargas parciais dos átomos do ligante. O NADH foi manualmente colocado dentro de seu sítio de ligação na enzima InhA, levemente deslocado de sua posição cristalográfica, com o programa MOE. O arquivo com a posição de partida do NADH e suas cargas atômicas parciais foi salvo na extensão .mol2, necessária para o programa AutoTors. Todas as rotações de ligações no NADH foram desativadas (*docking* rígido).

Malhas de afinidades de 61 x 61 x 61 pontos, calculadas pelo programa AutoGrid, foram definidas com um espaço de 0,375 Å entre os pontos da malha. Como o local de ligação do NADH na enzima InhA é conhecido, as malhas de afinidades foram centradas em seu sítio de ligação, apresentando dimensões 22,5 Å x 22,5 Å x 22,5 Å.

O conjunto de parâmetros para o SA otimizado inicialmente foi utilizado para os experimentos. Estes parâmetros incluem incrementos de 0,2 Å para as translações e de 5 graus para as orientações e torsões³⁶, temperatura inicial de anelamento de 37 °C (parâmetro $rt_0 = 616,0 \text{ cal mol}^{-1}$ – Apêndice 5) e um esquema de redução geométrica da temperatura.

O experimento de *docking* via SA foi realizado com 10 corridas de 50 ciclos cada, com critério de terminação de ciclo de um máximo de 25.000 passos aceitos, ou 25.000 passos rejeitados, o que vier primeiro. O estado de menor energia do ciclo anterior foi utilizado para o início do próximo ciclo. Uma energia inicial máxima de $1.000 \text{ kcal mol}^{-1}$ e um número máximo de 1.000 tentativas para gerar o estado inicial aleatório foram permitidos.

³⁶ Estes incrementos determinam o quanto cada variável de estado pode mover-se durante o *simulated annealing*.

Os resultados do *docking* (melhores conformações) foram classificados em relação crescente de energia. As coordenadas da conformação de menor energia de cada ciclo foram comparadas com aquelas de todas as outras conformações, e agrupadas de acordo com os valores de desvio médio quadrático (RMSD). Assim, uma lista de famílias de conformações similares e avaliações de energia livre de ligação foi produzida.

Na análise das conformações de *docking*, a tolerância do cluster para o RMSD foi de 1,0 Å, utilizando como referência as coordenadas do ligante na posição cristalográfica.

No protocolo de *docking* via LGA otimizado foram realizadas 10 corridas independentes, com uma população aleatória inicial de 50 indivíduos e um número máximo de 250.000 avaliações de energia. Foi utilizado um número máximo de 27.000 gerações com valor de elitismo de 1 (número de melhores indivíduos que sobrevivem para a próxima geração), taxa de mutação de 0,02 (que corresponde à probabilidade de um gene sofrer uma mutação aleatória) e uma taxa de cruzamento de 0,80. A busca local (*local search*) foi utilizada permitindo um máximo de 300 iterações em cada busca, com no máximo 4 sucessos ou fracassos consecutivos antes da alteração das condições de busca.

3.4.2 - *Docking* do NADH na estrutura cristalina do mutante S94A (1ENZ)

Para a verificação da capacidade do método para a comparação da diferença de afinidade da ligação do NADH na enzima espécie selvagem e mutantes, o mesmo protocolo de *docking* molecular foi testado na estrutura cristalina do mutante S94A (1ENZ)

A estrutura do mutante S94A 1ENZ foi inicialmente considerada sem as águas cristalográficas e sem o ligante. A adição dos átomos de hidrogênio e o assinalamento das cargas atômicas parciais foram realizados com o programa AMBER e os parâmetros de solvatação da macromolécula foram adicionados com o utilitário *addsol*.

A conformação inicial do NADH correspondeu àquela da estrutura cristalina 1ENZ. Os hidrogênios do ligante foram adicionados com o programa MOE, e as

cargas atômicas parciais atribuídas conforme descrito na seção anterior e detalhada no Apêndice 1. O NADH foi manualmente colocado dentro de seu sítio de ligação no mutante S94A, levemente deslocado de sua posição cristalográfica, e foi considerado rígido durante o experimento.

As malhas de afinidades e protocolo de *docking* via SA foram definidos conforme a seção anterior.

3.4.3 – Consideração das águas cristalográficas no *docking* do NADH na estrutura cristalina da enzima InhA (1ENY)

A estrutura cristalina da InhA espécie selvagem (1ENY) e suas 41 moléculas de águas cristalográficas também foram consideradas para os experimentos de *docking* molecular. As moléculas de água foram tratadas como parte da macromolécula.

A adição dos átomos de hidrogênio e o assinalamento das cargas atômicas parciais foram realizados de acordo com a seção 3.4.1.

O NADH foi considerado rígido durante o experimento, com conformação correspondente àquela da estrutura cristalina 1ENY e posição inicial levemente deslocada dentro do sítio de ligação. A preparação do ligante foi realizada como descrito na seção 3.4.1.

As malhas de afinidades e protocolo de *docking* via SA foram definidos conforme a seção 3.4.1.

3.4.4 - *Docking* do NADH nas estruturas dos mutantes S94A, I16T, I21V e I47T da enzima InhA

A comparação das afinidades de ligação do NADH nos sistemas mutantes foi realizada utilizando-se as estruturas dos mutantes S94A, I16T, I21V e I47T geradas a partir da mutação *in silico* da estrutura da InhA espécie selvagem (*wt* InhA) em 1,9 ns de simulação.

Após modeladas, cada uma das estruturas tiveram suas cargas atômicas parciais e parâmetros de solvatação atribuídos conforme a seção 3.4.1.

O NADH, com a conformação cristalográfica em 1ENY e preparado de acordo com a seção 3.4.1, foi manualmente colocado dentro de seu sítio de ligação nas estruturas modeladas, levemente deslocado de sua posição cristalográfica, e foi considerado rígido durante os experimentos.

Os *dockings* do NADH nas estruturas dos mutantes da enzima InhA foram realizado de acordo com o protocolo SA definido na seção 3.4.1.

3.5 - *Docking* do NADH nas trajetórias da enzima InhA espécie selvagem e mutantes S94A, I16T, I21V e I47T

A afinidade do NADH foi testada na trajetória das simulações da enzima InhA espécie selvagem (*wt* InhA-NADH) e dos mutantes S94A, I16T, I21V e I47T como forma de simular a flexibilidade da macromolécula no *docking* molecular.

As estruturas instantâneas das macromoléculas ao longo de cada uma das trajetórias geradas nas simulações por dinâmica molecular foram utilizadas nos experimentos de *docking* do NADH.

Para cada um dos sistemas, retirou-se as moléculas de NADH, as moléculas de água do sistema, e os contra-íons das estruturas dos *snapshots*. Estas estruturas foram então sobrepostas às respectivas estruturas médias das fases de produção para permitir que as mesmas coordenadas iniciais do ligante fossem utilizadas para todos os *dockings*.

O NADH, na conformação inicial correspondente à da estrutura cristalina, foi manualmente colocado dentro do sítio de ligação da estrutura média da fase de produção de cada um dos sistemas, levemente deslocado da posição cristalográfica. A preparação do arquivo do ligante seguiu as etapas descritas na seção 3.4.1.

Dois protocolos de *docking* do NADH nas trajetórias foram testados.

3.5.1 - *Docking* do NADH nas trajetórias das simulações por DM via *simulated annealing*

O *docking* do NADH via *simulated annealing* nas trajetórias das simulações utilizou o protocolo de SA anteriormente otimizado para o sistema em estudo (seção 3.4.1 e Apêndice 5), e considerou o ligante como uma molécula rígida, na conformação cristalográfica.

Como os *snapshots* da trajetória foram sobrepostos, a mesma orientação inicial do ligante em relação à macromolécula (levemente deslocado da posição cristalográfica) foi considerada para todos os experimentos de *docking*.

Para cada um dos sistemas, um programa de leitura seqüencial utilizou os *snapshots* das respectivas trajetórias da DM (de 1 em 1 ps) para fazer a preparação das macromoléculas, assim como o cálculo das malhas de afinidade e *docking* do ligante em cada um dos *snapshots*.

Um arquivo com os resultados de *docking* foi gerado para cada um dos *snapshots* considerados (Tabela 3.6), formando um conjunto de avaliações de energia livre de ligação do NADH em função do tempo para cada um dos diferentes sistemas. Somente as melhores energias de *docking* foram consideradas para análise.

Tabela 3.6. Resumo dos experimentos de *docking* do NADH nas trajetórias das simulações por DM via *simulated annealing*.

Sistemas	wt InhA	cWAT InhA	S94A InhA	I16T InhA	I21V InhA	I47T InhA
Tempo de simulação	3.100 ps	4.000 ps	3.100 ps	5.500 ps	6.500 ps	4.000ps
Nº total de <i>dockings</i> via SA	3.100	4.000	3.100	5.500	6.500	4.000

3.5.2 - *Docking* do NADH nas trajetórias das simulações por DM via algoritmos genéticos

O *docking* do NADH via algoritmos genéticos nas trajetórias das simulações utilizou o protocolo de LGA anteriormente otimizado para o sistema de estudo (seção 3.4.1 e Apêndice 5), e considerou a flexibilidade das 6 ligações centrais da molécula do ligante (porção pirofosfato), descritas na Tabela 3.4.

Como os *snapshots* da trajetória foram sobrepostos, a mesma orientação inicial do ligante em relação à macromolécula (levemente deslocado da posição cristalográfica) foi considerada para todos os experimentos de *docking*.

Para cada um dos sistemas, um programa de leitura seqüencial utilizou os *snapshots* das respectivas trajetórias da DM (de 2 em 2 ps) para fazer a preparação das macromoléculas, assim como o cálculo das malhas de afinidade e *docking* do ligante em cada um dos *snapshots*.

Um arquivo com os resultados do *docking* do NADH foi gerado para cada um dos *snapshots* considerados (Tabela 3.7)³⁷, formando um conjunto de avaliações de energia livre de ligação do NADH em função do tempo para cada um dos diferentes sistemas. Somente as melhores energias de *docking* foram consideradas para análise.

Tabela 3.7. Resumo dos experimentos de *docking* do NADH nas trajetórias das simulações por DM via algoritmos genéticos.

Sistemas	wt InhA	cWAT InhA	S94A InhA	I16T InhA	I21V InhA	I47T InhA
Tempo de simulação	3.100 ps	4.000 ps	3.100 ps	5.500 ps	6.500 ps	4.000 ps
Nº total de <i>dockings</i> via LGA	1.550	2.000	1.550	2.750	3.250	2.000

³⁷ O tempo de *docking* em cada *snapshot* de uma mesma trajetória pode variar de acordo com a conformação das cadeias laterais naquela estrutura instantânea e com o protocolo de *docking*. Nos experimentos via SA, onde o NADH é considerado rígido, o *docking* em cada *snapshot* em computador Dell Pentium IV, de 1GHz com 512 MB de memória RAM variou entre 5 e 10 min. Nos experimentos via GA, apesar do método permitir uma convergência mais rápida, o NADH é considerado flexível, e o *docking* em cada *snapshot* variou entre 7 e 12 min.

Resultados e Discussão

A enzima 2-*trans*-enoil-ACP-redutase de *Mycobacterium tuberculosis* (InhA) é reconhecidamente um alvo importante para a intervenção no tratamento anti-tuberculose (Dessen *et al.*, 1995; Quémard *et al.*, 1995; Roswarski *et al.*, 1998). A inibição desta enzima pela isoniazida, o principal fármaco utilizado nos esquemas de tratamento dessa doença, leva à inibição da síntese de ácidos micólicos, fundamentais para a estrutura do envelope celular do MTB.

Com o relato do aumento do número de casos de tuberculose causados por cepas de MTB resistentes à INH, a enzima InhA tem sido foco de estudo para o entendimento de seu mecanismo de ação e resistência. Grande parte destes estudos está centrado na investigação bioquímica de seu mecanismo de ação, inibição e resistência, assim como na identificação de mutantes resistentes à INH (Johnsson *et al.*, 1995; Quémard *et al.*, 1995; Quémard *et al.*, 1996; Zabinski & Blanchard, 1997; Basso *et al.*, 1996 e 1998; Basso & Blanchard, 1998; Parikh *et al.*, 1999; Ward *et al.*, 1999; Lei *et al.*, 2000; Parikh *et al.*, 2000; Nguyen *et al.*, 2001; Rawat *et al.*, 2003).

Detalhes da enzima InhA em nível atômico tornaram-se possíveis com a determinação das estruturas cristalinas da enzima espécie selvagem com seu cofator (Dessen *et al.*, 1995) e com um análogo do substrato (Rozwarski *et al.*, 1999), com o complexo inibidor NAD-INH (Rozwarski *et al.*, 1998), com os inibidores TCL e GEQ (Kuo *et al.*, 2003), e do mutante S94A complexado com o NADH (Dessen *et al.*, 1995). Estas estruturas, aliadas aos dados bioquímicos, nos fornece um bom indicativo dos eventos moleculares envolvidos na ligação do cofator, substrato e inibidores a essa enzima, assim como apontam para alguns fatores que podem ser responsáveis pela resistência à INH. Entretanto, à exceção de um trabalho que relata o estudo *ab initio* da interação de alguns resíduos de aminoácidos da InhA espécie selvagem com a porção pirofosfato do NADH e 1ns

de simulação por DM do complexo ternário da InhA com o complexo inibidor NAD-INH (Pantano *et al.*, 2002), nenhum trabalho de acompanhamento das propriedades dinâmicas do sistema espécie selvagem e mutantes foi relatado a fim de se observar as diferenças que podem levar à resistência.

Tendo em vista que o mecanismo de ação da InhA envolve a associação da enzima com seu cofator, e que as mutações que levam à resistência a INH estão situadas no sítio de ligação do NADH (Figura 4.1), a simulação por DM dos sistemas da enzima (espécie selvagem e mutantes) com o NADH parece ser uma alternativa apropriada para a investigação, em maiores detalhes, das principais interações entre o NADH e a InhA, possibilitando o conhecimento das propriedades dinâmicas destes sistemas, e contribuindo para a elucidação dos eventos moleculares envolvidos nos mecanismos de ação e resistência desta enzima.

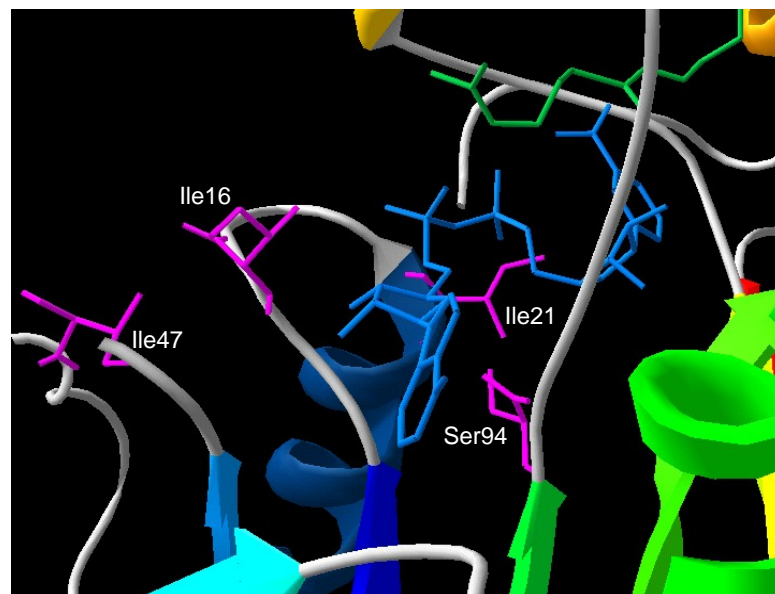


Figura 4.1. Visualização aproximada do sítio de ligação do NADH (em ciano) na enzima InhA (1BVR-C) com destaque para os aminoácidos (em magenta) do sítio de ligação do NADH que apresentam mutação em isolados clínicos resistentes à isoniazida. Enzima InhA colorida de acordo com sucessão de estruturas secundárias (azul no N-terminal e vermelho do C-terminal). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Inicialmente, foram realizadas simulações por DM do sistema InhA espécie selvagem associada ao NADH (sistema *wt* InhA-NADH) e da apo-enzima (apo-InhA) para a verificação das modificações estruturais decorrentes da formação do

complexo binário e observação dos resíduos de aminoácidos importantes para a ligação da enzima ao cofator.

Os mutantes S94A, I16T, I21V e I47T foram, então, modelados a partir da mutação *in silico* da estrutura instantânea do complexo *wt* InhA-NADH em 1,9 ns de simulação (mutação *on-the-fly*) para a verificação das mudanças que podem ocorrer nas características dinâmicas do sistema InhA-NADH quando resíduos de aminoácidos são substituídos, reproduzindo os mutantes observados em isolados clínicos resistentes à INH. Todas as simulações foram realizadas com consideração explícita do solvente (moléculas de água do tipo TIP3P - Jorgensen *et al.*, 1983), nas CNTP, em um *ensemble* isobárico-isotérmico (NPT).

Para uma avaliação mais quantitativa da diferença de afinidade das enzimas InhA (espécie selvagem e mutantes) pelo cofator e comparação com os estudos bioquímicos, estudos de *docking* molecular do NADH foram realizados nos diferentes mutantes modelados e em suas trajetórias das simulações.

4.1 – Avaliação da convergência e estabilidade das simulações por dinâmica molecular dos sistemas da InhA (espécie selvagem e mutantes)

O acompanhamento de uma simulação por dinâmica molecular pode ser realizado pelo monitoramento de diferentes parâmetros como: energia total do sistema, energias potencial e cinética, temperatura, pressão, volume, distância ou ângulo entre os átomos ou entre dois domínios, área acessível ao solvente, ângulos diedros de resíduos importantes, etc, dependendo das propriedades do sistema que se deseja observar.

Como uma medida geral da estabilidade das simulações por dinâmica molecular realizadas utilizando o conjunto de programas do pacote AMBER 6.0, as energias totais, energias cinéticas e potenciais, temperatura, volume e densidade dos respectivos sistemas foram acompanhados ao longo das trajetórias das simulações.

O monitoramento dos parâmetros energéticos das simulações mostrou que em todos os sistemas estes parâmetros alcançaram um patamar nos primeiros 150 ps das trajetórias, permanecendo estáveis após este período de equilíbrio, como pode ser visto na Figura 4.2.

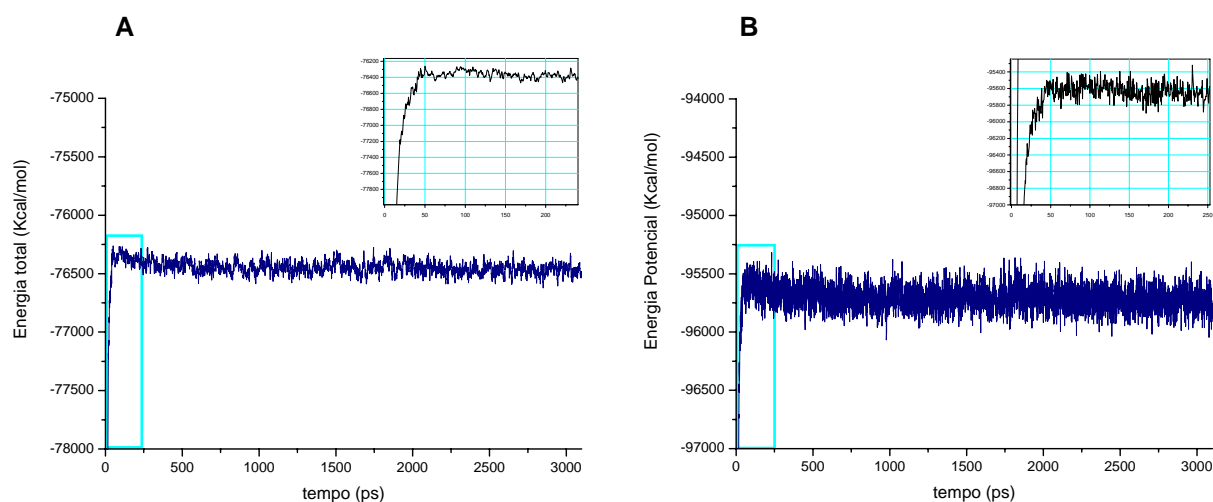


Figura 4.2. Gráficos de energia da simulação do sistema *wt* InhA-NADH. (A) Energia total e (B) energia potencial do sistema *wt* InhA-NADH ao longo de 3,1 ns de simulação, e aproximação das regiões correspondentes à fase de equilíbrio (marcadas em ciano nos gráficos maiores). Gráfico gerado com o programa Origin 6.1 (Microcal Software, Inc.).

As médias das energias totais e potenciais em todas as simulações dos sistemas InhA-NADH são da mesma ordem de grandeza, variando entre -76.140 a -76.533 kcal/mol e entre -95.269 a -95.788 kcal/mol, respectivamente. Os demais parâmetros energéticos correspondentes às ligações, ângulos diedros e interações de van der Waals e eletrostáticas também são semelhantes entre si. Da mesma forma, as temperaturas, volumes e densidades dos sistemas não apresentaram desvios significativos ao longo das trajetórias das simulações.

No estudo dos complexos da enzima InhA com o NADH, assim como da enzima em sua forma livre (apo-InhA), os movimentos estruturais que podem ser associados à maior ou menor afinidade pelo cofator devem ser observados. Para tal, os diferentes sistemas foram monitorados em relação a seus desvios médios quadráticos das posições atômicas (RMSD - *root mean square deviations*). O RMSD dos sistemas nos fornece informações sobre as mudanças conformacionais que ocorrem nestes sistemas ao longo de suas trajetórias. Assim, as estruturas instantâneas geradas nas simulações são sobrepostas a uma estrutura de referência (geralmente a estrutura cristalográfica, conformação inicial do sistema) e o desvio em relação a esta referência é calculado como forma de comparação entre as duas estruturas.

O acompanhamento do RMSD dos sistemas ao longo das trajetórias em função do tempo foi o principal parâmetro utilizado para avaliação da estabilidade das simulações. A apresentação e discussão deste parâmetro é feita ao longo deste capítulo, à medida em que os sistemas são tratados.

A avaliação da área total da superfície acessível ao solvente (ATSAS), assim como o raio de giro (Rg) também fornecem parâmetros estruturais importantes, sendo indicativos de como a estrutura da macromolécula varia em função do tempo no ambiente de simulação. O raio de giro, definido como a média ponderada da distribuição das massas atômicas em relação ao centro de massa da macromolécula, mede a compactação da macromolécula e, em última instância, nos fornece uma estimativa da variação do tamanho da proteína durante as simulações.

A manutenção das estruturas secundárias das proteínas ao longo das trajetórias também foram avaliadas em termos da manutenção do número total de pontes de hidrogênio intramoleculares (NTPH). Quando um sistema

macromolecular é simulado com consideração explícita das moléculas de solvente, algumas pontes de hidrogênio intramoleculares podem ser naturalmente substituídas por interações com moléculas de solvente ao longo da trajetória. Entretanto, de forma geral e se o sistema de simulação é estável, as estruturas secundárias devem ser mantidas.

Um resumo geral desses parâmetros encontra-se na tabela abaixo (Tabela 4.1) e foram acompanhados durante toda a trajetória das simulações por DM dos sistemas em estudo. Entretanto, para efeito de análise, somente os últimos 2 ns das trajetórias (fase de produção), quando os sistemas foram considerados estáveis de acordo com os critérios de convergência utilizados, são apresentados. Uma análise mais detalhada destes parâmetros estruturais foi colocada, quando necessário, na apresentação e discussão de cada um dos sistemas simulados.

Tabela 4.1. Resumo geral dos parâmetros estruturais da InhA observados nos últimos 2 ns das simulações por dinâmica molecular.

Parâmetros	<i>apo</i> -InhA	<i>wt</i> InhA-NADH	<i>cWAT</i> -NADH	S94A-NADH	I16T-NADH	I21V-NADH	I47T-NADH
RMSD_{ep} (Å) ^a	2,0 ± 0,10	1,6 ± 0,1	2,0 ± 0,1	1,7 ± 0,1	2,0 ± 0,1	2,1 ± 0,1	1,7 ± 0,1
Rg_{ep} (Å) ^b	17,9 ± 0,1	18,1 ± 0,1	17,9 ± 0,1	18,0 ± 0,1	18,3 ± 0,1	18,0 ± 0,2	18,0 ± 0,1
ATSAS (Å ²) ^c	12.977 ± 183	12.854 ± 136	13.110 ± 113	12.748 ± 150	13.250 ± 128	13.056 ± 177	12.823 ± 125
NTPH ^d	173 ± 4	178 ± 5	176 ± 4	177 ± 5	175 ± 4	174 ± 5	175 ± 4
Tempo total de simulação	3,95 ns	3,1 ns	4,0 ns	3,1 ns	4,0 ns	6,5 ns	5,5 ns

^aRMSD_{ep}: desvio médio quadrático das posições atômicas do esqueleto polipeptídico em relação à estrutura cristalográfica 1ENY, calculado com o programa ptraj-6.4 do AMBER 6.0.

^bRg_{ep}: Raio de giro do esqueleto polipeptídico, calculado com o programa Carnal do AMBER 6.0. Rg 1ENY = 17,9 Å.

^cATSAS: área total da superfície acessível ao solvente, calculado com o programa NACCESS (Hubbard & Thornton, 1993). ATSAS 1ENY = 12.643 Å².

^dNTPH: número total de pontes de H intramoleculares na estrutura da enzima, calculado com o programa DSSP (Kabsch & Sander, 1983). NTPH 1ENY = 203.

4.2 – Simulação do sistema *wt* InhA-NADH e da apo-InhA

As simulações da enzima InhA espécie selvagem na sua forma livre em solução (apo-InhA) e associada ao cofator NADH (*wt* InhA-NADH) foram realizadas com o pacote AMBER 6.0, com consideração explícita do solvente (água), por 3,1 ns e 3,95 ns, respectivamente, utilizando-se como estrutura inicial a estrutura cristalina 1ENY que apresenta resolução de 2,2 Å.

As simulações por DM das enzimas InhA espécie selvagem geraram trajetórias estáveis com valores de RMSD do esqueleto polipeptídico, calculados em relação à estrutura cristalina 1ENY, de $1,6 \pm 0,1$ Å (*wt* InhA-NADH) e $2,0 \pm 0,1$ Å (apo-InhA) para os últimos 2 ns de simulação (Figura 4.3).

A pequena variação no número total de pontes de hidrogênio (NTPH - Tabela 4.1), calculada para a série de estruturas instantâneas geradas, também indicam a conservação global das estruturas.

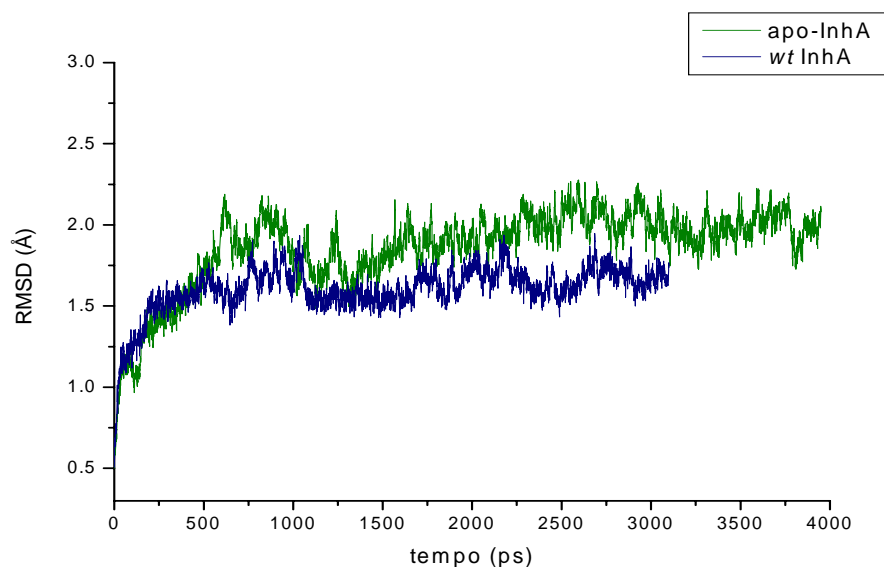


Figura 4.3. Gráfico do RMSD do esqueleto polipeptídico em função do tempo para as trajetórias dos sistemas *wt* InhA-NADH (azul escuro) e apo-InhA (verde). RMSD calculado em relação à estrutura cristalina 1ENY. Gráfico gerado com o programa Origin 6.1 (Microcal Software, Inc.).

De acordo com avaliações dos conjuntos de estruturas médias de cada 250 ps de simulação realizadas com o programa PROCHECK-NMR (Laskowski *et al.*, 1993), grande parte das estruturas secundárias que formam o *Rossmann fold* da InhA permanece intacta em ambas as trajetórias, principalmente as fitas β centrais que formam o núcleo de ligação do substrato (Figura 4.4). Na simulação do sistema *wt* InhA-NADH, a enzima apresenta a sua hélice $\alpha 5$ um pouco menos estável, e uma de suas voltas (onde se localiza a Lys165, um importante resíduo para a ligação do NADH) é momentaneamente desfeita. A hélice $\alpha 5$ é uma α -hélice externa e longa que no cristal faz contato com outras subunidades. O fato desta hélice não estar mais envolvida em interações entre subunidades no ambiente de simulação permite que seu esqueleto polipeptídico experimente conformações alternativas, provavelmente induzidas pela associação da cadeia lateral da Lys165 com o cofator. Na simulação da apo-InhA, este comportamento da hélice $\alpha 5$ não é observado (não há cofator para associação da cadeia lateral da Lys165), entretanto, a hélice $\alpha 6$ que pertence à alça de ligação do substrato é temporariamente desfeita, sendo recuperada adiante na trajetória.

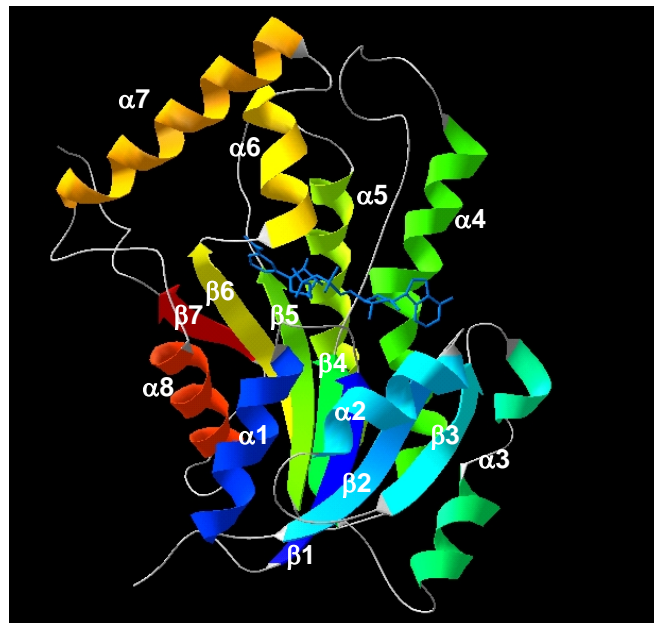


Figura 4.4. Identificação das estruturas secundárias na estrutura da enzima InhA (colorida de acordo com a sucessão de estruturas secundárias, em azul a porção N-terminal e em vermelho a porção C-terminal) e molécula de NADH (azul) sobre o leque de fitas β do *Rossmann fold*. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A flexibilidade de uma macromolécula é refletida na flutuação de cada átomo em relação à sua posição média, durante a fase de produção, e pode ser estimada a partir do cálculo dos fatores de temperatura (*B-factor*) (Hünenberger *et al.*, 1995). O *B-factor* calculado para o esqueleto polipeptídico das enzimas em relação às respectivas estruturas médias das fases de produção mostrou um perfil de flexibilidade semelhante entre a enzima livre e a enzima complexada (Figura 4.5). Em ambas as simulações, as regiões correspondentes às fitas β e à maioria das α -hélices apresentam um *B-factor* menor, enquanto nas alças e regiões de maior mobilidade estes valores são maiores³⁸. Em sua maioria, os aminoácidos do sítio de ligação do cofator possuem fatores de temperatura menores na *wt* InhA-NADH do que na apo-enzima, onde a restrição imposta pela ligação com o NADH não ocorre.

Entre as regiões de maior mobilidade estão: (i) a região dos resíduos 82-88, que corresponde a alça entre α 3- β 4, cujo movimento no cristal está parcialmente restrito pelo contato com a subunidade vizinha; (ii) as regiões dos resíduos 100-114

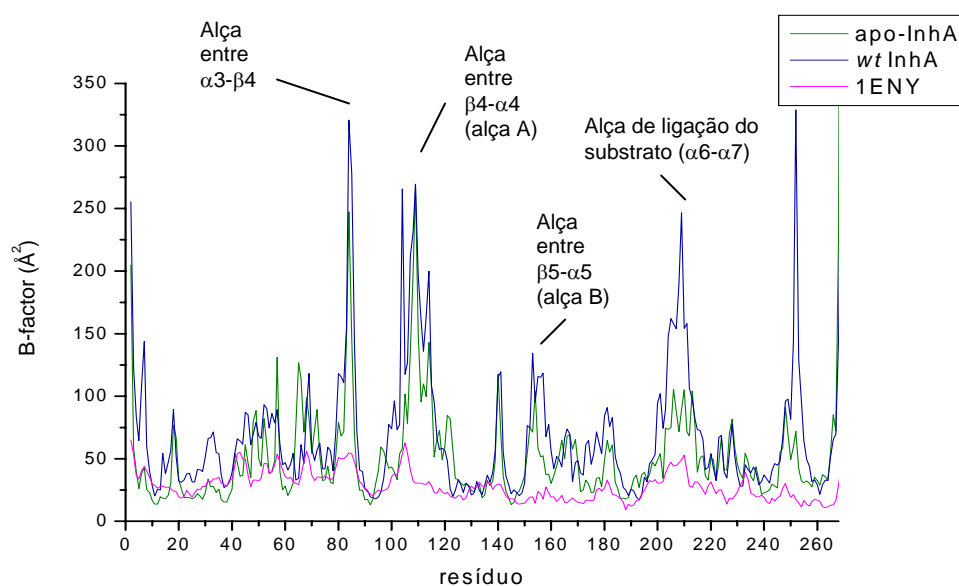


Figura 4.5. Fatores de temperatura (*B-factor*) da cadeia principal dos resíduos da enzima InhA nas simulações *wt* InhA-NADH (azul) e apo InhA (verde) calculados sobre os últimos 2 ns de simulação, e do cristal 1ENY (magenta). Gráfico gerado com o programa Origin 6.1 (Microcal Software, Inc.).

³⁸ Cabe ressaltar que não é esperada uma concordância numérica exata entre os valores de *B-factor* cristalográfico e os de DM devido a diferenças no tempo de amostragem e no cálculo deste parâmetro. Efeitos de empacotamento do retículo cristalino também podem influenciar a mobilidade de alguns resíduos (Bakowies & van Gusnteren, 2002; Hünenberger *et al.*, 1995; Stocker *et al.*, 2000).

(alça entre $\beta 4$ - $\alpha 4$) e 150-161 (alça entre $\beta 5$ - $\alpha 5$) que formam um dos lados da cavidade de ligação do substrato, e no cristal também fazem contato com a subunidade vizinha; e (iii) a região entre os resíduos 197 e 220 que correspondem à parte da alça de ligação do substrato e compreende as α -hélices $\alpha 6$ e $\alpha 7$, perpendiculares entre si.

A flexibilidade dos sistemas estudados, principalmente das alças próximas à $\alpha 4$ (alça A) e à $\alpha 5$ (alça B) e da alça de ligação do substrato, pode ser observada logo no início das simulações. Para ilustrar as mudanças conformacionais que ocorrem durante a trajetória, a Figura 4.6A apresenta a sobreposição de 4 estruturas instantâneas (*snapshots*) obtidas em diferentes momentos da simulação *wt* InhA-NADH³⁹.

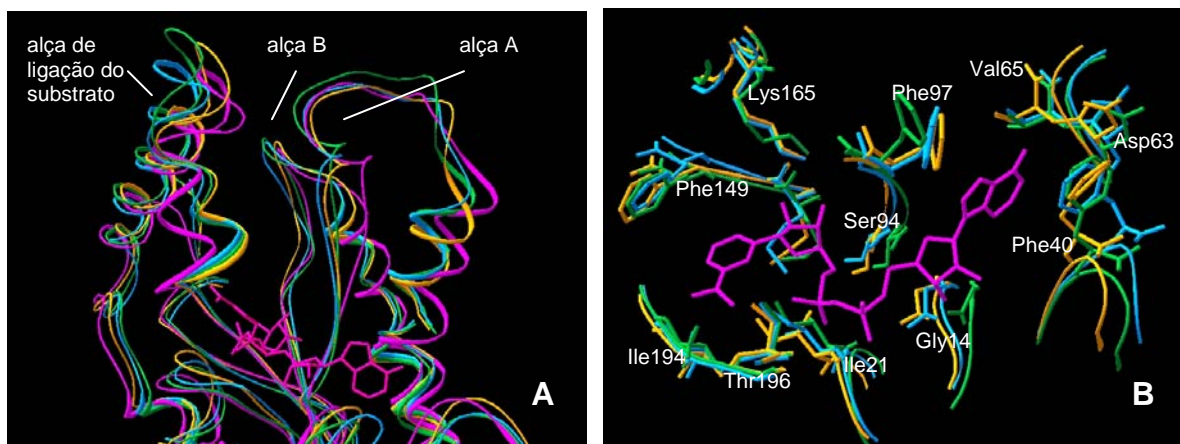


Figura 4.6. Flexibilidade do sistema *wt* InhA-NADH ao longo da simulação por dinâmica molecular. (A) Flexibilidade da alça de ligação do substrato (resíduos 196-219) e das alças A e B. (B) Flexibilidade dos aminoácidos do sítio de ligação do NADH. 1ENY em magenta, complexo *wt* InhA-NADH em 1ns de simulação (verde); em 2ns de simulação (ciano) e em 3ns de simulação (amarelo). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Na escala de tempo em que as simulações foram realizadas, não foi observado o movimento amplo de abertura da alça de ligação do substrato (resíduos 196-219), como o esperado para a associação do substrato (Rozwarski *et al.*, 1999). Movimentos deste tipo (*hinge-bending* motions) costumam ocorrer em uma escala de tempo superior à utilizada neste estudo, normalmente na ordem de

³⁹ Embora estas quatro estruturas não apresentem todas as conformações que a cadeia polipeptídica pode assumir durante a simulação, elas caracterizam o tipo de movimento durante a trajetória.

centenas de ns (Benkovic & Hammes-Schiffer, 2003). Entretanto, no complexo *wt* InhA-NADH, a distância dos centros de massa dos aminoácidos que compõem a hélice α_6 (resíduos 197-220) da alça de ligação do substrato e da alça A (resíduos 100-111) (entre β_4 - α_4) variou entre 12,4 Å e 16,2 Å (distância média = $14,5 \pm 0,5$ Å). O afastamento de mais de 3,7 Å indica que este movimento de abertura é possível, e a aproximação do substrato também deve ajudar na indução do afastamento efetivo destas alças.

A flexibilidade da cadeia principal do sistema *wt* InhA-NADH também é refletida no sítio de ligação do NADH, onde pode-se observar movimentos importantes nas cadeias do esqueleto polipeptídico, principalmente na Gly14 e nos aminoácidos próximos à região de ligação do anel adenina (aminoácidos Asp63 e Phe40 na Figura 4.6B). Esta flexibilidade é característica do sistema, não podendo ser atribuída a mutações e, portanto, não pode ser associada com a resistência à INH. Apesar da grande mobilidade, o número de aminoácidos no sítio de ligação do cofator (4 Å do NADH) permanece constante (16 ± 2 aminoácidos) ao longo da trajetória.

Pela análise do raio de giro (R_g) das duas simulações (Figura 4.7), pode-se notar que, enquanto a enzima do sistema *wt* InhA-NADH apresenta uma raio de

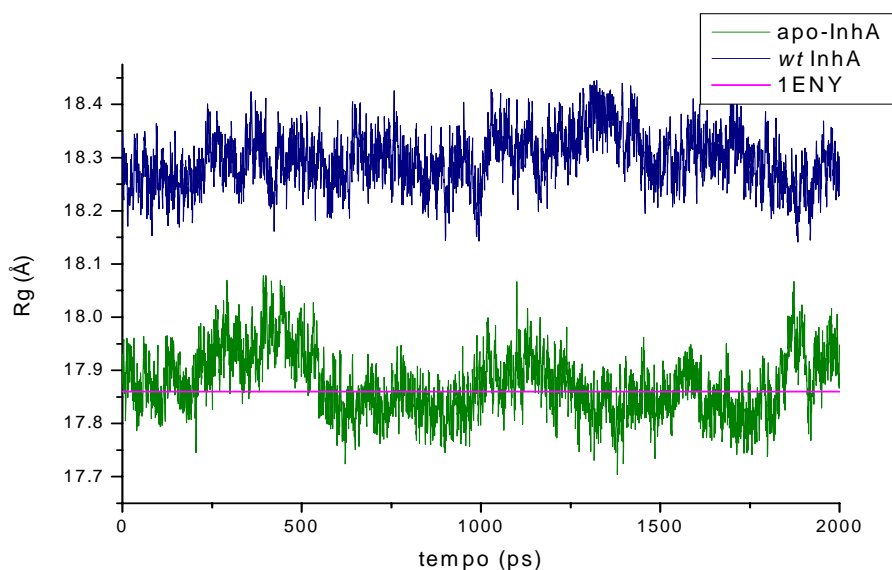


Figura 4.7. Gráfico do raio de giro (R_g) para os últimos 2 ns das simulações *wt* InhA-NADH (azul escuro) e apo-InhA (verde), e comparação com o R_g do cristal 1ENY (magenta). Gráfico gerado com o programa Origin 6.1 (Microcal Software, Inc.).

giro maior do que o apresentado pelo cristal 1 ENY (o que é esperado dada a flexibilidade do sistema e à ausência das interações do retículo cristalino, permitindo o relaxamento do sistema), a apo-InhA assume uma estrutura mais compacta devido à ausência do cofator.

Esta compactação foi mais tarde refletida nos estudos de *docking* do NADH na estrutura da InhA. Enquanto o modo de ligação do NADH foi facilmente reproduzido na enzima do sistema *wt* InhA-NADH, seu *docking* não foi possível nas estruturas geradas na trajetória do sistema apo-InhA, onde as cadeias laterais dos aminoácidos do sítio de ligação do NADH causam um impedimento estérico para o atracamento do ligante.

4.2.1 - Comportamento dinâmico dos resíduos conservados Lys165, Tyr158 e Phe149 no sítio de ligação do cofator

Como membro da família das SDRs que utiliza dinucleotídeos como cofator (Jörnvall *et al.*, 1995), a InhA apresenta os resíduos Tyr158 e Lys165 conservados (Duax *et al.*, 1996). Estes resíduos, juntamente com a Phe149, estão implicados no mecanismo de ação enzimático (Rozwarski *et al.*, 1999) (Tabela 4.2).

Tabela 4.2. Resíduos de aminoácidos implicados no mecanismo catalítico da enzima InhA de MTB.

Aminoácido	Importância
Lys 165	Associação com as OHs da ribose N do NADH
Tyr 158	Estabilização do enolato intermediário da reação
Phe149	Fixação da conformação em forma de “U” da cadeia hidrofóbica do substrato, e controle do acesso de moléculas de água no centro catalítico

Dadas suas importâncias, esses resíduos foram acompanhados ao longo das simulações para verificação de seus comportamentos dinâmicos.

Apesar do resíduo conservado Lys165 ter sua função inicialmente atribuída à estabilização do enolato intermediário formado na redução do substrato pela InhA

(Quémard *et al.*, 1995), estudos de mutação sítio-direcionada (Parikh *et al.*, 1999) e a determinação da estrutura do complexo ternário 1BVR (Rozwarski *et al.*, 1999) demonstraram que este resíduo está envolvido na ligação do cofator no sítio de ligação da InhA.

A comparação das trajetórias do complexo *wt* InhA-NADH e da apo-InhA mostra que, na forma não ligada, este resíduo é extremamente flexível, enquanto na forma ligada sua mobilidade é diminuída pela presença do cofator (Figura 4.8). Na *wt* InhA-NADH, a cadeia lateral da Lys165 forma uma ponte de hidrogênio forte com as hidroxilas da ribose N do NADH, permanecendo ligada ao O'N2 e O'N3 durante 83% e 99% do tempo, respectivamente, nos últimos 2 ns de simulação, confirmando seu envolvimento na ligação com o cofator.

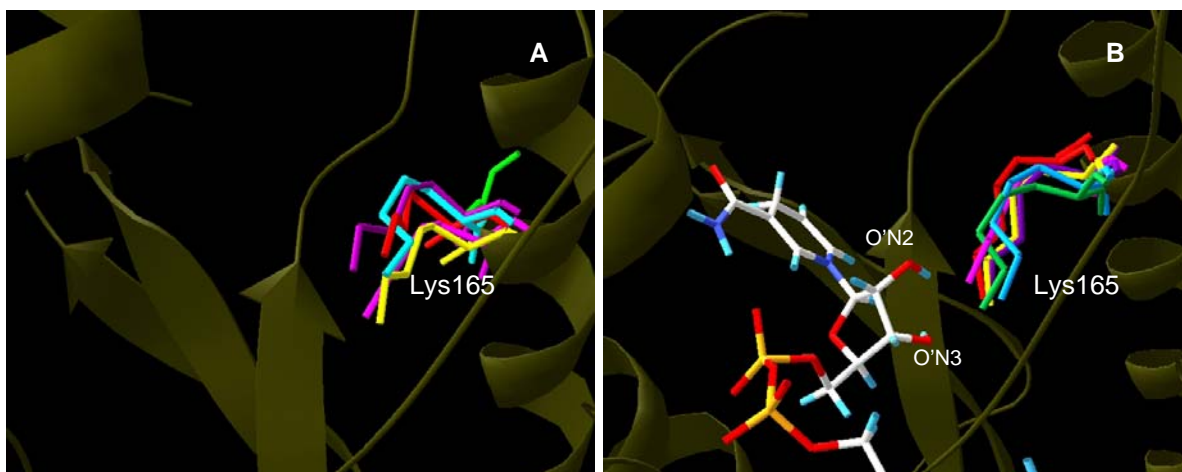


Figura 4.8. Flexibilidade da cadeia lateral da Lys165 nos últimos 2ns das simulações apo-InhA e *wt* InhA-NADH. (A) Sobreposição das estruturas instantâneas da cadeia lateral da Lys165 na simulação da apo-InhA (2 ns em ciano, 2,5 ns em verde, 3 ns em amarelo, 3,5ns em vermelho e 3,95 ns em violeta) e comparação com a estrutura cristalina 1ENY (Lys165 em magenta e estrutura terciária em ocre). (B) Sobreposição das estruturas instantâneas da cadeia lateral da Lys165 na simulação *wt* InhA-NADH (1 ns em ciano, 1,5 ns em verde, 2 ns em amarelo, 2,5 ns em vermelho e 3,0 ns em violeta), confirmando a importância deste resíduo na fixação do NADH na cavidade de ligação. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

O resíduo Tyr158, implicado na interação da enzima com o substrato (Rozwarski *et al.*, 1999; Parikh *et al.*, 1999), não experimenta rotação da cadeia lateral na simulação com o cofator e assume uma conformação próxima à observada na estrutura cristalográfica (diedro Tyr158 $Chi_1 = -75,7^0 \pm 8,6^0$ na *wt* InhA-NADH e $-67,9^0$ na 1ENY) (Figura 4.9). Este direcionamento da cadeia lateral

da Tyr158 ocorre devido à formação de uma ponte de hidrogênio entre sua hidroxila e uma molécula de água ordenada (wat 3662) que, na simulação *wt* InhA–NADH, também interage com a cadeia 3-carboxiamida da porção nicotinamida do NADH (anel N). Esta ponte de hidrogênio formada entre a Tyr158 e uma molécula de água no fundo da cavidade de ligação também é observada nas estruturas cristalográficas onde o sítio de ligação do substrato não está ocupado.

Como observado na estrutura do complexo ternário InhA–NAD⁺–C16 (Rozwarski *et al.*, 1999) e em estudos da inibição do InhA por triclosan via simulação por DM (Schroeder *et al.*, 2002), mais tarde confirmados pela determinação da estrutura cristalográfica deste complexo (Kuo *et al.*, 2003), a presença do substrato ou inibidor no sítio de ligação da enzima deslocam esta molécula de água e a ligação indireta com o cofator é desfeita, induzindo o rearranjo da cadeia lateral da Tyr158 para a associação com o ligante (Figura 4.10). Na ausência do cofator (apo-InhA) a mobilidade da cadeia lateral da Tyr158 não está restrita e pode sofrer rotação em relação à sua posição na 1ENY (diedro Tyr158 $Chi_1 = -93,6^\circ \pm 34,4^\circ$ para a simulação, e $-67,9^\circ$ na 1ENY), direcionando-se para o centro da cavidade da enzima (Figura 4.9).

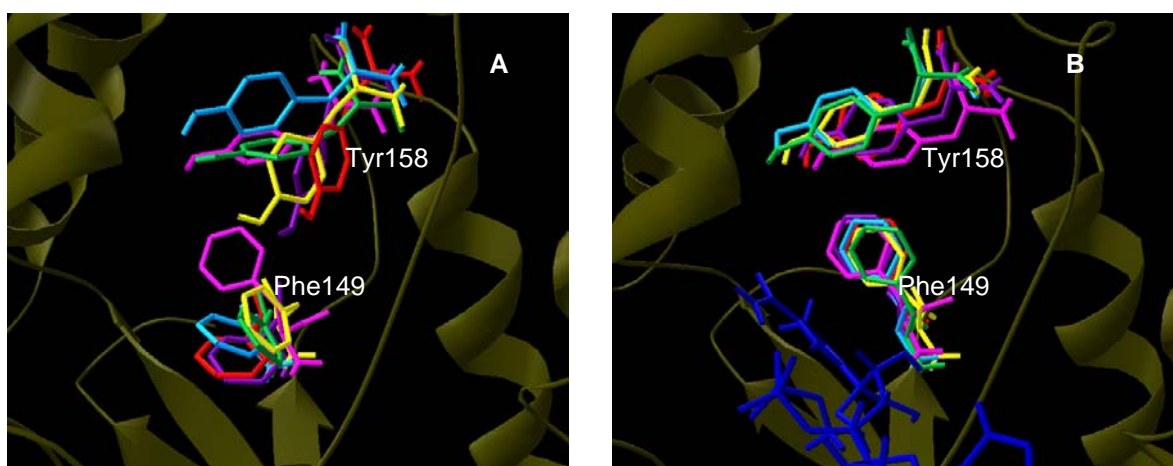


Figura 4.9. Flexibilidade das cadeias laterais da Tyr158 e Phe149 nos últimos 2 ns das simulações apo-InhA e *wt* InhA–NADH. (A) Sobreposição das estruturas instantâneas das cadeias laterais da Tyr158 e Phe149 na simulação da apo-InhA (2 ns em ciano, 2.5 ns em verde, 3 ns em amarelo, 3.5 ns em vermelho e 3.95 ns violeta) e comparação com a estrutura cristalina 1ENY (Tyr158 e Phe149 em magenta e estrutura terciária em ocre). (B) Sobreposição das estruturas instantâneas das cadeias laterais da Tyr158 e Phe149 na simulação *wt* InhA–NADH (1 ns em ciano, 1.5 ns em verde, 2 ns em amarelo, 2.5 ns em vermelho e 3,0 ns em violeta) e comparação com a estrutura cristalina 1ENY (Tyr158 e Phe149 em magenta e estrutura terciária em ocre). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

O resíduo Phe149, além de ser importante para a fixação da conformação do substrato no sítio ativo da InhA, também poderia desempenhar uma função importante no mecanismo de reação, permitindo o acesso de moléculas de água (localizadas em um canal de água delimitado pelos resíduos Asp150, Arg153, Trp222, Tyr159 e His265, logo atrás de seu anel de fenila – Figura 1.18) que poderiam fornecer o próton necessário para a redução do enolato intermediariamente formado (Rozwarski *et al.*, 1999). A possibilidade de rotação da cadeia lateral deste aminoácido foi verificada na estrutura cristalográfica da InhA inibida pelo bi-substrato NAD-INH (1ZID - Rozwarski *et al.*, 1998) (Figura 1.21).

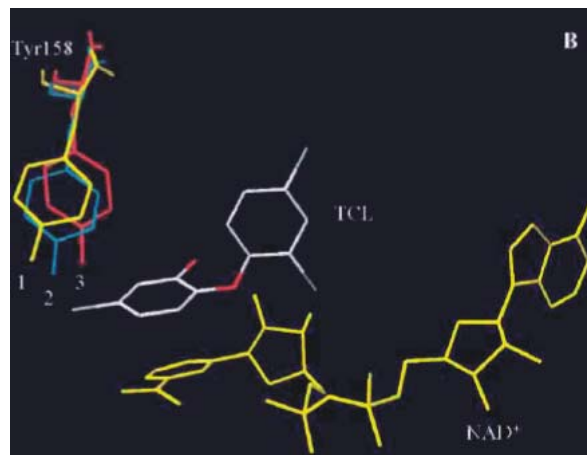


Figura 4.10. Movimentação da cadeia lateral da Tyr158 durante a simulação do complexo ternário formado entre a *wt* InhA, o NAD⁺ (amarelo) e o TCL (colorido de acordo com o tipo de átomo) no estudo do modo de ligação do TCL na Enoil Redutase de *M. tuberculosis*. Posição da Tyr158 na estrutura cristalográfica 1ENY (1), em 20ps (2) e em 28ps de simulação (3). Figura retirada de Schroeder *et al.*, 2002.

Ao longo da trajetória da *wt* InhA-NADH, esse canal de água contém de 1-2 moléculas de água próximas ao anel Phe149 que poderiam ter acesso ao interior da cavidade da enzima pela movimentação da cadeia lateral da Phe149. Entretanto, o monitoramento dos ângulos diedros Chi_1 e Chi_2 (C_α - C_β - C_γ - $C_{\delta 1}$) da Phe149 durante a trajetória: $-166,4^\circ \pm 7,5^\circ$ e $28,5^\circ \pm 16,8^\circ$, respectivamente para os últimos 2ns de simulação (na 1ENY: $-163,7^\circ$ e $63,4^\circ$), demonstrou que, ao contrário do que é observado na apo-InhA, este resíduo não sofreu rotação em sua cadeia lateral ao longo da simulação (Figuras 4.9A e B). Isto, em parte pode ser explicado pela proximidade da cadeia lateral da Tyr158 que se direciona para o fundo da

cavidade de ligação do NADH e tem sua mobilidade reduzida pela associação indireta com a 3-carboxiamida do NADH (via pontes de hidrogênios com uma molécula de água), permitindo também a aproximação da cadeia lateral da Met161. No sistema *wt* InhA-NADH, a proximidade destes dois resíduos causa um impedimento estérico para a movimentação da cadeia lateral da Phe149.

Ainda resta a questão se na presença do enolato intermediário esta rotação seria favorecida para a entrada de uma molécula de água que fornecesse o próton necessário para a redução do segundo carbono da ligação dupla. Tal possibilidade poderia ser verificada pela simulação do complexo ternário formado entre a enzima, o NAD⁺ e esse enolato intermediário. Nesta situação, a cadeia lateral da Tyr158 estaria impedida de se associar indiretamente à cadeia 3-carboxiamida do NADH e, assim como observado na simulação *wt* InhA-NAD⁺-TCL (Schroeder *et al.*, 2002 – Figura 4.10), se deslocaria para a associação com a carbonila do substrato, afastando também a cadeia lateral da Met161 e liberando espaço para a rotação da Phe149.

4.2.2 – Ligação do NADH no sítio ativo da enzima *wt* InhA

A InhA faz parte de um grupo de enzimas que ligam NAD(H). Quando ligado a estas enzimas, o NAD(H) normalmente adota uma conformação estendida com uma distância entre o C2 do anel nicotinamida (anel N) e o C6 da adenina (anel A) que pode variar de 8 a 18 Å (Bell *et al.*, 1997).

Em solução aquosa, o NAD⁺ assume uma conformação fechada com distância entre o anel N e o anel A de 5 Å devido a interações π -stacking entre esses anéis aromáticos (Smith & Tanner, 1999). O NADH, apesar de experimentar conformações estendidas na ordem de 15 Å de distância entre seus anéis, assume conformações mais variadas, com distâncias entre C2N e C6A de 5 a 10 Å (Apêndice 6).

Para a catálise, a conformação estendida é preferida, pois expõe uma maior área da molécula para a interação com enzima e substrato reduzindo, também, o

volume de solvente em torno do anel N para permitir a transferência de hidreto (Lau *et al.*, 2003; Smith & Tanner, 2000). A maioria das enzimas que ligam o NAD(H) na forma estendida apresenta um domínio central do tipo *Rossmann fold* (Richardson, 1981). Neste domínio, o NAD(H) pode se ligar de diferentes formas (em relação à face do anel N disponibilizada e à orientação do anel A), dependendo das características do receptor. A orientação do NAD(H) no sítio da enzima determina a estereoespecificidade da reação a ser catalisada (Fillgrove & Anderson, 2000).

Na simulação *wt* InhA-NADH, a molécula de NADH mantém uma conformação estendida, com distância C2N-C6A de $15,6 \pm 0,3$ Å para os 3,1 ns de simulação, e $15,7 \pm 0,3$ Å nos últimos 2 ns. Esta conformação estendida é mantida pela interação do cofator com cadeias laterais de aminoácidos polares, com a C=O ou NH do esqueleto peptídico e com moléculas de água bem orientadas.

A interação do NADH com a enzima InhA segue um padrão de ligação comum entre as enzimas da família das SDRs com estrutura do tipo *Rossmann fold*, como a presença de uma alça rica em resíduos glicina entre β_1 e α_1 (resíduos 14-21), envolvida na interação entre a proteína e a porção pirofosfato do cofator, um resíduo carboxilado (Asp 42) logo depois de β_2 , que pode interagir com a hidroxila da ribose próxima à adenina (ribose A) (Carugo & Argos, 1997) e um sítio de reconhecimento do anel adenina com a presença de um resíduo aspartato ou asparagina (na InhA: Asp64), seguido de um resíduo hidrofóbico (Val65) e um resíduo hidrofílico ou polar (Gln66) (Denessiouk *et al.*, 2001). As enoil-tioester redutases ainda possuem um resíduo treonina conservado que se localiza próximo ao sítio de ligação do anel N do NADH (Thr196) (Fillgrove & Anderson, 2000) (Figura 4.11).

Na InhA, esses e outros resíduos, assim como moléculas de água do sítio ativo da enzima, são responsáveis pela ligação do cofator em uma conformação estendida.

As pontes de hidrogênio formadas entre o NADH e seu sítio de ligação na enzima foram monitoradas ao longo da trajetória da simulação da *wt* InhA-NADH e aquelas pontes de hidrogênio que apresentam uma ocorrência de no mínimo 50% do tempo e, portanto, mais importantes para a interação, estão descritas na Tabela 4.3.

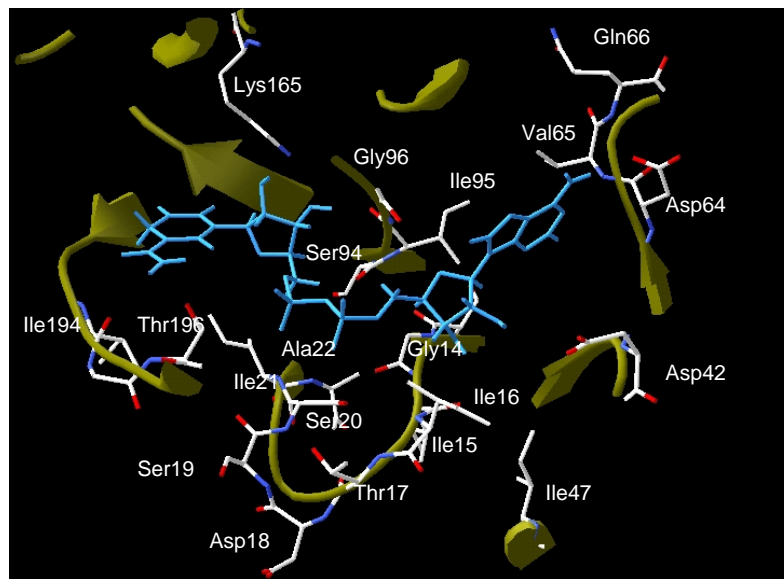


Figura 4.11. Estrutura média do sítio de ligação do NADH (azul) nos últimos 2 ns de simulação do complexo *wt* InhA-NADH (estrutura secundária em ocre e aminoácidos coloridos de acordo com o tipo de átomo). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Tabela 4.3. Pontes de hidrogênio formadas entre o NADH e aminoácidos do sítio de ligação da InhA espécie selvagem com ocorrência maior do que 50% durante os últimos 2 ns de simulação.

Resíduo	Grupo do aminoácido	Átomo do NADH ^a	Ocorrência ^b (%)
Ile15	C=O	O'A3	70
	N	O'A3	72
Ser20	OH	OPA2	100
Ile21	N	OPN2	95
Asp64	COO-	N6A	99
Val65	N	N1A	100
Ile95	C=O	O'N3	100
Gly96	N	O'A4	100
Lys165	NZ	O'N2	83
	NZ	O'N3	99
Ile194	C=O	N7N	95
	N	O7N	100
Thr196	OG1	N7N	67
	OG1	OPN1	61

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1.

^bOcorrência calculada com o programa Carnal do AMBER6.0.

No fundo da cavidade de ligação, o grupo 3-carboxiamida do anel N forma uma forte ponte de hidrogênio (distância O-N de $3,0 \pm 2$ Å com ocorrência de 100%) com o NH da cadeia principal da Ile194, e com a cadeia lateral do resíduo conservado Thr196. Este último resíduo também se liga ao oxigênio OPN1 da porção pirofosfato do NADH.

A ribose próxima ao anel nicotinamida (ribose N) interage com a enzima por pontes de hidrogênio entre suas hidroxilas e a cadeia lateral da Lys165 e cadeia principal da Ile95, durante toda a fase de produção.

Os resíduos Ser20 e Ile21 da alça de glicina interagem fortemente com a porção pirofosfato do NADH, enquanto a Ile15 se associa a uma das hidroxilas da ribose próxima à adenina (ribose A).

O oxigênio do anel ribose A também permanece ligado ao NH da cadeia principal da Gly96 durante a trajetória.

Na InhA, a cadeia lateral do resíduo Asp64 e a cadeia principal do resíduo Val65 são responsáveis pela interação do anel adenina (anel A) com seu sítio de reconhecimento na enzima InhA.

Durante a trajetória todos os resíduos que interagem com o cofator na estrutura cristalográfica foram mantidos na maior parte da simulação, exceto a Gly14, que sofre um dos maiores rearranjos conformacionais e só mantém sua interação com o O'A3 durante 22% do tempo. Outras interações diretas com a enzima, e importantes para a ligação do cofator, foram a associação do OH da Ser20 com o O'A5 do pirofosfato (ocorrência de 48%) e da C=O da Ser94 com o O'N3 da ribose N (ocorrência de 32%). De acordo com cálculos realizados com o programa HBPLUS (McDonald & Thornton, 1994), durante a trajetória do complexo *wt* InhA-NADH o número de pontes de hidrogênio entre a proteína e o NADH é mantido constante ($11,9 \pm 1,2$ pontes de hidrogênio).

A simulação do sistema com consideração explícita do solvente permitiu que as interações entre a enzima e o cofator mediadas por moléculas de água fossem avaliadas. Na fase de produção, a molécula de NADH do sistema *wt* InhA-NADH apresenta $19,1 \pm 1,4$ moléculas de água em sua 1ª esfera de solvatação (menos de 3,5 Å de distância) e $27,6 \pm 2,4$ moléculas de água em sua 2ª esfera de solvatação (menos de 5 Å de distância), valores consideravelmente menores do que os observados na simulação do NADH livre em solução ($62,8 \pm 5,7$ na 1ª esfera e

121,4 ± 9,1 na 2ª esfera de solvatação⁴⁰) (Apêndice 6). O menor número de moléculas de água em torno do NADH está de acordo com o fato da molécula do cofator estar associada à enzima, em uma cavidade predominantemente hidrofóbica.

Como esperado, a maior parte das moléculas de água que fazem parte da 1ª esfera de solvatação no sistema *wt* InhA-NADH está concentrada próximo à porção pirofosfato do NADH, que possui 8,1 ± 0,8 moléculas de água.

Pela característica hidrofóbica do fundo da cavidade de ligação do NADH, o anel N e a ribose N permanecem menos hidratados com 1,9 ± 0,4 e 1,7 ± 0,6 moléculas de água, respectivamente, a 3,5 Å de distância. A ribose A e o anel adenina, por estarem mais expostos ao solvente apresentaram 3,6 ± 0,9 e 4,4 ± 1,0 moléculas de água em sua 1ª camada de solvatação, respectivamente.

Algumas destas moléculas de água formam pontes de hidrogênio entre o NADH e resíduos de aminoácidos da proteína e contribuem significativamente para a interface de ligação do dinucleotídeo em todas as enzimas com estrutura do tipo *Rossmann fold* (Bottoms *et al.*, 2002).

Ao longo da trajetória moléculas de água do solvente entram e saem do sítio de ligação da enzima interagindo momentaneamente com a enzima e/ou com o cofator. Entretanto, algumas moléculas de água, que participam das redes de ponte de hidrogênio que são importantes para a interação do ligante no sítio do receptor, apresentam uma baixa mobilidade e um grande tempo de residência neste local.

Na simulação *wt* InhA-NADH, seis moléculas de água fazem importantes pontes de hidrogênio com o cofator (Tabela 4.4 e Figura 4.12) participando também de interações com a proteína. Duas delas localizam-se no fundo da cavidade de ligação (wat4149 e wat3662) interagindo com o grupo 3-carboxiamida do anel N do NADH.

A molécula de água wat3662 é responsável pela interação indireta da Tyr158 com o grupo amida do NADH neste sistema de simulação, direcionando a cadeia lateral deste aminoácido para o fundo do sítio de ligação, e reproduzindo a água observada na estrutura cristalina do 1ENY.

⁴⁰ Primeira esfera de solvatação do NADH durante a trajetória da simulação da molécula livre em solução: anel N = 12,2 ± 2,3 , ribose N = 10,1 ± 2,0, pirofosfato = 15,6 ± 1,7, ribose A = 10,8 ± 2,2, anel A = 13,3 ± 2,5 moléculas de água.

A molécula de água wat4149 está envolvida na formação de uma rede de ponte de hidrogênio com o NADH, interagindo com o NH₂ do grupo 3-carboxiamida, com os átomos OPN1 e O'N5 do pirofosfato e com a molécula de água wat4148. Esta última, por sua vez, é uma molécula de água bem ordenada e está localizada próxima à alça da glicina se ligando à porção pirofosfato do NADH (OPN1, OPN2 e OPA2) e interagindo com a cadeia lateral da Ser18 e com a cadeia principal da Leu197 da hélice α 6 da alça de ligação do substrato.

Tabela 4.4. Pontes de hidrogênio formadas entre o NADH e moléculas de água que medeiam a interação com o sítio de ligação da *wt* InhA com ocorrência maior do que 50% durante os últimos 2 ns de simulação.

Molécula de água	Átomo da água	Átomo do NADH ^a	Ocorrência (%) ^b
wat1156	O	N7N	87
wat3662	O-H1	O7N	66
wat4125	O-H2	OPN1	53
wat4128	O	O'2N	93
wat4149	O	N7N	97
	O-H2	OPN1	55
	O-H2	O'N5	52

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1.

^bOcorrência calculada com o programa Carnal do AMBER6.0. O programa Carnal diferencia os hidrogênios das moléculas de água como H1 e H2.

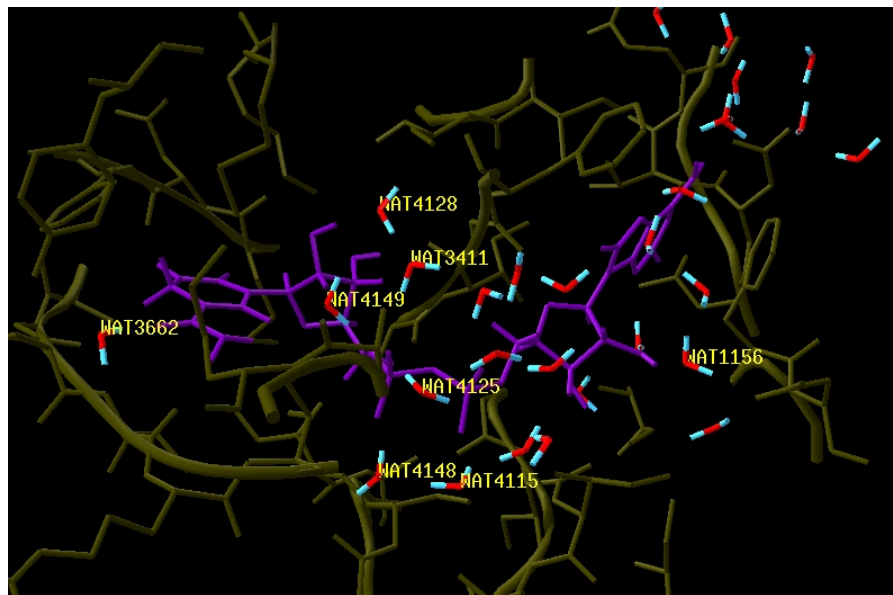


Figura 4.12. Moléculas de água conservadas durante a trajetória do sistema *wt* InhA-NADH no sítio de ligação do NADH. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A porção pirofosfato é responsável pelo maior número de interações com moléculas de água, formando pontes de hidrogênio estáveis com as moléculas de água *wat4125*, *wat4115* e *wat3411* criando uma malha de pontes de H nesta região.

Nas regiões da ribose A e anel adenina, mais expostas ao solvente, as moléculas de água que fazem ponte de hidrogênio com as hidroxilas da ribose A e com os nitrogênios do anel A são facilmente substituídas por outras moléculas de solvente, entretanto a água *wat1156* interage longamente com a hidroxila O'A2 da ribose A durante a fase de produção, mediando a interação da cadeia lateral do resíduo Asp42 com esta hidroxila. Nos últimos 200ps de simulação esta molécula de água é substituída por outra molécula de água do solvente (*wat2658*).

No anel A ocorre somente uma interação prolongada envolvendo a molécula de água *wat6639* e o nitrogênio N3A do anel, que também é observada no cristal da 1ENY. A *wat6639* também interage com o O'A2 da ribose A.

Durante a trajetória gerada na simulação do sistema *wt* InhA-NADH, o número de pontes de hidrogênio entre o NADH e a proteína mediadas por moléculas de água dentro do sítio de ligação (média de $13,0 \pm 1,9$) e de contatos

não ligantes (média de $28,2 \pm 5,9$) permaneceram estáveis, indicando a convergência e estabilidade destes parâmetros no sistema *wt* InhA-NADH.

A associação e orientação do cofator dentro do sítio catalítico da enzima também são importantes para a determinação do curso estereoquímico da reação (Fillgrove & Anderson, 2000). Durante a fase de produção, a média de $11,9 \pm 1,2$ pontes de hidrogênio entre o NADH e a enzima⁴¹, aliada às interações com moléculas de água bem ordenadas do sítio de ligação, garantem que a molécula do NADH permaneça estendida e dentro da cavidade da enzima (distância entre os centros de massa do NADH e da enzima na fase de produção de $6,3 \pm 0,2$ Å) na trajetória da *wt* InhA-NADH.

Apesar de enoil-tioésteres redutases utilizarem piridina-dinucleotídeos reduzido na catálise da redução de enoil-tioésteres, a diastereosseletividade da transferência de hidreto (H⁻) do C₄ do anel nicotinamida pode variar entre elas. Da mesma forma, a adição do próton (H⁺), provavelmente derivado do solvente, ocorre com diferentes especificidades.

Na InhA, a redução do substrato se dá pela transferência do hidrogênio *pro*-4S (face B) do anel nicotinamida ao carbono eletrofílico do substrato, e a adição dos dois átomos de hidrogênio ocorrem de forma *syn* (os dois H são adicionados na mesma face do tioéster insaturado (face *Re* da ligação C=O) (Fillgrove & Anderson, 2000). Isto implica que, antes da reação, o NADH deve estar ligado dentro do sítio ativo de forma a disponibilizar seu H_{*pro*-S} para a reação de redução (Quémard *et al.*, 1995).

O acompanhamento dos ângulos diedros do NADH (Apêndice 1) mostrou que o anel N mantém a conformação *syn* observada no cristal com o grupo 3-carboxiamida do mesmo lado da porção pirofosfato ($\chi_N = 67,3^0 \pm 11,3^0$), mantendo sempre sua face B direcionada para a cavidade hirofóbica de ligação do substrato, e, conseqüentemente, disponibilizando o H_{4S}. Esta conformação é mantida principalmente pela interação com três resíduos da cavidade de ligação: pelo OH da Thr196 que forma ponte de hidrogênio com grupo 3-carboxiamida, pela cadeia principal da Ile194, que também interage com a C=O do mesmo grupo, e pela cadeia lateral da Lys165, que “fixa” a hidroxila da ribose N (O’N3) via uma forte

⁴¹ O número de pontes de hidrogênio entre a enzima e o NADH foi calculado com o programa HBPLUS (McDonald & Thornton, 1994).

ponte de hidrogênio. Esta observação concorda com os dados cinéticos obtidos para mutantes na posição 165: a substituição da Lys165 da InhA por alanina ou metionina leva a enzimas que não se associam ao NADH, mas a mutação por glutamina ou arginina (que podem fazer pontes de H) apresentam efeitos mínimos sobre a atividade enzimática (Parikh *et al.*, 1999).

Durante a trajetória os diedros β_N , α_N , ζ_N , ζ_A , α_A e β_A sofrem variações em relação aos diedros da posição cristalográfica, sendo responsáveis pela flexibilidade observada para a molécula de NADH, mas flutuam em torno aos seus valores médios. O anel A também mantém a conformação *anti* ao longo da simulação ($\chi_A = -118,8^\circ \pm 13,5^\circ$), auxiliado pela interação com a Val65.

4.3 – Simulação por DM dos mutantes S94A, I21V, I16T e I47T da enzima InhA complexados com o NADH

As mutações *on the fly* para a construção dos mutantes S94A, I21V, I16T e I47T utilizou a estrutura instantânea do complexo *wt* InhA-NADH em 1,9 ns. Como todas essas mutações envolvem a substituição de um aminoácido maior por outro menor, a construção dos modelos foi realizada pela substituição dos respectivos aminoácidos nos arquivos de coordenadas do sistema da enzima espécie selvagem. Simulando uma mutação espontânea, as velocidades do sistema *wt* InhA-NADH foram editadas e reatribuídas aos sistemas mutantes, e as simulações foram continuadas sem a necessidade de nova termalização.

A substituição dos aminoácidos nos respectivos sistemas leva a uma desestabilização inicial na trajetória dos mutantes. Como pode ser observado nos gráficos dos RMSDs (Figura 4.13). Esta desestabilização é mais marcante para o sistema mutante I21V-NADH, que teve sua simulação estendida até 6,5 ns após a mutação, para a convergência do RMSD. Para efeito de comparação, somente os últimos 2 ns de simulação de cada um dos sistemas foram considerados nas análises.

A avaliação dos conjuntos de estruturas médias a cada 250 ps de simulação dos sistemas mutantes mostrou que, de modo geral, as estruturas secundárias, principalmente as fitas que formam as folhas β centrais, são conservadas durante as trajetórias. Em todos os quatro mutantes, a volta da hélice $\alpha 5$, onde encontra-se a Lys165, é temporariamente desfeita, principalmente pela variação dos diedros ϕ e φ da Ala167. Nos mutantes S94A, I16T e I21V, a organização da porção N-terminal da hélice $\alpha 7$, que faz parte da alça de ligação do substrato, é também desfeita ao longo da simulação, sendo mais marcante no mutante I21V, onde a organização da hélice $\alpha 6$ também é momentaneamente desfeita. Dentre os mutantes, o I47T apresenta a maior conservação das estruturas secundárias, com comportamento semelhante à enzima espécie selvagem.

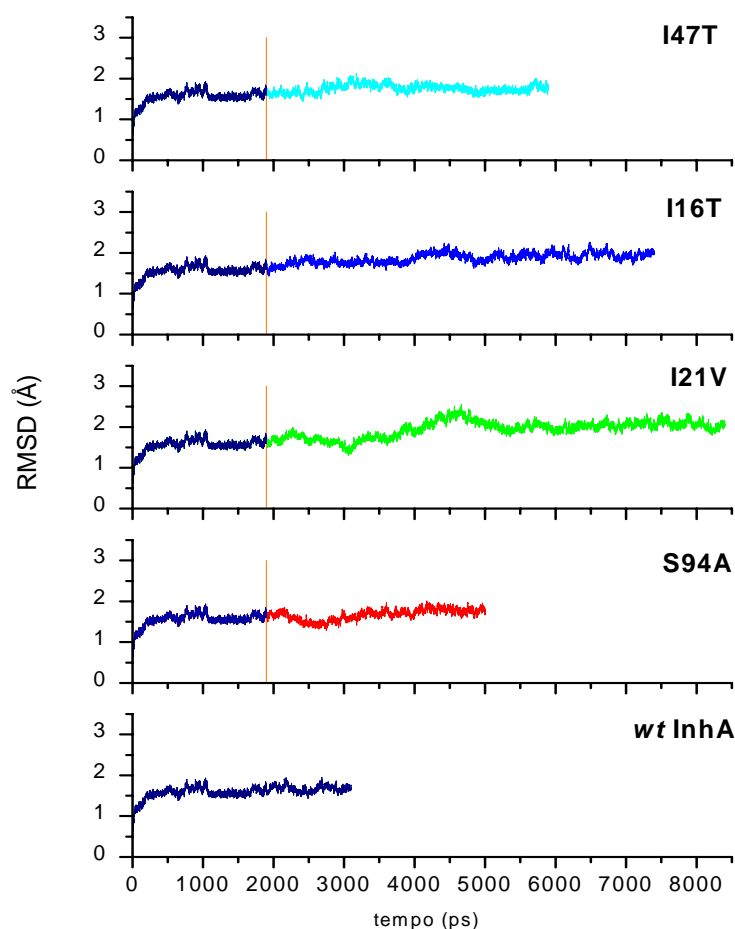


Figura 4.13. Gráficos do RMSD do esqueleto polipeptídico em função do tempo para as trajetórias dos sistemas mutantes S94A-NADH (vermelho), I21V-NADH (verde), I16T-NADH (azul) e I47T-NADH (ciano), com a identificação da mutação *on the fly* em 1,9 ns (barra vertical laranja), e comparação com o sistema *wt InhA*-NADH (azul escuro). Gráfico gerado com o programa Origin 6.1 (Microcal Software, Inc.).

A mutação de aminoácidos do sítio de ligação do NADH leva a enzimas mutantes com o mesmo perfil de *B-factor* (Figura 4.14) observado para a enzima espécie selvagem. Entretanto, as alças A e B nestes sistemas apresentam menor mobilidade. Os mutantes também possuem valores de *B-factor* mais altos para a região que corresponde à alça de ligação do substrato, mas a mutação I21V leva a uma enzima com maior mobilidade nesta alça, seguido da enzima I16T cujo valor de *B-factor* nesta região se compara ao da *wt InhA*.

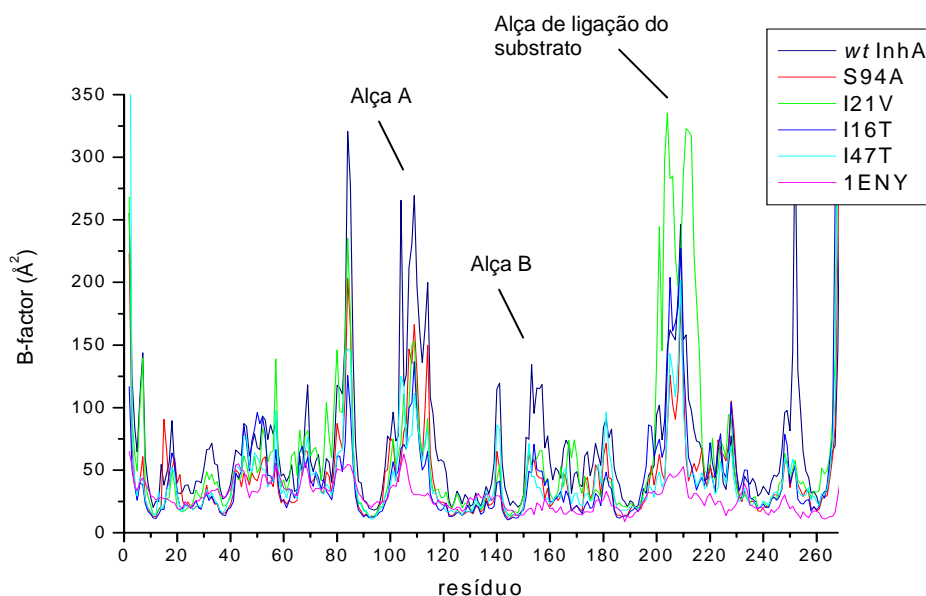


Figura 4.14. Fatores de temperatura (*B-factor*) da cadeia principal dos resíduos da enzima InhA nas simulações do sistema *wt* InhA-NADH (azul escuro) e mutantes S94A-NADH (vermelho), I21V-NADH (verde), I16T-NADH (azul) e I47T-NADH (ciano), calculados sobre os últimos 2 ns de simulação. *B-factor* da 1ENY em magenta. Gráfico gerado com o programa Origin 6.1 (Microcal Software, Inc.).

No entanto, a maior mobilidade da alça de ligação do substrato no mutante I21V e I16T não leva a uma maior abertura na entrada da cavidade de ligação do NADH (que se daria pelo afastamento em relação à alça A), mas sim na parte posterior da enzima, em relação à alça B (Figura 4.15).

O maior distanciamento entre a alça de ligação do substrato e a alça B em relação ao observado na *wt* InhA permite a entrada de mais moléculas de água no sítio hidrofóbico de ligação do anel N, o que é refletido no número de moléculas de água na primeira esfera de solvatação calculada para cada porção da molécula do NADH (Tabela 4.5).

Enquanto no sistema *wt* InhA-NADH a primeira esfera de solvatação do anel N é de $1,9 \pm 0,3$ moléculas de água, nos mutantes I21V e I16T ela passa para $3,3 \pm 1,1$ e $2,9 \pm 1,0$ moléculas de água, respectivamente.

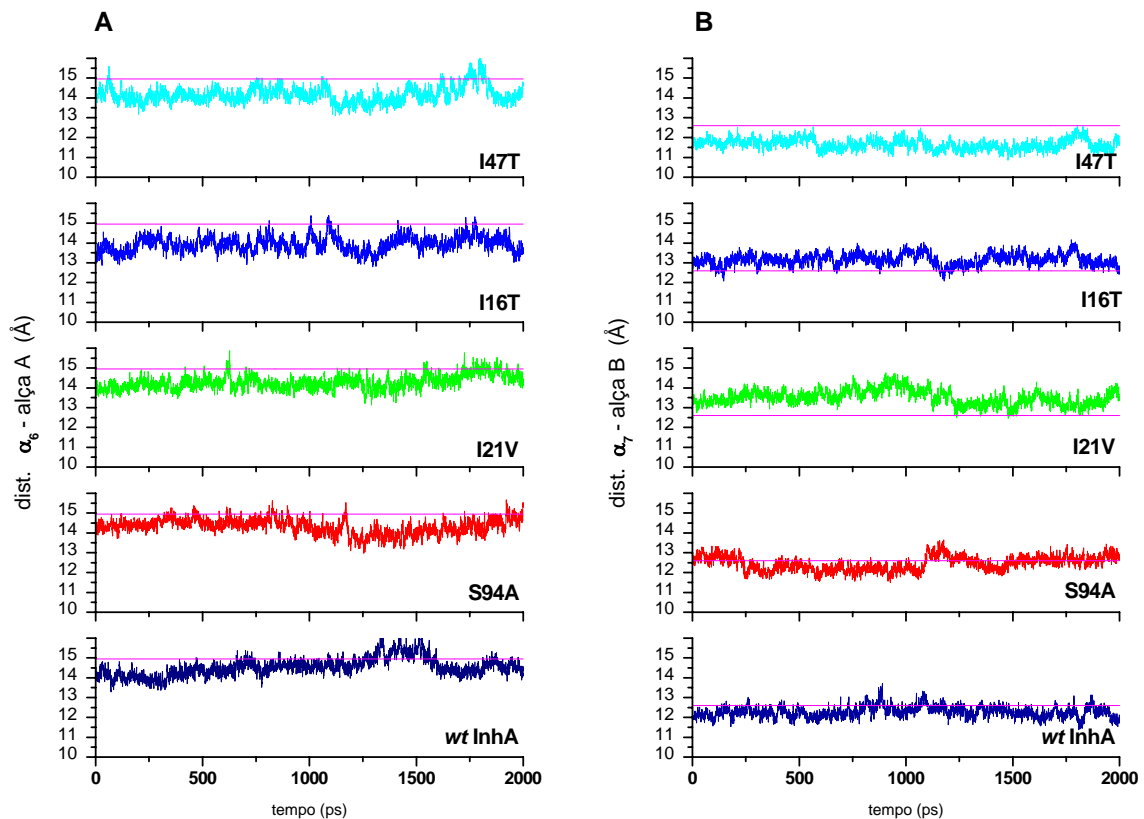


Figura 4.15. Comparação da abertura da alça de ligação do substrato (resíduos 196-219) em relação às alças A e B nas simulações do sistema *wt InhA-NADH* e mutantes *S94A-NADH* (vermelho), *I21V-NADH* (verde), *I16T-NADH* (azul) e *I47T-NADH* (ciano), calculados sobre os últimos 2 ns de simulação. (A) Distância entre o centro de massa da hélice α_6 da alça de ligação do substrato (resíduos 197-208) e da alça A (resíduos 100-111). (B) Distância entre o centro de massa da hélice α_7 da alça de ligação do substrato (resíduos 209-225) e da alça B (resíduos 149-160). Gráficos gerados com o programa Origin 6.1 (Microcal Software, Inc.).

Tabela 4.5. Número de moléculas de água na primeira esfera de solvatação (3,5 Å de distância) da molécula do NADH calculado sobre os últimos 2 ns de simulação dos sistemas *InhA-NADH*. Cálculo feito com o programa ptraj-6.4 do AMBER 6.0.

Porção do NADH	<i>wt InhA-NADH</i>	<i>S94A-NADH</i>	<i>I16T-NADH</i>	<i>I21V-NADH</i>	<i>I47T-NADH</i>
Anel N	1,9 ± 0,3	2,0 ± 0,5	2,9 ± 1,0	3,3 ± 1,1	1,8 ± 0,5
Ribose N	1,7 ± 0,6	1,8 ± 0,1	2,0 ± 0,7	1,5 ± 0,6	1,4 ± 0,6
Pirofosfato	8,1 ± 0,8	7,4 ± 1,7	11,9 ± 1,0	11,2 ± 1,0	7,5 ± 1,0
Ribose A	3,6 ± 0,9	3,7 ± 0,9	4,0 ± 1,2	4,0 ± 1,0	4,6 ± 1,1
Anel A	4,4 ± 1,0	4,4 ± 1,0	4,1 ± 1,1	4,8 ± 1,1	4,1 ± 1,2

A maior hidratação também é observada na porção pirofosfato da molécula de NADH nos sistemas mutantes, que aumenta de $8,1 \pm 0,8$ moléculas de água na *wt* InhA-NADH para $11,2 \pm 1,0$ e $11,9 \pm 1,0$ nos mutantes I21V e I16T, respectivamente.

As mudanças conformacionais decorrentes da mutação de aminoácidos da alça da glicina (característica de estruturas do tipo *Rossmann fold*), seja pela substituição de um aminoácido apolar por outro apolar de menor volume (no caso do I21V) ou de um aminoácido apolar por um levemente polar também de menor volume (no caso do I16T) levam a uma modificação no perfil de interação com a molécula de NADH, que é refletida em cada um dos sistemas em relação à sua interação com resíduos da proteína ou com moléculas de água ordenadas (Tabelas 4.6 e 4.7).

Tabela 4.6. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante I21V-InhA, com ocorrência maior do que 50% durante os últimos 2 ns de simulação. (A) Pontes de hidrogênio formadas diretamente entre o NADH e aminoácidos do sítio de ligação. (B) Pontes de hidrogênio entre o NADH e moléculas de água que medeiam a interação com o sítio de ligação.

(A)	Resíduo	Grupo do aminoácido	Átomo do NADH ^a	Ocorrência ^b (%)
	Ile95	C=O	O'N3	100
	Gly96	N	O'A5	74
		N	O'A4	59
	Lys165	NZ	O'N2	86
		NZ	O'N3	86
	Ile194	C=O	N7N	84
		N	O7N	100

(B)	Molécula de HOH	Átomo da HOH	Átomo do NADH ^a	Ocorrência (%) ^a
	wat1704	O-H1	O'A2	54
	wat3581	O	O'N2	98
	wat4069	O-H2	OPN2	58
	wat4115	O-H2	OPN2	71
		O-H1	OPA2	61
	wat5720	O	N6A	60
	wat7025	O	O'A3	65
	wat4189	O-H1	OPN1	52

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1.

^bOcorrência calculada com o programa Carnal do AMBER6.0

Na substituição da Ile21 por Val, dos dez resíduos que fazem pontes de hidrogênio com ocorrência superior a 50% na fase de produção no sistema espécie selvagem, somente quatro (Ile95, Gly96, Lys165 e Ile194) se conservam por mais da metade da trajetória deste mutante. Com exceção da Gly96, que interage com a metade adenina do NADH, todos os aminoácidos estão envolvidos em interações com o anel N e a ribose N (Figura 4.16).

Na mutação da Ile21 da alça rica em Gly, as interações diretas entre a proteína e a porção pirofosfato, feitas majoritariamente por aminoácidos desta alça (Gly14, Ile15, Ser20 e Ile21) são desfeitas, diminuindo a afinidade do mutante pelo NADH. Da mesma forma, as interações com aminoácidos do sítio de reconhecimento da adenina, formado pelo Asp64 e Val65 (que na espécie selvagem apresentam ocorrência de 99 e 100%, respectivamente), tornam-se menos importante, formando pontes de hidrogênio por apenas 21% e 46% do tempo da fase de produção, respectivamente.

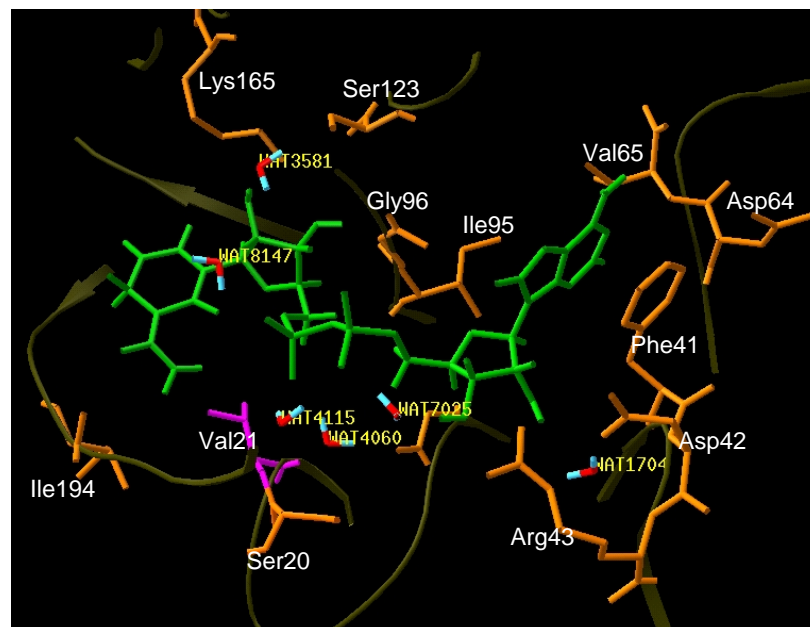


Figura 4.16. Sítio de ligação do NADH (verde) no mutante I21V InhA. Identificação do aminoácido mutado Val21 (magenta) e dos aminoácidos (laranja) e moléculas de água (vermelho e azul) importantes para a associação do cofator. Estrutura instantânea de menor RMSD em relação à estrutura média da fase de produção da simulação I21V-NADH. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A abertura em relação à alça B no fundo da cavidade de ligação do NADH permite uma troca maior de moléculas de água próximas ao anel N e, portanto, as moléculas de água que interagem com a porção 3-carboxiamida do anel N não apresentam grande tempo de residência. O mesmo ocorre em relação à ribose N, onde somente a wat3581 interage fortemente com o O'N2, intermediando a ligação indireta com a cadeia lateral da Ser123.

Outra molécula de água importante é a wat4115, que se liga ao OPN2 e OPA2, formando uma rede de pontes de hidrogênio com o NH da Leu197. Esta molécula de água apresenta equivalência com a wat4148 do sistema *wt* InhA-NADH. Entretanto, na espécie selvagem ela também se associa com a cadeia lateral da Ser20 da alça rica em glicina, aumentando a interação com o pirofosfato. No mutante I21V, esta interação com a Ser20 não ocorre, permitindo o distanciamento da porção pirofosfato do NADH de sua região de ligação.

A molécula de água wat4060 intermedeia parcialmente a interação do OPN2 com resíduos da alça da glicina, formando pontes de hidrogênio com o NADH em 58% do tempo, também se associando em parte do tempo à cadeia lateral da Ser20 e à C=O da Gly14.

A molécula de água wat1704 inicialmente está distante da enzima, e viaja durante a simulação se aproximando da região do NADH mais exposta ao solvente. Esta água permanece um pouco mais da metade do tempo da fase de produção participando de uma rede de pontes de hidrogênio com o O'A2 da ribose A, a cadeia lateral do Asp42, e os NH das cadeias principais da Phe41 e Arg43. E as demais moléculas de água com alto tempo de residência se associam ao NADH formando somente pontes de hidrogênio com outras moléculas de água do sítio da enzima.

Na substituição da Ile16 por Thr (mutante I16T), dos dez aminoácidos que ligam diretamente ao NADH via pontes de hidrogênio, somente sete são mantidos. Esta troca da Ile na posição 16 da alça rica em glicina leva à não associação direta dos aminoácidos Ser20 e Ile21 com a porção pirofosfato do NADH e à menor interação da cadeia principal da Ile15 com a ribose A (Tabela 4.7).

Tabela 4.7. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante I16T-InhA, com ocorrência maior do que 50% durante os últimos 2 ns de simulação. (A) Pontes de hidrogênio formadas diretamente entre o NADH e aminoácidos do sítio de ligação. (B) Pontes de hidrogênio entre o NADH e moléculas de água que medeiam a interação com o sítio de ligação.

(A)	Resíduo	Grupo do aminoácido	Átomo do NADH ^a	Ocorrência ^b (%)
	Ile15	C=O	O'A2	54
	Asp64	OD1	N6A	100
	Val65	N	N1A	92
	Gln66	OE1	N6A	64
	Ile95	C=O	O'N3	100
	Gly96	N	O'A4	100
	Tyr158	OH	O7N	98
	Lys165	NZ	O'N2	65
		NZ	O'N3	65
	Ile194	C=O	N7N	86
		N	O7N	100

(B)	Molécula de HOH	Átomo da HOH	Átomo do NADH ^a	Ocorrência (%) ^b
	wat6639	O-H2	N3A	55
	wat7916	O	O'N2	61

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1.

^bOcorrência calculada com o programa Carnal do AMBER6.0.

Embora a associação com os demais aminoácidos do sítio de ligação apresentem ocorrências semelhantes, há uma diminuição de 20% na ocorrência de associação da cadeia lateral da Lys165 com os OHs da ribose N.

A perda da interação da porção pirofosfato do NADH com a alça rica em glicina permite uma maior flexibilidade desta região e formação de novas interações via pontes de hidrogênio do NH₂ do anel A com a Gln66 e do O7N do anel N com a Tyr158. Entretanto, durante a fase de produção, somente duas moléculas de água não são rapidamente trocadas na trajetória: a wat7916, que se liga ao O'N2 da ribose N e ao OH da Ser123 e C=O da Phe97, e a wat6639 que interage com a N3A do anel A, com o O'A2 da ribose A e também faz ponte de hidrogênio com o NH da Phe41 (Figura 4.17).

A menor associação do ligante diretamente com os aminoácidos do sítio de ligação, ou mediados por pontes de hidrogênio com moléculas de água bem

ordenadas contribuem para a menor afinidade da enzima pelo NADH observada para os mutantes I21V e I16T (Tabela 1.2). De fato, a mutação de resíduos da alça rica em glicina já foi relacionada com a atenuação ou eliminação da atividade enzimática, e até mesmo com doenças, em proteínas com o domínio de ligação de dinucleotídeos (Rescigno & Perham, 1994; van Grunsven *et al.*, 1998; Eschenbrenner *et al.*, 2001) devido à modificação do padrão de interação com a região pirofosfato.

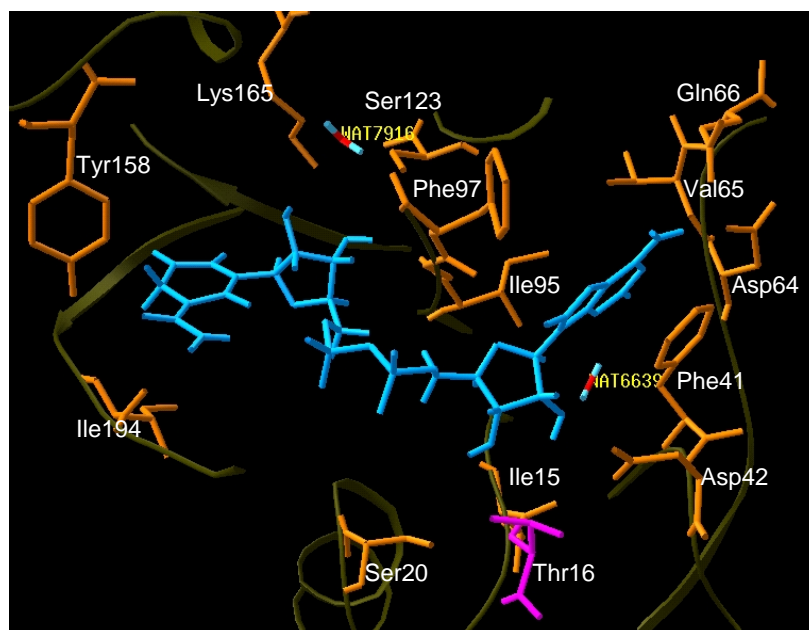


Figura 4.17. Sítio de ligação do NADH (ciano) no mutante I16T InhA. Identificação do aminoácido mutado Thr16 (magenta) e dos aminoácidos (laranja) e moléculas de água (vermelho e azul) importantes para a associação do cofator. Estrutura instantânea de menor RMSD em relação à estrutura média da fase de produção da simulação I16T-NADH. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A menor interação da porção pirofosfato do NADH com a alça rica em glicina é refletida nas mudanças conformacionais que ocorrem nesta porção do ligante nos mutantes I21V e I16T em relação à *wt* InhA, resultando no distanciamento desta região do NADH em relação à proteína (Figura 4.18).

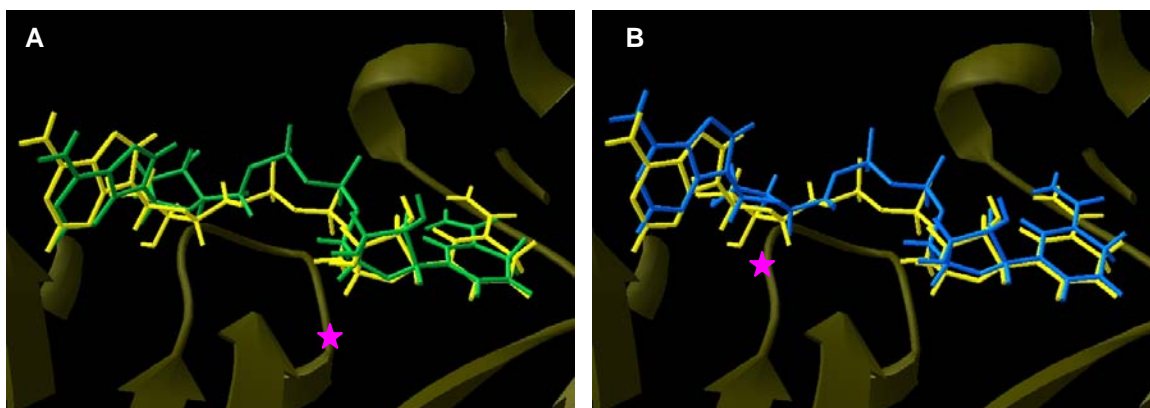


Figura 4.18. Comparação da conformação do NADH na estrutura média da fase de produção do sistema *wt* InhA-NADH (amarelo) com suas estruturas médias nos mutantes I21V-NADH (verde) e I16T-NADH (azul). As estrelas (magenta) indicam o local da mutação. Nesta figura o NADH é visto pelo lado oposto à alça rica em glicina, com o anel N para o lado direito. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

O monitoramento dos ângulos diedros do NADH ao longo das trajetórias mostrou que estas mudanças conformacionais são causadas pela modificação de diedros da porção pirofosfato, principalmente dos ângulos α_A e γ_A ⁴², como exemplificado para o mutante I21V na Figura 4.19. Enquanto na espécie selvagem estes ângulos flutuam próximos ao valor cristalográfico, nos mutantes I21V e I16T eles assumem valores bem diferentes.

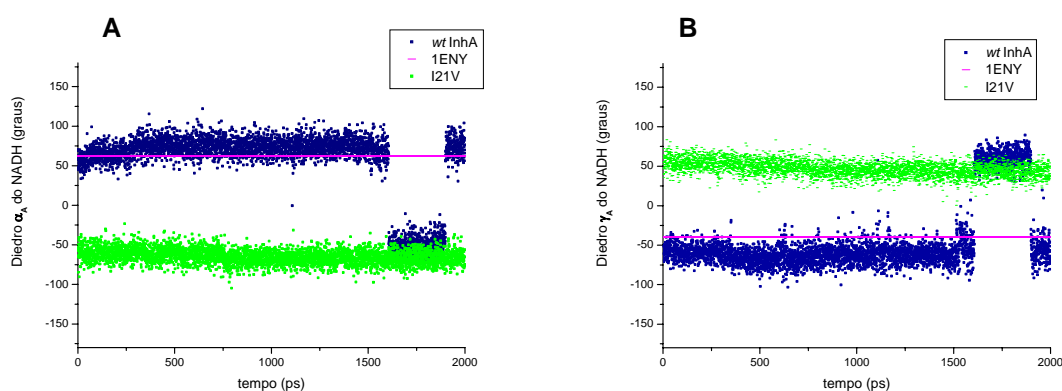


Figura 4.19. Comparação dos ângulos diedros α_A e γ_A do NADH na fase de produção das trajetórias do mutante I21V-NADH (verde), do sistema *wt* InhA-NADH (azul escuro) e da estrutura cristalina do NADH na 1ENY (magenta). (A) Diedro α_A (O3P-PA-O'A5-C'A5) do NADH. (B) Diedro γ_A (O'A5-C'A5-C'A4-C'A3) do NADH. Gráficos gerados com o programa Origin 6.1 (Microcal Software, Inc.).

⁴² Identificação dos ângulos diedros do NADH monitorados durante as simulações encontram-se no Apêndice 1.

É interessante notar que nas simulações, as torções α_A e γ_A da região pirofosfato do NADH estão intimamente relacionadas, mas β_A é conformacionalmente estável, e oscila em uma pequena faixa na região antiperiplanar (*trans*). Esta correlação entre α_A e γ_A , separados por β_A , é conhecido como *crankshaft motion*, onde a mudança em um dos diedros é compensada pelo outro diedro, semelhante ao observado na dinâmica da região fosfodiéster em DNAs (Swaminathan *et al.*, 1991).

A torção da região do pirofosfato do NADH nos mutantes I21V e I16T permite a entrada de um maior número de moléculas de água sobre esta região (Figura 4.20) em relação à enzima espécie selvagem (como anteriormente observado no cálculo das esferas de solvatação do NADH - Tabela 4.5) e a substituição das interações com aminoácidos do sítio de ligação.

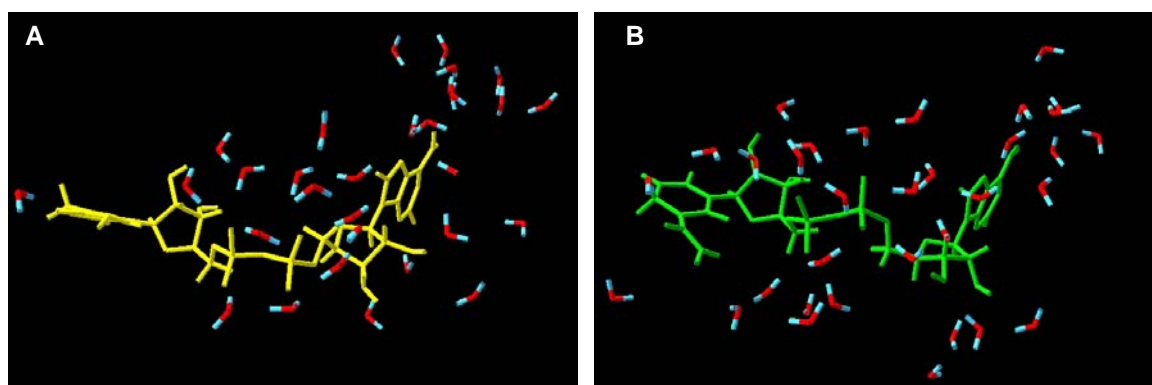


Figura 4.20. Comparação da distribuição das moléculas de água (vermelho e azul) a 5 Å do NADH nos sistemas *wt* InhA-NADH (amarelo em A) e mutante I21V-NADH (verde em B). Estruturas instantâneas com menor RMSD em relação às estruturas médias das fases de produção. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A exemplo do observado em estudos de simulação por DM de alguns sistemas ligante-receptor (Bakowies & van Gunsteren, 2002), este considerável movimento da porção pirofosfato do NADH no sítio de ligação poderia representar a fase inicial da expulsão do cofator da cavidade de ligação, causada pela menor afinidade pelo NADH.

A substituição de resíduos de aminoácidos da alça rica em glicina nos mutantes I21V e I16T visivelmente acrescenta também uma instabilidade no

sistema enzima-NADH, notada desde o início das simulações. Apesar das estruturas secundárias serem mantidas ao longo destas trajetórias, a instabilidade também se refletiu no longo tempo necessário para a convergência do RMSD destes sistemas, fazendo com que as simulações fossem prolongadas até 6,5 ns e 5,5 ns após as mutações, respectivamente.

Na mutação da Ile47 por treonina (mutante I47T) um aminoácido apolar é substituído por outro polar em uma região não tão próxima ao NADH quanto as mutações anteriores, entretanto esta substituição introduz algumas modificações discretas no padrão de interação da enzima com o ligante (Tabela 4.8).

Entre estas modificações, está a aproximação do NH da cadeia principal da Gly14 em relação à ribose A, formando uma ponte de hidrogênio com O'A3, com ocorrência de 75%. Na *wt* InhA, esta interação só ocorre durante 22% do tempo.

A aproximação da Gly14 causa uma movimentação da alça rica em glicina e a Ile21, que antes interagiu fortemente com o OPN2 do pirofosfato (ocorrência de 95% na enzima espécie selvagem), é afastada mantendo esta interação somente 26% do tempo da fase de produção (Figura 4.21). Este afastamento em relação à alça diminui a interação com o pifofosfato, permitindo uma torção em seus ângulos α_N e ζ_A que move o átomo de fósforo PN e seus oxigênios em direção ao centro da cavidade de ligação.

Na verdade, esta torção também é induzida por modificações na região entre a fita β_2 e a hélice α_2 . Com a substituição do aminoácido na posição 47, ocorre uma maior interação da cadeia lateral da Thr47 com a cadeia lateral da Arg43, que se direciona para o interior da cavidade de ligação da porção adenina, afastando o anel A do seu local original e a alça rica em glicina, que antes fazia contatos não ligantes com a Ile47 através de sua Ile16. A mudança conformacional na base da entrada da cavidade de ligação do cofator também aproxima o resíduo Phe41 do NADH em relação à espécie selvagem, e seu NH passa a interagir mais com o O'A2 da ribose A (ocorrência de 42% no mutante e 21% na *wt* InhA), fechando mais o sítio de ligação. Da mesma forma, o início da hélice α_6 (resíduos 194-198) da alça de ligação do substrato se aproxima do NADH. Entretanto esta contração parece ser bem localizada alterando somente a conformação do NADH e de cadeias laterais do sítio de ligação, não sendo notada em nível global na proteína,

que não apresenta diminuição significativa no raio de giro do esqueleto polipeptídico em relação à enzima espécie selvagem.

Tabela 4.8. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante I47T-InhA, com ocorrência maior do que 50% durante os últimos 2 ns de simulação. (A) Pontes de hidrogênio formadas diretamente entre o NADH e aminoácidos do sítio de ligação. (B) Pontes de hidrogênio entre o NADH e moléculas de água que medeiam a interação com o sítio de ligação.

(A)	Resíduo	Grupo do aminoácido	Átomo do NADH ^a	Ocorrência ^b (%)
	Gly14	N	O'A3	75
	Ile15	N	O'A3	91
	Ser20	OH	OPA2	100
	Asp64	COO-	N6A	67
	Val65	N	N1A	94
	Ile95	C=O	O'N3	100
	Gly96	N	O'A4	100
	Lys165	NZ	O'N2	96
		NZ	O'N3	100
	Ile194	C=O	N7N	89
		N	O7N	100
	Thr196	OG1	N7N	51
		OG1	OPN2	70

(B)	Molécula de HOH	Átomo da HOH	Átomo do NADH ^a	Ocorrência (%) ^b
	wat750	O	O'A2	58
	wat3334	O-H2	OPA1	66
		O-H1	OPA2	54
	wat3662	O-H2	O7N	51
	wat4128	O	O'2	53
	wat4148	O-H2	OPA2	62
	wat4149	O	N7N	56
		O-H2	OPN2	55
		O-H2	O'N5	55
	wat4189	O	O'A3	68
	wat6639	O-H2	N3A	59
	wat7159	O	O'A2	53

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1.

^bTempos de ocorrência calculados com o programa Carnal do AMBER6.0.

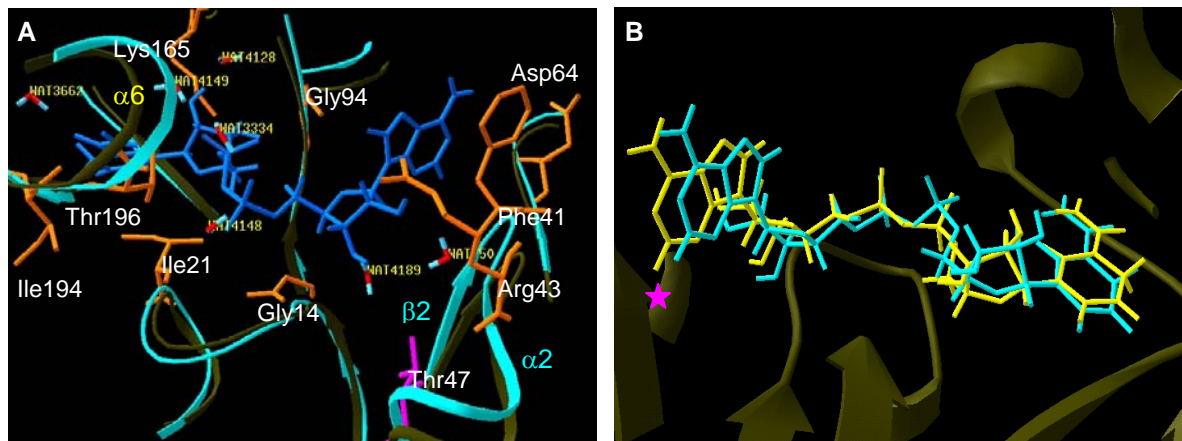


Figura 4.21. Sítio de ligação do NADH no mutante I47T InhA e comparação da conformação do NADH na estrutura média da fase de produção do sistema *wt* InhA-NADH e do mutante I47T-NADH. (A) Estrutura instantânea do sítio de ligação do NADH (azul) de menor RMSD em relação à estrutura média da fase de produção da simulação I21V-NADH. Identificação do aminoácido mutado Val21 (magenta) e dos aminoácidos (laranja) e moléculas de água (vermelho e azul) importantes para a associação do cofator. (B) Comparação da conformação do NADH nas estruturas médias da fase de produção do mutante I21V-NADH (ciano) e da *wt* InhA (amarelo). A estrela (magenta) indica o local da mutação. Nesta figura o NADH é visto pelo lado oposto à alça rica em glicina, com o anel N para o lado direito. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

No mutante I47T, um bom número de moléculas de água apresenta tempo de residência maior do que 50%, entretanto, nenhuma apresenta ocorrência maior do que 70%, e quatro ficam acima de 60%. A água *wat4189*, que apresenta tempo de residência de 68%, interage com o O'A3 por baixo da ribose A e fica “presa” nesta região pela contração do sítio de ligação, também formando pontes de hidrogênio com a C=O da Gly14 e da Thr39 da fita $\beta 2$. A *wat4148*, que se liga ao OPA2 do pirofosfato também mantém pontes de hidrogênio com o NH da Ile21 e a cadeia lateral da Thr17, ambos da alça rica em glicina, assim como com o NH da Leu197. Nesta mesma região a *wat3334*, que se associa a OPA1 e OPN2, forma uma rede de pontes de hidrogênio com o NH da Ala198, do início da alça de ligação do substrato (resíduos 194-198). Esta associação leva à aproximação desta região em relação ao NADH.

A molécula de água *wat4128*, que interage com o O'N2, faz a intermediação da interação do OH do Ser123 com a ribose N do ligante.

É interessante notar que o fechamento do sítio de ligação leva a expulsão de algumas moléculas de água que participam da interação entre a proteína e o ligante

(Figura 4.22). Enquanto no início da fase de produção existe uma média de 27 moléculas de água a 5 Å do NADH (comparável ao número de moléculas de água ao longo da trajetória da *wt* InhA, de $27,6 \pm 2,4$), nos últimos 250 ps este número cai para 23. Este fato também não permite que a porção pirofosfato do NADH no I47T, que também se distancia da alça de ligação rica em glicina, sofra maior hidratação como observado para os mutantes I21V e I16T.

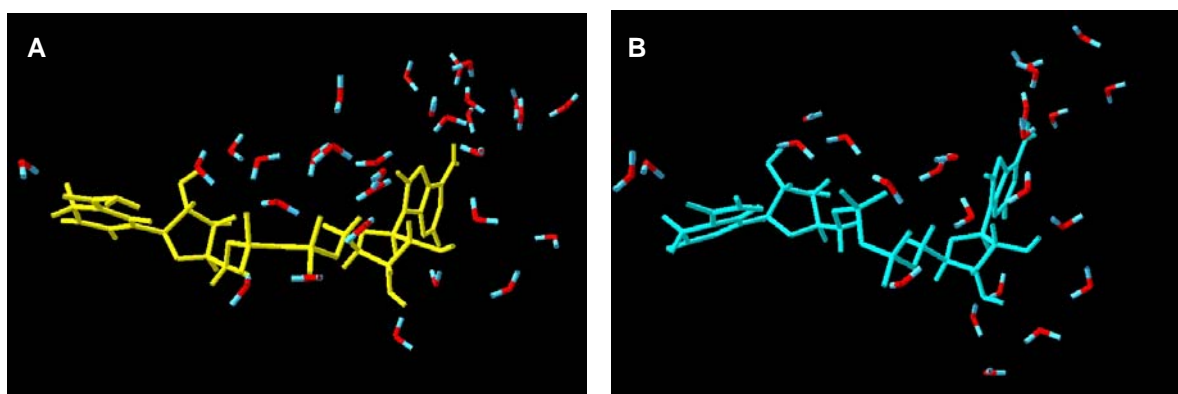


Figura 4.22. Comparação da conformação do NADH e distribuição das moléculas de água (vermelho e azul) a 5 Å do cofator nos sistemas *wt* InhA-NADH (amarelo em A) e mutante I47T-NADH (ciano em B). Estruturas instantâneas com menor RMSD em relação às estruturas médias das fases de produção. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A contração da molécula do NADH causada pelo fechamento do sítio de ligação no mutante I47T e torção de diedros da região do pirofosfato também é aparente no cálculo da distância entre C2N e C6A ao longo da trajetória (Figura 4.23). Enquanto na espécie selvagem o NADH apresenta uma distância média de $15,7 \pm 0,3$ Å entre C2N e C6A, no mutante I47T esta distância passa a $14,7 \pm 0,4$ Å.

A contração da molécula de NADH também foi observada no mutante I21V (distância C2N-C6A = $14,6 \pm 0,4$ Å), mas no mutante I16T ($15,3 \pm 0,3$ Å) esta contração ocorre em menor extensão.

O acompanhamento do número de pontes de hidrogênio formadas entre o NADH e aminoácidos do sítio ativo, e daquelas formadas com moléculas de água que medeiam a interação com as enzimas, permitem um melhor entendimento destas modificações que ocorrem nas interações do NADH com seus sítios de ligação nos diferentes mutantes (Figura 4.24).

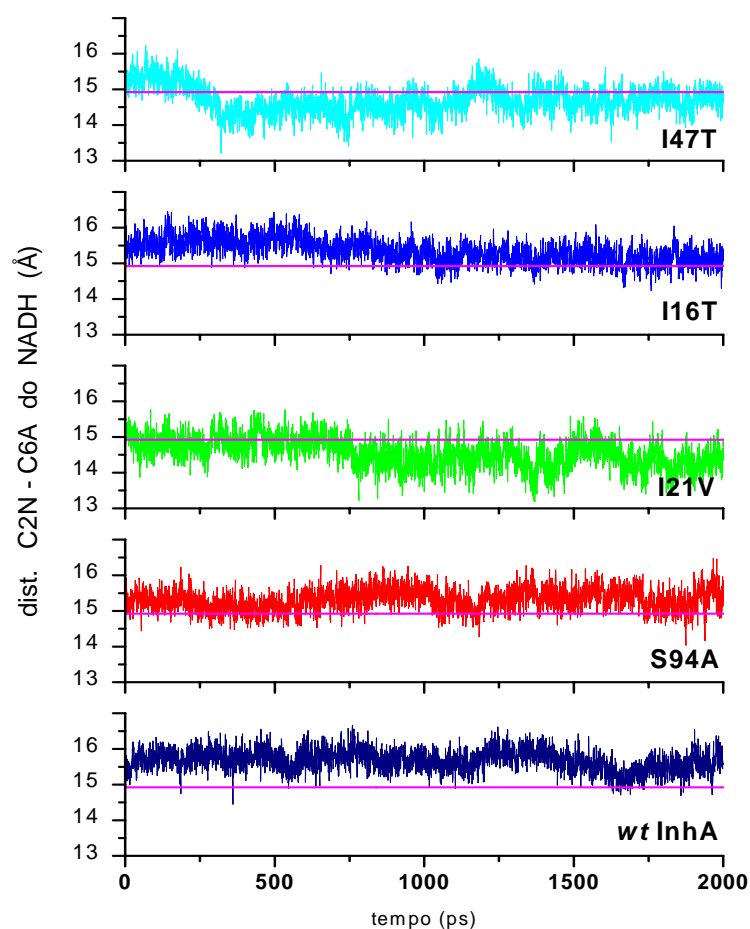


Figura 4.23. Tamanho do NADH nas simulações do sistema *wt* InhA-NADH e mutantes S94A-NADH (vermelho), I21V-NADH (verde), I16T-NADH (azul) e I47T-NADH (ciano), calculado como a distância entre C2N-C6A sobre os últimos 2 ns de simulação. Cálculo realizado com o programa ptraj-6.4 do AMBER 6.0.

Nos últimos 2 ns de simulação do sistema *wt* InhA-NADH as interações via pontes de hidrogênio diretamente com aminoácidos do sítio de ligação e aquelas mediadas por moléculas de água flutuam em torno de valores médios. Nos mutantes, estes valores nem sempre convergiram, mais uma vez atestando a instabilidade dos sistemas pela introdução de mutações. Para comparação, as médias foram então calculadas apenas para os últimos 250 ps de simulação de cada sistema.

Nos mutantes I21V e I16T, onde o NADH experimenta maiores mudanças conformacionais, sofrendo uma maior solvatação, ocorre uma grande diminuição do número de pontes de hidrogênio formadas entre o NADH e moléculas de água que também interagem com a proteína. Enquanto na espécie selvagem são realizadas

em média de $12,0 \pm 1,6$ pontes de hidrogênio com estas águas, nos mutantes I21V e I16T este valor passa a ser $5,5 \pm 0,7$ e $5,3 \pm 1,1$, respectivamente. O número de pontes de hidrogênio formadas com aminoácidos do sítio de ligação também diminui em relação ao sistema espécie selvagem (*wt* InhA = $11,6 \pm 1,3$, I21V = $5,4 \pm 1,0$ e I16T = $10,5 \pm 1,1$ pontes de hidrogênio), e o mutante I21V passa a fazer, inclusive, menos interações deste tipo do que o observado na estrutura cristalina 1ENY (que faz 8 pontes de H).

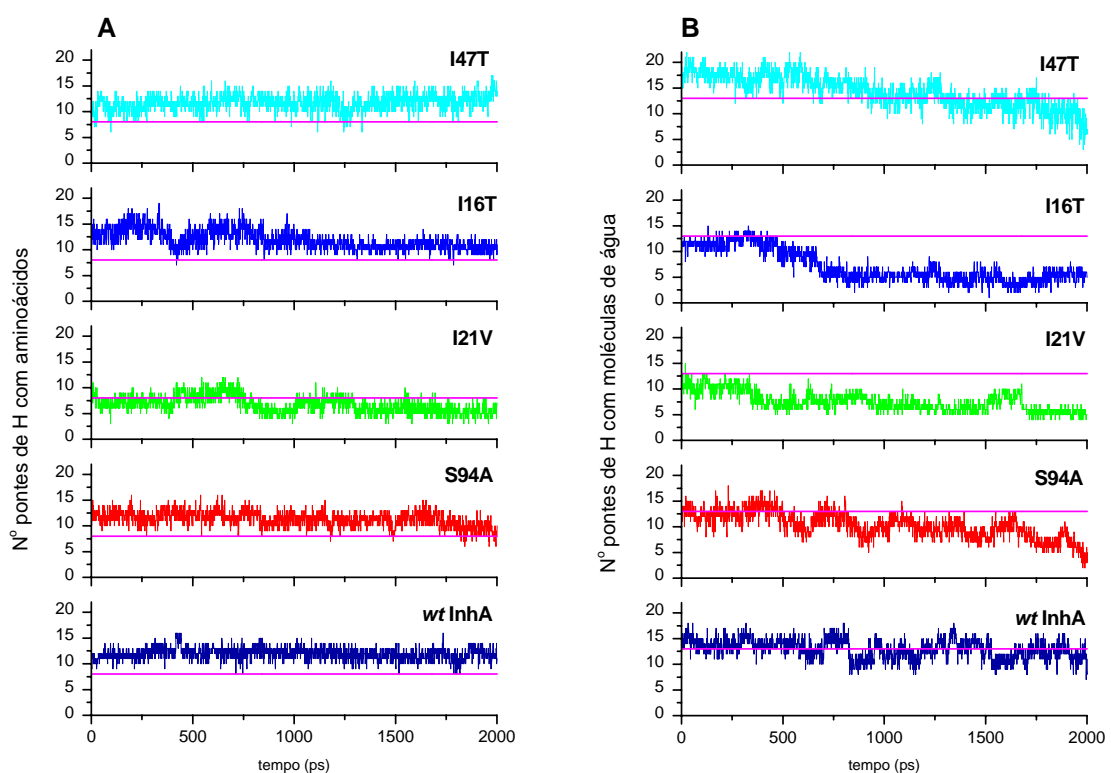


Figura 4.24. Comparação do número de pontes de hidrogênio formadas entre o NADH e seu sítio de ligação dos sistemas *wt* InhA-NADH (azul escuro) e mutantes S94A-NADH (vermelho), I21V-NADH (verde), I16T-NADH (azul) e I47T-NADH (ciano). (A) Número de pontes de H diretas entre o NADH e aminoácidos do sítio de ligação. Referência: estrutura cristalográfica 1ENY (magenta) – 8 pontes de H. (B) Número de pontes de H com moléculas de água que medeiam a interação com a proteína. Referência: número de pontes de hidrogênio médio no sistema *wt* InhA-NADH (magenta) – 12 pontes de H. Cálculos realizados com o programa HBPLUS (McDonald & Thornton, 1994) no LIGPLOT (Wallace *et al.*, 1995) sobre os últimos 2 ns das simulações.

No mutante I47T, pela contração do sítio de ligação do NADH, o número de interações via pontes de hidrogênio realizadas diretamente com aminoácidos aumenta levemente em relação à espécie selvagem (I47T = $12,6 \pm 1,2$ e *wt* InhA =

11,6 ± 1,3 pontes de H), mas a expulsão de moléculas de água do sítio e reorientação da molécula de NADH levou à diminuição progressiva do número de pontes de hidrogênio com moléculas de água que medeiam interações com a proteína (valor médio para os últimos 250 ps no I47T = 10,0 ± 2,1 pontes de H), mostrando a tendência deste mutante em perder estas interações.

O mutante S94A foi o sistema cujo comportamento dinâmico mais se assemelhou ao sistema espécie selvagem, e sua estrutura média na fase de produção apresentou o menor RMSD em relação à estrutura média da enzima espécie selvagem (RMSD = 1,0 Å)⁴². A interação do NADH com este mutante segue o mesmo perfil de interação da *wt* InhA (Tabela 4.9)

Tabela 4.9. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do mutante S94A-InhA, com ocorrência maior do que 50% durante os últimos 2 ns de simulação. (A) Pontes de hidrogênio formadas diretamente entre o NADH e aminoácidos do sítio de ligação. (B) Pontes de hidrogênio entre o NADH e moléculas de água que medeiam a interação com o sítio de ligação.

(A)	Resíduo	Grupo do aminoácido	Átomo do NADH ^a	Ocorrência ^b (%)
	Gly14	N	O'A3	51
	Ile15	N	O'A3	59
	Ser20	OH	OPA2	97
	Asp64	COO-	N6A	92
	Val65	N	N1A	100
	Ile95	C=O	O'N3	100
	Gly96	N	O'A4	96
	Lys165	NZ	O'N2	75
		NZ	O'N3	100
	Ile194	C=O	N7N	98
		N	O7N	100
	Thr196	OG1	N7N	50
		OG1	OPN1	64

(B)	Molécula de HOH	Átomo da HOH	Átomo do NADH ^a	Ocorrência (%) ^a
	wat3662	O-H2	O7N	59
	wat4189	O	O'A3	50
	wat4959	O	O'A3	55
	wat6639	O-H2	N3A	53

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1.

^bTempos de ocorrência calculados com o programa Carnal do AMBER6.0.

⁴² RMSD entre as estruturas médias dos últimos 2 ns de simulação, calculado com o programa SPDBV v3.7b2 em relação ao esqueleto polipeptídico: S94A = 1,0 Å, I21V = 1,6 Å, I16T = 1,4 Å e I47T = 1,1 Å

Os aminoácidos que interagem com ambos sistemas são os mesmos, e uma pequena diferença ocorre somente nos tempos de ocorrência de algumas pontes de hidrogênio, apesar da maioria das variações não passarem de 10%.

Na substituição da Ser94 pelo aminoácido apolar alanina, a cadeia lateral deste resíduo não pode mais formar pontes de hidrogênio diretamente com o NADH (ocorrência de 32% na *wt* InhA) e com a C=O da Ser13 do início da alça rica em glicina, que também se ligava ao O'A5 do NADH. Nesta troca de aminoácidos, a C=O da Ser13 é liberada pela quebra da ligação com a Ser94, permitindo o rearranjo da conformação da alça rica em glicina. Essa alteração faz com que a Gly14, que anteriormente interagía com o O'A3 da ribose A com 22% de ocorrência, passe a formar uma ponte de hidrogênio mais efetiva (ocorrência de 51%) (Figura 4.25).

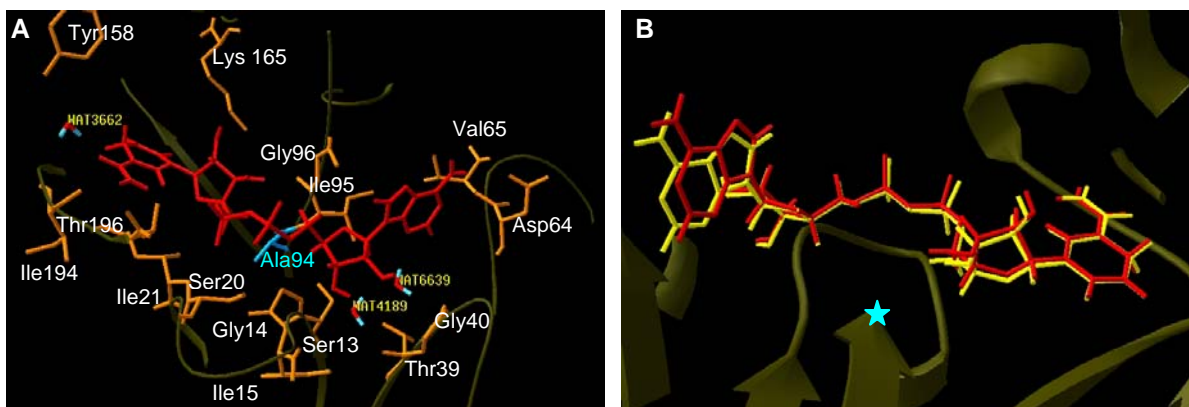


Figura 4.25. Sítio de ligação do NADH no mutante S94A InhA e comparação da conformação do NADH na estrutura média da fase de produção do sistema *wt* InhA-NADH e do mutante S94A-NADH. (A) Estrutura instantânea do sítio de ligação do NADH (vermelho) de menor RMSD em relação à estrutura média da fase de produção da simulação S94A-NADH. Identificação do aminoácido mutado Ala94 (ciano), dos aminoácidos (laranja) e moléculas de água (vermelho e azul) importantes para a associação do cofator. (B) Comparação da conformação do NADH nas estruturas médias da fase de produção do mutante S94A-NADH (vermelho) e da *wt* InhA (amarelo). A estrela (ciano) indica o local da mutação. Nesta figura o NADH é visto pelo lado oposto à alça rica em glicina, com o anel N para o lado direito. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Outras diferenças são notadas em relação a resíduos desta alça: a Ile15 diminui sua interação em 14% com o O'A3, e a Ile21 em 13% com o OPN2. A

interação da Thr196 com o N7N do anel N via ponte de hidrogênio também diminui sua ocorrência em quase 20%.

No sítio de ligação do NADH, somente quatro moléculas de água apresentam pontes de hidrogênio com o cofator com ocorrência maior do que 50%. As moléculas wat3662, wat4189 e wat6639 permanecem sempre na mesma região durante a simulação, enquanto a wat4959, que interage com o O'A2 é trocada por outras moléculas de água de fora do sítio da enzima. A wat3662, assim como sua equivalente nos sistema *wt* InhA, é responsável pela intermediação da ligação indireta da Tyr158 com o O7N da 3-carboxiamida do anel N no fundo da cavidade de ligação. A wat4189, forma uma rede de pontes de hidrogênio com o O'A3, a C=O da Thr39, o NH da Gly14 e com a wat6639. Esta última intermedeia a ligação indireta da Gly40 com o N3A do anel A.

As esferas de solvatação do NADH no mutante S94A são estatisticamente idênticas em relação à espécie selvagem (Tabela 4.5). Pouca diferença também é notada em relação aos ângulos diedros do NADH ao longo da trajetória, onde somente por um curto período de 200 ps é observado uma modificação dos diedros α_N e ζ_A do pirofosfato. Esta pouca diferença resulta em uma distância C2N-C6A do NADH ($15,3 \pm 0,3$ Å) também muito próxima à do sistema selvagem ($15,7 \pm 0,3$ Å).

Dos parâmetros acompanhados durante a trajetória do mutante S94A, a maior diferença é notada em relação ao número de pontes de hidrogênio formadas entre o NADH e moléculas de água que medeiam interações com o sítio de ligação. Enquanto nos últimos 2 ns de simulação a espécie selvagem forma uma média de $13,0 \pm 1,9$ pontes de hidrogênio com estas moléculas de água, estas interações tendem a diminuir no final da trajetória. O número médio de pontes de hidrogênio entre o NADH e estas moléculas de água para os últimos 250 ps das simulações mostram que, enquanto na *wt* InhA são feitas $12,3 \pm 1,6$ interações, no mutante S94A este número cai para $6,6 \pm 1,6$.

Uma pequena diminuição do número de pontes de hidrogênio diretamente com aminoácidos do sítio de ligação também é observada nos últimos 250 ps de simulação (*wt* InhA = $11,6 \pm 1,3$ e S94A = $9,7 \pm 1,4$ pontes de H com os aminoácidos).

Na escala de tempo em que a simulação do sistema S94A-NADH foi realizada (3,1 ns), as mudanças conformacionais experimentadas pela

macromolécula e pelo ligante foram muito sutis, como previsto pela quase perfeita sobreposição das estruturas cristalográficas 1ENY e 1ENZ (Dessen *et al.*, 1995). A discreta diferença no padrão e no tempo de ocorrência de pontes de hidrogênio entre o NADH e aminoácidos e entre o NADH e moléculas de água do sítio de ligação foi atribuída como responsável pela diferença de afinidade do mutante S94A pelo NADH observada experimentalmente (Dessen *et al.*, 1995; Basso *et al.*, 1998).

A comparação das trajetórias das simulações por DM dos sistemas formados pelos mutantes da enzima InhA ligados ao cofator NADH com o sistema espécie selvagem mostrou que em todos os mutantes existe uma tendência de diminuição do número de interações do tipo pontes de hidrogênio entre o cofator e moléculas de água que medeiam a ligação com o sítio da enzima.

Em enzimas do tipo *Rossmann fold*, as moléculas de água do sítio de ligação contribuem significativamente para a interface de ligação do dinucleotídeo (Bottoms *et al.*, 2002), e a perda deste tipo de interação pode levar à menor afinidade da enzima pelo cofator, como observado em ensaios enzimáticos nos mutantes da InhA estudados (Basso *et al.*, 1998; Parikh *et al.*, 2000 – Tabela 1.2). As mutações observadas em isolados clínicos resistentes à INH causam modificações discretas na interação destes mutantes com o cofator que parecem afetar a ligação do NADH de diferentes formas. Nos mutantes I21V e I16T, as mutações na alça rica em glicina causam uma menor interação da enzima com a porção pirofosfato do NADH, que nestes mutantes parece iniciar um movimento de dissociação. No mutante I47T, as mudanças no padrão de interação dos aminoácidos próximos à hélice α_2 , causada pela substituição de um aminoácido apolar por um polar, leva a uma contração do sítio de ligação do NADH e expulsão gradativa de algumas moléculas de água importantes para o reconhecimento do cofator. Tal contração também pode ser responsável pelo maior K_m observado para o NADH em experimentos de cinética enzimática ($K_m \text{ NADH}_{I47T} = 243 \pm 25 \mu\text{M}$, e $K_m \text{ NADH}_{wtInhA} = 56 \pm 4 \mu\text{M}$) (Basso *et al.*, 1998), pois a associação do cofator pode ser dificultada pelo impedimento estérico causado por cadeias laterais do sítio de ligação, principalmente pelo redirecionamento da cadeia lateral da Arg43 na entrada do sítio de ligação.

No mutante S94A, as diferenças de interação são mais sutis, mas igualmente levam à diminuição do número de interações entre o NADH e moléculas de água de seu sítio de ligação.

Recentemente, um estudo *in vitro* sobre a inibição dos mutantes da InhA pelo bi-substrato INH-NAD mostrou que as mutações I21V, I47T e S94A, relacionadas com a resistência à INH, causam apenas um pequeno impacto nas constantes de inibição enzimática quando o complexo inibidor é previamente formado em solução. Estes mutantes são inativados apenas 1,2-1,5 vezes mais lentamente do que a enzima espécie selvagem (Rawat *et al.*, 2003). De fato, a presença do anel aromático da INH deve aumentar a afinidade da enzima pelo aduto inibidor pela formação de interações hidrofóbicas com o sítio de ligação do substrato, compensando parcialmente as diferenças na ligação da porção NAD(H) do bi-substrato.

O local de formação do bi-substrato *in vivo* ainda permanece controverso. Entretanto, dada a semelhança de inibição quando o bi-substrato é formado antes da associação com a enzima, e aos dados de cinética enzimática que apontam uma diferença de associação do NADH nos mutantes em relação à enzima espécie selvagem, também observados na análise qualitativa das trajetórias das simulações por DM dos complexos enzimáticos com o cofator, é mais provável que a formação do bi-substrato inibidor se dê dentro do sítio da enzima. Desta forma, a menor afinidade dos mutantes pela molécula de NADH pode facilitar a substituição do complexo inibidor por outra molécula de NADH disponível na biofase.

4.4 – Análise da importância da consideração das águas cristalográficas na simulação do sistema *wt* InhA-NADH

Na escala de tempo em que as simulações foram realizadas, observou-se que moléculas de água importantes para a formação de pontes de hidrogênio abaixo do NADH no sítio ativo da enzima não tiveram suas posições cristalográficas reproduzidas. Enquanto a simulação por DM da *wt* InhA-NADH permitiu reproduzir bem a posição de moléculas de água bem orientadas acima do NADH, como a wat280 e a wat302, e aquelas moléculas abaixo do anel adenina (wat277, wat297 e wat298), as moléculas de água localizadas abaixo do plano do cofator, e importantes para a interação deste com a proteína, não foram reproduzidas (Figura 4.26).

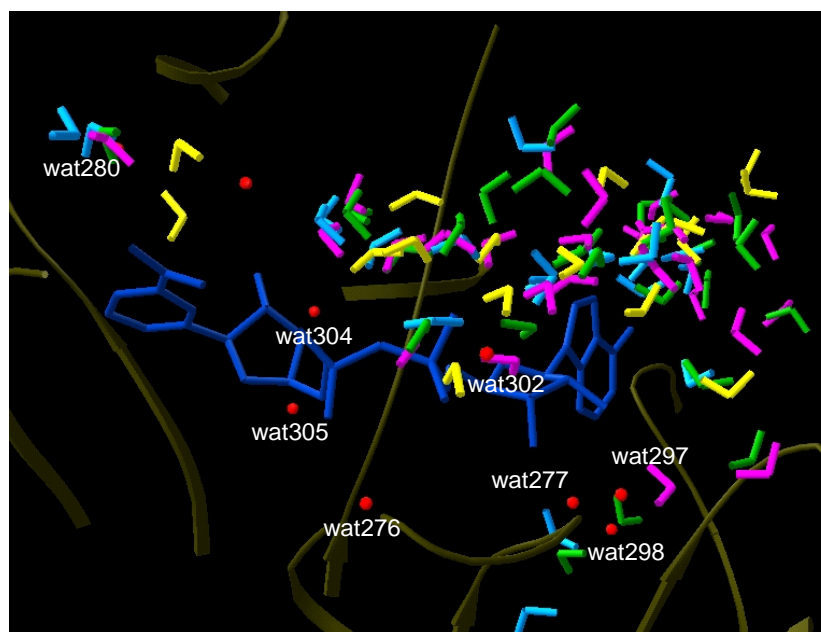


Figura 4.26. Moléculas de água cristalográficas (wat – representadas por seus átomos de oxigênio) não reproduzidas durante a simulação por DM do sistema *wt* InhA-NADH. Sobreposição das estruturas instantâneas da simulação *wt* InhA-NADH em 0,05 ns (amarelo), 1 ns (magenta), 2 ns (verde) e 3 ns (azul). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

Destas moléculas de água, a reprodução de uma molécula de água no local da água cristalográfica wat276 foi especialmente acompanhada. Esta é uma

molécula de água conservada, típica de estruturas do tipo *Rossmann fold* (Bottoms *et al.*, 2002), responsável pela formação de pontes de hidrogênio entre a alça da glicina e a porção pirofosfato do dinucleotídeo, que contribui significativamente para o reconhecimento na molécula de NADH. Apesar do pirofosfato não ser assimétrico, seus átomos de oxigênio são estereoquimicamente diferentes, e é quase sempre o OPN2 que interage com esta água conservada.

Na estrutura cristalina da InhA, esta água também faz pontes de hidrogênio com a C=O da Gly14 e com a cadeia lateral da Ser94, formando uma rede de pontes de hidrogênio. Foi sugerido que a menor afinidade do mutante S94A deva-se à substituição da cadeia lateral polar da serina pela apolar da alanina, incapaz de fazer pontes de hidrogênio com esta molécula de água. Assim, haveria a quebra desta rede de interação, levando à reorientação da cadeia da Gly14 (Dessen *et al.*, 1995; Rozwarski, *et al.*, 1999).

Para o estudo da importância desta molécula de água na ligação do NADH, um novo sistema InhA-NADH considerando as águas cristalográficas foi então construído e simulado durante 4 ns nas mesmas condições anteriores (sistema cWAT-NADH: 31.307 átomos, 41 moléculas de água cristalográfica, fazendo um total de 9.074 moléculas do tipo TIP3P).

A estrutura da proteína chegou a um patamar logo após os primeiros 0,5 ns de simulação, como pode ser visto no gráfico do RMSD (Figura 4.27), apresentando valores de RMSD para o esqueleto polipeptídico de aproximadamente 2,0 Å (média de $2,0 \pm 0,2$ Å para os últimos 2 ns de simulação). Entretanto, após 1,5 ns de simulação o RMSD aumentou lentamente, mas sistematicamente, e após 3,5 ns subiu para a faixa de 2,3 Å. Mudanças conformacionais associadas à movimentação da alça de ligação do substrato em relação às alças A e B são responsáveis pelo maior desvio em relação à estrutura cristalográfica. Após 3,5 ns ocorre uma aproximação da referida alça à alça A e afastamento em relação à alça B. Apesar destes movimentos, a avaliação dos conjuntos de estruturas médias geradas a cada 250 ps de simulação mostrou que a estrutura secundária é bem conservada ao longo da trajetória o que também é confirmado pelo acompanhamento do número total de pontes de hidrogênio, calculados de acordo com o método de Kabsch & Sander (1983) (NTPH = 176 ± 4). De modo geral, a flexibilidade do sistema é semelhante à observada para o sistema

sem a consideração das águas cristalográficas, apresentando fatores de temperatura de mesma amplitude.

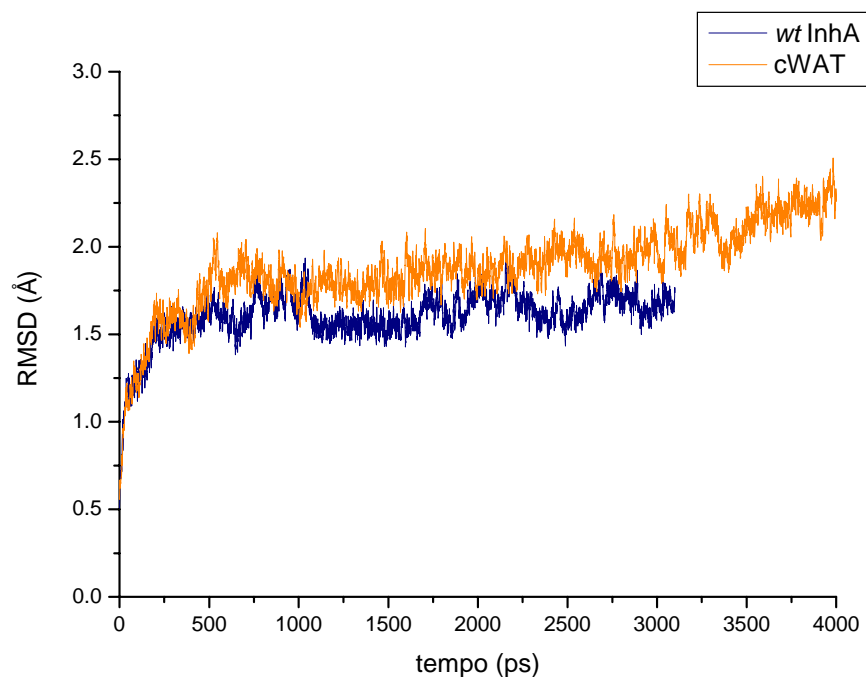


Figura 4.27. Gráfico do RMSD do esqueleto polipeptídico em função do tempo para as trajetórias do sistema cWAT-NADH (laranja) em relação à estrutura cristalográfica, e comparação com o sistema *wt* InhA-NADH (azul escuro).

Em comparação com a simulação *wt* InhA-NADH, a consideração das águas cristalográficas levou a um sistema mais compactado, com raio de giro para o esqueleto polipeptídico mais próximo ao observado para o cristal e para a apoenzima ($R_g = 17,9 \pm 0,1 \text{ \AA}$) (Figura 4.28A).

A presença de moléculas de água abaixo do NADH no sítio de ligação levou a uma associação mais efetiva do cofator no sítio da enzima, que também foi refletido em uma menor distância entre os centros de massa da proteína e ligante desde o início da fase de produção (Figura 4.28B).

Durante a simulação, o NADH permanece na sua forma estendida (distância C2N-C6A = $15,5 \pm 0,3 \text{ \AA}$) e com uma grande estabilidade em seus ângulos de torção, que flutuam em torno da média, sem apresentar desvios significativos.

No sistema cWAT-InhA, o NADH mantém um número estável de aminoácidos em seu sítio de ligação ($17,5 \pm 1,7$ aminoácidos a 4\AA do NADH), fazendo uma média de $11,9 \pm 1,1$ pontes de hidrogênio diretamente com os aminoácidos, e $11,1 \pm 1,8$ pontes com moléculas de água que medeiam a interação com a proteína.

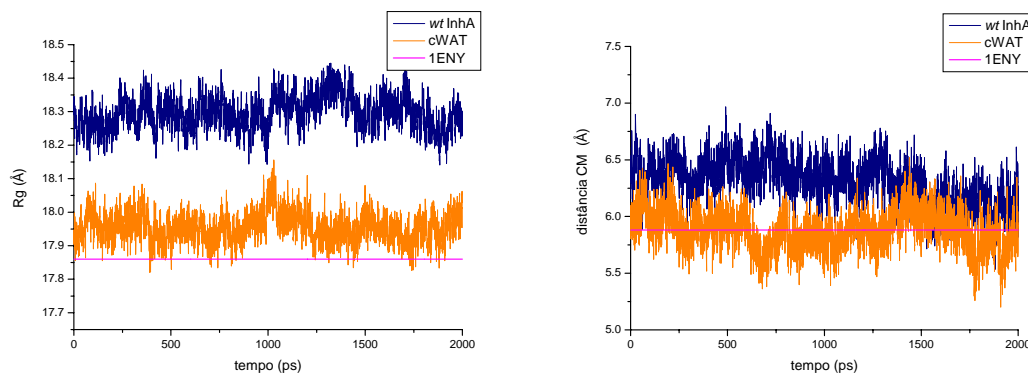


Figura 4.28. Gráficos dos raios de giro e distâncias entre centros de massa da enzima e do NADH nos sistemas *wt* InhA-NADH (azul escuro) e cWAT-NADH (laranja). (A) Raio de giro (Rg) do esqueleto polipeptídico em função do tempo para os últimos 2 ns das trajetórias. (B) Distância entre os centros de massa da enzima e do ligante em função do tempo para os últimos 2 ns das trajetórias. Referência: estrutura cristalográfica 1ENY (magenta).

As pontes de hidrogênio formadas entre o NADH e seu sítio de ligação na enzima foram monitoradas ao longo da trajetória da simulação da cWAT-NADH e aqueles resíduos que fazem pontes de hidrogênio com ocorrência de no mínimo 50% do tempo, e portanto mais importantes para a interação, estão colocados na Tabela 4.10.

O sítio de ligação do NADH no sistema contendo as águas cristalográficas apresenta algumas diferenças em relação ao sistema *wt* InhA-NADH. Naquele sistema, a Tyr158 que antes interagia com a porção 3-carboxiamida do NADH através de uma molécula de água agora se liga diretamente ao NADH formando uma forte ponte de hidrogênio (ocorrência de 99%). Esta aproximação da Tyr158 em relação ao NADH permite que a cadeia lateral da Met161 se direcione mais para o centro da cavidade de ligação, acima da porção da ribose N.

A Ile95, cuja C=O da cadeia principal se ligava diretamente à hidroxila da ribose N, sofre um rearranjo de sua cadeia polipeptídica para interagir com a molécula de água cristalográfica wat304. E a Gly14, que anteriormente só fazia uma interação fraca com o O'A3 (ocorrência de 22% na *wt* InhA-NADH), passa a interagir fortemente com aquela hidroxila (ocorrência de 99%), e com a água ordenada wat276.

Tabela 4.10. Pontes de hidrogênio formadas entre o NADH e o sítio de ligação do sistema cWAT-InhA, com ocorrência maior do que 50% durante os últimos 2 ns de simulação do sistema.

Resíduo	Grupo do aminoácido	Átomo do NADH ^a	Ocorrência ^b (%)
Gly14	C=O	O'A3	97
Ser20	OH	OPA2	100
Ile21	N	OPN2	100
Asp64	COO-	N6A	95
Val65	N	N1A	99
Gly96	N	O'A4	100
Tyr158	OH	O7N	99
Lys165	NZ	O'N2	71
Ile194	C=O	N7N	42
Ile194	N	O7N	100
Thr196	OG1	N7N	100
Thr196	OG1	OPN1	100

(B) Molécula de HOH	Átomo da HOH	Átomo do NADH ^a	Ocorrência (%) ^a
wat276	O-H2	OPN2	100
	O-H2	O3P	92
	O-H1	O'A5	63
wat298	O-H2	N3A	55
wat302	O-H1	OPA2	51
	O-H2	OPA2	52
wat304	O	O'N2	84
wat4105	O-H2	O'A3	58
	O-H1	O'A3	56
wat9292	O	O'N3	58

^aNome dos átomos da molécula de NADH de acordo com o Apêndice 1.

^bO cupância calculada com o programa Carnal do AMBER6.0

Como previsto anteriormente, as moléculas de água localizadas abaixo do plano da molécula de NADH ficaram restritas à suas regiões de ligação durante toda a fase de produção, correspondendo às moléculas de água com maior ocorrência. A wat276, que anteriormente não teve sua localização reproduzida, flutua em torno de sua posição média, formando uma rede de pontes de hidrogênio com a porção pirofosfato do NADH, resíduos da alça rica em glicina (Gly14 e Ala22) e cadeia lateral da Ser94. Outra água com grande ocorrência é a wat304, que interage com a ribose N, com as carbonilas da Ser94 e Ile95 e outras moléculas de água.

As águas wat302 e wat298 interagem com a metade adenina do NADH, mas também fazem ponte de hidrogênio com ocorrência maior do que 50%, mediando a interação desta região do NADH com a alça rica em glicina e com a hélice $\alpha 2$.

A wat4105 é uma molécula de solvente que fazia parte do sítio de ligação do NADH desde o início da simulação, se aproximando da ribose A e intermediando interação do O'A3 com o NH da Phe41 e C=O da Ile15. A wat9292 também é uma molécula de solvente, entretanto, no início da simulação ela está localizada fora do sítio de ligação do NADH, se aproximando somente nos últimos 2 ns, formando uma rede de pontes de hidrogênio com a região pirofosfato do cofator e outras moléculas de água.

O monitoramento do ângulo diedro da cadeia lateral do resíduo Ser94 (envolvido em uma das mutações) demonstrou que, também no sistema que considera as águas cristalográficas, o aminoácido Ser94 possui uma grande flexibilidade, e nem sempre está envolvido na formação da rede de pontes de hidrogênio entre o O'N3 da ribose N, a C=O da Gly14 e a molécula de água ordenada 276, como proposto pela análise da estrutura cristalográfica da 1ENY (Dessen *et al.*, 1995). A formação e quebra da ponte de hidrogênio entre o OH da Ser94 e a molécula wat276 foi acompanhada ao longo da trajetória (Figura 2.29). Enquanto a Gly14 permanece a uma distância da wat276 compatível com a formação de ponte de hidrogênio (menor do que 3,6 Å), o OH da serina desfaz periodicamente sua interação com esta molécula de água (estudos de DM *ab initio* também observaram o mesmo comportamento da cadeia lateral da Ser94 a enzima espécie selvagem, mostrando que à temperatura ambiente esta rotação da serina é energeticamente acessível - Pantano *et al.*, 2002). Portanto, baseado em nossa

análise e ao contrário do que foi proposto (Dessen *et al.*, 1995), a quebra desta rede de pontes de hidrogênio causada pela substituição da Ser94 por uma Ala não parece ser suficiente para justificar a menor afinidade deste mutante pelo NADH e nem sua resistência, uma vez que num processo dinâmico a quebra desta interação também ocorre com o aminoácido original.

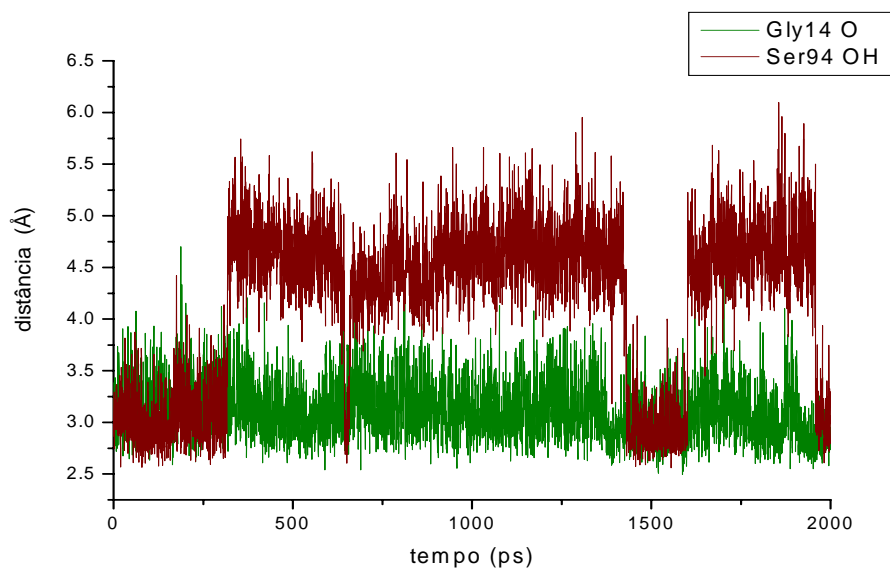


Figura 4.29. Distância entre a C=O da Gly14 (verde) e o OH da Ser94 (marrom) e a molécula de água wat276 durante a simulação por DM do sistema cWAT-NADH.

Na simulação por dinâmica molecular do sistema da enzima selvagem utilizando as moléculas de água do cristal, observou-se que aquelas águas cristalográficas localizadas abaixo do plano do NADH (principalmente as wat276 e wat304) foram preservadas, se movendo em torno da posição inicial (Figura 4.30).

As moléculas de água cristalográficas que inicialmente se localizavam em sítios mais expostos são rapidamente trocadas por moléculas de solvente ao longo da trajetória, inclusive a molécula de água wat297 localizada inicialmente abaixo do plano do anel A, devido, em grande parte, à acessibilidade ao solvente e flexibilidade da porção adenina do NADH. Aquelas moléculas de água mais internas, que não foram reproduzidas na primeira dinâmica, apresentaram uma ocorrência maior, fazendo pontes de hidrogênio efetivas com o ligante. Novamente,

as moléculas de água com maior ocorrência foram aquelas que se associaram à porção pirofosfato do NADH (Tabela 4.10).

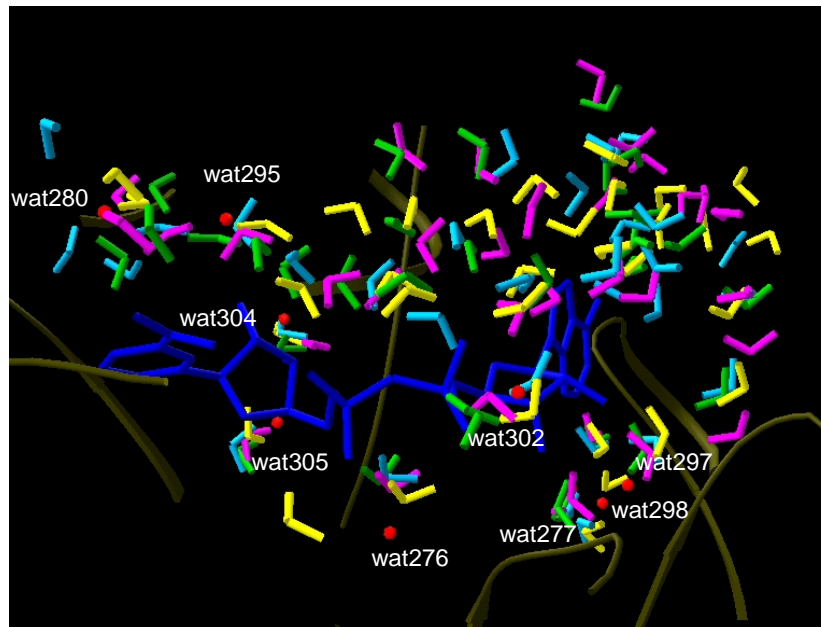


Figura 4.30. Moléculas de água cristalográficas (wat – representadas por seus átomos de oxigênio) reproduzidas durante a simulação por DM do sistema cWAT-NADH. Sobreposição das estruturas instantâneas da simulação cWAT-NADH em 0,05 ns (amarelo), 1 ns (magenta), 2 ns (verde) e 3 ns (azul). Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

A utilização das águas cristalográficas no sistema InhA-NADH mostrou ser importante para a reprodução de certas interações no sítio de ligação do NADH, principalmente no que diz respeito às interações entre a alça da glicina e a região pirofosfato do cofator. A importância destas águas também foi reforçada pela consideração destas moléculas de água em estudos de *docking* molecular do NADH onde as águas do sítio de ligação foram consideradas. Desta forma, a consideração destas águas cristalográficas nas simulações deveriam ser levadas em conta para uma avaliação mais precisa da afinidade enzima-cofator nos sistemas mutantes da InhA-NADH.

4.5 – Estudos de *docking* molecular para a avaliação da afinidade da enzima InhA pelo NADH

A avaliação da interação entre ligante e receptor em termos quantitativos é uma questão importante para o entendimento da especificidade do receptor por um determinado ligante. Uma das ferramentas atualmente mais utilizadas para avaliação deste reconhecimento molecular é o *docking* molecular (Brooijmans & Kuntz, 2003). Este método permite prever a estrutura de complexos ligante receptor e avaliar a energia de interação entre estas duas moléculas pela utilização de um algoritmo eficiente de busca conformacional e funções de avaliação de energia (Stoddard & Koshland, 1993; Sotriffer *et al.*, 2000).

Para a avaliação quantitativa da interação do sistema enzimático em estudo com o NADH realizamos estudos de *docking* molecular utilizando o programa AutoDock3.0 desenvolvido por Olson e colaboradores (Goodsell & Olson, 1990; Goodsell *et al.*, 1996; Morris *et al.*, 1996; Morris *et al.*, 1998). Este programa utiliza uma malha de afinidade pré-calculada para a avaliação rápida da energia de interação, reduzindo assim a demanda computacional da simulação. As energias de interação são calculadas em termos de energia livre de ligação (FEB – *free energy of binding*), e uma estimativa da constante de inibição (K_i) é realizada.

Inicialmente diversos experimentos de *docking* do NADH na estrutura cristalográfica da InhA espécie selvagem (1ENY) foram realizados para se testar a capacidade do algoritmo em reproduzir a ligação observada no cristal. A geometria inicial do ligante no sítio de ligação do receptor e a utilização de temperaturas mais elevadas não tiveram uma grande influência nos resultados finais, e RMSDs menores que 1,0 Å foram encontrados para as conformações com melhor avaliação de energia.

Experimentos de *docking* utilizando o protocolo de *simulated annealing* (SA) forneceu melhores resultados quando o ligante foi considerado rígido. Quando a flexibilidade do ligante foi considerada, o protocolo de algoritmos genéticos (GA) foi mais apropriado, pois convergiu mais rapidamente para a conformação de menor energia de ligação.

O programa AutoDock3.0 comporta bem a consideração de até 8 ângulos de torção. Entretanto, a consideração de mais de seis ligações flexíveis tornaram o cálculo via *simulated annealing* computacionalmente muito dispendioso. Além disso, quando a flexibilidade do NADH foi considerada nos experimentos via SA, a convergência para a ligação do cofator na forma estendida dentro do sítio de ligação necessitou de protocolos com maior número de passos de *docking*, mas nem sempre isso garantiu a obtenção de boas geometrias de ligação.

Para a verificação da capacidade do método na comparação da diferença de afinidade da ligação do NADH na InhA, diferentes condições de *docking* foram testadas na estrutura 1ENY e no mutante S94A (1ENZ) para a escolha do protocolo mais adequado para o sistema.

As melhores avaliações de FEB foram observadas com experimentos de *docking* do NADH via SA, embora este protocolo também gere conformações de alta energia em ambas estruturas cristalinas. O protocolo padrão de SA otimizado para o sistema consiste em 10 corridas de 50 ciclos cada, com critério de terminação máximo de 25.000 passos aceitos ou rejeitados, e está de acordo com o que é normalmente descrito na literatura (Lunney *et al.*, 1994; Coutinho *et al.*, 1997) (detalhes no Capítulo 3). Para o sistema em estudo, o aumento do número de corridas ou ciclos não resultou em melhores avaliações de energia, entretanto os valores dos critérios de terminação (número de passos aceitos ou rejeitados) devem ser mantido acima de 20.000. Em cada experimento de *docking* molecular uma série de 10 conformações finais do NADH rígido no sítio ativo do receptor foram geradas e classificadas de acordo com sua FEB (Figura 4.31).

É importante ressaltar que o programa AutoDock3.0 utiliza um gerador de números aleatórios para fazer o movimento inicial do ligante dentro do sítio de ligação. Por isso, a utilização do mesmo protocolo de *docking* para diferentes experimentos pode não resultar na mesma energia livre de interação. A fim de obter-se uma reprodução dos resultados nestes experimentos, o programa foi modificado de forma a utilizar um número fixo para a geração das conformações iniciais do ligante.

Nos experimentos envolvendo as estruturas cristalinas da InhA espécie selvagem e o mutante S94A, a energia livre de ligação da melhor conformação encontrada para a ligação do NADH com a enzima espécie selvagem foi mais

negativa (mais favorável) do que a registrada para o mutante S94A, como era de se esperar, e as respectivas conformações de melhor energia apresentaram RMSD menores do que 1,0 Å em relação à posição nos cristais. Apesar da pequena diferença de energia livre de ligação, e dos valores estimados de K_i não serem da mesma ordem de grandeza da constante de dissociação observada experimentalmente (na faixa de μM), a tendência experimental de maior afinidade da enzima espécie selvagem pelo NADH foi verificada (Dessen *et al.*, 1995; Quémard *et al.*, 1995).

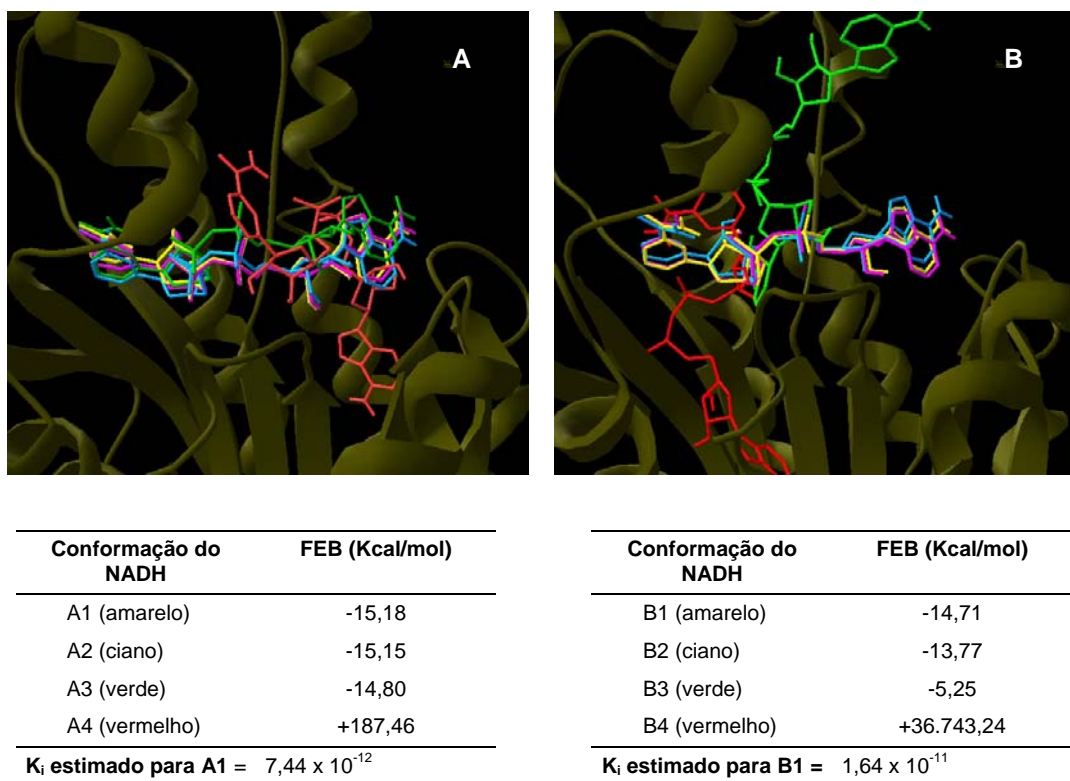


Figura 4.31. Comparação do *docking* molecular do NADH na estrutura cristalina da InhA espécie selvagem e do mutante S94A. (A) *Docking* molecular do NADH na estrutura cristalina da InhA espécie selvagem (1ENY). (B) *Docking* molecular do NADH na estrutura cristalina do mutante S94A (1ENZ). Referência: NADH na posição cristalográfica (magenta) e estrutura das enzimas em ocre. As conformações do NADH estão coloridas de acordo com a ordem decrescente de FEB: amarelo, ciano, verde e vermelho. Tabela com os valores de energia livre de ligação (FEB) para cada uma das conformações. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

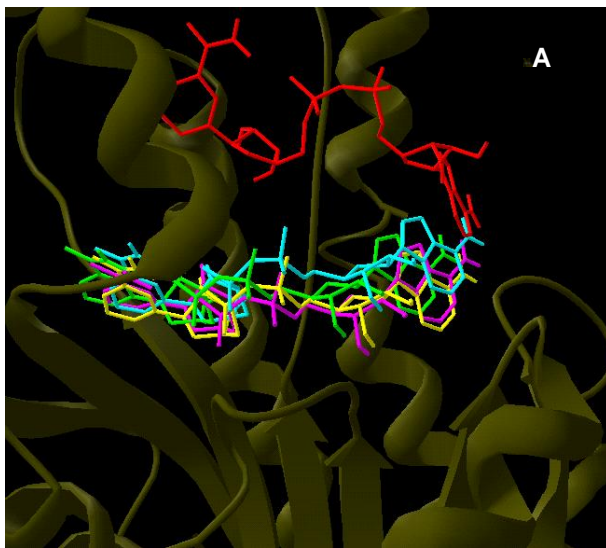
O *docking* molecular do NADH nas estruturas dos mutantes I21V, I16T e I47T modelados a partir da estrutura da enzima espécie selvagem, também

levaram a energias livres de interação menos favoráveis do que a encontrada para a *wt* InhA (Tabela 4.11). Portanto este método foi escolhido para a avaliação da diferença de afinidade dos diferentes sistemas da InhA pelo NADH.

Tabela 4.11. Comparação das energias livre de ligação (FEB) do NADH nas estruturas cristalinas da InhA espécie selvagem (1ENY), do mutante S94A (1ENZ) e dos modelos dos mutantes I21V, I16T e I47T calculadas a partir de experimentos de *docking* com o protocolo de *sumulated annealing*.

Sistema	FEB (Kcal/mol)
<i>wt</i> InhA	-15,18
S94A	-14,71
I21V	-10,65
I16T	-9,60
I47T	-10,67

A partir da observação das estruturas cristalinas da InhA e verificação nas simulações por DM, a interação do NADH com a enzima também é auxiliada pela presença de algumas moléculas de água bem localizadas. A fim de verificar a importância destas águas na ligação, os experimentos de *docking* foram repetidos na estrutura da 1ENY com a utilização das moléculas de água cristalográfica do sítio de ligação (Figura 4.32). A consideração destas moléculas de água levou a valores de energia livre de ligação bem mais favoráveis (-31,12 Kcal/mol), indicando, mais uma vez, que a consideração das águas conservadas do sítio de ligação do NADH são importantes na avaliação da afinidade da enzima pelo NADH.



Conformação do NADH	FEB (Kcal/mol)
A1	-31,12
A2	-22,92
A3	-19,09
A4	+800,31

K_i estimado para A1 = $1,56 \times 10^{-23}$

Figura 4.32. Docking do NADH na estrutura cristalina da InhA espécie selvagem (1ENY) com consideração das moléculas de águas cristalográficas. Referência: NADH na posição cristalográfica (magenta) e estrutura das enzimas em ocre. As conformações do NADH estão coloridas de acordo com a ordem decrescente de FEB: amarelo, ciano, verde e vermelho. Tabela com os valores de energia livre de ligação (FEB) para cada uma das conformações. Figura gerada com SwissPDB Viewer v3.7b2 (Guex & Peitsch, 1997).

4.6 – Utilização explícita da flexibilidade da enzima InhA nos experimentos de *docking*

A constatação da flexibilidade dos sistemas da enzima InhA espécie selvagem e mutantes nas simulações por DM indicou que a simples avaliação da interação do ligante em um único modelo estrutural da enzima InhA talvez não seja suficiente para a investigação de novos inibidores e também para justificar a diferença de associação verificada entre a espécie selvagem e mutantes.

Em seu estado nativo, as biomacromoléculas podem experimentar uma série de sub-estados de conformações com mínimos locais de energia causados por variações nas orientações das cadeias laterais e até mesmo na cadeia principal. Por isso, a utilização de uma única estrutura de receptor para a identificação de bons ligantes só é aplicável para um conjunto muito pequeno de conformações possíveis para este receptor, e para que uma única conformação seja representativa da estrutura de uma proteína, o sistema deve ser muito rígido (Carlson & McCammon, 2000).

Entretanto, a consideração da flexibilidade da macromolécula em experimentos de *docking* não é trivial. Enquanto a flexibilidade do ligante (geralmente uma molécula pequena) pode ser facilmente levada em conta nestes experimentos, a flexibilidade do receptor é difícil de ser tratada devido à complexidade do sistema (muitos graus de liberdade envolvidos).

Diferentes alternativas para a consideração da flexibilidade do receptor foram propostas, entre elas a utilização de bibliotecas de rotâmeros (Leach, 1994; Jackson *et al.*, 1998), o uso de características médias de um conjunto de estruturas (Knegtel *et al.*, 1997, Österberg *et al.*, 2002), otimização da conformação da ligação por minimização de energia ou rápidas dinâmicas moleculares (Schaffer & Verkhiver, 1998; Apostolakis *et al.*, 1998; Subramanian *et al.*, 1998). Entretanto, o método mais rigoroso de geração de subconjuntos de conformações é através da DM. Assim, uma série de conformações pode ser gerada e utilizada para estudos de *docking* (Pang & Kozikowski, 1994; Lin *et al.*, 2002).

Partindo do pressuposto que uma proteína existe em diversas conformações e que um bom ligante deve se ligar a ela em qualquer conformação disponível (Bouzida *et al.*, 1999), e conhecendo a flexibilidade dos sistemas da InhA,

utilizamos o conjunto de conformações das macromoléculas geradas nas simulações por DM para a incorporação da flexibilidade das enzimas InhA nos experimentos de *docking* e avaliação da interação da molécula de NADH nos diferentes sistemas. A afinidade do NADH foi avaliada para as estruturas instantâneas a cada 1 ps, utilizando-se o programa AutoDock3.0 e o mesmo protocolo de *simulated annealing* utilizado no *docking* nas estruturas cristalográficas. O NADH foi inicialmente considerado rígido na conformação estendida que assume na estrutura cristalina 1ENY, que corresponde a uma conformação intermediária entre as observadas durante as simulações. Somente a conformação de menor energia de ligação foi utilizada para análise.

O *docking* do NADH rígido nos diferentes sistemas simulados gerou um conjunto de energias de ligação que descrevem a afinidade do ligante em relação às diferentes conformações assumidas pelas macromoléculas ao longo de suas trajetórias dinâmicas. Este método se mostrou sensível a pequenas alterações nas conformações de cadeias laterais, e algumas energias muito altas foram encontradas para os mutantes I16T e I21V (apesar de terem sido observadas, com menos frequência, nos outros sistemas) (Figura 4.33A).

Desde o início dos *dockings* nas trajetórias da InhA espécie selvagem e mutantes foi observado que os valores de energia livre de ligação apresentou uma distribuição bimodal (Figura 4.33B), uma na faixa de -17 Kcal/mol e outra na faixa de -10 Kcal/mol. A inspeção das conformações de *docking* nestas duas regiões de FEB mostrou que, em todos os sistemas, as conformações finais do NADH com energia menos negativas correspondiam, em sua maioria, a avaliações de energia com o ligante fora de seu sítio, próximo à entrada da cavidade de ligação. Este comportamento indica que algumas modificações locais e momentâneas destes sítios de ligação não favoreceram a interação com o cofator na conformação estendida.

Para a avaliação e comparação das diferenças de interação entre os sistemas, somente aquelas energias livres de ligação onde o ligante apresentou um RMSD menor do que $5,0$ Å em relação à posição inicial de *docking* foram consideradas para garantir que estes valores de energia correspondessem à interação do NADH na cavidade de ligação.

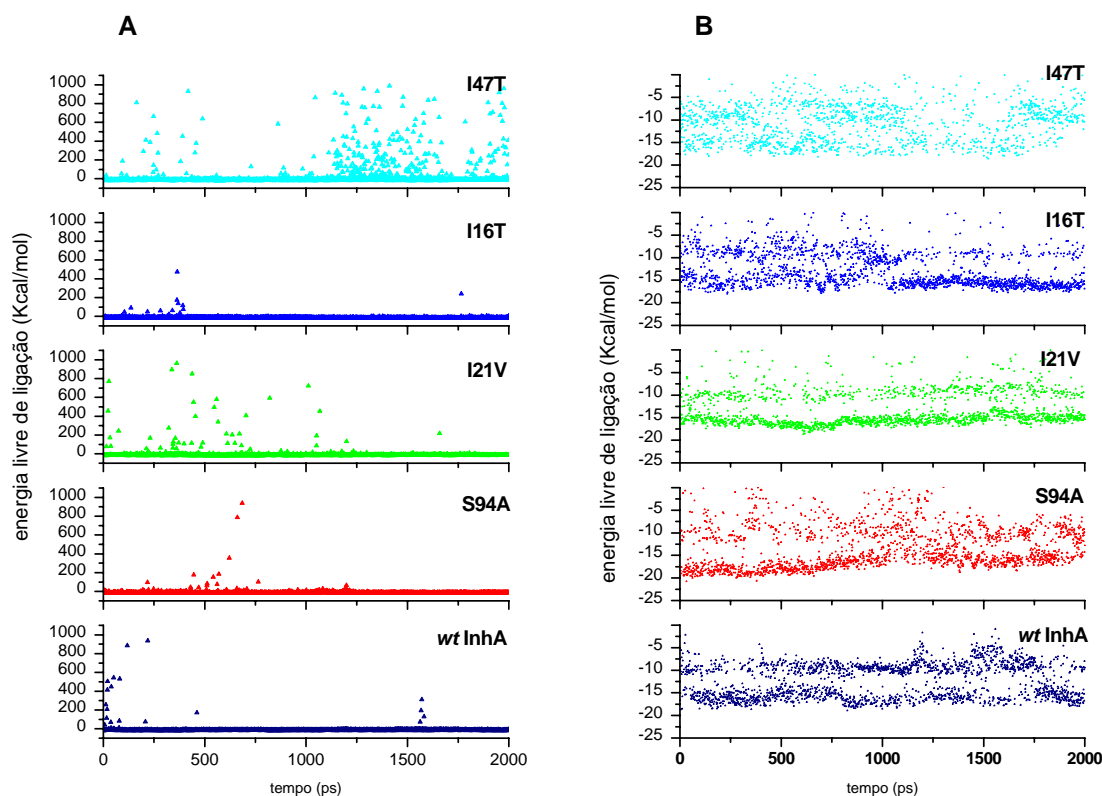


Figura 4.33. Avaliação de energia livre de ligação (FEB) calculadas por experimentos de *docking* molecular via *simulated annealing* do NADH na trajetória dos sistemas da InhA: *wt* InhA (azul escuro), S94A (vermelho), I21V (verde), I16T (azul) e I47T (ciano). (A) FEB durante os últimos 2 ns de simulação. (B) Redução da escala dos gráficos em A para visualização da região de melhor avaliação da FEB (região mais negativa).

A consideração de apenas aqueles valores de FEB correspondentes ao *docking* do ligante dentro do sítio de ligação resultou numa mudança do perfil da distribuição destes valores, aproximando-os da normalidade. Entretanto, alguns desvios ainda poderiam caracterizar uma distribuição bimodal ou uma distribuição com curtose (Apêndice 7).

A distribuição das energias livres de ligação entre NADH e enzimas foi então tratada estatisticamente pela análise de variância do tipo *One-Way Anova*. Esta análise revelou que o fator F, que avalia a relação de variação intra e inter conjuntos de dados, é extremamente elevado (65,28) e bem maior que o parâmetro F crítico definido para este caso (2,37) (Apêndice 7). A grande diferença entre o fator F calculado e o parâmetro F crítico permite que as análises sejam realizadas sem a necessidade de normalização dos dados. Esta análise de variância mostrou

que a hipótese h_1 de que pelo menos uma das médias é diferente não pode ser rejeitada. Ou seja, pelo menos uma das médias das energias livres dos diferentes sistemas difere das outras. O procedimento de Turkey foi então utilizado para a avaliação de quais das médias apresentam diferenças entre si (Apêndice 7). Este procedimento revelou que a média da energia livre de ligação no sistema *wt* InhA é estatisticamente igual à média no sistema S94A, mas difere das médias de FEB obtidas para os outros sistemas (Tabela 4.12).

Tabela 4.12. Valores médios e variância da energia livre de ligação (FEB) do NADH na trajetória dos sistemas da InhA, calculados por experimentos de *docking* molecular via *simulated annealing* considerando o ligante rígido na forma estendida.

Sistema enzimático	FEB média (Kcal/mol)	Variância
<i>wt</i> InhA	-15,53	3,40
S94A InhA	-15,33	9,23
I21V InhA	-15,18	3,43
I16T InhA	-14,58	5,61
I47T InhA	-14,00	7,36

Apesar de próximas, as médias das energias livres de ligação do NADH calculadas por experimentos de *docking* via SA nos diferentes sistemas mostraram uma pequena tendência de maior afinidade da enzima espécie selvagem em relação a seus mutantes I21V, I16T e I47T. Como observado qualitativamente durante as simulações por DM, a interação do mutante S94A com o NADH é muito semelhante à da espécie selvagem, indicando que para este mutante é necessário um tempo de simulação maior para que as mudanças conformacionais e de interação entre enzima e cofator sejam observadas.

A comparação visual das conformações de ligação do NADH propostas para aqueles valores de FEB de maior ocorrência mostrou que o modo de ligação do NADH no mutante S94A coincide com o da *wt* InhA, e ambos são próximos da estrutura cristalográfica.

Nas FEB de maior ocorrência nos mutantes I21V, I16T e I47T, a ligação do anel N e da ribose N coincide com a posição cristalográfica, entretanto, a porção adenina da molécula fica um pouco afastada do fundo da cavidade de ligação.

Contatos hidrofóbicos entre o anel A do NADH e o anel fenila da Phe97 se destacam nestas interações.

É interessante notar que no *docking* rígido do NADH nos sistemas I21V e I16T foram observadas conformações de ligação do NADH semelhantes às observadas durante as trajetórias destes sistemas, com a região do pirofosfato afastada da alça rica em glicina. Tais situações apresentam uma energia de interação em torno de $-9,5$ kcal/mol.

As conformações da enzima espécie selvagem geradas a partir da simulação do sistema cWAT-InhA também foram utilizadas nestes experimentos de *docking* para a avaliação da influência da utilização das águas cristalográficas na associação do cofator. Nestes experimentos somente a trajetória da macromolécula foi utilizada.

O *docking* do NADH na trajetória da cWAT-InhA levaram a avaliações de energia mais negativas para o sistema selvagem ($FEB = -17,14 \pm 2,05$ Kcal/mol) (Figura 4.34).

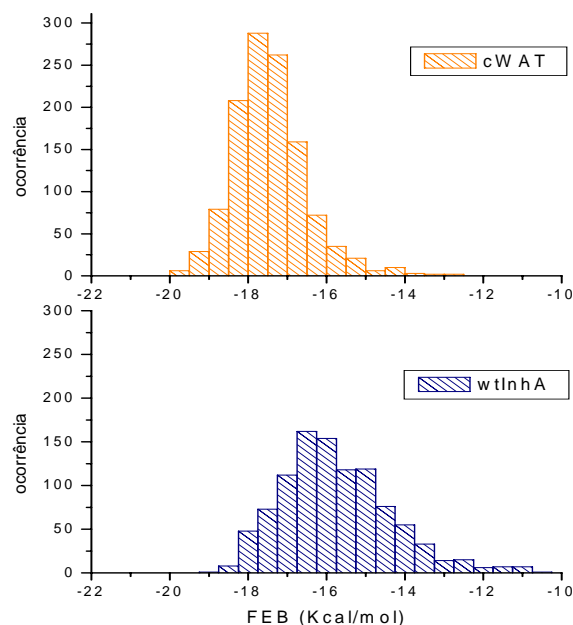


Figura 4.34. Distribuições da energia livre de ligação (FEB) calculadas por experimentos de *docking* molecular via *simulated annealing* do NADH na trajetória dos sistemas wTlnhA (azul escuro) e cWAT-InhA (laranja).

Apesar das energias livres de ligação do NADH neste sistema também apresentarem uma distribuição bimodal com a faixa de maior energia correspondente a conformações com RMSDs maiores do que 5 Å, a faixa de melhor FEB, que corresponde ao *docking* do ligante dentro do sítio do receptor, foi mais populada do que no sistema *wt* InhA. Isto indica que a utilização das águas cristalográficas durante a simulação por DM favorece a obtenção de conformações do sítio de ligação que acomodem melhor o ligante na sua forma estendida.

4.6.1 – Consideração da flexibilidade da enzima InhA e do ligante em experimentos de *docking*

Como foi observado que a flexibilidade dos sistemas simulados levam a uma variação da conformação do NADH durante as trajetórias, e que as diferentes conformações são decorrentes de rotações nos ângulos diedros do pirofosfato, a flexibilidade do ligante em experimentos de *docking* molecular também foi considerada, agora utilizando-se o protocolo de algoritmos genéticos otimizado para a estrutura cristalina.

Os 6 diedros centrais da porção pirofosfato do NADH foram considerados flexíveis para o *docking* deste ligante nas estruturas instantâneas geradas nas simulações por DM. A posição inicial do NADH foi considerada dentro do sítio de ligação.

A consideração da flexibilidade do ligante no *docking* via GA levou a uma distribuição mais homogênea nas FEB (Figura 4.35) na faixa de -15 Kcal/mol e poucas vezes foi observado o *docking* do ligante fora de seu sítio de ligação. Energias de ligação menos favoráveis foram encontradas para os mutantes I21V, I16T e I47T (Tabela 4.35). Entretanto, o mutante S94A novamente apresentou uma energia de ligação estatisticamente igual à da espécie selvagem (Tabela 4.13).

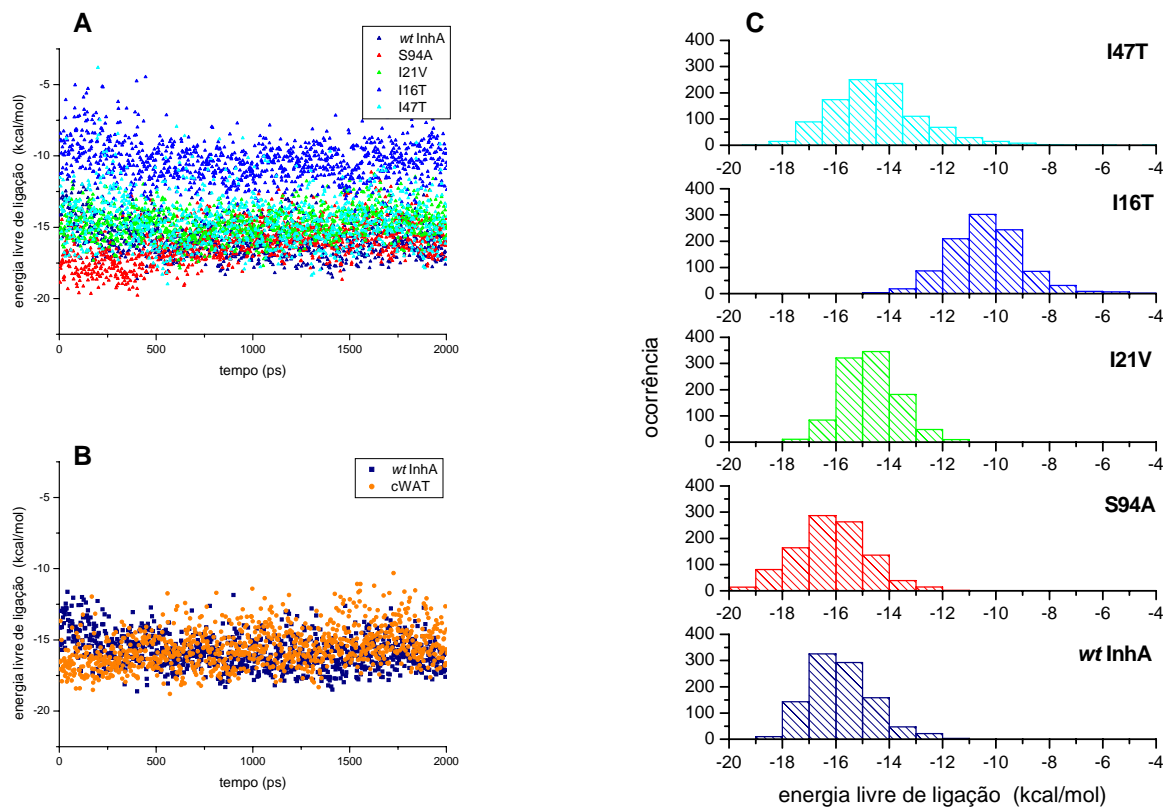


Figura 4.35. Avaliação da energia livre de ligação (FEB) do NADH calculados por experimentos de *docking* molecular nos sistemas da InhA via algoritmos genéticos. (A) Comparação das energias livre de ligação do NADH nos últimos 2 ns das trajetórias dos sistemas *wt*InhA (azul escuro), S94A (vermelho), I21V (verde), I16T (azul) e I47T (ciano). (B) Comparação das energias livre de ligação do NADH nos últimos 2 ns das trajetórias dos sistema *wt* InhA-NADH (azul escuro) e cWAT-INHA (laranja). (C) Histogramas da distribuição da FEB do NADH na *wt* InhA e mutantes.

Tabela 4.13. Valores de energia livre de ligação (FEB) do NADH na trajetória dos sistemas da InhA, calculados por experimentos de *docking* molecular via algoritmos genéticos, considerando a flexibilidade do ligante nas ligações do pirofosfato.

Sistema enzimático	FEB (Kcal/mol)
<i>wt</i> InhA	-15,81 ± 1,18
S94A InhA	-16,15 ± 1,36
I21V InhA	-14,72 ± 1,05
I16T InhA	-10,35 ± 1,43
I47T InhA	-14,46 ± 1,75

A comparação visual das conformações de ligação do NADH propostas para aqueles valores de FEB de maior ocorrência mostrou que o modo de ligação do NADH em todos os sistemas corresponde a sua forma estendida. Em casos isolados, a conformação fechada também foi observada, mas a porção nicotinamida do NADH apresenta-se sempre orientada para o fundo da cavidade de ligação.

Os valores de energia livre de ligação calculados nos experimentos de docking molecular via SA e GA apontaram para uma tendência de menor afinidade dos mutantes I21V, I16T e I47T pelo NADH do que a enzima espécie selvagem, embora as diferenças e a ordem de afinidade não tenham reproduzido os valores experimentais (Basso *et al.*, 1998). A associação menos efetiva destes sistemas com o NADH também foram notadas qualitativamente nas simulações por DM, onde as mutações na alça da glicina e a mutação na região de ligação da porção adenina levam a mudanças conformacionais no sítio de ligação que modificam o padrão de interação do cofator com a enzima.

Assim como observado na DM, a mutação Ser94 por alanina parece não alterar muito a afinidade pelo cofator calculada pelos experimentos de *docking* molecular, e talvez a trajetória do sistema S94A deva ser mais longa para que mudanças mais significativas possam ser observadas e refletidas no *docking* do NADH.

Experimentos de ligação do aduto inibidor NAD-INH indicaram que a associação do inibidor à InhA modula os contatos na interface da enzima no homotetrâmero (Rawat *et al.*, 2003), e ensaios enzimáticos mostraram que as mutações de InhA observadas em isolados clínicos resistentes à INH, além de diminuir a afinidade pelo NADH, também leva ao aparecimento de cooperatividade na ligação do cofator (Basso *et al.*, 1998). Pode ser que, para o sistema S94A, a existência de cooperatividade entre os sítios tenha uma maior importância para a modulação do sítio de interação do NADH no mutante S94A. Então a consideração de somente uma subunidade na simulação deste mutante pode não representar corretamente as interações que ocorrem no sistema *in vivo*.

Considerações Finais e Perspectivas

O trabalho aqui apresentado consiste no primeiro estudo teórico-computacional via simulação por dinâmica molecular e *docking* molecular para a comparação da associação do NADH à enzima 2-*trans*-Enoil-ACP (CoA) Redutase de *Mycobacterium tuberculosis* (espécie selvagem e mutantes observados em isolados clínicos resistentes à INH), e fornece informações importantes acerca da dinâmica destes sistemas e do mecanismo de resistência, demonstrando que:

- Apesar de mudanças conformacionais globais de grande amplitude não terem sido observadas, os sistemas estudados são bastante flexíveis. Esta flexibilidade por sua vez é refletida na orientação de alguns resíduos do sítio de ligação. Entretanto estes movimentos são característicos do sistema, e nem sempre estão relacionados com as mutações.

- Na simulação da apo-InhA, a ausência do ligante permite uma maior mobilidade dos aminoácidos do sítio ativo, que se orientam em direção ao centro da cavidade. Na presença do NADH a mobilidade de alguns aminoácidos do sítio de ligação é reduzida pela formação de pontes de hidrogênio importantes com a molécula do cofator.

- Na simulação do sistema *wt* InhA foi possível a verificação da importância do resíduo conservado Lys165 na ligação do cofator pela formação de pontes de hidrogênio com as hidroxilas da ribose N. A interação com este aminoácido e com a Ile 194 e Thr196 no fundo da cavidade de ligação, com a alça rica em glicina e com o Asp64 e a Val65 do sítio de reconhecimento da adenina são responsáveis pela manutenção da

conformação estendida do NADH com sua face B (Apêndice 1), e conseqüentemente o H_{4S} direcionado para a cavidade de ligação do substrato. Estas interações também são observadas nos sistemas mutantes, portanto devem ser levadas em conta no desenvolvimento de novos inibidores para esta enzima.

- A utilização das águas cristalográficas na simulação deste sistema enzimático (cWAT-InhA) mostrou ser importante para a reprodução de algumas interações entre enzima e cofator, uma vez que na simulação da *wt* InhA aquelas moléculas de água localizadas abaixo do plano da molécula do ligante não foram reproduzidas na escala de tempo das simulações.

- A mutação de aminoácidos do sítio ativo da enzima que levam à resistência parece afetar a ligação do cofator de diferentes formas. As mutações de aminoácidos da alça rica em glicina (mutantes I21V e I16T) é que levam a uma maior alteração no padrão de interação do NADH e instabilidade dos sistemas. Estas mutações diminuem a afinidade da enzima pela porção pirofosfato do NADH, que sofre mudanças em sua conformação, parecendo iniciar o processo de dissociação do cofator nestes sistemas.

- A mutação I47T, apesar de estar mais distante da molécula do ligante, altera o perfil de interação entre cadeias laterais da entrada da cavidade de ligação, próxima à porção adenina do NADH, causando uma contração deste sítio, da molécula do cofator, que se reflete na torção da região do pirofosfato, e a expulsão de algumas moléculas de água que intermedeiam a ligação com a enzima.

- Durante os 3,1 ns de simulação do sistema S94A-InhA, não foi detectada diferenças significativas no padrão de interação desta enzima com o ligante, que permaneceu com sua conformação muito próxima à da simulação da espécie selvagem. Dados experimentais indicaram a existência de cooperatividade na associação do ligante aos mutantes, que pode ter maior influência para este mutante, e devem ser consideradas para a

avaliação do comportamento desta enzima em relação à ligação com o cofator.

- Na simulação da enzima espécie selvagem com a consideração das águas cristalográficas, ficou demonstrado que o mecanismo de resistência proposto para o mutante S94A não é adequado, visto que também na espécie selvagem a rede de pontes de hidrogênio formada entre a cadeia lateral da Ser94, a Gly14, a ribose N e uma molécula de água ordenada (wat276) é constantemente formada e desfeita ao longo desta trajetória.

- Independentemente do sistema, as moléculas de água que interagem com a porção adenina do NADH são mais facilmente trocadas por moléculas do solvente, enquanto aquelas localizadas abaixo do plano do cofator apresentam um nível elevado de ocorrência.

- A utilização das estruturas instantâneas geradas nas simulações por DM permitiu a simulação da flexibilidade dos sistemas simulados durante os experimentos de *docking* molecular.

- A avaliação da energia livre de ligação (FEB) por experimentos de *docking* molecular demonstrou que, apesar da proximidade dos sistemas simulados, existe uma tendência de menor afinidade dos mutantes I21V, I16T e I47T pelo ligante. A grande semelhança entre o mutante S94A e a enzima espécie selvagem também foi atestada nestes experimentos.

- Os estudos preliminares de avaliação da energia livre de interação realizados pelo método MM-GBSA também mostraram a tendência de menor afinidade dos mutantes pelo NADH, entretanto, algumas inconsistências em relação aos dados obtidos em experimentos de *docking* e às observações nas simulações destes sistemas indicaram que, para a utilização deste método uma melhor avaliação da energia eletrostática de interação (baseada na solução da equação de Poisson-Boltzmann) e da contribuição da entropia devam ser consideradas.

Dentre os projetos futuros e perspectivas abertas por este trabalho podemos relacionar:

- A extensão do tempo de simulação dos sistemas mutantes para a melhor convergência dos parâmetros como o número de pontes de hidrogênio formadas diretamente entre o cofator e aminoácidos ou indiretamente através de moléculas águas, e a observação de maiores diferenças na ligação do cofator à enzima (principalmente no caso do mutante S94A).

- A consideração das águas cristalográficas também nas simulações dos mutantes para verificar a importância destas moléculas de água também na interação do cofator com os mutantes.

- Utilização daquelas moléculas de água importantes para a associação do NADH em seu sítio de ligação nos experimentos de *docking* molecular, para uma melhor avaliação das diferenças nas energias de ligação.

- Comparação e clusterização das estruturas geradas nas trajetórias, como 2D RMSD, para a agilização dos experimentos de *docking*.

Referências Bibliográficas

- Apostolakis, J.; Plückthun, A.; Caflish, A., 1998.** Docking small ligands in flexible binding sites. *J. Comput. Chem.*, **19**, 21-37.
- Bakowies, D. & vanGunsteren, W.F., 2002.** Simulation of apo and holo-fatty acid binding protein: structure and dynamics of protein, ligand and internal water. *J. Mol. Biol.*, **315**, 713-736.
- Banerjee, A.; Dubnau, E.; Quémard, A.; Balasubramanian, V.; Um, K.S.; Wilson, T.; Collins, D.; de Lisle, G.; Jacobs Jr., W.R., 1994.** *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science*, **263**, 227-230.
- Banerjee, A.; Sugantino, M.; Sacchettini, J.C.; Jacobs Jr., W.R., 1998.** The *mabA* gene from the *inhA* operon of *Mycobacterium tuberculosis* encodes a 3-ketoacyl reductase that fails to confer isoniazid resistance. *Microbiology*, **144**, 2697-2704.
- Barry III, C.E.; Lee, R.E.; Mdluli, K.; Sampson, A.E.; Schroeder, B.G.; Slayden, R.A.; Yuan, Y., 1998.** Mycolic acids: structure, biosynthesis and physiological functions. *Prog. Lipid Res.*, **37**, 143-179.
- Basso, L.A.; Zheng, R.; Blanchard, J.S., 1996.** Kinetics of inactivation of WT and C243S mutant of *Mycobacterium tuberculosis*: Enoyl Reductase by activated isoniazid. *J. Am. Chem. Soc.*, **118**, 11301-11302.
- Basso, L.A. & Blanchard, J.S., 1998.** Resistance to antitubercular drugs. *Adv. Exp. Med. Biol.*, **456**, 115-144.
- Basso, L.A.; Zheng, R.; Musser, J.M.; Jacobs Jr., W.R.; Blanchard, J.S., 1998.** Mechanisms of isoniazid resistance in *Mycobacterium tuberculosis*: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. *J. Infect.Dis.*, **178**, 769-775.
- Baxter, C.A. Murray, C.W.; Waszkowycz, B.; Li, J.; Sykes, R.A.; Bone, R.G.; Perkins, T.D.J.; Wylie, W., 2002.** New approach to molecular docking and its application to virtual screening of chemical databases. *J. Chem. Inf. Comput. Sci.*, **40**, 254-262.
- Bayly, C.I.; Cieplak, P.; Cornell, W.D.; Kollman, P.A., 1993.** A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model. *J. Phys. Chem.*, **97**, 10269-10280.
- Bell, C.E.; Yeates, T.O.; Eisenberg, D., 1997.** Unusual conformation of nicotinamide adenine dinucleotide (NAD) bound to diphtheria toxin: a comparison with NAD bound to the oxidoreductase enzymes. *Protein Sci.*, **6**, 2084-2096.
- Benkovic, S.J. & Hammes-Schiffer, S., 2003.** A perspective on enzyme catalysis. *Science*, **301**, 1196-1202.
- Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R., 1984.** Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81** (8), 3684-3690.

- Bernstein, J., Lott, W.A., Steinberg, B.A.; Yale, H.L., 1952.** Chemotherapy of experimental tuberculosis - V. Isonicotinic acid hydrazide (nydrazid) and related compounds. *Am. Rev. Tuberc.*, **65**, 357-364.
- Blanchard, J.S., 1996.** Molecular mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Annu. Rev. Biochem.*, **65**, 215-239.
- Bloom, B.R. & Murray, C.J.L., 1992.** Tuberculosis: Commentary on a reemergent killer. *Science*, **257**, 1055-1064.
- Böhm, H.J., 1994.** The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J. Comput. Aided Mol. Des.*, **8**, 243-256.
- Bottoms, C.A.; Smith, P.E.; Tanner, J.J., 2002.** A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci.*, **11**, 2125-2137.
- Bouzida, D.; Arthurs, S.; Colson, A.B.; Freer, S.T., Gehlaar, D.K.; Larson, V.; Luty, B.A.; Retjo, P.A.; Rose, P.W.; Verkhivker, G.M., 1999.** Thermodynamics and kinetics of ligand-protein binding studied with the weighted histogram analysis method and simulated annealing. *Pac. Symp. Biocomput.*, 426-437
- Brennan, P.J. & Nikaido, H., 1995.** The envelope of *Mycobacteria*. *Annu. Rev. Biochem.*, **64**, 29-63.
- Brooijmans, N. & Kuntz, I.D., 2003.** Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.*, **32**, 335-373.
- Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; Karplus, M., 1983.** CHARMM – a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, **4**, 187-217.
- Broussy, S.; Yannick, C.; Nguyen, M.; Bernadou, J.; Meunier, B., 2003.** ¹H and ¹³C NMR Characterization of hemiamidal isoniazid-NAD(H) adducts as possible inhibitors of InhA reductase of *Mycobacterium tuberculosis*. *Chem. Eur. J.*, **9**, 2034-3038.
- Camus, J-C.; Pryor, M.J.; Médigue, C.; Cole, S.T., 2002.** Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology*, **148**, 263-270.
- Carlson, H.A. & McCammon, A., 2000.** Accomodating protein flexibility in computational drug design. *Mol. Pharmacol.*, **57**, 213-218.
- Carugo, O. & Argos, P., 1997.** NADP-dependent enzymes. I: conserved stereochemistry of cofactor binding. *Proteins*, **28**, 10-28.
- Case, D.A; Pearlman, D.A.; Caldwell, J.W.; Cheatham, T.E. III; Ross, W.R.; Simmerling, C.L.; Darden, T.A.; Merz, K.M.; Stanton, R.V.; Cheng, A.L.; Vincent, J.J.; Crowley, M.; Tsui, V.; Radmer, R.J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G.L.; Singh, U.C.; Weiner, P.K. and Kollman, P.A., 1999.** AMBER 6.0, University of California, San Francisco.
- Cheatham, T.E.III & Kollman, P.A., 1996.** Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.*, **259**,434-444.
- Cheatham, T.E.III; Crowley, M.F.; Fox, T.; Kollman. P.A., 1997.** A molecular level picture of the stabilization of A-DNA in mixed ethanol-water solutions. *Proc. Natl. Acad. Sci. USA*, **94**, 9626-9630.
- Clemmens, D. L., 1996.** Characterization of the *Mycobacterium tuberculosis* phagosome. *Trends Microbiol.*, **4**, 113-118.

- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.-A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S. and Barrell, B.G., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, **393**, 537-544.
- Corbett, E.L.; Watt, C.J.; Walker, N.; Mahler, D.; Williams, B.G.; Raviglione, M.C.; Dye, C., 2003. The growing burden of tuberculosis. *Arch. Intern. Med.*, **163**, 1009-1021.
- Cornell, W.D.; Howard, A.E.; Kollman, P., 1991. Molecular mechanical potential functions and their application to study molecular systems. *Curr. Opin. Struct. Biol.*, **1**, 201-212.
- Cornell, W.D.; Cieplak, P. Bayly, C.I.; Gould, I.R.; Merz, K.M.Jr.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A., 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179-5197.
- Coutinho, P.M.; Dowd, M.K.; Reilly, P.J., 1997. Automated docking of monosaccharide substrates and analogues and methyl α -acarviosinide in the glucoamylase active site. *Proteins*, **27**, 235-248.
- Cronan, J.E. & Rock, C.O., 1996. In: *Escherichia coli and Salmonella: Cellular and Molecular Biology*, (Neidhardt, F.C. Ed.), ASM Press, Washington, pp 612-636.
- Darden, T.; York, D.; Pedersen, L., 1993. Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089-10092.
- Denessiouk, K.A.; Rantanen, V-V; Johnson, M.S., 2001. Adenine recognition: a motif present in ATP-, CoA-, NAD-, NADP-, and FAD-dependent proteins. *Proteins*, **44**, 282-291.
- Dessen, A.; Quémard, A.; Blanchard, J.S.; Jacobs Jr., W.R.; Sacchettini, J.C., 1995. Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. *Science*, **267**, 1638-1641.
- Dinadayala, P.; Laval, F.; Raynaud, C.; Lemassu, A.; Lanéelle, M.A.; Lanéelle, G.; Daffé, M., 2003. Tracking the pupative biosynthetic precursors of oxygenated mycolates of *Mycobacterium tuberculosis*. *J. Biol. Chem.*, **278**, 7310-7319.
- Draper, P., 1982. In: *The biology of the Mycobacteria*. (C.Ratledge and J. Stanford, Eds), Academic Press, London, p. 9-52.
- Duax WL, Griffin JF, Ghosh D. 1996. The fascinating complexities of steroid-binding enzymes. *Curr. Opin. Struct. Biol.*, **6**, 813-823
- Elber, R.; Ghosh, A.; Cárdenas, A., 2002. Long time dynamics of complex systems. *Acc. Chem. Res.*, **35**, 396-403.
- Eldridge, M.D.; Murray, C.W.; Auton, T.R.; Paoloni, G.V.; Mee, R.P.J., 1997. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.*, **11**, 425-445.
- Eschenbrenner, M; Chlumsky, L.J.; Khanna, P.; Strasser, F.; Jorns, M.S., 2001. Organization of the multiple coenzyme and subunits and role of the covalent flavin link in the complex heterotetrameric sarcosine oxidase. *Biochemistry*, **40**, 5352-5367.

- Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G., 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577-8593.
- Ewing, T.J.A.; Makino, S.; Skillman, A.G.; Kuntz, I.D., 2001. DOCK4.0: search strategies for automated molecular docking of flexible molecule databases. *J. Comp. Aided Mol. Des.*, **15**, 411-428.
- Faller, M.; Niederweis, M.; Schultz, G.E., 2004. The structure of a mycobacterial outer-membrane channel. *Science*, **303**, 1189-1192.
- Fillgrove, K.L. & Anderson, V.E., 2000. Orientation of Coenzyme A substrates, nicotinamide and active site functional groups in (di)enoyl-coenzyme A reductases. *Biochemistry*, **39**, 7001-7011.
- Fox, T. & Kollman, P.A., 1996. The application of different solvation and electrostatic models in molecular dynamics simulations of ubiquitin: how well is the X-ray structure "maintained"? *Proteins* **25**, 315-334.
- FUNASA, 2001. Guia Epidemiológico da Tuberculose, [http://www.funasa.gov.br/guia_epi/htm/doencas/tuberculose/index.htm], acessado em dezembro/2001.
- FUNASA, 2002. Situação da Prevenção e Controle das Doenças Transmissíveis no Brasil. Setembro/2002 [http://www.funasa.gov.br/guia_epi/], acessado em janeiro/2004.
- Gargallo, R.; Oliva, B.; Querol, E.; Avilés, F.X., 2000. Effect of the reaction field electrostatic term on the molecular dynamics simulation of the activation domain of procarboxypeptidase B. *Protein Eng.*, **13**, 21-26.
- Ginsburg, A.S.; Grosset, J.H.; Bishai, W.R., 2003. Fluorquinolones, tuberculosis, and resistance. *Lancet Infect. Dis.*, **3**, 432-442.
- Glickman, M.S. & Jacobs Jr., W.R., 2001. Microbial pathogenesis of *Mycobacterium tuberculosis*. *Cell*, **104**, 477-485.
- Gomez, J.E. & McKinney, J.D., 2004. *M. tuberculosis* persistence, latency, and drug tolerance. *Tuberculosis*, **84**, 29-44.
- Goodford, P.J., 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.*, **28**, 849-857.
- Goodsell, D.S. & Olson, A.J., 1990. Automated docking of substrates to proteins by simulated annealing. *Proteins*, **8**, 195-202.
- Goodsell, D.S.; Morris, G.M.; Olson, A.J., 1996. Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recognit.*, **9**, 1-5.
- Grosset, J., 1980. Bacteriologic basis of short course chemotherapy for tuberculosis. *Clin. Chest. Med.*, **1**, 231-241.
- Guex, N. & Peitsch, M.C., 1997. SWISS_MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.
- Heym, B.; Zhang, Y.; Poulet S.; Young, D.; Cole, S.T., 1993. Characterization of the *katG* gene encoding a catalase-peroxidase required for the isoniazid susceptibility of *Mycobacterium tuberculosis*. *J. Bacteriol.*, **175**, 4255-4259.
- Heym, B.; Alzari, P.M.; Honore, N.; Cole, S.T., 1995. Missense mutations in the catalase-peroxidase gene, *katG*, are associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Mol. Microbiol.*, **15**, 235-245.

- Hiriyanna, K.T. & Ramakrishnan, T., 1986. Deoxyribonucleic-acid replication time in *Mycobacterium tuberculosis*. *Arch. Microbiol.*, **144**, 105-109.
- Hou, T.; Wang, J.; Chen, L.; Xu, X., 1999. Automated docking of peptides and proteins by using a genetic algorithm combined with a tabu search. *Protein Eng.*, **12**, 639-647.
- Hubbard, S.J. & Thornton, J.M., 1993. NACCESS, Computer program, Department of Biochemistry and Molecular Biology, University College London.
- Hünenberger, P. H.; Mark, A.E.; vanGunsteren, W.F., 1995. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J. Mol. Biol.*, **252**, 492-503.
- Hünenberger, P.H. & McCammon, J.A., 1999a. Ewald artifacts in computer simulations of ionic solvation and ion-ion interaction: A continuum electrostatics study. *J. Chem. Phys.*, **110**, 1856-1872.
- Hünenberger, P.H. & McCammon, J.A., 1999b. Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: a continuum electrostatic study. *Biophys. Chem.*, **78**, 69-88
- Ibragimova, G.T. & Wade, R.C., 1998. Importance of explicit salt ions for protein stability in molecular dynamics simulation. *Biophys. J.*, **74**, 2906-2911.
- Jackson, R.M.; Gabb, H.A.; Sternberg, M.J.E., 1998. Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. *J. Mol. Biol.*, **276**, 265-285.
- Jayaram, B.; Sprous, D; Beveridge, D.L., 1998. S free energy of biomacromolecules: parameters for a modified generalized B model consistent with the AMBER force field. *J. Phys. Chem. B*, **102**, 9571-9576.
- Jensen, F., 1999. Introduction to computational chemistry. John Wiley & Sons Ltd, Chichester, England, p383-388.
- Johnsson, K. & Schultz, P.G., 1994. Mechanistic studies of the oxidation of isoniazid by the catalase peroxidase from *Mycobacterium tuberculosis*. *J. Am. Chem. Soc.*, **116**, 7425-7426.
- Johnsson, K.; King, D.S.; Schultz, P.G., 1995. Studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis. *J. Am. Chem. Soc.*, **117**, 5009-5010.
- Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M., 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926-935.
- Jorgensen, W.L. & Tirado-Rives, J., 1988. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, **110**, 1657-1666.
- Jorgensen, W.L.; Briggs, J.M.; Contreras, M.L., 1990. Relative partition coefficients for organic solutes from fluid simulations. *J. Phys. Chem.*, **94**, 1683-1686.
- Jornvall H, Persson B, Krook M, Atrian S, Gonzalez-Duarte R, Jeffery J, Ghosh D., 1995. Short-chain dehydrogenases/reductases (SDR). *Biochemistry*, **18**, 6003-6013.
- Kabsch, W. & Sander, C., 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-2637.
- Karplus, M. & Petsko, G.A., 1990. Molecular dynamics simulations in biology. *Nature*, **347**, 631-639.

- Karplus, M., 2002.** Molecular dynamics simulations of biomolecules. *Acc. Chem. Res.*, **35**, 321-323.
- Knegtel, R.M.A.; Kuntz, I.D.; Oshiro, C.M., 1997.** Molecular docking to ensembles of protein structures. *J. Mol. Biol.*, **266**, 424-440.
- Kolattukudy, P.E.; Fernandes, N.D.; Azad, A.K.; Fitzmaurice, A.M.; Sirakova, T.D., 1997.** Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. *Mol. Microbiol.*, **24**, 263-270.
- Kremer, L.; Dover, L.G.; Morbidoni, H.R.; Vilchèze, C.; Maughan, W.N.; Baulard, A.; Tu, S.-C.; Honoré, N.; Deretic, V.; Sacchettini, J.C.; Locht, C.; Jacobs Jr., W.R.; Besra, G.S., 2003.** Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in *Mycobacteria*. *J. Biol. Chem.*, **278**, 20547-20554.
- Krumrine, J.; Raubacher, F.; Brooijmans, N.; Kuntz, I., 2003.** Principles and methods of docking and ligand design. In: Structural Bioinformatics, Edited by Phipipe E. Bourne & Helge Weissig, Wiley-Liss Inc., New Jersey, USA.
- Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; Langridge, R.; Ferrin, T.E., 1982.** A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.*, **161**:269-288.
- Kuntz, I.D., 1992.** Structure-based strategies for drug design and discovery, *Science*, **257**, 1078-1082.
- Kuntz, I.D.; Meng, E.C.; Shoichet, B.K., 1994.** Structure-based molecular design. *Acc. Chem. Res.*, **27**, 117-123.
- Kuo, M.R.; Morbidoni, H.R.; Alland, d.; Sneddon, S.F.; Gourlie, B.B.; Staveski, M.M.; Leonard, M.; Gregory, J.S.; Janjigian, A.D.; Yee, C.; Musser, J.M.; Kreiswirth, B.; Iwamoto, H.; Perozzo, R.; Jacobs Jr., W.R.; Sacchettini, J.C.; Fodock, D.A., 2003.** Targeting tuberculosis and malaria through inhibition of enoyl reductase: compound activity and structural data. *J. Biol. Chem.*, **278** (23), 20851-20859.
- Larsen, M.H.; Vilchèze, C.; Kermer, L.; Besra, G.S.; Parsons, L.; Salfinger, M.; Heifets, L.; Hazbom, M.H.; Alland, D.; Sacchettini, J.C.; Jacobs Jr., W.R., 2002.** Overexpression of *inhA*, but not *kasA*, confers resistance to isoniazid and ethionamide in *Mycobacterium smegmatis*, *M. bovis* BCG and *M. tuberculosis*. *Mol. Microbiol.*, **46**, 453-466.
- Laskowski, R.A.; McArthur, M.W.; Moss, D.S.; Thornton, J.M., 1993.** PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**, 283-291.
- Lau, S.K.; Chass, G.A.; Lovas, S.; Penke, B.; Csizmadia, I.G., 2003.** An exploratory ab initio conformational analysis of selected fragments of nicotinamide adenine dinucleotide (NAD⁺). Part I: 5-deoxyribose nicotinamide *N*-glycoside. *J. Mol. Struct. (Theochem)*, **666-667**, 415-429.
- Leach, A.R., 1994.** Ligand docking to proteins with discrete side-chain flexibility. *J. Mol. Biol.*, **235**, 345-356.
- Leach, A.R., 1996.** Molecular Modelling: principles and applications. A.W.Longman Ltd. Ed., Essex, England, p131-182.
- Lei, B.; Wei, C-J; Tu, S-C, 2000.** Action mechanism of antitubercular isoniazid: activation by *Mycobacterium tuberculosis* KatG, isolation, and characterization of InhA inhibitor. *J. Biol. Chem.*, **275**, 2520-2526.
- Lengauer, T. & Rarey, M., 1996.** Computational methods for biomolecular docking. *Curr. Opin. Struct. Biol.*, **6**, 402-406.

- Lin, J.-H.; Perryman, A.L.; Schames, J.R.; McCammon, J.A., 2002. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J. Am. Chem. Soc.*, **124**, 5632-5633.
- Lunney, E.A.; Hagen, S.E.; Damagala, J.M.; Humblet, C.; Kosinski, J.; Tait, B.D.; Warmus, J.S.; Wilson, M.; Ferguson, D.; Hupe, D.; Tummino, P.J.; Baldwin, E.T.; Bhat, T.N.; Liu, B.; Erickson, J.W., 1994. A novel non peptide HIV-1 protease inhibitor: elucidation of the binding model and its application in the design of related analogs. *J. Med. Chem.*, **37**, 2664-2677.
- Lybrand, T.P., 1995. Ligand-protein and rational drug design. *Curr. Opin. Struct. Biol.*, **5**, 224-228.
- Manabe, Y.C. & Bishai, W.R., 2000. Latent *Mycobacterium tuberculosis* – persistence, patience, and winning by waiting. *Nat. Med.*, **6**, 1327-1329.
- Marrakchi, H.; Ducase, S.; Labesse, G.; Montrozier, H.; Margeat, E.; Emorine, L.; Charpentier, X.; Daffé, M.; Quémard, A., 2002. MabA (FabG1), a *Mycobacterium tuberculosis* protein involved in the long-chain fatty acid elongation system FAS-II. *Microbiology*, **148**, 951-960.
- McDonald, I.K. & Thornton, J.M., 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.*, **238**, 777-793.
- McKinney, J.D.; Jacobs Jr., W.R.; Bloom, B.R., 1998. Persisting problems in tuberculosis, in *Emerging Infection* (Krause, Richard M., Ed.), Academic Press, New York, p. 51-146.
- McKinney, J.D., 2000. *In vivo veritas*: The search for TB drug targets goes live. *Nat. Med.*, **6**, 1330-1333.
- Mdluli, K.; Slayden, R.A.; Zhu, Y.; Ramaswamy, S.; Pan, X.; Mead, D.; Crane, D.D.; Musser, J.M.; Barry III, C.E., 1998. Inhibition of a *Mycobacterium tuberculosis* β -ketoacyl ACP synthase by isoniazid. *Science*, **280**, 1607-1610.
- Middlebrook, G., 1952. Sterilization of the tubercle bacilli by isonicotinic acid hydrazide and the incidence of variants resistant to the drug *in vitro*. *Am. Rev. Tuberc.*, **65**, 765-767.
- Miesel, L.; Weisbrod, T.R.; Marcinkeviciene, J.A.; Bittman, R.; Jacobs Jr., W.R., 1998. NADH dehydrogenase defects confer isoniazid resistance and conditional lethality in *Mycobacterium smegmatis*. *J. Bacteriol.*, **180** (9), 2459-2467.
- Mitchison, D.A., 1985. The action of antitubercular drugs in short-course chemotherapy. *Tubercle*, **66**, 219-225.
- MOE (Molecular Operating Environment), 1998. Chemical Computing Group Inc., Montreal, Quebec, Canada. [<http://www.chemcomp.com>], acessado em dezembro/2002.
- Morris, G.M.; Goodsell, D.S.; Huey, R.; Olson, A.J., 1996. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock2.4. *J. Comput. Aided Mol. Des.*, **10**, 293-304.
- Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J., 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*, **19**, 1639-1662.
- Morris, G.M.; Goodsell, D.S.; Huey, R.; Hart, W.E.; Halliday, S.; Belew, R.; Olson, A.J., 2001. AutoDock User Guide - AutoDock: Automated docking of flexible ligands to receptors. Version 3.0.5. [<http://www.scripps.edu/pub/olson-web/doc/autodock/documentation.html>], acessado em março/2004.
- Muegge, I. & Martin, Y.C. 1999. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *J. Med. Chem.*, **42**, 791-804.

- Musser, J.M.; Kapur, V.; Williams, D.L.; Kreiswirth, B.N.; van Soolingen, D.; van Embden, J.D., 1996.** Characterization of the catalase-peroxidase gene (*katG*) and *inhA* locus in isoniazid-resistant and –susceptible strains of *Mycobacterium tuberculosis* by automated DNA sequencing: restricted array of mutations associated with drug resistance. *J. Infect. Dis.*, **173**, 196-202.
- Nakajima, N.; Makamura, H.; Kidera, A., 1997.** Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides. *J. Phys. Chem. B*, **101**, 817-824.
- Nelson, D.L. & Cox, M.M., 2000.** Lehninger principles of biochemistry. 3rd ed. Worth Publishers, New York, USA.
- Nguyen, M.; Quemard, A.; Marrakchi, H.; Bernadou, J.; Meunier, B., 2001.** The nonenzymatic activation of isoniazid by MnII-pyrophosphate in the presence of NADH produces the inhibition of the enoyl-ACP reductase *InhA* from *Mycobacterium tuberculosis*. *C. R. Acad. Sci. Paris, Série Iic, Chimie / Chemistry*, **4**, 35-40.
- Niederweis, M., 2003.** Mycobacterial porins – new channel proteins in unique outer membranes. *Mol. Microbiol.*, **45**, 1167-1177.
- Norberto de Souza, O., 1994.** Intrinsic curvature of DNA: a molecular dynamic study. *Ph.D. Thesis*. Department of Crystallography, Birbeck College, University of London, London, England.
- Norberto de Souza, O. & Ornstein, R.L., 1997.** Effect of periodic box size on aqueous molecular dynamics simulations of a DNA dodecamer with Particle-Mesh Ewald method. *Biophys. J.*, **72**, 2395-2397.
- Norberto de Souza, O. & Ornstein, R.L., 1999.** Molecular dynamic simulations of a protein-protein dimer: Particle-Mesh Ewald electrostatic model yields far superior results to standard cutoff model. *J. Biomol. Struct. Dyn.*, **16**, 1205-1218.
- Oppermann, U.; Filling, C.; Hult, M.; Shafqat, N.; wu, X.; Lindh, M.; Shafqat, J.; Nordling, E.; Kallberg, Y.; Persson, B.; Jornvall, H., 2003.** Schort-chain dehydrogenase/reductase (SDR): the 2002 update. *Chem. Biol. Interact.*, **143**, 247-253.
- Österberg, F.; Morris, G.M.; Sanner, M.F.; Olson, A.J.; Goodsell, D.S., 2002.** Automated docking to multiple target Structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins*, **46**, 34-40.
- Pak, Y. & Wang, S., 2000.** Application of a molecular dynamics simulation method with generalized effective potential to the flexible molecular docking problems. *J. Phys. Chem. B*, **104**, 354-359.
- Pang, Y-P. & Kozikowski, A.P., 1994.** Prediction of the binding site of huperizine A in acetylcholinesterase by docking studies. *J. Comput. Aided Mol. Des.*, **8**, 669-681.
- Pantano, S.; Alber, F.; Lamba, D.; Carloni, P., 2002.** NADH interactions with WT- and S94A-Acyl Carrier Protein Reductase from *Mycobacterium tuberculosis*: an *ab initio* study. *Proteins*, **46**, 34-40.
- Parikh, S.; Moynihan, D.P.; Xiao, G.; Tonge, P.J., 1999.** Roles of Tyrosine 158 and Lysine 165 in the catalytic Mechanism of *InhA*, the Enoyl-ACP Reductase from *Mycobacterium tuberculosis*. *Biochemistry*, **38**, 13623-13634.
- Parikh, S.L.; Xiao, G.; Tonge, P.J., 2000.** Inhibition of *InhA*, the Enoyl Reductase from *Mycobacterium tuberculosis*, by triclosan and isoniazid. *Biochemistry*, **39**, 7645-7650.
- Payne, D.J.; Warren, P.V.; Holmes, D.J.; Ji, Y.; Lonsdale, J.T., 2001.** Bacterial fatty-acid biosynthesis: a genomics-driven target for antibacterial drug discovery. *Drug Discov. Today*, **6**, 537-544.

- Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham, T.E. III; DeBolt, S.; Ferguson, D.; Seibel, G.L.; Kollman, P. A., 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.*, **91**, 1-41.
- Quémard, A.; Sacchettini, J.C.; Dessen, A.; Vilcheze, C.; Bittman, R.; Jacobs Jr., W.R.; Blanchard, J.S., 1995. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochemistry*, **34**, 8235-8241.
- Quémard, A.; Dessen, A.; Sugantino, M.; Jacobs Jr., W.R.; Sacchettini, J.C.; Blanchard, J.S., 1996. Binding of catalase-peroxidase-activated isoniazid to wild-type and mutant *Mycobacterium tuberculosis* Enoyl-ACP Reductases. *J. Am. Chem. Soc.*, **118**, 1561-1562.
- Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G., 1996. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.*, **261**, 470-489.
- Rawat, R.; Whitty, A.; Tonge, P.J., 2003. The isoniazid-NAD adduct is a slow, tight-binding inhibitor of InhA, the *Mycobacterium tuberculosis* enoyl reductase: adduct affinity and drug resistance. *Proc. Natl. Acad. Sci. USA*, **100**, 13881-13886.
- Rescigno, M. & Perham, R.N., 1994. Structure of the NADPH-binding motif of glutathione reductase: efficiency determined by evolution. *Biochemistry*, **33**, 5721-5727.
- Reyes, C.M. & Kollmann, P.A., 2000. Structure and thermodynamics of RNS-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. *J. Mol. Biol.*, **297**, 1145-1158.
- Rhee, J.T.; Piatek, A.S.; Small, P.M.; Harris, L.M.; Chaparro, S.V.; Kramer, F.R.; Alland, D., 1999. Molecular epidemiologic evaluation of transmissibility and virulence of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.*, **37**, 1764-1770.
- Richardson, J.S., 1981. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.*, **34**, 167-339.
- Rossmann, M.G.; Liljas, A.; Branden, C.-I.; Banaszak, L.J., 1975. Evolutionary and structural relationships among dehydrogenases. *Enzymes*, **11A**, 61-102.
- Rosny, P.J. & Karplus, M., 1979. Solvation. A molecular dynamics study of a dipeptide in water. *J. Am. Chem. Soc.*, **101**, 1913-1937.
- Rozwarski, D.A.; Grant, G.A.; Barton, D.H.; Jacobs Jr., W.R.; Sacchettini, J.C., 1998. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science*, **279**, 98-102.
- Rozwarski, D.A.; Vilchèze, C.; Sugantino, M.; Sacchettini, J.C., 1999. Crystal structure of the *Mycobacterium tuberculosis* Enoyl-ACP reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate. *J. Biol. Chem.*, **274**, 15582-15589.
- Russell, D.G., 2001. *Mycobacterium tuberculosis*: here today, and here tomorrow. *Nat. Rev. Mol. Cell Biol.*, **2**, 569-77.
- Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C., 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327-341.
- Ryde, U., 1995. Molecular dynamic simulations of alcohol dehydrogenase with a four- or five-coordinate catalytic zinc ion. *Proteins*, **21**, 40-56.

- Sacchettini, J.C. & Blanchard, J.S., 1996.** The structure and function of the isoniazid target in *Mycobacterium tuberculosis*. *Res. Microbiol.*, **147**, 36-43.
- Šali, A., 1998.** 100,000 protein structures for the biologist. *Nat. Struct. Biol.*, **5**, 1029-1032.
- Salman, M.; Brennan, P.J.; Lonsdale, J.T., 1999.** Synthesis of mycolic acids of mycobacteria: an assessment of the cell-free system in light of the whole genome. *Biochim. Biophys. Acta*, **1437**, 325-332.
- Schaffer, L. & Verkhivier, GM, 1998.** Predicting structural effects in HIV-1 protease mutant complexes with flexible ligand docking and protein side-chain optimization. *Proteins*, **33**, 295-310.
- Schaeffer, M.L.; Agnihotri, G.; Volker, C.; Kallender, H.; Brennan, P.J.; Lonsdale, J.T., 2001.** Purification and Biochemical Characterization of the *Mycobacterium tuberculosis* β -Ketoacyl-acyl Carrier Protein Synthases KasA and KasB. *J. Biol. Chem.*, **276**, 47029 – 47037.
- Schreiber, H. & Steinhauser, O., 1992a** Cutoff size does strongly influence molecular dynamics results on solvated polypeptides. *Biochemistry.*, **31**(25), 5856-5860.
- Schreiber, H. & Steinhauser, O., 1992b.** Taming cut-off induced artifacts in molecular dynamics studies of solvated polypeptides. The reaction field method. *J. Mol. Biol.*, **228**, 909-923.
- Schroeder, E.K.; de Souza, O.N.; Santos, D.S.; Blanchard, J.S.; Basso, L.A., 2002.** Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. *Curr. Pharm. Biotechnol.*, **3**, 197-225.
- Scior, T.; Morales, I.M.; Eisele, S.J.G.; Domeyer, D.; Laufer, S., 2002.** Antitubercular isoniazid and drug resistance of *Mycobacterium tuberculosis* – a review, *Arch. Pharm. Pharm. Med. Chem.*, **11**, 511-525.
- Sharp, K.A. & Honig, B., 1990.** Electrostatic interactions in macromolecules – theory and applications *Annu. Rev. Biophys. Biophys. Chem.*, **94**, 7684-7692.
- Sherman, D.R.; Mdluli, K.; Hickey, M.J.; Arain, T.M.; Morris, s.L.; Barry III, C.E.; Strover, C.K., 1996.** Compensatory *aphC* gene expression in isoniazid-resistant *Mycobacterium tuberculosis*. *Science*, **272**, 1641-1643.
- Slayden, R.A. & Barry III, C.E., 2002.** The role of KasA and KasB in the biosynthesis of meromycolic acids and isoniazid resistance in *Mycobacterium tuberculosis*. *Mol. Microbiol.*, **38** (3), 514-525.
- Slayden, R.A.; Lee, R.E.; Barry III, C.E., 2000.** Isoniazid affects multiple components of the type II fatty acid synthase system of *Mycobacterium tuberculosis*. *Mol. Microbiol.*, **38** (3), 514-525.
- Smith, P.E. & Pettit, B.M., 1991.** Peptides in ionic solutions: a comparison of the Ewald and switching function technics. *J. Chem. Phys.*, **95**, 8430-8441.
- Smith, P.E. & Pettit, B.M., 1994.** Modeling solvent in biomolecular systems. *J. Chem. Phys.*, **98**, 9700-9711.
- Smith, P.E & Tanner, J.J., 1999.** Molecular dynamics simulations of NAD⁺ in solution. *J. Am. Chem. Soc.*, **121**, 8637-8644.
- Smith, P.E & Tanner, J.J., 2000.** Conformations of nicotinamide adenine dinucleotide (NAD⁺) in various environments. *J. Mol. Recognit.*, **13**, 27-34.
- Smith, P.E.; Pettit, B.M.; Karplus, M., 1993.** Stochastic dynamics simulations of the alanine dipeptide using a solvent-modified potential energy surface. *J. Phys. Chem.*, **97**, 6907-6913.

- Srinivasan, J.; Miller, J.; Kollmann, P.A.; Case, D.A., 1998.** Continuum solvent studies of the stability of RNS hairpin loops and helices. *J. Biomol. Struct. Dyn.*, **16**, 671-682.
- Srinivasan, J.; Trevathan, M.W.; Beroza, P.; Case, D.A., 1999.** Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salts effects. *Theor. Chem. Acc.*, **101**, 426-434.
- Sotriffer, C.A.; Flader, W.; Winger, R.H.; Rode, B.M.; Liedl, K.R.; Varga, J.M., 2000.** Automated docking of ligands to antibodies: methods and applications. *Methods*, **20**, 280-291.
- Stocker, U.; Spiegel, K.; van Gunsteren, W.F., 2000.** On the similarity of propretirs in solution or in the crystalline state: A molecular dynamics study of hen lysosyme. *J. Biomol. NMR*, **18**, 1-12.
- Stoddard, B.L. & Koshland, D.E. Jr., 1993.** Molecular recognition analyzed by docking simulations: the aspartate receptor and isocitrate dehydrogenase from Escherichia coli. *Proc. Natl. Acad. Sci. USA*, **90**, 1146-1153.
- Subramanian, G.; Paterlini, M.G.; Larson, P.S.; Ferguson, D.M., 1998.** Conformational analysis and automated docking of selective arylacetamide-based *k*-opiod agonists. *J. Med. Chem.*, **41**, 4777-4789.
- Swaminathan, S.; Ravishanker, G.; Beveridge, D.L., 1991.** Molecular dynamics of B-DNA including water and counterions: a 140-ps trajectory for d(CGCGAATTCGCG) base on the GROMOS forcr field. *J. Am. Chem. Soc.*, **113**, 5027-5040.
- Takayama, K.; Wang, L.; David, H.L., 1972.** Effect of isoniazid on the *in vivo* mycolic acid biosynthesis, cell growth and viability of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.*, **2**, 29-35.
- Takayama, K.; Wang, L.; Merkal, R.S., 1973.** Scanning electron-mycroscopy of H37RA Strain of *Mycobacterium tuberculosis* exposed to isoniazid. *Antimicrob. Agents Chemother.*, **4**, 62-65.
- Takayama, K.; Schnoes, H.K.; Armstrong, E.L.; Byle, R.W., 1975.** Site of inhibitory action of isoniazid in the synthesis of mycolic acids in *Mycobacterium tuberculosis*. *J. Lipid Res.*, **16**, 308-317.
- Tasaki, K.; McDonald, S.; Brady, J.W., 1993.** Observations concerning the treatmente of long-range interactions in molecular dynamics simulations. *J. Comput. Chem.*, **14**, 278-284.
- Taylor, R.D.; Jewsbury, P.J.; Essex, J.W., 2002.** A review of protein-small molecule docking methods. *J. Comput. Aided Mol. Des.*, **16**, 151-166.
- Thormann, M. & Pons, M., 2001.** Massive docking of flexible ligands using environmental niches in parallelized genetic algorithms. *J. Comput. Chem.*, **22**, 1971-1982.
- Tironi, I.G.; Sperb, R.; Smith, P.E.; van Gunsteren, W.F., 1995.** A generalized reaction field methos for molecular dynamics simulations. *J. Chem. Phys.*, **102**, 5451-5459.
- Trosset, J.Y. & Scheraga, H.A., 1998.** Reaching the global minimum in docking simulations: a Monte Carlo energy minimization approach using Bezier splines. *Proc. Natl. Acad. Sci. USA*, **95**, 8011-8015.
- Tuckerman, M.E & Martyna, G.J., 2000.** Understanding Modern Molecular Dynamics: Techniques and Applications. *J. Phys. Chem. B*, **104**, 159-178.
- van Gunsteren, W.F. & Berendsen, H.J.C., 1977.** Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.*, **34**, 1311-1327.

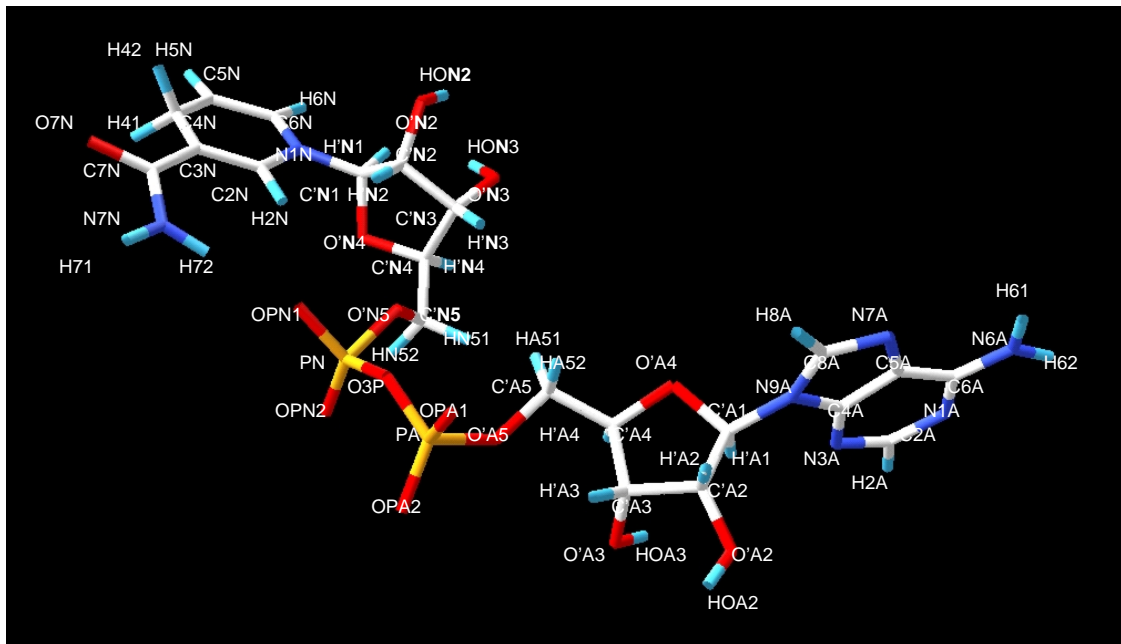
- van Gunsteren, W.F. & Karplus, M., 1982.** Protein dynamics in solution and in a crystalline environment: a molecular dynamics study. *Biochemistry*, **21**, 2259-2274.
- van Gunsteren, W.F. & Berendsen, H.J.C., 1987.** Groningen molecular simulation (GROMOS) library manual, BIOMOS: Nijenborgh, Groningen, The Netherlands.
- van Gunsteren, W.F. & Berendsen, H.J.C., 1990.** Computer simulation of molecular dynamics: methodology, applications and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.*, **29**, 992-1023.
- van Gunsteren, W.F. & Mark, A.E., 1992.** On the interpretation of biochemical data by molecular dynamics computer simulation. *Eur. J. Biochem.*, **204**, 947-961.
- van Gunsteren, W.F.; Hünenberger, P.H.; Mark, A.E.; Smith, P.E.; Tironi, I.G., 1995.** Computer simulation of protein motion. *Comp. Phys. Commun.*, **91**, 305-319.
- van Grunsven, E.G.; vanBerkel, E.; Ijlst, L.; Vreken, P.; de Klerk, J.B.; Adamski, J.; Lemonde, H.; Clayton, P.T.; Cuebas, D.A.; Wanders, R.J., 1998.** Peroxisomal D-hydroxyacyl-CoA dehydrogenase deficiency: resolution of the enzyme defect and its molecular basis in bifunctional protein deficiency. *Proc. Natl. Acad. Sci. USA*, **95**, 2128-2133.
- Verlet, L., 1967.** Computer 'experiments' on classical fluids I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, **159**, 98-103.
- Vieth, M.; Hirst, J.D.; Kolinski, A.; Brooks, C.L. III, 1998a.** Assessing energy functions for flexible docking. *J. Comput. Chem.*, **19**, 1612-1622.
- Vieth, M.; Hirst, J.D.; Dominy, B.N.; Daigler, H.; Brooks, C.L. III, 1998b.** Assessing search strategies for flexible docking. *J. Comput. Chem.*, **19**, 1623-1631.
- Voet, D. & Voet, J.G., 1995.** Biochemistry. 2nd ed. John Wiley & Sons, Inc., New York, USA.
- Wakil, S.J.; Stoops, J.K.; Joshi, V.C., 1983.** Fatty acid synthesis and its regulation. *Annu. Rev. Biochem.*, **52**, 537-579.
- Wallace, A.C.; Laskowski, R.A.; Thornton, J.M., 1995.** LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.*, **8**, 127-134.
- Wang, J.; Morin, P.; Wang, W.; Kollman, P.A., 2001.** Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT to TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J. Am. Chem. Soc.*, **123**, 5221-5230.
- Ward, W.H.J.; Holdgate, G.A.; Rowsell, S.; McLean, E.G.; Pauptit, R.A.; Clayton, E.; Nichols, W.W.; Colls, J.G.; Minshull, C.A.; Jude, D.A.; Mistry, A.; Timms, D.; Cample, R.; Hales, N.J.; Britton, C.J.; Taylor, I.W.F., 1999.** Kinetic and structural characteristics of the inhibition of Enoyl (Acyl Carrier Protein) Reductase by triclosan. *Biochemistry*, **38**, 12514-12525.
- Watanabe, M.; Aoyagi, Y.; Ridell, M.; Minnikin, D.E., 2001.** Separation and characterization of individual mycolic acids in representative mycobacteria. *Microbiology*, **147**, 1825-1837.
- Watanabe, M.; Aoyagi, Y.; Mitome, H.; Fujita, T.; Naoki, H.; Ridell, M.; Minnikin, D.E., 2002.** Location of functional groups in mycobacterial meromycolate chains: the recognition of new structural principles in mycolic acids. *Microbiology*, **148**, 1881-1902.
- Weber, W.; Hünenberger, P.H.; McCammon, J.A., 2000.** Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions: influence of artificial periodicity on peptide. *J. Phys. Chem. B*, **104**, 3668-3675.

- Weiner, S.J.; Kollman, P.A.; Case, D.A.; Singh, U.C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P., 1984.** A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, **106**, 765-784.
- WHO, 2001.** Stop TB Annual Report 2001. World Health Organization, Geneva, Switzerland, WHO/CDS/STB/2002.17. Reprint: [<http://www.who.int/gtb/publications/TBCatalogue.htm#2002>], acessado em fevereiro/2004.
- WHO, 2002.** Tuberculosis. Fact Sheet N° 104. World Health Organization, Geneva, Switzerland. August/2002. [<http://www.who.int/mediacentre/factsheets/who104/en/print.html>], acessado em março/2004.
- WHO, 2003.** The global plan to stop tuberculosis. World Health Organization, Geneva, Switzerland, WHO/CDS/STB/2003.23. Reprint: [<http://www.who.int/gtb/publications/TBCatalogue.htm#2003>], acessado em fevereiro/2004.
- WHO, 2004a.** Tuberculosis. Fact Sheet N° 104. World Health Organization, Geneva, Switzerland. March/2004. [<http://www.who.int/mediacentre/factsheets/fs104/en/print.html>], acessado em março/2004.
- WHO, 2004b.** Global tuberculosis control. In *WHO Report 2004*. World Health Organization, Geneva, Switzerland, WHO/HTM/TB/2004.331. Reprint [<http://www.who.int/gtb/publications/TBCatalogue.htm#2004>], acessado em março/2004.
- Wilming, M. & Johnsson, K., 1999.** Spontaneous formation of the bioactive form of the tuberculosis drug isoniazid. *Angew. Chem. Int. Ed. Engl.*, **38**, 2588-2590.
- Winder, F.G. & Collins, P.B., 1970.** Inhibition by isoniazid of synthesis of mycolic acids in *Mycobacterium tuberculosis*. *J. Gen. Microbiol.*, **63**, 41-48.
- Wüthrich, K., 2000.** Protein recognition by NMR. *Nat. Struct. Biol.*, **7**, 188-189.
- Young, D.B., 2001.** A post-genomic perspective. *Nat. Med.*, **7** (1), 11-12.
- Yuan, Y.; Crane, D.C.; Musser, J.M.; Sreevatsan, S.; Barry III, C.E., 1997.** MMAS-1, the branch point between *cis*- and *trans*-cyclopropane-containing oxygenated mycolates in *Mycobacterium tuberculosis*. *J. Biol. Chem.*, **272**, 10041-10049.
- Zabinski, R.F. & Blanchard, J.S., 1997.** The requirement for manganese and oxygen in the isoniazid-dependent inactivation of *Mycobacterium tuberculosis* Enoyl Reductase. *J. Am. Chem. Soc.*, **119**, 2331-2332.
- Zhang, Y.; Heym, B.; Allen, B.; Young, D.; Cole, S.T., 1992.** The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature*, **13**, 591-593.
- Zhang, Y.; Garble, T.; Young, D., 1993.** Transformation with *katG* restores isoniazid-sensitivity in *Mycobacterium tuberculosis* isolates resistant to a range of drug concentrations. *Mol. Microbiol.*, **8**, 521-524.
- Zhang, Y.; Wade, M.M.; Scorpio, A.; Zhang, H.; Sun, Z., 2003.** Mode of action of pyrazinamide: disruption of *Mycobacterium tuberculosis* membrane transport and energetics by pyrazinoic acid. *J. Antimicrob. Chemother.*, **52**, 790-795.
- Zhao, S.; Morris, G.M.; Olson, A.J.; Goodsell, D.S., 2001.** Recognition templates for predicting adenylate -binding sites in proteins. *J. Mol. Biol.*, **314**, 1245-1255.

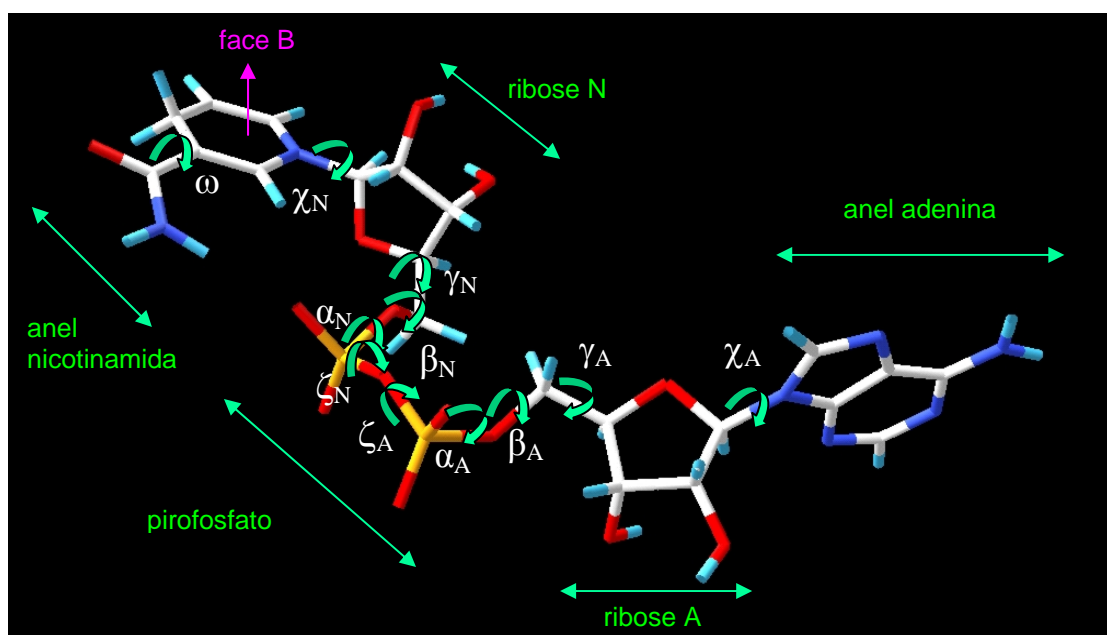
Apêndices

Apêndice 1

Molécula de NADH usada na simulação:



Ângulos diedros do NADH monitorados durante as simulações:



Parâmetros do NADH (Ryde, 1995)

NADH, AMBER 5.0 compatible; U. Ryde, to be published, 22/4-98

nah.dat

NAH	INT	1								
CORR	OMIT	DU	BEG							
0.00000										
1	DUMM	DU	M	0	-1	-2	.000	.000	.000	.00000
2	DUMM	DU	M	1	0	-1	1.000	.000	.000	.00000
3	DUMM	DU	M	2	1	0	1.000	90.000	.000	.00000
4	C6N	CA	M	3	2	1	1.000	80.257	304.544	-.35520
5	H6N	H4	E	4	3	2	1.086	86.845	113.956	.22190
6	C5N	CA	M	4	3	2	1.403	39.816	260.797	-.17300
7	H5N	HA	E	6	4	3	1.080	118.831	322.829	.12570
8	C4N	CT	M	6	4	3	1.395	120.143	142.477	.13480
9	H41	HC	E	8	6	4	1.093	106.309	122.831	.01960
10	H42	HC	E	8	6	4	1.093	107.814	239.002	.01960
11	C3N	CA	M	8	6	4	1.407	120.163	2.543	-.22270
12	C7N	C	B	11	8	6	1.505	119.504	173.166	.82850
13	O7N	O	E	12	11	8	1.233	117.812	21.930	-.60990
14	N7N	N	B	12	11	8	1.352	117.707	203.001	-.99350
15	H71	H	E	14	12	11	1.004	120.682	179.872	.40470
16	H72	H	E	14	12	11	1.009	120.477	1.302	.40470
17	C2N	CA	M	11	8	6	1.382	119.053	355.721	-.18400
18	H2N	H4	E	17	11	8	1.084	118.597	179.611	.14220
19	N1N	N*	M	17	11	8	1.387	120.593	1.805	.22930
20	C'N1	CT	M	19	17	11	1.488	116.524	174.865	-.03840
21	H'N1	H2	E	20	19	17	1.094	110.618	20.227	.16170
22	C'N2	CT	M	20	19	17	1.508	117.258	142.452	.06700
23	H'N2	H1	E	22	20	19	1.088	113.807	46.086	.09720
24	O'N2	OH	S	22	20	19	1.436	111.340	277.899	-.61390
25	HON2	HO	E	24	22	20	.959	106.795	89.056	.41860
26	C'N3	CT	M	22	20	19	1.520	99.733	162.298	.20220
27	H'N3	H1	E	26	22	20	1.091	114.651	198.096	.06150
28	O'N3	OH	S	26	22	20	1.423	109.790	79.530	-.65410
29	HON3	HO	E	28	26	22	.960	105.431	40.651	.43760
30	C'N4	CT	M	26	22	20	1.521	103.900	322.360	.10650
31	H'N4	H1	E	30	26	22	1.090	111.514	142.236	.11740
32	O'N4	OS	E	30	26	22	1.435	104.712	23.472	-.35480
33	C'N5	CT	M	30	26	22	1.521	114.250	264.249	.05580
34	HN51	H1	E	33	30	26	1.029	113.342	288.300	.06790
35	HN52	H1	E	33	30	26	1.092	118.256	168.168	.06790
36	O'N5	OS	M	33	30	26	1.440	108.747	48.875	-.56610
37	PN	P	M	36	33	30	1.620	117.474	190.938	1.37290
38	OPN1	O2	E	37	36	33	1.473	105.958	176.012	-.85580
39	OPN2	O2	E	37	36	33	1.481	109.032	308.271	-.85580
40	O3P	OS	M	37	36	33	1.574	100.139	62.417	-.57600
41	PA	P	M	40	37	36	1.611	136.620	195.390	1.37290
42	OPA1	O2	E	41	40	37	1.485	106.319	337.227	-.85580
43	OPA2	O2	E	41	40	37	1.477	109.295	207.094	-.85580
44	O'A5	OS	M	41	40	37	1.580	103.587	92.769	-.56610
45	C'A5	CT	M	44	41	40	1.420	121.317	75.897	.05580
46	HA51	H1	E	45	44	41	1.024	108.046	267.459	.06790
47	HA52	H1	E	45	44	41	1.093	115.908	35.989	.06790
48	C'A4	CT	M	45	44	41	1.503	100.744	154.082	.10650
49	H'A4	H1	E	48	45	44	1.104	106.529	59.863	.11740
50	O'A4	OS	E	48	45	44	1.414	106.546	175.953	-.35480
51	C'A3	CT	M	48	45	44	1.523	112.631	292.482	.20220
52	H'A3	H1	E	51	48	45	1.092	111.048	-18.541	.06150
53	O'A3	OH	S	51	48	45	1.434	110.551	135.535	-.65410
54	HOA3	HO	E	53	51	48	.962	106.326	273.873	.43760
55	C'A2	CT	M	51	48	45	1.511	103.289	252.113	.06700
56	H'A2	H1	E	55	51	48	1.091	107.136	88.678	.09720
57	O'A2	OH	S	55	51	48	1.429	112.235	212.910	-.61390
58	HOA2	HO	E	57	55	51	.961	108.105	20.361	.41860
59	C'A1	CT	M	55	51	48	1.500	102.172	330.379	.03940
60	H'A1	H2	E	59	55	51	1.096	110.904	284.037	.20070
61	N9A	N*	M	59	55	51	1.481	115.254	161.267	-.02510
62	C8A	CK	M	61	59	55	1.339	126.756	321.228	.20060
63	H8A	H5	E	62	61	59	1.081	125.157	359.651	.15530
64	N7A	NB	M	62	61	59	1.344	111.176	179.494	-.60730
65	C5A	CB	M	64	62	61	1.335	102.981	.509	.05150
66	C6A	CA	M	65	64	62	1.375	123.335	180.071	.70090
67	N6A	N2	B	66	65	64	1.334	120.288	352.032	-.90190
68	H61	H	E	67	66	65	1.001	123.250	184.224	.41150
69	H62	H	E	67	66	65	1.004	122.140	3.982	.41150
70	N1A	NC	M	66	65	64	1.334	115.111	179.506	-.76150
71	C2A	CQ	M	70	66	65	1.345	122.289	359.037	.58750
72	H2A	H5	E	71	70	66	1.079	119.198	180.231	.04730
73	N3A	NC	M	71	70	66	1.345	121.528	1.544	-.69970
74	C4A	CB	M	73	71	70	1.341	119.342	358.936	.30530


```

CHARGE NADH
-0.3552 0.2219 -0.1730 0.1257 0.1348
 0.0196 0.0196 -0.2227 0.8285 -0.6099
-0.9935 0.4047 0.4047 -0.1840 0.1422
 0.2293 -0.0384 0.1617 0.0670 0.0972
-0.6139 0.4186 0.2022 0.0615 -0.6541
 0.4376 0.1065 0.1174 -0.3548 0.0558
 0.0679 0.0679 -0.5661 1.3729 -0.8558
-0.8558 -0.5760 1.3729 -0.8558 -0.8558
-0.5661 0.0558 0.0679 0.0679 0.1065
 0.1174 -0.3548 0.2022 0.0615 -0.6541
 0.4376 0.0670 0.0972 -0.6139 0.4186
 0.0394 0.2007 -0.0251 0.2006 0.1553
-0.6073 0.0515 0.7009 -0.9019 0.4115
 0.4115 -0.7615 0.5875 0.0473 -0.6997
 0.3053

```

```

IMPROPER
C6N C2N N1N C'N1
N1N C3N C2N H2N
C2N C4N C3N C7N
C4N C6N C5N H5N
C5N N1N C6N H6N
C3N N7N C7N O7N
C7N H71 N7N H72
C8A C4A N9A C'A1
C6A H61 N6A H62
N7A N9A C8A H8A
N1A N3A C2A H2A
C5A N1A C6A N6A

```

```

LOOP
N1N C6N
C'N1 O'N4
O'A4 C'A1
C4A C5A
C4A N9A

```

DONE

Parâmetros do campo de força do NADH:

Special force constants for NAD⁺, NADH, NADP⁺, and NADPH; Ulf Ryde, 24 April, 1998

BOND

CA-N* 448.00 1.344 NAD: N1-C2 and N1-C6

ANGL

CA-C -O 80.00 119.80 NAD, Vanhommerig, BBA 1295(96)125
CA-C -N 70.00 115.70 NAD, Vanhommerig, BBA 1295(96)125
C -CA-CT 63.00 120.00 NAD
CA-CA-N* 70.00 119.50 NAD, Vanhommerig, BBA 1295(96)125
CT-CA-HA 35.00 120.00 NAD
H4-CA-N* 35.00 114.00 NAD
CA-CT-CA 63.00 120.00 NAD
CA-N*-CA 70.00 121.20 NAD, Vanhommerig, BBA 1295(96)125
CA-N*-CT 70.00 120.60 NAD, Vanhommerig, BBA 1295(96)125

DIHE

X -CA-N*-X 4 7.40 180.0 2. NAD; =X-CM-N*-X
CA-CA-C -O 1 0.02 180.0 1. NAD, Vanhommerig, BBA 1295(96)125
CA-CA-C -O 1 0.45 180.0 2. NAD, Vanhommerig, BBA 1295(96)125
CA-CA-C -N 1 0.02 180.0 1. NAD, Vanhommerig, BBA 1295(96)125
CA-CA-C -N 1 0.45 180.0 2. NAD, Vanhommerig, BBA 1295(96)125

IMPR

CA-CA-N*-CT 1.0 180. 2. NAD
CA-CA-CA-C 1.1 180. 2. NAD
CA-CT-CA-C 1.1 180. 2. NAD
CT-CA-CA-C 1.1 180. 2. NAD

END

Apêndice 2

Arquivos de entrada para as simulações da dinâmica molecular:**mini_00.in:**

```
# Belly Minimization. NTF=NTC=2
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=1, ntr=0, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=100, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
        a=78, b=65, c=70,
&end
Belly Dynamics.Only solvent (WAT + CIO) Atoms move.
RES 270 9407
END
END
```

mdcp_00.in:

```
# 20 ps NPT. NTF+NTC=2. Equilibrating Solvent (WAT + CIO) from 10 to 298.16K. NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntr=1,
        ntso=1, ntp=50,
        ntw=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=1, ntr=0, cut=9.0,
        nstlim=10000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
        temp0=298.16, tempi=10.0, ig=555333, heat=0.0001, ntt=1,
        dtemp=10.0, tautp=1.0, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=2.0, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=78, b=65, c=70,
&end
Belly Dynamics.Only solvent (WAT + CIO) Atoms move.
RES 270 9407
END
END
```

mini_01.in:

```
# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
        a=74.7676060, b=62.3063383, c=67.0991336,
&end
Restraining Protein+NAH Only
1000.000
RES 1 269
END
END
```

mini_02.in:

```
# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
```

```
&end
&ewald
      a=74.7676060, b=62.3063383, c=67.0991336,
&end
Restraining Protein+NAH Only
500.000
RES 1 269
END
END
```

mini_03.in:

```
# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
      ntr=10,
      ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=1, cut=9.0,
      ntc=2, tol=0.00001,
      maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
      a=74.7676060, b=62.3063383, c=67.0991336,
&end
Restraining Protein+NAH Only
250.000
RES 1 269
END
END
```

mini_04.in:

```
# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
      ntr=10,
      ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=1, cut=9.0,
      ntc=2, tol=0.00001,
      maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
      a=74.7676060, b=62.3063383, c=67.0991336,
&end
Restraining Protein+NAH Only
100.000
RES 1 269
END
END
```

mini_05.in:

```
# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
      ntr=10,
      ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=1, cut=9.0,
      ntc=2, tol=0.00001,
      maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
      a=74.7676060, b=62.3063383, c=67.0991336,
&end
Restraining Protein+NAH Only
50.000
RES 1 269
END
END
```

mini_06.in:

```
# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
      ntr=10,
      ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=1, cut=9.0,
      ntc=2, tol=0.00001,
      maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
      a=74.7676060, b=62.3063383, c=67.0991336,
&end
Restraining Protein+NAH Only
```

```

10.000
RES 1 269
END
END

```

mini_07.in:

```

# Unrestrained Minimization.NTF=NTC=2.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
        a=74.7676060, b=62.3063383, c=67.0991336,
&end

```

mdcp_01.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 10K to 50K.NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntr=1,
        ntso=1, ntp=50,
        ntw=250, iwrap=0, ntw=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
        temp0=50.0, tempi=10.0, ig=555333, heat=0.0001, ntt=1,
        dtemp=10.0, tautp=1.0, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=2.0, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=74.7676060, b=62.3063383, c=67.0991336,
&end

```

mdcp_02.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 50K to 100K.NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntr=1,
        ntso=1, ntp=50,
        ntw=250, iwrap=0, ntw=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
        temp0=100.0, tempi=50.0, ig=555333, heat=0.0001, ntt=1,
        dtemp=10.0, tautp=0.2, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=72.7648821, b=60.6374017, c=65.3018173,
&end

```

mdcp_03.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 100K to 150K.NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntr=1,
        ntso=1, ntp=50,
        ntw=250, iwrap=0, ntw=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
        temp0=150.0, tempi=100.0, ig=555333, heat=0.0001, ntt=1,
        dtemp=10.0, tautp=0.2, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=72.9464506, b=60.7887088, c=65.4647634,
&end

```

mdcp_04.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 150K to 200K.NPSCAL=1

```

```

&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
        nt xo=1, ntpr=50,
        ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=1000, ndfmin=0, ntc m=0, nscm=0, t=0.0, dt=0.002,
        temp0=200.0, tempi=150.0, ig=555333, heat=0.000001, ntt=1,
        dtemp=10.0, tautp=0.2, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=73.2435078, b=61.0362565, c=65.7313532,
&end

```

mdcp_05.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 200K to 250K.NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
        nt xo=1, ntpr=50,
        ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=1000, ndfmin=0, ntc m=0, nscm=0, t=0.0, dt=0.002,
        temp0=250.0, tempi=200.0, ig=555333, heat=0.000001, ntt=1,
        dtemp=10.0, tautp=0.2, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=73.5176566, b=61.2647139, c=65.9773842,
&end

```

mdcp_06.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 250K to 298.16K.NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
        nt xo=1, ntpr=50,
        ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=1000, ndfmin=0, ntc m=0, nscm=0, t=0.0, dt=0.002,
        temp0=298.16, tempi=250.0, ig=555333, heat=0.000001, ntt=1,
        dtemp=10.0, tautp=0.2, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=73.8630713, b=61.5525595, c=66.2873718,
&end

```

mdcp_prod_02.in:

```

# 50ps NPT. NTF+NTC=2. taup=0.8,taup=1.0. NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=7, irect=1, ntrx=1,
        nt xo=1, ntpr=50,
        ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=25000, ndfmin=0, ntc m=0, nscm=0, t=0.0, dt=0.002,
        temp0=298.16, tempi=298.16, ig=555333, heat=0.0001, ntt=1,
        dtemp=10.0, tautp=0.6, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=1.0, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=74.7676060, b=62.3063383, c=67.0991336,
&end

```

Apêndice 3

Sistemas das simulações da enzima InhA de *Mycobacterium tuberculosis*:**wt InhA:**

- Coord. Iniciais: 1ENY.pdb
- Tempo total de simulação: 3950ps
- Cluster: Dell PC (0-1000ps) e Ombrófila (1050-3950ps)

wt InhA	RESÍDUOS	ÁTOMOS
InhA	1-268	1-4008
4 Na+	269-272	4009-4012
TIP3P (8229 molec.)	273-8501	4013-28699
Total	8501 resíduos	28699 átomos

wt InhA_NADH:

- Coord. Iniciais: 1ENY.pdb
- Tempo total de simulação: 3100ps
- Cluster: Amazônia (0-2100ps) e Ombrófila (2150-3100ps)

wt InhA_NADH	RESÍDUOS	ÁTOMOS
InhA	1-268	1-4008
NADH	269	4009-4079
6 Na+	270-275	4080-4085
TIP3P (9682 molec.)	276-9407	4086-31481
Total	9407 resíduos	31481 átomos

cWAT InhA_NADH :

- Coord. Iniciais: 1ENY.pdb e suas 41 moléculas de águas cristalográficas
- Tempo total de simulação: 3000ps
- Cluster: Ombrófila (0-3600ps)

cWAT_InhA_NADH	RESÍDUOS	ÁTOMOS
InhA	1-268	1-4008
NADH	269	4009-4079
6 Na+	270-275	4080-4085
41 WAT cristalográficas	276-316	4086-4208
TIP3P (9033 molec.)	317-9349	4209-31307
TIP3P _{total} (9074 molec.)	276-9349	4086-31307
Total	9349 resíduos	31307 átomos

S94A InhA_NADH:

- Coord. Iniciais: INHA_NADH 1900ps (mdcp_1900ps.cav)
- Tempo total de simulação: 3100ps
- Cluster: Ombrófila e Amazônia

S94A_InhA_NADH	RESÍDUOS	ÁTOMOS
InhA	1-268	1-4007
NADH	269	4008-4078
6 Na+	270-275	4079-4084
TIP3P (9132 molec.)	276-9407	4085-31480
Total	9407 resíduos	31480 átomos

I21V InhA NADH:

- Coord. Iniciais: INHA_NADH 1900ps (mdcp_1900ps.cav)
- Tempo total de simulação: 6500ps
- Cluster: Ombrófila (0-6500ps)

I21V_InhA_NADH	RESÍDUOS	ÁTOMOS
InhA	1-268	1-4005
NADH	269	4006-4076
6 Na+	270-275	4077-4082
TIP3P (9132 molec.)	276-9407	4083-31478
Total	9407 resíduos	31478 átomos

I16T InhA NADH:

- Coord. Iniciais: INHA_NADH 1900ps (mdcp_1900ps.cav)
- Tempo total de simulação: 3500ps
- Cluster: Ombrófila (0-3500ps)

I16T_InhA_NADH	RESÍDUOS	ÁTOMOS
InhA	1-268	1-4003
NADH	269	4004-4074
6 Na+	270-275	4075-4080
TIP3P (9132 molec.)	276-9407	4081-31476
Total	9407 resíduos	31476 átomos

I47T InhA NADH:

- Coord. Iniciais: INHA_NADH 1900ps (mdcp_1900ps.cav)
- Tempo total de simulação: 4000ps
- Cluster: Ombrófila (0-4000ps)

I47T_InhA_NADH	RESÍDUOS	ÁTOMOS
InhA	1-268	1-4003
NADH	269	4004-4074
6 Na+	270-275	4075-4080
TIP3P (9132 molec.)	276-9407	4081-31476
Total	9407 resíduos	31476 átomos

Apêndice 4

Construção dos mutantes S94A, I16T, I21V e I47T da enzima InhA, a partir das coordenadas do sistema *wt* InhA-NADH em 1,9 ns de simulação:

Mutante S94A:

```
----- Ser94 -----
ATOM 1419 N SER 93 2.078 1.883 0.039 0.00 0.00
ATOM 1420 H SER 93 2.681 2.572 0.465 0.00 0.00
ATOM 1421 CA SER 93 0.780 1.591 0.628 0.00 0.00
ATOM 1422 HA SER 93 0.159 1.007 -0.052 0.00 0.00
ATOM 1423 CB SER 93 1.062 0.713 1.916 0.00 0.00
ATOM 1424 HB2 SER 93 1.643 1.293 2.633 0.00 0.00
ATOM 1425 HB3 SER 93 0.121 0.529 2.434 0.00 0.00
ATOM 1426 OG SER 93 1.733 -0.457 1.635 0.00 0.00
ATOM 1427 HG SER 93 1.963 -0.508 0.704 0.00 0.00
ATOM 1428 C SER 93 0.084 2.928 0.964 0.00 0.00
ATOM 1429 O SER 93 -0.624 2.942 1.975 0.00 0.00
```

```
----- Ser94Ala -----
(OG -> HB3) / (HG -> deletado)
ATOM 1419 N ALA 93 2.078 1.883 0.039 0.00 0.00
ATOM 1420 H ALA 93 2.681 2.572 0.465 0.00 0.00
ATOM 1421 CA ALA 93 0.780 1.591 0.628 0.00 0.00
ATOM 1422 HA ALA 93 0.159 1.007 -0.052 0.00 0.00
ATOM 1423 CB ALA 93 1.062 0.713 1.916 0.00 0.00
ATOM 1424 HB1 ALA 93 1.643 1.293 2.633 0.00 0.00
ATOM 1425 HB2 ALA 93 0.121 0.529 2.434 0.00 0.00
ATOM 1426 HB3 ALA 93 1.733 -0.457 1.635 0.00 0.00
ATOM 1428 C ALA 93 0.084 2.928 0.964 0.00 0.00
ATOM 1429 O ALA 93 -0.624 2.942 1.975 0.00 0.00
```

Mutante I16T:

```
----- Ile16 -----
ATOM 214 N ILE 15 0.835 6.217 9.237 0.00 0.00
ATOM 215 H ILE 15 0.103 5.526 9.156 0.00 0.00
ATOM 216 CA ILE 15 0.503 7.368 10.089 0.00 0.00
ATOM 217 HA ILE 15 1.468 7.719 10.453 0.00 0.00
ATOM 218 CB ILE 15 -0.182 8.432 9.128 0.00 0.00
ATOM 219 HB ILE 15 0.331 8.615 8.184 0.00 0.00
ATOM 220 CG2 ILE 15 -1.592 8.076 8.760 0.00 0.00
ATOM 221 1HG2 ILE 15 -2.153 7.985 9.690 0.00 0.00
ATOM 222 2HG2 ILE 15 -1.993 8.907 8.181 0.00 0.00
ATOM 223 3HG2 ILE 15 -1.638 7.134 8.214 0.00 0.00
ATOM 224 CG1 ILE 15 -0.243 9.797 9.750 0.00 0.00
ATOM 225 2HG1 ILE 15 -0.852 10.516 9.202 0.00 0.00
ATOM 226 3HG1 ILE 15 -0.646 9.765 10.762 0.00 0.00
ATOM 227 CD1 ILE 15 1.203 10.386 9.738 0.00 0.00
ATOM 228 1HD1 ILE 15 1.279 11.284 10.351 0.00 0.00
ATOM 229 2HD1 ILE 15 1.815 9.620 10.215 0.00 0.00
ATOM 230 3HD1 ILE 15 1.538 10.590 8.721 0.00 0.00
ATOM 231 C ILE 15 -0.206 7.026 11.436 0.00 0.00
ATOM 232 O ILE 15 -0.030 7.732 12.411 0.00 0.00
```

```
----- Ile16Thr -----
(CG1 -> OG1) / (2HG1 -> HG1) / (3HG1,CD1,1HD1,2HD1,3HD1 -> deletados)
ATOM 214 N THR 15 0.835 6.217 9.237 0.00 0.00
ATOM 215 H THR 15 0.103 5.526 9.156 0.00 0.00
ATOM 216 CA THR 15 0.503 7.368 10.089 0.00 0.00
ATOM 217 HA THR 15 1.468 7.719 10.453 0.00 0.00
ATOM 218 CB THR 15 -0.182 8.432 9.128 0.00 0.00
ATOM 219 HB THR 15 0.331 8.615 8.184 0.00 0.00
ATOM 220 CG2 THR 15 -1.592 8.076 8.760 0.00 0.00
ATOM 221 1HG2 THR 15 -2.153 7.985 9.690 0.00 0.00
ATOM 222 2HG2 THR 15 -1.993 8.907 8.181 0.00 0.00
ATOM 223 3HG2 THR 15 -1.638 7.134 8.214 0.00 0.00
ATOM 224 OG1 THR 15 -0.243 9.797 9.750 0.00 0.00
ATOM 225 HG1 THR 15 -0.852 10.516 9.202 0.00 0.00
ATOM 231 C THR 15 -0.206 7.026 11.436 0.00 0.00
ATOM 232 O THR 15 -0.030 7.732 12.411 0.00 0.00
```


Mutante I21V:

```

----- Ile21 -----
ATOM 281 N ILE 20 -1.908 0.311 6.858 0.00 0.00
ATOM 282 H ILE 20 -2.545 0.912 6.354 0.00 0.00
ATOM 283 CA ILE 20 -1.722 -1.077 6.465 0.00 0.00
ATOM 284 HA ILE 20 -2.364 -1.654 7.131 0.00 0.00
ATOM 285 CB ILE 20 -2.451 -1.408 5.160 0.00 0.00
ATOM 286 HB ILE 20 -3.457 -0.993 5.210 0.00 0.00
ATOM 287 CG2 ILE 20 -1.680 -0.816 3.979 0.00 0.00
ATOM 288 1HG2 ILE 20 -0.928 -1.588 3.811 0.00 0.00
ATOM 289 2HG2 ILE 20 -2.306 -0.658 3.101 0.00 0.00
ATOM 290 3HG2 ILE 20 -1.167 0.124 4.184 0.00 0.00
ATOM 291 CG1 ILE 20 -2.685 -2.912 5.019 0.00 0.00
ATOM 292 2HG1 ILE 20 -1.732 -3.435 4.942 0.00 0.00
ATOM 293 3HG1 ILE 20 -3.230 -3.376 5.841 0.00 0.00
ATOM 294 CD1 ILE 20 -3.421 -3.344 3.773 0.00 0.00
ATOM 295 1HD1 ILE 20 -2.715 -3.307 2.944 0.00 0.00
ATOM 296 2HD1 ILE 20 -3.665 -4.403 3.855 0.00 0.00
ATOM 297 3HD1 ILE 20 -4.329 -2.783 3.553 0.00 0.00
ATOM 298 C ILE 20 -0.262 -1.570 6.338 0.00 0.00
ATOM 299 O ILE 20 0.058 -2.716 6.513 0.00 0.00

```

```

----- Ile21Val -----
(2HG1 -> 1HG1) / (3HG1 -> 2HG1) / (CD1 -> 3HG1) / (1HD1,1HD1,1HD1 -> deletados)

```

```

ATOM 281 N VAL 20 -1.908 0.311 6.858 0.00 0.00
ATOM 282 H VAL 20 -2.545 0.912 6.354 0.00 0.00
ATOM 283 CA VAL 20 -1.722 -1.077 6.465 0.00 0.00
ATOM 284 HA VAL 20 -2.364 -1.654 7.131 0.00 0.00
ATOM 285 CB VAL 20 -2.451 -1.408 5.160 0.00 0.00
ATOM 286 HB VAL 20 -3.457 -0.993 5.210 0.00 0.00
ATOM 287 CG2 VAL 20 -1.680 -0.816 3.979 0.00 0.00
ATOM 288 1HG2 VAL 20 -0.928 -1.588 3.811 0.00 0.00
ATOM 289 2HG2 VAL 20 -2.306 -0.658 3.101 0.00 0.00
ATOM 290 3HG2 VAL 20 -1.167 0.124 4.184 0.00 0.00
ATOM 291 CG1 VAL 20 -2.685 -2.912 5.019 0.00 0.00
ATOM 292 1HG1 VAL 20 -1.732 -3.435 4.942 0.00 0.00
ATOM 293 2HG1 VAL 20 -3.230 -3.376 5.841 0.00 0.00
ATOM 294 3HG1 VAL 20 -3.421 -3.344 3.773 0.00 0.00
ATOM 298 C VAL 20 -0.262 -1.570 6.338 0.00 0.00
ATOM 299 O VAL 20 0.058 -2.716 6.513 0.00 0.00

```

Mutante I47T:

```

----- Ile47 -----
ATOM 702 N ILE 46 5.333 8.819 13.803 0.00 0.00
ATOM 703 H ILE 46 5.599 9.694 13.374 0.00 0.00
ATOM 704 CA ILE 46 5.828 7.592 13.142 0.00 0.00
ATOM 705 HA ILE 46 5.049 6.834 13.226 0.00 0.00
ATOM 706 CB ILE 46 6.072 7.847 11.630 0.00 0.00
ATOM 707 HB ILE 46 6.761 8.692 11.626 0.00 0.00
ATOM 708 CG2 ILE 46 6.825 6.668 10.851 0.00 0.00
ATOM 709 1HG2 ILE 46 7.659 6.268 11.428 0.00 0.00
ATOM 710 2HG2 ILE 46 6.124 5.849 10.689 0.00 0.00
ATOM 711 3HG2 ILE 46 7.284 7.056 9.942 0.00 0.00
ATOM 712 CG1 ILE 46 4.797 8.182 10.866 0.00 0.00
ATOM 713 2HG1 ILE 46 4.229 7.284 10.624 0.00 0.00
ATOM 714 3HG1 ILE 46 4.149 8.837 11.449 0.00 0.00
ATOM 715 CD1 ILE 46 5.089 8.957 9.540 0.00 0.00
ATOM 716 1HD1 ILE 46 5.904 9.680 9.568 0.00 0.00
ATOM 717 2HD1 ILE 46 5.347 8.187 8.812 0.00 0.00
ATOM 718 3HD1 ILE 46 4.213 9.468 9.141 0.00 0.00
ATOM 719 C ILE 46 7.065 7.058 13.843 0.00 0.00
ATOM 720 O ILE 46 7.247 5.853 13.865 0.00 0.00

```

```

----- Ile47Thr -----
(CG1 -> OG1) / (2HG1 -> HG1) / (3HG1,CD1,1HD1,2HD1,3HD1 -> deletados)

```

```

ATOM 702 N I47T 46 5.333 8.819 13.803 0.00 0.00
ATOM 703 H I47T 46 5.599 9.694 13.374 0.00 0.00
ATOM 704 CA I47T 46 5.828 7.592 13.142 0.00 0.00
ATOM 705 HA I47T 46 5.049 6.834 13.226 0.00 0.00
ATOM 706 CB I47T 46 6.072 7.847 11.630 0.00 0.00
ATOM 707 HB I47T 46 6.761 8.692 11.626 0.00 0.00
ATOM 708 CG2 I47T 46 6.825 6.668 10.851 0.00 0.00
ATOM 709 1HG2 I47T 46 7.659 6.268 11.428 0.00 0.00
ATOM 710 2HG2 I47T 46 6.124 5.849 10.689 0.00 0.00
ATOM 711 3HG2 I47T 46 7.284 7.056 9.942 0.00 0.00
ATOM 712 OG1 I47T 46 4.797 8.182 10.866 0.00 0.00
ATOM 713 HG1 I47T 46 4.229 7.284 10.624 0.00 0.00
ATOM 719 C I47T 46 7.065 7.058 13.843 0.00 0.00
ATOM 720 O I47T 46 7.247 5.853 13.865 0.00 0.00

```

Apêndice 5

Protocolos de *docking* molecular com o programa AutoDock 3.0**Docking com *simulated annealing* (SA): arquivo de entrada (SA.dpf)**

```

seed      88664422 # for random number generator
types    CANOPH      # atom type names
fld      ENYmacro.maps.fld      # grid data file
map      ENYmacro.C.map      # C-atomic affinity map file
map      ENYmacro.A.map      # A-atomic affinity map file
map      ENYmacro.N.map      # N-atomic affinity map file
map      ENYmacro.O.map      # O-atomic affinity map file
map      ENYmacro.P.map      # P-atomic affinity map file
map      ENYmacro.H.map      # H-atomic affinity map file
map      ENYmacro.e.map      # electrostatics map file

move     NAH_moved.pdbq      # small molecule file
about    -3.281 32.468 13.357      # small molecule center

# Initial Translation, Quaternion and Torsions
tran0    random      # initial coordinates/A or "random"
quat0    random      # initial quaternion or "random"
ndihe    0           # number of initial torsions
dihe0    random      # initial torsions

torsdof  0 0.3113 # num. non-Hydrogen torsional DOF & coeff.

# Initial Translation, Quaternion and Torsion Step Sizes and Reduction Factors
tstep    0.2         # translation step/A
qstep    5.0         # quaternion step/deg
dstep    5.0         # torsion step/deg
trnrf    1.          # trans reduction factor/per cycle
quarf    1.          # quat reduction factor/per cycle
dihrf    1.          # tors reduction factor/per cycle

# Hard Torsion Constraints
#hardtorcon 1 -180. 30. # constrain torsion, num., angle(deg), range(deg)

# Internal Non-Bonded Parameters
intnbp_r_eps 4.00 0.0222750 12 6      #C-C lj
intnbp_r_eps 4.00 0.0222750 12 6      #C-A lj
intnbp_r_eps 3.75 0.0230026 12 6      #C-N lj
intnbp_r_eps 3.60 0.0257202 12 6      #C-O lj
intnbp_r_eps 4.00 0.0222750 12 6      #C-C lj
intnbp_r_eps 3.00 0.0081378 12 6      #C-H lj
intnbp_r_eps 4.00 0.0222750 12 6      #A-A lj
intnbp_r_eps 3.75 0.0230026 12 6      #A-N lj
intnbp_r_eps 3.60 0.0257202 12 6      #A-O lj
intnbp_r_eps 4.00 0.0222750 12 6      #A-C lj
intnbp_r_eps 3.00 0.0081378 12 6      #A-H lj
intnbp_r_eps 3.50 0.0237600 12 6      #N-N lj
intnbp_r_eps 3.35 0.0265667 12 6      #N-O lj
intnbp_r_eps 3.75 0.0230026 12 6      #N-C lj
intnbp_r_eps 2.75 0.0084051 12 6      #N-H lj
intnbp_r_eps 3.20 0.0297000 12 6      #O-O lj
intnbp_r_eps 3.60 0.0257202 12 6      #O-C lj
intnbp_r_eps 1.90 0.3280000 12 10      #O-H hb
intnbp_r_eps 4.00 0.0222750 12 6      #C-C lj
intnbp_r_eps 3.00 0.0081378 12 6      #C-H lj
intnbp_r_eps 2.00 0.0029700 12 6      #H-H lj

# Simulated Annealing Parameters
rt0 616.      # SA: initial RT
rtrf 0.95     # SA: RT reduction factor/per cycle
#linear_schedule # SA: do not use geometric cooling
runs 10       # SA: number of runs
cycles 50     # SA: cycles
accs 25000    # SA: steps accepted
rejs 25000    # SA: steps rejected
select m      # SA: minimum or last

# Trajectory Parameters (Simulated Annealing Only)
trjfrq 100    # trajectory frequency
trjbeg 1      # start trj output at cycle
trjend 50     # end trj output at cycle
trjout NAHcristal.trj # trajectory file
trjsel A      # A=acc only;E=either acc or rej

```

```
#watch NAH_moved.watch.pdb      # real-time monitoring file

outlev 1                          # diagnostic output level

# Docked Conformation Clustering Parameters for "analysis" command
rmstol 1.0                        # cluster tolerance (Angstroms)
rmsref NAH_moved.pdbq           # reference structure file for RMS calc.
#rmsnosym                        # do no symmetry checking in RMS calc.
write_all                        # write all conformations in a cluster

extnrg 1000.                     # external grid energy
e0max 0.10000 # max. allowable initial energy, max. num. retries

# Perform Dockings
simanneal                        # do as many SA runs as set by the "runs" command above

# Perform Cluster Analysis
analysis                          # do cluster analysis on results
```

Docking com algoritmos genéticos Lamarkianos (LGA): arquivo de entrada (LGA.dpf)

```
seed time pid # for random number generator
types CANOPH # atom type names
fld ENYmacro.maps.fld # grid data file
map ENYmacro.C.map # C-atomic affinity map file
map ENYmacro.A.map # A-atomic affinity map file
map ENYmacro.N.map # N-atomic affinity map file
map ENYmacro.O.map # O-atomic affinity map file
map ENYmacro.P.map # P-atomic affinity map file
map ENYmacro.H.map # H-atomic affinity map file
map ENYmacro.e.map # electrostatics map file

move NAH_moved.pdbq # small molecule file
about -3.281 32.468 13.357 # small molecule center

# Initial Translation, Quaternion and Torsions
tran0 random # initial coordinates/A or "random"
quat0 random # initial quaternion or "random"
ndihe 16 # number of initial torsions
dihe0 random # initial torsions

torsdof 16 0.3113 # num. non-Hydrogen torsional DOF & coeff.

# Initial Translation, Quaternion and Torsion Step Sizes and Reduction Factors
tstep 0.2 # translation step/A
qstep 5.0 # quaternion step/deg
dstep 5.0 # torsion step/deg
trnrf 1. # trans reduction factor/per cycle
quarf 1. # quat reduction factor/per cycle
dihrf 1. # tors reduction factor/per cycle

# Hard Torsion Constraints
#hardtcon 1 -180. 30. # constrain torsion, num., angle(deg), range(deg)

# Internal Non-Bonded Parameters
intnbp_r_eps 4.00 0.0222750 12 6 #C-C lj
intnbp_r_eps 4.00 0.0222750 12 6 #C-A lj
intnbp_r_eps 3.75 0.0230026 12 6 #C-N lj
intnbp_r_eps 3.60 0.0257202 12 6 #C-O lj
intnbp_r_eps 4.00 0.0222750 12 6 #C-C lj
intnbp_r_eps 3.00 0.0081378 12 6 #C-H lj
intnbp_r_eps 4.00 0.0222750 12 6 #A-A lj
intnbp_r_eps 3.75 0.0230026 12 6 #A-N lj
intnbp_r_eps 3.60 0.0257202 12 6 #A-O lj
intnbp_r_eps 4.00 0.0222750 12 6 #A-C lj
intnbp_r_eps 3.00 0.0081378 12 6 #A-H lj
intnbp_r_eps 3.50 0.0237600 12 6 #N-N lj
intnbp_r_eps 3.35 0.0265667 12 6 #N-O lj
intnbp_r_eps 3.75 0.0230026 12 6 #N-C lj
intnbp_r_eps 2.75 0.0084051 12 6 #N-H lj
intnbp_r_eps 3.20 0.0297000 12 6 #O-O lj
intnbp_r_eps 3.60 0.0257202 12 6 #O-C lj
intnbp_r_eps 1.90 0.3280000 12 10 #O-H hb
intnbp_r_eps 4.00 0.0222750 12 6 #C-C lj
intnbp_r_eps 3.00 0.0081378 12 6 #C-H lj
intnbp_r_eps 2.00 0.0029700 12 6 #H-H lj

outlev 1                          # diagnostic output level
```

```
# Docked Conformation Clustering Parameters for "analysis" command
rmstol 1.0 # cluster tolerance (Angstroms)
rmsref NAHcristal.pdbq # reference structure file for RMS calc.
#rmsnosym # do no symmetry checking in RMS calc.
write_all # write all conformations in a cluster

extnrg 1000. # external grid energy
e0max 0. 10000 # max. allowable initial energy, max. num. retries

# Genetic Algorithm (GA) and Lamarckian Genetic Algorithm Parameters (LGA)
ga_pop_size 50 # number of individuals in population
ga_num_evals 250000 # maximum number of energy evaluations
ga_num_generations 27000 # maximum number of generations
ga_elitism 1 # num. of top individuals that automatically survive
ga_mutation_rate 0.02 # rate of gene mutation
ga_crossover_rate 0.80 # rate of crossover
ga_window_size 10 # num. of generations for picking worst individual
ga_cauchy_alpha 0 # ~mean of Cauchy distribution for gene mutation
ga_cauchy_beta 1 # ~variance of Cauchy distribution for gene mutation
set_ga # set the above parameters for GA or LGA

# Local Search (Solis & Wets) Parameters (for LS alone and for LGA)
sw_max_its 300 # number of iterations of Solis & Wets local search
sw_max_succ 4 # number of consecutive successes before changing rho
sw_max_fail 4 # number of consecutive failures before changing rho
sw_rho 1.0 # size of local search space to sample
sw_lb_rho 0.01 # lower bound on rho
ls_search_freq 0.06 # probability of performing local search on an indiv.
set_pswl # set the above pseudo-Solis & Wets parameters

# Perform Dockings
#do_local_only 50 # do only local search

ga_run 10 # do this many GA or LGA runs

# Perform Cluster Analysis
analysis # do cluster analysis on results
```

Apêndice 6

Simulação por dinâmica molecular do NADH livre em solução:

A preparação do sistema para simulação da molécula de NADH livre em solução foi realizada com os programas Prep, Link, Edit e Parm do pacote AMBER 4.1. As coordenadas e conformação inicial do NADH foram obtidas da estrutura 1ENY, e o conjunto com os parâmetros do campo de força foram gerados pelo Prof. Dr. Ulf Ryde (Apêndice 1).

Todos os átomos do sistema foram tratados explicitamente (modelo atômico), e o NADH foi considerado com carga – 2 (mesma carga das simulações dos complexos com a InhA), por isso o sistema simulado foi neutralizado pela adição de 2 íons de sódio (Na^+) próximos à porção pirofosfato (5Å de distância).

O NADH foi solvatado em um sistema cúbico, adicionando-se caixas de 216 moléculas de água (TIP3P) equilibradas pelo método de Monte Carlo, desprezando-se aquelas moléculas muito próximas aos átomos da molécula (distância de até 2,5 Å entre átomos pesados), e também aquelas moléculas de água a mais de 10 Å, resultando em uma caixa retangular de dimensões iniciais 39,5 Å x 34,5 Å x 27,1 Å (Figura Ap-6.1) com a seguinte composição (Tabela Ap-6.1):

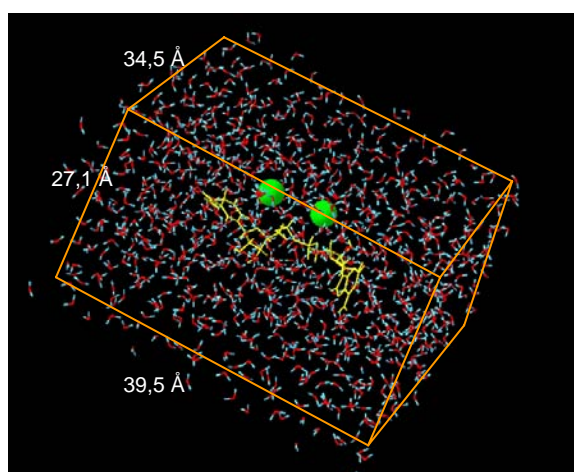


Figura Ap-6.1. Definição do sistema de simulação: NADH (amarelo) e contra-íons Na^+ (superfície molecular em verde) solvatados por moléculas de água (TIP3P - ciano e vermelho) em uma caixa retangular de dimensões 39,5 Å x 34,5 Å x 27,1 Å. Este sistema foi replicado periodicamente em todas as direções. Figura gerada com SPDBViewer v3.7b2.

Tabela Ap-6.1. Composição do sistema de simulação do NADH livre em solução (NADH_WATbox):

NADH (Ulf charged / AMBER compatible):

- Coord. Iniciais: NADH da 1ENY.pdb
- Tempo total de simulação: 7000ps
- Cluster: Ombrófila (0-7000ps)

NADH_WATbox	RESÍDUOS	ÁTOMOS
NADH	1	1-71
2 Na+	2-3	72-73
TIP3P (999 wat)	4-1002	74-3070
Total	1002 resíduos	3070 átomos

A simulação foi realizada sob condições periódicas de contorno, num *ensemble* isobárico-isotérmico (NPT), numa temperatura de 298,16K (25 °C) e pressão de 1 atm, com o programa SANDER do pacote AMBER 6.0 (Case *et al.*, 1999).

A preparação do sistema de simulação, assim como as análises da trajetória foram feitas em um computador Dell Pentium IV, de 1 GHz com 526 MB de memória RAM e 80 GB de disco rígido. A simulação foi rodada com o pacote AMBER 6.0 no cluster OMBRÓFILA (cluster com 16 máquinas HP e-pc, cada uma com um processador Pentium III de 1GHz, 256MB de memória principal, e HD de 20GB).

A relaxação das interações desfavoráveis entre as moléculas de solvente e entre estas e o soluto foi realizada por 100 passos de minimização de energia (com o algoritmo *steepest descent*), seguido de 20 ps de dinâmica molecular (com passos de 2 fs), para a elevação da temperatura do solvente de 10 a 298,16 K e equilibração da densidade da caixa nesta última temperatura. Ambos procedimentos foram realizados com restrição (*constrains*) dos movimentos da molécula do NADH, utilizando a opção BELLY do SANDER. Os contra-íons foram tratados como parte do solvente.

O NADH foi então relaxado, com etapas sucessivas de minimização de energia de 500 passos (100 primeiros passos com *steepest descent* e os 400 passos seguintes com o método gradiente conjugado), onde seus movimentos foram gradualmente liberados (pela diminuição da constante de força de restrição), até que fossem completamente liberados¹.

A seguir, o aquecimento gradual do sistema até 298,16K foi realizado em pequenas dinâmicas de 2 ps cada, totalizando 12 ps de aquecimento (10→50 K, 50→100 K, 100→150 K, 150→200 K, 200→250 K, 250→298,16 K). As velocidades iniciais do sistema foram atribuídas de acordo com a distribuição de Maxwell-Boltzman a 10 K. A cada etapa de aquecimento as velocidades foram reatribuídas na nova temperatura, e os valores das dimensões da caixa de simulação foram atualizados².

Depois de alcançada a temperatura desejada, a simulação do sistema foi realizada à temperatura e pressão constantes de 298,16 K e 1 atm pelo acoplamento do sistema a um banho térmico (Berendsen *et al.*, 1984) e regulação da pressão por reescalamento isotrópico (em relação a cada átomo individualmente). O algoritmo SHAKE (Ryckaert *et al.*, 1977), com uma tolerância de 10^{-5} Å, foi utilizado para a correção das ligações contendo átomos de hidrogênio, permitindo assim o uso de um passo de 2 fs para a integração das equações do movimento. Nenhuma outra medida restritiva foi aplicada ao sistema após a fase de equilibração. As interações eletrostáticas entre átomos não-ligados foram avaliadas pelo método PME, que utilizou uma malha de aproximadamente 1 Å para a interpolação do somatório no espaço direto. As interações de Lennard-Jones foram avaliadas utilizando uma distância máxima de 9 Å³.

Os dados da simulação por dinâmica molecular do sistema NADH_WATbox foram coletados por 7,0 ns, e as estruturas para análise foram salvas a cada 0,5 ps.

Arquivos de entrada para as simulações da dinâmica molecular:

mini_00.in:

```
# Belly Minimization.NTF=NTC=2
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=1, ntr=0, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=100, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
        a=40, b=35, c=28,
&end
Belly Dynamics.Only solvent (WAT + ClO) Atoms move.
RES 2 1002
END
END
```

mdcp_00.in:

```
# 20 ps NPT. NTF+NTC=2. Equilibrating Solvent (WAT + ClO) at 298.16K.NPSCAL=1
```

¹ Arquivos de entrada para a minimização do soluto: mini_01.in a mini_07.in.

² Arquivos de entrada para a termalização/equilibração do soluto: mdcp_01.in a mdcp_06.in.

³ Arquivo de entrada para a simulação da DM: mdcp_prod_02.in.

```

&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
        ntxo=1, ntp=50,
        ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=1, ntr=0, cut=9.0,
        nstlim=10000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
        temp0=298.16, tempi=10.0, ig=555333, heat=0.0001, ntt=1,
        dtemp=10.0, tautp=1.0, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=2.0, npscal=1,
        ntc=2, tol=0.00001,
&end
&ewald
        a=40, b=35, c=28,
&end
Belly Dynamics.Only solvent (WAT + CIO) Atoms move.
RES 2 1002
END
END

```

mini_01.in:

```

# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
&end
Restraining NAH Only
1000.000
RES 1
END
END

```

mini_02.in:

```

# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
&end
Restraining NAH Only
500.000
RES 1
END
END

```

mini_03.in:

```

# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
&end
&ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
&end
Restraining NAH Only
250.000
RES 1
END
END

```

mini_04.in:

```

# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,

```

```

        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
    &end
    &ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
    &end
    Restraining NAH Only
    100.000
RES 1
END
END

```

mini_05.in:

```

# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
    &end
    &ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
    &end
    Restraining NAH Only
    50.000
RES 1
END
END

```

mini_06.in:

```

# Restrained Minimization.NTF=NTC=2.NTR=1.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=1, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
    &end
    &ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
    &end
    Restraining NAH Only
    10.000
RES 1
END
END

```

mini_07.in:

```

# Unrestrained Minimization.NTF=NTC=2.
&cntrl  imin=1,
        ntp=10,
        ntf=2,dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        ntc=2, tol=0.00001,
        maxcyc=500, ncyc=100, ntmin=1, dx0=0.01, dxm=0.5, drms=0.0001,
    &end
    &ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
    &end

```

mdcp_01.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 10K to 298.16K.NPSCAL=1
&cntrl  timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
        ntso=1, ntp=50,
        ntw=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
        ioutfm=0, ntwprt=0,
        ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
        ibelly=0, ntr=0, cut=9.0,
        nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
        temp0=50.0, tempi=10.0, ig=555333, heat=0.0001, ntt=1,
        dtemp=10.0, tautp=1.0, vlimit=10.0,
        ntp=1, pres0=1.0, comp=44.6, taup=2.0, npscal=1,
        ntc=2, tol=0.00001,
    &end
    &ewald
        a=36.8294450, b=32.2257643, c=25.7806115,
    &end

```


mdcp_02.in:

```
# 2ps NPT. NTF+NTC=2. Equilibrating the system from 50K to 100K. NPSCAL=1
&cntrl timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
      ntso=1, ntpr=50,
      ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
      ioutfm=0, ntwprt=0,
      ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=0, cut=9.0,
      nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
      temp0=100.0, tempi=50.0, ig=555333, heat=0.0001, ntt=1,
      dtemp=10.0, tautp=0.2, vlimit=10.0,
      ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
      ntc=2, tol=0.00001,
&end
&ewald
      a=36.1654601, b=31.6447776, c=25.3158221,
&end
```

mdcp_03.in:

```
# 2ps NPT. NTF+NTC=2. Equilibrating the system from 100K to 150K. NPSCAL=1
&cntrl timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
      ntso=1, ntpr=50,
      ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
      ioutfm=0, ntwprt=0,
      ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=0, cut=9.0,
      nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
      temp0=150.0, tempi=100.0, ig=555333, heat=0.0001, ntt=1,
      dtemp=10.0, tautp=0.2, vlimit=10.0,
      ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
      ntc=2, tol=0.00001,
&end
&ewald
      a=36.0077570, b=31.5067874, c=25.2054299,
&end
```

mdcp_04.in:

```
# 2ps NPT. NTF+NTC=2. Equilibrating the system from 150K to 200K. NPSCAL=1
&cntrl timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
      ntso=1, ntpr=50,
      ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
      ioutfm=0, ntwprt=0,
      ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=0, cut=9.0,
      nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
      temp0=200.0, tempi=150.0, ig=555333, heat=0.000001, ntt=1,
      dtemp=10.0, tautp=0.2, vlimit=10.0,
      ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
      ntc=2, tol=0.00001,
&end
&ewald
      a=36.131653, b=31.6151967, c=25.2921574,
&end
```

mdcp_05.in:

```
# 2ps NPT. NTF+NTC=2. Equilibrating the system from 200K to 250K. NPSCAL=1
&cntrl timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
      ntso=1, ntpr=50,
      ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
      ioutfm=0, ntwprt=0,
      ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
      ibelly=0, ntr=0, cut=9.0,
      nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
      temp0=250.0, tempi=200.0, ig=555333, heat=0.000001, ntt=1,
      dtemp=10.0, tautp=0.2, vlimit=10.0,
      ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
      ntc=2, tol=0.00001,
&end
&ewald
      a=36.3099474, b=31.7712043, c=25.4169635,
&end
```

mdcp_06.in:

```
# 2ps NPT. NTF+NTC=2. Equilibrating the system from 200K to 250K. NPSCAL=1
&cntrl timlim=99999900, imin=0, ntx=1, irect=0, ntrx=1,
      ntso=1, ntpr=50,
      ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
      ioutfm=0, ntwprt=0,
      ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
```

```

ibelly=0, ntr=0, cut=9.0,
nstlim=1000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
temp0=298.16, tempi=250.0, ig=555333, heat=0.000001, ntt=1,
dtemp=10.0, tautp=0.2, vlimit=10.0,
ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
ntc=2, tol=0.00001,
&end
&ewald
a=36.4890481, b=31.9279174, c=25.5423340,
&end

```

mdcp_prod_01.in:

```

# 2ps NPT. NTF+NTC=2. Equilibrating the system from 200K to 250K.NPSCAL=1
&cntrl timlim=99999900, imin=0, ntx=7, irect=1, ntrx=1,
ntxo=1, ntpr=50,
ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
ioutfm=0, ntwprt=0,
ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
ibelly=0, ntr=0, cut=9.0,
nstlim=19000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
temp0=298.16, tempi=298.16, ig=555333, heat=0.0001, ntt=1,
dtemp=10.0, tautp=0.2, vlimit=10.0,
ntp=1, pres0=1.0, comp=44.6, taup=0.2, npscal=1,
ntc=2, tol=0.00001,
&end
&ewald
a=36.4890481, b=31.9279174, c=25.5423340,
&end

```

mdcp_prod_01.in:

```

# 50ps NPT. NTF+NTC=2.taup=0.6,taup=1.0. NPSCAL=1
&cntrl timlim=99999900, imin=0, ntx=7, irect=1, ntrx=1,
ntxo=1, ntpr=50,
ntwr=250, iwrap=0, ntwx=250, ntwv=0, ntwe=0,
ioutfm=0, ntwprt=0,
ntf=2, ntb=2, dielc=1.0, igb=0, scnb=2.0, scee=1.2,
ibelly=0, ntr=0, cut=9.0,
nstlim=25000, ndfmin=0, ntc=0, nscm=0, t=0.0, dt=0.002,
temp0=298.16, tempi=250.0, ig=555333, heat=0.0001, ntt=1,
dtemp=10.0, tautp=0.6, vlimit=10.0,
ntp=1, pres0=1.0, comp=44.6, taup=1.0, npscal=1,
ntc=2, tol=0.00001,
&end
&ewald
a=36.9812991, b=32.3586371, c=25.8869098,
&end

```

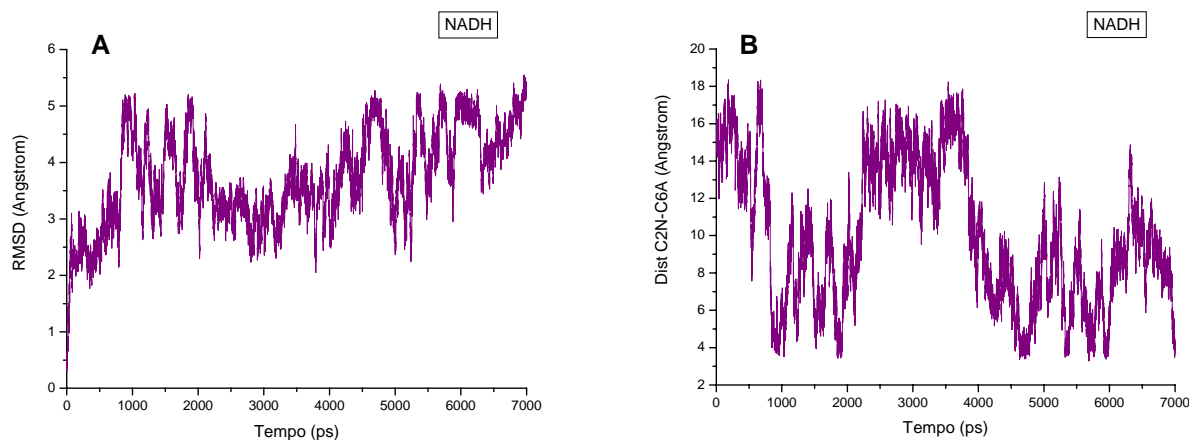
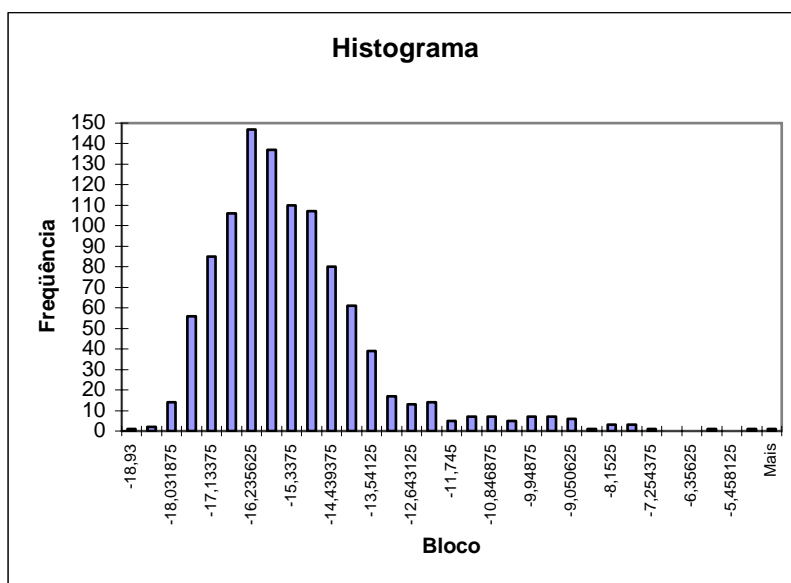
Acompanhamento da simulação da dinâmica molecular do NADH livre em solução:

Figura Ap-6.2. (A) Gráfico de RMSD do NADH em relação a sua estrutura cristalográfica na 1ENY ao longo dos 7,0 ns de trajetória da simulação NADH_WATbox. (B) Gráfico da distância entre os anéis nicotinamida e adenina (medidos como a distância entre C2N-C6A) ao longo dos 7,0 ns de trajetória da simulação NADH_WATbox.

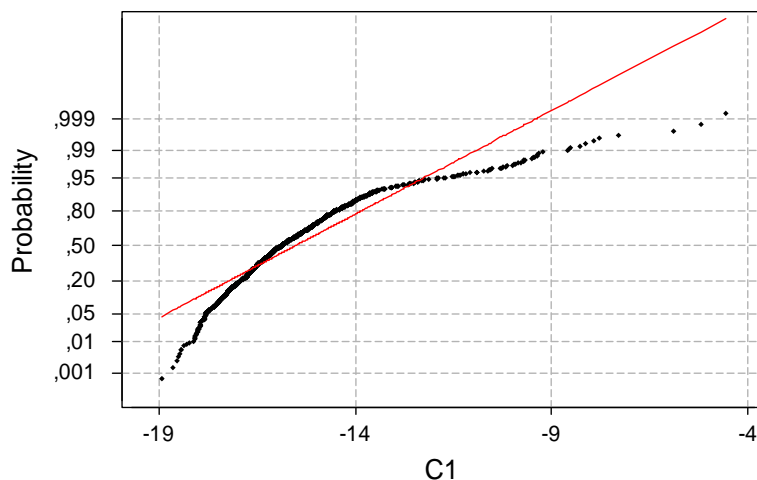
Apêndice 7

Análise estatística das distribuições das energias livres de ligação (FEB) do NADH nos sistemas *wt* InhA, S94A, I21V, I16T e I47T, obtidas com experimentos de *docking* via *simulated annealing*:

***wt* InhA:**



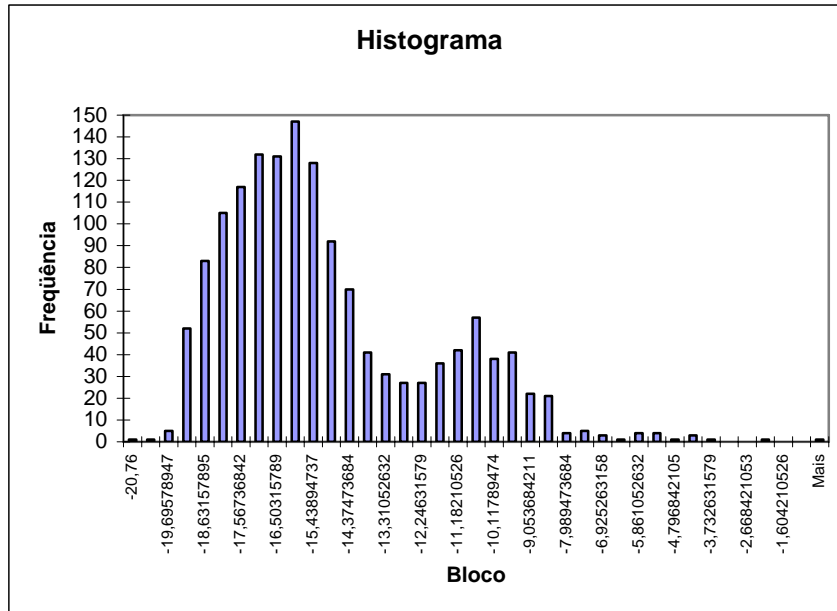
Normal Probability Plot



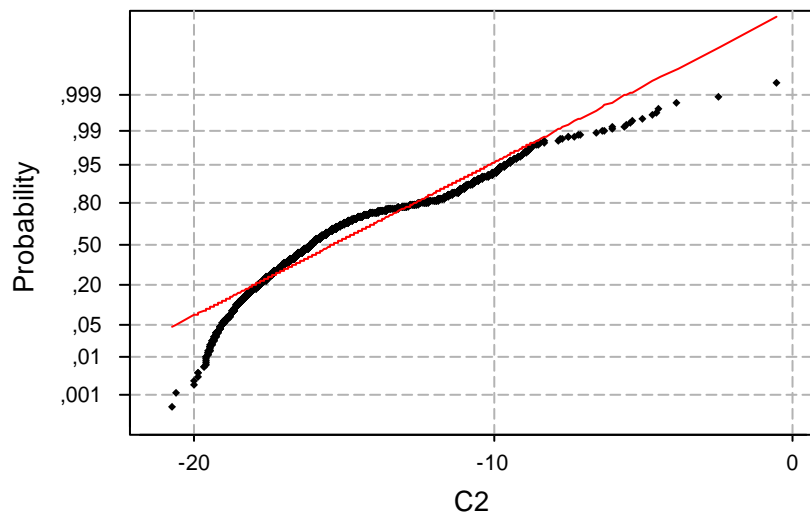
Average: -15,5330
StDev: 1,84477
N: 1044

Anderson-Darling Normality Test
A-Squared: 27,672
P-Value: 0,000

S94A-InhA:



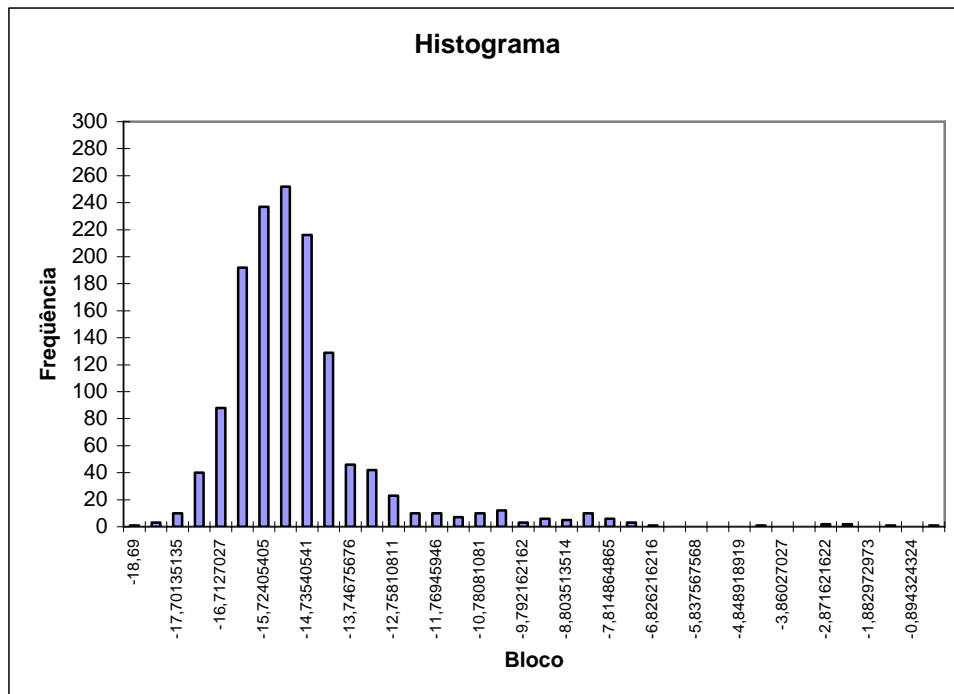
Normal Probability Plot



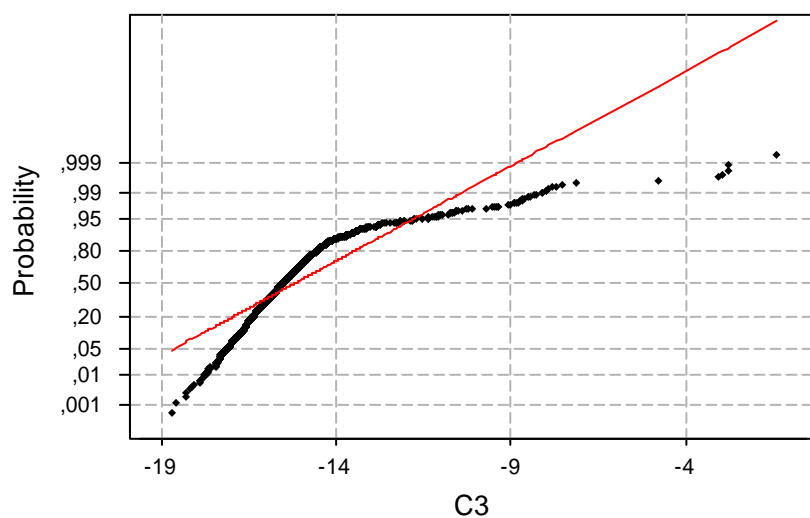
Average: -15,3351
 StDev : 3,03868
 N: 1475

Anderson-Darling Normality Test
 A-Squared: 40,063
 P-Value: 0,000

I21V-InhA:



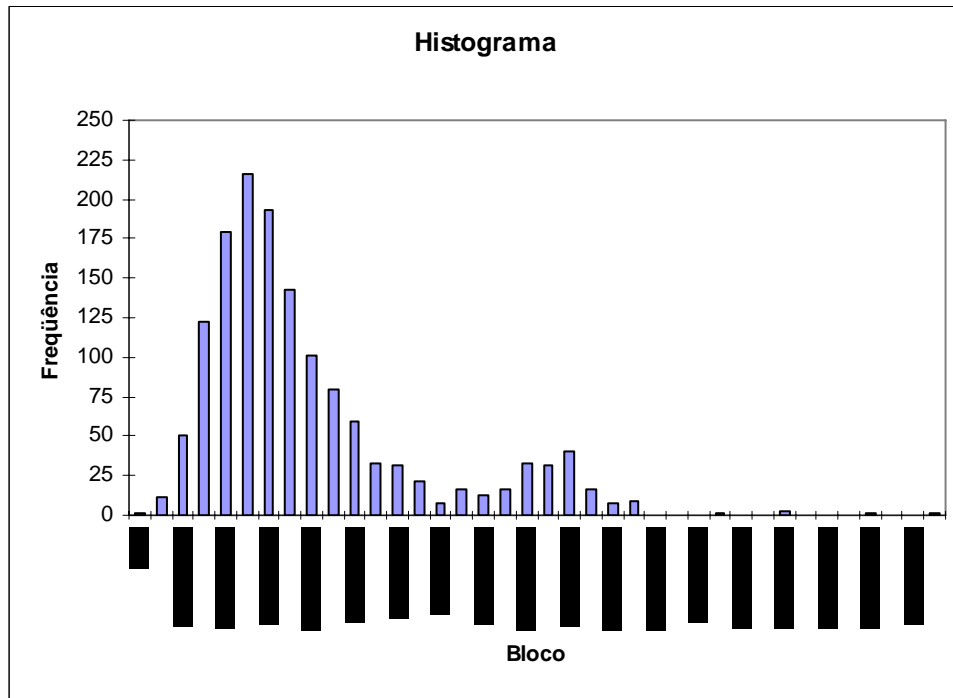
Normal Probability Plot



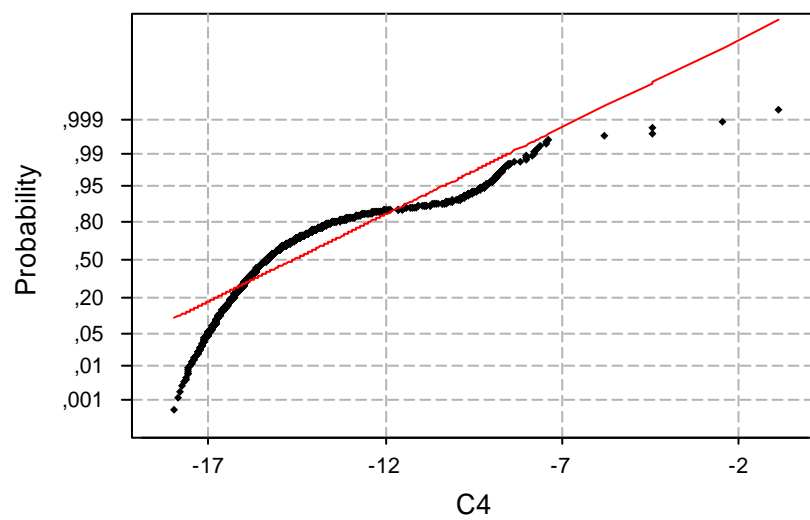
Average: -15,1882
 StDev: 1,80944
 N: 1368

Anderson-Darling Normality Test
 A-Squared: 71,207
 P-Value: 0,000

I16T-InhA:



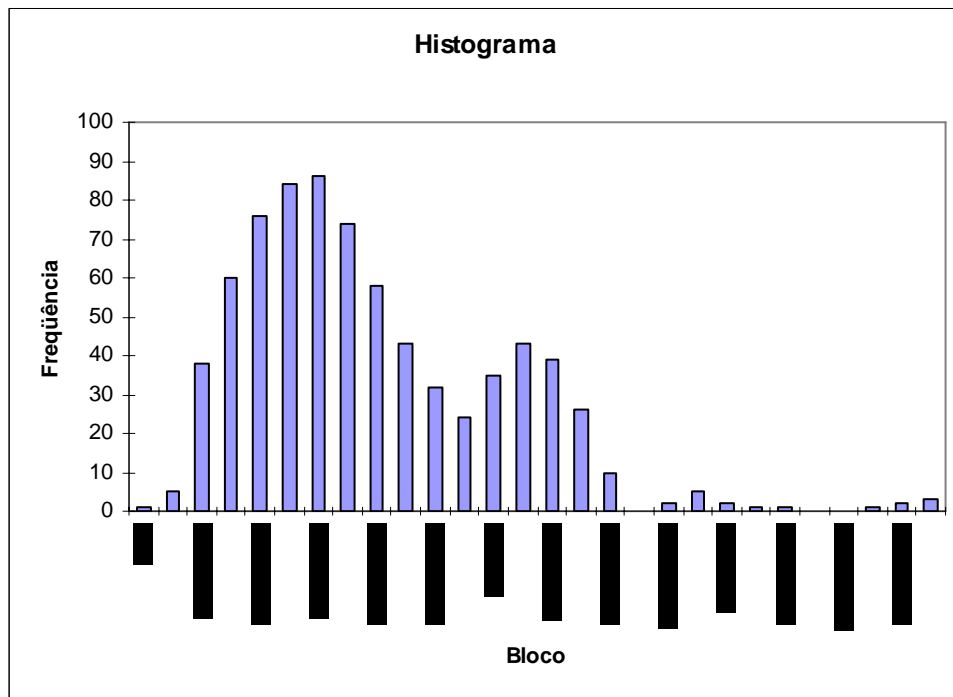
Normal Probability Plot



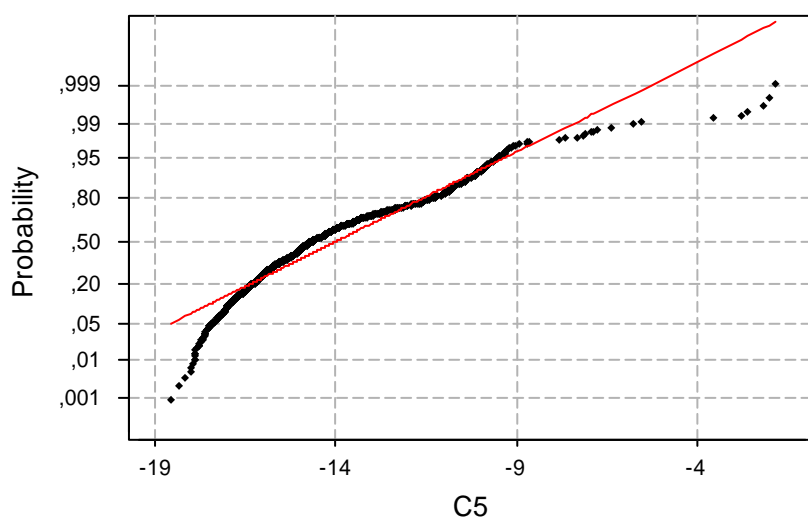
Average: -14,5820
 StDev: 2,36908
 N: 1436

Anderson-Darling Normality Test
 A-Squared: 83,030
 P-Value: 0,000

I47T-InhA:



Normal Probability Plot



Average: -13,9981
 StDev: 2,71403
 N: 751

Anderson-Darling Normality Test
 A-Squared: 13,208
 P-Value: 0,000

Análise One-Way Anova:

Anova: fator único

teste de hipóteses:

Ho: todas as médias (população) são iguais

H1: pelo menos uma das médias é diferente

RESUMO

<i>Grupo</i>	<i>Contagem</i>	<i>Soma</i>	<i>Média</i>	<i>Variância</i>
wtINHA	1044	-16216,49	-15,53304	3,403168
S94A	1475	-22619,29	-15,33511	9,233571
I21V	1368	-20777,42	-15,18817	3,274061
I16T	1436	-20939,8	-14,58203	5,612556
I47T	751	-10512,54	-13,99806	7,365961

ANOVA

<i>Fonte da variação</i>	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>valor-P</i>	<i>F crítico</i>
Entre grupos	1515,172	4	378,793	65,2837	3,82E-54	2,373397
Dentro dos grupos	35213,92	6069	5,80226			
Total	36729,09	6073				

ou seja, H1 não pode ser rejeitada (pelo menos uma das médias difere das outras)

Para definir quais médias diferem entre si, foi utilizado o procedimento de Tukey

Procedimento de Turkey

<i>Grupo</i>	<i>Contagem</i>	<i>Soma</i>	<i>Média</i>	<i>Variância</i>
wtINHA	1044	-16216,49	-15,53304	3,403168
S94A	1475	-22619,29	-15,33511	9,233571
I21V	1368	-20777,42	-15,18817	3,274061
I16T	1436	-20939,8	-14,58203	5,612556
I47T	751	-10512,54	-13,99806	7,365961

<i>Fonte da variação</i>	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>valor-P</i>	<i>F crítico</i>
Entre grupos	1515,172	4	378,793	65,2837	3,82E-54	2,373397
Dentro dos grupos	35213,92	6069	5,80226			
Total	36729,09	6073				

Tukey's Multiple Comparison Procedure

$$(Y_r - Y_s) \pm (q(\alpha; k, v)/\text{raiz}(2)) * \text{raiz}(M_{\text{error}}) * \text{raiz}(1/n_r + 1/n_s)$$

$$q(0,05; 5, 6069) = 3,86$$

$$(3,86/\text{raiz}(2)) * \text{raiz}(5,80226) = 6,5746$$

5 = número de grupos

6069 = 6074 - 5 = número total de observações - número de grupos

wtINHA e S94A

$$(15,533 - 15,3351) \pm 6,5746 * \text{raiz}(1/1044 + 1/1475)$$

$$0,1979 \pm 0,2659 = (-0,068, 0,4638)$$

ou seja, wtINHA = S94A (intervalo contém o zero)

wtINHA e I21V

$$(15,533 - 15,1882) \pm 6,5746 * \text{raiz}(1/1044 + 1/1368)$$

$$0,3448 \pm 0,2702 = (0,0746, 0,6150)$$

ou seja, wtINHA é diferente (menor) do que I21V

wtINHA e I16T

$$(15,533 - 14,582) \pm 6,5746 * \text{raiz}(1/1044 + 1/1436)$$

$$0,9510 \pm 0,2674 = (0,6836, 1,2184)$$

ou seja, wtINHA é diferente (menor) do que I16T

wtINHA e I47T

$$(15,533 - 13,9981) \pm 6,5746 * \text{raiz}(1/1044 + 1/751)$$

$$1,5349 \pm 0,3146 = (1,2203, 1,8495)$$

ou seja, wtINHA é diferente (menor) do que I47T

Anexos

Drugs that Inhibit Mycolic Acid Biosynthesis in *Mycobacterium tuberculosis*

E.K. Schroeder^{§†}, O. N. de Souza[†], D.S. Santos[§], J.S. Blanchard[‡], and L.A. Basso[§]

[§]Rede Brasileira de Pesquisa em Tuberculose, Departamento de Biologia Molecular e Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil

[†]Programa de Pós-Graduação em Ciência da Computação, Faculdade de Informática, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil

[‡]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY, USA



Abstract: Tuberculosis resurged in the late 1980s and now kills more than 2 million people a year. The reemergence of tuberculosis as a potential public health threat, the high susceptibility of human immunodeficiency virus-infected persons to the disease, and the proliferation of multi-drug-resistant (MDR) strains have created much scientific interest in developing new antimycobacterial agents to both treat *Mycobacterium tuberculosis* strains resistant to existing drugs, and shorten the duration of short-course treatment to improve patient compliance. Bacterial cell-wall biosynthesis is a proven target for new antibacterial drugs. Mycolic acids, which are key components of the mycobacterial cell wall, are α -alkyl, β -hydroxy fatty acids, with a species-dependent saturated "short" arm of 20-26 carbon atoms and a "long" meromycolic acid arm of 50-60 carbon atoms. The latter arm is functionalized at regular intervals by cyclopropyl, β -methyl ketone, or β -methyl methylethers groups. The mycolic acid biosynthetic pathway has been proposed to involve five distinct stages: (i) synthesis of C20 to C26 straight-chain saturated fatty acids to provide the α -alkyl branch; (ii) synthesis of the meromycolic acid chain to provide the main carbon backbone, (iii) modification of this backbone to introduce other functional groups; (iv) the final Claisen-type condensation step followed by reduction; and (v) various methyltransferase processes to cellular lipids. The drugs shown to inhibit mycolic acid biosynthesis are isoniazid, ethionamide, isoxyl, thiolactomycin, and triclosan. In addition, pyrazinamide was shown to inhibit fatty acid synthase type I which, in turn, provides precursors for fatty acid elongation to long-chain mycolic acids by fatty acid synthase II. Here we review the biosynthesis of mycolic acids and the mechanism of action of antimicrobial agents that act upon this pathway. In addition, we describe molecular modeling studies on InhA, the bona-fide target for isoniazid, which should improve our understanding of the amino acid residues involved in the enzyme's mechanism of action and, accordingly, provide a rational approach to the design of new drugs.

1. INTRODUCTION

1.1. Epidemiology

Tuberculosis remains one of the deadliest diseases in the world. The fifth annual report on global tuberculosis control of the World Health Organization found that there were an estimated 8.4 million new cases in 1999, up from 8.0 million in 1997 [1]. It is expected that there will be 10.2 million new cases in 2005 if the present trend continues. Approximately 3 million persons die from the disease each year [2]. Ninety-percent of tuberculosis cases occur in developing countries, where few resources are available to ensure proper treatment and where human immunodeficient virus (HIV) infection may be common. In 1997, an estimated 1.87 million people died of TB and the global case fatality rate was 23% but exceeded 50% in some African

countries with high HIV rates [3]. In 1990, of the 50.5 million deaths of all causes worldwide, approximately 2 million were due to tuberculosis, with 98% of these deaths occurring in the demographically developing nations [4]. The concentration of tuberculosis and mortality in the age range from 25 to 54 years, the most economically fruitful years of life, causes substantial losses in productivity and contributes to the impoverishment of third-world countries [5]. Tuberculosis is generally thought to develop in an individual as a consequence of one of three processes: progression of primary infection, exogenous reinfection, or endogenous reactivation. Individuals with immunosuppressive conditions such as HIV infection, diabetes, corticosteroid therapy, silicosis, gastrectomy, and intravenous drug abuse may result in reactivation of the bacilli in the latently infected host. Tuberculosis in adults, whether due to a primary infection, reinfection, or reactivation, is sputum positive in about 50% of cases, and these cases are the main source of infection or reinfection in the population [6]. It is for this reason that attempts to reduce the transmission of tubercle bacilli in a population through treatment focuses primarily on the identification and treatment of sputum-positive patients so as to render them sputum negative as rapidly as possible and thus reduce the risk of infecting others. The

(Address correspondence to this author at the Rede Brasileira de Pesquisa em Tuberculose, Departamento de Biologia Molecular e Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil; LAB: Tel: +55-51-33166234; Fax: +55-51-33167309; E-mail: labasso@dna.cbiot.ufrgs.br and JSB: Tel: +1-718-4303096; Fax: +1-718-4308565; E-mail, blanchar@aecom.yu.edu

acid-fast smear plays an important role in early diagnosis of mycobacterial infection because of the relatively long time required for mycobacteria to be detected by culture methods. However, various studies have indicated that 50 to 80% of patients with pulmonary tuberculosis will have positive sputum smears of concentrated clinical specimens [2]. In addition, acid-fast bacilli seen on smears may represent either *M. tuberculosis* or nontuberculous mycobacteria. Therefore, the development of new methods of diagnosing tuberculosis that would be applicable, acceptable and affordable by developing countries, require minimal skills or supervision and be readily integrated into the health services of these countries are decidedly needed.

Development of modern multi-drug antibiotic therapy permitted a continuous decrease in tuberculosis mortality rates. This trend continued from 1953, when U.S. surveillance programs were instituted, to the mid-1980s, when the number of new reported cases began to increase in the United States. Possible factors underlying the resurgence of tuberculosis have been proposed [7]: the HIV epidemic, increase in the homeless population, decline in health care structures and national surveillance. The resurgence of tuberculosis caused by the emergence of drug-resistant strains throughout the world [8-9] has led to an increased need for understanding the molecular mechanisms of drug action and drug resistance. A multi-drug resistant (MDR) strain is defined as one that has resistance to rifampicin and isoniazid (and occasionally other drugs). The emergence of MDR-TB in different parts of the world is particularly worrisome, given the poor therapeutic outcomes when isolates are resistant to both isoniazid and rifampicin [10-11]. Furthermore, the cost of treating MDR-TB can approach US\$ 250,000, about 10-15 times the cost of treating a case of fully drug-sensitive tuberculosis [4].

Robert Koch, in 1882, identified an acid-fast bacterium, *Mycobacterium tuberculosis*, as the causative agent of tuberculosis [12]. In 1921, the BCG (bacille Calmette-Guerin) vaccine was first used to immunize a patient, and is currently the most widely used vaccine in the world. Discovery of the antibacterial activity of a drug released for general medical use under the trade name of Prontosil rubrum in 1935 by Gerhard Domagk marked a fundamental change from pathogen-specific therapy to non-pathogen-specific therapy that allowed prompt and effective treatment of bacterial infections without the necessity of identifying the pathogens involved [13]. The discovery of the antibacterial and antitubercular properties of streptomycin in 1944 [14], and both isoniazid and pyrazinamide in 1952 [15-16], led to effective chemotherapies that decreased tuberculosis mortality rate worldwide. The later introduction of ethionamide, rifampicin, ethambutol and ciprofloxacin to the arsenal used to treat tuberculosis seemed to provide an adequate number of effective antimicrobial agents.

The response of the pharmaceutical industry to the progression of antibiotic resistance has involved three primary approaches to the development of chemotherapeutic agents: modification of existing agent classes, interference with resistance mechanisms to increase target access, and search for agents with novel mechanisms of action [17]. All antibiotics introduced since 1960 are chemical modifications

of existing basic structures [18]. Future approaches may include exploitation of potential antibacterial targets, such as DNA or RNA methylation, RNA processing, protein processing, lipid synthesis, membrane assembly, cell division, and cytoplasmic and membrane transfer steps of peptidoglycan synthesis [19-21]. Moreover, investigation of the molecular interactions between specific mycobacterial products and the host cell should provide attractive targets for therapeutic intervention [22].

1.2. Physiology of Tubercle Bacilli

Tuberculosis is caused by *Mycobacterium tuberculosis*, which is a rod-shaped obligate aerobe, that survives best in an environment with a pO₂ of 100 to 140 mm Hg, 5% CO₂ and a physiologic pH of 7.4 [23]. The bacilli colonize the walls of pulmonary cavities where oxygen tension is high and the medium is lightly alkaline. *M. tuberculosis* is a facultative, intracellular bacterial pathogen that survives and multiplies within phagosomes whose maturation has been arrested by the bacilli and that resist acidification [24]. This altered phagosomal maturation seems to be associated with modifications in the protein content of the vacuole including persistence of functional Rab5 GTPase with ensuing maintenance of phagosome-early endosome interaction [25], exclusion of the vacuolar proton-ATPase with consequent lack of acidification [26], and retention of a phagosomal coat protein designated TACO (for Tryptophane Aspartate-containing Coat Protein) that is normally released prior to phagosome fusion with, or maturation into, lysosomes [27]. Uptake of mycobacteria into macrophages and subsequent retention of TACO on the phagosome appears to be dependent on the accumulation of host cell-derived cholesterol in the plasma membrane at the point of bacterial entry [28]. A recent report has suggested that the exported repetitive protein (ERP) is also essential for *M. tuberculosis* growth in macrophages [29]. Although some of the biochemical events that accompany mycobacterial vacuolar trafficking within the macrophage have recently been revealed, the specific bacterial proteins or cell envelope components that interact with the host cell to perturb these processes are unknown.

The disease is spread from person to person almost exclusively by small-particle aerosols of sufficiently small size (1-5 μm in diameter that may contain from 1 to 10 tubercle bacilli) so that they are inhaled into the alveolar space, where they are deposited and the organism begins to replicate. Tubercle bacilli divide slowly with a doubling time of approximately 24 hours [30]. This slow doubling time increases the length of time required for chemotherapy and makes drug susceptibility determinations time-consuming. This slow growth is also the reason for an initially intensive drug therapy, useful in eliminating larger bacterial populations, followed by fewer drugs and less frequent doses for an extended period (total six months) in order to completely eradicate the bacilli. Mitchison [31] has suggested that pathogenic mycobacterial populations be divided into four components: actively metabolizing and rapidly growing, semidormant in an acidic intracellular environment, semidormant in a nonacidic intracellular environment, and dormant. The latter categories appear to be

unique properties of mycobacterial infections, with the organisms able to remain quiescent for years, or decades.

1.3. Chemotherapy

Effective tuberculosis chemotherapy must include early bactericidal action against rapidly growing organisms and subsequent sterilization of the semidormant and dormant populations of bacilli [31]. The first-line drugs isoniazid, rifampicin, streptomycin, and ethambutol exhibit early bactericidal activity against actively metabolizing bacilli [32]. Pyrazinamide is active against the semidormant bacilli in acidic intracellular environments. The bacteriostatic second-line drugs ethionamide, cycloserine and *para*-aminosalicylic acid (PAS) are reserved to strengthen the treatment of drug-resistant disease or when bactericidal drugs are prohibited because of toxicity [33]. The newer fluoroquinolones (sparfloxacin, ciprofloxacin and ofloxacin) are bactericidal against *M. tuberculosis*. Although usually not indicated as first-line therapeutic agent for the treatment of tuberculosis, the fluoroquinolones have an important role in the treatment of resistant organisms. The modern, standard “short-course” therapy [31, 34] for tuberculosis is based on a four-drug regimen of isoniazid, rifampicin, pyrazinamide, and ethambutol or streptomycin for two months, followed by treatment with a combination of isoniazid and rifampicin for an additional four months. This combination therapy must be strictly followed to prevent drug resistance and relapse, and direct observation of patient compliance is the most reliable way to ensure effective treatment and prevent the acquisition of resistance [35].

Mycobacteria show a high degree of intrinsic resistance to most antibiotics and chemotherapeutic agents due to the low permeability of the cell wall [36]. Hydrophilic agents traverse the cell wall slowly because of mycobacterial porin inefficiency in permeation of solutes, and low concentration of porins. Lipophilic agents are retarded by the lipid bilayer which is of unusually low fluidity. Nevertheless, the cell wall barrier alone cannot produce significant levels of drug resistance.

Resistant mutants of *M. tuberculosis* occur naturally, irrespective of exposure to drugs [37]. *M. tuberculosis* mutants resistant to any single effective drug are present in any large bacterial population at a frequency of 1×10^{-6} [38]. The frequency of mutants resistant to rifampicin is 1×10^{-8} , and 1×10^{-6} for isoniazid and streptomycin. Therefore the large extracellular population (10^7 - 10^9) of actively metabolizing and rapidly growing tubercle bacilli in cavitory lesions will contain approximately 10 to 1000 organisms which are resistant to a single effective drug. Consequently, monotherapy or improperly administered two-drug therapy will select for drug-resistant mutants which ultimately may lead to drug resistance in the entire population of bacilli.

The factors that most influence the emergence of drug-resistant strains include inappropriate treatment regimens and patient noncompliance in completing the prescribed courses of therapy [39]. Scrutiny of treatment strategies has resulted in recommendations for inclusion of as many as three or four

additional drugs to the standard four-drug regimen of isoniazid, rifampicin, ethambutol, and pyrazinamide [40].

2. MYCOLIC ACID STRUCTURE AND BIOSYNTHESIS

2.1. The Mycobacterial Cell Wall

The model of mycobacterial cell wall structure, which has emerged largely from biochemical studies of Patrick Brennan and colleagues [reviewed in 41-48], is depicted in Fig. (1). The mycobacterial cell wall is comprised of three covalently linked macromolecules: peptidoglycan, arabinogalactan, and the hallmark of mycobacteria, mycolic acid. This covalently linked skeleton of cell wall is often described as mycolyl-arabinogalactan-peptidoglycan complex (mAGP). The cytoplasmic membrane of mycobacteria is encapsulated by a rigid layer of peptidoglycan. The peptidoglycan component of the cell wall is a polysaccharide chain composed of repeating *N*-acetyl- β -D-glucosaminyl-(1-4)-*N*-glycolylmuramic acid units crosslinked by *L*-alanyl-*D*-isoglutaminyl-*meso*-diaminopimelyl-*D*-alanine peptide chains linked to the lactoyl groups of muramic acid residues (Fig. 2). This type of peptidoglycan, Ala [49], is one of the commonest found in bacteria. However, the mycobacterial peptidoglycan differs in two respects. First, the muramic acid is *N*-glycolylated in *Mycobacterium* and *Nocardia*, in contrast to the *N*-acetylation found in all other bacteria, which may provide additional hydrogen bonding resulting in enhanced rigidity of the peptidoglycan structure [43]. Second, the peptide cross-links in the mycobacterial peptidoglycan include a proportion of bonds between two residues of *meso*-diaminopimelic acid (*meso*-DAP) instead of the more usual *meso*-DAP-*D*-alanine linkages [50].

The mycobacterial peptidoglycan structure is attached to arabinogalactan through a phosphodiester linkage between the C6 of about 10-12 % of muramic acid residues and a disaccharide linker unit attached to the nonreducing terminus of the galactan in the following arrangement: β -D-Galf-(1-4)- β -L-Rhap-(1-3)- β -D-GlcNAc-(1-6)-Mur-*N*-glycolyl. Arabinogalactan is a unique polysaccharide consisting of linear galactan chains composed of alternating 5- and 6-linked β -D-Galf (β -D-Galactofuranose) units which, in turn, are attached through C5 of some of the 6-linked β -D-Galf units to the extensively branched chains of *D*-Araf-containing arabinan. The nonreducing termini of arabinan consist of a branched hexaarabinofuranosyl structure [β -D-Araf-(1-2)- β -D-Araf] $_2$ -3,5- β -D-Araf-(1-5)- β -D-Araf, of which, about two-thirds are esterified to mycolic acids (Fig. 3).

In addition to lipids in the covalently linked mAGP skeleton, several types of extractable lipids are present in various mycobacterial species, including trehalose-containing glycolipids, phenolic glycolipids, glycopeptidolipids (GLPs), lipooligosaccharides (LOSs), phosphatidylinositol mannosides (PIMs), phosphatidylethanolamine (PE), and triacylglycerols (TAGs). Among the most prominent glycolipids non-covalently associated to the mAGP skeleton in *M. tuberculosis* is lipoarabinomannan (LAM). LAM is

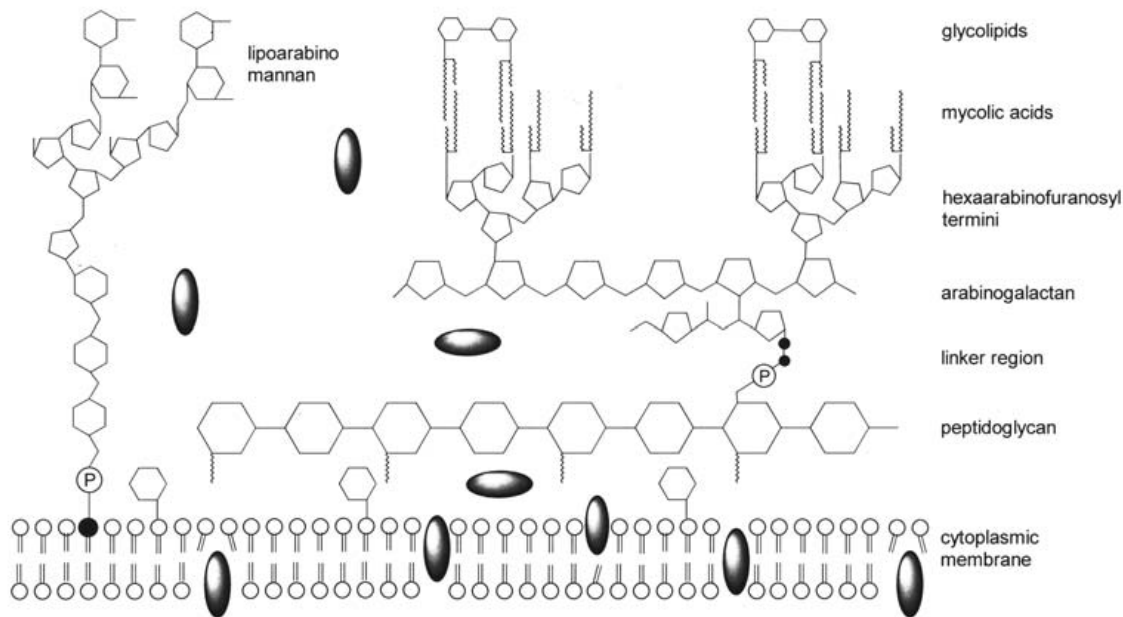


Fig. (1). Schematic representation of the mycobacterial cell wall. The cytoplasmic membrane is encapsulated by a layer of peptidoglycan. The peptidoglycan backbone is attached to arabinogalactan through an unusual disaccharide phosphate linker region. The arabinogalactan is a branched-chain polysaccharide consisting of a proximal galactose chain linked to a distal arabinose chain. The hexaarabinofuranosyl termini of arabinogalactan are esterified to mycolic acids. The mycolic acid chains are shown perpendicular to the cytoplasmic membrane with the exposed chains interacting with the mycolic chains of trehalose dimycolate. Another major component non-covalently associated to the mycobacterial cell wall is the immunogenic lipoarabinomannan, which is attached to the cytoplasmic membrane by a phosphatidylinositol anchor.

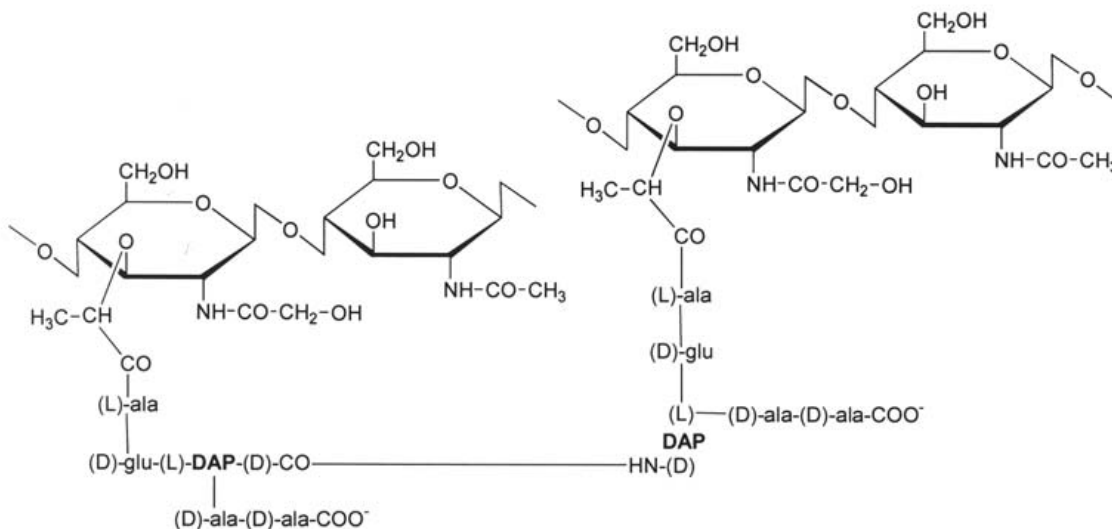


Fig. (2). The peptidoglycan component of the cell wall. It is a polysaccharide of repeating *N*-acetyl- β -D-glucosaminyl-(1 \rightarrow 4)-*N*-glycolylmuramic acid units crosslinked by L-alanyl-D-isoglutaminyl-*meso*-diaminopimelyl-D-alanine peptide chains linked to the lactoyl groups of muramic acid residues.

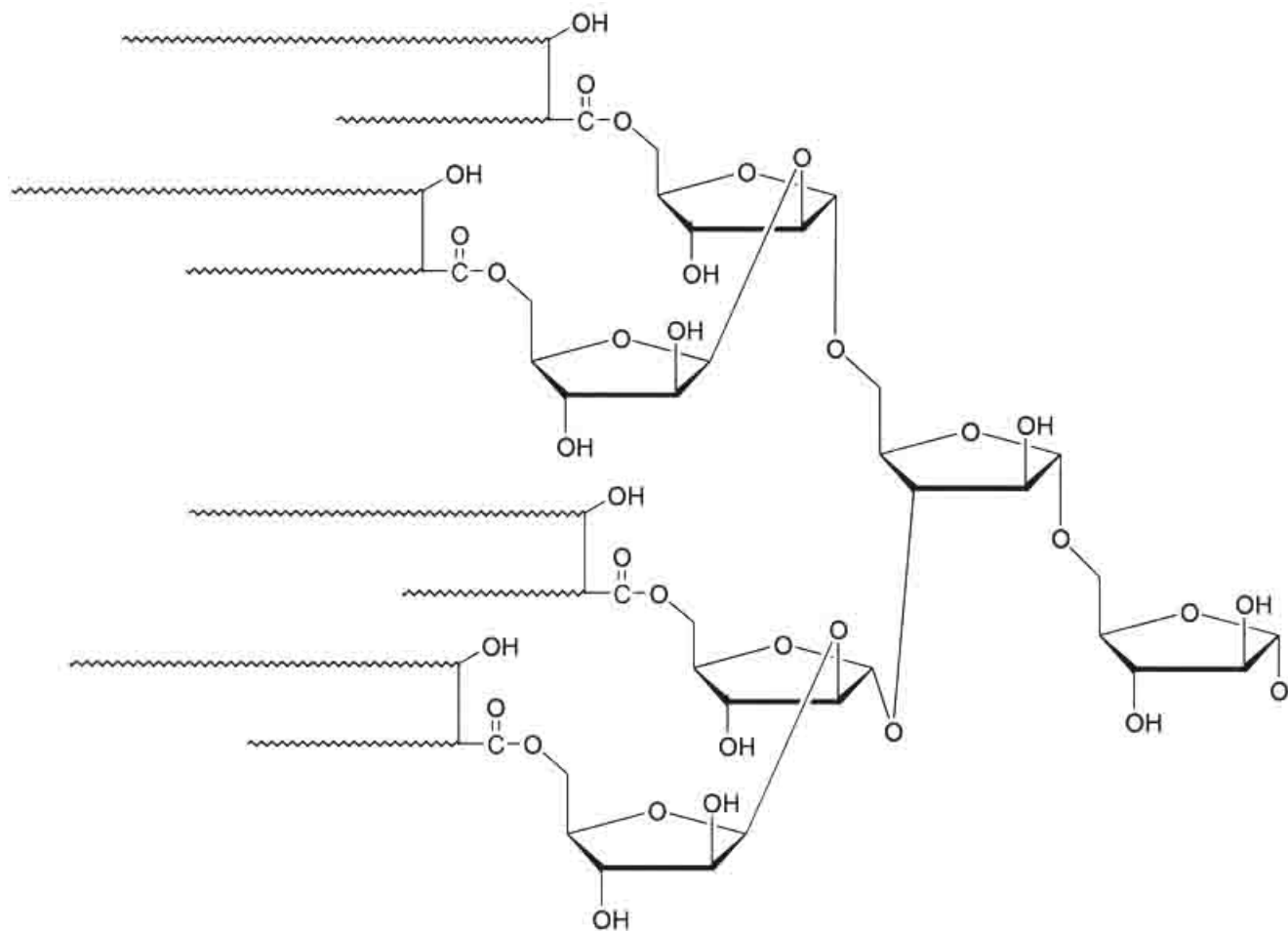


Fig. (3). Hexaarabinosyl structure of non-reducing termini of arabinan. They consist of a branched hexaarabinofuranosyl structure of which about two-thirds are esterified to mycolic acids.

anchored in the plasma membrane by a phosphatidylinositol-diacylglycerol domain at the proximal end, linked to a branched-chain arabinomannan polysaccharide. LAM displays a number of biological activities relevant to tuberculosis pathogenesis, including inhibition of T-cell activation and of macrophage activation by interferon- γ , scavenging of reactive oxygen intermediates, inhibition of protein kinase C activity, and induction of macrophages to secrete several cytokines including TNF [4]. The total lipid content of the mycobacterial cell wall is over 60% in weight, of which the covalently linked skeleton of cell wall comprises over 40% [50].

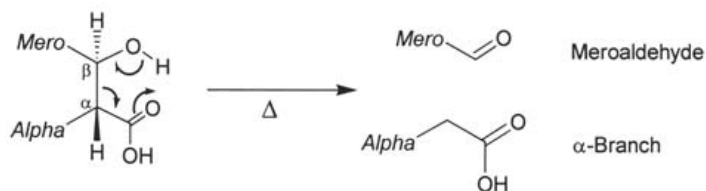
2.2. Mycolic Acids: Chemical Structure

Mycolic acids are high-molecular-weight α -alkyl, β -hydroxy fatty acids, appearing mostly as bound esters in tetramethylpentaarabinosyl clusters (Fig. 3). Mycolic acids are also present as extractable lipids, mainly as trehalose 6,6'-dimycolate (cord factor). In the pyrolytic cleavage of mycolic acids the intact fatty acid released is often referred to as the α branch since it occupies the alpha position with respect to the carboxylic acid group (Fig. 4A). The aldehyde released is referred to as the meroaldehyde and the

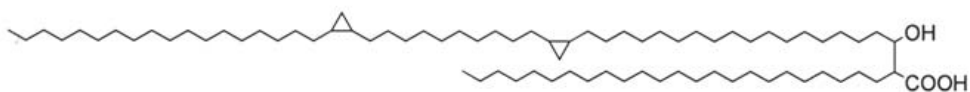
corresponding segment of the intact mycolate is often referred to as the meromycolate branch (Fig. 4A). Mycolic acids possessing only desaturation and/or cyclopropanation in the meromycolate branch are known, rather confusingly, as β -mycolic acids, where in this case simply refers to their position on thin layer chromatography (TLC). In order to distinguish between these two uses of the greek letter, the term " β branch" will be used to refer to the shorter acid released on pyrolysis. Accordingly, mycolic acids are α -branched lipids, with a species-dependent saturated "short" arm of 20-26 carbon atoms and a "long" meromycolic acid arm of 50-60 carbon atoms, functionalized at regular intervals by cyclopropyl (β -mycolates), β -methyl ketone (ketomycolates), or β -methyl methylethers (methoxymycolates) groups (Fig. 4B). In *M. tuberculosis*, β -mycolic acids are predominant followed closely by methoxymycolates with significantly less ketomycolates [51]. The stereochemistry of the asymmetric centers at C-2 (C-2) and C-3 (C-3) of a range of mycolic acids has been shown to be *R,R*, thus these two centers are *erythro* to each other [51].

Treatment of intact cells with mixtures of chloroform and methanol extracts the smaller quantity of non-covalently attached mycolate, yielding delipidated cells that are treated under alkaline condition to extract mycolates of the cell wall

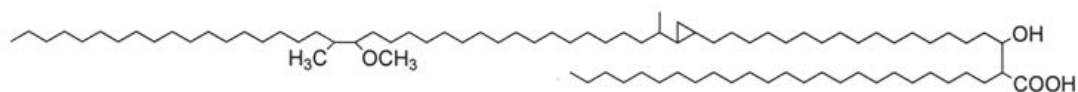
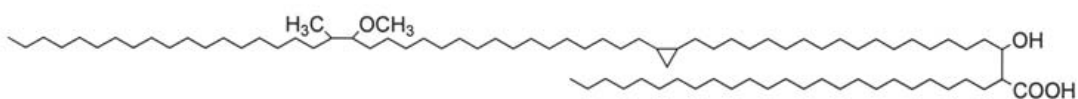
(A)



(B)

 α -Mycolates

Methoxymycolates



Ketomycolates

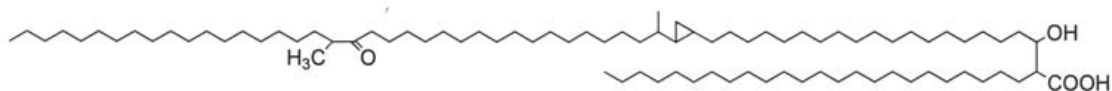
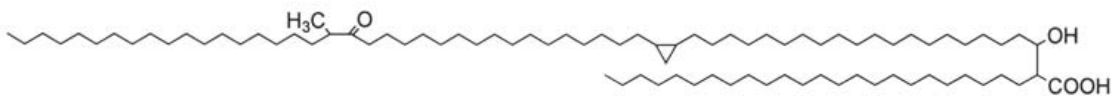


Fig. (4). A). Pyrolysis of β -hydroxy carboxylic acid, yielding the α -branch and meroaldehyde. B). The structures of mycolic acids identified in *M. tuberculosis*. α -Mycolates are the most abundant mycolic acid type (57%) in *M. tuberculosis* and its mero chain contains two *cis*-cyclopropanes. Methoxymycolates, the second most abundant class of mycolic acid in *M. tuberculosis* (32%), contain an β -methyl methyl-ether moiety in the distal position and a *cis*-cyclopropane or an β -methyl *trans*-cyclopropane in the proximal position. Ketomycolates, which are less abundant (11%), contain an β -methyl ketone moiety in the distal position and proximal functionalities as in the methoxy series.

[46]. Separation, identification and structural characterization of mycolic acids were resolved through the use of a variety of techniques such as TLC, IR, proton and carbon NMR, MS, pyrolytic GC as well as basic chemical analysis. Mycobacterial mycolic acids present the following features that distinguish them from those of other genera (such as *Corynebacterium*, *Dietzia*, *Nocardia*, *Rhodococcus*, and *Tsukamurella*): (a) They are the largest (C₇₀ to C₉₀); (b) they have the largest β -branch (C₂₀ to C₂₆); (c) their meromycolic acid moiety contains one or two groups, which may be double bonds or cyclopropane rings, that are capable of introducing "kinks" (bends) in the molecule; (d) they may contain oxygen functions additional to the β -hydroxy group; and (e) they may have methyl branches in the meromycolate chain. Unlike mycobacteria, desaturation is the only form of functionalization found in the other members of mycolate-containing genera.

The functional groups present in mycobacterial mycolic acids occur at two points in the meromycolate chain and are referred to as distal (closest to the β -end of the chain) and proximal (closest to the β -hydroxy acid). Polar modifications are usually restricted to the distal position, including such functional groups as methyl ethers, ketones, esters, and epoxides. Non-polar modifications occur at both the distal and the proximal positions and include *cis* or *trans* double bonds and *cis* or *trans* cyclopropanes. Whenever double bonds and cyclopropanes occur in a *trans* configuration, methyl groups occur adjacent to them (*cis*-cyclopropanes and *cis* double bonds in mycolic acids do not occur adjacent to methyl groups). The structures of mycolic acids identified in *M. tuberculosis* are shown in Fig. (4B). A comprehensive list of mycolic acids found in several mycobacterial species was given by Barry III *et al.* [46].

It should be pointed out that there is considerable heterogeneity within a particular class of mycolic acids, which is a result of β -branch and meromycolate lengths variation in *M. tuberculosis* [52-53]. In addition to this heterogeneity, the term "microheterogeneity" has been used for a population of mycolic acids with the same total number of carbons but with functional groups at different positions [46]. This microheterogeneity contributes significantly to the total number of molecules in each class of mycolic acids. It was estimated that for the five major classes of mycolic acids in *M. tuberculosis* there may be 500 structural isomers [46].

2.3. Cell-Wall Bound Mycolic Acids: Structure's Physiological Roles

The totally saturated and fairly long structure of the β -branch as well as the exceptional length of the meromycolate chain should favor the regular, parallel packing of the hydrocarbon chains of mycolic acid. On the other hand, the oxygen functions, *cis* double bonds, and *cis* cyclopropane rings are expected to introduce kinks in the chains, thereby modulating the mycolic acid tight packing. There usually are 12-13 *Ar* β residues between the mycolate residues and the central galactan chain (Fig. 3). Owing to the exceptional flexibility of *Ar* β -(1 \rightarrow 5)-*Ar* β linkages of the arabinogalactan, tight packing should be facilitated by the

lateral movement of mycolic acids afforded by these linkages [43]. Moreover, the (*R*)- β -OH group appears to stabilize the mycolic acid monolayer (Fig. 3) by intermolecular or intramolecular hydrogen bonding [54].

X-ray diffraction studies of the purified *M. chelonae* cell wall showed that a large part of the hydrocarbon chains in the cell wall are tightly packed in a parallel array in a direction perpendicular to cell surface [55]. Moreover, a large fraction of the lipid is in a nearly crystalline arrangement of presumably very low fluidity, whereas the rest show the diffuse diffraction expected of fluid domains. The coexistence of these domains of different fluidities is consistent with the mycolic acid structure: The β -branch has no double bond or cyclopropane groups; functional groups, double bond or cyclopropane, at the proximal position of the meromycolate chain are often *trans*, thereby allowing tight packing. The presence of double bonds and cyclopropane ring in the *cis* configuration and oxygen-containing groups at the distal position would disrupt the tight lateral packing, consistent with an increase in fluidity. A primary physiological function of the proximal substituent in mycolic acids has been proposed to be the maintenance of a membrane environment of constant fluidity through *cis/trans* ratio changes [56]. The importance of *trans*-mycolate content to many macroscopic qualities of *M. tuberculosis*, such as increased cell wall rigidity, increased drug resistance, potentially a decreased growth rate, and altered colony morphology, has been demonstrated [56]. More recently, a gene coding for *S*-adenosyl-*L*-methionine dependent enzyme activity, which converts a *cis* olefin to a *trans* olefin with an allylic methyl branch, was isolated [57]. Overexpression of this gene product in *M. tuberculosis* resulted in an increase of *trans* substitutions in the form of cyclopropane and double bond at the proximal position of meromycolate chain. However, compensatory mechanisms in which there was a decrease in the amount of methoxy substituent at the distal position and a corresponding increase in the amount of ketones were observed. The authors pointed out that the correlation between *cis/trans* ratio and fluidity should be considered with caution. Although cyclopropanation of mycolic acids is observed only in pathogenic mycobacteria of the TB complex (*M. tuberculosis*, *M. bovis*, *M. microti*, *M. africanum*), little is known about its physiological role. Since cyclopropanes are relatively inert to oxidation and overexpression of cyclopropane mycolic acid synthase (CMAS-1) resulted in both an increase in cyclopropanated mycolate and an increase in resistance to killing by hydrogen peroxide [58], it has been suggested that cyclopropanation of mycolic acids may play an important role in the survival and pathogenesis of *M. tuberculosis* in macrophages. Accordingly, it has more recently been shown that a mutant of the *M. tuberculosis* *pcaA* gene, which encodes for a *cis*-cyclopropane synthase, is essential for persistence in mice [59]. Oxygenated mycolates have been proposed to provide hydrophilic attachment points for the association of peripheral cell surface molecules [51]. Consistent with this proposal, overexpression in *M. smegmatis* of the *M. tuberculosis* MMAS-4 protein, which introduces a hydroxyl group with an adjacent methyl branch at the distal position of β -mycolates, resulted in an extremely hydrophilic surface morphology [60]. Moreover, overexpression of MMAS-1, an

enzyme encoded in the gene cluster responsible for the biosynthesis of methoxymycolates, resulted in decreased thermal stability of the *M. tuberculosis* cell wall and increased permeability to chenodeoxycholate, thereby suggesting that increasing the proportion of polar ketomycolates may exert significant fluidizing effect on the cell wall [57]. In addition, comparison between clinical isolates and two laboratory strains (H37Rv and H37Ra) of mycolate composition determined by NMR has shown that there is five times more *trans*-cyclopropane groups in the clinical isolates [57]. The potential functional roles of mycolic acids are, therefore, dependent on two basic properties: chain length and functional group (including oxygenated or nonoxygenated as well as configuration about olefinic or cyclopropyl groups).

2.4. Mycolic Acids: Biosynthesis

An understanding of the overall metabolic process leading to synthesis of mycolic acids should provide important targets for the development of chemotherapeutic agents. The mycolic acid biosynthetic pathway has been proposed to involve five distinct steps [61]: (1) synthesis of 24- to 25-carbon straight-chain saturated fatty acids to provide the β -alkyl branch; (2) synthesis of the meromycolic acid chain to provide the main carbon backbone; (3) introduction of other functional groups to the meromycolic acid chain; (4) the final Claisen-type condensation between the meromycolate chain and the β -branch, followed by reduction; and (5) mycolyltransferase processes to cellular lipids.

2.4.1. Short-Chain Fatty Acid Synthesis: β -Branch and Mero Primers (FAS-I)

Fatty acid elongation occurs through repetitive cycles of condensation, β -keto reduction, dehydration, and enoyl reduction which are catalyzed by, respectively, β -ketoacyl synthase (KAS, condensing enzyme), β -ketoacyl reductase (KAR), β -hydroxyacyl dehydrase (DE), and enoyl reductase (ENR). These chemical reactions are catalyzed by two types of fatty acid synthase systems (FAS). The FAS-I system is a multidomain polyprotein that encodes all the enzymes necessary for fatty acid synthesis in one large polypeptide and is generally present in most eukaryotes, except in plants [62]. FAS-II systems, which are present in bacteria and plants, catalyze the individual reactions by separate proteins readily purified independently of the other enzymes of the pathway and are encoded by unique genes. The fatty acid substrate is activated by a thioester linkage to the phosphopantetheine group of either the required cofactor (acyl carrier protein or coenzyme A) or the enzyme complex KAS. The coupling of the growing chain, which is transiently attached to a cysteine at the KAS active site, with malonate to yield a β -ketoacyl product is the first step in elongation and is catalyzed by the KAS component of the FAS system [63]. The reacting malonate molecule is attached to a thiol group of a prosthetic phosphopantetheine which, in turn, is covalently attached to a serine residue of either an acyl carrier protein (ACP) in the case of FAS-II system or an ACP-domain in the FAS-I system. The condensation occurs as a result of nucleophilic attack by the

C2-carbanion ion of decarboxylated malonyl-ACP on the acyl carbon of the thiol ester of the growing fatty acid chain. Reduction of the resulting β -ketoacyl-ACP by KAR yields a secondary alcohol which, in turn, is dehydrated to an α,β -unsaturated enoyl-ACP (*trans*-2-enoyl-ACP) by the DE enzyme activity. The final basic reaction of fatty acid elongation is the reduction of the *trans*-2-enoyl-ACP to a saturated hydrocarbon chain by the ER enzyme activity.

Unlike most organisms, mycobacteria have both FAS-I and FAS-II systems [64]. The mycobacterial FAS-I system catalyzes not only the synthesis of C₁₆ and C₁₈ fatty acids, the normal products of *de novo* synthesis, but also elongation to produce C₂₄ and C₂₆ fatty acids [65-66]. The bimodal distribution of products arises from the ability of the multifunctional synthase to use the very short primers as is usual for *de novo* fatty acid synthesis, and the additional ability to use the longer primers that are the usual end-products of *de novo* synthesis [67]. A *M. bovis* BCG multifunctional FAS protein of 301 kDa has been purified and characterized [67]. Cloning and sequencing of the synthase gene revealed an 8,389-bp open reading frame (ORF) encoding a protein of 2,797 amino acids [68]. Amino acid sequence analysis revealed that the domain organization is as follows: acyl-transferase, enoyl reductase, dehydratase, malonyl/palmitoyl transferase, ACP, β -keto reductase, and β -acyl synthase. In *M. tuberculosis*, the *fas* gene encodes for a FAS-I system that generates several shorter CoA esters from acetyl-CoA primers and probably provides precursors for elongation by all of the other fatty acid and polyketide systems [69].

2.4.2. Long-Chain Fatty Acid Synthesis (FAS-II)

The Type II fatty acid synthase of mycobacteria is analogous to other bacterial FAS-II systems, with the notable exception of primer specificity [64, 70]. The mycobacterial FAS-II is not capable of *de novo* synthesis from acetate but instead elongates palmitoyl-ACP to fatty acids ranging from 24 to 56 carbons in length [71-72]. NMR studies of mycolic acids of *M. tuberculosis* grown in the presence of [1-¹³C]- and [2-¹³C]-acetate suggest that the biosynthetic pathway from palmitate is composed of a series of condensations of malonyl-CoA units onto this primer, rather than head-to-head type condensations of smaller subunits [46]. The complete genome sequence of *M. tuberculosis* H37Rv reveals the organization of the genes encoding the components of the FAS-II system [69]. The main locus is an operon composed of five genes transcribed in the same orientation. The first ORF in the operon, *fadD*, encodes a malonyl-CoA:ACP acyl transferase. The second ORF, *acpM*, encodes an ACP which is the carrier of the growing acyl chain from reactive center to reactive center of each enzyme. The recognition of fatty acids as substrates by the enzymes of FAS-II system is dependent on them being tethered to an ACP [73]. It has recently been suggested that AcpM is the acyl carrier of FAS-II [72]. The third and fourth ORFs, respectively *kasA* and *kasB*, encode for β -ketoacyl-ACP synthases (KasA and KasB). It has recently been suggested that KasA may participate in the synthesis of C₁₈₋₃₄ fatty acids, whereas KasB may be involved in later steps of mycolic acid biosynthesis [48]. Cloning and biochemical characterization of KasA and KasB showed that both

enzymes condense an acyl-ACP with malonyl-ACP and they have specificity for long-chain acyl-ACPs containing at least 16 carbons [74]. It should be pointed out that there are many differences between AcpM (*M. tuberculosis* acyl carrier protein) and ACP (*E. coli* acyl carrier protein) including size and the substitution of various conserved residues which have been suggested to play key roles in Kas recognition [75]. Accordingly, comparison of steady-state kinetic parameters of KasA and KasB indicate that acyl-AcpMs are better substrates than acyl-ACPs [74]. The specific roles of KasA and KasB are still unclear. However, a dissimilar resistance pattern to thiolactomycin upon overexpression of *M. tuberculosis* KasA and KasB enzymes in *M. bovis* BCG [76] and a lower specific activity of KasB as compared to KasA using C_{16:0} and C_{20:0} acyl primers as substrates [74] may indicate that they play unique roles in mycolic acid synthesis. It has been suggested that KasA catalyzes the condensation steps (from C₁₈ to C₃₄ long-chain fatty acids) of the acyl primers passed to the FAS-II system from FAS-I via mtFabH, which would, in turn, be used as KasB substrates for further elongation, leading to longer chain fatty acids, perhaps to meromycolates [76]. The last ORF of the operon, *accD6*, encodes for an acetyl/proprionyl CoA carboxylase subunit of an enzyme that catalyzes the transfer of biotin-bound CO₂ to acetyl-CoA to yield malonyl-CoA. The latter is the primary substrate of the KAS enzymes. The genome reveals a third KasC enzyme, which is a more distant homologue related to ketoacyl synthase III system [69]. Two other genes whose products are required to complete the fatty acid cycle are located in the *inhA* operon [77]. The *inhA* gene codes for a *trans* 2-enoyl reductase (InhA) and the *mabA* (*fabG* in *E. coli*) gene for a β -ketoacyl reductase (MabA). The dehydrase is the only component enzyme of FAS-II system required to complete fatty acid biosynthesis that is missing from these two operons. However, numerous dehydratase-like coding sequences have been identified in the *M. tuberculosis* genome [69]. More recently, mutation at the β -hydroxyacyl-ACP dehydratase gene has been suggested to be responsible for the nearly complete lack of mycolic acids of a *M. smegmatis* mutant [78].

Inactivation of the *inhA*-encoded FAS-II enoyl-ACP reductase was shown to induce accumulation of the FAS-I end products (C_{24:0}) and cell lysis in *M. smegmatis* [79]. The inhibition of mycolic acid biosynthesis demonstrates that the FAS-II system, which InhA is part of, is primarily responsible for the elongation of fatty acids to mycolic acids *in vivo*. Recently, the *M. tuberculosis fabH* (ORF Rv0533c) gene was cloned and shown to encode for a β -ketoacyl-ACP synthase III, termed mtFabH, that has a marked preference for long-chain acyl-CoAs, in which the upper limit appears to be 16 carbons, as primers over long-chain acyl-ACPs, and is inactive with acetyl-CoA by a conformational-sensitive gel electrophoresis assay [80]. Bacterial condensing enzymes are divided into two broad categories. The FabH (β -ketoacyl-ACP synthase III, mtFabH in *M. tuberculosis*) class is responsible for initiation of type II pathways and prefers acyl-CoA primers over acyl-ACP whereas the FabB/FabF (β -ketoacyl-ACP synthase I/ β -ketoacyl-ACP synthase II in *E. coli*) class of condensing enzymes uses acyl-ACP primers and has broad substrate specificity that allows catalysis of all of the subsequent elongation reactions in the pathway. The three-dimensional structure determination of the 76.6 kDa

homodimeric mtFabH offered a molecular basis for understanding the unusual substrate specificity of mtFabH [81]. The mtFabH substrate specificity is consistent with this enzyme playing a role in funneling acyl-CoA formed by the type I system into the type II chain-elongation system, which is ultimately responsible for the formation of mycolic acids. The mycobacterial FAS-I produces a bimodal (C_{14:0}-C_{16:0} to C_{24:0}-C_{26:0}) distribution of acyl-CoA fatty acids. The substrate specificity of mtFabH and crystal structure indicate that only the shorter acyl-CoA products (C_{14:0}-C_{16:0}) from FAS-I are elongated by FAS-II. Accordingly, the mycobacterial FAS-I system would provide the long-chain acyl-CoA precursors which are condensed with malonyl-ACP by mtFabH enzyme activity whose products, in turn, would be elongated by the FAS-II system, yielding the long carbon chain of the meromycolate branch (50-60 carbons) of mycolic acids (Fig. 5). The longer chain acyl-CoA products (C_{24:0}-C_{26:0}) of FAS-I would be excluded from chain elongation and remain available to be utilized, presumably in the CoA form, as substrates for formation of the β -alkyl branch (20-26 carbons) of mycolic acids.

2.4.3. Modification Reactions of the Meromycolic Acid Chain: Desaturation, and Methyl Transfer and Oxidation Reactions

The introduction of functional groups in the meromycolate chain is thought to begin with desaturations of the saturated alkyl chain to yield a *cis* double bond at the distal position and a *cis* double bond at the proximal position. Desaturation may follow formation of the full-length saturated meromycolate branch or occur during fatty acid elongation. *M. tuberculosis* H37Ra cell extracts have been shown to synthesize C₃₀ to C₅₆ fatty acids containing various numbers of double bonds [82]. Identification of unsaturated and cyclopropanated fatty acids shorter than full-length meromycolate chains that are possible precursors of the full-length meromycolate suggests that desaturation may be coupled to elongation [83]. A *M. tuberculosis* exported protein (Des) has been identified by gene fusion methodology and shown to share significant sequence identity to soluble plant stearyl-ACP desaturase [84]. Amino acid sequence comparison showed the presence of conserved motifs found only in class II diiron-oxo (Fe-O-Fe) proteins [84-85]. Three potential ORFs for aerobic acyl-ACP desaturase activity (*desA1*, *desA2*, and *desA3*) have more recently been identified in the *M. tuberculosis* genome [69]. These proteins show similarity to plant ACP-dependent desaturases instead of vertebrate or yeast enzymes that act on CoA esters, suggesting that unsaturation of the meromycolate branch may occur while the acyl group is bound to AcpM [76].

It has been shown by incorporation experiments of either [¹⁴C-methyl]methionine or [³H-methyl]methionine present in mycobacterial growth media that the bridging methylenes of the cyclopropane rings, the carbon of the methoxy functional group, and the methyl branches adjacent to *trans*-olefins, methoxy and keto moieties are all derived from methionine, probably via SAM [86-88]. The mycolic acid methyl transferases of *M. tuberculosis* are a large family of highly homologous SAM-dependent proteins that modify the mycolic acids of the cell wall with cyclopropane rings and

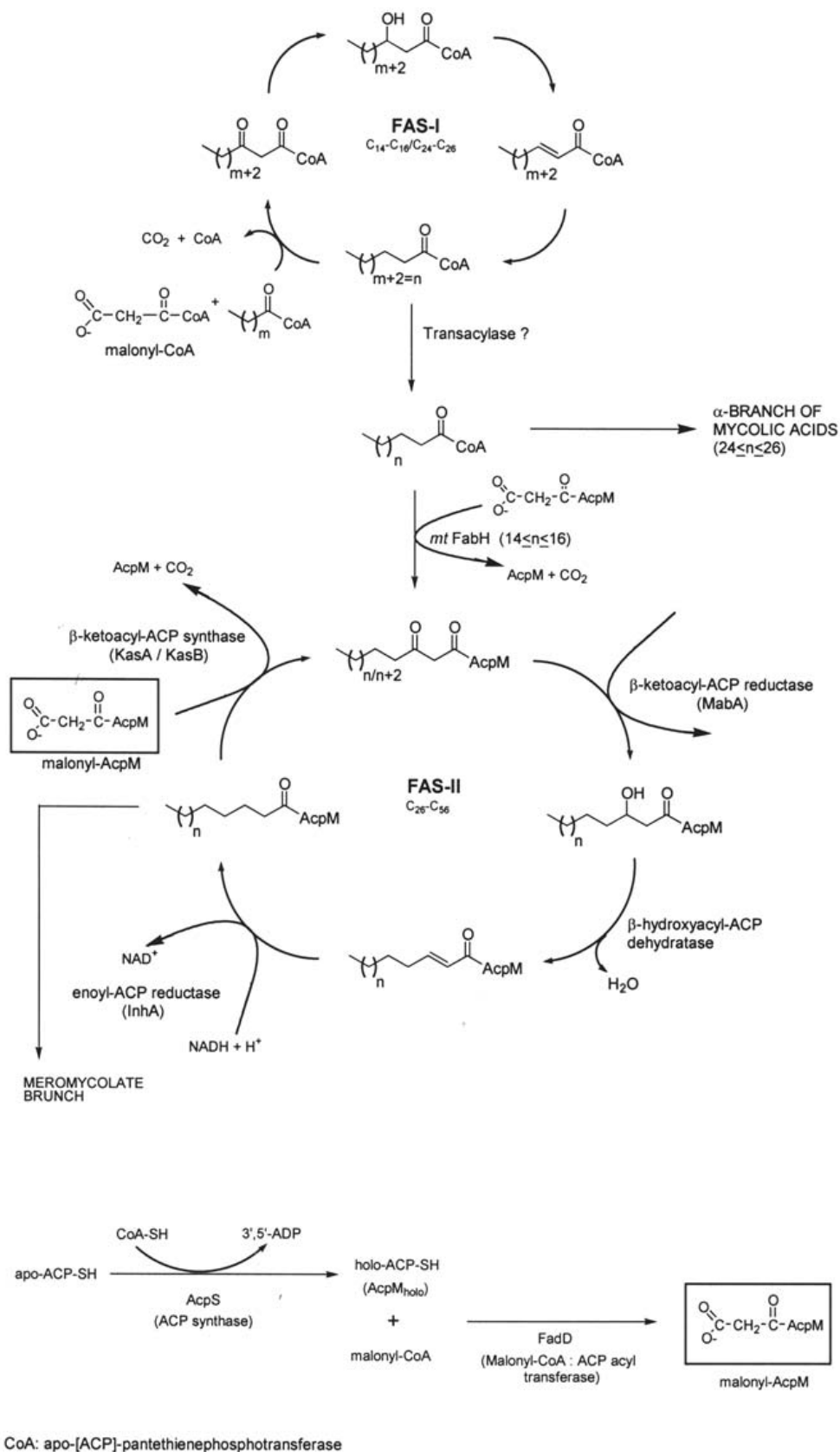


Fig. (5). Fatty acid biosynthesis in mycobacteria. Fatty acid synthase I (FAS-I) carries out the synthesis of C_{16:0}/C_{18:0} and C_{24:0}/C_{26:0}. One or both FAS-I products act as substrates for the synthesis of mycolic acids by fatty acid synthase II (FAS-II). The fatty acids produced by FAS-I are removed from the complex by an unidentified transacylase yielding long-chain acyl-CoAs. These acyl-CoAs are used by mtFabH, which initiates a round of elongation by FAS-II that eventually produces the mycolic acids.

methyl branches, and as many as 8 discrete homologs of this family have been identified in the *M. tuberculosis* genome [69]. Despite striking amino acid sequence conservation of the members of this family, the individual cyclopropane synthases appear to have distinct catalytic function that cannot be compensated by another member of the family [22].

The product of a *M. tuberculosis* gene, termed *cmal* for cyclopropane mycolic acid synthase (CMAS-1), has been shown to confer upon *M. smegmatis* the ability of acting upon olefinic precursor to introduce *cis*-cyclopropane functions at distal position of the meromycolate branch of mycolic acids [58]. Sequence alignment of the *E. coli* cyclopropane fatty acid synthase and *M. tuberculosis* CMAS-1 proteins showed two segments of noteworthy homology: an S-adenosyl-L-methionine consensus binding motif (from amino acid 79 to 87) [VL]-L-[ED]-X-G-X-G-X-G [89], and the cysteine-290 residue implicated in catalysis by the *E. coli* enzyme [90].

By sequence homology to *cmal*, a second gene from *M. tuberculosis* termed *cmal2* was identified. The corresponding protein, CMAS-2, has been shown by heterologous expression in *M. smegmatis* to introduce a cyclopropane ring at the proximal position in α -mycolic acids [91]. Since this enzyme can cyclopropanate the α but not the β mycolic acids from *M. smegmatis*, it was concluded that CMAS-2 is specific for *cis* double bonds. More recently, a mycobacterial gene, *pcaA*, has been established as essential for *M. tuberculosis* pathogenesis because a mutant of *pcaA* cannot establish a chronic persistent *M. tuberculosis* infection in mice [59]. Surprisingly, *pcaA* was also biochemically shown to be required for proximal *cis*-cyclopropanation of α -mycolates [59], a function previously attributed to the *cmal2*-gene product [91]. In *M. tuberculosis*, *cmal2*-gene deletion studies have demonstrated that CMAS-2 (or CmaA2) is the *trans*-cyclopropane synthase for both the methoxy- and ketomycolates [22].

A cluster of four genes encoding homologous methyltransferases, termed *mma1* to *mma4* (methoxy mycolic acid), have been identified by DNA hybridization with a *cmal* probe [60]. Heterologous expression of this *M. tuberculosis* gene cluster conferred upon *M. smegmatis* the ability to synthesize methoxymycolates (*M. smegmatis* synthesizes α -, β -, and epoxy-mycolic acids). The *mma4* gene product (MMAS-4) catalyzes an unusual SAM-dependent transfer to the distal *cis*-olefin, yielding a secondary alcohol with an adjacent methyl group. MMAS-3 catalyzes O-methylation of the hydroxyl group resulting in the corresponding methyl ether, whereas MMAS-2 introduces a *cis*-cyclopropane in the proximal position of the methoxymycolate [60, 92]. Overexpression of MMAS-1 in *M. tuberculosis* resulted in both an increase in the amount of mycolate containing *trans*-cyclopropane and the appearance of mycolates containing *trans*-olefin with an allylic methyl branch in the proximal position [57]. Thus the function of MMAS-1 seems to be the isomerization of the proximal *cis* double bond to a *trans* double bond in the oxygenated meromycolate precursor with simultaneous introduction of an adjacent methyl branch, suggesting that the *trans*-olefinic mycolate with an allylic methyl branch is the precursor of

the *trans*-cyclopropane with an adjacent methyl branch. In addition, MMAS-1 action appears to be only on oxygenated mycolic acids, since MMAS-1 expression showed no *trans*-cyclopropane or *trans*-olefin functions present in β -mycolates. Interestingly, labeling experiments of mycolic acids produced by *M. tuberculosis* and *M. smegmatis* overexpressing the *mma1* gene product have shown that the methyl transferase MMAS-1 of *M. smegmatis* operates in an inverted manner to that of *M. tuberculosis* [93]. UmaA1 is another member of the *M. tuberculosis* mycolic acid methyl transferases family which has no known function at present [22]. It should be pointed out that β -mycolate, which contains double bonds or cyclopropane rings as its only functional groups, is not a precursor of the more complex oxygenated mycolates [94]. Thus, any common intermediate must be at an early stage in biosynthesis. A scheme for functionalization of the meromycolate chain is shown in Fig. (6). A mechanism has been proposed by Barry III *et al.* [46], which involves the generation of an unstable cationic intermediate bound to the active site of the methyl transferase that can decompose in a number of ways to yield the observed meromycolate products. Based on the proposal of cation- π interaction having a stabilizing effect on cationic intermediates [95], a conserved tryptophan residue located in the SAM-binding motif [22] has been proposed to be involved in stabilization of the cationic intermediate of methyltransferase mechanism of action [46]. This carbonium ion can then be deprotonated to form a cyclopropane ring or a *trans* olefin with an adjacent methyl branch. The *trans* olefin could be the substrate for a second SAM-dependent methylation to form the *trans* cyclopropane series seen in the oxygenated mycolates of *M. tuberculosis*. Hydration of the carbonium intermediate can form the hydroxymethyl mycolate precursor to keto and methoxymycolates.

Antibodies raised against AcpM, which has recently been shown to be the carrier of full-length meromycolate produced by a unique type II fatty acid synthase system, inhibited the methyl transfer reactions of these acids in *M. tuberculosis* cell extracts [96]. These results suggest that mycolate modification reactions occur simultaneously to elongation of the AcpM-bound meromycolate chain.

Relatively little is known about either the enzymology or the genetics of oxidative modifications of the mycobacterial meromycolic acid chain. The biosynthesis of the ketomycolate and methoxymycolate appears to be intimately linked through the common hydroxymycolate precursor, which is the product of the H₂O transfer to the distal olefin yielding a secondary alcohol with an adjacent methyl group incorporated by MMAS-4 enzyme (Fig. 6). It has been suggested that there may be a specific oxidase in *M. tuberculosis* that catalyzes the conversion of hydroxymycolate into ketomycolate, even though the ability to easily alter keto and methoxy mycolate production suggests either that this oxidase is not very active or is poorly expressed [46].

2.4.4. The Final Claisen-Type Condensation

The Claisen-type condensation between the meromycolate chain and the β -branch, followed by reduction are the final steps in mycolic acid biosynthesis. This Claisen-type

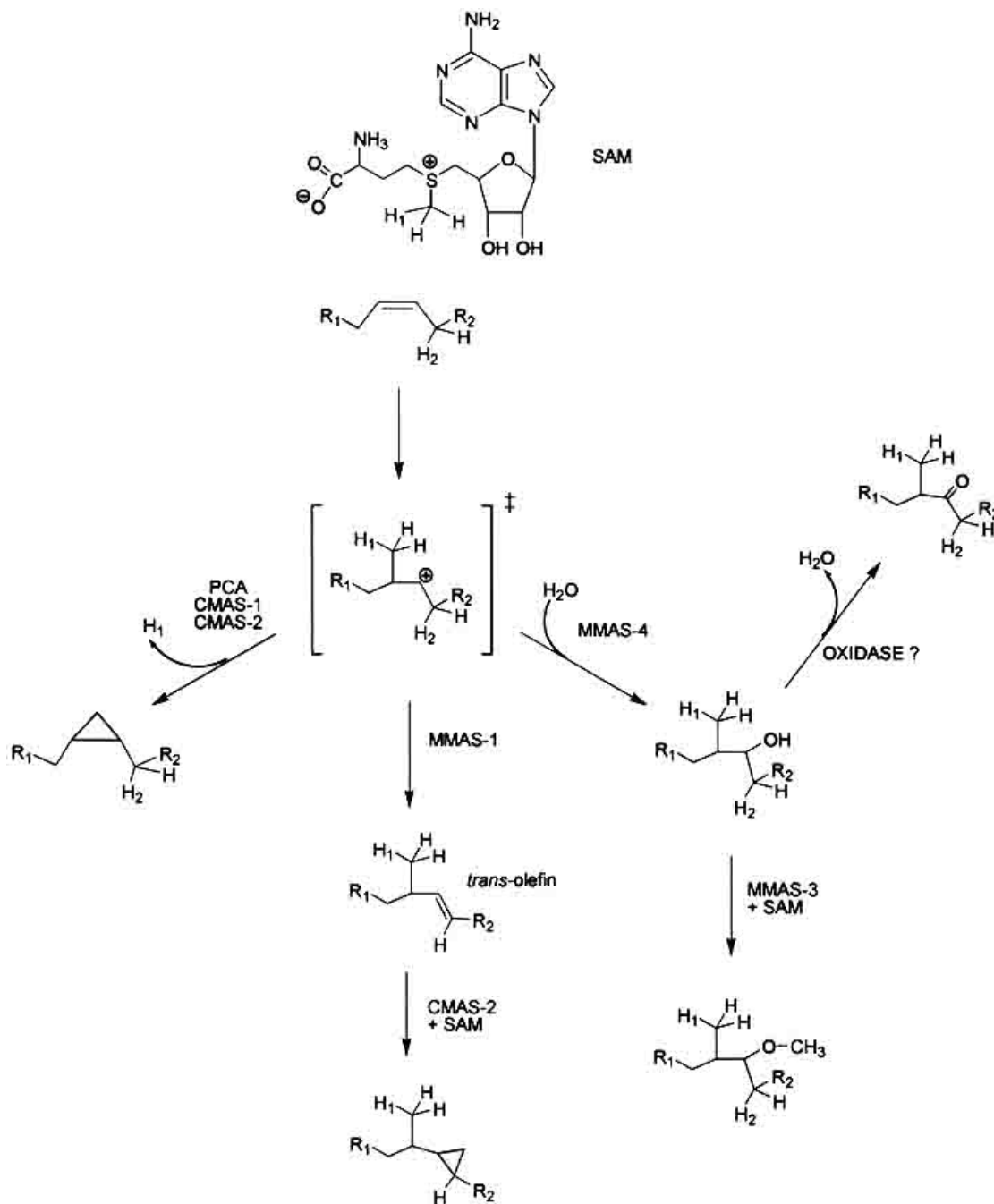


Fig. (6). A scheme for functionalization of the meromycolate chain.

condensation involves the carbon of one (thio)ester molecule and the carbonyl group of another leading to a substitution reaction. A thioester (R₁-CO-S-R₂) resonance structure with a negative charge on the carbonyl oxygen and a carbonyl carbocation makes the carbonyl group more susceptible to nucleophilic attack. Since the -S-R₂ group is a good leaving group, these two factors make thiol esters effective acylating agents. Mechanistically, the -carbon of the -branch would make a nucleophilic attack on the meromycolate carbonyl carbon, which would act as an electrophile, followed by departure of the -S-ACP group, which

was previously bound to the meromycolate chain, and reduction of the carbonyl oxygen (Fig. 7).

The Claisen-type condensation of mycolic acids in *M. tuberculosis* has been postulated to involve the interaction of an activated C₅₆-meromycolate chain with an activated C₂₆-branch, yielding a 3-oxomycolate intermediate that is subsequently reduced to form the mature mycolic acid prior to its transfer to the various cell wall components [97-98]. The Claisen-type condensation step has been demonstrated in corynebacteria, which synthesize a smaller version of

granuloma formation when injected into mice [102] in addition to the inhibition of phospholipid vesicles fusion [103]. A protein from *M. smegmatis* was purified to homogeneity and shown to catalyze the exchange of a mycolyl group between trehalose, trehalose 6-monomycolate (TMM) and TDM [104]. More recently, an enzyme from *M. tuberculosis* was purified and shown to catalyze mycolyl transfer between TDM and TMM [105]. The N-terminal amino acid sequence of the purified mycolyl transferase was, unexpectedly, similar to that of the antigen85 (Ag85) complex (the A, B, and C components of *M. tuberculosis* Ag85), which are the dominant exported proteins and protective antigens of *M. tuberculosis*, and have been implicated in disease pathogenesis through their fibronectin-binding capacity and involvement in complement receptor-mediated phagocytosis [106]. The abundance of TMM and TDM in the cell wall indicates that these molecules are important to the integrity of the cellular envelope and that TMM and TDM biosynthesis is a viable drug target. Accordingly, inhibition of all three members of Ag85 complex activity by 6-azido-6-deoxy- α -D-trehalose resulted in complete suppression of *M. aurum* growth [105]. Moreover, a *M. tuberculosis* mutant lacking a functional Ag85C gene showed a 40% decrease in the amount of mycolate transferred to the cell wall with no change in the types of mycolates, suggesting that the enzyme was not specific for a given type of mycolate [107]. In addition, the mutant strain showed increased diffusion of chenodeoxycholate and glycerol through the cell wall, indicating that the cell wall-linked mycolate layer represents a barrier for the diffusion of small hydrophobic and hydrophilic molecules, even though no effect on growth could be observed in murine macrophages. Disruption of the 85A component of antigen85 in *M. tuberculosis* H37Rv resulted in a mutant that lacked the ability to grow in macrophage cell lines and in mice [108]. These results suggest that although both antigens 85A and 85C share mycolyltransferase activities, 85A may also be involved in mycobacterial virulence. The crystal structure of Ag85C solved to 1.5Å resolution has been proposed to allow for the rational design of peptide mimetics of Ag85 proteins that stimulate antibody production against the fibronectin binding site, which would impair one of the gateways *M. tuberculosis* uses to enter host macrophages [109]. A further homologue to Ag85 complex, Ag85C', has been identified in the genome sequence of *M. tuberculosis* coded by *fbpC1* gene [69]. However, further studies are required to show whether this new protein exhibits mycolyltransferase activity and/or the reason behind the apparent redundancy.

3. DRUGS THAT INHIBIT MYCOLIC ACID BIOSYNTHESIS

Owing to the pivotal role of mycolic acids for the mycobacterial cell wall, inhibitors of the proteins involved in the biosynthetic pathway of mycolates should have potential antimycobacterial activity. A few examples of drugs used in the treatment of tuberculosis that are known to inhibit mycolic acid biosynthesis as well as compounds that inhibit enzyme components of this pathway are given in this section.

3.1. Isoniazid (INH)

Isoniazid (INH, isonicotinic acid hydrazide, Fig. 9) was first reported to be effective in the treatment of tuberculosis in 1952 [110]. Both *M. tuberculosis* and *M. bovis* BCG are susceptible to isoniazid in the range of 0.02-0.2 µg/ml [32]. Isoniazid is the oldest synthetic antitubercular, and the most prescribed drug for active infection and prophylaxis. Strains of *M. tuberculosis* resistant to isoniazid were reported shortly after the introduction of this drug [16]. A major target for isoniazid has recently been shown to be the product of the *inhA* structural gene, which is located downstream of an ORF coding for a β -ketoacyl reductase (MabA) in the *inhA* operon [77]. InhA was subsequently identified as an NADH-dependent enoyl-ACP (acyl carrier protein) reductase [111]. This enzyme exhibits specificity for long-chain (C₁₈ > C₁₆) enoyl thioester substrates which is consistent with its suggested involvement in mycolic acid biosynthesis [111].

Some of the earliest studies of isoniazid-resistant clinical isolates of *M. tuberculosis* noted a correlation between resistance and attenuated catalase-peroxidase activity [112-113]. The *katG* gene, which encodes the mycobacterial catalase-peroxidase enzyme, was cloned from *M. tuberculosis* [114]. Introduction of the wild-type *katG* gene into either *M. smegmatis* or *E. coli*, which are naturally less susceptible to isoniazid, or isoniazid-resistant strains of *M. tuberculosis* enhanced the sensitivity of these organisms to the drug [115]. Deletions of, or missense mutations in, the *katG* gene have been associated with decreased susceptibility to isoniazid in approximately 50% of clinical isolates of *M. tuberculosis* [116-119, 9]. Peroxidases have been shown to react with isoniazid to generate a number of oxidized products similar to those observed *in vivo* [120]. A mechanism for the action of isoniazid has been proposed that involves the conversion of isoniazid, by the mycobacterial *katG*-encoded catalase-peroxidase, into a number of electrophilic intermediates capable of either oxidizing or acylating amino acid residues in proteins [121]. The amino acid residue of the enoyl reductase modified by the electrophilic drug metabolite was proposed to be cysteine-243 [122], although subsequent studies suggested this not to be likely [123]. Although isoniazid does not bind to the *inhA*-encoded enoyl reductase [111], the catalase-peroxidase-activated drug intermediate binds to, and inhibits the reductase activity in the presence of NADH or NAD⁺ [122].

The mycobacterial *inhA*-encoded enoyl reductase shares 40% amino acid sequence identity with the *E. coli* *envM*-encoded enoyl reductase [77]. The serine residue at position 94 of the enoyl reductase is substituted by alanine in isoniazid-resistant mutants of *M. smegmatis* [77] and the glycine 93 residue of the enoyl reductase is substituted by a serine residue in diazaborine-resistant mutants of *E. coli* [124] and both residues are located in the nucleotide binding site. The three dimensional structure has been determined for the ternary complex of the wild-type enoyl reductase of *M. tuberculosis*, NADH and activated isoniazid [125]. In this complex, generated by the *in vitro* reaction of manganese, isoniazid, oxygen, NADH and InhA [126], the acylpyridine fragment of isoniazid is covalently attached to the C4

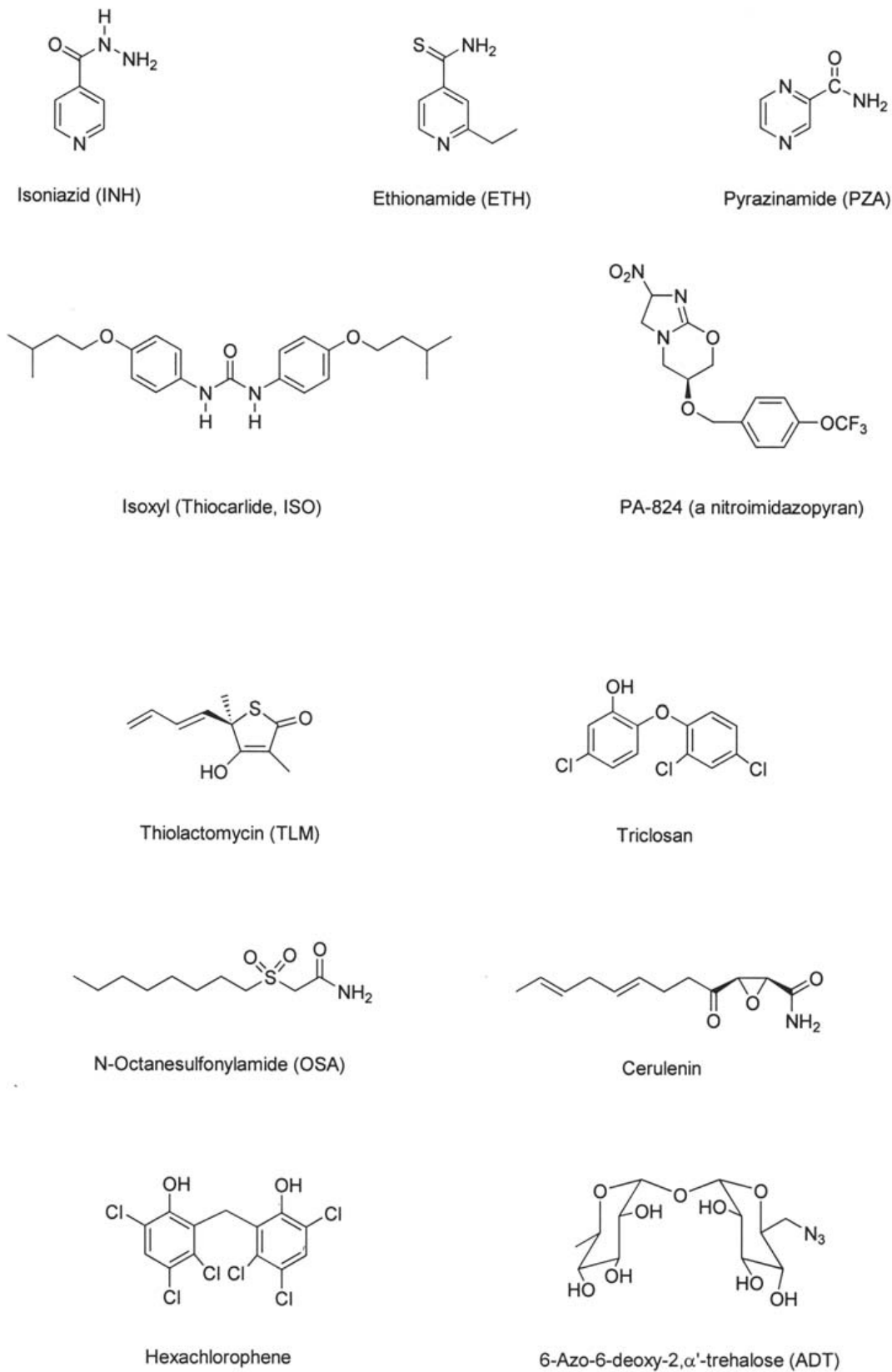


Fig. (9). Structures of several antimycobacterial agents which inhibit enzymes of the mycolic acid biosynthetic pathway.

position of NADH. This isonicotinyl-NAD⁺ adduct has been characterized spectroscopically and shown to bind to InhA with a dissociation constant value lower than 0.4 nM [127]. Catalase-peroxidase-activated isoniazid intermediates were shown to bind exclusively to the enoyl reductase-NADH binary complex [128]. More recently, however, efficient formation of the isonicotinyl-NAD⁺ adduct could be achieved in the absence of InhA [129]. Moreover, although catalase-peroxidase KatG is required for oxidation of isoniazid to form the isonicotinyl radical, KatG is not required for addition of the free radical compound to NAD⁺ to form the adduct [129]. It is likely that the *in vivo* oxidation of isoniazid by the catalase-peroxidase to generate the isonicotinyl intermediate that reacts with NAD(H) would result in binding to, and inhibition of, the enoyl reductase. Isoniazid is, therefore, a pro-drug, that upon formation of an isonicotinyl-NAD⁺ adduct inhibits the *M. tuberculosis* enoyl reductase, resulting in reduction of mycolic acid synthesis.

Mutations in the *inhA* structural gene and in the *inhA* locus promoter region have been associated with isoniazid resistance [9]. The *in vitro* generated S94A mutant of the enoyl reductase [77] displayed a higher K_m value for NADH [111]. Comparison of the crystal structure of binary complex of wild-type and the mutant enoyl reductases with NADH showed that disruption of a hydrogen bond network could account for the observed higher K_m value [130]. INH-resistant clinical isolates of *M. tuberculosis* harboring *inhA*-structural gene mutations were shown to have higher K_m values for NADH and increased dissociation constants for NADH from InhA-NADH binary complexes [131]. Consistent with these results, all mutations in the enoyl reductase conferring INH resistance were located in the nucleotide binding site. More recently, inactivation of the *M. tuberculosis* *inhA*-encoded enoyl reductase and INH treatment resulted in similar morphological changes to the mycobacterial cell wall, leading to cell lysis [79]. These results strongly suggest that InhA is the target for isoniazid action.

The mechanism of action of isoniazid is complex, as mutations in at least five different genes (*katG*, *inhA*, *ahpC*, *kasA*, and *ndh*) have been found to correlate with isoniazid resistance. It has been suggested that increased expression of an *ahpC*-like component of the alkyl hydroperoxidase complex can confer isoniazid resistance to some *M. tuberculosis* strains lacking mutations in *katG* and *inhA* genes [132-133]. However, overexpression of this gene in *M. tuberculosis* results in no significant isoniazid resistance in the absence of additional *katG* mutations [134]. Moreover, the virtual absence of mutations in the *ahpC* coding region or mutations in the *ahpC* promoter region in isoniazid-resistant clinical isolates with lesions at the *katG* or *inhA* loci [135-136] suggest a secondary role of the *ahpC* gene in clinical isoniazid resistance. Mutations mapping to *ndh*, which encodes for an NADH dehydrogenase, have been shown to be involved in the mechanism of INH resistance in *M. smegmatis* [137]. These mutations had pleiotropic phenotypes including coresistance to INH and ethionamide (ETH), temperature sensitivity for growth, and auxotrophies as a result of increased NADH/NAD⁺ ratios in the cell [137]. Although the significance of these results to the mechanism

of action and/or resistance to INH in *M. tuberculosis* awaits further experimental evidence, it has been proposed that NAD⁺-binding proteins might contribute to isoniazid resistance [129]. Since accumulation of C_{26:0} is observed following INH treatment of *M. tuberculosis* [138] instead of monounsaturated intermediates due to InhA inhibition by INH, the proposal of *inhA*-encoded enoyl reductase as the primary target in *M. tuberculosis* has been questioned [139, 72]. Formation of a covalent complex in *M. tuberculosis* between radioactive INH, AcpM (the ACP of FAS-II) and the *kasA*-encoded β -ketoacyl synthase (KasA), as well as mutations in the KasA protein in INH-resistant clinical isolates that lacked other mutations associated with resistance to the drug have prompted the proposal of KasA as an alternative INH target [72]. Two INH-resistant clinical isolates of *M. tuberculosis* harboring *kasA* gene mutations were shown to have no mutations in other loci involved in INH resistance: *katG*, *inhA*, and *ahpC* [72]. However, binding to a protein does not establish inhibition. Moreover, *inhA*-operon promoter region and *ndh* structural gene mutations have also been correlated with INH resistance in mycobacteria. In addition, two clinical isolates of *M. tuberculosis* harboring G269S KasA mutant, which is thought to be associated with INH resistance [72], were shown to be INH sensitive [76]. A drug target for a bactericidal drug has been suggested to be an enzyme of the bacterium (i) that binds the drug, (ii) that is inhibited by the drug, and (iii) whose inhibition induces the death of the bacterium [79]. Based on these criteria, genetic and biochemical results point to InhA as the bona fide target for INH in *M. tuberculosis*.

The manganese-dependent peroxidase activity of KatG catalyzes the conversion of Mn²⁺ to Mn³⁺ [140]. Since KatG is required for oxidation of isoniazid to form the isonicotinyl radical that reacts with NAD⁺ to form the bisubstrate inhibitor of InhA, a method for nonenzymatic activation of isoniazid by Mn³⁺-pyrophosphate has been proposed to allow for simple chemical synthesis of various INH derivatives [141]. These derivatives could be assessed as *M. tuberculosis* enoyl reductase inhibitors, including the INH-resistant clinical isolates of *M. tuberculosis* harboring *inhA*-structural gene mutations and KatG mutants, and tested as antitubercular agents. Moreover, the crystal structure determination of *M. tuberculosis* InhA in complex with NAD⁺ and a C₁₆ fatty acyl substrate [142] should facilitate the design of inhibitors to be tested as antitubercular drugs. Ultimately, the knowledge that InhA inactivation induces the lysis of mycobacteria suggests that all the FAS-II enzymes (β -ketoacyl-ACP reductase, β -hydroxyacyl-ACP dehydratase, and KasA and KasB) may represent potential targets for the development of novel antimycobacterial agents.

3.2. Ethionamide (ETH)

The discovery of the potent antitubercular activity of isoniazid was followed by a search for structural analogs, of which ethionamide (ETH, Fig. 9) was found to be the most potent. ETH is an important component of second-line therapy for the treatment of multidrug-resistant tuberculosis. INH and ETH have almost identical inhibitory effects on

mycolic acid biosynthesis [143-144]. The *inhA*-encoded enoyl reductase, which is the target for isoniazid, had been suggested to also be a target of ethionamide because a single mutation in the *inhA* gene conferred resistance to INH and ETH [77]. However, the one notable distinction between the actions of ethionamide and isoniazid is the lack of cross-resistance: strains resistant to ethionamide are still sensitive to isoniazid while strains resistant to isoniazid show slightly increased sensitivity to ethionamide [143]. Moreover, there are some significant differences in their activation mechanisms [122]. Ethionamide, like isoniazid, is a prodrug that is metabolized to its *S*-oxide which has been proposed to be capable of undergoing addition reactions to nucleophilic protein side chains [122]. Whereas mutations in the enoyl reductase generate resistance to both isoniazid and ethionamide [77], isoniazid-resistant strains containing mutations in the catalase-peroxidase gene remain susceptible to ethionamide. It is, therefore, tempting to speculate the presence of a second enzyme that activates ethionamide in mycobacteria to convert this prodrug into a form capable of binding to, and presumably inhibiting, an enzyme of mycolic acid biosynthetic pathway. A KatG-independent mechanism of ETH activation has recently been reported [145]. The product of *M. tuberculosis* Rv3855 gene, named *ethR*, conferred resistance to ETH, but not to INH (or isoxyl, thiolactomycin, erythromycin, crystal violet, and streptomycin), when overexpressed in either *M. smegmatis*, *M. bovis* BCG, or *M. tuberculosis* [145]. *M. bovis* BCG grown in the presence of ETH displayed specific inhibition of mycolic acid synthesis. Genetic and biochemical results suggest that EthR is a transcriptional regulator that is not directly implicated in mycolic acid biosynthesis, but plays an important role in the regulation of a second ORF (*Rv3854c*, termed *ethA*), which is responsible for the activation of ETH. EthA is a putative monooxygenase that has been suggested to be implicated in the activation of ETH. Ethionamide inhibited all classes of mycolic acids in *M. bovis*. On the other hand, ETH treatment of *M. bovis* expressing EthR showed no effect on the synthesis of mycolic acids (ETH resistance). Construction of an *M. bovis* BCG *ethR* gene knock out mutant confirmed EthR to be specific for ETH resistance. Overexpression of *ethA* gene in *M. smegmatis* resulted in hypersensitivity to ETH and a dramatic inhibition of mycolic acid synthesis at low concentrations of ETH as compared to wild-type *M. smegmatis*. Thus, EthR is a transcriptional repressor that regulates the production of EthA, which is an ETH activator. Accordingly, repression of *ethA* gene expression results in impaired ETH activation and ensuing drug resistance. Any agent able to block EthR may favor the production of EthA and lead to the activation of substantial amounts of ETH, thereby increasing the sensitivity of the bacilli to ETH.

Examination of metabolites formed by whole cells of *M. tuberculosis* using radiolabeled ETH has allowed a proposal for a chemical mechanism for ETH activation [146]. ETH is activated via a thioamide *S*-oxide in an oxygen-dependent process by EthA action, yielding (2-ethyl-pyridin-4-yl)methanol intermediate that is incorporated into macromolecules. Wild-type *inhA* gene conferred INH and ETH resistance when transferred on a multicopy plasmid vector to *M. smegmatis* and *M. bovis* BCG [77] and oxidation of ETH by horseradish peroxidase and H₂O₂ led to

inactivation of InhA [122]. Although these results suggest that both ETH and INH target the *inhA*-encoded enoyl reductase, further work is needed to establish the mechanism of inhibition of ETH intermediates on InhA enzyme activity.

ETH is but an example of a thiocarbonyl-containing antitubercular drug: thioacetazone and thiocarlide (isoxyl) are two other examples among the second-line tuberculosis therapeutics that might be activated by EthA-catalyzed *S*-oxidation. Consistent with this proposal, aminoacid-altering mutations, deletion and addition of 1 nucleotide in *ethA* gene were found in *M. tuberculosis* clinical isolates resistant to thiocarbonyl-containing drugs, thus establishing cross-resistance between these compounds [146].

3.3. Pyrazinamide (PZA)

The weak antitubercular activity of nicotinamide began a search for more powerful analogs and resulted in the discovery of pyrazinamide (PZA, Fig. 9). Although shown to have substantial activity in 1952 [15], pyrazinamide was not used extensively in the treatment of tuberculosis until the mid-1980s. Pyrazinamide is now recommended in essentially every combination therapy for the treatment of the disease because of its strong synergistic, and accelerating, effect in combination with isoniazid and rifampicin [31, 32]. The introduction of pyrazinamide combination chemotherapy allowed treatment to be reduced from 9-12 months to 6 months. It appears to be most effective in the first 2 months of therapy with modest additional incremental effects, and is often discontinued after 2 months of therapy, with continued treatment with other agents [147]. The MIC for pyrazinamide varies from 8 to 60 µg/ml depending on the assay method and media [32]. Pyrazinamide does not appear to be bactericidal, even at concentrations significantly greater than its MIC, and its *in vitro* effect is presently termed "sterilizing" [148] to distinguish it from the effects of other drugs such as rifampicin. The drug is most active against *M. tuberculosis* at acidic pH values (below 6) but has little, if any, activity at higher pH. Pyrazinamide appears to exhibit its *in vivo* sterilizing activity on the semidormant populations of *M. tuberculosis* in acidic intracellular compartments, such as the macrophage phagolysosomes [149]. However, the intracellular pH surrounding live *M. tuberculosis* was shown not to be acidic [150]. These puzzling experimental observations will require further studies to determine the factors necessary for pyrazinamide activation. Although pyrazinamide is effective against *M. tuberculosis* infection, the compound is not effective in the treatment of other mycobacterial infections, including *M. bovis* and fast-growing mycobacteria.

Pyrazinamide is a pro-drug that requires amide hydrolysis by mycobacterial pyrazinamidase (a nicotinamidase) to pyrazinoic acid, the presumed active form [151-152]. This proposal is supported by the finding that pyrazinoic acid is active *in vitro* against pyrazinamide- and nicotinamide-resistant strains of *M. tuberculosis*, as well as the naturally resistant *M. bovis*, which lacks pyrazinamidase activity [151]. Further support for this proposal is the loss of pyrazinamidase activity in strains of *M. tuberculosis* resistant to pyrazinamide [153]. A series of pyrazinoic acid

esters have been shown to exhibit better *in vitro* activity against *M. tuberculosis* than pyrazinamide [154], also supporting the view that the acid form is the active component. The *M. tuberculosis* gene that encodes pyrazinamidase and nicotinamidase activities (*pncA*) has been cloned and sequenced [155]. A remarkable experimental observation was that *M. bovis* species, known to lack pyrazinamidase activity, has a *pncA* gene. The defective pyrazinamidase in three *M. bovis* and three BCG strains studied is due to the same point mutation in the *pncA* gene at position 169 (C G), which caused the substitution of histidine for aspartic acid at position 57 [155]. The genetic basis of pyrazinamide resistance, and the correlation between pyrazinamide resistance and *pncA* mutation, has been reviewed elsewhere [9].

5-chloro-pyrazinamide (5-Cl-PZA), a pyrazinamide analogue, was shown to possess *in vitro* activity against *M. tuberculosis*, *M. bovis*, and several nontuberculous mycobacteria [156]. More recently, a genetic approach was used to identify the target of 5-Cl-PZA [157]. The *M. tuberculosis* *fasI*-encoding multifunctional fatty acid synthase type I (FAS-I) gene, present in a multi-copy vector, conferred resistance to 5-chloro-pyrazinamide in *M. smegmatis*. Moreover, 5-chloro-pyrazinamide treatment of *M. smegmatis* and *M. tuberculosis* resulted in inhibition of [¹⁴C]acetate incorporation into the organic solvent soluble fraction containing primarily C₁₆ - C_{24/26} fatty acids. Similarly, PZA inhibited [¹⁴C]acetate incorporation into fatty acid biosynthesis by FAS-I, whereas [¹⁴C]palmitate incorporation into mycolic acids by FAS-II was unaffected. These results strongly suggest that PZA inhibits only FAS-I despite the similarities between the chemical reactions catalyzed by FAS-I and FAS-II in lipid synthesis. Although pyrazinamide was shown not to inhibit mycolic acid synthesis, it was deemed pertinent its inclusion in this manuscript because FAS-I provides precursors for FAS-II fatty acid elongation to long-chain mycolic acids. Analogs of PZA may have an additional activity against extracellular bacilli because they are active in neutral pH [157], while PZA is only active at acidic pH. More potent inhibitors of FAS-I may not only shorten the length of tuberculosis treatment but also provide more options for treatment against drug resistant *M. tuberculosis* as well as *M. avium*.

3.4. Isoxyl (Thiocarlide, ISO)

Isoxyl (ISO, 4,4'-diisoamyloxydiphenylthiourea, 4,4'-diisoamyloxythiocarbanilide, thiocarlide, Fig. 9) is a diaryl thiourea that was used for the clinical treatment of tuberculosis in the 1960s [143]. Isoxyl, like INH and ETH, was shown to inhibit fatty acid and mycolic acid synthesis in *M. bovis* during a 6 hour exposure using 10 µg mL⁻¹, which is the minimum inhibitory concentration [158]. The minimum inhibitory concentration determined on 7H11 agar medium of ISO for *M. tuberculosis* was determined to be 2 µg mL⁻¹ and of a number of ISO derivatives substituted in the *para* and *para'* positions by alkyl, alkoxy, or sulfur functional groups were in the range of < 0.1 to > 20 µg mL⁻¹ [159]. Among these derivatives, the butyl derivative of ISO possessed a low MIC value (0.1 to 0.5 µg mL⁻¹), and replacement of oxygen with sulfur in the side chain(s)

provided extremely high antibacterial activity against *M. tuberculosis*. Since the thiourea nucleus of ISO is required for antimycobacterial activity, substitutions solely to the side chain(s) may render even more potent agents. The use of *in vivo* [1,2-¹⁴C]acetate pulse-labeling of *M. tuberculosis* H37Rv in the presence of ISO at its MIC resulted in inhibition of synthesis of both short-chain fatty acids and mycolic acids (-mycolates, 91.6%; methoxymycolate, 94.3%; and ketomycolate 91.1%), unlike INH and ETH that are specific for mycolic acid synthesis. In addition, ISO and its butyl derivative demonstrated cidal action in an *in vitro* macrophage model [159]. The target(s) for ISO action is(are) not known at the present, and they may lie at the points shared in the synthesis by both short-chain fatty acids and mycolic acids. An understanding of the mode of action of ISO should aid not only in the development of new antimycobacterial agents but also unveil new fatty acid biosynthesis and mycolic acid biosynthesis genes.

3.5. Thiolactomycin (TLM)

Thiolactomycin [TLM, (4S)(2E,5E)-2,4,6-trimethyl-3-hydroxy-2,5,7-octatriene-thiolide, Fig. 9] is a unique thiolactone antibiotic produced as a secondary metabolite of the soil fungi *Nocardia* sp. [160-161] which exhibits potent chemotherapeutic activity in mice [162] and *in vitro* activity against a broad spectrum of pathogens, including Gram-positive cocci and enteric, acid-fast, and anaerobic bacteria [163]. TLM inhibits bacterial and plant FAS-II but not mammalian and yeast FAS-I [164]. Recent findings have shown that TLM possesses *in vivo* antimycobacterial activity against *M. smegmatis* and *M. tuberculosis* [165]. In addition, killing of viable intracellular *M. tuberculosis* in a dose-dependent manner has also been shown by an *in vitro* murine macrophage model. TLM was shown to inhibit the synthesis of both fatty acids and mycolic acids. However, synthesis of the shorter-chain -mycolates of *M. smegmatis* was not inhibited by TLM, whereas synthesis of characteristic longer-chain -mycolates and epoxy mycolates was almost completely inhibited. The use of *M. smegmatis* cell extracts demonstrated that TLM specifically inhibited the mycobacterial FAS-II but not the multifunctional FAS-I [165]. Overexpression of *M. tuberculosis* KasA and KasB (-ketoacyl-ACP synthases) individually or co-expression of both enzymes in *M. bovis* BCG resulted in increased *in vivo* and *in vitro* resistance against TLM, suggesting that TLM targets both KasA and KasB [76]. In addition, synthesis of several TLM derivatives led to compounds more potent both *in vitro* against fatty acid and mycolic acid biosynthesis and *in vivo* against *M. tuberculosis* [76]. More recently, KasA and KasB were shown to catalyze the condensation of acyl-ACP with malonyl-ACP, as opposed to acyl-CoAs, and to have specificity for long-chain acyl-ACPs, indicating that they function as elongation enzymes [74]. Consistent with *in vivo* results [76], kinetic measurements demonstrated that KasA and KasB are inhibited by TLM [74].

Although TLM has *in vivo* efficacy in a mouse model, exhibiting rapid absorption and wide distribution into tissues; synthesis and stability issues have prevented the compound from making significant progress as a therapeutic agent [166]. Nevertheless, TLM, or its analogues, remains an

interesting antibiotic because it shares little or no cross-resistance with other classes of antibiotics, thereby having added therapeutic value since it may be used to treat patients with resistant strains of *M. tuberculosis*. The inhibitory *in vitro* effects of TLM on KasA and KasB activity [74] and TLM resistance upon KasA and KasB overexpression [76] provide compelling evidence that one or both of these enzymes may lead to the compounds with antitubercular activities. The recent crystal structure determination of the binary complex formed between *E. coli* β -ketoacyl-ACP synthase I and TLM [167] should provide a structural framework for the rational design of new chemotherapeutic agents.

3.6. Triclosan

Triclosan (2,4,4'-trichloro-2'-hydroxyphenyl ether; Fig. 9) was originally patented as a herbicide, but was soon realized to possess a broad spectrum antibacterial action [166]. Triclosan has been used in a wide variety of medical and surgical devices, as a handwash for both laboratories and hospital surgical/medical units, and as an antiseptic due to its favorable safety profile [168]. It is also used in numerous household products such as toothpaste, mouthwash, deodorant, hand soaps, and cutting boards. Antiseptics are chemical agents that destroy or inhibit microorganisms thus limiting or preventing the harmful results of infection. They are applied principally to cutaneous and mucous surfaces and infected wounds; however, applications to wounds are not recommended because they retard the healing and may injure the tissues [169]. Disinfectants are chemical agents that destroy microorganisms, but not usually bacterial spores; they do not necessarily kill all microorganisms, but reduce them to a level that is harmful neither to health nor the quality of perishable goods. They are employed to treat inanimate objects and materials (including excreta) and may also be applied to agents used to treat skin and other body membranes and cavities. Triclosan is not an antibiotic. Instead, it is a chlorophor acting as a broad spectrum (bacteria, fungi, and viruses) antiseptic and disinfectant.

Isolation of *E. coli* resistant to triclosan and mapping the resistance in one strain to the enoyl-ACP reductase-encoding *fabI* gene provided the first indication that triclosan inhibits lipid synthesis [170]. Isolation of a S93V variant of *E. coli* FabI resistant to triclosan, and the demonstration that the purified mutant protein is resistant to triclosan *in vitro*, whereas the wild-type protein is sensitive, provided direct evidence that FabI is the biochemical target of this drug [171]. Binding assays established that triclosan increases the affinity of FabI for NAD⁺ by approximately 4 orders of magnitude, suggesting formation of a stable FabI-NAD⁺-triclosan complex [172]. Kinetic studies showed that triclosan is a slow onset inhibitor of *E. coli* FabI and that inhibition is promoted by NAD⁺ with an estimated value of 20-40 pM for the dissociation constant of triclosan from enzyme-NAD⁺ complex [173]. Determination of the X-ray crystal structure of *E. coli* FabI-NAD⁺-triclosan ternary complex showed that the diphenyl ether of triclosan adopted a conformation with a dihedral angle of about 90° between the two phenyl rings and that extensive π - π interactions is formed between the phenyl ring of triclosan and the nicotinamide ring [172, 174]. Analysis of the FabI-NAD⁺-

triclosan structure revealed that the aminoacids altered by mutations in the *E. coli* *FabI* gene (G93V, M159T, and F203L) that confer resistance to triclosan make direct contacts with triclosan [174], thus providing an understanding at molecular level of triclosan mode of action and resistance. It had been assumed that FabI was the only bacterial enoyl-ACP reductase; however, FabK and FabL are recently discovered enoyl-ACP reductase isoforms from *S. pneumoniae* and *B. subtilis* [175-176]. FabK is resistant to triclosan [175], while FabL is reversibly inhibited by this drug [176]. The finding that *E. coli* cells expressing either FabK or FabL have unimpaired growth in the presence of triclosan provided evidence that FabI is the only target of triclosan in *E. coli* and that triclosan does not possess membrane disruptive properties.

In *M. smegmatis* strains selected for resistance to triclosan, mutations were mapped to *inhA* structural gene, the *fabI* homologue, thus providing evidence that mycobacterial enoyl-ACP reductase is a target for triclosan [177]. The aminoacid substitutions in M161V, A124V, and M103T *InhA* mutants resistant to triclosan are located in the active site of the enzyme [177]. *In vitro* inhibition studies demonstrated that triclosan is a submicromolar uncompetitive inhibitor of *M. tuberculosis* enoyl-ACP reductase [178]. Consistent with these results, the M161V and A124V *InhA* mutations resulted in reduced sensitivity to triclosan. The S94A *InhA* mutant, selected for isoniazid resistance [77], was shown to be cross-resistant to triclosan [177]. However, despite the 95% identity of *M. smegmatis* and *M. tuberculosis* *InhA* enzymes [166], triclosan was shown not to inhibit the *M. tuberculosis* S94A *InhA* mutant enzyme [178]. Moreover, no enzyme activity could be detected for the *M. tuberculosis* M103T *InhA* mutant. Thus, these results suggest that M103T and S94A mutations cannot account for the triclosan resistance observed in *M. smegmatis*.

Since, unlike isoniazid, the maximum obtainable resistance of *M. smegmatis* to triclosan reaches a threshold, it has been suggested that *M. smegmatis* may have a second target for this drug [166]. *M. smegmatis* is susceptible to triclosan, whereas *M. tuberculosis* is resistant to this drug [179]. The four residues identified in *M. smegmatis* *InhA* that confer triclosan resistance (S94, M103, A124, and M161) are conserved in *M. tuberculosis* [177]. They would therefore not represent residues unique to *M. tuberculosis* *InhA* that could account for the intrinsic resistance to triclosan. *S. pneumoniae* FabK is an enoyl-ACP reductase isoform resistant to triclosan [175]. An open reading frame encoding a FabK homologue sharing 33% protein sequence identity has been identified in *M. tuberculosis* [175], which could account for triclosan resistance in this pathogen. Mechanisms unrelated to *InhA* may play a role in triclosan resistance in *M. tuberculosis*, such as the activity of an endogenous efflux pump(s) analogous to those which operate on triclosan in *E. coli* [180] and *P. aeruginosa* [181]. Mutations leading to triclosan resistance (M161V and A124V) were shown to reduce the sensitivity of *M. tuberculosis* *InhA* to isoniazid [178] raising the prospect that triclosan use as an antitubercular agent could result in the emergence of isoniazid-resistant *M. tuberculosis* strains harboring *InhA* mutant enzymes.

3.7. *N*-Octanesulfonylacetamide (OSA)

N-Octanesulfonylacetamide (OSA, Fig. 9) is a β -sulfonyl carboxamide designed to mimic the transition state of the reaction catalyzed by the β -ketoacyl synthase. Different species of mycobacteria were shown to be susceptible to OSA, including *M. tuberculosis*, with MICs ranging from 6.25 to 12.5 $\mu\text{g mL}^{-1}$ [182]. Of considerable importance to tuberculosis treatment is the finding that OSA showed no cross-resistance to isoniazid, rifampicin, ethambutol, streptomycin and pyrazinamide [182]. OSA treatment of *M. bovis* BCG resulted in inhibition of all mycolic acids commonly found in this mycobacterial species detected by two-dimensional thin layer chromatography. Moreover, OSA treatment of *M. bovis* BCG disrupted cell wall synthesis and there was incomplete septum formation as observed by transmission electron microscopy. The inhibition of INH-resistant *M. tuberculosis* growth suggests that inhibition of mycolic acid synthesis by OSA may be due to interaction with an enzymatic target different from INH. Novel and possibly unexploited mechanism of action of OSA in *M. tuberculosis* may be found, hence unveiling new approaches to antitubercular drug design. Inhibition of FAS-I and FAS-II systems interfere with phospholipid synthesis. Since OSA treatment of *M. bovis* BCG did not result in inhibition of phospholipid synthesis, it has been suggested that the target of OSA may lie in an ACP-dependent FAS-II system [182]. However, the target for OSA mode of action remains to be identified.

3.8. Cerulenin

Cerulenin (2R,3S-epoxy-4-oxo-7,10-trans,trans-dodecanoic acid amide; Fig. 9) was isolated from the culture broth of the fungus *Cephalosporium caerulens* [183]. Cerulenin is active against a wide variety of yeast, fungi and bacteria [184]. Since this compound inhibits both FAS-I and FAS-II systems, interest in the development of this lead as a therapeutic agent has waned. More recently, however, cerulenin has been shown to inhibit growth of several species of mycobacteria, including *M. tuberculosis*, *M. bovis* BCG, and multi-drug resistant isolates of *M. tuberculosis* [185]. Absence of cross-resistance to cerulenin by *M. tuberculosis* clinical isolates resistant to rifampicin, ethambutol, streptomycin, pyrazinamide, and isoniazid suggests a novel and unexploited mechanism of drug action. The primary fatty acid products of FAS-I are either $\text{C}_{16,18}$ or $\text{C}_{24,26}$ [70]. The former are commonly incorporated into phospholipids and triglycerides. Thus, incorporation of fatty acids into these lipids serve as an indirect indicator of FAS-I activity. Cerulenin treatment of *M. bovis* BCG resulted in the inhibition of phospholipid and mycolic acid synthesis [185], suggesting that this drug inhibits both mycobacterial FAS-I and FAS-II systems. *E. coli* FAS-II is composed of three β -ketoacyl-ACP synthases [62]. Two of these synthases are elongation condensing enzymes. FabB (β -ketoacyl-ACP synthase I) is required for a critical step in the elongation of unsaturated fatty acids. FabF (β -ketoacyl-ACP synthase II) controls the temperature-dependent regulation of fatty acid composition. FabH (β -ketoacyl-ACP synthase III) functions in the first, initiating, condensation step in the pathway. Unlike FabB and FabF, FabH enzymes use an acyl-CoA

rather than an acyl-ACP as primer. FabH is further distinguished by a His-Asn-Cys [186] catalytic triad in contrast to the His-His-Cys triad in the FabB [187] and FabF [188] enzymes. The activity of purified FabB and FabF enzymes were shown to be inhibited by cerulenin, with FabB being the most sensitive, whereas FabH is resistant to cerulenin [167]. Determination of the crystal structure of FabB-cerulenin complex showed that there is formation of a covalent complex between the C2 of the ring-opened form of cerulenin and the active site Cys-163 [167]. Moreover, the two histidines in the triad of FabB are important for this enzyme sensitivity toward cerulenin, since replacement of His333 for Asn produced an *E. coli* FabB with a FabH active site architecture that is resistant to cerulenin. Although cerulenin appears not to be a candidate drug, as a result of its lack of stability in mammalian systems, the growth inhibitory effect of cerulenin against multi-drug resistant mycobacteria suggests that it may unveil novel target(s) for drug development.

3.9. Nitroimidazopyrans (NAPs)

PA-824 is a 3-substituted nitroimidazopyran (NAP, Fig. 9) that has been shown to possess highly specific activity against *M. tuberculosis* complex and is safe for oral delivery [189]. Structure activity relationship of over 100 3-substituted NAPs indicated that optimal activity was obtained with lipophilic substituents and *S* enantiomeric forms at C3. PA-824 inhibits protein and lipid synthesis. The latter lead to the accumulation of hydroxymycolic acid with a concomitant reduction in ketomycolate. Since hydroxymycolate is a known biosynthetic precursor of cell wall ketomycolate (Fig. 6), PA-824 is thought to inhibit an enzyme or deplete a cofactor responsible for the oxidation of hydroxymycolate to ketomycolate [189]. Metabolic activation of PA-824 appears to be necessary for activity and may involve a nitro-reduction step catalyzed by an F420-dependent glucose-6-phosphate dehydrogenase (Fgd). Of great importance for tuberculosis treatment is the finding that PA-824 is active against both actively replicating and static *M. tuberculosis*, and is also active against multidrug-resistant strains of *M. tuberculosis*. PA-824 may, therefore, represent an important lead compound with potential for reducing the duration of current therapy that targets growing organisms and for treating drug resistant strains of *M. tuberculosis*.

3.10. Other Potential Inhibitors

Hexachlorophene (Fig. 9) is structurally related to triclosan that possesses antimicrobial activity and is used in disinfectants and surgical scrubs [171]. Hexachlorophene is active against Gram-positive bacteria, especially strains of *Staphylococcus*, but less effective against Gram-negative organisms [169]. Hexachlorophene has been shown to inhibit the *S. aureus* *fabI*-encoded NADPH-dependent enoyl-ACP reductase [190] and the *H. influenzae* *fabI*-encoded NADH-dependent enoyl ACP reductase [191]. However, it has been suggested that FabI is unlikely to be a biologically relevant target for the antibacterial activity of hexachlorophene due to the inability of this drug to form a ternary complex with

FabI and the nucleotide cofactor [190]. The inhibition of *M. tuberculosis* InhA enzyme and/or antimycobacterial activity by hexachlorophene remains to be demonstrated. It is unlikely that hexachlorophene will be useful as an antitubercular drug due to its neurotoxic side effects [192].

Antigen85 complex (section 2.4.5), which was shown to have mycolyl transferase activity, is inhibited by 6-azido-6-deoxy- α -D-trehalose (ADT, Fig. 9), demonstrating marked inhibition of the synthesis of trehalose 6-monomycolate and 6, 6'-trehalose dimycolate [105]. The three-dimensional structure determination of the Antigen85C component [109] should allow for the rational design of inhibitor with potential antitubercular activity.

4. MOLECULAR MODELING OF TARGET ENZYMES

InhA catalyzes the NADH-dependent stereospecific reduction of long-chain 2-*trans*-enoyl-ACP fatty acids, and is the bona fide target for isoniazid (section 3.1). It has recently been demonstrated that InhA is also the target for the broad-spectrum biocide triclosan (section 3.6). Triclosan also inhibits the enoyl reductase (EnvM) from *E. coli*, of which structure in complex with NAD⁺ and triclosan has been solved [193-194]. *M. tuberculosis* InhA and *E. coli* EnvM share 40% sequence identity [77] and structural comparison of the two enzymes showed that the amino acid residues forming the active site are highly conserved.

We are performing molecular modeling and simulations studies of triclosan docked into InhA structures (1ENY [130] and 1BVR [142]) in order to address the molecular basis of InhA inhibition by this compound.

Triclosan was docked in the substrate binding site of InhA in the open (1ENY) and closed (1BVR) forms of the enzyme, followed by energy minimization of the complex using the AMBER forcefield [195].

The preliminary results show that in the 1BVR structure, the phenol ring of triclosan stacks above the nicotinamide ring of the NAD⁺ cofactor allowing π - π stacking interactions. Furthermore, the ring makes van der Waals contacts with the side-chains of Phe149, Tyr158, Pro193 and Met199. In the open structure (1ENY), the Tyr158 side-chain does not point toward the binding site, even after energy minimization. However, in the closed structure (1BVR), after minimization, this side-chain rotates and forms a hydrogen bond with the hydroxyl group of triclosan (2.89Å). These results are illustrated in Fig. (10A).

It is likely that triclosan will bind to InhA after release of the reduced fatty acid substrate. Consistent with this proposal is the greater affinity of triclosan for the InhA-NAD⁺ than to NADH-InhA complex [178]. Thus, the NAD⁺-InhA complex would be in an open form as described in the 1ENY structure [130], allowing triclosan access to the substrate binding site. In order to address the role of InhA flexibility (transition between the open and closed forms) for triclosan inhibition, we also started molecular dynamics simulations of the fully solvated NAD⁺-InhA-triclosan

complex. Preliminary results, illustrated in Fig. (10B), show the Tyr158 side-chain is moving towards the phenolic oxygen of triclosan. This interaction seems to be essential for the inhibitor binding to InhA [178].

The molecular modeling and simulation studies of *M. tuberculosis* InhA in complex with NAD⁺ and triclosan showed that triclosan is located in the substrate binding pocket. These results are consistent with the three-dimensional structure determination of the ternary complexes between NAD⁺, triclosan and *E. coli* and *B. napus* enoyl reductases, which showed that triclosan is located in the substrate binding site of these enzymes [193]. Thus, owing to occupancy of the enoyl reductase active site by the noncovalent NAD⁺-triclosan complex, this binding mode is capable of preventing the reduction reaction. This inhibition mode of action indicates that inhibitors targeted at the substrate binding site may be an alternative to the treatment of resistant tuberculosis.

5. CONCLUDING REMARKS

The Institute for Genomic Research (TIGR) published the first complete genome sequence for a free-living organism, *Haemophilus influenzae*, only six years ago [196]. Since that first report there has been an exponential increase in complete genome sequences published (for details see <http://www.tigr.org/tdb/mdb/mdb.html>). One of the expected benefits of genome analysis of pathogenic bacteria is in the area of human health, particularly in the design of more rapid diagnostic reagents and development of new vaccines and antimicrobial agents. These goals have become more urgent with the continuing spread of antibiotic resistance in important human pathogens [197]. However, there are mechanisms for generating antigenic variation in proteins expressed on the cell surface that are encoded in the genome sequence of an organisms. For example, slipped-strands mispairing within DNA sequence repeats have been found in 5'-intergenic regions and coding sequences in *M. tuberculosis* [69]. The ability of a human pathogen to alter its antigenic potential and thereby evade the immune system has the potential to hinder vaccine development by conventional methods [197].

The challenge to the scientific community is to eliminate the comparative lack of modern biochemical and genetic information about mycobacteria in general. *M. tuberculosis*, in particular, is not an organism whose large scale culture is achievable in any but a handful of laboratories. The recent development of strains of the nonpathogenic, fast-growing *M. smegmatis* that can be genetically manipulated [198-199] has allowed classical genetic studies to be performed in this nonpathogenic organism. These approaches, with PCR-based techniques and genetic manipulations of *M. bovis* and *M. tuberculosis*, have clarified the mechanism of action of, and resistance to, first- and second-line antitubercular drugs [9]. Moreover, the ability to genetically manipulate *M. tuberculosis* and generate defined mutants with specific phenotypes has paved the way for investigations of *M. tuberculosis* pathogenesis [22]. This will also allow the molecular interactions between specific microbial products

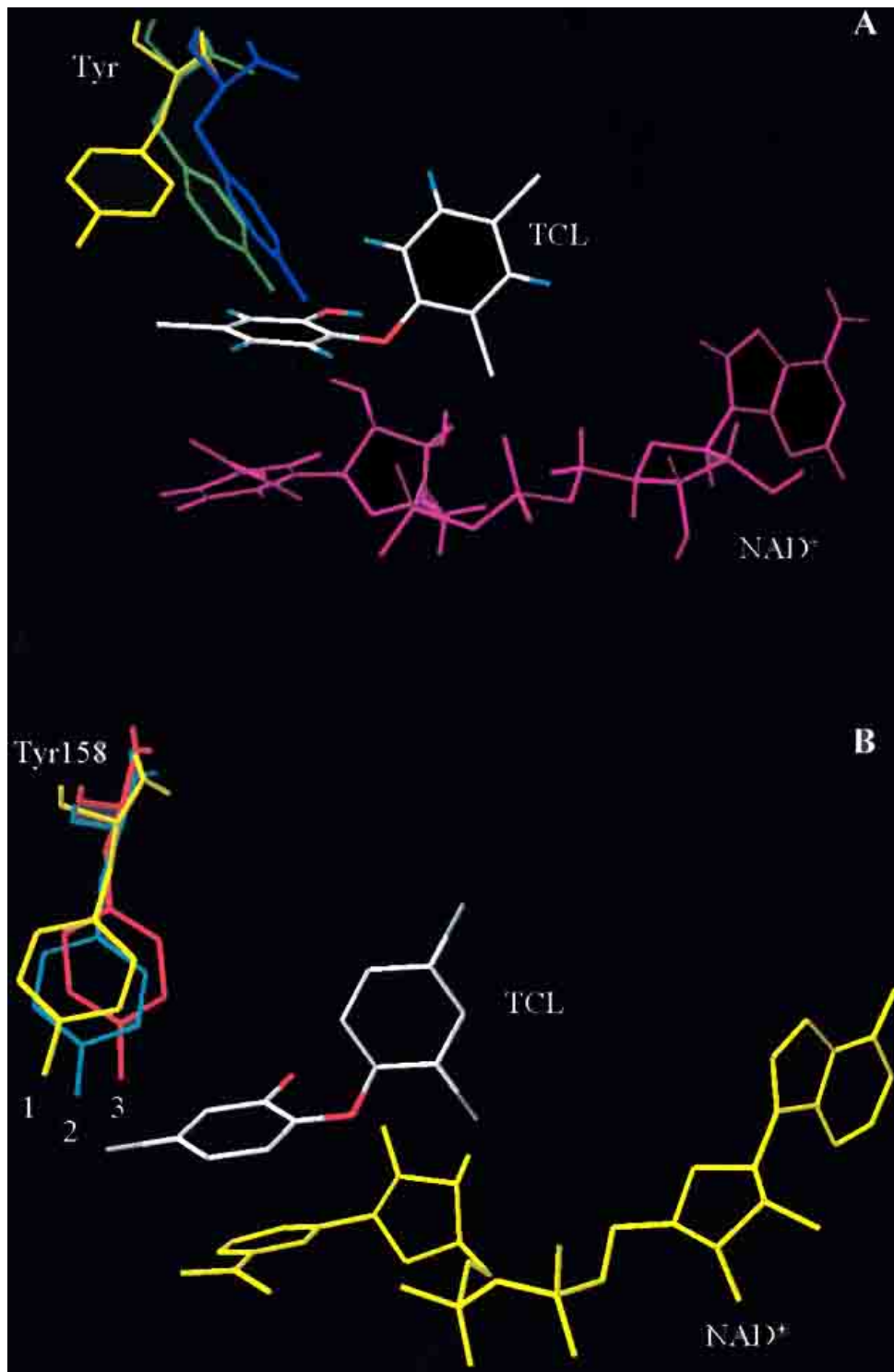


Fig. (10). A). Triclosan-enoyl reductase interactions in *E. coli* EnvM (blue), and *M. tuberculosis* minimized 1BVR (green), and 1ENY (yellow). Only the hydrogen-bond donor/acceptor Tyr residue is displayed. B). The molecular dynamics trajectory of the Tyr158 side-chain of *M. tuberculosis* 1ENY showing its movement towards the phenolic oxygen of triclosan. Hydrogen atoms were removed. (1) the initial, minimized, conformation of Tyr158 found in the 1ENY structure. (2) and (3) conformations at 20ps and 28ps, respectively. In going from 1 to 3, the distance between the phenolic oxygen of Tyr158 and the phenolic oxygen of triclosan is 6.6Å, 5.8Å, and 4.4Å, respectively. All energy minimization and molecular dynamics simulations were performed with the AMBER version 6 suite of programs [209] and the Cornell *et al.* forcefield [195]. Visualization and geometric analysis were performed with Swiss-PdbViewer [210].

and the host cell to be discovered. It has been pointed out that *M. tuberculosis* employs novel strategies for both replication and persistence *in vivo*, many of which involve novel lipid mediators in its cell envelope [22]. In particular, the identification of genes that are critical to mycobacterial persistence should provide targets that may have a positive effect on therapeutic regimen and allow the duration of therapy to be reduced.

The serendipitous discovery of isoniazid, ethionamide, and ethambutol has now provided clues into critical and unique biosynthetic pathways in mycobacteria. The chemical simplicity of these molecules, and a decade of synthetic endeavors after their initial discovery, suggest that more potent analogs, prepared by classic organic synthesis or novel combinatorial synthetic methods, will be hard to find. Given what is now known about the mechanism of action of isoniazid, the discovery of homologs that are both actively accumulated and oxidatively activated seems unlikely. However, as the mechanism of action of these drugs is clarified, and their molecular targets identified, more rational mechanism-based and structure-based approaches to inhibitor design will be possible. These latter approaches are being implemented in our laboratory through the use of computer modeling and simulation, enzymology, organic synthesis, crystallography, and mycobacterial genetics as part of our efforts to develop new antitubercular agents.

Very few organisms produce such a diverse array of lipophilic molecules as *M. tuberculosis*, ranging from simple fatty acids and isoprenoids to very-long-chain, highly complex molecules such as mycolic acids [69]. An understanding of the biochemistry and molecular genetics of the unique cell wall lipids should provide a rational basis for seeking drugs targeted at the biosynthetic pathways of these lipids. Accordingly, the enzymes involved in mycolic acid biosynthesis should represent potential targets for the development of antitubercular agents.

Alternative solutions that should also be examined include the reevaluation of existing antibacterials. The example of fluoroquinolone inhibitors of DNA gyrase was discussed above. A second example could include the β -lactam inhibitors of peptidoglycan biosynthesis. Mycobacteria are naturally insensitive to β -lactams, because of their extremely hydrophobic cell wall [200] and an active β -lactamase [201-202]. However, the presence of a penicillin-binding protein, which possesses transpeptidase activity, has been demonstrated in *M. smegmatis* [203]. Accordingly, the combined administration of β -lactams and β -lactamase inhibitors has recently been shown to be effective in inhibiting the growth of mycobacteria [204, 202]. Given their oral availability and favorable toxicology profile, the thousands of β -lactams that have been synthesized seem worthy of re-examination, in combination with inhibitors of mycobacterial β -lactamases [8]. Unfortunately, acquired β -lactam resistance in *M. smegmatis* has been shown to arise from a combination of decreased affinities of high-molecular-weight penicillin-binding protein (PBP 1) for β -lactams and reduced cell wall permeability [205].

The restoration of antibiotic sensitivity to resistant strains of *E. coli* by external guide sequences (EGS) of RNA

nucleotides may represent an alternative approach for treatment of drug-resistant bacilli [206]. EGSs form complexes with the mRNAs encoded by genes responsible for drug resistance, and these complexes are recognized as substrates by endoribonuclease RNase P which cleaves and inactivates the targeted mRNA, thereby rendering the host cells drug-sensitive. This technique has recently been used to inhibit bacterial growth by reducing the level of expression of two different essential proteins that exist in fewer than 1,500 copies in *E. coli*, RNase P Cr protein and gyrase [207]. Moreover, the inhibition of bacterial growth via ESG-targeted technology exhibits species specificity and dose-response; thus suggesting that it may be used for narrow spectrum antimicrobial activity. Another interesting approach is the use of "peptoid" molecules made from nonnatural, identical amino acids, that have been shown to serve as effective transporters for the molecular delivery of drugs [208]. These "peptoids" are chemically modified versions of a nine-aminoacid oligopeptide responsible for membrane crossing of HIV type I nuclear transcription activator protein. Whether or not such approaches can be applied to *M. tuberculosis* remains to be seen. For now, the best solution to the problems with tuberculosis chemotherapy and the explosion of multi-drug resistance is patient surveillance and treatment compliance by directly observed therapy.

ACKNOWLEDGMENTS

Financial support for this work was provided by Millenium Initiative Program MCT-CNPq, Ministry of Health - Secretary of Health Policy (Brazil) to DSS, LAB and ONdS. ONdS also acknowledges FAPERGS for financial support. EKS is a recipient of a fellowship from CAPES (Brazil). JSB is a grant career award from NIH (USA).

ABBREVIATIONS

ADT	=	6-Azido-6-deoxy- α -D-galactopyranosyl- β -D-trehalose
CMAS	=	Cyclopropane mycolic acid synthase
EGS	=	External guide sequence
ENR	=	Enoyl Reductase
ETH	=	Ethionamide
GLP	=	Glycopeptidolipids
INH	=	Isoniazid
ISO	=	Isoxyl [4,4'-diisoamyloxydiphenylthiourea or 4,4'-diisoamyloxythiocarbamide or thiocarbide]
LAM	=	Lipoarabinomannan
LOS	=	Lipooligosaccharides
mAGP	=	Mycolyl-arabinogalactan-peptidoglycan complex
MMAS	=	Methoxy mycolic acid synthase

NAP	= Nitroimidazopyran
OSA	= <i>N</i> -Octanesulfonylacetamide
PE	= Phosphatidylethanolamine
PIM	= Phosphatidylinositol mannoside
PZA	= Pyrazinamide
SAM	= <i>S</i> -adenosyl methionine
TAG	= Triacylglycerol
TDM	= 6,6'-trehalose dimycolate
TLM	= Thiolactomyacin [(4 <i>S</i>)(2 <i>E</i> ,5 <i>E</i>)-2,4,6-trimethyl-3-hydroxy-2,5,7-octatriene-thiolide]
TMM	= Trehalose 6-monomycolate

REFERENCES

- [1] World Health Organization. *Global Tuberculosis Control. WHO Report 2001*. Geneva, Switzerland, WHO/CDS/TB/2001.287.
- [2] Dunlap, N.E.; Bass, J.; Fujiwara, P.; Hopewell, P.; Horsburgh, C.R.Jr.; Salfinger, M. and Simone, P.M. (2000) *Am. J. Respir. Crit. Care Med.* **161**, 1376-1395.
- [3] Dye, C.; Scheele, S.; Dolin, P.; Pathania, V. and Raviglione, M.C. (1999) *JAMA* **282**, 677-686.
- [4] McKinney, J.D.; Jacobs, W.R.Jr. and Bloom, B.R. (1998) in *Emerging Infections* (Krause, R.M. Ed.), Academic Press, New York, pp. 51-146.
- [5] Murray, C.J.L. (1994) in *Tuberculosis: Pathogenesis, Protection, and Control*, (Bloom, B.R.; Ed.), ASM, Washington, pp. 583-622.
- [6] Smith, P.G. and Moss, A.R. (1994) in *Tuberculosis: Pathogenesis, Protection, and Control*, (Bloom, B.R.; Ed.), ASM, Washington, pp. 47-59.
- [7] Bloom, B.R. and Murray, C.J.L. (1992) *Science* **257**, 1055-1064.
- [8] Blanchard, J.S. (1996) *Annu. Rev. Biochem.* **65**, 215-239.
- [9] Basso, L.A. and Blanchard, J.S. (1998) *Adv. Exp. Med. Biol.* **456**, 115-144.
- [10] Goble, M.; Iseman, M.D.; Madsen, L.A.; Waite, D.; Ackerson, L. and Horsburgh C.R. (1993) *N. Engl. J. Med.* **328**, 527-532.
- [11] Fischl, M.A.; Daikos, G.L.; Uttamchandani, R.B.; Poblete, R.B.; Moreno, J.N.; Reyes, R.R.; Boota, A.M.; Thompson, L.M.; Cleary, T.J.; Oldham, S.A.; Saldana, M.J. and Lai, S. (1992) *Ann. Intern. Med.* **117**, 184-190.
- [12] Koch, R. (1932) *Am. Rev. Tuberc.* **25**, 285-323. [Translated from the original 1882 article by Berna Pinner and Max Pinner.]
- [13] Casadevall, A. (1996) *Clin. Infect. Dis.* **23**, 790-794.
- [14] Schatz, A. and Waksman, S.A. (1944) *Proc. Soc. Exp. Biol. Med.* **57**, 244-248.
- [15] Kushner, S.; Dalalian, H.; Sanjurjo, J.L.; Bach, F.L.; Safir, S.R.; Smith, V.K.Jr.; and Williams, J.H. (1952) *J. Am. Chem. Soc.* **74**, 3617-3621.
- [16] Middlebrook, G. (1952) *Am. Rev. Tuberc.* **65**, 765-767.
- [17] Kessler, R.E. (1997) *Clin. Infect. Dis.* **24**:S146-147.
- [18] Bax, R.P. (1997) *Clin. Infect. Dis.* **24**, S151-153.
- [19] Silver, L. and Bostian, K. (1990) *Eur. J. Clin. Microbiol. Infect. Dis.* **9**, 455-461.
- [20] Coleman, K.; Athalye, M.; Clancey, A.; Davison, M.; Payne, D.J.; Perry, C.R.; and Chopra, I. (1994) *J. Antimicrob. Chemother.* **33**, 1091-1116.
- [21] Ghuysen, J.M. (1994) *Trends Microbiol.* **2**, 372-380.
- [22] Glickman, M.S.; Cahill, S.M. and Jacobs, W.R.Jr. (2001) *J. Biol. Chem.* **276**, 2228-2233.
- [23] Grosset, J. (1980) *Clin. Chest Med.* **1**, 231241.
- [24] Clemens, D.L. (1996) *Trends Microbiol.* **4**, 113-118.
- [25] Clemens, D.L.; Lee, B.Y. and Horwitz, M.A. (2000) *Infect. Immun.* **68**, 2671-2684.
- [26] Sturgill-Koszycki, S.; Schlesinger, P.H.; Chakraborty, P.; Haddix, P.L.; Collins, H.L.; Fok, A.K.; Allen, R.D.; Gluck, S.L.; Heuser, J. and Russell, D.G. (1994) *Science* **263**, 678-681.
- [27] Ferrari, G.; Langen, H.; Naito, M. and Pieters, J. (1999) *Cell* **97**, 435-447.
- [28] Gatfield, J. and Pieters, J. (2000) *Science* **288**, 1647-1650.
- [29] Berthet, F.X.; Lagranderie, M.; Gounon, P.; Laurent-Winter, C.; Ensergueix, D.; chavarot, P.; Thouron, F.; Maranghi, E.; Pelicic, V.; Portnoi, D.; Marchal, G. and Gicquel, B. (1998) *Science* **282**, 759-762.
- [30] Hiriyan, K.T. and Ramakrishnan, T. (1986) *Arch. Microbiol.* **144**, 105-109.
- [31] Mitchison, D.A. (1985) *Tubercle* **66**, 219-225.
- [32] Heifets, L.B. (1994) *Semin. Respir. Infect.* **9**, 84-103.
- [33] Dutt, A. K. and Stead W. (1997) *Dis. Month* **43**, 247-274.
- [34] Stratton, M.A. and Reed, M.T. (1986) *Clin. Pharm.* **5**, 977-987.
- [35] Weis, S.E.; Slocum, P.C.; Blais, F.X.; King, B.; Nunn, M.; Matney, G.B.; Gomez, E. and Foresman, B.H. (1994) *N. Engl. J. Med.* **330**, 1179-1184.
- [36] Jarlier, V. and Nikaïdo, H. (1994) *FEMS Microbiol. Lett.* **123**, 11-18.
- [37] Canetti, G. (1965) *Am. Rev. Resp. Dis.* **92**, 687-703.
- [38] David, H. L. (1980) *Clin. Chest. Med.* **11**, 227-230.
- [39] Huebner, R.E. and Castro, K.G. (1995) *Annu. Rev. Med.* **46**, 47-55.

- [40] Weltman, A.C. and Rose, D.N. (1994) *Arch. Intern. Med.* **154**, 2161-2167.
- [41] Brennan, P.J. and Draper, P. (1994) in *Tuberculosis. Pathogenesis, Protection, and Control*, (B.R. Bloom, Ed.), ASM Press, Washington, D.C.; pp. 271-284.
- [42] Besra, G.S. and Chatterjee, D. (1994) in *Tuberculosis. Pathogenesis, Protection, and Control*, (B.R. Bloom, Ed.), ASM Press, Washington, D.C.; pp. 285-306.
- [43] Brennan, P.J. and Nikaido, H. (1995) *Annu. Rev. Biochem.* **64**, 29-63.
- [44] Chatterjee, D. (1997) *Current Opinion in Chemical Biology*, **1**, 579-588.
- [45] Kolattukudy, P.E.; Fernandes, N.D.; Azad, A.K.; Fitzmaurice, A.M. and Sirakova, T.D. (1997) *Mol. Microbiol.* **24**, 263-270.
- [46] Barry III, C.E.; Lee, R.R.; Mdluli, K.; Sampson, A.E.; Schroeder, B.G.; Slayden, R.A. and Yuan, Y. (1998) *Prog. Lipid Res.* **37**, 143-179.
- [47] Belanger, A.E. and Inamine, J.M. (2000) in *Molecular Genetics of Mycobacteria*, (G.F. Hatfull and W.R. Jacobs, Eds.), ASM Press, Washington, D.C.; pp.191-202.
- [48] Kremer, L.; Baulard, A.R. and Besra, G.S. (2000) in *Molecular Genetics of Mycobacteria*, (Hatfull, G.F. and Jacobs, W.R.Jr.; Eds.), ASM Press, Washington, pp. 173-190.
- [49] Schleifer, K.H. and Kandler, O.B. (1970) *Bacteriol. Rev.*; **36**, 407-477.
- [50] Draper, P. (1982) in *The Biology of the Mycobacteria*, (C. Ratledge and J. Stanford, Eds.), Academic Press, London, pp. 9-52.
- [51] Minnikin, D.E. (1982) in *The Biology of the Mycobacteria*, (C. Ratledge and J. Stanford, Eds.), Academic Press, London, pp. 95-184.
- [52] Kusaka, T. and Mori, T. (1986) *J. Gen. Microbiol.* **132**, 3403-3406.
- [53] Kaneda, K.; Naito, S.; Imaizumi, S.; Yano, I.; Mizuno, S.; Tomiyasu, I.; Baba, T.; Kusunose, E. and kusunose, M. (1986) *J. Clin. Microbiol.* **24**, 1060-1070.
- [54] Durand, E.; Welby, M.; Laneelle, G. and Toccane, J.F. (1979) *Eur. J. Biochem.* **93**, 103-112.
- [55] Nikaido, H.; Kim, S.H. and Rosenberg, E.Y. (1993) *Mol. Microbiol.* **8**, 1025-1030.
- [56] Liu, J.; Barry, III,C.E.; Besra, G.S. and Nikaido, H. (1996) *J. Biol. Chem.* **271**, 29545-29551.
- [57] Yuan, Y.; Crane, D.C.; Musser, J.M.; Sreevatsan, S. and Barry, C.E.;III. (1997) *J. Biol. Chem.* **272**, 10041-10049.
- [58] Yuan, Y.; Lee, R.E.; Besra, G.S.; Belisle, J.T. and Barry, III,C.E. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 6630-6634.
- [59] Glickman, M.S.; Cox, J.S. and Jacobs, W.R.Jr. (2000) *Mol. Cell* **5**, 717-727.
- [60] Yuan, Y. and Barry, III,C.E. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 12828-12833.
- [61] Takayama, K. and Qureshi, N. (1984) in *The Mycobacteria: a Sourcebook*, (Kubica, G.P. and Wayne, L.G.; Eds), Marcel Dekker, New York, pp. 315-344.
- [62] Cronan, J.E. and Rock, C.O. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, (Neidhardt, F.C.; Ed.), ASM Press, Washington, pp. 612-636.
- [63] Wakil, S.J.; Stoops, J.K. and Joshi, V.C. (1983) *Ann. Rev. Biochem.* **52**, 537-579.
- [64] Brindley, D.N.; Matsumura, S. and Bloch, K. (1969) *Nature* **224**, 666-669.
- [65] Bloch, K. (1975) *Methods. Enzymol.* **35**, 84-90.
- [66] Wood, W.I.; Peterson, D.O. and Bloch, K. (1977) *J. Biol. Chem.* **252**, 5745-5749.
- [67] Kikuchi, S.; Rainwater, D.L. and Kolattukudy, P.E. (1992) *Arch. Biochem. Biophys.* **295**, 318-326.
- [68] Fernandes, N.D. and Kolattukudy, P.E. (1996) *Gene* **170**, 95-99.
- [69] Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M.A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J.E.; Taylor, K.; Whitehead, S. and Barrell, B.G. (1998) *Nature* **393**, 537-544.
- [70] Bloch, K. (1977) *Adv. Enzymol. Relat. Areas Mol. Biol.* **45**, 1-84.
- [71] Ratledge, C. (1982) in *The Biology of Mycobacteria*, (Ratledge, C. and Stanford, J.; Eds.), Academic Press, San Diego, pp. 53-94.
- [72] Mdluli, K.; Slayden, R.A.; Zhu, Y.; Ramaswamy, S.; Pan, X.; Mead, D.; Crane, D.D.; Musser, J.M. and Barry, C.E.III (1998) *Science* **280**, 16071610.
- [73] Magnuson, K.; Jackowski, S.; Rock, C.O. and Cronan, J.E.Jr. (1993) *Microbiol. Rev.* **57**, 522-542.
- [74] Schaeffer, M.L.; Agnihotri, G.; Volker, C.; Kallender, H.; Brennan, P.J. and Lonsdale, J.T. (2001) *J. Biol. Chem. Papers in press*. Published on October 12, 2001 as Manuscript M108903200
- [75] Schaeffer, M.L.; Agnihotri, G.; Kallender, H.; Brennan, P.J. and Lonsdale, J.T. (2001) *Biochim. Biophys. Acta* **1532**, 67-78.
- [76] Kremer, L.; Douglas, J.D.; Baulard, A.R.; Morehouse, C.; Guy, M.R.; Alland, D.; Dover, L.G.; Lakey, J.H.; Jacobs, W.R.Jr.; Brennan, P.J.; Minnikin, D.E. and Besra, G.S. (2000) *J. Biol. Chem.* **275**, 16857-16864.
- [77] Banerjee A, Dubnau E, Quémar A, Balasubramanian, V.; Um, K.S.; Wilson, T.; Collins, D.; de Lisle, G. and Jacobs, W.R.Jr. (1994) *Science* **263**, 227-230.

- [78] Liu, J. and Nikaido, H. (1999) *Proc. Natl. Acad. Sci.* **96**, 4011-4016.
- [79] Vilchèze, C.; Morbidoni, H.R.; Weisbrod, T.R.; Iwamoto, H.; Kuo, M.; Sacchettini, J.C. and Jacobs, W.R.Jr. (2000) *J. Bacteriol.* **182**, 4059-4067.
- [80] Choi, K.W.; Kremer, L.; Besra, G.S. and Rock, C.O. (2000) *J. Biol. Chem.* **275**, 28201-28207.
- [81] Scarsdale, J.N.; Kazanina, G.; He, X.; Reynolds, K.A. and Wright, H.T. (2001) *J. Biol. Chem.* **276**, 20516-20522.
- [82] Qureshi, N.; Sathyamoorthy, N. and Takayama, K. (1984) *J. Bacteriol.* **157**, 46-52.
- [83] Qureshi, N.; Takayama, K. and Schnoes, H.K. (1980) *J. Biol. Chem.* **255**, 182-189.
- [84] Jackson, M.; Portnoi, D.; Catheline, D.; Dumail, L.; Rauzier, J.; Legrand, P. and Gicquel, B. (1997) *Infect. Immun.* **65**, 2883-2889.
- [85] Fox, B.G.; Shanklin, J.; Ai, J.; Loerh, T.M. and Sanders-Loerh, J. (1994) *Biochemistry* **33**, 12776-12786.
- [86] Wong, M.Y.H.; Steck, P.A. and Gray, G.R. (1979) *J. Biol. Chem.* **254**, 5734-5740.
- [87] Danielson, S.J. and Gray, G.R. (1982) *J. Biol. Chem.* **257**, 12196-12203.
- [88] Lacave, C.; Lanéelle, M.A.; Daffé, M.; Montrozier, H.; Rols, M.P. and Asselineau, C. (1987) *Eur. J. Biochem.* **163**, 369-378.
- [89] Ingrosso, D.; Fowler, A.V.; Bleibaum, J. and Clarke, S. (1989) *J. Biol. Chem.* **264**, 20131-20139.
- [90] Wang, A.-Y.; Grogan, D.W. and Cronan, J.E.Jr. (1992) *Biochemistry* **31**, 11020-11028.
- [91] George, K.M.; Yuan, Y.; Sherman, D.R.; Sherman, D.R. and Barry, III, C.E. (1995) *J. Biol. Chem.* **45**, 27292-27298.
- [92] Dubnau, E.; Lanéelle, M.A.; Soares, S.; Benichou, A.; Vaz, T.; Prome, D.; Prome, J.C.; Daffé, N. and Quemard, A. (1997) *Mol. Microbiol.* **23**, 313-322.
- [93] Schroeder, B.G. and Barry, C.E.III. (2001) *Bioorg. Chem.* **29**, 164-177.
- [94] Wheeler, P.R. and Ratledge, C. (1994) in *Tuberculosis. Pathogenesis, Protection, and Control*, (B.R. Bloom, Ed.), ASM Press, Washington, D.C.; pp. 353-385.
- [95] Dougherty, D.A. (1996) *Science* **271**, 163-168.
- [96] Yuan, Y.; Mead, D.; Schroeder, B.G.; Zhu, Y. and Barry, C.B.III. (1998) *J. Biol. Chem.* **273**, 21282-21290.
- [97] Takayama, K. and Davidson, L.A. (1979) *Trends Biochem. Sci.* **14**, 280-282.
- [98] Lee, R.E.; Armour, J.W.; Takayama, K.; Brennan, P.J. and Besra, G.S. (1997) *Biochim. Biophys. Acta* **1346**, 275-284.
- [99] Gastambide-Odier, M. and Lederer, E. (1959) *Nature* **184**, 1563-1564.
- [100] Walker, R.W.; Promer, J.C. and Lacave, C.S. (1973) *Biochim. Biophys. Acta* **326**, 52-62.
- [101] Etémadi, A.-H. (1967) *Bulletin de la Société de Chimie Biologique* **49**, 695-706.
- [102] Ozeki, Y.; Kaneda, K.; Fujiwara, N.; Morimoto, M.; Oka, S. and Yano, I. (1997) *Infect. Immun.* **65**, 1793-1799.
- [103] Spargo, B.J.; Crowe, L.M.; Ioneda, T.; Beaman, B.L. and Crowe, J.H. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 737-740.
- [104] Sathyamoorthy, N. and Takayama, K. (1987) *J. Biol. Chem.* **262**, 13417-13423.
- [105] Belisle, J.T.; Vissa, V.D.; Sievert, T.; Takayama, K.; Brennan, P.J. and Besra, G.S. (1997) *Science* **276**, 1420-1422.
- [106] Wiker, H.G. and Harboe, M. (1992) *Microbiol. Rev.* **56**, 648-661.
- [107] Jackson, M.; Raynaud, C.; Lanéelle, M.A.; Guillot, C.; Laurent-Winter, C.; Ensergueix, D.; Gicquel, B. and Daffé, M. (1999) *Mol. Microbiol.* **31**, 1573-1587.
- [108] Armitige, L.Y.; Jagannath, C.; Wanger, A.R. and Norris, S.J. (2000) *Infect. Immun.* **68**, 767-778.
- [109] Ronning, D.R.; Klabunde, T.; Besra, G.S.; Vissa, V.D.; Belisle, J.T. and Sacchettini, J.C. (2000) *Nature Struct. Biol.* **7**, 141-146.
- [110] Berstein J.; Lott W.A.; Steinberg B.A. and Yale H.L. (1952) *Am. Rev. Tuberc.* **65**, 357-364.
- [111] Quémard, A.; Sacchettini, J.C.; Dessen, A.; Vilchèze, C.; Bittman, R.; Jacobs, W.R.Jr. and Blanchard, J.S. (1995) *Biochemistry* **34**, 8235-8241.
- [112] Middlebrook, G. (1954) *Am. Rev. Tuberc.* **69**, 471-472.
- [113] Middlebrook, G.; Cohn, M.L. and Schaefer, W.B. (1954) *Am. Rev. Tuberc.* **70**, 852-872.
- [114] Zhang, Y.; Heym, B.; Allen, Bryan, Young, D. and Cole, S. (1992) *Nature* **358**, 591-593.
- [115] Zhang, Y.; Garbe, T. and Young, D. (1993) *Mol. Microbiol.* **8**, 521-524.
- [116] Heym, B. and Cole, S.T. (1992) *Res. Microbiol.* **143**, 721-730.
- [117] Cockerill, F.R.; Uhl, J.R.; Temesgen, Z.; Zhang, Y.; Stockman, L.; Roberts, G.D.; Williams, D.L. and Kline, B.C. (1995) *J. Infect. Dis.* **171**, 240-245.
- [118] Heym, B.; Alzari, P.M.; Honore, N. and Cole, S.T. (1995) *Mol. Microbiol.* **15**, 235-245.
- [119] Musser, J.M.; Kapur, V.; Williams, D.L.; Kreiswirth, B.N.; Soolingen, D.V. and van Embden, J.D.A. (1996) *J. Infect. Dis.* **173**, 196-202.
- [120] Shoeb, H.A.; Bowman, B.U.; Ottolenghi, A.C. and Merola, A.J. (1985) *Antimicrob. Agents Chemother.* **27**, 399-403.
- [121] Johnsson, K. and Schultz, P.G. (1994) *J. Am. Chem. Soc.* **116**, 7425-7426.

- [122] Johnsson, K.; King, D.S. and Schultz, P.G. (1995) *J. Am. Chem. Soc.* **117**, 5009-5010.
- [123] Basso, L.A.; Zheng, R. and Blanchard, J.S. (1996) *J. Am. Chem. Soc.* **118**, 11301-11302.
- [124] Bergler, H.; Högenauer, G. and Turnowsky, F. (1992) *J. Gen. Microbiol.* **269**, 2093-2100.
- [125] Rozwarski, D.A.; Grant, G.A.; Barton, D.H.R.; Jacobs, W.R.Jr. and Sacchettini, J.C. (1998) *Science* **279**, 98-102.
- [126] Zabinski, R.F. and Blanchard, J.S. (1997) *J. Am. Chem. Soc.* **119**, 2331-2332.
- [127] Lei, B.; Wei, C.-J. and Tu, S.-C. (2000) *J. Biol. Chem.* **275**, 2520-2526.
- [128] Quémar, A.; Dessen, A.; Sugantino, M.; Jacobs, W.R.Jr.; Sacchettini, J.C. and Blanchard, J.S. (1996) *J. Am. Chem. Soc.* **118**, 1561-1562.
- [129] Wilming, M. and Johnsson, K. (1999) *Angew. Chem. Int. Ed.* **38**, 2588-2590.
- [130] Dessen, A.; Quémar, A.; Blanchard, J.S.; Jacobs, W.R.Jr. and Sacchettini, J.C. (1995) *Science* **267**, 1638-1641.
- [131] Basso, L.A.; Zheng, R.; Musser, J.M.; Jacobs, W.R.Jr. and Blanchard, J.S. (1998) *J. Infect. Dis.* **178**, 769-775.
- [132] Dhandayuthapani, S.; Zhang, Y.; Mudd, M.H. and Deretic, V. (1996) *J. Bacteriol.* **178**, 3641-3649.
- [133] Wilson, T.M. and Collins, D.M. (1996) *Mol. Microbiol.* **19**, 1025-1034.
- [134] Sherman, D.R.; Mdluli, K.; Hickey, M.J.; Arain, T.M.; Morris, S.L.; Barry, C.E.III and Stover, C.K. (1996) *Science* **272**, 1641-1643.
- [135] Sreevatsan, S.; Pan, X.; Zhang, Y.; Deretic, V. and Musser, J.M. (1997) *Antimicrob. Agents Chemother.* **41**, 600-606.
- [136] Kelley, C.L.; Rouse, D.A. and Morris, S.L. (1997) *Antimicrob. Agents Chemother.* **41**, 2057-2058.
- [137] Miesel, L.; Weisbrod, T.R.; Marcinkeviciene, J.A.; Bittman, R. and Jacobs, W.R.Jr. (1998) *J. Bacteriol.* **180**, 2459-2467.
- [138] Takayama, K.; Schnoes, H.K.; Armstrong, E.L. and Boyle, R.W. (1975) *J. Lipid Res.* **16**, 308-317.
- [139] Mdluli, K.; Sherman, D.R.; Hickey, M.J.; Kreiswirth, B.N.; Morris, S.; Stover, C.K. and Barry, C.E.III (1996) *J. Infect. Dis.* **174**, 1085-1090.
- [140] Magliozzo, R.S. and Marcinkeviciene, J.A. (1997) *J. Biol. Chem.* **272**, 8867-8870.
- [141] Nguyen, M.; Quemard, A.; Marrakchi, H.; Bernadou, J. and Meunier, B. (2001) *C. R. Acad. Sci. Paris, Série Iic, Chimie* **4**, 35-40.
- [142] Rozwarski, D.A.; Vilchèze, C.; Sugantino, M.; Bittman, R. and Sacchettini, J.C. (1999) *J. Biol. Chem.* **274**, 15582-15589.
- [143] Winder, F.G. (1982) in *The Biology of the Mycobacteria*, (Ratledge, C. and Stanford, J. Eds.), Academic Press, New York, pp. 353-438.
- [144] Quémar, A.; Lanéelle, G. and Lacave, C. (1992) *Antimicrob. Agents Chemother.* **36**, 1316-1321.
- [145] Baulard, A.R.; Betts, J.C.; Engohang-Ndong, J.; Quan, S.; McAdam, R.A.; Brennan, P.J.; Loch, C. and Besra, G.S. (2000) *J. Biol. Chem.* **275**, 28326-28331.
- [146] DeBarber, A.E.; Mdluli, K.; Bosman, M.; Bekker, L.-G. and Barry, C.E.III. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 9677-9682.
- [147] Davidson, P.T. and Quoc, L.H. (1992) *Drugs* **43**, 651-673.
- [148] Heifets, L.B. and Lindholm-Levy, P.J. (1992) *Am. Rev. Respir. Dis.* **145**, 1223-1225.
- [149] Mackaness, G.B. (1956) *Am. Rev. Tuberc.* **74**, 718-728.
- [150] Crowle, A.J.; Dahl, R.; Ross, E. and May, M.H. (1991) *Infect. Immun.* **59**, 1823-1831.
- [151] Konno, K.; Feldmann, F. and McDermott, W. (1967) *Am. Rev. Resp. Dis.* **95**, 461-469.
- [152] Butler, W.R. and Kilburn, J.O. (1983) *Antimicrob. Agents Chemother.* **24**, 600-601.
- [153] Miller, M.A.; Thibert, L.; Desjardins, F.; Siddiqi, S.H. and Dascal, A. (1995) *J. Clin. Microbiol.* **33**, 2468-2470.
- [154] Yamamoto, S.; Toida, I.; Watanabe, N. and Ura, T. (1995) *Antimicrob. Agents Chemother.* **39**, 2088-2091.
- [155] Scorpio, A. and Zhang, Y. (1996) *Nature Med.* **2**, 662-667.
- [156] Cynamon, M.H.; Speirs, R.J. and Welch, J.T. (1998) *Antimicrob. Agents Chemother.* **42**, 462-463.
- [157] Zimhony, O.; Cox, J.S.; Welch, J.T.; Vilchèze, C. and Jacobs, W.R.Jr. (2000) *Nature Med.* **6**, 1043-1047.
- [158] Winder, F.G.; Collins, P.B. and Whelan, D. (1971) *J. Gen. Microbiol.* **66**, 379-380.
- [159] Phetsuksiri, B.; Baulard, A.R.; Cooper, A.M.; Minnikin, D.E.; Douglas, J.D.; Besra, G.S. and Brennan, P.J. (1999) *Antimicrob. Agents Chemother.* **43**, 1042-1051.
- [160] Oishi, H.; Noto, T.; Sasaki, H.; Suzuki, K.; Hayashi, T.; Okazaki, H.; Ando, K. and Sawada, M. (1982) *J. Antibiot. (Tokyo)* **35**, 391-395.
- [161] Sasaki, H.; Oishi, H.; Hayashi, T.; Matsuura, I.; Ando, K. and Sawada, M. (1982) *J. Antibiot. (Tokyo)* **35**, 396-400.
- [162] Miyakawa, S.; Suzuki, K.; Noto, T.; Harada, Y. and Okazaki, H. (1982) *J. Antibiot. (Tokyo)* **35**, 411-419.
- [163] Noto, T.; Miyakawa, S.; Oishi, H.; Endo, H. and Okazaki, H. (1982) *J. Antibiot. (Tokyo)* **35**, 401-410.
- [164] Hayashi, T.; Yamamoto, O.; Sasaki, H.; Kawaguchi, A. and Okazaki, H. (1983) *Biochem. Biophys. Res. Commun.* **115**, 1108-1113.
- [165] Slayden, R.A.; Lee, R.E.; Armour, J.W.; Cooper, A.M.; Orme, I.M.; Brennan, P.J. and Besra, G.S. (1996) *Antimicrob. Agents Chemother.* **40**, 2813-2819.
- [166] Heath, R.J.; White, S.W. and Rock, C.O. (2001) *Prog. Lipid Res.* **40**, 467-497.

- [167] Price, A.C.; Choi, K.-H.; Heath, R.J.; Li, Z.; White, S.W. and Rock, C.O. (2001) *J. Biol. Chem.* **276**, 6551-6559.
- [168] Tierno, P.M.Jr. (1999) *Am. J. Infect. Control* **27**, 71-72.
- [169] Korolkovas, A. (1988) in *Essentials of Medicinal Chemistry*, John Wiley & Sons, Inc.; New York.
- [170] McMurry, L.M.; Oethinger, M. and Levy, S.B. (1998) *Nature* **394**, 531-532.
- [171] Heath, R.J.; Yu, Y.-T.; Shapiro, M.A.; Olson, E. and Rock, C.O. (1998) *J. Biol. Chem.* **273**, 30316-30321.
- [172] Heath, R.J.; Rubin, J.R.; Holland, D.R.; Zhang, E.; Snow, M.E. and Rock, C.O. (1999) *J. Biol. Chem.* **274**, 11110-11114.
- [173] Ward, W.H.J.; Holdgate, G.A.; Rowsell, S.; McLean, E.G.; Pauptit, R.A.; Clayton, E.; Nichols, W.W.; Colls, J.G.; Minshull, C.A.; Jude, D.A.; Mistry, A.; Timms, D.; Camble, R.; Hales, N.J.; Britton, C.J. and Taylor, I.W.F. (1999) *Biochemistry* **38**, 12514-12525.
- [174] Levy, C.W.; Roujeinikova, A.; Sedelnikova, S.; Baker, P.J.; Stuitje, A.R.; Slabas, A.R.; Rice, D.W. and Rafferty, J.B. (1999) *Nature* **398**, 383-384.
- [175] Heath, R.J. and Rock, C.O. (2000) *Nature* **406**, 145-146.
- [176] Heath, R.J.; Su, N.; Murphy, C.K. and Rock, C.O. (2000) *J. Biol. Chem.* **275**, 40128-40133.
- [177] McMurry, L.M.; McDermott, P.F. and Levy, S.B. (1999) *Antimicrob. Agents Chemother.* **43**, 711-713.
- [178] Parikh, S.L.; Xiao, G. and Tonge, P.J. (2000) *Biochemistry* **39**, 7645-7650.
- [179] Vischer, W.A. and Regos, J. (1974) *Zentbl. Bakteriologie Hyg. Abt. 1 Orig. A.* **226**, 376-389.
- [180] McMurry, L.M.; Oethinger, M. and Levy, S.B. (1998) *FEMS Microbiol. Lett.* **166**, 305-309.
- [181] Schweizer, H.P. (1998) *Antimicrob. Agents Chemother.* **42**, 394-398.
- [182] Parrish, N.M.; Houston, T.; Jones, P.B.; Townsend, C. and Dick, J.D. (2001) *Antimicrob. Agents Chemother.* **45**, 1143-1150.
- [183] Sato, Y.; Nomura, S.; Kamio, Y.; Omura, S. and Hata, T. (1967) *J. Antibiot. (Tokyo)* **20**, 344-348.
- [184] Matsumae, A.; Nomura, S. and Hata, T. (1964) *J. Antibiot. (Tokyo)* **17**, 1-7.
- [185] Parrish, N.M.; Kuhajda, F.; Heine, H.; Bishai, W. and Dick, J.D. (1999) *J. Antimicrob. Chemother.* **43**, 219-226.
- [186] Davies, C.; Heath, R.J.; White, S.W. and Rock, C.O. (2000) *Structure* **8**, 185-195.
- [187] Olsen, J.G.; Kadziola, A.; Wettstein-Knowles, P.; Siggaard-Andersen, M.; Lindqvist, Y. and Larsen, S. (1999) *FEBS Lett.* **460**, 46-52.
- [188] Huang, W.; Jia, J.; Edwards, P.; Dehesh, K.; Schneider, G. and Lindqvist, Y. (1998) *EMBO J.* **17**, 1183-1191.
- [189] Stover, C.K.; Warrener, P.; VanDevanter, D.R.; Sherman, D.R.; Arain, T.M.; Langhorne, M.H.; Anderson, S.W.; Towell, J.A.; Yuan, Y.; McMurray, D.N.; Kreiswirth, B.N.; Barry, C.F.III and Baker, W.R. (2000) *Nature* **405**, 962-966.
- [190] Heath, R.J.; Li, J.; Roland, G.E. and Rock, C.O. (2000) *J. Biol. Chem.* **275**, 4654-4659.
- [191] Marcinkeviciene, J.; Jiang, W.; Kopcho, L.M.; Locke, G.; Luo, Y. and Copeland, R.A. (2001) *Arch. Biochem. Biophys.* **390**, 101-108.
- [192] Evangelista de Duffard, A.M. and Duffard, R. (1996) *Environ. Health Perspect.* **104** (Suppl.2), 353-360.
- [193] Roujeinikova, A.; Levy, C.W.; Rowsell, S.; Sedelnikova, S.; Baker, P.J.; Minshull, C.A.; Mistry, A.; Colls, J.G.; Camble, R.; Stuitje, A.R.; Slabas, A.R.; Rafferty, J.B.; Pauptit, R.A.; Viner, R. and Rice, D.W. (1999) *J. Mol. Biol.* **294**, 527-535.
- [194] Qiu, X.; Janson, C.; Court, R.; Smyth, M.; Payne, D. and Abdel-Meguid, S. (1999) *Protein Sci.* **8**, 2529-2532.
- [195] Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.Jr, Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W. and Kollman, P.A. (1995) *J. Am. Chem. Soc.* **117**, 5179-5197.
- [196] Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.R.; Bult, C.J.; Tomb, J.F.; Dougherty, B.A.; Merrick, J.M.; et al. (1995) *Science* **269**, 496-512.
- [197] Fraser, C.M.; Eisen, J.A. and Salzberg, S.L. (2000) *Nature* **406**, 799-803.
- [198] Shinnick, T.M.; King, C.H. and Quinn, F.D. (1995) *Am. J. Med. Sci.* **309**, 92-98.
- [199] Snapper, S.B.; Melton, R.E.; Mustafa, S.; Kieser, T. and Jacobs, W.R.Jr. (1990) *Mol. Microbiol.* **4**, 1911-1919.
- [200] Jarlier, V.L.; Gutmann, L. and Nikaido, H. (1991) *Antimicrob. Agents Chemother.* **35**, 1937-1939.
- [201] Fattorini, L.; Orefici, G.; Jin, S.H.; Scardaci, G.; Amicosante, G.; Franceschini, N. and Chopra, I. (1992) *Antimicrob. Agents Chemother.* **36**, 1068-1072.
- [202] Zhang, Y.; Steingrube, V.A. and Wallace, R.J.Jr. (1992) *Am. Rev. Respir. Dis.* **145**, 657-660.
- [203] Basu, J.; Chattopadhyay, R.; Kundu, M. and Chakrabarti, P. (1992) *J. Bacteriol.* **174**, 4829-4832.
- [204] Prabhakaran, K.; Harris, E.B.; Randhawa, B.; Adams, L.B.; Williams, D.L. and Hastings, R.C. (1993) *Microbios* **76**, 251-261.
- [205] Mukhopadhyay, S.; and Chakrabarti, P. (1997) *Antimicrob. Agents Chemother.* **41**, 1721-1724.
- [206] Guerrier-Takada, C.; Salavati, R. and Altman, S. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 8468-8472.
- [207] McKinney, J.; Guerrier-Takada, C.; Wesolowski, D. and Altman, S. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 6605-6610.

- [208] Wender, P.A.; Mitchell, D.J.; Pattabiraman, K.; Pelkey, E.T.; Steinman, L. and Rothbard, J.B. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 13003-13008.
- [209] Case, D.A.; Pearlman, D.A.; Caldwell, J.W.; Cheatham, T. E.; Ross, W.R.; Simmerling, C.L.; Darden, T.A.; Merz, K.M.; Stanton, R.V.; Cheng, A.L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R.J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U.C.; Weiner, P. K. and Kollman, P. A. (1999) AMBER 6.0, University of California, San Francisco.
- [210] Guex, N. and Peitsch, M.C. (1997) *Electrophoresis* **18**, 2714-2723.

Molecular Dynamics Simulation Studies of the Wild-Type, I21V, and I16T Mutants of Isoniazid-Resistant *Mycobacterium tuberculosis* Enoyl Reductase (InhA) in Complex with NADH: Toward the Understanding of NADH-InhA Different Affinities

Evelyn Koeche Schroeder,^{*†} Luiz Augusto Basso,^{†‡} Diógenes Santiago Santos,^{†§} and Osmar Norberto de Souza^{*¶}

^{*}Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas—LABIO, PPGCC, Faculdade de Informática, PUCRS, 90619-900, Porto Alegre, RS, Brazil; [†]PPGBCM, Centro de Biotecnologia, UFRGS, 91501-970, Porto Alegre, RS, Brazil; [‡]Dep. Biologia Molecular e Biotecnologia, UFRGS, 91501-970, Porto Alegre, RS, Brazil; [§]Centro de Pesquisa em Biologia Molecular e Funcional, TECNOPUC, PUCRS, 90619-900, Porto Alegre, RS, Brazil; and [¶]Instituto de Pesquisas Biomédicas, PUCRS, 90610-000, Porto Alegre, RS, Brazil

ABSTRACT The increasing prevalence of tuberculosis in many areas of the world, associated with the rise in drug-resistant *Mycobacterium tuberculosis* (MTB) strains, presents a major threat to global health. InhA, the enoyl-ACP reductase from MTB, catalyzes the nicotinamide adenine dinucleotide (NADH)-dependent reduction of long-chain *trans*-2-enoyl-ACP fatty acids, an intermediate in mycolic acid biosynthesis. Mutations in the structural gene for InhA are associated with isoniazid resistance in vivo due to a reduced affinity for NADH, suggesting that the mechanism of drug resistance may be related to specific interactions between enzyme and cofactor within the NADH binding site. To compare the molecular events underlying ligand affinity in the wild-type, I21V, and I16T mutant enzymes and to identify the molecular aspects related to resistance, molecular dynamics simulations of fully solvated NADH-InhA (wild-type and mutants) were performed. Although very flexible, in the wild-type InhA-NADH complex, the NADH molecule keeps its extended conformation firmly bound to the enzyme's binding site. In the mutant complexes, the NADH pyrophosphate moiety undergoes considerable conformational changes, reducing its interactions with its binding site and probably indicating the initial phase of ligand expulsion from the cavity. This study should contribute to our understanding of specific molecular mechanisms of drug resistance, which is central to the design of more potent antimycobacterial agents for controlling tuberculosis.

INTRODUCTION

Despite more than five decades of effective chemotherapy, tuberculosis (TB) caused by *Mycobacterium tuberculosis* (MTB) still remains an important problem of public health. The World Health Organization estimates that there were 8.8 million new cases of TB in 2002. The global incidence rate of TB was growing at $\sim 1.1\%$ per year and the number of cases at 2.4% per year (1).

Infection with drug-sensitive strains of MTB can be effectively treated with a combination of drugs including isoniazid (INH), rifampicin, pyrazinamide, and ethambutol or streptomycin (2–5). However, the emergence of multidrug-resistant strains of MTB, defined as resistant to at least INH and rifampicin (the two most powerful anti-TB drugs), makes it urgent to identify the mechanisms of drug action and resistance whose understanding should contribute to the design of more potent antimycobacterial agents for controlling drug-sensitive and drug-resistant TB.

INH has been the cornerstone in TB chemotherapy since it was first found to exhibit a powerful bactericidal activity against MTB in 1952 (6,7). This drug causes the loss of acid-fastness of the tubercle bacilli as a result of the inhibition of mycolic acid synthesis, α -alkyl β -hydroxy long-chain fatty

acids (60–90 carbons in length), important components of the mycobacterial cell wall (8,9).

Using a genetic approach, Jacobs and co-workers identified the protein product of the mycobacterial *inhA* gene as the putative target for INH inhibition (10). The InhA enzyme, or *trans*-2-enoyl-ACP-reductase (EC 1.3.1.9), carries out the stereospecific reduction of α,β -unsaturated fatty acids bound to the acyl carrier protein in an NADH-dependent reaction (11) and is a member of the type II dissociated fatty acid biosynthesis pathway (FAS II) in MTB.

INH itself does not inhibit the enzyme directly but requires conversion to an activated form of the drug by a mycobacterial catalase-peroxidase (KatG) (12–15) and becomes covalently attached to the nicotinamide ring of the NADH bound within the active site of the enzyme (16).

Studies of clinical isolates from MTB resistant to INH were found to have mutations in the *inhA* structural gene, leading to an increase in dissociation constant for NADH cofactor binding to the purified enzyme (17). Consistent with these results, x-ray crystallographic studies (16,18,19) have revealed that all mutations in the enoyl reductase conferring INH resistance were located in the nucleotide binding site. This observation suggested that the mechanism of drug resistance may be related to specific interactions between enzyme and cofactor within the NADH binding site.

Submitted October 5, 2004, and accepted for publication May 9, 2005.

Address reprint requests to Osmar Norberto de Souza, E-mail: osmarns@inf.pucrs.br.

© 2005 by the Biophysical Society

0006-3495/05/08/876/09 \$2.00

doi: 10.1529/biophysj.104.053512

Although INH resistance has also been attributed to another FAS II enzyme (KasA) (20,21), thermal inactivation of InhA in *Mycobacterium smegmatis* resulted in identical phenotypic response to that treated with INH (22), therefore validating InhA as the primary target of INH and as an important target for drug discovery.

InhA is a member of the short chain dehydrogenase/reductase (SDR) family of enzymes (23,24), with a polypeptide backbone Rossmann fold topology, in which a parallel β -sheet is flanked by α -helices forming the central core of the NADH binding site.

Although dinucleotide-binding shows very low overall sequence homology, their protein backbones superimpose fairly well (25,26) and have some common sequence features, one of which is a glycine-rich phosphate binding loop that connects the C-terminus of $\beta 1$ with the N-terminus of $\alpha 1$ (23,26,27).

In a computational comparison of 102 high-resolution Rossmann fold enzymes, Tanner and co-workers (28) identified that even though the majority of the enzymes display the pyrophosphate binding loop sequence having three Gly, the two enoyl-ACP-reductases analyzed (from *Brassica napus* and *Escherichia coli*) have a G-X₇-G pattern. In InhA, the last Gly is substituted by an Ala, but it has the same NADH mode of binding described elsewhere (28–30).

As mutations in the glycine-rich loop have been correlated with attenuation or elimination of enzyme activity (31–33) and two mutations in this loop (I21V and I16T) are related to INH resistance and decrease of cofactor affinity (wild-type InhA: K_m NADH = $56 \pm 4 \mu\text{M}$, K_d NADH = $0,57 \pm 0,04 \mu\text{M}$; I21V InhA: K_m NADH = $104 \pm 7 \mu\text{M}$, K_d NADH = $13,9 \pm 1,7 \mu\text{M}$; I16T InhA: K_m NADH = $149 \pm 10 \mu\text{M}$, K_d NADH = $5,95 \pm 0,33 \mu\text{M}$) (17), in this work we report the molecular dynamics (MD) simulations of wild-type InhA and isoniazid-resistant clinical isolates I21V and I16T mutants in complex with the NADH molecule to identify the molecular aspects of NADH affinity differences and predict the events related to INH resistance.

MATERIALS AND METHODS

MD simulations

MD simulations were performed with the SANDER module of AMBER 6.0 (34) using the Cornell et al. all-atom model force field (35), except for the atomic charges in the NADH molecule which were assigned by ab initio calculations in the HF 6-31G* level (36) and fitted with the RESP procedure (37).

The initial structure for the wild-type InhA-NADH complex was taken from the 2,2 Å crystal structure 1ENY (18) without the crystallographic water molecules. The NADH pyrophosphate moiety was considered deprotonated at physiological pH. The all-atom model of the protein in complex with the cofactor contains 4079 atoms and a net molecular charge of -6 . Hence, six sodium counterions were added near surface Asp residues (5 Å apart) to neutralize the negative charge density of the complex, which was then immersed in a rectangular box containing a total of 9132 TIP3P (38) water molecules. The initial simulation cell dimensions were $77.5 \text{ \AA} \times$

$64.3 \text{ \AA} \times 70.0 \text{ \AA}$ and had the complex solvated by a layer of water molecules of at least 10 \AA in all directions, as previously described (39). The simulation cell contained a total of 31,481 atoms.

In the MD protocol, all hydrogen atoms, ions, and water molecules were first subjected to 100 steps of energy minimization by steepest descent to remove close van der Waals contacts. The system was then relaxed with sequential steps of energy minimization (100 steps of steepest descent plus 400 steps of conjugate gradient) where the solute atoms were gradually allowed to move until no restraints were applied. The temperature of the system was then increased from 10 K to 298,16 K in 6 steps (10 K to 50 K, 50 K to 100 K, 100 K to 150 K, 150 K to 200 K, 200 K to 250 K, and 250 K to 298,16 K), and the velocities at each step were reassigned according to the Maxwell-Boltzmann distribution at that temperature and equilibrated for 2 ps.

Energy minimization and MD were carried out under periodic boundary conditions. The simulation was computed in the NPT ensemble at 298.16 K with the Berendsen temperature coupling and constant pressure of 1 atm with isotropic molecule-based scaling (40). The SHAKE algorithm (41), with a 10^{-5} \AA tolerance, was applied to fix all bonds containing a hydrogen atom, allowing the use of a time step of 2.0 fs in the integration of the equations of motion. No extra restraints were applied after the equilibration phase. The electrostatic interactions between nonligand atoms were evaluated by the particle-mesh Ewald method with a charge grid spacing of $\sim 1.0 \text{ \AA}$, and the charge grid was interpolated on a cubic grid with the direct sum tolerance set to 4.0×10^{-6} . The Lennard-Jones interactions were evaluated using a 9.0 \AA atom-based cutoff (42). Data were collected for 3.1 ns, and structures for analysis were saved every 0.5 ps. The initial system preparation and results analysis were performed on a 2.4 GHz Pentium IV PC, and the MD simulations were performed on a Pentium III PC cluster with 16 single nodes of 1GHz and 256 Mb memory each.

The I21V and I16T mutants were constructed by in silico mutations of the 1.9 ns instantaneous wild-type InhA-NADH conformation, in a way that simulates a spontaneous mutation. As the velocities were taken from the wild-type InhA-NADH simulation, the mutants' MD proceeded with no new thermalization, thus simulating what we called on-the-fly mutation.

The convergence of the different simulations were analyzed in terms of the energy components, secondary structure, root mean-square deviation (RMSD) from the initial x-ray structure (1ENY), and root mean-square-fluctuation (RMSF) to estimate the B-factor. For the B-factor calculation, the RMSFs were calculated relative to the last 2 ns averaged backbone structures, and all coordinate frames from the trajectories were first superimposed on the initial conformation to remove any effect of overall translation and rotation. Atomic isotropic B-factors were calculated from the trajectories using the equation

$$\text{B-factor}_i = (8\pi^2/3)(\langle r_i^2 \rangle - \langle r_i \rangle^2),$$

where $\langle r_i^2 \rangle - \langle r_i \rangle^2$ is the mean-square positional fluctuation of atom i (43,44).

Hydrogen bonds analysis

Because the MD simulations include hydrogen atoms, it was possible to make an explicit analysis of hydrogen bond properties, such as donor-acceptor assignments and hydrogen bond occupancies.

For the occupancies calculations, the NADH hydrogen bonding interactions were assumed to be present if the participating heavy atoms were $< 3.9 \text{ \AA}$ apart and the angle formed between the heavy atoms and donated hydrogen was larger than 120° and was performed with the carnal module of AMBER 6.0. The strongest hydrogen bonds have the highest occupancies, the smallest standard deviations, and the shortest heavy atoms separations. For this reason, we assumed an occupancy cutoff of 50% for the most important H-bonds.

The total number of direct H-bonds between the cofactor and the protein, or those mediated by water molecules, were calculated with the HBPLUS program and default parameters (distance $D-A = 3.9 \text{ \AA}$ and minimum angle $= 90^\circ$) (45).

Dictionary of secondary structure of proteins (46) and PROCHECK-NMR (47) were used to evaluate the secondary structure pattern during the simulations.

Figures were generated with the SwissPDB Viewer v3.7 b2 program (48). Graphics, averages, and standard deviations were calculated with Origin 6.1 Scientific Graphing and Analysis Software (Microcal Software, OriginLab, Northampton, MA).

RESULTS AND DISCUSSION

Protein structural features

To obtain an estimate of the MD trajectory quality and convergence, the backbone RMSD from the starting crystal structure (1ENY) were calculated (Fig. 1). After a rapid increase during the first 250 ps, the protein backbone RMSD average and standard deviation over the last 2 ns of the wild-type InhA-NADH trajectory was $1.6 \pm 0.1 \text{ \AA}$. In the I21V InhA-NADH and I16T InhA-NADH complexes, mutations of the 1.9 ns instantaneous wild-type structure caused a perturbation in these systems, and the simulation had to be prolonged for another 6.5 and 5.5 ns for I21V and I16T, respectively. The I21V InhA-NADH system was most affected by mutation, and only 4.0 ns after the I21V mutation, the backbone RMSD reached a plateau around $2.1 \pm 0.1 \text{ \AA}$ during the last 2 ns. The instability caused by mutation was not so evident in the I16T mutant, but its backbone RMSD average over the last 2 ns was $2.0 \pm 0.1 \text{ \AA}$.

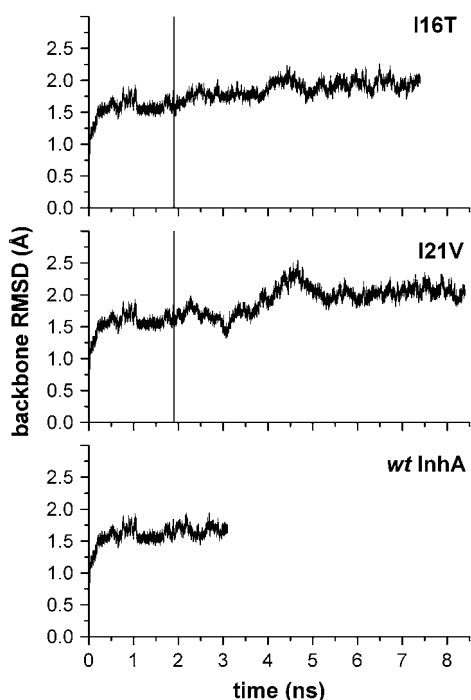


FIGURE 1 Backbone RMSD of wild-type InhA, I21V, and I16T mutants from the starting crystal structure (1ENY). The on-the-fly mutation moment is identified by a vertical line at 1.9 ns.

The total number of protein hydrogen bonds (TNHB) was also monitored along the trajectories and compared with the crystal structure as a way to access the overall stabilization of the structures. Dictionary of Secondary Structure of Proteins (45) analysis of the last 2 ns snapshots of each enzyme showed that there were no significant differences among the systems ($TNHB_{\text{wild-type}} = 178 \pm 5$, $TNHB_{\text{I21V}} = 174 \pm 5$, $TNHB_{\text{I16T}} = 175 \pm 5$). This lower number of hydrogen bonds related to the crystal structure ($TNHB_{\text{1ENY}} = 203$) was expected as, in solution, some residue-residue interactions can be substituted by residue-water interactions.

Because the overall Rossmann fold is important for the cofactor binding and InhA activity, the last eight averaged structures over a 250 ps window of each trajectory were also examined in terms of their secondary structure with PROCHECK-NMR (47) to characterize conformational changes during the MD simulations (data not shown). In all systems, the secondary structures were well maintained, especially the β -sheet core (Fig. 2). In the I21V and I16T mutants, the helical organization of the N-terminal positions of both $\alpha 7$ were lost, being more noticeable for the I21V mutant whose $\alpha 6$ was also less stable.

As indicated by the B-factors calculated over the last 2 ns time window (Fig. 3), the most flexible regions correspond to the substrate binding loop ($\alpha 6$ and $\alpha 7$ helices, residues 196–219), the A loop (residues 100–114), and the B loop (residues 150–161). The large mobility of these loops is consistent with the necessity of this region to undergo conformational changes upon fatty acyl substrate binding (19). Even though the substrate binding loop also has higher B-factors in the crystal structure, they are much lower than observed for the dynamic structures, in which crystal packing is not present. The A and B loops, which were involved in subunit contacts in the crystallographic tetrameric struc-

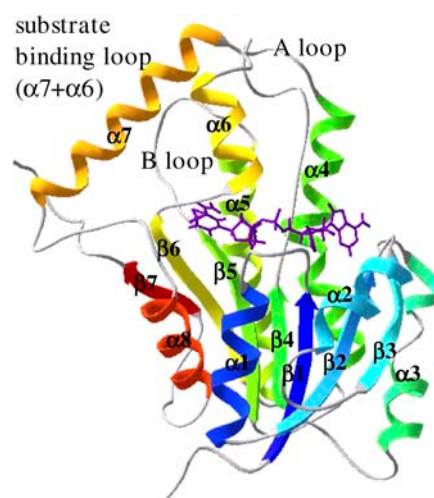


FIGURE 2 Secondary structure identification of the wild-type InhA initial structure colored according to secondary structure succession, from blue (N-terminal) to red (C-terminal). The NADH molecule is in an extended conformation over the Rossman fold β -sheet core.

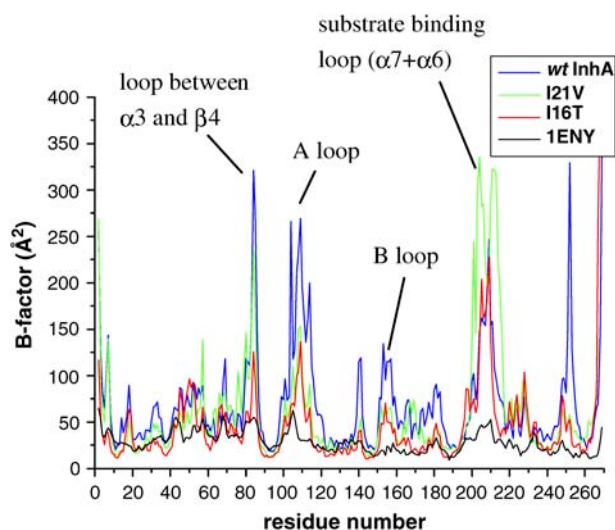


FIGURE 3 Backbone residue-based B-factors calculated over the last 2 ns time window for wild-type (blue), I21V (green), and I16T (red) InhA enzymes. Crystallographic B-factors (black).

ture, were also shown to have higher B-factors. The same is true for residues 80–90 from the loop between $\alpha 3$ and $\beta 4$. Yet, the mobility of the core residues remain fairly close to those derived from the crystallographic data.

As the substrate binding loop and A and B loops movements are involved in cofactor and substrate entrance and exit, the distance between $\alpha 6$ helix and loop A and $\alpha 7$ helix and loop B were monitored during the trajectories. From this analysis it was observed that although all $\alpha 6$ -A loop distances fluctuate near or below the crystal distance (except for a 250 ps period for the wild-type InhA), the $\alpha 7$ -B loop distance on the I21V ($13.5 \pm 0.4 \text{ \AA}$) and I16T ($13.2 \pm 0.3 \text{ \AA}$) mutants were above that observed for the crystal structure (12.6 \AA), probably induced by the instability introduced by these mutations.

NADH association

Even though the overall Rossmann fold structure was well maintained during the MD trajectories, the NADH association within each binding site was different.

The NADH binding mode within the InhA enzyme is similar to that of the other SDR enzymes with a Rossmann fold structure (23,26,49–52). The NADH molecule binds to the enzyme in an extended conformation along the top of the C-termini of the core β -sheet, with the nicotinamide ring pointing deep into a hydrophobic cavity formed by the strands $\beta 4$, $\beta 5$, and $\beta 6$ and by the helices $\alpha 5$, $\alpha 6$, and $\alpha 7$ (Fig. 2) (18).

As mutations in the InhA enzyme are associated to an increase in the NADH dissociation constant (17), of particular interest are the hydrogen bonds between the NADH molecule (Fig. 4) and the enzyme since they monitor position and movements of the ligand within the binding site. These

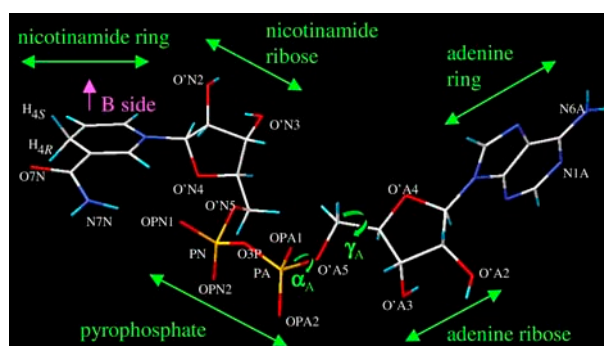


FIGURE 4 Atom names and dihedral angles for the NADH molecule.

H-bonds were monitored during all MD simulations, and Table 1 summarizes their occupancies during the last 2 ns.

For the wild-type enzyme, 10 residues make conserved hydrogen bonds with the cofactor molecule (occupancies $\geq 50\%$). In the nicotinamide cavity, residues Ile¹⁹⁴ and Thr¹⁹⁶ interact with the 3-carboxamide group of the nicotinamide ring (N ring) fixing the B-side of the ring (Fig. 5) and thus the H_{4S} hydrogen, facing the substrate cavity, consistent with the proposed InhA mechanism (11,49).

Residues Lys¹⁶⁵ and Ile⁹⁵ interact with the nicotinamide ribose (NR). Initially, the role of the conserved Lys¹⁶⁵ residue was attributed to stabilization of the intermediate enolate formed during substrate reduction (11). During the wild-type InhA-NADH dynamics, the Lys¹⁶⁵ side-chain hydrogen bonds to the RN hydroxy groups lasted almost the whole trajectory, reinforcing the site-directed mutagenesis studies that indicated that this residue has a primary role in cofactor binding (53).

TABLE 1 Hydrogen bond contacts between the NADH ligand and InhA enzymes with $>50\%$ occupancies during the production phase

Residue	Amino acid group	NADH atom [†]	Occupancy* (%)		
			Wild-type InhA	I21V	I16T
Ile ¹⁵	C=O	O'A3	70	—	54
		O'A3	72	—	—
Ser ²⁰	OH	OPA2	100	—	—
Ile ²¹	N	OPN2	95	—	—
Asp ⁶⁴	COO ⁻	N6A	99	—	100
Val ⁶⁵	N	N1A	100	—	92
Gln ⁶⁶	OE1	N6A	—	—	64
Ile ⁹⁵	C=O	O'N3	100	100	100
Gly ⁹⁶	N	O'A4	100	59	100
Tyr ¹⁵⁸	OH	O7N	—	—	98
Lys ¹⁶⁵	NZ	O'N2	83	86	65
		O'N3	99	86	63
Ile ¹⁹⁴	C=O	N7N	95	84	86
		O7N	100	100	100
Thr ¹⁹⁶	OG1	N7N	67	—	—
		OPN1	61	—	—

*Occupancies were calculated with the CARNAL module from AMBER 6.0.

[†]Atom names from the NADH molecule in Fig. 4.

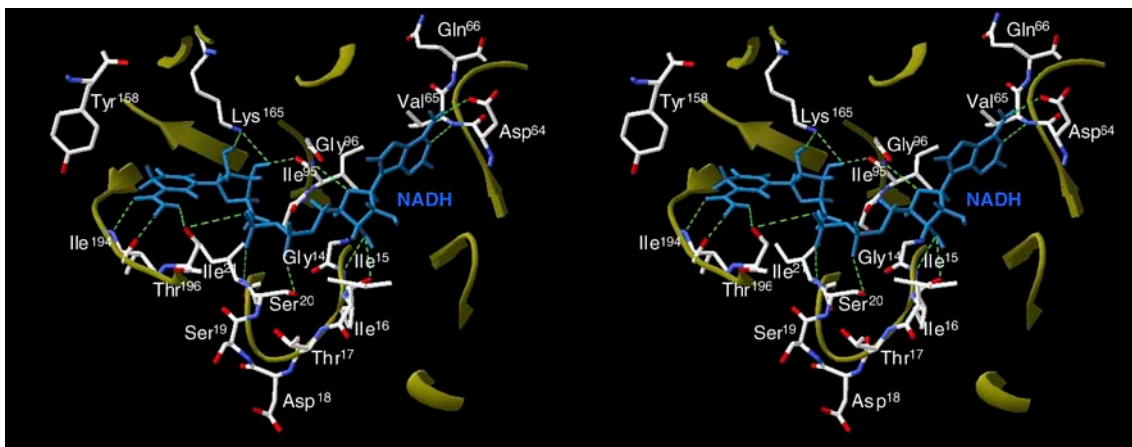


FIGURE 5 Stereo view of the NADH association in its binding pocket, including the glycine-rich loop, in the production phase average structure of the wild-type InhA simulation. NADH is blue, and hydrogen bonds are represented as dashed green lines. NADH atoms and amino acid groups involved in the interactions are listed in Table 1.

Next to the Ile⁹⁵ residue, the Gly⁹⁶ H-bond to the oxygen atom of the adenine ribose (AR) was maintained for the whole trajectory. The adenine moiety of the NADH cofactor molecule also makes important hydrogen bonds with Asp⁶⁴ and Val⁶⁵ located in the adenine ring (A ring) recognition site (54).

In the wild-type enzyme, the glycine-rich loop residues Ser²⁰ and Ile²¹ are involved in the pyrophosphate moiety interactions and are responsible for keeping this flexible cofactor region bound to the enzyme, whereas the Ile¹⁵ residue hydrogen bonds to the AR hydroxy group next to the pyrophosphate moiety (Fig. 5).

Conformational changes caused by the substitution of an apolar residue for a less bulky one (I21V mutation) or for a slightly polar but also less bulky one (I16T mutant) result in a different pattern for NADH interaction.

In the I21V mutant simulation, only four (Ile⁹⁵, Gly⁹⁶, Lys¹⁶⁵, and Ile¹⁹⁴) of the 10 residues that make important bonds with the cofactor molecule are maintained with occurrence $\geq 50\%$. Of these residues, only Gly⁹⁶ is involved in H-bonding with the adenine portion. The other three interact with the nicotinamide portion, and only Ile⁹⁵ and Ile¹⁹⁴ H-bond to O'N3 and O7N, respectively, keeping 100% occurrence.

In the I21V mutant, all direct interactions between the protein and the pyrophosphate moiety became less important during the trajectory. The same is true for those residues that form the adenine recognition region. The occurrence of H-bonds with residues Asp⁶⁴ and Val⁶⁵ changed from 99% and 100% to 21% and 46%, respectively.

In the I16T mutant simulation, there were no such changes in the H-bonds pattern between the NADH molecule and the enzyme binding site. From the 10 residues that originally make H-bonds with occurrence $\geq 50\%$, 7 are present in this mutant. In the adenine recognition moiety, Asp⁶⁴, Val⁶⁵, and now Gln⁶⁶ are H-bonded to the A ring nitrogens.

Interactions with residues Ile⁹⁵, Gly⁹⁶, and Ile¹⁹⁴ remained within the same level as in the wild-type enzyme, but both

NR hydroxy groups make weaker H-bonds with Lys¹⁶⁵ residues.

Because of the NADH conformational changes, the Thr¹⁹⁶ H-bonds to the 3-carboxy group from the N ring were substituted by the Tyr¹⁵⁸ residue. Originally, the Tyr¹⁵⁸ OH group makes an indirect H-bond to this cofactor region through a conserved water molecule that is lost with the I16T mutation. In this mutant, all original interactions between the glycine-rich loop and the pyrophosphate moiety also became less important during the trajectory, and there are no strong H-bonds between the protein and this cofactor region. From the glycine-rich loop, only Ile¹⁵ makes a 54% occupancy H-bond with the other hydroxy group from AR (O'A3).

From the MD simulations, it became clear that mutation in the glycine-rich loop residues I21V and I16T change the pattern of direct H-bonds contacts with the pyrophosphate moiety from the NADH molecule. Being less attached to its binding site, this flexible cofactor region undergoes conformational changes, becoming more distant from the glycine-rich loop (Fig. 6).

The NADH dihedral angles were monitored during the trajectories and showed that while attached to the wild-type enzyme, the cofactor molecule was flexible but its dihedrals fluctuated around values observed in the crystal structures. In the mutant enzymes, NADH undergoes considerable conformational changes in the pyrophosphate moiety, being more prominent for α_A and γ_A dihedrals for both mutants (Fig. 7). The transitions in α_A and γ_A dihedrals, separated by β_A , are highly correlated and characteristic of a "crankshaft motion". This movement is similar to that observed for the phosphodiester torsions in MD trajectories of DNA (55).

The conformational changes experienced by the cofactor molecule and the opening movement of the B loop region allowed a greater interaction with water molecules. As a consequence, in the mutant enzymes the number of water

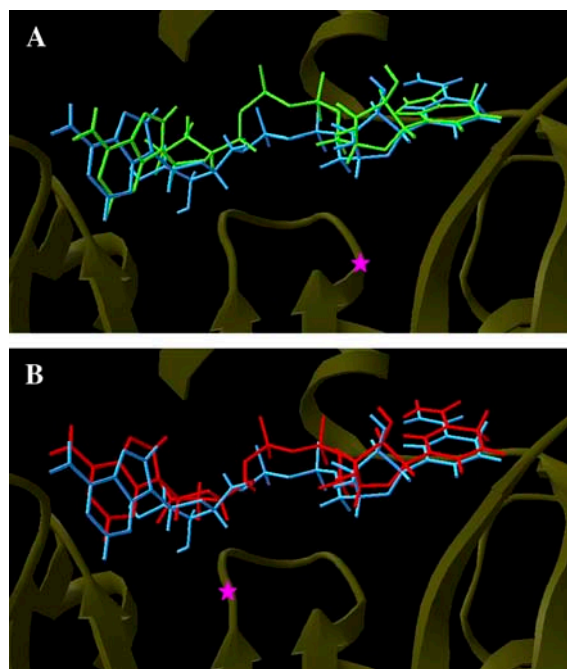


FIGURE 6 Superposition of the NADH production phase average structures from wild-type (blue) and (A) I21V InhA-NADH (green) complex and (B) I16T InhA-NADH complex (red). Magenta stars indicate the locations of the mutated residues.

molecules in the NADH first solvation shell is greater than in the wild-type system (Table 2).

The larger opening between the substrate-binding loop and the B loop in the mutant enzyme as compared to the wild-type enzyme allows more water molecules to enter the nicotinamide binding site. This site is a hydrophobic cavity and in the wild-type enzyme, only 1.9 ± 0.3 water molecules are within a 3.5 Å distance from this cofactor region. In the mutant enzymes, the average number of water molecules became 3.3 ± 1.1 and 2.0 ± 1.0 for I21V and I16T mutant, respectively.

Another significant difference among wild-type and mutant enzymes was in the pyrophosphate solvation shell. This moiety of cofactor moves apart from its binding site, becoming more hydrated, and the average number of water molecules surrounding it changes from 8.1 ± 0.8 to 11.2 ± 1.0 and 11.9 ± 1.0 for the I21V and I16T mutants, respectively.

Because the AR and ring are close to the cofactor binding crevice entrance and accessible to bulk water even in the wild-type enzyme, no differences in this region's solvation shells were observed among the enzymes.

With a larger number of water molecules surrounding the NADH within the mutant binding sites, some direct interactions between cofactor and protein amino acids are substituted by these water molecules. This fact can be observed in the H-bond analysis between the cofactor and the enzymes' amino acids performed with the HBPLUS program (56) for the last 2 ns of each trajectory (Fig. 8).

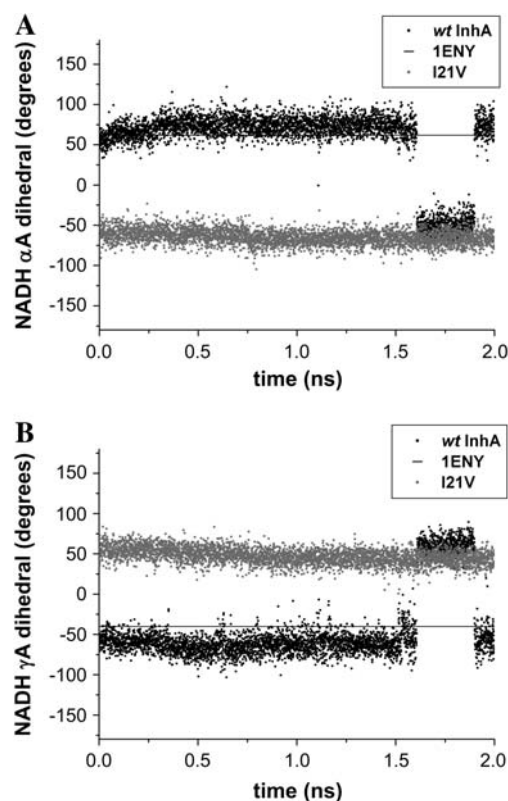


FIGURE 7 Comparison of the NADH dihedral angles (A) α_A and (B) γ_A for the last 2 ns of the wild-type (black) and I21V InhA-NADH (shaded) trajectories. The horizontal line represents the crystallographic values from 1ENY.

In the wild-type system, the number of direct H-bonds between the NADH molecule and the enzyme binding site are almost constant, and fluctuations around the average of 11.6 ± 1.3 are slightly higher than in 1ENY crystal structure (8 NADH-enzyme H-bonds). In the mutated enzymes, the large conformational changes make the total number of direct H-bonds decrease. Only in this analysis, the average number was calculated over the last 500 ps of each trajectory where the I21V and I16T mutants make only 5.6 ± 1.1 and 10.5 ± 1.1 direct H-bonds, respectively.

It is known that water molecules are important for cofactor binding within Rossmann fold proteins (28). In the InhA crystal structures (16,18,19), there are at least seven conserved water molecules within the binding site. H-bond

TABLE 2 Number of water molecules in the first solvation shell (3.5 Å) calculated for the last 2 ns of the wild-type, I21V, and I16T InhA-NADH trajectories using the program ptraj-6.4 from AMBER6.0

NADH moiety	Wild-type InhA-NADH	I21V-NADH	I16T-NADH
N ring	1.9 ± 0.3	3.3 ± 1.1	2.9 ± 1.0
N ribose	1.7 ± 0.6	1.5 ± 0.6	2.0 ± 0.7
Pyrophosphate	8.1 ± 0.8	11.2 ± 1.0	11.9 ± 1.0
A ribose	3.6 ± 0.9	4.0 ± 1.0	4.0 ± 1.2
A ring	4.4 ± 1.0	4.8 ± 1.1	4.1 ± 1.1

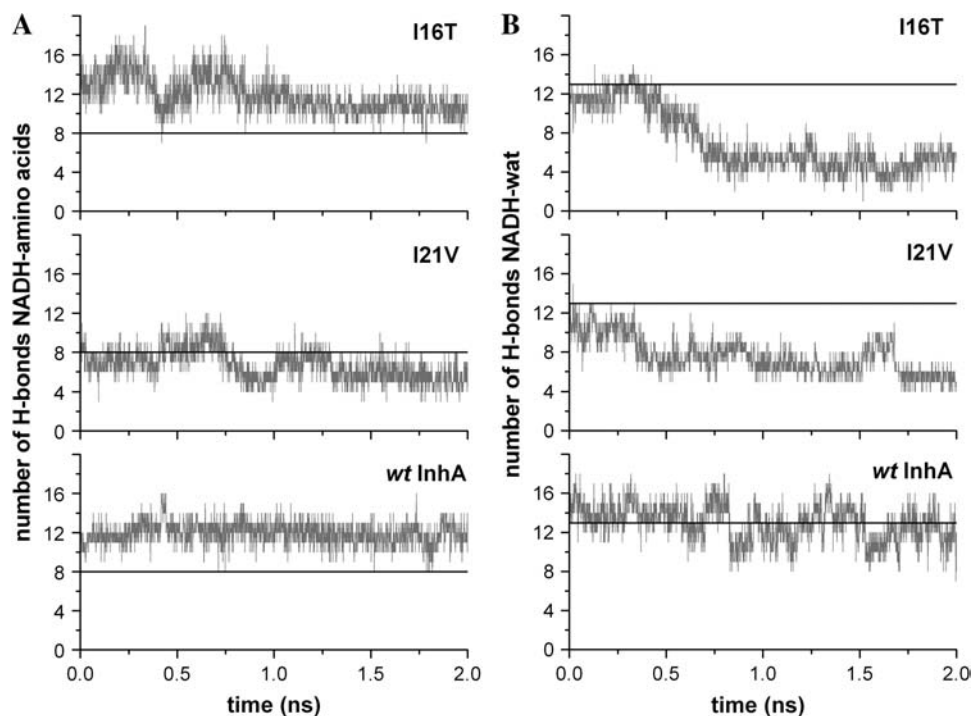


FIGURE 8 H-bond analysis of wild-type, I21V, and I16T InhA-NADH, in the production phase, using the HBPLUS program: (A) Total number of H-bonds between the NADH molecule and amino acids from the enzyme binding sites (crystallographic reference in *black*). (B) Total number of H-bonds with water molecules that mediate cofactor interactions with the enzyme binding sites (average number of water mediated H-bonds in wild-type-InhA system in *black*).

analysis between cofactor and bridging water molecules revealed that there is a good number of high occupancy H-bond interactions with these molecules (data not shown), but analyses of the number of H-bonds between the NADH molecule and water molecules that mediate the interaction with the enzymes are different among the systems. Whereas in the wild-type complex the cofactor molecules make an average of 12.0 ± 1.6 H-bonds with water molecules that mediate interactions with the enzyme amino acids, in the I21V and I16T mutants, only 5.5 ± 0.7 and 5.3 ± 1.1 H-bonds are formed, respectively, between these molecules.

Based on this analysis of wild-type InhA-NADH, I21V InhA-NADH, and I16T InhA-NADH complex trajectories, the increase in the NADH dissociation constant in the I21V and I16T mutant enzymes observed in experimental studies (17) can be attributed to the decrease in the number of direct H-bond interactions between the cofactor molecule and amino acids in the binding pocket and between the cofactor molecule and water molecules that mediate interactions with the enzyme. These structural variations, which lead to different affinities for NADH by wild-type and mutant enzymes, were also quantitatively attested by molecular docking experiments with the AutoDock 3.0 program (57). The NADH cofactor was flexibly docked to the average structures of each system using a genetic algorithm protocol (57). As a result we obtained estimates of the NADH free energy of binding (ΔG_{bind}). The docking experiments showed that wild-type InhA has a higher affinity for NADH than the mutant enzymes (wild-type InhA $\Delta G_{\text{bind}} = -15.8 \pm 1.2$ kcal/mol, I21V InhA $\Delta G_{\text{bind}} = -14.7 \pm 1.0$ kcal/mol, I16T

InhA $\Delta G_{\text{bind}} = -10.3 \pm 1.4$ kcal/mol), indeed corroborating the structural observations.

CONCLUSIONS

In this work, we reported fully solvated, all-atom MD simulations of the wild-type InhA enzyme and two INH-resistant clinical mutants (I21V and I16T) associated with the cofactor NADH molecule to investigate the molecular events related to drug cofactor affinities that could lead to drug resistance. It was noticed from the MD simulations that these mutations cause macromolecular instabilities that were reflected in the long time necessary for trajectory convergence.

The mutations caused a different pattern of cofactor interaction within the enzyme binding site. Whereas in the wild-type enzyme the NADH remains tightly bounded to its pocket, in an extended conformation in the I21V and I16T it undergoes considerable conformational changes. Mutations in the glycine-rich loop affect cofactor pyrophosphate moiety binding and allow its movement apart from its binding site.

Since in aqueous solution the NADH molecule preferentially adopts a more folded conformation with an interring distance between 5 and 10 Å (58,59; E. Schroeder and O. Norberto de Souza, unpublished data) and has to unfold to bind to the enzyme in an extended conformation, it is likely that the conformational changes experienced by the pyrophosphate moiety inside the mutant enzymes could be indicative of an initial phase of NADH folding and expulsion from its binding cavity as a result of the reduced affinity for the mutants' binding pocket.

As InhA inhibition by INH depends on the covalent INH-NAD adduct and mutations in the glycine-rich loop affects the NADH mode of binding within these enzymes, we have predicted that the lower affinity for the pyrophosphate moiety should contribute to INH resistance.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at <http://www.biophysj.org>.

We are grateful to Dr. Ulf Ryde for providing the NADH parameters, and Dr. César A. F. De Rose and the CPAD/PUCRS team for the Ombrófila cluster administration.

This project was supported by grants from FINEP and Millennium Institute, CNPq/MCT to D.S.S. and FAPERGS to O.N.S. D.S.S. and L.A.B. are CNPq research fellow awardees. E.K.S. was supported by a PhD scholarship from CAPES.

REFERENCES

- World Health Organization Ninth Annual Report on Tuberculosis. 2004. Global tuberculosis control: surveillance, planning, financing. <http://www.who.int/tb/publications/2004/en/index1.html>.
- Mitchison, D. A. 1985. The action of antitubercular drugs in short-course chemotherapy. *Tubercle*. 66:219–225.
- Stratton, M. A., and M. T. Reed. 1986. Short-course drug therapy for tuberculosis. *Clin. Pharm.* 5:977–987.
- Combs, D. L., R. J. O'Brien, and L. J. Geiter. 1990. USPHS tuberculosis short-course chemotherapy trial-21—effectiveness, toxicity, and acceptability—the report of final results. *Ann. Intern. Med.* 122:397–406.
- Schroeder, E. K., O. Norberto de Souza, D. S. Santos, J. S. Blanchard, and L. A. Basso. 2002. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. *Curr. Pharm. Biotechnol.* 3:197–225.
- Bernstein, J., W. A. Lott, B. A. Steinberg, and H. L. Yale. 1952. Chemotherapy of experimental tuberculosis—V. Isonicotinic acid hydrazide (nydrazid) and related compounds. *Am. Rev. Tuberc.* 65:357–364.
- Fox, H. H. 1952. The chemical approach to the control of tuberculosis. *Science*. 116:129–134.
- Takayama, K., L. Wang, and H. L. David. 1972. Effect of isoniazid on the *in vivo* mycolic acid biosynthesis, cell growth and viability of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* 2:29–35.
- Winder, F. G., and P. B. Collins. 1970. Inhibition by isoniazid of synthesis of mycolic acids in *Mycobacterium tuberculosis*. *J. Gen. Microbiol.* 63:41–48.
- Banerjee, A., E. Dubnau, A. Quémard, V. Balasubramanian, K. S. Um, T. Wilson, D. Collins, G. de Lisle, and W. R. Jacobs Jr. 1994. *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science*. 263:227–230.
- Quémard, A., J. C. Sacchettini, A. Dessen, C. Vilchère, R. Bittman, W. R. Jacobs Jr., and J. S. Blanchard. 1995. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochemistry*. 34:8235–8241.
- Johnsson, K., and P. G. Schultz. 1994. Mechanistic studies of the oxidation of isoniazid by the catalase peroxidase from *Mycobacterium tuberculosis*. *J. Am. Chem. Soc.* 116:7425–7426.
- Johnsson, K., D. S. King, and P. G. Schultz. 1995. Studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis. *J. Am. Chem. Soc.* 117:5009–5010.
- Basso, L. A., R. Zheng, and J. S. Blanchard. 1996. Kinetics of inactivation of WILD-TYPE and C243S mutant of *Mycobacterium tuberculosis*: enoyl reductase by activated isoniazid. *J. Am. Chem. Soc.* 118:11301–11302.
- Quémard, A., A. Dessen, M. Sugantino, W. R. Jacobs Jr., J. C. Sacchettini, and J. S. Blanchard. 1996. Binding of catalase-peroxidase-activated isoniazid to wild-type and mutant *Mycobacterium tuberculosis* enoyl-ACP reductases. *J. Am. Chem. Soc.* 118:1561–1562.
- Rozwarski, D. A., G. A. Grant, D. H. R. Barton, W. R. Jacobs Jr., and J. C. Sacchettini. 1998. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science*. 279:98–102.
- Basso, L. A., R. Zheng, J. M. Musser, W. R. Jacobs Jr., and J. S. Blanchard. 1998. Mechanisms of isoniazid resistance in *Mycobacterium tuberculosis*: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. *J. Infect. Dis.* 178:769–775.
- Dessen, A., A. Quémard, J. S. Blanchard, W. R. Jacobs Jr., and J. C. Sacchettini. 1995. Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. *Science*. 267:1638–1641.
- Rozwarski, D. A., C. Vilchère, M. Sugantino, R. Bittman, and J. C. Sacchettini. 1999. Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate. *J. Biol. Chem.* 274:15582–15589.
- Mdluli, K., D. R. Sherman, M. J. Hickey, B. N. Kreiswirth, S. Morris, C. K. Stover, and C. E. Barry III. 1996. Biochemical and genetic data suggest that InhA is not the primary target for activated isoniazid in *Mycobacterium tuberculosis*. *J. Infect. Dis.* 174:1085–1090.
- Mdluli, K., R. A. Slayden, Y. Zhu, S. Ramaswamy, X. Pan, D. Mead, D. D. Crane, J. M. Musser, and C. E. Barry III. 1998. Inhibition of a *Mycobacterium tuberculosis* β -ketoacyl ACP synthase by isoniazid. *Science*. 280:1607–1610.
- Vilchère, C., H. R. Morbidoni, T. R. Weisbrod, H. Iwamoto, M. Kuo, J. C. Sacchettini, and W. R. Jacobs Jr. 2000. Inactivation of the *inhA*-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of *Mycobacterium smegmatis*. *J. Bacteriol.* 182:4059–4067.
- Jornvall, H., B. Persson, M. Krook, S. Atrian, R. Gonzalez-Duarte, J. Jeffery, and D. Ghosh. 1995. Short-chain dehydrogenases/reductases (SDR). *Biochemistry*. 18:6003–6013.
- Duax, W. L., J. F. Griffin, and D. Ghosh. 1996. The fascinating complexities of steroid-binding enzymes. *Curr. Opin. Struct. Biol.* 6:813–823.
- Branden, C. I., and J. Toose. 1999. Introduction to Protein Structure, 2nd ed. Garland Publishing, New York.
- Oppermann, U., C. Filling, M. Hult, N. Shafiqat, X. Wu, M. Lindh, J. Shafiqat, E. Nordling, Y. Kallberg, B. Persson, and H. Jörnvall. 2003. Short-chain dehydrogenase/reductases (SDR): the 2002 update. *Chem. Biol. Interact.* 143:247–253.
- Wierenga, R. K., M. C. H. De Maeyer, and W. G. J. Hol. 1985. Interaction of pyrophosphate moieties with α -helices in dinucleotide binding proteins. *Biochemistry*. 24:1346–1357.
- Bottoms, C. A., P. E. Smith, and J. J. Tanner. 2002. A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci.* 11:2125–2137.
- Rossmann, M. G., A. Liljas, C.-I. Branden, and L. J. Banaszak. 1975. Evolutionary and structural relationships among dehydrogenases. *Enzymes*. 11A:61–102.
- Lesk, A. M. 1995. NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* 5:775–783.
- Rescigno, M., and R. N. Perham. 1994. Structure of the NADPH-binding motif of glutathione reductase: efficiency determined by evolution. *Biochemistry*. 33:5721–5727.
- Nishiya, Y., and T. Imanaka. 1996. Analysis of interaction between the *Arthrobacter* sarcosine oxidase and the coenzyme flavin adenine

- dinucleotide by site-directed mutagenesis. *Appl. Environ. Microbiol.* 62:2405–2410.
33. Eschenbrenner, M., L. J. Chlumsky, P. Khanna, F. Strasser, and M. S. Jorns. 2001. Organization of the multiple coenzyme and subunits and role of the covalent flavin link in the complex heterotetrameric sarcosine oxidase. *Biochemistry.* 40:5352–5367.
 34. Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham III, S. DeBolt, D. Ferguson, G. L. Seibel, and P. A. Kollman. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* 91:1–41.
 35. Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117: 5179–5197.
 36. Ryde, U. 1995. Molecular dynamic simulations of alcohol dehydrogenase with a four- or five-coordinate catalytic zinc ion. *Proteins.* 21: 40–56.
 37. Bayly, C. I., P. Cieplak, W. D. Cornell, and P. A. Kollman. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges—the RESP model. *J. Phys. Chem.* 97:10269–10280.
 38. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
 39. Norberto de Souza, O., and R. L. Ornstein. 1997. Effect of periodic box size on aqueous molecular dynamics simulations of a DNA dodecamer with particle-mesh Ewald method. *Biophys. J.* 72:2395–2397.
 40. Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
 41. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23: 327–341.
 42. Norberto de Souza, O., and R. L. Ornstein. 1999. Molecular dynamics simulations of a protein-protein dimer: particle-mesh Ewald electrostatic model yields far superior results to standard cutoff model. *J. Biomol. Struct. Dyn.* 16:1205–1218.
 43. Hünenberger, P. H., A. E. Mark, and W. F. van Gunsteren. 1995. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J. Mol. Biol.* 252:492–503.
 44. van Gunsteren, W. F., and A. E. Mark. 1998. Validation of molecular dynamics simulation. *J. Chem. Phys.* 108:6109–6116.
 45. Baker, E. N., and R. E. Hubbard. 1984. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44:97–179.
 46. Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.
 47. Laskowski, R. A., M. W. McArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.
 48. Guex, N., and M. C. Peitsch. 1997. SWISS_MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis.* 18:2714–2723.
 49. Fillgrove, K. L., and V. E. Anderson. 2000. Orientation of coenzyme A substrates, nicotinamide and active site functional groups in (di)enoyl-coenzyme A reductases. *Biochemistry.* 39:7001–7011.
 50. Grimm, C., E. Maser, E. Mobs, G. Klebe, K. Reuter, and R. Ficner. 2000. The crystal structure of 3 α -hydroxysteroid dehydrogenase/carbonyl reductase from *Comamonas testosteroni* shows a novel oligomerization pattern within the short chain dehydrogenase/reductase family. *J. Biol. Chem.* 275:41333–41339.
 51. Gerratana, B., W. W. Cleland, and P. A. Frey. 2001. Mechanistic roles of thr134, tyr160 and Leu 164 in the reaction catalyzed by did-glucose 4,6-dehydratase. *Biochemistry.* 40:9187–9195.
 52. Kavanagh, K. L., M. Klimacek, B. Nidetzky, and D. K. Wilson. 2002. Crystal structure of *Pseudomonas fluorescens* mannitol 2-dehydrogenase binary and ternary complexes. *J. Biol. Chem.* 277:43433–43442.
 53. Parikh, S., D. P. Moynihan, G. Xiao, and P. J. Tonge. 1999. Roles of tyrosine 158 and lysine 165 in the catalytic mechanism of InhA, the enoyl-ACP reductase from *Mycobacterium tuberculosis*. *Biochemistry.* 38:13623–13634.
 54. Denessiouk, K. A., V.-V. Rantanen, and M. S. Johnson. 2001. Adenine recognition: a motif present in ATP-, CoA-, NAD-, NADP-, and FAD-dependent proteins. *Proteins.* 44:282–291.
 55. Swaminathan, S., G. Ravishanker, and D. L. Beveridge. 1991. Molecular dynamics of B-DNA including water and counterions: a 140-ps trajectory for d(CGCGAATTCGCG) based on the GROMOS force field. *J. Am. Chem. Soc.* 113:5027–5040.
 56. McDonald, I. K., and J. M. Thornton. 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* 238:777–793.
 57. Morris, G. M., D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson. 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* 19:1639–1662.
 58. Smith, P. E., and J. J. Tanner. 1999. Molecular dynamics simulations of NAD⁺ in solution. *J. Am. Chem. Soc.* 121:8637–8644.
 59. Smith, P. E., and J. J. Tanner. 2000. Conformations of nicotinamide adenine dinucleotide (NAD⁺) in various environments. *J. Mol. Recognit.* 13:27–34.