Pontifícia Universidade Católica do Rio Grande do Sul
Faculdade de Biociências
Programa de Pós-Graduação em Biologia Celular e Molecular

Ivani Pauli

Orientador: Dr. Osmar Norberto de Souza

Estudos *in silico* da interação da enzima InhA de *Mycobacterium tuberculosis* com pequenas moléculas do tipo fármaco

Porto Alegre, RS
2011

Ivani Pauli

Estudos *in silico* da interação da enzima InhA de *Mycobacterium tuberculosis* com pequenas moléculas do tipo fármaco

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Celular e Molecular da Pontifícia Universidade Católica do Rio Grande do Sul como requisito para a obtenção do título de Mestre em Biologia Celular e Molecular.

Porto Alegre, RS
2011

AGRADECIMENTOS

Ao Prof. Dr. Osmar Norberto de Souza agradeço pela oportunidade de ter trabalhado no seu laboratório e por ter dividido comigo valiosos conhecimentos em nossas longas conversas sobre os mais diversos assuntos. Também agradeço pela confiança em mim e no meu trabalho e pela dedicação e cuidado com que conduziu a minha orientação durante estes dois anos.

Aos Professores Dr. Diógenes Santiago Santos Dr. Luiz Augusto Basso pelo apoio indispensável em momentos importantes para que este trabalho se concretizasse, assim como pelos conhecimentos compartilhados que contribuíram sobremaneira para o aumento da qualidade científica desta pesquisa.

Ao Professor Dr. Adriano D. Andricopulo, ao Professor Rafael V. C. Guido, ao doutorando Ricardo N. Santos, à doutoranda Lívia de Barros Salum e aos demais membros do Laboratório de Química Medicinal e Computacional do Instituto de Física de São Carlos da Universidade de São Paulo pelo apoio e conhecimentos compartilhados durante o período em que passei na USP fazendo parte deste trabalho.

Aos meus pais, Isidoro e Valesca, não tenho palavras suficientes para agradecer a compreensão em momentos de ausência, pelas palavras de ânimo e coragem em momentos difíceis, pelo apoio incondicional durante toda a minha vida acadêmica e por nunca terem medido esforços para que “no final tudo desse certo!” Sem vocês certamente eu não teria conseguido chegar até aqui.

Às minhas irmãs, Vânia, Sílvia e Cíntia pelo carinho, pelo companheirismo e pelo apoio durante estes anos de estudo, os quais tornaram esta jornada mais prazeirosa.

Ao Luís Fernando, companheiro sem igual, pelo carinho, pelo apoio incondicional em todos os momentos, pela compreensão e discussões sobre todos os assuntos, novas metodologias e novidades em geral que surgiam ao longo deste período e que foram imprescindíveis para que este trabalho fosse realizado. Enfim, não tenho palavras suficientes para expressar sua importância.

Aos meus colegas do LABIO pelo companheirismo durante este período e a todas às pessoas que de alguma forma, direta ou indiretamente, contribuíram para que este trabalho tivesse êxito.

RESUMO

O gene *inh*A de *Mycobacterium tuberculosis* (Mtb) codifica a enzima enoil redutase, InhA, uma enzima chave no ciclo de alongamento de ácidos graxos tipo II e têm sido validada como um alvo efetivo para o desenvolvimento de agentes anti-microbianos. A InhA catalisa a redução NADH-dependente da ligação dupla trans entre as posições C2 e C3 de substratos de ácidos graxos. Ela é o alvo da isoniazida, uma droga de primeira linha no tratamento da tuberculose. Mutações no gene estrutural da InhA estão associadas com a resistência à isoniazida *in vivo*.

Mesmo que mutações no gene *inh*A sejam conhecidas por facilitar a resistência à isoniazida, InhA ainda é uma excelente candidata a alvo para o planeamento de fármacos porque: (i) a grande maioria das mutações encontradas em isolados clínicos de cepas resistentes à isoniazida estão associadas com o ativador da isoniazida (KatG catalase-peroxidase); (ii) apenas uma enoil-ACP redutase é encontrada em *M. tuberculosis*, ao contrário de outras enzimas dos sistemas FAS-II de bactérias; (iii) a especificidade da InhA por substratos de cadeia mais longa a distingue das enoil-ACP redutases de outros tipos.

Nosso objetivo com este trabalho foi de analisar em detalhe as informações estruturais e físico-químicas disponíveis sobre a InhA de Mtb utilizando ferramentas de bioinformática. Como resultado, foi desenvolvido um modelo farmacofórico baseado na estrutura do sítio de ligação ao substrato da InhA, permitindo a aplicação de uma metodologia de triagem virtual focada em selecionar ligantes que satisfizessem essas características, permitindo assim, a melhor complementariedade com a proteína alvo. Além disso testamos a habilidade de quatro algoritmos de docagem molecular em encontrar conformações semelhantes para uma mesma molécula, fornecendo indícios de que esta seja a conformação mais próxima àquela adotada *in vivo*. Por fim, simulações de dinâmica molecular foram empregadas para melhor compreensão da interação da enzima InhA com um inibidor já conhecido desta enzima.

ABSTRACT

The *inhA* gene from *Mycobacterium tuberculosis* (Mtb), encodes for an enoyl acyl carrier protein reductase, InhA, a key enzyme of the mycobacterial type II fatty acid elongation cycle and has been validated as an effective target for the development of anti-microbial agents. InhA catalyzes the NADH-dependent reduction of *trans* double bond between positions C2 and C3 of fatty acyl substrates. It is the target of isoniazid, a first line drug in the tuberculosis treatment. Mutations in InhA structural gene are associated with isoniazid resistance *in vivo*.

Even though mutations within the *inh*A gene are known to facilitate isoniazid resistance, InhA remains a good candidate for drug design because: (i) the vast majority of the mutations found in isoniazid-resistant clinical isolates are associated with the isoniazid activator (KatG catalase-peroxidase); (ii) only one enoyl-ACP reductase is found in Mtb, unlike some of the other enzymes of bacterial FAS-II systems; (iii) the longer substrate chain length specificity of InhA distinguishes it from the enoyl-ACP reductases from other sources.

Our goal with this work was to analyze in detail the structural and physicochemical available information about Mtb InhA using bioinformatics tools. As a result, we developed a pharmacophoric model based on the InhA substrate binding cavity that allowed the application of a virtual screening methodology focused in selecting ligands that satisfied these features, allowing so, a best complementarity with the target protein. Besides we tested the hability of four docking algorithms to find similar conformation to a molecule, providing clues that this would be the conformation closest that adopted *in vivo*. Finally, molecular dynamics simulations were employed to achieve a better comprehension of the interaction between InhA and a known inhibitor.

LISTA DE ILUSTRAÇÕES

LISTA DE SIGLAS E ABREVIATURAS

AIDS: Síndrome da Imunodeficiência Humana

Apo enzima: enzima livre, não complexada com ligante

CPBMF: Centro de Pesquisas em Biologia Molecular e Funcional

DOTS: *Directly Observed Treatment, Short Course*, tratamento padrão de "curta duração preconizado contra a TB pela OMS.

ENRs: enoil redutases

FAS: sistema de síntese de ácidos graxos (*Fatty Acid Synthase*)

HIV: Vírus da Imunodeficiência Humana

INCT-TB: Instituto Nacional de Ciência e Tecnologia em Tuberculose

InhA: 2-*trans*-enoil-ACP (CoA) Redutase

LABIO: Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas

MDR-TB: TB multi-resistente à drogas

Mtb: *Mycobacterium tuberculosis*

OMS: Organização Mundial da Saúde

PDB: Banco de Dados de Proteínas (*Protein Data Bank*)

SDR: Família das desidrogenases de cadeia curta (*short chain dehydrogenase*)

TB: Tuberculose.

TDR-TB: TB totalmente resistente à drogas.

TECNOPUC: Parque Tecnológico da PUCRS

XDR-TB: TB extensivamente resistente à drogas

WHO: *World Health Organization*

SUMÁRIO

1 INTRODUÇÃO

Tendo em vista a gravidade do problema da Tuberculose (TB) a nível mundial, a pesquisa para o desenvolvimento de novos fármacos para o seu tratamento é de extrema relevância e urgência, uma vez que desde a década de 1980 não é desenvolvido nenhum medicamento novo contra essa doença.

A enzima InhA (2-*trans*-enoil-ACP (CoA) Redutase) é um dos alvos mais bem estabelecidos e validados para o planejamento de fármacos anti-TB (Agüero *et al*., 2008) e foi a enzima alvo do presente Projeto de Pesquisa.

Nos capítulos seguintes apresentaremos a enzima InhA, seu papel no metabolismo do *Mycobacterium tuberculosis* (Mtb) e a relevância da sua escolha como alvo para o desenvolvimento de novos fármacos anti-TB. Mais especificamente, ainda no capítulo 1, apresentaremos a descrição da enzima InhA de Mtb, seu papel bioquímico, sua importância como alvo terapêutico e a motivação para a realização deste trabalho. No capítulo 2 apresentaremos os objetivos da nossa pesquisa. Em seguida, o capítulo 3, descreverá a justificativa para o trabalho. A metodologia a ser empregada será apresentada no capítulo 4. Para finalizar, o capítulo 5 descreverá as considerações finais sobre o trabalho desenvolvido.

1.1 A enzima InhA

A enzima InhA, ou 2-*trans*-enoil-ACP (CoA) Redutase (EC 1.3.1.9) é uma das enzimas essenciais envolvidas no ciclo de alongamento de ácidos graxos em Mtb. Ela é responsável por catalisar a última reação da sequência de alongamento da cadeia de ácidos graxos do sistema FASII (seção 1.2, **Figura 8**).

A InhA catalisa a 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**), cuja porção não lipídica pode ser tanto uma proteína carreadora de acilas (ACP) como a coenzima A (CoA) (Quémard *et al*., 1995). Neste processo, a enzima InhA

catalisa a transferência de um hidreto (H^-) do NADH para o substrato, resultando na redução da dupla ligação (entre C_2 e C_3) da porção lipídica do substrato.

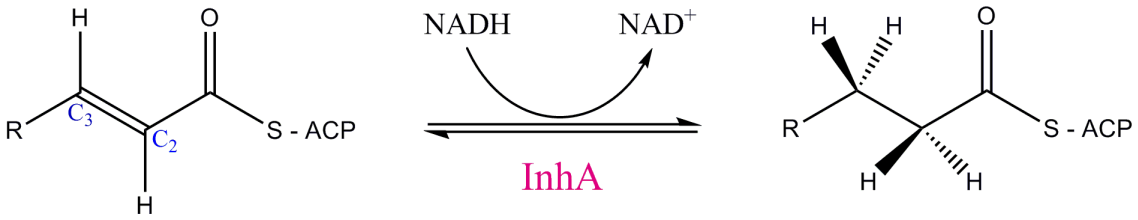


Figura 1. Reação de redução de um tioéster α,β-insaturado catalisada pela enzima InhA. A imagem foi gerada com o programa ChemDraw (Mills, 2006).

A InhA tem preferência por substratos de cadeia carbônica longa (12-24 carbonos), característica consistente com sua função na síntese de ácidos graxos saturados precursores de ácidos micólicos (Blanchard, 1996).

Estudos do mecanismo enzimático em estado estacionário da InhA de Mtb indicaram que, apesar do mecanismo cinético não ser estritamente ordenado, o NADH se liga preferencialmente à enzima seguido pela ligação do substrato (Quémard *et al*., 1995).

1.1.1 A estrutura da InhA

A enzima InhA é codificada pelo gene *inhA* e possui 268 aminoácidos com massa molecular de ~29 kD. Pertence à família de enzimas desidrogenase/redutase de cadeia curta, as SDR (*short chain dehydrogenase/reductase*), que utilizam como coenzima uma molécula de NAD(H) ou NADP(H) (Parikh *et al*., 2000). Essa família de proteínas é caracterizada por apresentar uma topologia onde cada subunidade consiste em um domínio único com o núcleo central do tipo *Rossmann fold*, onde se encontra o sítio de ligação da coenzima (Rossmann *et al*., 1975). De forma geral, a estrutura da enzima se parece com uma cadeira e é composta por sete fitas β e oito hélices α. A coenzima se liga em uma cavidade entre o “encosto” e o “assento” da

estrutura. O sítio de ligação do substrato localiza-se em uma cavidade no “encosto” (**Figura 2**) (Dessen *et al*., 1995).

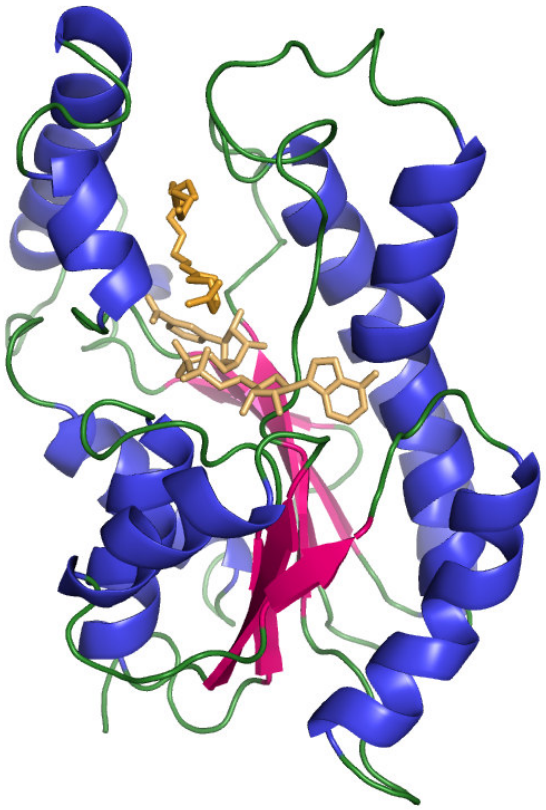


Figura 2. Estrutura monomérica do complexo ternário da enzima InhA com o NADH e o análogo C_{16} do substrato (código PDB: 1BVR). Representação do tipo *Ribbons* da cadeia principal da subunidade C. As hélices α estão representadas em azul, as fitas da folha β em magenta (caracterizando a topologia do tipo *Rossmann fold*), as voltas e alças estão destacadas em verde. O NADH (na cor salmão) está posicionado sobre o leque de fitas β e o análogo do substrato (em laranja) encontra-se na cavidade de ligação, logo acima do NADH. A Imagem foi gerada com o programa PyMol (DeLano, 2002).

Várias hélices α e fitas β do *Rossmann fold* se prolongam acima do sítio de ligação do NADH, criando uma fenda profunda onde se liga o substrato (Sacchettini & Blanchard, 1996).

Apesar de as estruturas da InhA com códigos de acesso PDB 1ENY (Dessen, *et al*., 1995), 1ENZ (Dessen, *et al*., 1995) e 1ZID (Rozwarski *et al*., 1998) terem sido resolvidas anteriormente e já indicarem a região de ligação do derivado de ácido graxo, o modo de ligação do substrato só foi determinado em 1999 com a resolução da estrutura cristalina do complexo ternário da InhA com a coenzima NADH e um

análogo de substrato; um tioéster formado por um derivado de ácido graxo α,β-insaturado de 16 carbonos ligado ao grupo N-acetilcisteamina ou C_{16} (Rozwarski *et al.*, 1999). Por representar de uma forma mais completa a estrutura da enzima InhA, devido a presença tanto da coenzima quanto do substrato, utilizamos para ilustrar a estrutura da InhA nas imagens deste trabalho, as coordenadas do arquivo PDB 1BVR.

Análises de cromatografia por exclusão de tamanho demonstraram que a InhA é um homotetrâmero (**Figura 3**) em solução (Quémard *et al.*, 1995).

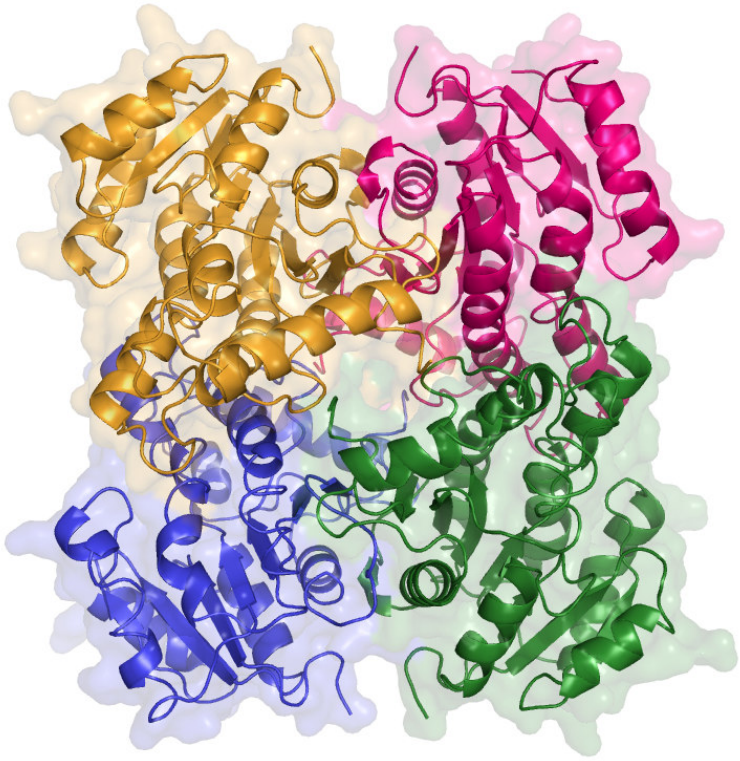


Figura 3. Unidade biológica funcional (homotetrâmero) da InhA de Mtb (código PDB: 1BVR). Representação do tipo *Ribbons* da cadeia principal das subunidades C (rosa), D (verde), E (azul) e F (laranja) com a respectiva representação de suas superfícies moleculares (em transparência). Imagem gerada com o programa PyMol (DeLano, 2002).

1.1.2 O sítio ativo da InhA

Os aminoácidos que delimitam o lado mais externo (em relação ao tetrâmero) da cavidade de ligação, formam duas hélices α transversais sustentadas por alças e são chamados de “alça de ligação ao substrato” (resíduos 196-219) (Dessen, *et al.*,

1995) (**Figura 4**). A alça de ligação ao substrato é um motivo estrutural da InhA e é maior nesta enzima do que nas demais enoil redutases (ENRs), o que está de acordo com a especificidade da ENR de Mtb de reduzir substratos maiores (Rozwarski *et al*., 1999). Os aminoácidos do lado oposto formam duas grandes alças igualmente importantes para a ligação do substrato no sítio ativo (alças A e B). Estas encontram-se do lado da enzima que faz contato com as demais subunidades do homotetrâmero, enquanto as duas hélices da alça de ligação ao substrato estão voltadas para fora do tetrâmero, em direção ao solvente.

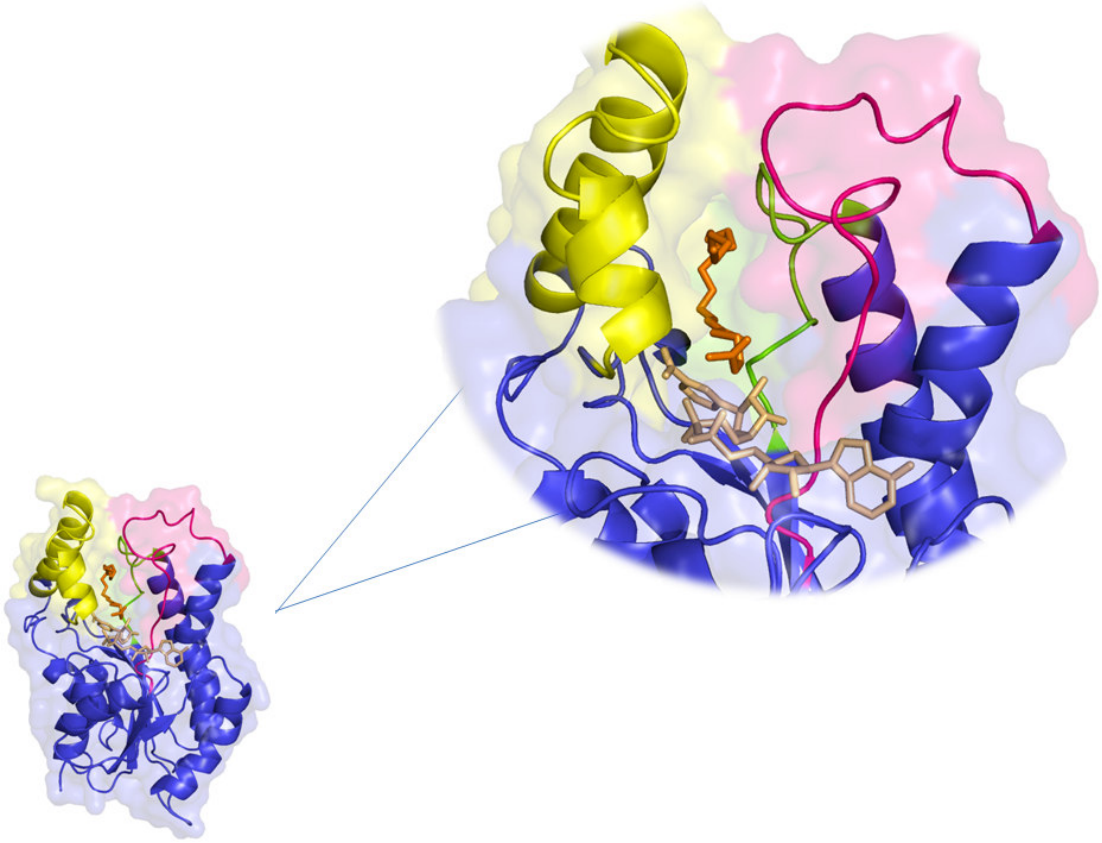


Figura 4. Sítio ativo da enzima InhA de Mtb. Representação do tipo *Ribbons* da cadeia principal da subunidade C (código PDB: 1BVR). Em amarelo destaca-se o motivo estrutural da alça de ligação ao substrato. As alças A e B são mostradas em magenta e verde, respectivamente. O NADH está colorido na cor salmão e o análogo de substrato em laranja, ambos representados pelo modelo de palitos. Figura gerada com o programa PyMol (DeLano, 2002).

Os sítios ativos das quatro subunidades, distantes entre si mais de 40 Å considerando-se subunidades vizinhas, estão voltados para lados opostos na estrutura quaternária da InhA (**Figura 5**). Essa distância torna os sítios ativos de

cada monômero independentes juntamente com a falta de evidências existentes sobre efeitos alostéricos entre subunidades adjacentes na enzima selvagem, justificando a utilização da estrutura monomérica da InhA em simulações computacionais.

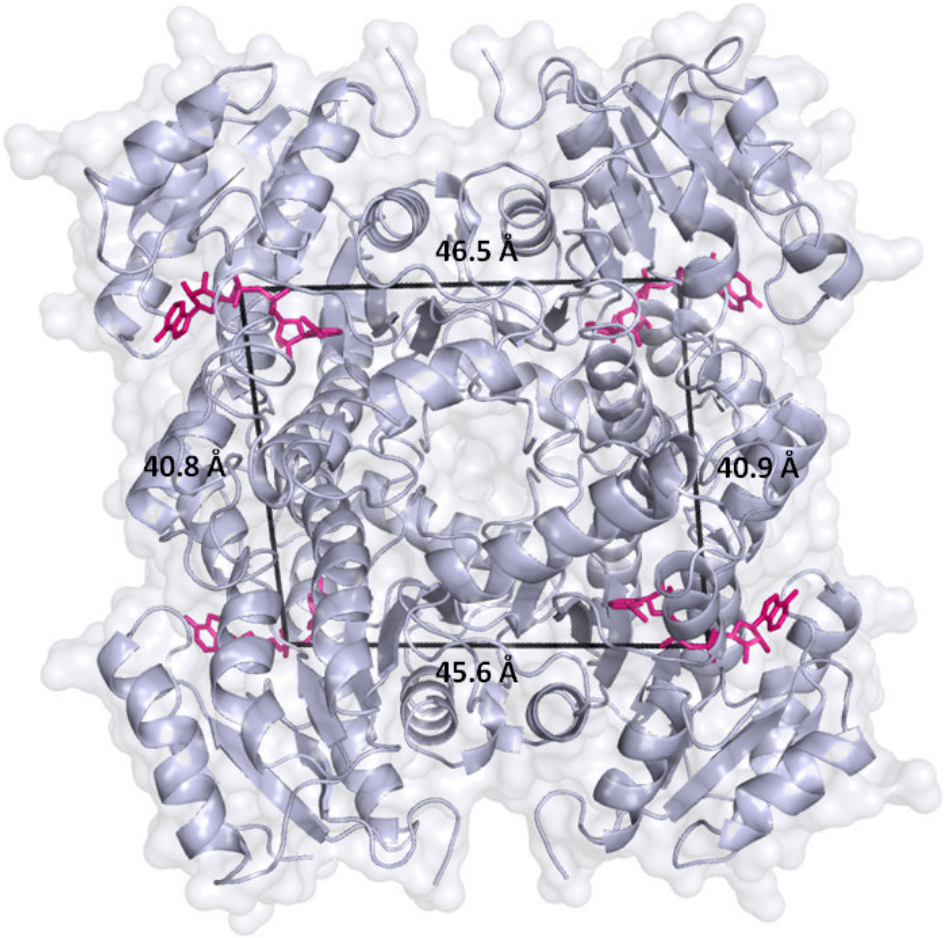


Figura 5. Distância entre os sítios ativos presentes em cada monômero da InhA de Mtb. Representação do tipo *Ribbons* da cadeia principal, a superfície molecular da enzima InhA e as medições foram geradas com o programa PyMol (DeLano, 2002). Na figura, as distâncias entre os monômeros adjacentes foram medidas tendo-se como referência o átomo PA do NADH (magenta), presente no sítio ativo de cada monômero da InhA. Esse átomo foi escolhido por estar localizado praticamente no centro do sítio ativo.

Em cada uma das subunidades, o NADH assume uma conformação estendida dentro da sua cavidade de ligação, ao longo das porções carbóxi-terminais do núcleo de fitas β (com exceção de β4 e β5, cujas porções carboxi-terminais se prolongam além do NADH). O anel nicotinamida do NADH se liga ao fundo da cavidade de ligação, próximo ao encosto da “cadeira” (Figura 2), enquanto

a porção adenina se direciona para o lado oposto. Acima do NADH, sobre o anel da nicotinamida, o substrato C_{16} assume uma conformação em forma de "U" e é fixado pela alça de ligação ao substrato e outros resíduos de aminoácidos do seu sítio de ligação. Os resíduos de cada monômero da InhA que fazem algum tipo de interação tanto com o NADH quanto com o análogo C_{16} do substrato (a distância de até 4 Å dos seus átomos) são a GLY14, ILE15, ILE16, SER20, ILE21, PHE41, LEU63, ASP64, VAL65, GLN66, SER94, GLY96, GLN100, MET103, ILE122, MET147, ASP148, PHE149, ALA157, TYR158, MET161, LYS165, ALA191, GLY192, PRO193, ILE194, THR196, ALA198, MET199, ILE202 e ILE 215. Essas informações encontram-se resumidas na **figura 6**.

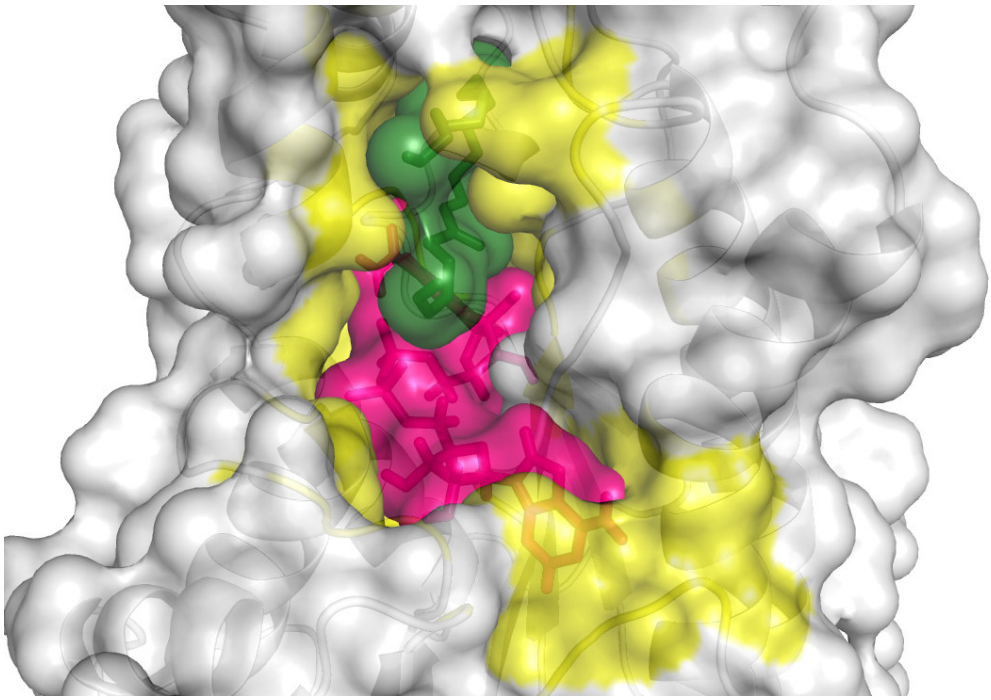


Figura 6. Sítio ativo da enzima InhA de Mtb. Representação do tipo *Ribbons* da cadeia principal e superfície molecular (em branco) da subunidade C (código PDB: 1BVR). A superfície molecular do NADH está representada em magenta e a do análogo C_{16} do substrato em verde. Em amarelo estão destacados os resíduos do sítio ativo que fazem interações com o NADH e com o C_{16}. A imagem foi gerada com o programa PyMol (DeLano, 2002).

A cavidade de ligação ao 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 C_{16} fica na direção da abertura principal, o que é consistente com a necessidade de essa região do substrato natural ficar voltada para fora, em direção ao solvente, quando ligada à InhA (Rozwarski *et al*., 1999).

Observando a cavidade de ligação dos ligantes NADH e C_{16} na InhA (**Figura 6**), pode-se supor que a entrada de um substrato grande do tipo C_{16} não seja fácil. Conforme proposto por Rozwarski *et al*., 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 ao substrato (formado por essas duas hélices) é externa e não faz contato com outras subunidades do tetrâmeto e a abertura em direção ao solvente está desimpedida (Rozwarski *et al*., 1999).

Além disso, na parte superior da InhA, em cima do sítio de ligação do C_{16}, na região de contato entre as duas hélices à esquerda e as alças da direita do substrato, há o predomínio de aminoácidos não carregados e as interações entre estes dois segmentos da molécula são do tipo hidrofóbicas.

O movimento de abertura da alça de ligação para a entrada do substrato não causaria o rompimento de ligações de hidrogênio, por exemplo, tornando possível o movimento de abrir e fechar sugerido para a entrada do susbstrato no sítio ativo.

1.2 A parede celular do *Mycobacterium tuberculosis*

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 7**). Apesar de várias moléculas serem importantes para a formação da estrutura da parede celular, os ácidos micólicos são os componentes majoritários. Eles 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}) (Barry *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. A natureza dos ácidos

micólicos tem um papel fundamental na determinação da fluidez e permeabilidade do envelope celular do Mtb. Modificações em sua estrutura ou composição levam a grandes alterações das características da parede celular, podendo resultar até mesmo na inviabilidade da bactéria (Barry *et al*., 1998). 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 a sobrevivência do Mtb (Kuo *et al*., 2003).

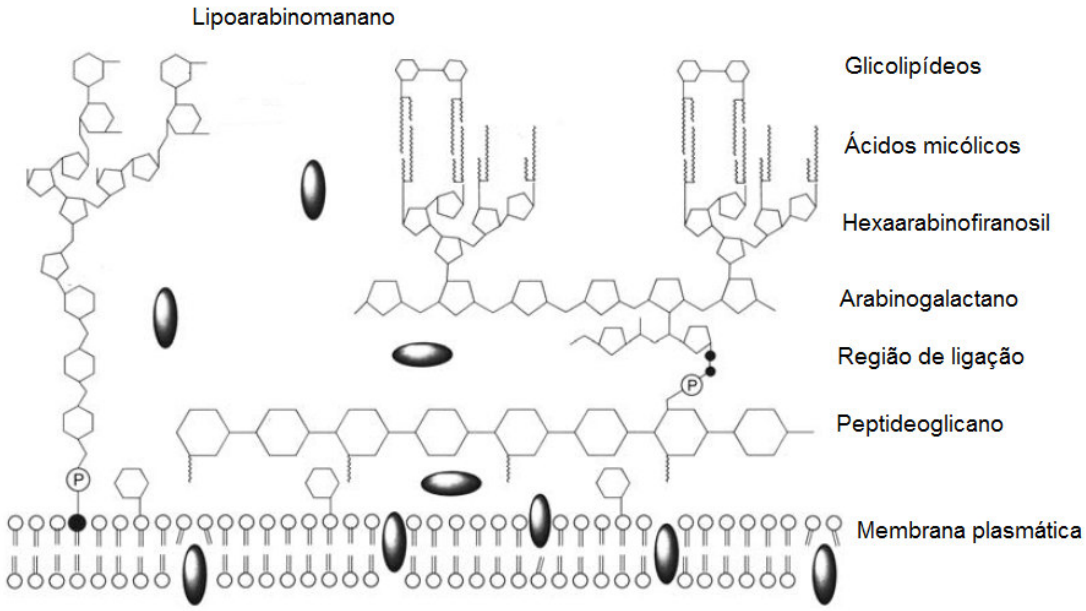


Figura 7. Representação esquemática da parede celular do Mtb (Schroeder *et al*., 2002).

No próximo sub-capítulo falaremos sobre a biossíntese de ácidos graxos em Mtb, onde a InhA desempenha um papel essencial no ciclo de alongamento de ácidos graxos, substratos para a síntese de ácidos micólicos.

1.3 A biossíntese de ácidos graxos em *Mycobacterium tuberculosis*

Dentre os genes envolvidos na biossíntese 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. O sistema FAS é subdividido em dois subtipos. No sistema FASI todas as atividades enzimáticas são codificadas por um único ou por dois polipeptídeos multifuncionais (Wakil *et al*., 1983; Kolattukudy *et al*., 1981), enquanto o sistema FASII é composto por um conjunto de enzimas distintas e monofuncionais (Rock & Cronan, 1996; Kater *et al*., 1994; Fulco, 1983). O sistema FASI está presente em mamíferos, aves e leveduras e o sistema FASII é mais comum em bactérias e plantas.

O Mtb apresenta ambos os sistemas, FASI e FASII. O sistema FASI produz ácidos graxos de cadeias curtas (C_{16} e C_{18}) e longas (C_{24} e C_{26}), os produtos usuais da síntese *de novo* de ácidos graxos (Kolattukudy *et al*., 1997). Tanto a enzima polifuncional de FASI, quanto o sistema multienzimático de FASII são constituídos de proteínas citoplasmáticas (Salman *et al*., 1999). O sistema multienzimático FASII, ACP-dependente é incapaz de realizar síntese *de novo*. Ele é responsável pelo alongamento do palmitoil-ACP (um C_{16}) a ácidos graxos de cadeias longas entre C_{24} e C_{56}.

Com o sequenciamento completo do genoma do Mtb H37Rv (Cole *et al*., 1998; Camus *et al*., 2002), os genes que codificam as enzimas do sistema FASII foram identificados e descobriu-se que o alongamento da cadeia de ácidos graxos em Mtb é feito por quatro enzimas diferentes. 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, 2000), liberando o AcpM e CO_2.

O grupo ceto do β-cetoester resultante é então reduzido pela enzima **MabA**, uma β-cetoacil-ACP redutase NADPH-dependente (Marrachi *et al*., 2002; Banerjee *et al*., 1998), formando β-hidroxiacil-ACP. A eliminação do grupo β-OH por uma β-hidroxiacil-ACP desidratase resulta na formação de 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 NADH-dependente, levando ao acil-AcpM correspondente, que novamente é condensado à malonil-AcpM, até que o ácido graxo de tamanho desejado seja sintetizado (**Figura 8**).

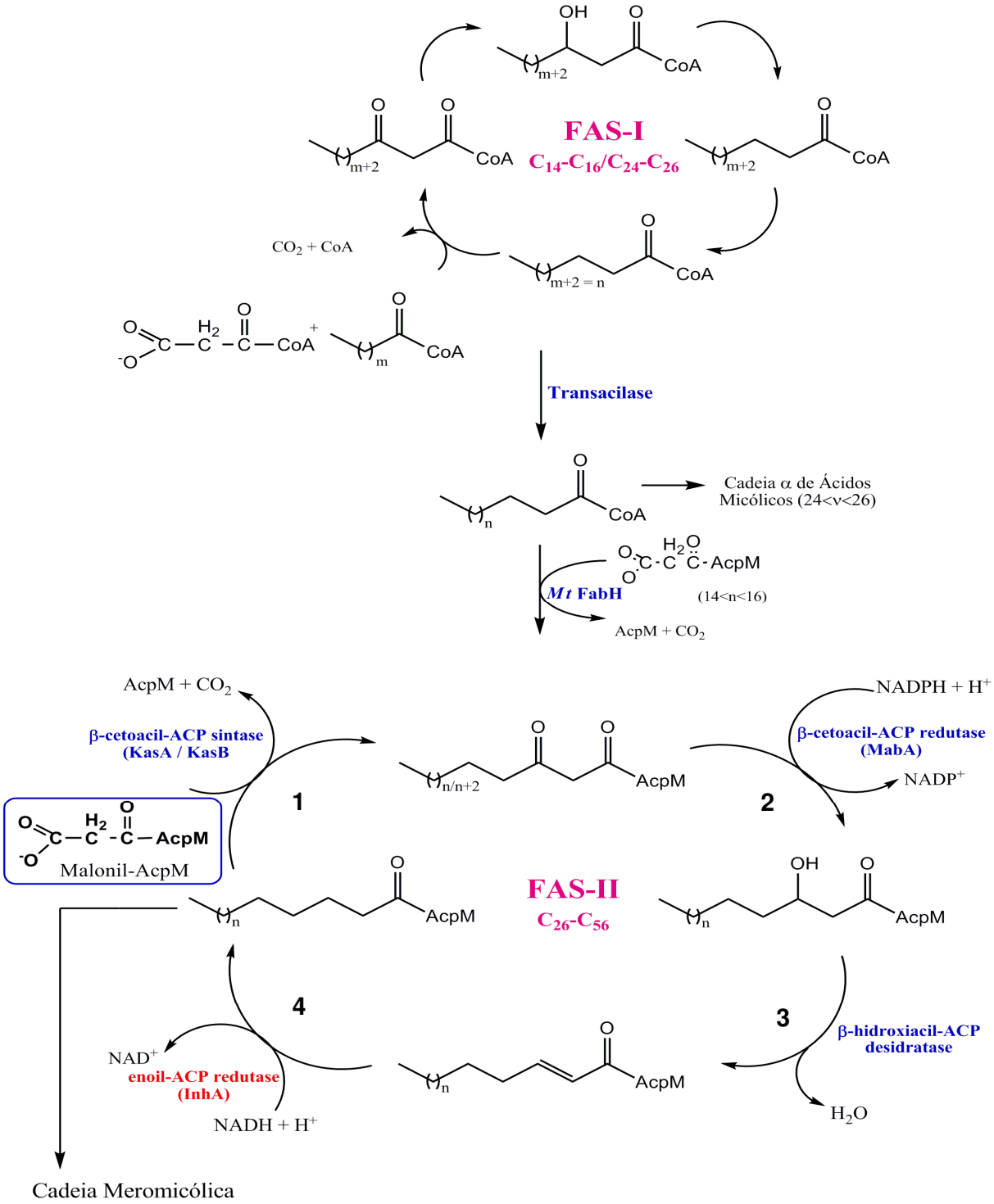


Figura 8. Biossíntese de ácidos graxos em Mtb. FASI sintetiza ácidos graxos numa distribuição bimodal $C_{16\text{-}18}$ e $C_{24\text{-}28}$, que são removidos do complexo enzimático por uma transacetilase na forma de acil-CoA de cadeia longa. A enzima FabH, com especificidade pelos produtos de cadeia menor do sistema FASI, catalisa a condensação desses precursores acil-CoA com o malonil-ACP. O produto é alongado por FASII, resultando em cadeias carbônicas longas, precursoras da cadeia meromicólica ($C_{50\text{-}60}$) dos ácidos micólicos. Os produtos mais longos do sistema FASI ($C_{24\text{-}26}$) são utilizados (provavelmente ligados à CoA) como substrato na formação da cadeia α dos dos ácidos micólicos (Schroeder *et al.*, 2002).

1.4 Motivação social: a Tuberculose

A tuberculose (TB) humana é uma doença infecto-contagiosa causada principalmente pelo Mtb, uma bactéria patogênica micro-aerofílica que estabelece sua infecção comumente nos pulmões. É a doença infecciosa mais prevalente mundialmente e permanece sendo a segunda maior causa de morte por doença infecciosa, seguindo HIV/AIDS (Frieden *et al*., 2003). A Organização Mundial da Saúde (OMS) estimou que um terço da população mundial, cerca de dois bilhões de pessoas, principalmente nos países em desenvolvimento, estão infectadas com o Mtb. Dentre os indivíduos infectados, oito milhões desenvolvem TB ativa e aproximadamente dois milhões de pessoas morrem dessa doença anualmente (WHO, 2006).

Durante vários anos, o desenvolvimento de medicamentos e o estabelecimento de hospitais exclusivos para o combate à TB trouxeram uma considerável redução na incidência da doença (Ducati *et al*., 2006). Porém, nos últimos anos, o ressurgimento da TB vem ocorrendo tanto em países em desenvolvimento quanto em países do primeiro mundo e isso tem sido atribuído a vários fatores, como o aumento da resistência aos fármacos existentes; ao aumento de usuários de drogas injetáveis; à mudanças na estrutura social; ao aumento de imigrantes de nações de alta prevalência de TB para nações desenvolvidas; ao envelhecimento e aumento da expectativa de vida da população humana; à transmissão ativa em ambientes de acumulação humana (prisões, hospitais, abrigos); à degradação dos sistemas de saúde (Ducati *et al*., 2006; Fätkenheuer *et al*., 1999) e principalmente à pandemia do HIV/AIDS e à resistência aos fármacos existentes.

1.4.1 A co-infecção Tuberculose – HIV/AIDS

O problema da TB aumentou muito nos últimos anos devido à pandemia da AIDS. Indivíduos com imunodeficiências, como os portadores do vírus da imunodeficiência humana (HIV), apresentam uma chance muito maior de desenvolver a TB (Morrison *et al.,* 2008). Outro problema importante é a interação medicamentosa entre os fármacos utilizadas para o tratamento dessas doenças. O uso combinado de um inibidor de protease, um fármaco anti-HIV, com rifampicina é

contraindicado para o tratamento da TB, pois a rifampicina induz fortemente o citocromo P-450, aumentando o metabolismo dos inibidores de protease e transcriptases reversas não-nucleosídicas diminuindo marcadamente suas concentrações no sangue (Di Perri *et al*., 2005).

1.4.2 O tratamento atual e o surgimento de cepas resistentes

O tratamento padrão de "curta duração" DOTS (*Directly Observed Treatment, Short Course*) recomendado pela OMS desde 1995 (WHO, 1995) consiste em uma associação de medicamentos de uso regular por um período longo o suficiente para evitar a resistência e persistência bacteriana. Os dois primeiros meses consistem na administração combinada de isoniazida, rifampicina, pirazinamida e estreptomicina (ou etambutol), seguido da combinação de isoniazida e rifampicina por pelo menos quatro meses adicionais (Ducati *et al*., 2006). O longo período do tratamento da TB envolve efeitos colaterais indesejáveis, levando os pacientes a abandonarem a quimioterapia. Esse comportamento tem um potencial muito grande de provocar o surgimento de cepas de Mtb resistentes a esses fármacos.

Em 2007, havia-se estimado a existência de meio milhão de casos de TB multi-resistente à fármacos (MDR-TB). A MDR-TB ocorre quando uma cepa de Mtb é resistente a isoniazida e rifampicina, dois dos fármacos de primeira linha mais potentes contra a TB (WHO, 2009). Já no final de 2008, 55 países e territórios haviam reportado ao menos um caso de cepas de TB extensivamente resistentes à fármacos (XDR-TB). Cepas XDR foram definidas como sendo resistentes a isoniazida e rifampicina (MDR-TB) e também a fluoroquinolonas e a pelo menos uma das três fármacos injetáveis de segunda linha, geralmente utilizadas no tratamento da MDR-TB (capreomicina, canamicina e amicacina) (Jain & Mondal, 2008). O tratamento da MDR-TB e da XDR-TB consiste em fármacos de segunda linha, mais caros, mais tóxicos, menos eficazes e utilizados por mais tempo do que os fármacos de primeira linha (Jain & Mondal, 2008).

Mais recentemente, foi reportado na literatura o surgimento de cepas de Mtb totalmente resistentes à fármacos (TDR-TB) (Velayati *et al*., 2009). As cepas reportadas apresentaram resistência *in vitro* à todos os fármacos de primeira e segunda linha testados.

1.5 Desenvolvimento de novos fármacos anti-TB utilizando a enzima InhA como alvo

A isoniazida tem sido utilizada como fármaco de primeira linha no tratamento da TB desde 1952 (Middlebrook, 1952). Ela é um pró-fármaco ativado oxidativamente *in vivo* por uma enzima catalase-peroxidase codificada pelo gene *katG*. Essa molécula altamente reativa reage não-enzimaticamente com o NAD+, gerando um aduto isonicotinoil-NAD (INH-NAD) (Argyrou *et al.*, 2007), um inibidor competitivo que se liga fortemente à InhA com uma constante de inibição de 0,75 nM (Rawat, *et al.*, 2003).

Outro grupo de fármacos, as tioamidas, representadas pela etionamida e protionamida, são geralmente considerados fármacos de segunda linha no tratamento da tuberculose. Elas possuem uma estrutura muito parecida com a isoniazida e, assim como ela, são pró-fármacos ativados por uma enzima monooxigenase dependente de flavina, codificada pelo gene *ethA*. Os produtos resultantes formam adutos com o NAD e estes adutos é que são os inibidores da InhA (Wang *et al.*, 2004).

O surgimento de mutações nos genes ativadores tanto da isoniazida (*katG*) quanto da etionamida e da protionamida (*ethA*) está relacionado com a maioria dos casos de resistência a esses fármacos. Por isso, compostos que não requerem essa ativação representam candidatos promissores a driblar esses mecanismos de resistência desenvolvidos pelo Mtb (Freundlich *et al.*, 2009). Listaremos a seguir pesquisas por novas moléculas com essa característica.

Inicialmente podemos citar o triclosano, um inibidor direto da enzima InhA de Mtb (McMurry *et al.*, 1998). Entretanto, ele causa uma inibição relativamente fraca (K_i = 0.2 μM) (Parikh *et al.*, 2000) e não possui uma biodisponibilidade muito boa (Wang *et al.*, 2004). Apesar disso, essa pequena molécula tem sido utilizada como estrutura inicial no planejamento de fármacos com base em estrutura (Freundlich *et al.*, 2009).

A primeira geração de éteres difenil foi desenvolvida com base na estrutura do triclosano e compostos inibidores na faixa de nM e valores de concentração inibitória mínima (MIC) de 1-2 μg/ml contra ambas as cepas de Mtb, sensíveis e resistentes à fármacos (Sullivan *et al.*, 2006) foram identificados. Além disso Freundlich e colaboradores recentemente relataram o desenvolvimento de potentes

derivados de triclosano que demonstraram inibição da InhA de Mtb na faixa de nM com valores de MIC de 5-10 µg/ml (Freundlich *et al*., 2009). Entretanto, todos os compostos descritos são inibidores rapidamente reversíveis da InhA.

Recentemente, Luckner e colaboradores (Luckner *et al*., 2010) desenvolveram um composto derivado do triclosano, 2-(o-*Tolyloxy*)-5-*hexylphenol* (PT70), que possui uma constante de inibição bastante alta para a InhA de Mtb (K_i de 22 pM) e inibição do tipo *slow onset*, com um tempo de residência de 24 minutos.

Além dos derivados de triclosano, também foi relatada a identificação de um análogo de isoniazida contendo uma porção cianoferrato capaz de inibir o Mtb do tipo selvagem porém não apresenta inibição satisfatória para o mutante I21V (Oliveira *et al*., 2004). Além disso, modificações químicas da unidade hidrazina da isoniazida, com o objetivo de bloquear a acetilação, foram realizadas e obtiveram sucesso como inibidores da InhA (Hearn *et al*., 2004).

He e colaboradores (He *et al*., 2006) descreveram a descoberta de uma nova classe de inibidores da InhA de Mtb, as carboxamidas de pirolidina e no ano seguinte o mesmo grupo relatou também o uso de arilamidas como outra nova classe de potentes inibidores para InhA (He *et al*., 2007). Moléculas derivadas do aduto Benzoilhidrazina-NADH e derivadas de NADH também foram estudadas como inibidoras da InhA de Mtb (Bonnaca et al., 2007; Broussy et al., 2005).

Além das abordagens *in vitro* para o descobrimento de novos inibidores para a InhA de Mtb, as abordagens *in silico* tem se tornado muito importantes. É o caso da "*chemical systems biology*" que visa a identificação de novos alvos para fármacos já aprovados pelo FDA (*Food and Drug Administration*) e comercialmente disponíveis para outros fins. Essa abordagem permitiu a identificação do Comtan (entacapona e tolcapona), um fármaco utilizado normalmente para o tratamento do mal de Parkinson, como inibidor da InhA de Mtb, impedindo a ligação do seu substrato (MIC_{99} = 260 µM) (Kinnings *et al*., 2009).

Lu e colaboradores (Lu *et al*., 2009) utilizaram como base em seu trabalho de triagem virtual, o modelo da conformação bioativa de um composto ativo potente complexado à InhA. Eles construíram um modelo baseado na estrutura do farmacóforo e o aplicaram na identificação de conformações bioativas, alinhando-o a carboxamidas de pirolidina para construir modelos 3D de QSAR (*Quantitative Structure Activity Relationship,* 3D-QSAR). Em seguida, o modelo do farmacóforo com a conformação característica foi usado como modelo de busca 3D em uma

base de dados (SPECS) para o descobrimento de novos inibidores diretos da InhA. Os resultados foram selecionados com base na conservação das interações entre a InhA e seus inibidores, mapeados no modelo do farmacóforo, e avaliados utilizando o modelo 3D-QSAR.

Subba Rao e colaboradores (Subba Rao, *et al*., 2008) utilizaram a InhA como alvo para o planejamento de pequenos peptídeos inibidores diretos desta enzima. Utilizando estruturas cristalinas disponíveis da InhA em complexo com um inibidor carboxamida de pirolidina (He *et al*., 2006), foi utilizada a modelagem computacional baseada em estrutura para o planejamento de um inibidor tripeptídico (WYW). Estudos de docagem molecular indicam que este peptídeo tem uma potência 100 vezes maior do que inibidores de InhA conhecidos, sugerindo que este peptídeo pode ser utilizado como composto líder para o desenvolvimento de novos fármacos anti-TB.

Kumar e Siddiqi (Kumar e Siddiqi, 2008) relataram estudos de análise comparativa de campos moleculares (*Comparative Molecular Field Analysis* – CoMFA) e subsequente planejamento *de novo* de ligantes carboxamida de pirolidina utilizando o programa LeapFrog. O modelo de CoMFA desenvolvido a partir dos inibidores neste estudo foi utilizado com sucesso para racionalizar a relação estrutura atividade das carboxamidas de pirolidina.

Apesar de todos os esforços desempenhados por diversos grupos de pesquisa no mundo inteiro na busca por novos inibidores contra a enzima InhA de Mtb, não temos conhecimento, até o momento, de algum composto que tenha chegado a fases avançadas em testes clínicos e prestes a se tornar um fármaco a ser utilizado na terapia anti-TB.

2 OBJETIVOS

2.1 Objetivo geral

Dentro do contexto do Instituto Nacional de Ciências e Tecnologia em Tuberculose (INCT-TB), Área de Drogas, o objetivo geral deste projeto foi descobrir novos compostos químicos líderes, pequenas moléculas do tipo fármaco, para desenvolver novos medicamentos para o combate à tuberculose, utilizando a enzima 2-*trans*-enoil-ACP (CoA) redutase (EC 1.3.1.9) de Mtb como alvo terapêutico.

2.2 Objetivos específicos

- Construção de um modelo farmacofórico 3D baseado nas características físico-químicas da estrutura tridimensional do receptor, a enzima InhA de Mtb, afim de identificar moléculas que satisfaçam interações importantes para a afinidade entre enzima-ligante;
- Realização de triagem (experimentos *in silico*, *dry-lab*) de pequenas moléculas do tipo fármaco ou ligantes em bibliotecas virtuais, como os bancos de dados ZINC;
- Execução de experimentos de docagem molecular destes compostos, utilizando os softwares GOLD, AutoDock 4.2, FlexX e Surflex;
- Identificação dos compostos mais promissores, baseados nos escores da interação InhA-ligante e na inspeção visual das estruturas dos complexos obtidos, resultantes da docagem molecular;
- Execução de simulações de dinâmica molecular, afim de compreender melhor o processo de interação proteína-ligante;
- Proposição da compra dos ligantes descobertos, ou a síntese dos que não podem ser obtidos comercialmente.

3 JUSTIFICATIVA

A necessidade de haver novas opções de tratamento para os pacientes infectados pela TB é evidente. O desenvolvimento de novos agentes quimioterápicos mais eficientes e menos tóxicos para reduzir a duração do tratamento da TB poderia levar a otimização do tratamento da MDR-TB, XDR-TB e TDR-TB e de pacientes co-infectados com o HIV.

Dentre os alvos mais promissores para o planejamento de novos agentes antibacterianos estão as enzimas envolvidas na via metabólica bacteriana de biossíntese de ácidos graxos (FAS) (White *et al*., 2005; Zhang *et al*., 2004; Campbell, 2001). Além do mais, a diferença distintiva na organização molecular entre os sistemas FAS de Mtb e dos mamíferos possibilita o planejamento de inibidores específicos com maior seletividade e menor toxicidade (Kumar & Siddiqi, 2008). Dentre as enzimas dessa via metabólica, a InhA tem recebido grande destaque e é um alvo muito bem estabelecido para o desenvolvimento de fármacos anti-TB (Agüero *et al*., 2008). Experimentos anteriores demonstraram que a InhA é essencial para a produção de ácidos micólicos em *Mycobacteria*, pois a sua inativação, sozinha, é suficiente para induzir o acúmulo de ácidos graxos saturados, alterações e lise da parede celular (Vilchèze *et al*., 2000).

Apesar de mutações que ocorrem no gene *inh*A facilitarem o desenvolvimento de resistência a um dos fármacos de primeira linha mais utilizados para o tratamento da TB, a isoniazida (Mdluli *et al*., 1998), a InhA continua sendo um excelente candidato a alvo para o planejamento de novos fármacos devido: (i) à maioria das mutações encontradas em isolados clínicos resistentes à isoniazida estarem relacionados com o ativador desse pró-fármaco (a enzima catalase-peroxidase, codificada pelo gene *kat*G); (ii) à existência de apenas uma enoil-ACP redutase em Mtb, diferentemente de algumas das outras enzimas do sistema FAS II de bactérias; e (iii) à especificidade da InhA por substratos de cadeias carbônicas mais longas, o que a distingue de outras enoil-ACP recutases, como a enoil-ACP redutase que faz parte do sistema FAS I de humanos (Quémard *et al*., 1996).

Além disso a InhA apresenta a característica de ser "*druggable*", ou seja, é suscetível à ação de inibição por fármacos (Agüero *et al*., 2008), o que a torna ainda mais atrativa como alvo para o desenvolvimento de novos medicamentos.

4 METODOLOGIA

4.1 Local de execução

Para as simulações e análises computacionais foi utilizada a estrutura do Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas (LABIO, integrante do INCT-TB) da Faculdade de Informática da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) e do Laboratório de Química Medicinal e Computacional do Instituto de Física da Universidade de São Paulo (USP), na cidade de São Carlos, SP, sob a supervisão do Professor Dr. Adriano D. Andricopulo.

4.2 Proteína alvo para as simulações computacionais

Um estudo mais aprofundado do mecanismo de ação e inibição da InhA se tornou possível por meio do conhecimento das suas características estruturais, pela resolução de suas estruturas cristalinas. Hoje, existem no Banco de Dados de Proteínas (PDB) (Bermann et al., 2000; Bernstein *et al*., 1977) 36 estruturas cristalinas da InhA de Mtb. Dentre estas estruturas encontram-se a InhA do tipo selvagem e seus mutantes relacionados com resistência à fármacos, tanto na forma apo, quanto complexada com um análogo de substrato, com a coenzima NADH e diversos tipos de ligantes e inibidores.

4.3 Identificação dos resíduos que fazem contato com os ligantes

A identificação dos resíduos que interagem com cada um dos ligantes presentes nas 36 estruturas do PDB, é de extrema importância para entendermos a dinâmica do seu sítio ativo. Para a identificação dos aminoácidos em contato com os ligantes utilizamos o programa LIGPLOT (Wallace *et al*., 1995), um programa amplamente utilizado e bem estabelecido para esta finalidade.

4.4 Avaliação dos volumes das principais cavidades da InhA

Sítios de ligação e sítios ativos de proteínas e DNA são muitas vezes associados com cavidades e bolsões na sua estrutura. O cálculo do volume é outra análise importante na caracterização do sítio ativo de uma enzima. Os volumes das cavidades de ligação da InhA, neste trabalho, foram calculados com CASTp (http://sts.bioengr.uic.edu/castp/) (Dundas *et* al., 2006).

4.5 Região do sítio ativo a ser utilizada

Neste trabalho nosso objetivo foi a busca por novos inibidores que sejam seletivos para a InhA de Mtb. Tendo em mente que o NADH é ligante de diversas outras enzimas no organismo humano e participa de muitas outras vias metabólicas, nosso foco foi o sítio de ligação do substrato. Por isso, nas simulações computacionais de triagem, deixamos o NADH fazendo parte da proteína "bloqueando" seu sítio de ligação, deixando livre apenas o sítio de ligação do substrato.

4.6 Banco de dados de pequenas moléculas

O ZINC (Irwin & Shoichet, 2005) é um banco de dados de acesso livre, disponível em http://zinc.docking.org, que contém informações tridimensionais sobre compostos comercialmente disponíveis em formato compatível com a maioria dos programas de docagem molecular. Atualmente ele se encontra na versão 8.0 e possui uma biblioteca de aproximadamente 19.500.000 moléculas (este número continua em constante crescimento), das quais 13.000.000 são disponíveis comercialmente.

Baseado numa análise prévia da estrutura e características físico-químicas da InhA e de seus inibidores conhecidos, uma pré-seleção de compostos para a triagem virtual foi feita a partir do banco de dados ZINC. Filtros moleculares 2D foram utilizados para identificar moléculas com características favoráveis à interação com a

cavidade de ligação do substrato da InhA: 4 < *LogP* < 7; *ligações rotacionáveis* < 6; *aceitadores de ligações de hidrogênio* < 8; *área da superfície polar* < 40; e 250 < *peso molecular* < 400. O conjunto final de moléculas que possuem essas características foi composto por aproximadamente um milhão de moléculas e foi utilizado para todas as simulações computacionais deste trabalho.

4.7 Abordagens utilizadas

Duas metodologias distintas foram utilizadas neste trabalho com o objetivo de identificar moléculas com potencial inibitório para a enzima InhA de Mtb. Na primeira abordagem, um modelo farmacofórico 3D foi construído com base nas características de todas as 36 estruturas cristalográficas disponíveis no PDB. Também foram utilizadas informações sobre suas características físico-químicas e outras peculiaridades importantes para a interação com ligantes já identificados e descritos na literatura.

Na segunda abordagem, utilizamos quatro programas de docagem molecular já bem estabelecidos na comunidade científica para este fim, cada um com um algoritmo característico para a busca e o ranqueamento da melhor conformação de cada ligante no sítio ativo da enzima.

Também utilizamos a metodologia de dinâmica molecular afim de compreender melhor as interações entre a enzima InhA de Mtb com uma molécula já caracterizada como inibidora desta enzima, *pentacyano(isoniazid)ferrate(II)* (PIF) (Oliveira *et* al., 2006)

4.8 Construção do modelo farmacofórico

Um dos mais modernos conceitos define o farmacóforo como “um conjunto de características estéricas e eletrônicas necessárias para assegurar interações supramoleculares ótimas com um alvo biológico específico para modular (ou bloquear) sua resposta biológica” (Wermuth *et* al., 1998).

A utilização das informações provenientes da estrutura tridimensional de proteínas, para o desenvolvimento de novos compostos que apresentem atividade

biológica, metodologia esta denominada de planejamento de fármacos baseado na estrutura da enzima (SBDD, do inglês *structural-based drug design*), é uma estratégia atrativa, de sucesso e bem estabelecida que vem sendo amplamente utilizada por laboratórios acadêmicos e empresas farmacêuticas (Henry, 2001; Kitchen *et* al., 2004; Jorgensen, 2004; Cavasotto *et* al, 2007; Andricopulo *et* al., 2009).

O SBDD é uma abordagem que requer um conhecimento aprofundado sobre as peculiaridades físicoquímicas da enzima a ser estudada, como, por exemplo, a "drogabilidade" do alvo, tamanho e flexibilidade do sítio ativo, geometria, polaridade, estados de protonação dos resíduos pertencentes ao bolsão de ligação e também o conhecimento de possíveis sítios de ancoramento de moléculas de água, as quais podem estar mediando ligações entre enzima-ligante. A partir destas características, programas como o UNITY (Sybyl 8.0), possibilitam a elaboração de um modelo farmacofórico com base nas características físicoquímicas do sítio ativo da enzima em questão.

Neste trabalho abordaremos este problema, utilizando o programa UNITY, inserido no pacote de programas SYBYL (Tripos Inc., St. Louis, MO). Primeiramente, alguns pontos farmacofóricos foram escolhidos com base em análises prévias das características da enzima alvo, essenciais para a interação enzima-ligante.

Em seguida, as estruturas 2D obtidas de bancos de dados de pequenas moléculas foram convertidas em suas respectivas estruturas 3D, por meio de parâmetros geométricos padrão do pacote SYBYL (Tripos Inc., St. Louis, MO).

Por útimo, utilizaremos o programa UNITY (Tripos Inc., St. Louis, MO) para efetuar a filtragem da biblioteca de pequenas moléculas, selecionando somente aquelas que satisfizerem as restrições impostas pelo modelo farmacofórico.

Após a realização desta triagem, foi necessário que os ligantes selecionados fossem submetidos a novos ciclos de minimização de energia, pois o processo de encaixe dos mesmos nos pontos farmacofóricos pode deformar sua geometria, uma vez que o peso atribuído aos pontos é muito alto no algoritmo do programa.

Para avaliar a interação e o modo de ligação das moléculas com a enzima foi necessária a realização de simulações de docagem molecular destes compostos minimizados.

4.9 Docagem molecular

As propriedades biológicas de uma proteína dependem das suas interações físicas com outras moléculas (Alberts *et al*., 2002). O reconhecimento molecular é a base dos sistemas biológicos e de muitas tecnologias químicas. A habilidade em predizer as conformações de ligação e a energia de interação entre pequenas moléculas orgânicas e macromoléculas biológicas, como proteínas e DNA, é de extrema importância fisiológia e farmacológica (Leach *et al*., 2006).

Nos últimos anos, técnicas computacionais como a docagem molecular foram desenvolvidas para a predição da estrutura de complexos intermoleculares formados entre duas ou mais moléculas e tem sido amplamente empregadas no estudo das interações proteína-ligante (Jiang, 2008).

Nesse tipo de simulação, a macromolécula é chamada de “receptor” e consiste, geralmente em uma proteína ou DNA (macromolécula, proteína e receptor são sinônimos em docagem molecular). A pequena molécula que se associa ao receptor (substrato, inibidor ou outra macromolécula, geralmente outra proteína), é denominado de “ligante”. A região da proteína na qual o ligante se associa é conhecido como sítio de ligação e consiste, geralmente, de uma cavidade na superfície da proteína formada por um arranjo específico dos seus aminoácidos. Esses aminoácios podem pertencer a diferentes porções da cadeia polipeptídica, que se aproximam quando a proteína se enovela. O sítio ativo é, geralmente, uma porção bastante dinâmica da proteína, no sentido de que sofre adaptações que dependem do tipo de ligante que nele se associa (Alberts *et al*., 2002).

O planejamento de fármacos baseado em estrutura tem feito importantes contribuições para a descoberta de novos fármacos nos últimos anos. Recentemente, o aumento da disponibilidade de dados estruturais e o acesso a plataformas computacionais de alta performance tem aumentado significativamente a aplicabilidade desses métodos (Waszkowycz, 2002). Métodos computacionais baseados em docagem de ligantes em proteínas com estrutura tridimensional conhecida desmpenham um papel cada vez mais importante na descoberta de compostos líderes e no desenvolvimento de novos fármacos (Cavasotto & Orry, 2007).

No campo de modelagem molecular, a docagem é um método que prediz a orientação preferida de uma molécula em outra quando estas se ligam uma a outra

para formar um complexo estável. O conhecimento dessa orientação preferida, por sua vez, pode ser usada para predizer a força de associação ou afinidade de ligação entre duas moléculas usando, por exemplo, funções escore.

As simulações de docagem molecular são realizadas pela implementação de algoritmos de docagem em programas de computador específicos. Esses algoritmos são capazes de gerar um grande número de complexos receptor-ligante, classificando-os de acordo com sua energia de interação, identificando aqueles cuja interação receptor-ligante 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 (Krumrine *et al*., 2003; Kuntz, 1992).

Qualquer método de docagem deve explorar o espaço conformacional disponível para a interação entre as duas moléculas e avaliar/classificar as conformações geradas pelo processo de busca (Sotriffer *et al*., 2000).

Vários tipos de interação entre ligante e receptor como: complementariedade geométrica, interações específicas (entre cargas, por exemplo), interações de solvatação/dessolvatação, interações hidrofóbicas e ligações de hidrogênio devem ser modelados (Vieth *et al*., 1998). Quanto mais tipos de interações forem contemplados, melhor o grau de predição da formação do complexo receptor-ligante.

O grande custo computacional de uma função de avaliação rigorosa, leva geralmente à sua simplificação e, consequentemente, a uma perda de precisão no resultado. Os métodos de avaliação podem variar desde os campos de força da mecânica molecular como AMBER (Cornell *et al*., 1995) e CHARMM (Brooks *et al*., 1983), até funções empíricas de avaliação de energia livre (Böhm, 1994) ou funções derivadas da análise de um grande conjunto de complexos conhecidos (*knowledge based functions*) (Muegge & Martin, 1999).

Após estudarmos cuidadosamente as características da docagem molecular descritas resumidamente acima e as características implementadas em diferentes programas de docagem molecular, escolhemos, para este trabalho, os programas GOLD (Jones *et* al., 1997) AutoDock (Morris *et al*., 1998; Morris *et al*., 1996; Goodsell *et al*., 1996; Goodsell & Olson, 1990), FlexX (Rarey et al., 1996) e Surflex-Dock (Jain, 2003) para a realização das simulações computacionais de docagem molecular e triagem virtual de diversos ligantes com a enzima InhA. Todos são programas muito bem estabelecidos pela comunidade científica para a realização

desse tipo de simulações computacionais. Com isso, nosso objetivo foi de verificar se os diferentes algoritmos implementados em cada um dos programas seria capaz de gerar a mesma conformação fornecendo indícios de que este é o modo de interação correto entre a proteína e o ligante.

GOLD (Jones *et* al., 1997), "Genetic Optimization for Ligand Docking", é um algoritmo genético para a docagem de ligantes flexíveis em sítios de ligação de proteínas. O programa dá algumas opções em relação à função de escore e, dentre elas foi escolhida a GoldScore para as simulações deste trabalho.

O AutoDock (Morris *et al.*, 1998; Morris *et al.*, 1996; Goodsell *et al.*, 1996; Goodsell & Olson, 1990) utiliza um campo de força semiempírico de energia livre para avaliar conformações durante simulações de docagem molecular e oferece vários métodos para a busca conformacional. O algoritmo genético lamarckiano, uma variação do algoritmo genétido tradicional foi a escolha para este estudo.

FlexX (Rarey *et* al., 1996) tem implementado um algoritmo de construção incremental onde o fragmento base é automaticamente selecionado e posicionado no sítio ativo utilizando uma abordagem algotítmica baseada numa técnica de reconhecimento de padrões, chamada de clusterização de poses. Após um bom conjunto de posicionamentos ter sido obtido para a base, as porções restantes do ligante são divididas em pequenos fragmentos e incrementalmente 'crescidos' nas bases alternativas.

O algoritmo do programa Surflex-Dock (Jain, 2003) emprega a função escore empírica Hammerhead e utiliza um ligante idealizado do sítio ativo, um protomol, para gerar poses para o ligante por meio de construção incremental e crossover que combina porções de poses distintas.

4.10 Triagem virtual

Convencionalmente, a procura por compostos líderes para o posterior desenvolvimento de fármacos envolve técnicas de triagem experimental em larga escala (*high-throughput screening* – HTS) de todos os espaços químicos possíveis dentro dos limites de um conjunto particular de compostos existentes (Mayr & Bojanic, 2009).

Nos últimos anos o desenvolvimento de métodos de triagem virtual (TV) aumentou enormemente as possibilidades de incluir nessa busca moléculas que, não necessariamente, existem fisicamente mas, podem ser obtidas rapidamente por meio de síntese ou compra.

A TV pode ser definida como o uso de computação de alto desempenho, por meio da execução de *softwares* de simulação de docagem molecular, para analisar grandes bancos de dados de compostos químicos visando a identificação de possíveis candidatos a fármacos (Walters *et al.*,1998). Ela apresenta algumas vantagens importantes como a diminuição (em até ~10^6 a ~10^3) do número de moléculas de uma biblioteca de compostos químicos que precisam ser testadas *in vitro*, o que proporciona uma economia de tempo e dinheiro importante nas pesquisas por novos fármacos.

4.11 Prediçao *in silico* da toxicidade das moléculas

Níveis inaceitáveis de toxicidade ainda é um dos maiores gargalos no processo de descoberta de novas drogas (Helma, 2004). Com o objetivo de abordar este problema, técnicas de predição de toxicidade *in silico* são alternativas (ou suplementos) baratas e eficientes aos ensaios biológicos para a identificação de efeitos tóxicos num estágio anterior do desenvolvimento do produto. Por isso, como uma análise *in silico* adicional para todas as moléculas selecionadas pelas duas abordagens, foi utilizado o programa OSIRIS (Sander *et*. al, 2009) para predizer características tóxicas. OSIRIS (disponível em: http://www.organic-chemistry.org/prog/peo/) é parte integrante do sistema de registro de substâncias da empresa Actelion's. Ele calcula várias propriedades relevantes, enquanto o usuário desenha a molécula em uma interface apropriada.

4.12 Dinâmica Molecular

A simulação por Dinâmica Mlecular (DM) é uma das técnicas mais versáteis para o estudo de macromoléculas biológicas, no âmbito das técnicas *in silico*. Conceitualmente a DM é uma abordagem computacional, na qual são empregadas

equações Newtonianas para a resolução de representações atomísticas, de um sistema molecular baseado na Mecânica Clássica, com o intuito de obter informações a respeito de suas propriedades em função do tempo (Alonso *et al.*, 2006). Os algoritmos utilizados nos programas de DM consistem da solução numérica destas equações do movimento fornecendo uma trajetória (coordenadas e momentos conjugados em função do tempo) do sistema em estudo.

Em 1977, McCammon e colaboradores realizaram a primeira simulação de DM envolvendo proteínas. Esta simulação foi realizada *in vacuo* e o tempo de simulação foi de 8,8 x 10^{-12} s (McCammon *et al.*, 1977). Desde então, a técnica de DM foi se aprimorando, deixando os sistemas construídos *in silico* mais realísticos. Este progresso deve-se tanto a avanços na área da química, com o melhoramento dos parâmetros dos campos de força, quanto da computação, com o desenvolvimento de máquinas mais robustas, o que permite a realização de simulações mais longas, chegando a 10^{-9} e 10^{-8} s.

A DM é empregada em várias áreas, desde o refinamento de estruturas cristalográficas, predição de estruturas protéicas, otimização de parâmetros geométricos, avaliação da interação ligante-receptor, entre outras.

Nesta dissertação, a DM foi empregada para compreender melhor as interações entre a enzima InhA de Mtb com um inibidor já descrito para esta enzima (Oliveira *et* al., 2006). Para isso, foi utilizado o pacote de programas AMBER, versão 9 (Case *et* al., 2006).

4.13 Programas para visualização, manipulação e geração de figuras

Programas como o SwissPdbViewer (Guex & Peitsch*,* 1997), PyMol (DeLano *et al.,* 2002), VMD (Humphrey *et al.,* 1996) e Chimera (Pettersen *et al.*, 2004) serão utilizados para manipulação das estruturas, vizualização e análises de resultados, como também para a geração de imagens.

5 ARTIGOS CIENTÍFICOS

5.1 Artigo em fase de finalização a ser submetido para a revista "*Journal of Medicinal Chemistry*"

Discovery of New Inhibitors of *Mycobacterium tuberculosis* InhA by Pharmacophore Based Virtual Screening

Ivani Pauli[1,2], Ricardo Nascimento dos Santos [3], Luiz Augusto Basso [2], Diogenes Santiago Santos[2], Rafael V. C. Guido[3], Adriano D. Andricopulo [3] and Osmar Norberto de Souza [1,2]*

[1]Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas - LABIO, Faculdade de Informática, PUCRS; [2]Centro de Pesquisas em Biologia Molecular e Funcional - CPBMF, Instituto Nacional de Ciência e Tecnologia em Tuberculose – INCT-TB, PUCRS; [3]Laboratório de Química Medicinal e Computacional - LQMC, Centro de Biotecnologia Molecular Estrutural - CBME, Instituto de Física d Carlos – IFSC-USP.

INTRODUCTION

Tuberculosis (TB) is more prevalent in the world today than at any other time in human history (Koul *et*. Al, 2011). The ressurgence of TB as a public health threat was mainly due to the arising of multi (MDR-TB), extenxively (XDR-TB) (Jain & Mondal, 2008), and recently, totally (TDR-TB) (Velayati *et al*., 2009) drug resistant *Mycobacterium tuberculosis* strains. TB co-infections in HIV/AIDS patients, also contributed significantly to increase this problem (Di Perri *et al*., 2005; Morrison *et al.,* 2008). Besides, *M. tuberculosis*, the TB ethiological agent, is known to use diverse strategies to survive in a variety of host lesions and to evade immune surveillance (Koul *et*. Al, 2011).

The urgency to pursue novel alternative treatments to TB infected patients is clear. To achieve global control of this epidemic, there is a need for new TB drugs, which can: (1) shorten treatment duration; (2) target drug resistant strains; (3) simplify treatment by reducing the daily pill burden; (4) lower dosing frequency; and (5) be co-administered with HIV medications (Koul *et*. Al, 2011).

Among the most attractive molecular targets to the design of novel antibacterial agents are the Fatty Acid Synthase (FAS) pathway enzymes (Agüero *et al*., 2008; White *et al*., 2005; Zhang *et al*., 2004; Campbell & Cronan, 2001). The *Mycobacterium tuberculosis* InhA (*Mt*InhA) or 2-*trans*-enoil-ACP (CoA) reductase (E.C.1.3.1.9), the fourth enzyme of the type II fatty acid synthase system (FAS II), is one of the key enzymes involved in the elongation cycle of fatty acids in *M. tuberculosis*. Its biological role includes the preferentially reduction of long chain enoyl thioester substrates (*e.g.*, containing 16 or more carbon atoms) yielding the long carbon chain of the meromycolate branch from mycolic acids (C_{40-60}), α-branched fatty acids, the hallmark of mycobacteria (Schroeder *et al*., 2002).

Previously, it has been shown that InhA is essential to the mycolic acid biosynthesis in *Mycobacterium* (Vilchèze *et al*., 2000). Its inactivation induces the accumulation of saturated fatty acids, leading to cell wall alterations, lysis and, consequently, to the pathogen dead. Therefore, due to its biological importance as well as structural differences between *M. tuberculosis* and the human homolog, the InhA enzyme has been investigated as an attractive target for drug design (Kumar & Siddiqi, 2008, Agüero *et al*., 2008).

In the present work, we describe the development of a pharmacophore model for ligands of *Mt*InhA, as well as a pharmacophore-based virtual screening approach, which resulted in the identification of compounds with substantial *in vitro* inhibitory activity against *Mt*InhA.

MATERIALS AND METHODS

Molecule library selection. Based on the structural and physicochemical analysis of *Mt*InhA, a preselection of compounds to be screened was carried out from the ZINC Database (Irvin et al., 2005). Molecular 2D filters identified molecules with favorable features to interact with InhA binding cavity: 4 < LogP < 7; rotatable bonds < 6;

hydrogen bond acceptors < 8; Polar Surface Area < 40; and 250 < molecular weight < 400. The final set of compounds gathering those properties was comprised by 1,310,127 molecules and was used in both, first and second approaches which are described below.

APPROACH: 1 Structure Based Pharmacophore Generation

Computational Approach. The pharmacophore search was performed using the SYBYL 8.0 package (Tripos Inc., St. Louis, MO) running on Red Hat Enterprise Linux workstations. The 3D structures of the inhibitors were generated using standard geometric parameters of the molecular modeling software package SYBYL 8.0. Each single optimized conformation of a molecule in the data set was energetically minimized employing the Tripos force field (Clark et al., 1989) and the Powell conjugate gradient algorithm (Powell, 1977) with a convergence criterion of 0.005 kcal/mol·Å. Partial atomic charges were calculated by the Gasteiger-Hückel method (Gasteiger & Marcili, 1980). The analyses, calculations, and visualizations were performed using the programs SYBYL 8.0 and Pymol (DeLano Scientific, Palo Alto, CA).

Pharmacophore model generation. Currently, 36 *Mt*InhA crystal structures are available at the Protein Data Bank (PDB) (Bermann et al., 2000; Bernstein *et al*., 1977). Among them, we found InhA wild type and its mutants related to drug resistance, in apo form (without a ligand) and in complex with NADH (cofactor), substrate analog, and several ligands (**Table 1**).

The 36 crystallographic structures were superimposed by least-squares fit of the Cα atoms with TopMatch (Sippl & Wiederstein, 2008; Sippl, 2008) available at (http://topmatch.services.came.sbg.ac.at/). The binding cavity volumes were calculated with CASTp (http://sts.bioengr.uic.edu/castp/) (Dundas et al., 2006). Hydrogen bond and hydrophobic contacts were also analyzed with LIGPLOT (Wallace *et al*., 1995) in order to get a more robust profile of the interaction pattern between *Mt*InhA and its crystallographic ligands. The structural analysis of *Mt*InhA in complex with the ligands bound to the substrate-binding cavity indicated several key amino acid residues that might be involved in the molecular recognition process. On the basis of that, a four-point 3D pharmacophore model was developed.

Table 1. Relation of the 36 *Mt*InhA structures deposited at PDB.

PDB Code (Mutant)	Resolution (Å)	Ligand	Reference
1ENY	2.2	NADH	Dessen *et al.*, 1995
1ENZ (S94A)	2.7	NADH	Dessen *et al.*, 1995
1ZID	2.7	INH-NADH Adduct	Rozwarski *et al.*, 1998
1BVR	2.8	NADH + Substrate Analog (C_{16})	Rozwarski *et al.*, 1999
1P44	2.7	NADH + GEQ	Kuo *et al.*, 2003
1P45	2.6	NADH + TCL	Kuo *et al.*, 2003
2B35	2.3	NADH + TCL	Sullivan *et al.*, 2006
2B36	2.8	NADH + 5PP	Sullivan *et al.*, 2006
2B37	2.6	NADH + 8PS	Sullivan *et al.*, 2006
2AQH (I21V)	2.0	NADH	Oliveira *et al.*, 2006
2AQI (I47T)	2.2	NADH	Oliveira *et al.*, 2006
2AQ8	1.9	NADH	Oliveira *et al.*, 2006
2AQK (S94A)	2.3	NADH	Oliveira *et al.*, 2006
2NV6 (S94A)	1.9	INH-NADH Adduct	Vilchèze *et al.*, 2006
2H7I	1.6	NADH + 566	He *et al.*, 2006
2H7L	1.7	NADH + 665	He *et al.*, 2006
2H7M	1.6	NADH + 641	He *et al.*, 2006
2H7N	1.9	NADH + 744	He *et al.*, 2006
2H7P	1.8	NADH + 468	He *et al.*, 2006
2NTJ	2.6	PTH-NADH Adduct	He *et al.*, 2006
2H9I	2.2	ETH-NADH Adduct	He *et al.*, 2006
2IDZ	2.0	INH-NADH Adduct	Dias *et al.*, 2007
2IE0 (I21V)	2.0	INH-NADH Adduct	Dias *et al.*, 2007
2IEB (S94A)	2.2	INH-NADH Adduct	Dias *et al.*, 2007
2IED (S94A)	2.1	APO	Dias *et al.*, 2007
2PR2	2.5	INH-NADH Adduct	Argyrou *et al.*, 2007
2NSD	1.9	NADH + 4PI	He *et al.*, 2007
3FNG	1.9	NADH + JPL	Freundlich *et al.*, 2009
3FNH	2.8	NADH + JPJ	Freundlich *et al.*, 2009
3FNE	1.9	NADH +8PC	Freundlich *et al.*, 2009
3FNF	2.3	NADH + JPM	Freundlich *et al.*, 2009
2X22	2.1	NADH + TCU	Luckner *et al.*, 2010
2X23	2.8	NADH + TCU	Luckner *et al.*, 2010
3OEW	2.1	NAD	Molle et al., 2010
3OEY (T266E)	2.0	NAD	Molle et al., 2010
3OF2 (T266D)	1.75 Å	NAD	Molle et al., 2010

GEQ: 5-{[4-(9H-fluoren-9-yl)piperazine-1-yl]carbonyl}-1H-indol; **TCL**:Triclosan; **5PP**: 5-pentyl-2-phenoxyphenol; **8PS**: 5-octyl-2-phenoxyphenol; **INH**: Isoniazid; **566**: (3S)-1-cicloexyl-5-oxo-N-phenylpirrolidine- 3-carboxamide; **665**: (3S)-N-(3-bromophenyl)-1-ciclohexyl-5-oxopyrrolidine-3-carboxamide; **641**: (3S)-1-ciclohexyl-N-(3,5-diclorofenil)-5-oxopyrrolidine-3-carboxamide; **744**: (3S)-N-(5-Cloro-2-metilphenyl)-1-ciclohexyl-5-oxopyrrolidine-3-carboxamide; **468**: (3S)-N-(3-cloro-2-metilphenyl)-1-ciclohexyl-5-oxopirrolidine-3-carboxamide; **PTH**: Prothionamide; **ETH**: Ethionamide; **4PI**: N-(4-metilbenzoyl)-4-benzylpiperidine; **JPL**: 5-(cyclohexa-1,5-dien-1-ylmethyl)-2-(2,4- dichlorophenoxy)phenol; **JPJ**: 2-(2,4-diclorophenoxy)-5-(2-phenylethyl)phenol; **8PC** 2-(2,4-diclorophenoxy)-5-(pyridine-2-ylm etil)phenol; **JPM**: 5-benzyl-2-(2,4-dichlorophenoxy)phenol; **TCU**: 5-hexyl-2-(2-methylphenoxy)phenol.

3D pharmacophore generation and database searching. The molecule library 3D structures were transferred into an UNITY database in SYBYL 8.0. The 3D search with UNITY was based on the generated pharmacophore model. In order to employ the pharmacophore model as 3D constraints, we used the crystallographic spatial coordinates of the ligand 5-{[4-(9H-fluoren-9-yl)piperazin-1-yl]carbonyl}-1H-indole, (Genz-10850), bound to *Mt*InhA (PDB ID: 1P44). The donor and acceptor atoms were defined by connecting the donor and the acceptor atoms of the ligand via the partial match utility. The hydrophobic feature was partially characterized based on two rings present in the ligand structure. The spheres diameters were adjusted in such a way that the test sample, composed of known InhA crystallographic ligands, could be retrieved. To screen the one million molecules library we performed a UNITY 3D flexible search.

Molecular modeling. UNITY algorithm for flexible searches assigns high score to the molecules that match the pharmacophore points, because of that, the optimized geometry of the compounds under evaluation may be affected. Therefore, the molecular structures of the selected set of compounds were minimized in order to avoid false positive results. To energetically minimize the molecules we used the same protocol as described in Computational Approach section. Subsequently, the selected ligand candidates were docked into *Mt*InhA substrate-binding cavity. The GOLD suite of programs (Jones *et* al., 1997) was used to perform the molecular docking simulations. The *Mt*InhA crystallographic structure used as the receptor (PDB ID, 1P44) was chosen among all the 36 available structures, because of the best relation between binding cavity volume (3,007.8 $Å^3$) and protein resolution (2.7Å). For the calculations, hydrogen atoms were added to the protein molecule in standard geometry, using the Biopolymer module as implemented in SYBYL 8.0. The active site was defined incorporating all amino acid residues within a radius sphere of 8.0 Å centered on the crystallographic bound ligand (Genz-10850). The NADH coenzyme was treated as part of the protein in all docking simulations. Molecular modeling studies were carried out using GOLD default parameters. The GoldScore fitness function was employed to select the representative pose for each compound. The solutions were visually checked to find if the poses obtained by docking still matched the original pharmacophore hypothesis. The visual inspection is a common practice important to avoid testing false positives due to assumptions and

shortcomings in docking methods and scoring functions (Klebe, 2005, Postigo 2010). In the visual inspection we critically assessed the suggested binding conformation, the mutual surface complementarity of ligand and protein, and the possible presence of unfilled space along the protein ligand interface. Only those molecules that matched all the four pharmacophore points after the docking simulations were selected to the next step (*in silico* toxicity prediction).

APPROACH 2: Exhaustive docking simulations using different algorithms

Docking Algorithms. Parallel to the structure-based pharmacophore approach, molecular docking simulations were performed using four different and well established docking programs: GOLD (Jones *et* al., 1997), AutoDock (Morris *et al*., 1998; Morris *et al*., 1996; Goodsell *et al*., 1996; Goodsell & Olson, 1990), FlexX (Rarey et al., 1996), and Surflex-Dock (Jain, 2003).

GOLD (Jones *et* al., 1997), the Genetic Optimization for Ligand Docking, is a genetic algorithm for docking flexible ligands into protein binding sites, using the Genetic Algorithm. The program gives some options in relation to the scoring function, and among them we elected GoldScore to be used in our simulations.

AutoDock (Morris *et al*., 1998; Morris *et al*., 1996; Goodsell *et al*., 1996; Goodsell & Olson, 1990) uses a semi empirical free energy force field to evaluate conformations during docking simulations and provides several methods to do the conformational search. The Lamarckian Genetic Algorithm, a modification of the traditional Genetic Algorithm, was our choice for this study.

FlexX (Rarey et al., 1996) has implemented an Incremental Construction Algorithm where the base fragment (the ligand core) is automatically selected and placed into the active site using an algorithmic approach based on a pattern recognition technique called pose clustering. After a good set of placements has been obtained for the base, the remaining portions of the ligand are divided into small fragments and incrementally “grown” onto the base alternatives.

The Surflex-Dock algorithm (Jain, 2003) employs the empirical Hammerhead (Jain, 1996) scoring function and uses an idealized active site ligand, or protomol, to generate ligand poses by incremental construction and a crossover procedure that combines pieces from distinct poses. Protomols were computed using the position of the cognate ligand to define the binding site.

Protein receptor. The protein receptor chosen for the simulations was *Mt*InhA with PDB access code 1P44, and NADH coenzyme was treated as part of the protein. In all docking simulations the receptor structure was kept rigid, while the ligands were treated as totally flexible.

Molecule data set. The same used for the pharmacophore based approach.

Docking simulations. Firstly, all 1,310,127 molecules in the initial dataset were docked into *Mt*InhA using GOLD (Jones *et* al., 1997). The active site was defined incorporating all amino acid residues within a radius sphere of 8.0 Å centered on the 1P44 crystallographic bound ligand (Genz-10850). The 100 molecules top-scored by GOLD were selected to be submitted to the other three programs, AutoDock, FlexX and Surflex-Dock, using the same parameters as in GOLD, being different only the algorithms for searching and scoring the best ligand poses, which are characteristic of each program. The ligands that achieved a very similar conformation, in terms of functional groups positioning, by at least three of the four program algorithms, were selected to the next step (*in silico* toxicity prediction).

***In silico* predictive toxicology.** Unacceptable toxicity is still a major bottleneck in the drug discovery process (Helma, 2004). Aiming at addressing this problem, *in silico* predictive toxicology techniques are fast and cost efficient alternatives (or supplements) to bioassays for the identification of toxic effects at an early stage of product development. Therefore, as an additional *in silico* analysis for all the molecules selected both by the first approach as well as for the second one, we used the OSIRIS drug discovery informatics system (Sander *et*. al, 2009) to predict toxicity features. The goal of the OSIRIS project is to develop integrated testing strategies (ITS) fit for use in the REACH (Registration, Evaluation and Authorization of Chemicals) system, that would enable a significant increase in the use of non-testing information for regulatory decision making, and thus minimize the need for animal testing (Vonk *et* al., 2009). The OSIRIS Property Explorer (available at: http://www.organic-chemistry.org/prog/peo/) is an integral part of Actelion's in house substance registration system. It calculates on-the-fly, while the user draws a molecule, various drug-relevant properties. Prediction results are valued and color coded. Properties with high risks of undesired effects like mutagenicity or a poor

intestinal absorption are shown in red. Whereas a green color indicates drug-conform behavior. The molecules selected from both approaches used in this work were submitted to OSIRIS and all those that presented unfavorable features were eliminated from the final set of molecules to be tested *in vitro*.

Similarity Ensemble Approach (SEA). For the ligands that passed the toxicity analysis we did a search at the SEA (Kaiser et al., 2007) online tool (http://sea.bkslab.org/search/) to have an idea of the relationships between our hits and molecules with described properties and targets.

RESULTS AND DISCUSSION

Approaches 1 and 2. Two distinct methodologies were used in order to identify the most promising inhibitors for the enzyme InhA from *M. tuberculosis*. As a first approach, a 3D pharmacophore model was built based on the features of all the 36 *Mt*InhA crystallographic structures available at PDB and on information from the literature. In the second approach we used four molecular docking programs (described in the materials and methods session), each one with its particular algorithm to search and score the best pose for each ligand.

Protein target peculiarities. InhA is encoded by the *inh*A gene and is composed by 269 amino acids with a molecular weight of ~29 kD. Size exclusion chromatography analysis demonstrates that InhA is a homo-tetramer in solution and this is the biologically active structure (Quémard *et al*., 1995). The active sites of the four monomers, far apart more than 40 Å considering neighboring subunits, are facing opposite sides in the InhA quaternary structure (**Figure 1**), justifying the use of the monomer in computational simulations.

Overall, the InhA monomer structure (**Figure2**) seems like a chair and is composed by seven β sheets and eight α helices (Dessen *et al*., 1995). Each monomer has two binding sites, one for the NADH coenzyme and another for the substrate (fatty acid). Several α helices and β sheets of the *Rossmann fold* extend over the NADH binding site, creating a crevasse where the substrate binds to. Three

loops, the substrate-binding loop, loop A and loop B are important components of the substrate-binding site.

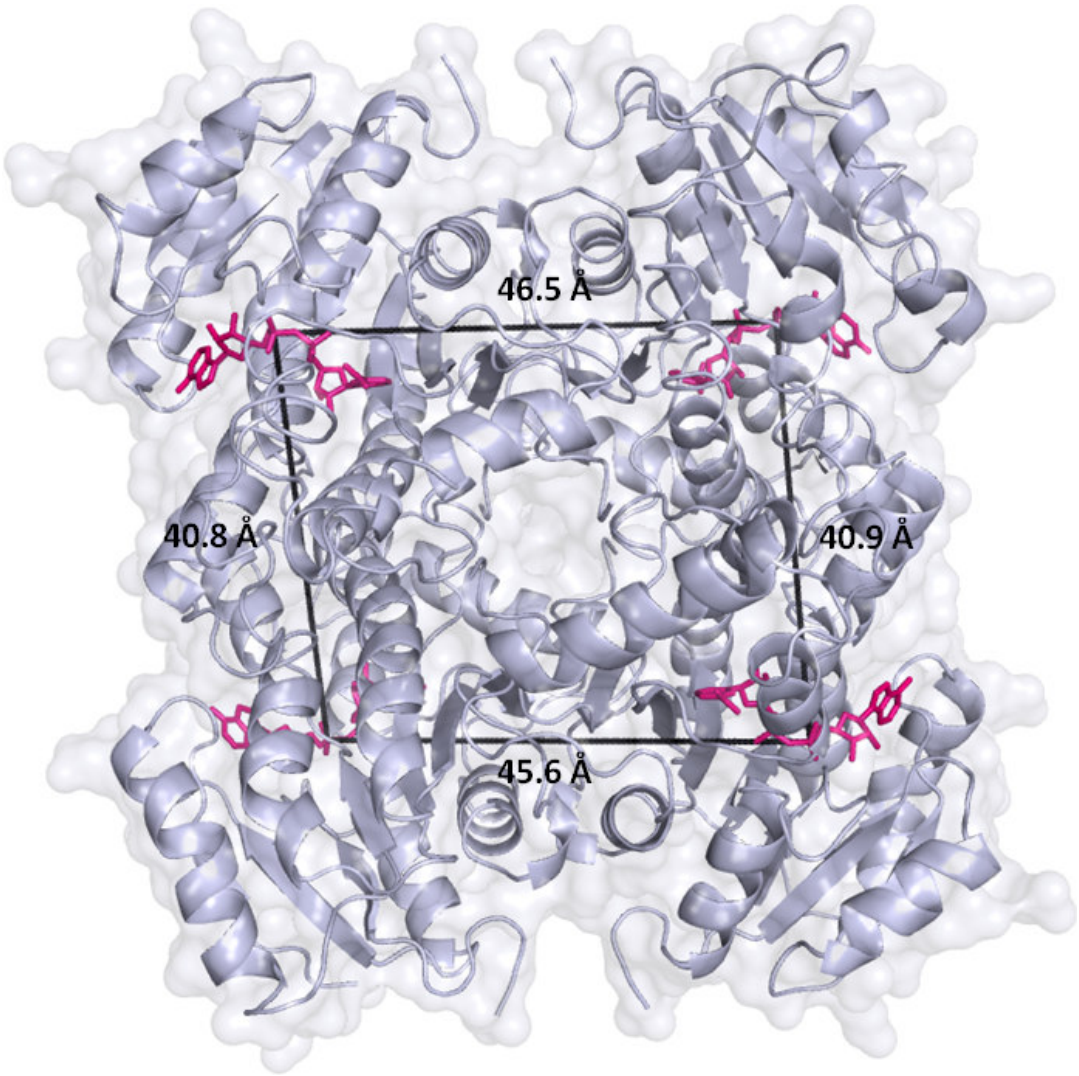


Figure 1. Distance between each monomer active site. *Ribbons* representation of the main chain, the *Mt*InhA molecular surface and the measures were generated with PyMol (DeLano, 2002). The distances between adjacent monomers were measured with reference to the NADH PA atom (magenta), present in each monomer active site. This atom was chosen because of it central position in the active site.

Modeling studies aim to identify competitive inhibitors with the substrate, avoiding molecules that bind to the NADH cavity, since NADH is also a coenzyme for many other proteins and has important roles in other metabolic pathways as well. In this way, a molecule that binds to the substrate-binding cavity would be much more selective for InhA. Therefore, in the present study, only the substrate-binding cavity was explored as the target site to identify new inhibitors.

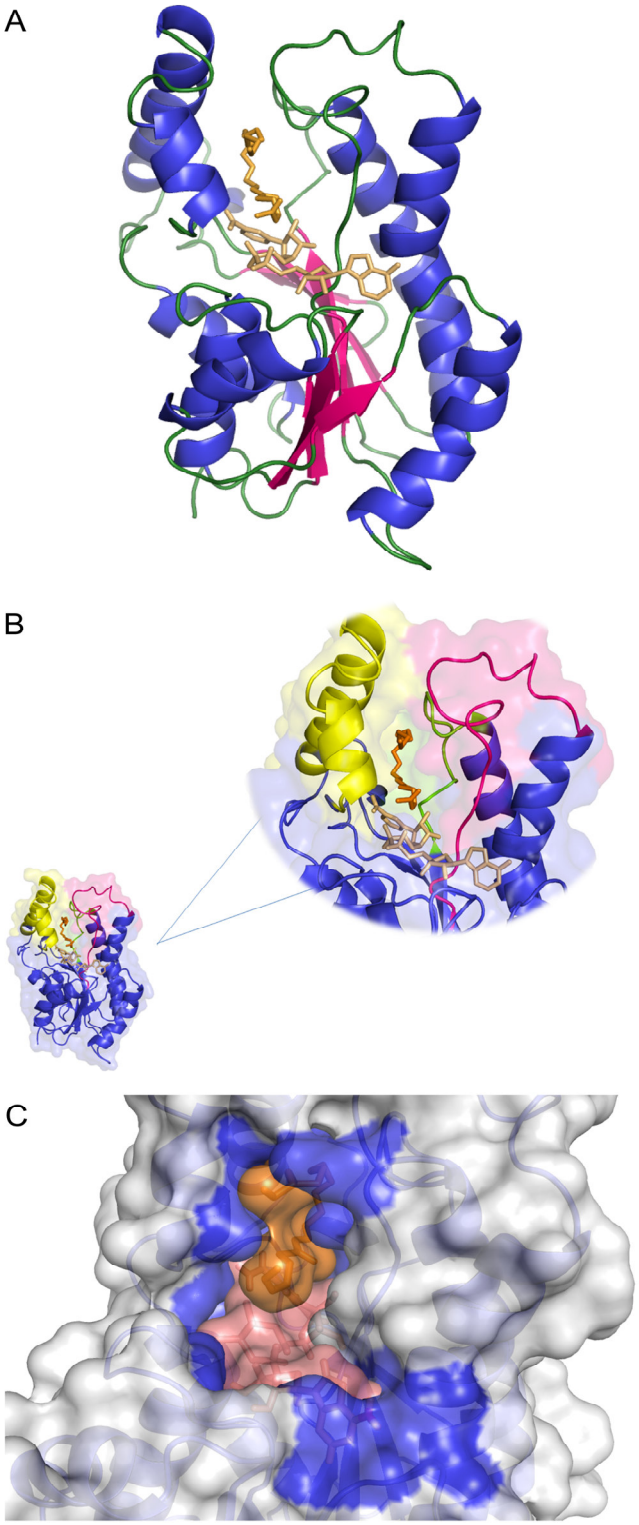


Figure 2. (**A**) Monomer structure of the ternary complex of InhA enzyme in complex with NADH and the C_{16} substrate analog. Ribbons representation of subunit C: α-helices in blue, β-strands in magenta (characterizing the *Rossmann fold* topology), and loops in green. NADH (salmon) is positioned over the *Rossmann fold* and the substrate analog (orange) is found into its binding cavity, right above NADH. (**B**) InhA active site. Highlighted in yellow, is the substrate-binding loop. Loops A and B are shown in magenta and green, respectively. (**C**) Molecular surface detail of the substrate-binding cavity. In blue are seconded the binding site residues which interact with NADH and with C_{16}. The images were generated by PyMol (DeLano, 2002), using InhA crystallographic structure with the PDB code 1BVR.

APPROACH 1

Construction of the pharmacophore model. One of the most modern concepts defines pharmacophore as “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response” (Wermuth *et* al., 1998).

To obtain success using a structure-based pharmacophore approach, it is essential to have available a great amount of information about the protein target of interest, since the molecular recognition phenomenon relies on the properties and features of the binding pocket, which are determined by the amino acids composing the binding cavity.

So, our goal was to design a robust 3D pharmacophore-based virtual screening approach in order to identify molecules able to inhibit *Mt*InhA. In this way, a detailed structural analysis was conducted on the 36 *Mt*InhA crystallographic structures to generate a receptor-based pharmacophore model (Table 1).

The superposition of the C□ atoms of the 36 monomers (in an all against all manner) revealed that the overall structure topology root mean square deviation (RMSD) among different complexes ranges from 0.1 to 1.4 Å, indicating that the overall structure topology of the different complexes is significantly different. This difference is also evident by analyzing the binding cavity volumes, ranging from 1,597.3 $Å^3$ to 3,046.7 $Å^3$ for the whole cavity (NADH + substrate cavities). We also demonstrate that the major contribution to flexibility comes from the substrate-binding cavity, with volumes ranging from 327.6 $Å^3$ to 2,109.8 $Å^3$, while the NADH binding site volume is more conserved (ranging from 1,292.8 $Å^3$ to 1525.7 $Å^3$).

No crystallographic water molecule was found to be conserved into the substrate-binding cavity, maybe due to its high hydrophobicity level (among 66% of all amino acids that established at least one contact with a ligand in the *Mt*InhA crystallographic structures, have a hydrophobic side chain).

Among the protein residues making hydrophobic contacts with the ligands in the crystallographic structures, Gly96, Phe97, Met103, Phe149, Met155, Ile215 and the three residues from the substrate-binding loop described in the literature as important for slow binding inhibition (Ala198, Met199 and Ile202) are the ones that were more recurrent. Therefore, we choose two pharmacophore points and placed them in regions favorable to hydrophobic interactions, one (point 1) nearby the

substrate-binding loop and another (point 2) near loops A and B. The spheres diameters were adjusted in order to cover the interaction surface of these residues (**Figure 3**).

In relation to the hydrogen bond pattern into the substrate-binding cavity, the role of a particular residue, Tyr158, deserves to be highlighted. Parikh and collaborators (Parikh *et* al., 1999) described Tyr158 as an electrophilic catalyst, stabilizing the transition state for hydride transfer by hydrogen bonding to the substrate carbonyl. In addition Rozwarski *et* al. (Rozwarski et al., 1999) reported a structure of a C16 fatty acid substrate analog bound to InhA that shows Tyr158 hydrogen bonded to the substrate carbonyl group and rotated from the position it occupies in the InhA-NADH binary complex. We also found in our analyses that all the crystallographic ligands in the InhA substrate-binding cavity are hydrogen bonded to Tyr158. Considering its importance, we placed a third pharmacophore point (point 3) next to Tyr158 side chain (**Figure 3**). This same point is also nearby the O2D atom of the NADH nicotinamide ring, another important recurrent hydrogen bond site to ligands into the substrate-binding cavity.

The forth pharmacophore point (point 4) was placed in another region favorable to hydrogen bond interactions, close to the phosphate groups of the NADH molecule (**Figure 3**), specifically, enclosing the favorable region for interaction with atoms the O1A, O3 and O3D of the phosphates groups.

Structure-based virtual screening. To highlight the spatial arrangement of chemical features that represent the essential interactions of small molecule ligands and *Mt*InhA, the 3D pharmacophore model was translated into a flexible search query appropriate for a 3D search with UNITY (Tripos Inc., St. Louis, MO). The tolerance spheres for the four pharmacophore points were placed by picking the appropriate atoms of the ligand Genz-10850 bound to the InhA active site (PDB ID, 1P44), and their diameters were manually adapted in order to cover the favorable areas indicated by the structural analysis.

A database containing 1,310,127 molecules was screened, and all the four pharmacophore points were required in the 3D searches. From the initial dataset (1,310,127 molecules), a total of 108,705 (10.87%) molecules were retrieved by satisfying the pharmacophore restrictions. These molecules were docked into the

*Mt*InhA substrate-binding cavity using GOLD (Jones *et* al., 1997). The coenzyme NADH was considered as part of the protein.

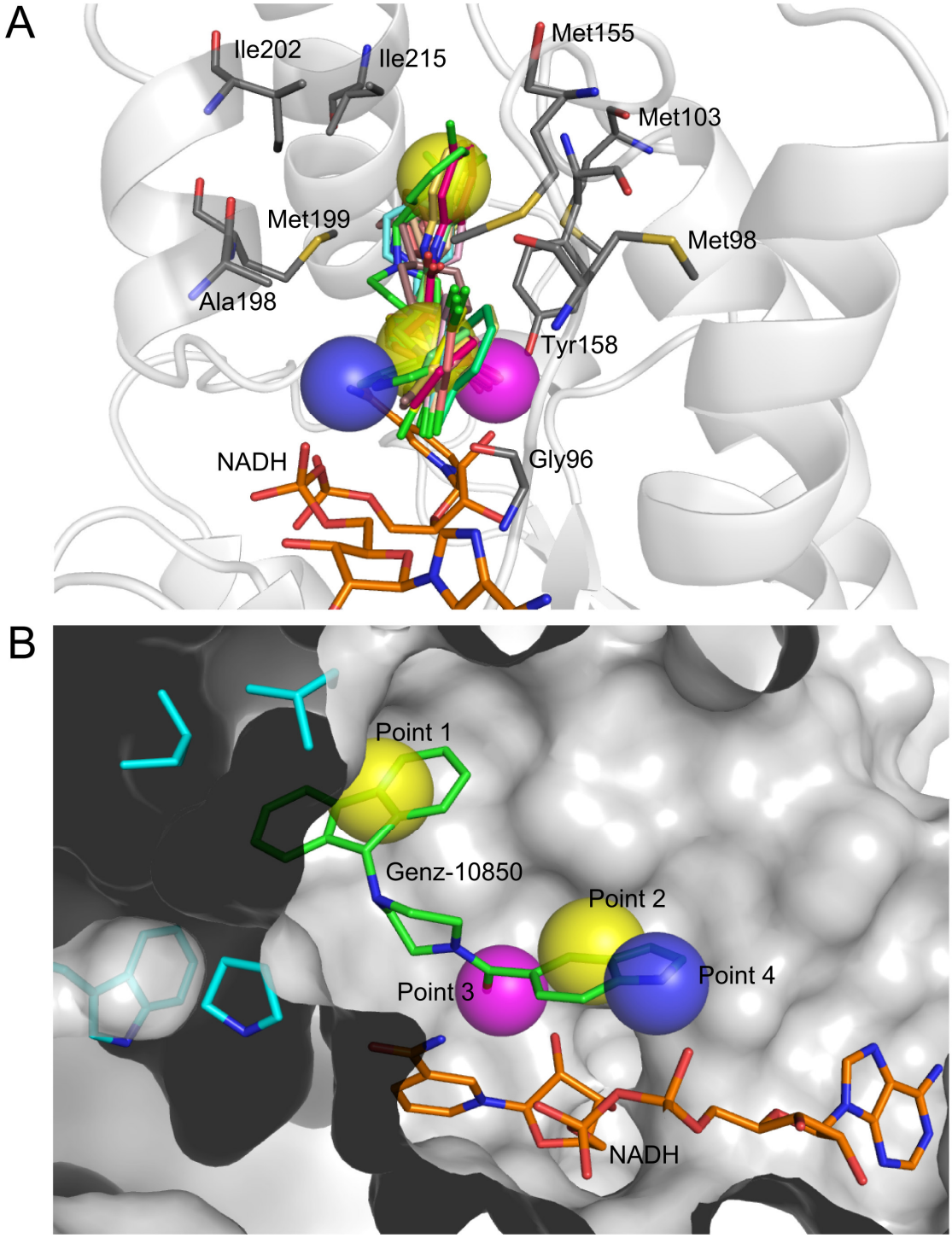


Figure 3. (**A**) Superposition of all ligands present into the InhA crystallographic structures substrate-binding cavity. Highlighted are the important amino acids to interactions with different ligands. (**B**) 3D pharmacophore model obtained from the receptor structures and the known ligands features. Yellow = regions favorable to hydrophobic groups; Magenta = favorable region to hydrogen bond donor or acceptor groups; Blue = region propitious to hydrogen bond donor groups. The images were generated by PyMol (DeLano, 2002).

The predicted binding modes into *Mt*InhA substrate-binding site were ranked based on their relative affinity using the GoldScore scoring function implemented in GOLD. The top 300 ranked molecules were selected for visual inspections to select the final set of compounds. The criteria used in the visual inspection included: (i) matching of the 4 points of the pharmacophore model after docking; (ii) overall matching of the hydrogen-bonding network, with emphasis on the formation of hydrogen bonds to Tyr158; (iii) overall matching of hydrophobic contacts to the regions assigned by the pharmacophore model; and (iv) complementarity between ligands and protein surfaces in terms of spatial occupancy. After this analysis the final subset of compounds consisted of 34 molecules.

APPROACH 2

Docking Programs. Molecular docking programs are commonly used to orient small molecules within a three-dimensional representation of the protein structure. There is a large number of docking programs currently available, with a wide diversity of algorithms to generate ligand conformations and scoring functions to rank the pose solutions obtained. Therewith our aim was to access whether the different algorithms implemented in each of the well established and widely used programs GOLD (Jones *et* al., 1997), AutoDock (Morris *et al*., 1998; Morris *et al*., 1996; Goodsell *et al*., 1996; Goodsell & Olson, 1990), FlexX (Rarey et al., 1996), and Surflex-Dock (Jain, 2003), could be able to generate the similar equivalent poses for a particular ligand, indicating that this may be the favorable mode of interaction between the protein and the ligand.

Docking simulations. After docking the whole dataset, composed by 1,310,127 using GOLD, the 100 top ranked molecules were resubmitted to molecular docking simulations to AutoDock, FlexX and Surflex, using the same parameters as used for the docking with GOLD. In this approach, the main difference relied on the implemented algorithms for searching and scoring the best ligand poses.

Molecule conformation analysis. The best pose obtained by each program for each of the 100 GOLD top-ranked molecules, was compared in terms of conformational geometry and positioning into the substrate-binding cavity. We

visually analyzed the results obtained by the four docking programs and selected the molecules for which, at least, three of the four programs assigned the same conformation when bound to *Mt*InhA substrate-binding cavity. These analyses provided us with a final subset of 21 molecules.

Toxicity evaluation. The final set of 55 compounds (34 from approach 1, and 21 from approach 2), where submitted to OSIRIS toxicity prediction server (Sander *et*. al, 2009). For the final set of hit candidates we selected just the molecules with no toxicity risks (mutagenicity, tumorigenicity, irritability, reproductive effectiveness) and with the most favorable cLogP, solubility, molecular weight, druglikeness and drug-score features. According to these criteria, 19 molecules were selected.

Similarity Ensemble Approach (SEA). Chemically similar drugs often bind biologically diverse protein targets, and proteins with similar sequences or structures do not always recognize the same ligands. SEA considers proteins from a chemocentric point of view, relating them through the chemical similarity of their ligands. The idea is that similar molecules have similar biological profiles and bind similar targets (Keiser et al., 2007). In general, the authors classified similarity according to the E-value. On one hand, for E-value < 1×10^{-10} the similarity is significant, on the other hand for E-value > 1.0 it is insignificant. The most interesting and significant results to our set of molecules were obtained for the molecules ZINC12509636 and ZINC14986320, referenced as antimalarial with E-values of 9.38×10^{-29} and 4.61×10^{-19}, respectively.

Lack of correlation between results from both approaches. Comparing the set of molecules resultant from both approaches, we could not identify a superposition of compounds. In other words, these approaches were unable to reach a common end result. One explanation could be the large size of the initial dataset (one million molecules) and the selection of just the 100 best scored molecules by GOLD in approach 2, where all the other molecules were eliminated from the subsequent steps. Still it is interesting due to promoting an enhancement to the compounds chemical diversity.

Selection of the molecules to *in vitro* evaluation. The final selection of the molecules to be tested *in vitro* for its inhibitory activity against *Mt*InhA was based on all the analyses described above. From the 19 remaining molecules 6 were chosen to be tested *in vitro* aiming to cover different modes of interaction with *Mt*Inha substrate-binding cavity. From the first pharmacophore based approach, we chose ZINC22559057 which interacts with the binding cavity mainly through hydrophobic contacts. ZINC12759934 interacts via a hydrogen bond to NADH and several hydrophobic interactions with the protein, including Ala198, important to slow tight binding. ZINC09137707 is not hydrogen bonded do NADH, however, establishes two hydrogen bonds with Tyr158 and several hydrophobic interactions (**Figure 4** and **Table 2**). From the second approach, we selected ZINC12242826 which binds to InhA through a hydrogen bond with NADH, one with Tyr158, one with Glu219 and several hydrophobic interactions. ZINC12509636 hydrogen bonds to Tyr158 and Glu219 and makes many hydrophobic interactions. Finally, ZINC02931014 makes two hydrogen bonds to NADH, two with Tyr158 and hydrophobic interactions also with Ala198 (**Figure 5** and **Table 3**).

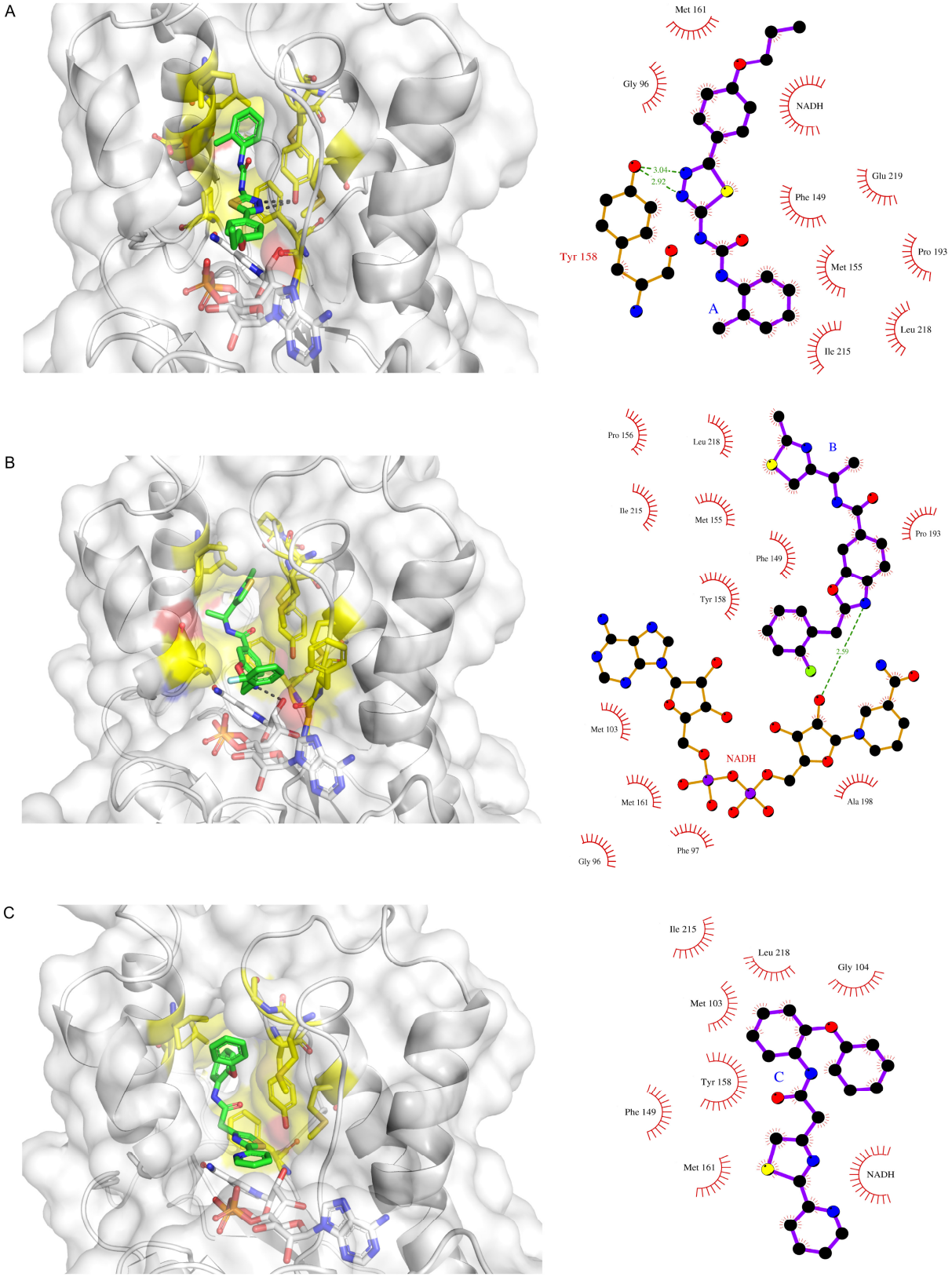


Figure 4. This figure shows the mode of interaction of the molecules selected using the first approach described in this work, where a pharmacophore model was used to search new potential inhibitor molecules. (**A**) ZINC22559057; (**B**) ZINC12759934; (**C**) ZINC09137707. The images were generated by PyMol (DeLano, 2002). The molecules are represented in green and the protein residues making hydrophobic contacts with them are shown in yellow.

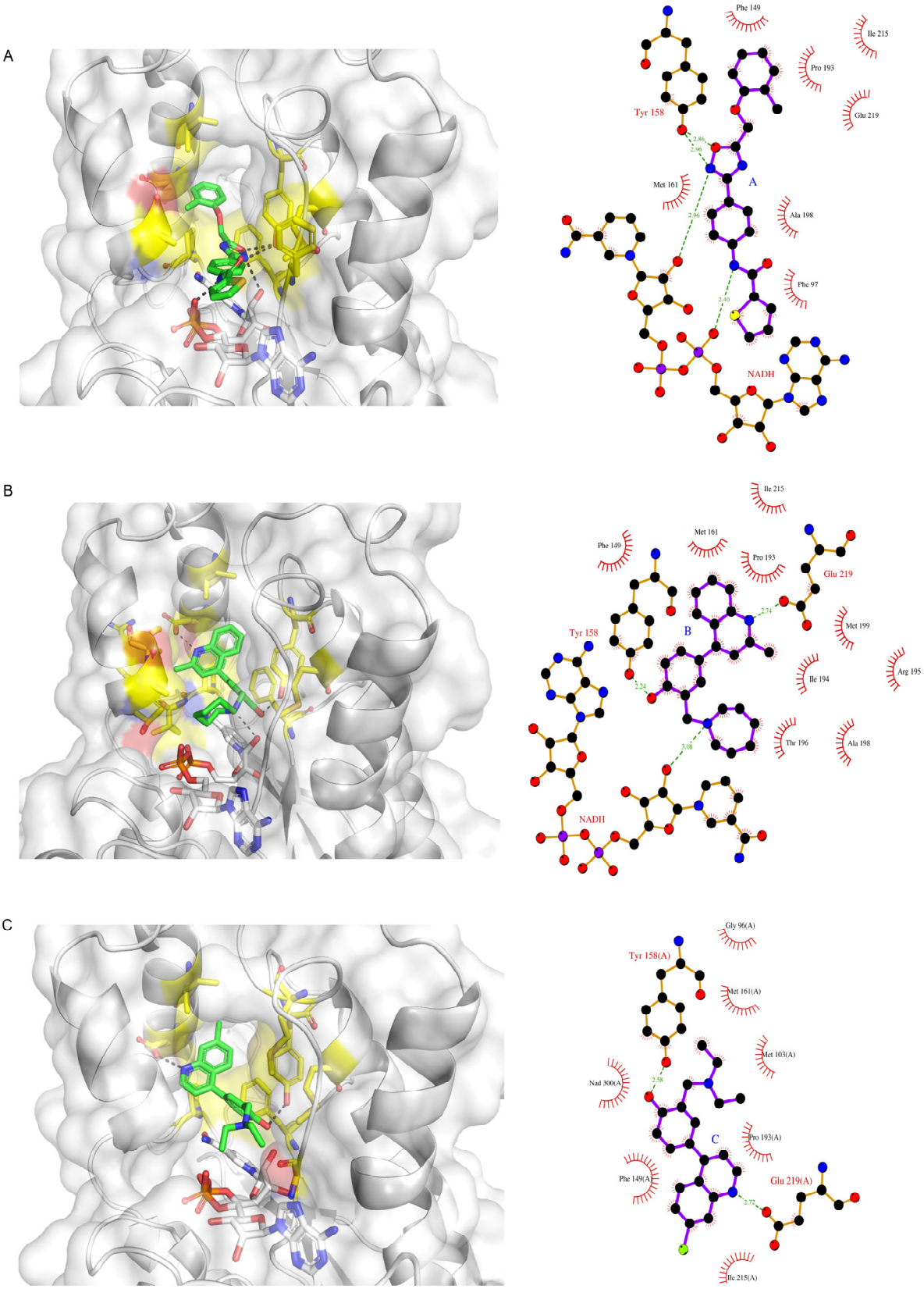


Figure 5. This figure shows the mode of interaction of the molecules selected using the second approach where four docking programs were used to identify the most promising molecules; (**A**) ZINC12242826; (**B**) ZINC12509636; (**C**) ZINC02931014. The images were generated by PyMol (DeLano, 2002). The molecules are represented in green and the protein residues making hydrophobic contacts with them are shown in yellow.

Table 2. Information of the molecules selected from Approach 1.

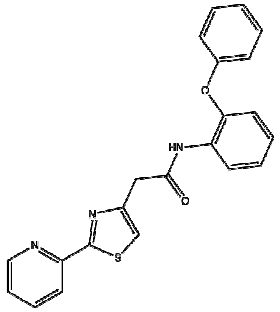
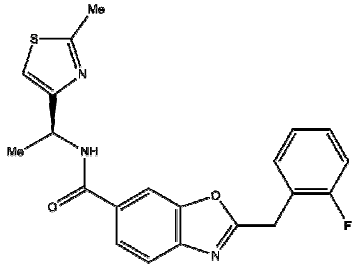
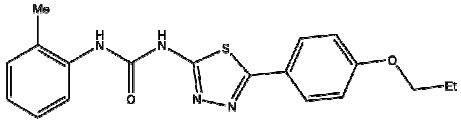
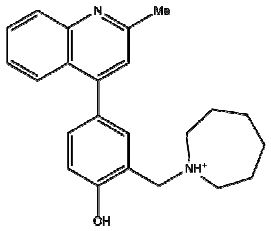
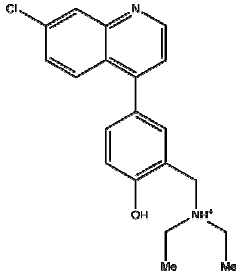
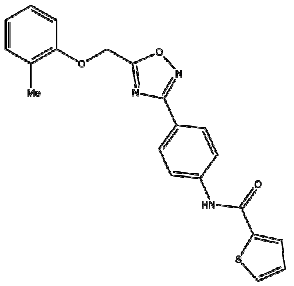
Molecule ZINC code	Structure	IUPAC Name	xLogP	HBD	HBA	Charge	Mwt	NRB
APPROACH 1								
ZINC22559057		N-(2-phenoxyphenyl)-2-[2-(2-pyridyl)thiazol-4-yl]acetamide	4.57	1	5	0	387.464	6
ZINC12759934		2-[(2-fluorophenyl)methyl]-N-[(1S)-1-(2-methylthiazol-4-yl)ethyl]-1,3-benzoxazole-6-carboxamide	4.06	1	5	0	395.459	5
ZINC09137707		3-(o-tolyl)-1-[5-(4-propoxyphenyl)-1,3,4-thiadiazol-2-yl]urea	4.74	1	6	-1	367.454	7

Table 3. Information of the molecules selected from Approach 2.

Molecule ZINC code	Structure	IUPAC Name	xLogP	HBD	HBA	Charge	Mwt	NRB
APPROACH 2								
ZINC12242826		2-(azepan-1-ylmethyl)-4-(2-methyl-4-quinolyl)phenol	5.54	2	3	1	347.482	3
ZINC12509636		4-(7-chloro-4-quinolyl)-2-(diethylaminomethyl)phenol	5.1	3	3	2	342.87	5
ZINC02931014		N-(4-{5-[(2-methylphenoxy)methyl]-1,2,4-oxadiazol-3-yl}phenyl)thiophene-2-carboxamide	4.89	1	6	0	391.452	6

OBS.:

Biochemical evaluation. Para a conclusão desta estapa, estamos aguardando a chegada das moléculas para realização dos testes de inibição da enzima InhA *in vitro*. Após a realização destes testes, será finalizada a redação deste manuscrito que será submetido para a revista "*Journal of Medicinal Chemistry*".

REFERENCES

Agüero, F.; Al-Lazikani, B.; Aslett, M.; Berriman, M.; Buckner, F.S.; Campbell, R. K.; Carmona, S.; Carruthers, I.M.; Chan, A.W.E.; Chen, F.; Crowther, G.J.; Doyle, M.A.; Hertz-Fowler, C.; Hopkins, A.L.; McAllister, G.; Nwaka, S.; Overington, J.P.; Pain, A.; Paolini, G.V.; Pieper, U.; Ralph, S.A.; Riechers, A.; Roos, D.S.; Sali, A.; Shanmugam, D.; Suzuki, T.; Van Voorhis, W.C.; Verlinde C.L.M.J. (2008) *Nature Reviews in Drug Discovery*, 900-907.

Bermann, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne P.E. (2000) *Nucleic Acids Research* 28, 235-242.

Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.B.; Meyer Jr., E.F.; Brice, M.D.; Rodgers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. (1977) *Journal of Molecular Biology* 112, 535-542.

Campbell, J.W.; Cronan, J.E.Jr. (2001) *Annual Review of Microbioogy.,* 55, 305-332.

Clark, M.; Cramer, R. D.; Van Opdenbosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982–1012.

DeLano, W. L. (2002) *DeLano Scientific*, San Carlos, CA, USA, [http://www.pymol.org].

Dessen, A.; Quémard, A.; Blanchard, J.S.; Jacobs, Jr. W.R.; Sacchettini, J.C. (1995) *Science*, 267, 1638-1641.

Di Perri, G.; Aguilar Marucco, D.; Mondo, A.; Gonzalez de Requena, D.; Audagnotto, S.; Gobbi, F.; Bonora, S. *Expert Opinion in Drug Safety*, 4, 821-836.

Dundas,J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. (2006) *Nucleic Acid Research*, 34, W116-W118.

Gasteiger, J.; Marsili, M. *Tetrahedron* (1980), *36*, 3219-3228.

Goodsell, D.S; Olson, A.J. (1990) *Protein Structure and Functional Genomics*, 8, 195-202.

Goodsell, D.S.; Morris, G.M.; Olson, A.J. (1996) *Journal of Molecular Recognition*, 9, 1-5.

Helma, C. (2004) In silico Predictive Toxicology: The State of the Art and Strategies to Predict Human Health Effects. Inst. f. Computer Science Univ. Freiburg October 19.

Irwin, J. J.; Shoichet, B. K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.***2005**, *45*, 177–182.

Jain, A. N. Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 427–440.

Jain, A. N. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46*, 499–511.

Jain, A.; Mondal, R. (2008) *FEMS Immunology and Medicinal Microbiology,* 53, 145-50.

Jones, G.; Willett, P.; Glen, R.C.; Leach, A.R.; Taylor, R. (1997) *Journal of Molecular Biology*, 267, 727-748.

Keiser, M.J.; Roth B.L.; Armbruster, B.N.; Ernsberger, P.; Irwin, J.J.; Shoichet, B.K. (2007) *Nature Biotechnology*, 25(2), 197-206.

Klebe, G. Virtual screening: Scope and limitation. In *Virtual Screening in Drug Disco*V*ery*, 1st ed.; Alvarez, J., Shoichet, B., Eds.; CRC Press: Boca Raton, FL, 2005, Vol. 1, pp 3-24.

Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K. (2011) *Nature*, 469, 483-490.

Kumar, A.; Siddiqi, M.I. (2008) *Journal of Molecular Modeling*, 14, 923-935.

Morris, G.M.; Goodsell, D.S.; Huey, R.; Olson, A.J. (1996) *Journal of Computer-Aided Molecular Design*, 10, 293-304.

Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. (1998) *Journal of Computational Chemistry*, 19, 1639-1662.

Morrison, J.; Pai, M.; Hopewell, P.C. (2008) *The Lancet Infectious Diseases,* 8, 359-68.

Parikh, S.; Moyniham D.P.; Xiao, G.; Tonge, P. (1999) *Biochemistry*, 38, 13623-13634.

Powell, M.J.D.; (1977) Restart Procedures for the Conjugate Gradient Method. *Math. Program.* 12, 241-254.

Quémard, A.; Sacchettini, J.C.; Dessen, A.; Vilchèze, C.; Bittman, R.; Jacobs, W.R. Jr.; Blanchard, J.S. (1995) *Biochemistry*, 34, 8235-8241.

Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. (1996) *Journal of Molecular* Biology, 261, 470-489.

Rozwarski, D.A.; Vilchèse, C.; Sugantino, M.; Sacchettini, J.C., (1999) *Journal of Biological Chemistry*, 274, 15582-15589.

Thomas, S.; Freyss, J.; von Korff M.; Reich J.R.; Rufen, C. (2009) *Journal of Chemical Information and Modeling*, 49, 232-246

Sippl, M.J.; Wiederstein, M. (2008) *Bioinformatics*, 24(3), 426-427.

Sippl, M.J. (2008) *Bioinformatics*, 24(6), 872-873.

Schroeder, E.K.; de Souza, O.N.; Santos, D.S.; Blanchard, J.S.; Basso, L.A. (2002) *Current Pharmaceutical Biotechnology*, 3, 197-225.

UNITY Chemical Information Software, version 4.1; Tripos Inc.: St. Louis, MO, 2006.

Vilchèze, C.; Morbidoni, H.R.; Weisbrod, T.R.; Iwamoto, H.; Kuo, M.; Sacchettini, J.S.; Jacobs Jr., W.R. (2000) *Journal of Bacteriology*, 182, 4059-4067.

Velayati, A.A.; Farnia, P.; Masjedi, M.R.; Ibrahim, T.A.; Tabarsi, P.; Haroun, R.Z.; Kuan, H.O.; Ghanavi, P.; Farnia, P.; Varahram, M. (2009) *European Respiratory Journal*, 34, 1202-1203

Vonk J.A.; Benigni, R.; Hewitt M.; Nendza, M.; Segner, H.; van de Meent, D.; Cronin M.T. (2009) *Alternatives to Laboratory Animals*., 37(5), 557-71.

Wallace, A.C.; Laskowski, R.A.; Thornton J.M. (1995) *Protein Engineering*, 8, 127-134.

Wermuth, C. G., Ganellin, C. R., Lindberg, P. & Mitscher L. A. (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). Pure Appl. Chem., *70*, 1129–1143.

White, S.W.; Zheng, J.; Zhang, Y.M.; Rock, C.O. (2005) *Annual Reviews of Biochemistry*, 74, 791-831.

Zhang, Y.M.; Lu, Y.J.; Rock, C.O. (2004) *Lipids*, 39, 1055-1060.

5.2 Artigo em fase de finalização

"The state of the art of the knowledge about the enzyme InhA from *Mycobacterium tuberculosis*"

Ivani Pauli[1,2], Luis Fernando Saraiva Macedo Timmers [1], and Osmar Norberto de Souza [1,2]*

[1]*Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas - LABIO, Faculdade de Informática, PUCRS;* [2]*Centro de Pesquisas em Biologia Molecular e Funcional - CPBMF, Instituto Nacional de Ciência e Tecnologia em Tuberculose – INCT-TB, PUCRS.*

INTRODUCTION

Human tuberculosis (TB) is a contagious-infectious disease mainly caused by *Mycobacterium tuberculosis*, an intracellular pathogenic bacterium that establishes its infection mainly in the lungs. TB is the most prevalent infectious disease worldwide and remains the second most common cause of death from infectious illness in the world (following HIV/AIDS) killing nearly two million people each year [1]. The World Health Organization (WHO) has estimated that one third of the world's population, nearly 2 billion people, mostly in developing countries, has been infected with *Mtb*. Among the infected individuals, 8 million develop active TB, and nearly 2 million people die from the disease annually [2].

TB resurgence has been attributed to several factors, such as the increase in drug resistance; the HIV/AIDS pandemic (currently, TB is the most common cause of death in patients with HIV); the increase of injectable drug users; changes in social structure; the increase of immigrants from high prevalence nations to developed ones; the aging of the world's population; the active transmission amongst

environments of human accumulation (prisons, hospitals, homeless shelters); and the degradation of health care systems [3,4].

Drug-susceptible strains of *Mtb* can be treated with a cocktail of inhibitors, including isoniazid (INH), rifampicin (RIF), ethambutol (ETH), and pyrazinamide (PZA). However, the escalating numbers of patients infected with multi-drug-resistant (MDR) mycobacterial strains poses a substantial public health risk [5].

According to Koul and collaborators, in a recent review about the challenge of new drug discovery for TB [6], the disease is more prevalent in the world today than at any other time in human history. They also recognize that, to achieve global control of this pandemic, there is a need for new TB drugs, which can: (1) shorten treatment duration; (2) target drug resistant strains; (3) simplify treatment by reducing the daily pill burden; (4) lower dosing frequency; and (5) be co-administered with HIV medications.

Among the most promising targets to design novel antibacterial agents are the Fatty Acid Synthase (FAS) pathway enzymes [7, 8, 9]. In most bacteria and plants, fatty acid biosynthesis is catalyzed by a set of distinct, monofunctional enzymes collectively known as the type II Fatty Acid Synthase (FAS-II) [10, 11, 12]. These enzymes differ significantly from the type I FAS (FAS-I) in mammalians, birds and yeast, in which all of the enzymatic activities are encoded in one or two multifunctional polypeptides [13, 14, 15].

Some bacteria, such as mycobacteria, possess both, a FAS-I and a FAS-II systems [16] (**Figure 1**).

The mycobacteria FAS-I system displays a bimodal distribution of products centered on C16 and C24-C26 [16]. FAS-II system prefers C16 as a starting substrate and can extend up to C56 [17], indicating that the mycobacterial FAS-II system utilizes the products of the FAS-I system as primers to extend fatty acyl chain lengths even further. The longer chain products of the FAS-II system are the precursors of mycolic acids, long chain alpha-alkyl-beta-hydroxy fatty acids, which are the major components of mycobacteria cell walls [18, 19] (**Figure 2**).

This distinctive difference in the FAS molecular organization between most bacteria and mammals makes possible to design inhibitors of increased selectivity and lower toxicity [20].

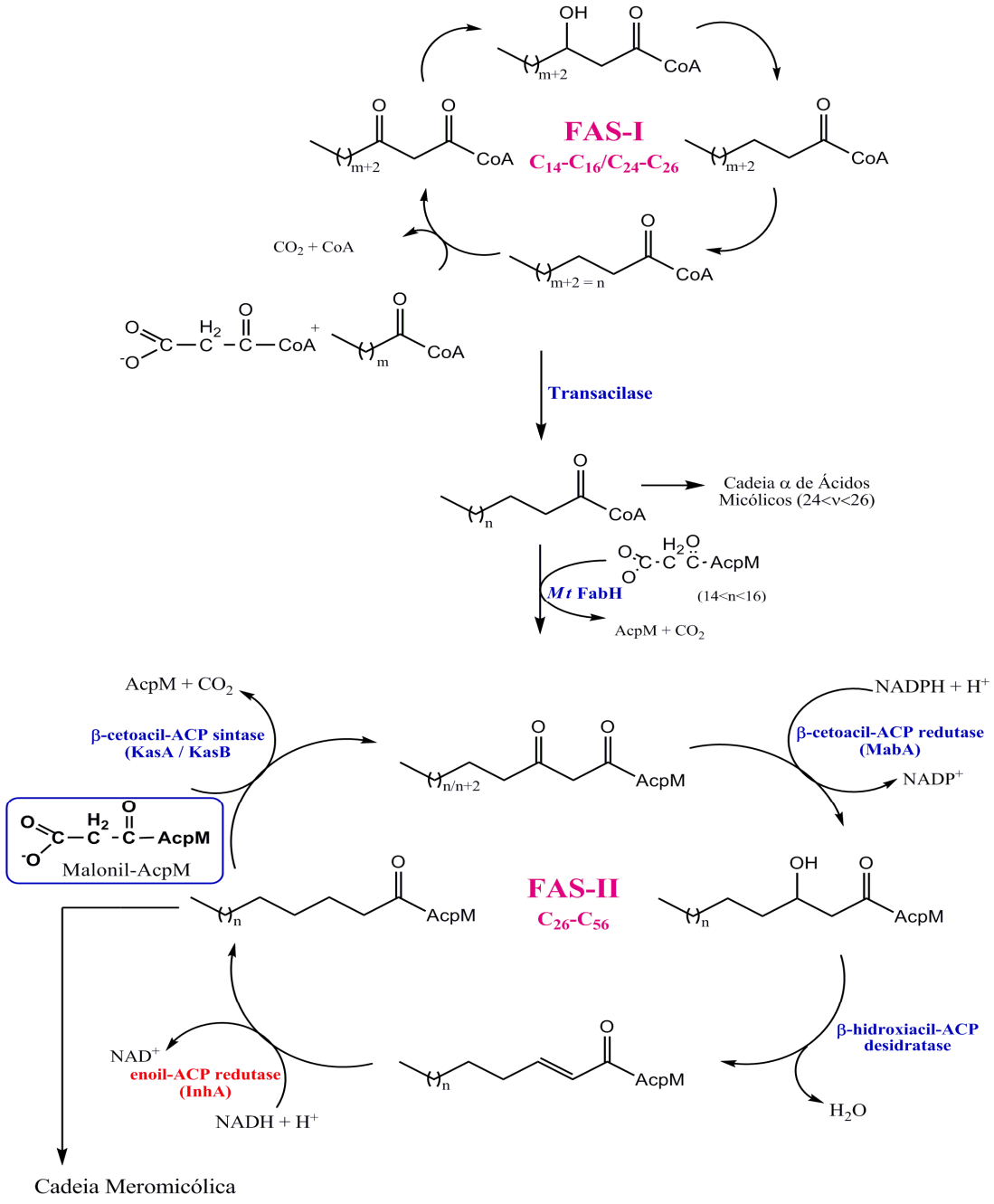


Figure 1. Fatty acid biosynthesis in Mtb.

InhA, or 2-*trans*-enoyl-ACP (CoA) reductase (E.C.1.3.1.9), is one of the key enzymes involved in the elongation cycle of fatty acids in *Mtb*. It is the fourth and last enzyme of the type II fatty acid synthase system (FAS II) and reduces preferentially long chain enoyl thioester substrates (those containing 16 or more carbon atoms) yielding the long carbon chain of the meromycolate branch from mycolic acids (C40-60), α-branched fatty acids, the hallmark of mycobacteria [21].

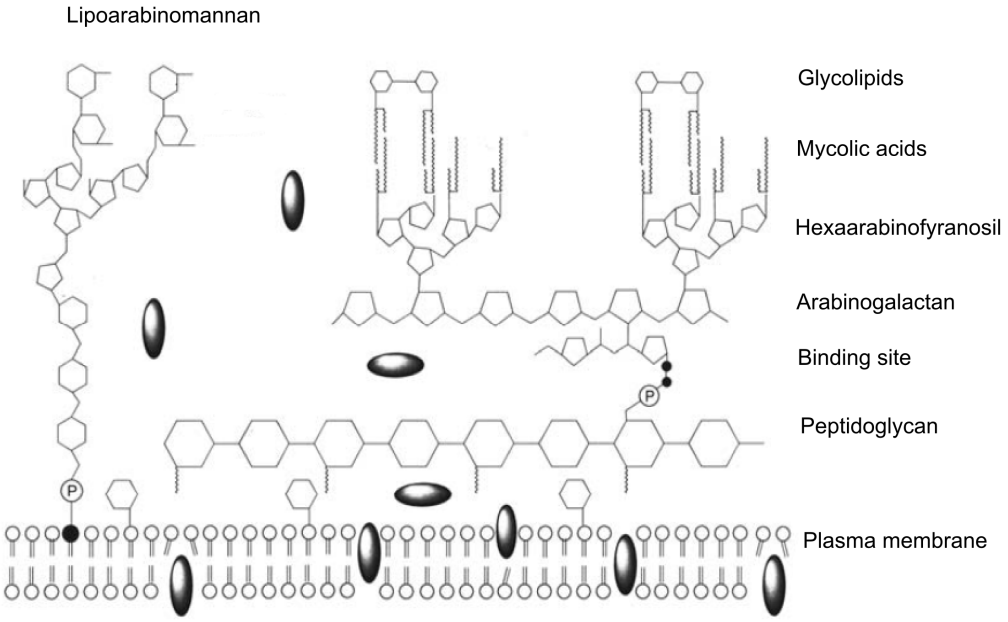


Figure 2. Mtb cell wall schematic representation (Schroeder *et al.*, 2002).

InhA catalyzes the nicotinamide adenine dinucleotide (NADH) dependent reduction of long chain 2-*trans*-enoyl-ACP fatty acids in its saturated correspondent, resulting in the stereo-specific reduction of the α,β-unsaturated thioester double bond, whose non-lipid portion could be ACP or CoA. Steady-state kinetic studies showed that the two substrates bind to InhA via a sequential, but not rigid, kinetic mechanism, starting preferentially with the addition of NADH followed by the binding of the enoyl substrate. The chemical mechanism involves stereospecific hydride transfer of the NADH 4S hydrogen to the substrate C3 position, followed by protonation at the C2 atom of an enzyme-stabilized enolate intermediate [22] (**Figure 3**).

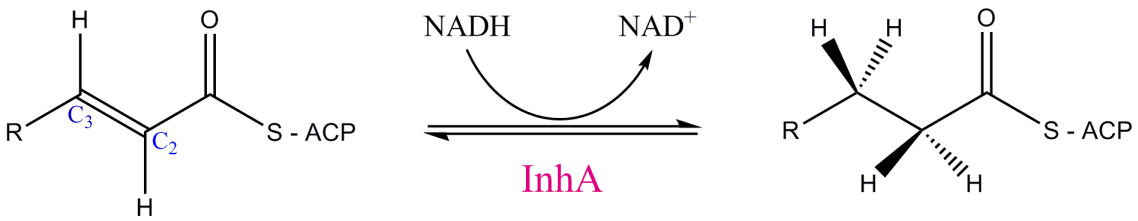


Figure 3. Reaction catalyzed by Mtb InhA.

In this review we describe the enzyme InhA from *Mtb* aiming to gather the available information that will be useful to understand its sequence-structure-dynamics-function relationships. We also performed a series of analyses using structural information, in order to identify peculiar characteristics, which could be essential or at least important for the rational drug design initiatives having this enzyme as a target.

InhA tertiary and quaternary structures

InhA is encoded by the *inh*A gene and is composed by 268 amino acids with a molecular weight of ~29kD. It belongs to the Short Chain Dehydrogenase/Reductase (SDR) family, that uses a NAD(H) or NADP(H) molecule as a coenzyme [23]. This family of proteins is characterized by having a topology in which each subunit consists of a single domain with a *Rossmann fold* core, where the coenzyme-binding site is found [24]. Overall, the InhA structure seems like a chair and is composed by seven β strands and eight α helices. The coenzyme binds into a cavity between the "back" and the "seat" of this structure while the substrate-binding site is localized into a cavity at the "back" [25]. Several α helices and β strands of the *Rossmann fold* extend over the NADH binding site, creating a crevasse where the substrate binds to [26] (**Figure 4**).

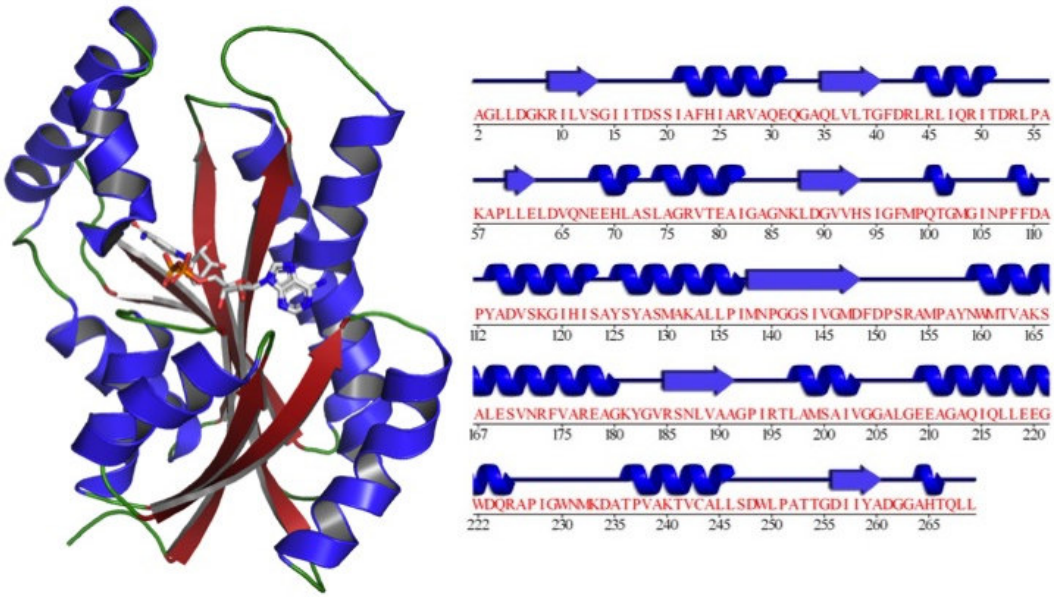


Figure 4. Primary and secondary structures of the InhA (PDB code: 1BVR).

Size exclusion chromatography analysis demonstrates that InhA is a homo-tetramer (**Figure 5**) in solution and this is the biologically active structure [22].

There is a lack of information about InhA quaternary structure, and most works just mention that InhA is biologically active as a tetramer, but they all use the monomer structure to perform simulations.

This may happen due to the assumption that each monomer active site works independently from the others, once they are located in opposite sides, far away from the monomer interface regions, being the distance between two cavity centers around 40 Å. This is reasonable when the receptor is kept rigid, but for simulations where the aim is to incorporate some level of protein flexibility the use of the tetramer may be more advisable due to the fact that two loops (A and B) are located in the interface region between adjacent subunits, which could decrease its flexibility.

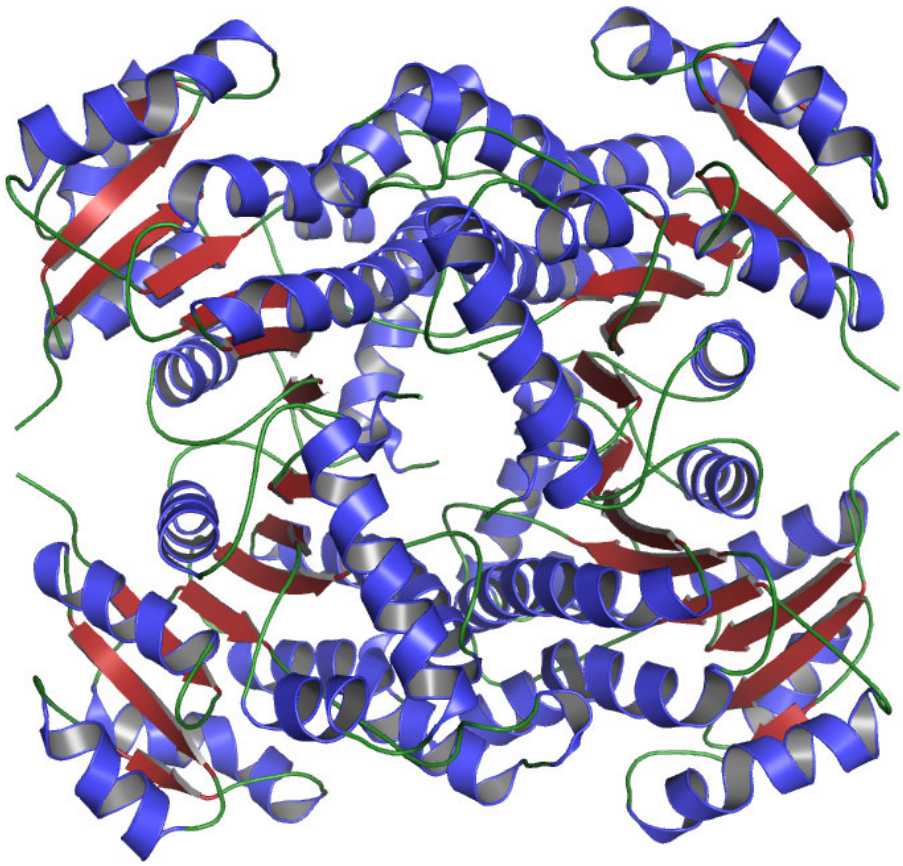


Figure 5. Quaternary structure of the Mtb InhA.

One of the few studies published relates conformational changes in the InhA quaternary structure with its inhibition [27]. Based on the knowledge that InhA

interacts *in vivo* with other components of the FAS II pathway, Kruh and collaborators investigated the structural changes that could affect protein-protein interactions involving InhA, and how these ligand-induced conformational changes are modulated in the InhA mutants. A significant result shows that NADH binding to wild-type InhA is hyperbolic, while the mutations bind the cofactor with positive cooperativity, suggesting that they permit access to a second conformational state of the protein (also observed by Oliveira *et* al., 2006 [28]. They also demonstrate by cross-linking studies that InhA inhibition causes dissociation of the tetramer into dimmers, and by analytical ultracentrifugation and size exclusion chromatography, that ligand binding causes a conformational change in the protein that prevents cross-linking across one of the dimmer-dimmer interfaces in the InhA tetramer. Interestingly, a similar ligand-induced conformational change is also observed for the InhA mutants, indicating that the mutations modulate communication between the subunits without affecting the two conformational states of the protein that are present.

Available structural data

A better and deeper understanding of the InhA mechanism of action and inhibition became possible through the knowledge of its structural features by the resolution of many crystallographic structures. Actually there are thirty-six *Mtb* InhA structures deposited in the Protein Data Bank (PDB) [29, 30]. Among them we can found the wild type enzyme and its mutants related to drug resistance in both, apo form and in complex with a substrate analog (C16), with the coenzyme NADH and with a variety of ligands and inhibitors (**Table 1**), making possible a better characterization of the binding pocket and the pivotal interactions between the protein-ligand complexes.

The active site structure

Each InhA monomer is characterized by having two binding cavities, one to bind the cofactor NADH, and another to bind the fatty acid substrate (**Figure 6**).

Table 1. Relation of the 36 *Mt*InhA structures deposited at PDB.

PDB Code (Mutant)	Resolution (Å)	Ligand	Reference
1ENY	2.2	NADH	Dessen *et al.*, 1995
1ENZ (S94A)	2.7	NADH	Dessen *et al.*, 1995
1ZID	2.7	INH-NADH Adduct	Rozwarski *et al.*, 1998
1BVR	2.8	NADH + Substrate Analog (C_{16})	Rozwarski *et al.*, 1999
1P44	2.7	NADH + GEQ	Kuo *et al.*, 2003
1P45	2.6	NADH + TCL	Kuo *et al.*, 2003
2B35	2.3	NADH + TCL	Sullivan *et al.*, 2006
2B36	2.8	NADH + 5PP	Sullivan *et al.*, 2006
2B37	2.6	NADH + 8PS	Sullivan *et al.*, 2006
2AQH (I21V)	2.0	NADH	Oliveira *et al.*, 2006
2AQI (I47T)	2.2	NADH	Oliveira *et al.*, 2006
2AQ8	1.9	NADH	Oliveira *et al.*, 2006
2AQK (S94A)	2.3	NADH	Oliveira *et al.*, 2006
2NV6 (S94A)	1.9	INH-NADH Adduct	Vilchèze *et al.*, 2006
2H7I	1.6	NADH + 566	He *et al.*, 2006
2H7L	1.7	NADH + 665	He *et al.*, 2006
2H7M	1.6	NADH + 641	He *et al.*, 2006
2H7N	1.9	NADH + 744	He *et al.*, 2006
2H7P	1.8	NADH + 468	He *et al.*, 2006
2NTJ	2.6	PTH-NADH Adduct	He *et al.*, 2006
2H9I	2.2	ETH-NADH Adduct	He *et al.*, 2006
2IDZ	2.0	INH-NADH Adduct	Dias *et al.*, 2007
2IE0 (I21V)	2.0	INH-NADH Adduct	Dias *et al.*, 2007
2IEB (S94A)	2.2	INH-NADH Adduct	Dias *et al.*, 2007
2IED (S94A)	2.1	APO	Dias *et al.*, 2007
2PR2	2.5	INH-NADH Adduct	Argyrou *et al.*, 2007
2NSD	1.9	NADH + 4PI	He *et al.*, 2007
3FNG	1.9	NADH + JPL	Freundlich *et al.*, 2009
3FNH	2.8	NADH + JPJ	Freundlich *et al.*, 2009
3FNE	1.9	NADH +8PC	Freundlich *et al.*, 2009
3FNF	2.3	NADH + JPM	Freundlich *et al.*, 2009
2X22	2.1	NADH + TCU	Luckner *et al.*, 2010
2X23	2.8	NADH + TCU	Luckner *et al.*, 2010
3OEW	2.1	NAD	Molle *et* al., 2010
3OEY (T266E)	2.0	NAD	Molle *et* al., 2010
3OF2 (T266D)	1.75	NAD	Molle *et* al., 2010

GEQ: 5-{[4-(9H-fluoren-9-yl)piperazine-1-yl]carbonyl}-1H-indol; TCL:Triclosan; 5PP: 5-pentyl-2-phenoxyphenol; 8PS: 5-octyl-2-phenoxyphenol; INH: Isoniazid; 566: (3S)-1-cicloexyl-5-oxo-N-phenylpirrolidine- 3-carboxamide; 665: (3S)-N-(3-bromophenyl)-1-ciclohexyl-5-oxopyrrolidine-3-carboxamide; 641: (3S)-1-ciclohexyl-N-(3,5-diclorofenil)-5-oxopyrrolidine-3-carboxamide; 744: (3S)-N-(5-Cloro-2-metilphenyl)-1-ciclohexyl-5-oxopyrrolidine-3-carboxamide; 468: (3S)-N-(3-cloro-2-metilphenyl)-1-ciclohexyl-5-oxopirrolidine-3-carboxamide; PTH: Prothionamide; ETH: Ethionamide; 4PI: N-(4-metilbenzoyl)-4-benzylpiperidine; JPL: 5-(cyclohexa-1,5-dien-1-ylmethyl)-2-(2,4- dichlorophenoxy)phenol; JPJ: 2-(2,4-diclorophenoxy)-5-(2-phenylethyl)phenol; 8PC 2-(2,4-diclorophenoxy)-5-(pyridine-2-ylm etil)phenol; JPM: 5-benzyl-2-(2,4-dichlorophenoxy)phenol; TCU: 5-hexyl-2-(2-methylphenoxy)phenol.

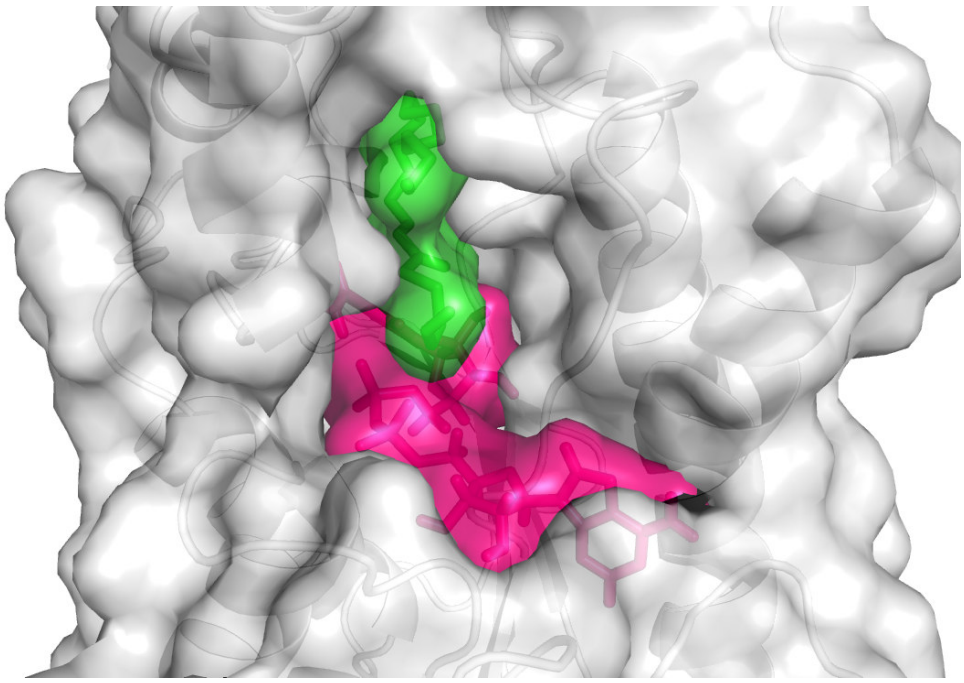


Figure 6. Molecular surface representation of the Mtb InhA binding site. In green a substrate analog, C_{16}, and in pink, the NADH coenzyme (PDB code: 1BVR).

At each subunit, NADH assumes an extended conformation into its binding cavity, along the carboxy-terminal portions of the β-sheet core (except β4 and β5, whose carboxy-terminal portions extend beyond NADH). The nicotinamide ring binds at the bottom of the binding cavity, near the back of the chair, while the adenine portion is directed to the opposite side. Above NADH, over the nicotinamide ring, the substrate assumes an "U" like conformation and is fixed by the substrate-binding loop and other amino acid residues of its binding site.

The amino acids that delimit the outer side of the binding cavity (in relation to the tetramer structure), forming two transverse α helices sustained by loops are called "substrate-binding loop" (residues 196-219) [25]. This loop is an InhA structural motif and is a bit larger in this enzyme in relation to other enoyl reductases (ENRs), which agrees with its the specificity in reducing larger substrates [31]. The amino acids at the opposite side form two big loops equally important to the substrate binding at the active site, the loops A and B. They are situated at the side of the enzyme in contact with the other tetramer subunits, while the two helices that compose the "substrate-binding loop" are pointed out the tetramer, towards the solvent (**Figure 7**). The substrate-binding cavity has an oval format, with the approximated dimensions 16 Å x 13 Å x 7 Å [31]. One side stays completely open

and exposed to the solvent, while the other side has only a small crevice. The terminal portion of the long lipid chain is oriented outside the binding site, enabling the reduction of longer substrate chains. The N-acetylcysteamine portion of the fatty acid substrate (C16 in the InhA structure with PDB access code 1BVR) is directed to the main opening, which is consistent with the necessity of this region of the natural substrate to be turned out, towards the solvent, when bind to InhA [31].

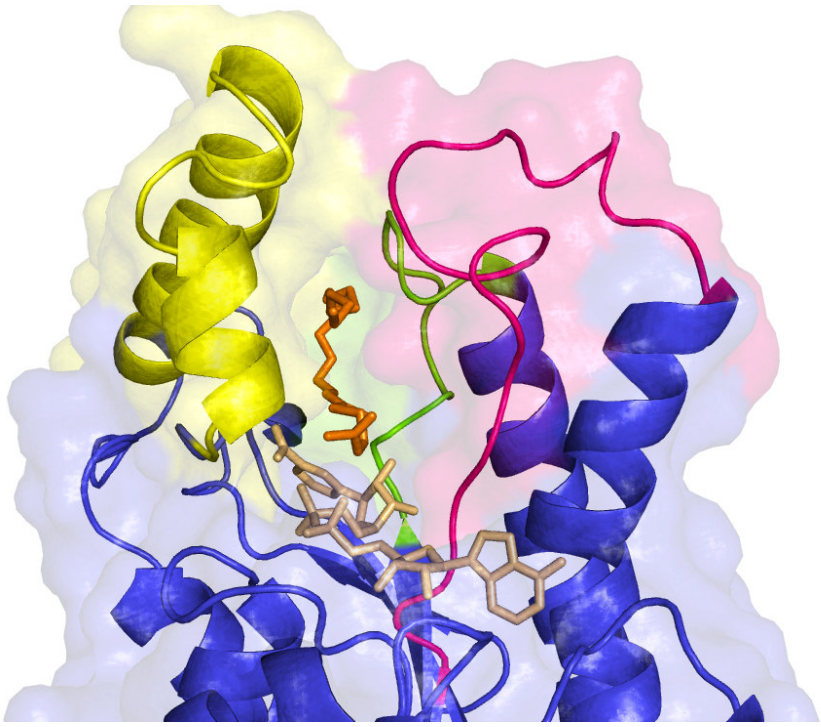


Figure 7. In detail the 3 characteristic structures of the Mtb InhA binding loop. In yellow the substrate binding loop, in magenta the A loop and in green, the B loop (PDB code: 1BVR).

Observing the NADH and the substrate-binding cavities in InhA, one can suppose that the entry of a big substrate would not be easy. As proposed by Rozwarsky and co-workers, the two helices of the “substrate-binding loop” at the left (which are sustained by loops) must move, opening space for the substrate entry. This move may be possible because the substrate-binding loop is external and do not make any contact with another tetramer subunit and so, the opening towards the solvent is unimpeded [31].

Another extremely important feature of the InhA binding cavity is its high level of hydrophobicity. By putting a CH4 probe atom in the center of the protein binding cavity and selecting the neighbor residues in a radius of 6 Å, we could find that all (100%) of the protein amino acids in this region are hydrophobic. When considering a

radius of 7 Å, 73.3% are hydrophobic. Increasing the neighborhood radius to 8 Å, we still have 71.42% of hydrophobic residues. At 9 Å it decreases to, 62.07% coming up to 54.5% in a radius of 10 Å. Nevertheless analyzing the crystallographic structures binding site we found shared features among all structures and some peculiarities that are unique when InhA is associated with an adduct. All InhA structures, when associated with a NADH or an analog, present three water molecules that seem to be important for adenine moiety stabilization. The Gly14 residue appears contributing to the hydrogen bond involving the nitrogen (N3A) of the adenine moiety and the others interactions are, in the most of times, involving the N6A and N7A. These complexes (PDB access code: 2NV6; 2NTJ; 2IEB; 2IE0; 2IDZ and 2H9I) preserve a water molecule next to the isonicotinic ring interacting with the N1Z, and another molecule is mediating a hydrogen bond between the nitrogen atom (N7N) of the nicotinamide ring and the Tyr158. **Figure 8** shows the interactions commented above.

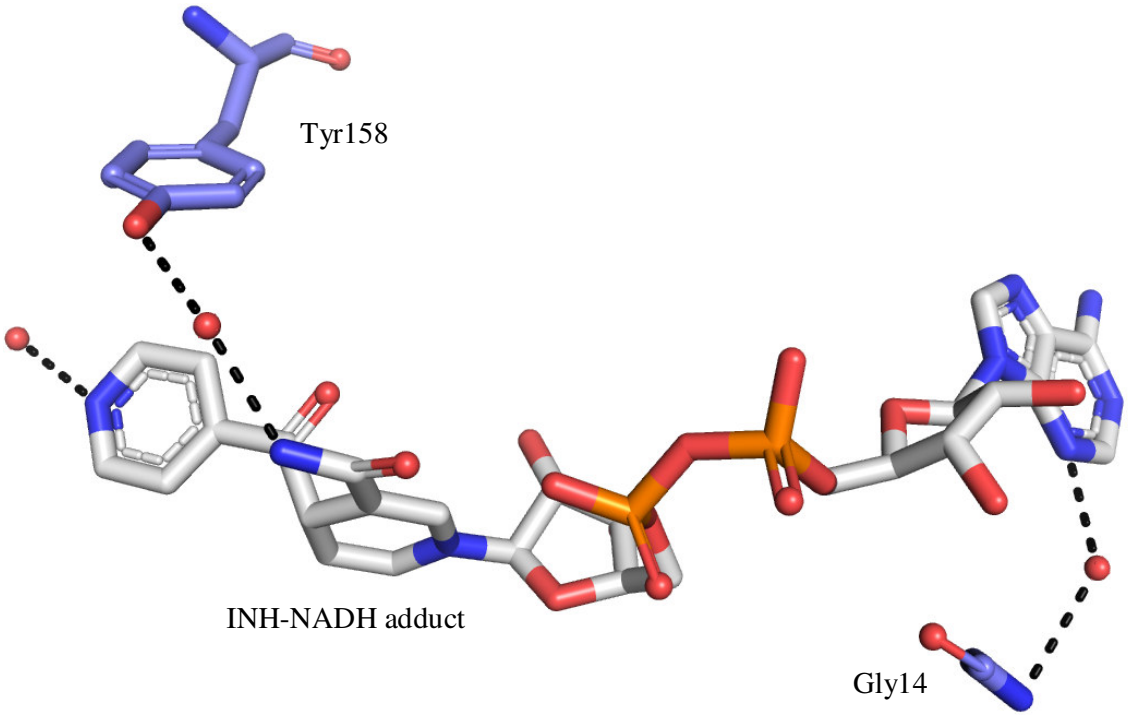


Figure 8. Hydrogen bonds mediated by crystal water molecules (PDB code: 2NV6).

By overlapping the Cα atoms of the monomers of all the 36 crystallographic structures (in an all against all manner) the overall structure topology root mean square deviation (RMSD) among different complexes ranges from 0.1 to 1.4 Å, indicating that the overall structure topology of the different complexes is quite

different. This difference is also evident by analyzing the binding cavity volumes, which range from 1597.3 Å3 up to 3046.7 Å3 for the whole cavity (NADH and substrate cavities). We also demonstrate that the major contribution to flexibility comes from the substrate-binding cavity, with volumes ranging from 327.6 Å3 up to 2109.8 Å3, while the NADH binding site volume is more conserved (ranging from 1292.8 Å3 up to 1525.7 Å3).

In relation to the hydrogen bond pattern into the substrate-binding cavity, the role of a particular residue, Tyr158, deserves to be highlighted. Parikh and collaborators [34] described, in 1999, Tyr158 as an electrophilic catalyst, stabilizing the transition state for hydride transfer by hydrogen bonding to the substrate carbonyl. In addition, also in 1999, Rozwarski *et* al. [31] published a structure of a C16 fatty acid substrate analogue bound to InhA that shows Tyr158 hydrogen bonded to the substrate carbonyl group and rotated from the position it occupies in the InhA-NADH binary complex (**Figure 9**).

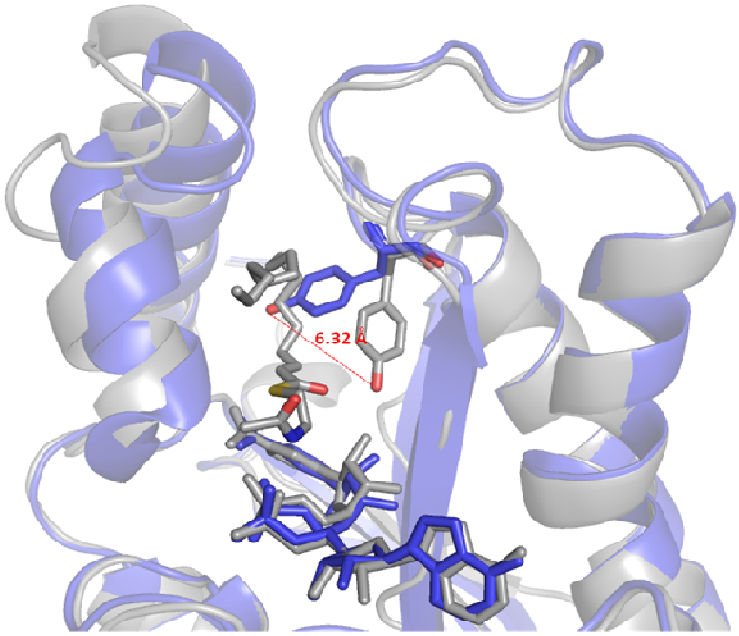


Figure 9. Illustration of the rotation of Tyr158 upon ligand binding. In gray is the InhA-NADH binary complex (PDB code: 1ENY) and in blue the InhA-NADH-C_{16} ternary complex.

We also found in our analyses that all the crystallographic ligands in the InhA substrate-binding cavity hydrogen bond to Tyr158. Another region favorable to hydrogen bond interactions is situated aboard the NADH phosphates. Hydrogen-bonding to atoms O1A, O3 and O3D is, as well, very common among ligands from

crystallographic structures. Still into the NADH binding cavity, Lys165 was proven to play a primary role in cofactor binding by site directed mutation experiments [33].

InhA as a target for drug development

InhA, the fourth enzyme of the *Mtb* FAS II system, has received great attention and is a very well established target to anti-TB drug design initiatives [34]. In order to access the importance of InhA for *Mycobacterium* survival, Vilchèze and co-workers [35], demonstrated that the inactivation of InhA alone was sufficient to inhibit mycolic acid biosynthesis, inducing cell lysis very rapidly after the bacteria exposure to a potent inhibitor, isoniazid.

Despite *inh*A gene mutations facilitate resistance development to one of the most used first line drugs in TB treatment, isoniazid [36], InhA is still an excellent target candidate to the design of novel drugs because: (i) most of the mutations found in isoniazid resistant clinical isolates were related with this pro-drug activator (the enzyme catalase-peroxidase, encoded by *kat*G gene); (ii) *Mtb* has only one enoyl-ACP reductase, differently from other bacterial FAS II systems; and (iii) InhA has specificity for long chain fatty acids, which distinguishes it from other enoyl-ACP reductases, such as the human FAS I enoyl-ACP reductase [37]. Moreover, InhA has the feature of being “druggable”, in other words, it is susceptible to drug inhibition action [34] and this makes this enzyme a target even more attractive to new drug design initiatives.

Available drugs and recent advances in InhA inhibitors design

Isoniazid (INH) has been used as a first line drug against TB since 1952 [38]. It is a pro-drug activated by a catalase-peroxidase encoded by the *kat*G gene. The resulting molecule forms an adduct with NADH which than binds to InhA. This is the same mechanism of action for Ethionamide (ETH) and Prothionamide (PTH), two-second line drugs in the TB treatment, which are activated by a flavin-dependent monooxygenase encoded by the *eth*A gene. INH, ETH and PTH are active only

against sensitive and growing TB, but fail against multi-drug resistant and resting TB due to selection for mutants of the activator proteins.

The emergence of mutations in the INH, ETH and PTH activator genes are related with most of the drug resistance cases. Therefore, compounds that directly inhibit InhA, without requiring previous activation, would be very promising candidates as novel effective drugs for combating resistant *Mtb* strains.

In this way, triclosan (TCN) has been reported to target InhA directly and based on its mechanism of action, many other classes of potent InhA direct inhibitors were designed. Among those are diphenyl ethers, designed by using structure-based drug design [39]. These compounds developed by Peter Tonge and collaborators, are rapid reversible inhibitors of the enzyme, and based on the knowledge that long drug-target residence times are an important factor for in vivo drug activity, they further set out to generate a slow onset inhibitor of InhA using structure based drug design. 2-(o-Tolyloxy)-5-hexylphenol (PT70) is a slow, tight binding inhibitor of InhA with a Ki value of 22pM (**Table 2**). PT70 binds preferentially to the InhA-NAD+ complex and has a residence time of 24 minutes [32].

The great success of this approach prompted scientists to apply high-throughput screening technology to discover novel InhA direct inhibitors. Meanwhile, the popularity of InhA as a screening target may derive from the facts that screens may usually lead to hits and at least two commercially useful compounds act by inhibiting it. Based on this, some novel classes of InhA direct inhibitors have been identified, such as indole-5-amides [40], pyrazole derivatives [40], pyrrolidine carboxamides [41], arylamides [42] and imidazopiperidines [43] among others. A more detailed review on these structures was recently published by Lu and collaborators [39].

In 2004, Oliveira and co-workers reported the identification of an inorganic complex [FeII(CN)5(INH)]3, an INH derivative which binds InhA directly, requiring no previous activation, and no need for the presence of NADH. This molecule may represent a new class of lead compounds to develop new anti-tubercular agents. *In vitro* assays indicated inactivation of both, wild type and I21V mutant strains [44].

The application of a chemical systems biology approach to identify off-targets of major pharmaceuticals on a proteome-wide scale allowed the discovery that existing commercially available drugs, prescribed for the treatment of Parkinson's disease, have the potential to treat MDR and XDR TBs [45]. These drugs,

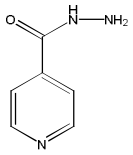
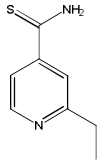
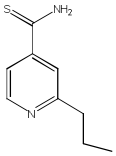
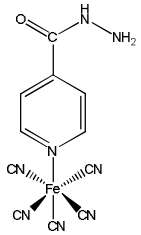
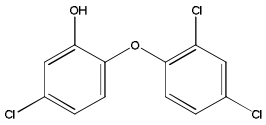
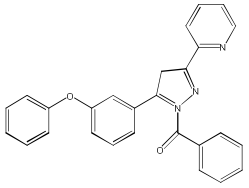
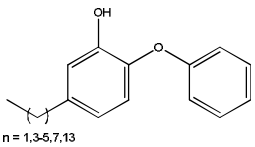
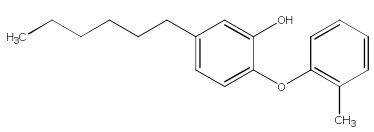
entacapone and tolcapone, are predicted to bind to the enzyme InhA and directly inhibit substrate-binding. Kinetic assays indicate that Comtan inhibits InhA activity by 47%, thus its active component represents a promising lead compound for developing a new class of anti-tubercular therapeutics with excellent safety profiles.

Recently, Vilchèze and collaborators described two compounds (CD39 and CD117) that presented inhibitory activity, both *in vitro* and *in vivo*, against Mtb. These molecules, known inhibitors of *Plasmodium falciparum* enoyl reductase, exhibited increased bactericidal activity when used in combination with INH or rifampin. However, these inhibitors presented inhibitory activity not only to InhA enzyme but also to fatty acid synthase type 1, and the compound CD117 is active against tuberculostearic acid synthesis. Due to the polypharmacology exhibited by this drugs, the authors suggested that this compounds could be used as scaffolds to chemical derivatization [46]. The structures of the most representative molecule of each of these classes of inhibitors, its inhibition levels and molecular structures are summarized in **Table 2**.

Active site residues involved in interactions with the different kind of ligands

To identify the protein amino acid residues interacting with each ligand complexed with the structures available at the PDB, it will be crucial to understand the active site dynamics. This analysis enabled us to know which amino acids did make contacts with which class of ligand as well as the frequency that this occurred. The residues we identified doing interactions with the ligands, including any pair of atoms, plus hydrogen bonds, at a distance less than 4 Å, in at least 50% of the crystallographic structures, are described below. All these interactions are presented in **Table 3**. These analyses indicated the residues most commonly found in contact with the majority of the ligands present in the available *Mtb* InhA crystallographic structures. At the substrate-binding site Met103 and Tyr158 were present doing interactions with almost all the ligands, which is expected due to the importance of the Tyr158 in the ligand stabilization process at the substrate-binding site. The residues Phe149 and Met199 were in contact with a variety kind of ligands and also with the coenzyme NADH when it was present.

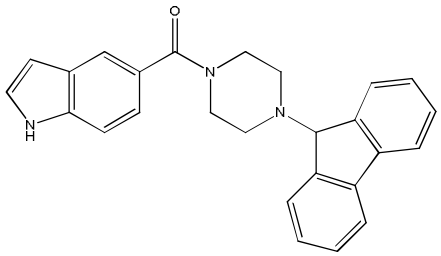
Table 2. Information about the structures of the most representative molecules of each class of inhibitors.

	Compound class	Structure	Inhibition	Reference
INDIRECT INHIBITORS	**1. Isoniazid and annalogs:**			
	Isoniazid (INH)		K_i = 5nM	[1]
	Ethionamide (ETH)		K_i = 7nM	[2]
	Prothionamide (PTH)		K_i = 2nM	[2]
DIRECT INHIBITORS	[FeII(CN)5(INH)]32 (PIF)		MIC = 0.2µg mL-1	[3]
	3. Triclosan and annalogs:			
	Triclosan (TCL)		K_i = 0.2µM	[4]
	3.1 Alkyl Diphenyl Ethers:			
	5-octyl-2-phenoxy-phenol, (8PP)		K_i = 1.1nM	[4]
	Pyrazoline		MIC = 2.4µM	[5]
	5-hexyl-2-(2-methylphenoxy) phenol, (PT70)		K_i = 22pM	**[6]**

DIRECT INHIBITORS

4. Indole-5-Amides:

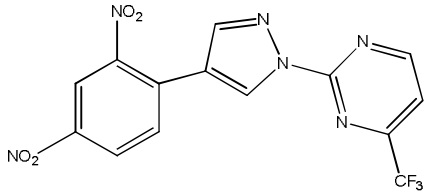
Genz10850



IC_{50} = 0.16μM [7]

5. Pyrazoles

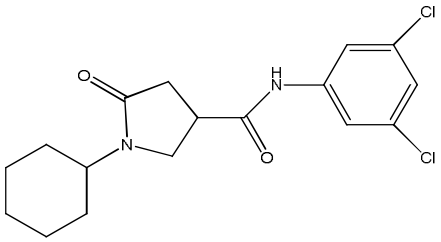
Genz8575



IC_{50} = 2.4μM [7]

6. Pyrrolidine Carboxamides:

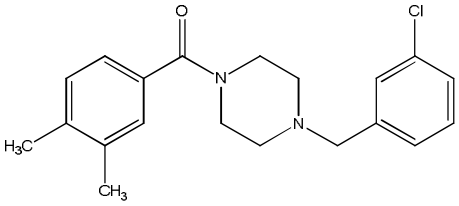
Compound d11



IC_{50} = 0.39μM [8]

7. Arylamides:

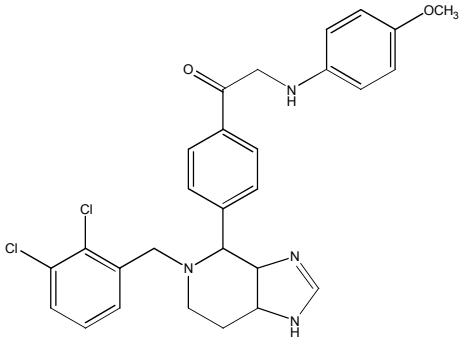
Compound a7



IC_{50} = 0.99μM [9]

8. Imidazopiperidines:

Compound 8e

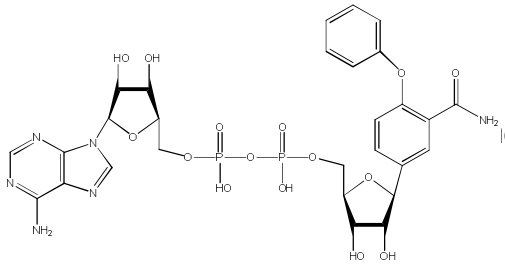


IC_{50} = 0.24μM [10]

DIRECT INHIBITORS

9. INH-NADH adduct annalog:

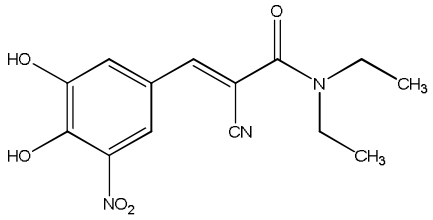
4-phenox benzamide riboside (BAD)



IC_{50} = 27μM [11]

10. Drug discovery using Chemical Systems Biology:

Comtan - Entacapone



MIC_{99} = 260μM [12]

Gly96 and Met161 were the more frequent residues from the NADH binding pocket found in contact with the ligands, besides the residues from the substrate-binding site, when NADH was absent from the structure. Based on the interactions analysis between InhA structures and the different classes of ligands, a variety of residues participate in the ligand accommodation process and do not permit a clear distinguish of specific residues to these different compound classes.

The relation between slow-onset inhibition and the substrate-binding site loop structure maintenance

The majority of lead compounds discovered recently are rapid reversible InhA inhibitors. This is significant given the increasing importance attached to compounds that have long residence times on their targets [47]. This was demonstrated by the highly successful INH-NADH adduct, which was shown to be a slow onset inhibitor of InhA [49].

Table 3. Interactions between Mtb InhA and the different classes of inhibitors.

Residues\Binding pocket	NADH	Pro-drugs	TCN and derivates	Indole-5-Amides	Pyrrolidine carboxamides	Arylamides	substrate analog
Gly15	X	X					
Ile15	X	X					
Ile16	X	X					
Ser20	X	X					
Ile21	X	X					
Phe41	X	X					
Leu63	X	X					
Asp64	X	X					
Val65	X	X					
Ser94	X	X					
Gly95	X	X					
Gly96	X	X	X	X	X	X	
Phe97			X		X		
Met98			X				
Gln100							X
Met103			X	X	X	X	X
Ile122	X	X					
Met147	X	X					
Asp148	X	X					
Phe149	X	X	X	X	X	X	X
Met155		X	X	X			
Pro156			X	X	X		
Ala157			X	X	X		X
Tyr158		X	X	X	X	X	X
Met161		X	X	X			
Lys165	X	X	X		X		
Ala191	X	X					
Gly192	X	X					
Pro193	X	X		X		X	
Ile194	X	X					
Thr195	X	X					
Ala198			X				X
Met199		X	X		X	X	X
Ile202						X	X
Val203						X	
Ile215				X	X		X
Leu218			X	X		X	
Glu219						X	
Trp222		X				X	

The crystal structures of InhA bound to INH-NADH adduct (PDB access code: 2NV6) [49] and the *Escherichia coli* enoyl reductase, ecFabI, (PDB access code: 1QG6) [50], support the hypothesis that slow onset inhibition is coupled to ordering of the binding site loop (residues 195-210 in InhA), which leads to a closure of the substrate-binding pocket.

Towards the design of a slow onset *Mtb* InhA inhibitor, Luckner and collaborators [32] speculated that there must be an entropic penalty for loop ordering. They introduced a methyl group to a diphenyl ether compound, resulting in PT70, a slow onset inhibitor of InhA with a Ki value of 22pM. The crystal structure of InhA bound to PT70 (PDB access codes: 2X22 and 2X23) corroborates that slow onset inhibition is coupled to ordering of the substrate-binding loop, providing information on the mechanistic imperatives for slow onset inhibition of enoyl ACP reductases.

REFERENCES

[1] Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C. Tuberculosis. *Lancet* (2003) 362, 887-99.

[2] WHO. World Health Organization. Factsheet on tuberculosis, (2006).

[3] Ducati RG, Ruffino-Netto A, Basso LA, Santos DS. The resumption of consumption - a review on tuberculosis. *Mem Inst Oswaldo Cruz.* (2006) 101, 697-714.

[4] Fätkenheuer G, Taelman H, Lepage P, Schwenk A, Wenzel R. The return of tuberculosis. *Diagn. Microbiol. Infect. Dis.* (1999) 34, 139-146.

[5] Heymann DL. Resistance to anti-infective drugs and the threat to public health. *Cell*. (2006) 124(4), 671-675.

[6] Koul A, Arnoult E, Lounis N, Guillemont J, Andries K. The challenges of new drug discovery for tuberculosis (2011), *Nature*, 469, 483-490.

[7] White SW, Zheng J, Zhang YM, Rock CO. The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem.* (2005) 74, 791-831.

[8] Zhang YM, Lu YJ, Rock CO. The reductase steps of the type II fatty acid synthase as antimicrobial targets. *Lipids*. (2004) 39, 1055-1060.

[9] Campbell JW, Cronan JE Jr. Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. *Annu. Rev. Microbiol.* (2005) 55, 305-332.

[10] Fulco AJ. Fatty acid metabolism in bacteria. *Prog. Lipid Res.* (1983) 22**,** 133-160.

[11] Kater MM, Koningstein GM, Nijkamp HJJ, Stuitje AR. The use of a hybrid genetic system to study the functional relationship between prokaryotic and plant multi-enzyme fatty acid synthetase complexes. *Plant Mol. Biol.* (1994) 25, 771-790.

[12] Rock CO, Cronan JE. Escherichia coli as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim. Biophys. Acta*. (1996) 1302, 1-16.

[13] Smith S. The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *FASEB J.* (1994) 8, 1248-1259.

[14] Kolattukudy PE, Poulose AJ, Buckner JS. Fatty acid synthase from the uropygial gland of goose. *Methods Enzymol.* (1981) 71, 103-109.

[15] Wakil SJ, Stoops JK, Joshi VC. Fatty acid synthesis and its regulation. *Annu. Rev. Biochem.* (1983) 52, 537-579.

[16] Bloch K. Control mechanisms for fatty acid synthesis in Mycobacterium smegmatis. *Adv. Enzymol.* (1977) 45, 1-84.

[17] Qureshi N, Sathyamoorthy N, Takayama K. Biosynthesis of C30 to C56 fatty acids by an extract of Mycobacterium tuberculosis H37Ra. *J. Bacteriol.* (1984) 157, 46-52.

[18] Brennan PJ, Nikaido H. The envelope of mycobacteria. *Annu. Rev. Biochem.* (1995) 64, 29-63.

[19] Lee RE, Brennan PJ, Besra GS. Mycobacterium tuberculosis cell envelope. *Curr. Top. Microbiol. Immunol.* (1996) 215, 1-27.

[20] Kumar A, Siddiqi MI. CoMFA based de novo design of Pyrrolidine Carboxamides as Inhibitors of Enoyl Acyl Carrier Protein Reductase from Mycobacterium tuberculosis. *J. Mol. Model*. (2008) 14, 923-935.

[21] Schroeder EK, de Souza ON, Santos DS, Blanchard JS, Basso LA. Drugs that inhibit mycolic acid biosynthesis in Mycobacterium tuberculosis. *Curr. Pharm. Biotechnol*. (2002) 3, 197-225.

[22] Quémard A, Sacchettini JC, Dessen A, Vilchèze C, Bittman R, Jacobs WR Jr, Blanchard JS. Enzymatic characterization of the target for isoniazid in Mycobacterium tuberculosis. *Biochemistry*, (1995) 34, 8235-8241.

[23] Parikh SL, Xiao G, Tonge PJ. Inhibition of InhA, the enoyl reductase from Mycobacterium tuberculosis, by triclosan and isoniazid. *Biochemistry*, (2000) 39, 7645-7650.

[24] Rossmann MG, Liljas A, Branden CI, Banaszak LJ. *Enzymes*, (1975) 11, 61-102.

[25] Dessen A, Quémard A, Blanchard JS, Jacobs Jr. WR, Sacchettini JC. Crystal structure and function of the isoniazid target of Mycobacterium tuberculosis. *Science*, (1995) 267, 1638-1641.

[26] Sachettini JC, Blanchard JS. The structure and function of the isoniazid target in M. tuberculosis. *Res. Microbiol*. (1996) 147(1-2), 36-43.

[27] Kruh NA, Rawat R, Ruzsicska BP, Tonge PJ, Probing mechanisms of resistance to the tuberculosis drug isoniazid: Conformational changes caused by inhibition of InhA, the enoyl reductase from *Mycobacterium tuberculosis*, *Protein Science*, 16(8), 1617-1627.

[28] Oliveira JS, Pereira JH, Canduri F, Rodrigues NC, de Souza ON, de Azevedo WF Jr, Basso LA, Santos DS. Crystallographic and pre-steady-state kinetics studies on binding of NADH to wild-type and isoniazid-resistant enoyl-ACP(CoA) reductase enzymes from Mycobacterium tuberculosis. *J. Mol. Biol.* (2006) 359, 646-66.

[29] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res.* (2000) 28, 235-242.

[30] Bernstein FC, Koetzle TF, Williams GJB, Meyer Jr. EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* (1977) 112, 535-542.

[31] Rozwarski DA, Vilchèse C, Sugantino M, Sacchettini JC. Crystal structure of the Mycobacterium tuberculosis enoyl-ACP reductase, InhA, in complex with NAD+ and a C16 fatty acyl substrate. *J. Biol. Chem.* (1999) 274, 15582-15589.

[32] Luckner SR, Liu N, am Ende CW, Tonge PJ, Kiske C. A Slow, Tight Binding Inhibitor of InhA, the Enoyl-Acyl Carrier Protein Reductase from *Mycobacterium tuberculosis. J. Biol. Chem*. (2010) 285(19), 14330-14337.

[33] Parikh SL, Xiao G, Tonge PJ. Inhibition of InhA, the enoyl reductase from Mycobacterium tuberculosis, by triclosan and isoniazid. *Biochemistry*. (2000) 39, 7645-7650.

[34] Agüero F, Al-Lazikani B, Aslett M, Berriman M, Buckner FS, Campbell RK, Carmona S, Carruthers IM, Chan AWE, Chen F, Crowther GJ, Doyle MA, Hertz-Fowler C, Hopkins AL, McAllister G, Nwaka S, Overington JP, Pain A, Paolini GV, Pieper U, Ralph SA, Riechers A, Roos DS, Sali A, Shanmugam, D, Suzuki T, Van

Voorhis WC, Verlinde CLMJ, Genomic-scale prioritization of drug targets: the TDR Targets database (2008) *Nature Reviews in Drug Discovery*, 900-907.

[35] Vilchèze C, Morbidoni HR, Weisbrod TR, Iwamoto H, Kuo M, Sacchettini JS, Jacobs Jr. W.R. 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 (2000) *Journal of Bacteriology*, 182, 4059-4067.

[36] Mdluli K, Slayden RA, Zhu Y, Ramaswamy S, Pan X, Mead D, Crane DD, Musser JM, Barry CE III. Inhibition of a Mycobacterium tuberculosis beta-ketoacyl ACP synthase by isoniazid (1998) *Science*, 280, 1607-1610.

[37] Quémard A, Dessen A, Sugantino M, Jacobs WRJr, Sacchettini JC, Blanchard JS (1996) *Journal of the American Chemical Society*, 118, 1561-1562.

[38] Middlebrook G. Sterilization of tubercle bacilli by isonicotinic acid hydrazide and the incidence of variants resistant to the drug in vitro. *The American review of tuberculosis and pulmonary diseases*, (1952) 65, 765-770.

[39] Lu XY, You QD, Chen ID. Recent progress in the identification and development of InhA direct inhibitors of Mycobacterium tuberculosis. *Mini Rev Med Chem*. 2010 10(3):181-92.

[40] Kuo MR, Morbidoni HR, Alland D, Sneddon SF, Gourlie BB, Staveski MM, Leonard M, Gregory JS, Janjigian AD, Yee C, Musser JM, Kreiswirth B, Iwamoto H, Perozzo R, Jacobs WR Jr, Sacchettini JC, Fidock DA. Targeting tuberculosis and malaria through inhibition of Enoyl reductase: compound activity and structural data. *J Biol Chem*., (2003) 278(23):20851-9.

[41] He X, Alian A, Stroud R, Ortiz de Montellano PR. Pyrrolidine carboxamides as a novel class of inhibitors of enoyl acyl carrier protein reductase from Mycobacterium tuberculosis. *J Med Chem*., (2006) 49(21):6308-23.

[42] He X, Alian A, Ortiz de Montellano PR. Inhibition of the Mycobacterium tuberculosis enoyl acyl carrier protein reductase InhA by arylamides. *Bioorg Med Chem*., (2007) 15(21):6649-58.

[43] Wall MD, Oshin M, Chung GAC, Parkhouse T, Gore A, Herreros E, Cox B, Duncan K, Evans B, Everetta M, Mendoza, A. Evaluation of N-(phenylmethyl)-4-[5-(phenylmethyl)-4,5,6,7-tetrahydro-1H-imidazo[4,5-c]pyridin-4-yl]benzamide inhibitors of Mycobacterium tuberculosis growth. *Bioorganic & Medicinal Chemistry Letters*

(2007) 17, 2740–2744.

[44] Oliveira JS, Sousa EHS, Basso LA, Palaci M, Dietze R, Santos DS, Moreira IS. An inorganic iron complex that inhibits wild-type and an isoniazid-resistant mutant 2-trans-enoyl-ACP (CoA) reductase from Mycobacterium tuberculosis. *Chem. Commun*. (2004) 3, 312-313.

[45] Kinnings SL, Liu N, Buchmeier N, Tonge PJ, Xie L, Bourne PE. Drug discovery using chemical systems biology: repositioning the safe medicine Comtan to treat multi-drug and extensively drug resistant tuberculosis *PLOS Comput. Biol*. (2009) 5, 1-10.

[46] Vilchèze C, Baughn AD, Tufariello J, Leung L, Kuo M, Basler C, Alland D, Sacchettini JC, Freundlich JS, Jacobs WR Jr. Novel inhibitors of InhA efficiently kill Mycobacterium tuberculosis under aerobic and anaerobic conditions. *Antimicrob. Agents Chemother*. (2011) May 31 (Epub ahead of print).

[47] Copeland RA, Pompliano DL, Meek TD. Drug-target residence time and its implications for lead optimization. *Nat. Rev Drug Discov.* (2006) **5**(9), 730-739.

[48] Rawat R, Whitty A, Tonge P. 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 U.S.A*. (2003) 100, 13881-13886.

[49] Vilchèze C, Wang F, Arai M, Hazbón MH, Colangeli R, Kremer L, Weisbrod TR, Alland D, Sacchettini JC, Jacobs WR Jr., Transfer of a point mutation in Mycobacterium tuberculosis inhA resolves the target of isoniazid. *Nat Med*. (2006), 12(9):1027-9.

[50] Oliveira JS, Sousa EHS, Basso LA, Palaci M, Dietze R, Santos DS, Moreira IS. An inorganic iron complex that inhibits wild-type and an isoniazid-resistant mutant 2-trans-enoyl-ACP(CoA)reductase from Mycobacterium tuberculosis. *Chem . Commun*. (2004), 312-313.

6 CONSIDERAÇÕES FINAIS

A pesquisa e o desenvolvimento de compostos mais efetivos contra a TB representam uma necessidade urgente à saúde pública mundial. O ressurgimento de diferentes cepas de Mtb resistentes à terapia medicamentosa existente reforça a necessidade de gerar novas formas de combater a doença e melhorar as condições do tratamento utilizado atualmente.

Nesse contexto, as enzimas que compõem as rotas metabólicas envolvidas em processos bioquímicos essenciais à viabilidade do bacilo, apresentam-se como alvos atrativos para o desenvolvimento racional de novos fármacos. Devido a diferenças distintivas na organização molecular em relação à via metabólica humana (possibilitando o planejamento de fármacos mais seletivos e menos tóxicos), as enzimas que participam da biossíntesse de ácidos graxos em Mtb são alguns dos alvos mais promissores para o planejamento de novos agentes antibacterianos. Dentre as enzimas dessa via metabólica, a InhA tem recebido grande destaque e é um alvo muito bem estabelecido para o desenvolvimento de fármacos anti-TB.

Neste trabalho empregamos metodologias computacionais ainda não utilizadas até o momento para a busca de novos inibidores para a InhA de Mtb, resultando na identificação de potenciais compostos líderes, para o desenvolvimento de novos medicamentos para o combate à TB, conforme objetivo proposto no capítulo 2.

Os objetivos específicos relacionados à construção de um modelo farmacofórico 3D, execução de experimentos de docagem molecular e triagem virtual foram satisfeitos utilizando duas abordagens distintas.

Na primeira, a análise estrutural da InhA em complexo com diversos inibidores ligados à cavidade de ligação do substrato permitiu a identificação de resíduos de aminoácidos importantes para o reconhecimento molecular.

Essas informações guiaram o desenvolvimento de um modelo farmacofórico 3D de quatro pontos baseado na estrutura do receptor (InhA de Mtb) e que incorporou: (i) duas regiões favoráveis para interações hidrofóbicas: uma na região da alça de ligação ao substrato, essencial para a efetividade da inibição e outra, próxima às alças A e B; (ii) uma região favorável para a interação com grupos doadores ou aceitadores de ligação de hidrogênio (cadeia lateral da Tyr158); (iii) uma região

favorável para interação com grupos doadores de ligação de hidrogênio próximo aos grupamentos fosfato da coenzima NADH. Estes pontos foram utilizados como restrições para a triagem virtual dos ligantes da biblioteca selecionada do banco de dados ZINC com o programa UNITY. Para serem selecionados, os ligantes deveriam satisfazer os quatro pontos farmacofóricos estabelecidos após processos de minimização de energia e docagem molecular com o programa GOLD.

Na segunda abordagem, foi testada a habilidade de quatro diferentes programas de docagem molecular e seus diferentes algoritmos em encontrar a mesma conformação para um determinado ligante, o que pode fornecer indícios sobre o modo mais correto de interação entre a proteína e o ligante.

O conjunto final obtido pela aplicação das duas metodologias (satisfazendo ao objetivo específico de identificação dos compostos mais promissores) foi de 55 moléculas (34 da primeira abordagem e 21 da segunda). Estas foram submetidas a testes de toxicidade *in silico* afim de eliminar previamente aquelas não promissoras. Destas 55, restaram 19.

Para satisfazer o último objetivo específico descrito no capítulo 2, a seleção final das moléculas a serem testadas *in vitro* para sua capacidade de inibir a InhA de Mtb foi baseada em todas as análises descritas acima. Das 19 moléculas restantes, 6 foram escolhidas para testes *in vitro* com o objetivo de satisfazer diferentes modos de interação com o sítio de ligação do substrato da InhA.

Este trabalho, bem como a redação do seu manuscrito estão em fase de finalização, aguardando a chegada das moléculas selecionadas pelos métodos *in silico*, para que sua atividade inibitória seja testada *in vitro* no Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF), integrante do INCT-TB, localizado no Parque Tecnológico da PUCRS (TECNOPUC). Estes experimentos estão sendo coordenados pelos professores Diógenes Santiago Santos e Luiz Augusto Basso.

Estamos trabalhnando também, num artigo de revisão do estado da arte sobre o conhecimento estrutura/função da enzima InhA de Mtb. O manuscrito encontra-se em fase de finalização e será submetido para uma revista de impacto na área em breve. Este artigo será importante para a comunidade científica, uma vez que reunirá informações para o planejamento racional de fármacos tendo esta enzima como alvo.

Foi demonstrado que a enzima InhA é o alvo primário da isoniazida (INH), um dos mais antigos fármacos sintéticos anti-TB. A INH é uma pró-droga biologicamente

ativada pela enzima catalase-peroxidase (KatG). A reação de ativação promove a formação de um aduto isonicotinil-NAD que inibe a enzima InhA, resultando na redução da biossíntese dos ácidos micólicos. O mecanismo de ação da INH é complexo, uma vez que mutações em pelo menos cinco genes diferentes (*katG, inhA, ahpC, kasA, and ndh*) estão correlacionadas com a resistência a este fármaco. Como resultado de pesquisas envolvendo o desenvolvimento racional de fármacos, para o planejamento de moléculas alternativas capazes de inibir a InhA de Mtb, o complexo inorgânico pentaciano(isoniazida)ferrato(II), o PIF, foi desenvolvido. Esta molécula foi capaz de inibir tanto a InhA do tipo selvagem quanto o mutante Ile21Val e esta inativação não necessitou de ativação pela KatG. No trabalho, que resultou no artigo submetido à revista "*Journal of molecular modeling*" e descrito detalhadamente no **Anexo 1**, foram relatados os resultados, obtidos por meio de simulações de dinâmica molecular, da interação entre a InhA e o PIF. Neste trabalho relatamos em detalhes as perturbações causadas na estrutura tridimensional da proteína pela presença deste inibidor. Com isto o objetivo específico referente à execução de simulações por dinâmica molecular para compreender melhor o processo de interação proteína-ligante foi satisfeito.

Em resumo, podemos dizer que este trabalho de Mestrado forneceu novidades no âmbito do conhecimento científico sobre a enzima InhA de *Mycobacterium tuberculosis* que serão úteis para o desenvolvimento de projetos futuros tanto para o LABIO, INCT-TB, quanto para os demais interessados neste campo da pesquisa.

REFERÊNCIAS

Agüero, F.; Al-Lazikani, B.; Aslett, M.; Berriman, M.; Buckner, F.S.; Campbell, R. K.; Carmona, S.; Carruthers, I.M.; Chan, A.W.E.; Chen, F.; Crowther, G.J.; Doyle, M.A.; Hertz-Fowler, C.; Hopkins, A.L.; McAllister, G.; Nwaka, S.; Overington, J.P.; Pain, A.; Paolini, G.V.; Pieper, U.; Ralph, S.A.; Riechers, A.; Roos, D.S.; Sali, A.; Shanmugam, D.; Suzuki, T.; Van Voorhis, W.C.; Verlinde C.L.M.J. (2008) *Nature Reviews in Drug Discovery*, 900-907.

Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. (2002) *Molecular Biology of the Cell*, New Yourk and London: Garland Science.

Alonso H.; Bliznyuk A.A.; Gready J.E., (2006) *Med. Res. Reviews*., 26, 531.

Andricopulo, A.D.; Salum, L.B.; Abraham, D.J. (2009) *Current Topics in Medicinal Chemistry*, 9, 771-790.

Argyrou, A.; Vetting, M.W.; Blanchard, J.S. (2007) *Journal of the Ammerican Chemical Society*., 129, 9582-3.

Banerjee, A.; Sugantino, M.; Sacchettini, J.C.; Jacobs, W.R. Jr. (1998) *Microbiology*, 144, 2697-2704.

Banerjee, R.; Schecter, G.F.; Flood, J.; Porco T.C. (2008) *Expert Review of Anti-infective Therapy* 6, 713-724.

Barry, C.E. 3rd.; Lee, R.E.; Mdluli, K.; Sampson, A.E.; Schroeder, B.G.; Slayden, R.A.; Yuan, Y. (1998) *Progress in Lipid Research* 37, 143-79.

Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne P.E. (2000) *Nucleic Acids Research* 28, 235-242.

Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.B.; Meyer Jr., E.F.; Brice, M.D.; Rodgers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. (1977) *Journal of Molecular Biology* 112, 535-542.

Blanchard, J.S. (1996) *Annual Reviews of Biochemistry*, 65, 215-239.

Bonnac, L.; Gao, G.Y.; Chen, L.; Felczak, K.; Bennett, E.M.; Xu, H.; Kim, T.S.; Liu, N.; Oh, H.W.; Tonge, P.J.; Pankiewicza, K.W. (2007) *Bioorganic & Medicinal Chemistry Letters*, 17, 4588-4591.

Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; Karplus, M. (1983) *Journal of Computational Chemistry*, 4, 187-217.

Broussy, S.; Bernardes-Ge´nisson, G.; Que´mard, A.; Meunier, B.; Bernadou, J.; (2005) *The Journal of Organic Chemistry*, 70, 10502-10510.

Campbell, J.W.; Cronan, J.E.Jr. (2001) *Annual Review of Microbioogy.,* 55, 305-332.
Camus, J.C.; Pryor, M.J.; Médigue, C.; Cole, S.T. (2002) *Microbiology*, 148, 263-270.

Case, D.A.; Darden T.A.; Cheatham III, T.E.; Simmerling, C.L.; Wang J.; Duke, R.E.; Luo, R.; Merz, K.M.; Pearlman, D.A.; Crowley, M.; Walker, R.C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K.F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D.H.; Schafmeister, C.; Ross W.S.; Kollman, P.A., (2006), AMBER 9, University of California, San Francisco.

Cavasotto, C.N.; Orry, A.J. (2007) *Current. Topics in Medicinal Chemistry*, 7, 1006-1014.

Cole, S. T.; Barrell B. G. (1998) *Novartis Foundation Symposium*, 160-172.

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). *Journal of the American Chemical Society*, 117, 5179-5197.

DeLano, W. L. (2002) *DeLano Scientific*, San Carlos, CA, USA, [http://www.pymol.org].

Dessen, A.; Quémard, A.; Blanchard, J.S.; Jacobs, Jr. W.R.; Sacchettini, J.C. (1995) *Science*, 267, 1638-1641.

Dias, M.V.; Vasconcelos, I.B.; Prado, A.M.; Fadel, V.; Basso, L.A.; de Azevedo, W.F.Jr.; Santos, D.S. (2007) *Journal of Structtural Biology*, 159, 369-80.

Di Perri, G.; Aguilar Marucco, D.; Mondo, A.; Gonzalez de Requena, D.; Audagnotto, S.; Gobbi, F.; Bonora, S. *Expert Opinion in Drug Safety*, 4, 821-836.

Ducati, R.G.; Ruffino-Netto, A.; Basso, L.A.; Santos, D.S. (2006) *Memórias do Instituto Oswaldo Cruz*, 101, 697-714.

Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. (2006) *Nucleic Acids Research*, 34, 116-118.

Dunkel, M.; Fullbeck, M.; Neumann, S.; Preissner, R. (2006) *Nucleic Acids Research,* 34, 678-683.

Fätkenheuer, G.; Taelman, H.; Lepage, P.; Schwenk, A.; Wenzel, R. (1999) *Diagnostic Microbiology and Infectious Disease,* 34, 139-146.

Fillgrove, K.L.; Anderson, V.E. (2000) *Biochemistry*, 39, 7001-7011.

Freundlich, J.S.; Wang, F.; Vilchèze, C.; Gulten, G.; Langley, R.; Schiehser, G.A.; Jacobus, D.P.; Jacobs, W.R.Jr.; Sacchettini, J.C. (2009) *ChemMedChem*, 4, 241-248.

Frieden, T.R.; Sterling, T.R.; Munsiff, S.S.; Watt, C.J.; Dye, C. (2003) *Lancet*, 362,

887-899.

Fulco, A.J. (1983) *Progress in Lipid Research,* 22**,** 133-160.

Gagneux, S. (2009) *Clinical Microbiology and Infection* 1, 66-68.

Global tuberculosis **1.** control: epidemiology, strategy, financing: WHO report 2009 (Publication no. WHO/HTM/TB/2009.411).

Goodford, P.J. (1985) *Journal of Medicinal Chemistry*, 28, 849-857.

Goodsell, D.S; Olson, A.J. (1990) *Protein Structure and Functional Genomics*, 8, 195-202.

Goodsell, D.S.; Morris, G.M.; Olson, A.J. (1996). *Journal of Molecular Recognition*, 9, 1-5.

Guex, N and Peitsch, M. C. (1997) *Electrophoresis*, 18, 2714-2723.

He, X.; Alian, A.; Stroud, R.; Ortiz de Montellano, P.R. (2006) *Journal of Medicinal Chemistry*, 49, 6308-23.

He, X.; Alian, A.; Ortiz de Montellano, P.R. (2007) *Bioorganic and Medicinal Chemistry*, 15, 6649-6658.

Hearn, M.J.; Cynamon, M.H. (2004) *Journal of Antimicrobial Chemotherapy*, 53, 185-191.

Helma, C. In silico Predictive Toxicology: The State of the Art and Strategies to Predict Human Health Effects. Inst. f. Computer Science Univ. Freiburg October 19, 2004.

Henry, C.M. (2001). *Chemical & Engineering News*, 79, 69-74.

Huey, R.; Morris, G.M.; Olson, A.J.; Goodsell, D.S.J. (2007) *Journal of Computational Chemistry*, 28(6), 1145-1152.

Humphrey, W.; Dalke, A. and Schulten, K. (1996) *Journal of Molecular Graphics,* 14, 33-38.

Imramovsky, A.; Polanc, S.; Vinsˇova, J.; Kocˇevar, M.; Jampı´lek, J.; Recˇkova, Z.; Kaustova, J.; (2007) *Bioorganic & Medicinal Chemistry*, 15, 2551-2559.

Irwin, J.; Shoichet, B.K. (2005) *Journal of Chemical Information and Modeling*, 45, 177-182.

Jain, A. N. (2003) *Journal of Medicinal Chemistry*, *46*, 499-511.

Jain, A.; Mondal, R. (2008) *FEMS Immunology and Medicinal Microbiology,* 53, 145-50.

Jiang, Z. (2008) *Current Pharmaceutical Design*, 6, 588-592.

Jones, G.; Willett, P.; Glen R.C.; Leach A.R.; Taylor, R. (1997) *Journal of Molecular Modeling*, 267, 727-748.

Jorgensen, W.L. (2004) *Science*, 303, 1813-1818.

Kater, M.M.; Koningstein, G.M.; Nijkamp, H.J.J.; Stuitje, A.R. (1994) *Plant Molecular Biology,* 25, 771-790.

Kinnings, S.L.; Liu, N.; Buchmeier, N.; Tonge, P.J.; Xie, L.; Bourne, P.E. (2009) *PLOS Computational Biology*, 5, 1-10.

Kolattukudy, P.E.; Poulose, A.J.; Buckner, J.S. (1981) *Methods in Enzymology,* 71, 103-109.

Kitchen, D.B.; Decormez, H.; Furr, J.R.; Bajorath, J. (2004) *Nature Reviews in Drug Discovery*, 3, 935-949.

Kolattukudy, P.E.; Fernandes, N.D.; Azad, A.K.; Fitzmaurice, A.M.; Sirakova, T.D. (1997) *Molecular Microbiology*, 24, 263-270.

Krumrine, J.; Raubacher, F.; Brooijimans, N.; Kuntz, I. (2003). Edited by Philipe E. Bourne & Helge Weissig, Wiley-Liss, New Jersey, USA.

Kumar, A.; Siddiqi, M.I. (2008) *Journal of Molecular Modeling*, 14, 923-935.

Kuntz, I.D. (1992) *Science*, 257, 1078-1082.

Kumar, A.; Siddiqi, M.I. (2008) *Journal of Molecular Modeling*, 14, 923-935.

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, W.R.Jr.; Sacchettini, J.C.; Fidock, D.A. (2003) *The Journal of Biological Chemistry,* 278, 20851-20859.

Lazo, J.S. (2006) *Molecular Interventions*, 6, 240-243.

Lindstrom, W.; Morris, G.M.; Weber, C.; Huey, R. (2008) *Using AutoDock 4 for Virtual Screening*. The Scripps Research Institute Molecular Graphics Laboratory 10550 N. Torrey Pines Rd. La Jolla, California 92037-1000, USA, v2.

Leach, A.R.; Shoichet, B.K.; Peishoff, C.E. (2006) *Journal of Medicinal Chemistry*, 49, 5851-5855.

Lu, X.Y.; Chen, Y.D.; Jiang, Y.J.; You, Q.D. (2009) *European Journal of Medicinal Chemistry*, 44, 3718-3730.

Luckner, S.R.; Liu, N.; Am Ende, C.W.; Tonge, P.J.; Kisker, C. (2010) The *Journal of Biological Chemistry*, 285, 14330-14337.

Marrachi, H.; Ducase, S.; Labesse, G.; Montrozier, H.; Margeat, E.; Emorine, L.; Charpentier, X.; Daffé, M.; Quémard, A. (2002) *Microbiology*, 148, 951-960.

Mayr, L.M.; Bojanic, D. (2009) *Current Opinion in Pharmacology*, 9, 580-588.

McCammon J.A.; Gelin B.R.; Karplus M., (1977) *Nature*, 267(5612), 585.

McMurry, L.M.; Oethinger, M.; Levy, S.B. (1998) *Nature*, 394, 531-532.

Mdluli, K.; Slayden, R.A.; Zhu, Y.; Ramaswamy, S.; Pan, X.; Mead, D.; Crane, D.D. Musser, J.M.; Barry, C.E. III. (1998) *Science*, 280, 1607-1610.

Middlebrook, G. (1952) *The American review of tuberculosis and pulmonary diseases*, 65, 765-770.

Mills, N. (2006). *Journal of the American Chemical Society,* 128, 13649-13650.

Molle, V.; Gulten G.; Vilchèze, C.; Veyron-Churlet, R.; Zanella-Cléon, I.; Sacchettini, J.C.; Jacobs Jr. W.R.; Kremer L. (2010) *Molecular Microbiology*, 78(6), 1591-1605.

Morris, G.M.; Goodsell, D.S.; Huey, R.; Olson, A.J. (1996) *Journal of Computer-Aided Molecular Design*, 10, 293-304.

Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. (1998) *Journal of Computational Chemistry*, 19, 1639-1662.

Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. (2008) *Journal of Computational Chemistry*, 30(16), 2785-2791.

Morrison, J.; Pai, M.; Hopewell, P.C. (2008) *The Lancet Infectious Diseases,* 8, 359-68.

Muegge, I.; Martin, Y.C. (1999) *Journal of Medicinal Chemistry,* 42, 791-804.

Oliveira, J.S.; Sousa E.H.S.; Basso, L. A.; Palaci, M.; Dietze, R.; Santos, D.S.; Moreira, I.S. (2004) *Chemical. Communications*, 312-313.

Oliveira, J.S.; Pereira, J.H.; Canduri, F.; Rodrigues, N.C.; de Souza, O.N.; de Azevedo, W.F. Jr.; Basso, L.A.; Santos, D.S. (2006) *Journal of Molecular Biology*, 359, 646-66.

Oliveira, J.S.; Souza, E.H.S; Norberto de Souza, O.; Moreira, I.S.; Santos, D.S.; Basso, L.A. (2006) Curr. Pharm. Design, 12, 2409-2424.

Österberg, F.; Garret, M.M.; Sanner, M.F.; Olson A.J.; Goodsell, D.S. (2002) *PROTEINS: Structure, Function and Genetics*, 46, 34-40.

Parikh, S.L.; Xiao, G.; Tonge, P.J. (2000) *Biochemistry*, 39, 7645-7650.

Payne, D.J.; Warren, P.V.; Holmes, D.J.; Ji, Y.; Lonsdale, J.T. (2001) *Drug Discovery Today*, 6, 537-544.

Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. (2004) *Journal of Computational Chemistry,* 25, 1605-1612.

Quémard, A.; Sacchettini, J.C.; Dessen, A.; Vilchèze, C.; Bittman, R.; Jacobs, W.R. Jr.; Blanchard, J.S. (1995) *Biochemistry*, 34, 8235-8241.

Quémard, A.; Dessen, A.; Sugantino, M.; Jacobs, W.R.Jr.; Sacchettini, J.C. Blanchard, J.S. (1996) *Journal of the American Chemical Society*, 118, 1561-1562.

Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. (1999) *Journal of Molecular Biology*, 261, 470-489.

Rozwarski, D.A.; Vilchèse, C.; Sugantino, M.; Sacchettini, J.C., (1999) *Journal of Biological Chemistry*, 274, 15582-15589.

Rawat, R.; Whitty, A.; Tonge, P. (2003) *Proceedings of the National Academy of Sciences*, 100, 13881-13886.

Rock, C.O.; Cronan, J.E. (1996*) Biochimica et Biophysica Acta* 1302, 1-16.

Rossmann, M.G.; Liljas, A.; Branden, C.I.; Banaszak, L.J. (1975) *Enzymes*, 11, 61-102.

Rozwarski, D.A.; Grant, G.A.; Barton, D.H.; Jacobs, W.R. Jr.; Sacchettini, J.C. (1998) *Science*, 279, 98-102.

Rozwarski, D.A.; Vilchèse, C.; Sugantino, M.; Sacchettini, J.C., (1999) *Journal of Biological Chemistry*, 274, 15582-15589.

Salman, M.; Brennan, P.J.; Lonsdale, J.T. (1999) *Biochemica et Biophysica Acta*, 1437, 325-332.

Sander, T.; Freyss J.; von Korff, M.; Reich J.R.; Rufen, C. (2009) *Journal of Chemical Information and Modeling*, 49, 232-246.

Schaeffer, M.L.; Agnihotri, G.; Volker, C.; Kallender, H.; Brennan, P.J.; Lonsdale, J.T. (2001) *Journal of Biological Chemistry*, 276, 47029-47037.

Schroeder, E.K.; de Souza, O.N.; Santos, D.S.; Blanchard, J.S.; Basso, L.A. (2002) *Current Pharmaceutical Biotechnology*, 3, 197-225.

Slayden, R.A.; Barry III, C.E. (2000) *Molecular Microbiology*, 38, 514-525.

Sotriffer, C.A.; Flader, W.; Winger, R.H.; Rode, B.M.; Liedl, K.R.; Varga, J.M. (2000) *Methods*, 20, 280-291.

Subba Rao, G.; Vijayakrishnan, R.; Kumar, M. (2008) *Chemical Biology and Drug Design*, 72, 444-449.

Sullivan, T.J.; Truglio, J.J.; Boyne, M.E.; Novichenok, P.; Zhang, X.; Stratton, C.F.; Li H.J.; Kaur, T.; Amin, A.; Johnson, F.; Slayden, R.A.; Kisker, C.; Tonge, P.J. (2006) *ACS Chemical Biology*., 1, 43-53.

UNITY Chemical Information Software, version 4.1; Tripos Inc.: St. Louis, MO, 2006.

Velayati, A.A.; Farnia, P.; Masjedi, M.R.; Ibrahim, T.A.; Tabarsi, P.; Haroun, R.Z.; Kuan, H.O.; Ghanavi, P.; Farnia, P.; Varahram, M. (2009) *European Respiratory Journal*, 34, 1202-1203

Vieth, M.; Hirst, J.D.; Dominy, B.N.; Daigler, H.; Brooks, C.L.III. (1998) *Journal of Computational Chemistry*, 19, 1623-1631.

Vilchèze, C.; Morbidoni, H.R.; Weisbrod, T.R.; Iwamoto, H.; Kuo, M.; Sacchettini, J.S.; Jacobs Jr., W.R. (2000) *Journal of Bacteriology*, 182, 4059-4067.

Vilchèze, C.; Wang, F.; Arai, M.; Hazbón, M.H.; Colangeli, R.; Kremer, L.; Weisbrod, T.R.; Alland, D.; Sacchettini, J.C.; Jacobs, W.R. Jr. (2006) *Nature Medicine*, 12, 1027-1029.

Wakil, S.J.; Stoops, J.K.; Joshi, V.C. (1983) *Annual Review of Biochemistry,* 52, 537-579.

Wallace, A.C.; Laskowski, R.A.; Thornton J.M. (1995) *Protein Engineering*, 8, 127-134.

Walters W.P.; Stahl M.T.; Murcko, M.A. (1998) *Drug Discovery Today*, 3, 160-178.

Wang L.Q.; Falany, C.N.; James, M.O. (2004) *Drug Metabolism and Disposition*, 32, 1162-1169.

Waszkowycz, B. (2002) *Curr Opin Drug Discov Devel*, 5, 407-413.

Weiner, S.J.; Kollman, P.A.; Case, D.A.; Singh, U.C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. (1984). *Journal of the American Chemical Society*, 106, 765-784.

Wermuth, C.G.; Ganellin, C.R.; Lindberg, P.; Mitscher L.A. (1998), *Pure Applied Chemistry*, *70*, 1129-1143.

White, S.W.; Zheng, J.; Zhang, Y.M.; Rock, C.O. (2005) *Annual Reviews of Biochemistry*, 74, 791-831.

WHO report on the tuberculosis epidemic, 1995. Stop TB at the source. Geneva: WHO, 1995. Report No WHO/TB/95.1

WHO. (2006) World Health Organization. Factsheet on tuberculosis.

WHO report 2009 (Publication no. WHO/HTM/TB/2009.411.).

Zhang, Y.M.; Lu, Y.J.; Rock, C.O. (2004) *Lipids*, 39, 1055-1060.

Zhao, S.; Morris, G.M.; Olson, A.J.; Goodsell, D.S. (2001) *Journal of Molecular Biology*, 314, 1245-1255.

Anexo 1

Artigo submetido para a revista "*Journal of Molecular Modeling*"

"Conformational changes in 2-*trans*-enoyl-ACP (CoA) Reductase (InhA) from *Mycobacterium tuberculosis* induced by an inorganic complex: a molecular dynamics simulation study"

Conformational changes in 2-*trans*-enoyl-ACP (CoA) Reductase (InhA) from *M. tuberculosis* induced by an inorganic complex: a molecular dynamics simulation study

André L. P. da Costa [1,2**], Ivani Pauli[1,2,3**], Márcio Dorn[1,§], Evelyn K. Schroeder[1], Chang-Guo Zhan[4], and Osmar Norberto de Souza[1,2,3*]

[1] LABIO - Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas. PPGCC, Faculdade de Informática, PUCRS, Av. Ipiranga, 6681 – Prédio 32, sala 602, 90619-900, Porto Alegre, RS, Brasil

[2] Programa de Pós-Graduação em Biologia Celular e Molecular, Faculdade de Biociências, PUCRS, Av. Ipiranga, 6681 – Prédio 12, Bloco A, Sala 204, 90619-900, Porto Alegre, RS, Brasil

[3]INCT-TB - Instituto Nacional de Ciência e Tecnologia em Tuberculose - Av. Ipiranga 6681, Tecnopuc, Prédio 92 A, 90619-900, Partenon, Porto Alegre, RS Brasil

[4]College of Pharmacy, University of Kentucky,741 South Limestone, BBSRB Building 353, Lexington, KY 40536, USA

[**] These authors contributed equally to the work.

[§]Current address: Instituto de Informática, Universidade Federal do Rio Grande do Sul – UFRGS, Av. Bento Gonçalves, 9500, Prédio 67, Sala 202, Porto Alegre, RS, Brasil.

*Corresponding author: Osmar Norberto de Souza. Address: LABIO - Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas. PPGCC, Faculdade de Informática, PUCRS, Av. Ipiranga, 6681 – Prédio 32, Sala 602, 90619-900, Porto Alegre, RS, Brasil.
Phone: +55-51-3320-3611 ext. 8608
Fax: +55-51-3320-3621

E-mail: osmar.norberto@pucrs.br

Abstract

InhA, the NADH-dependent 2-trans-enoyl-ACP reductase enzyme from *Mycobacterium tuberculosis* (MTB), is involved in the biosynthesis of mycolic acids, the hallmark of mycobacterial cell wall. InhA has been shown to be the primary target of isoniazid (INH), one of the oldest synthetic antitubercular drugs. INH is a prodrug which is biologically activated by the MTB catalase-peroxidase KatG enzyme. The activation reaction promotes the formation of an isonicotinyl-NAD adduct which inhibits the InhA enzyme, resulting in reduction of mycolic acid biosynthesis. As a result of rational drug design efforts to design alternative drugs capable of inhibiting MTB's InhA, the inorganic complex pentacyano(isoniazid)ferrate(II) (PIF) was developed. PIF inhibited both wild-type and INH-resistant Ile21Val mutants of InhA and this inactivation did not require activation by KatG. Since no three-dimensional structure of the InhA-PIF complex is available to confirm the binding mode and to assess the molecular interactions with the protein active site residues. Here we report the results of molecular dynamics simulations of PIF interaction with InhA. We found that PIF strongly interacts with InhA and that these interactions lead to macromolecular instabilities reflected in the long time necessary for simulation convergence. These instabilities were mainly due to perturbation of the substrate binding loop, particularly the partial denaturation of helices α6 and α7. We were also able correlate the changes in the SASAs of Trp residues with the recent spectrofluorimetric investigation of the InhA-PIF complex and confirm their suggestion that the changes in fluorescence are due to InhA conformational changes upon PIF binding. The InhA-PIF association is very strong in the first 20.0 ns, but becomes very week at the end of the simulation, suggesting that the PIF binding mode we simulated may not reflect that of the actual InhA-PIF complex.

Keywords: Molecular Dynamics Simulations – Mycobacterium tuberculosis – FAS-II Pathway – InhA – Pentacyano(isoniazid)ferrate(II) – Conformational Changes.

1 Introduction

The 2-*trans*-enoyl-ACP (acyl carrier protein) reductase enzyme (InhA or ENR, EC number: 1.3.1.9) from *Mycobacterium tuberculosis* (MTB) is a member of type II dissociated fatty acid biosynthesis system (FAS-II) in MTB [1, 2, 3, 4]. This pathway, consisting of monofunctional enzymes and ACP [1], elongates FAS-I acyl fatty acid precursors to produce the long carbon chains (50-60 carbons) [2,5] of the meromycolate branch of mycolic acids, the hallmark of mycobacterial cell wall [6]. InhA, a NADH-dependent reductase, has specificity for long-chain substrates (12-24 carbons) [7], consistent with its involvement in mycolic acid biosynthesis [8]. InhA has been shown to be the primary target of isoniazid [9]. Isoniazid (INH, isonicotinic acid hydrazide), one of the oldest synthetic antitubercular drugs [10], is a prodrug [11] which is biologically activated by the MTB catalase-peroxidase KatG enzyme [12]. The activation reaction promotes the formation of an isonicotinyl-NAD adduct which inhibits the InhA enzyme, resulting in reduction of mycolic acid biosynthesis [4, 5]. Drugs such as INH, ethionamide (ETH), and pyrazinamide (PZA) require activation for activity against MTB. Interestingly, both activated forms of INH and ETH target the InhA enzyme, despite their different activation processes [13-15]. Based on the mechanism of activation proposed for INH, via electron transfer reaction [8, 16-18], an alternative to the self-activation route was proposed for the design of new drugs for the treatment of wild-type (WT) and INH-resistant tuberculosis, through the nonenzymatic INH activation method that

mimic the isonicotinyl-NAD adduct [19]. Efforts to reproduce this mechanism of activation are in progress and new small molecule compounds have been suggested, for example: SQ109 diamine-based [20], alkyl diphenyl ethers [21], pyrrolidine carboxamide analogues [22], and others [23-25]. These compounds are not expected to be activated prior to interacting with its cellular target. Within this approach Basso, Moreira, Santos and collaborators [26-29] have proposed an INH analog, namely pentacyano (isoniazid) ferrate II, that contains a cyanoferrate moiety, a metal center and the INH ligand [26]. This class of compounds constitutes a suitable model system for new perspectives of novel drug development for the treatment of MDR-TB [27].

The small molecule pentacyano (isoniazid) ferrate II (PIF) is the result of a rational drug design effort to find alternative drugs capable of inhibiting InhA [4, 5, 26]. PIF was found to inhibit both WT and INH-resistant Ile21Val mutants of InhA and this inactivation does not require activation by KatG [26]. Since crystal structures of the InhA-PIF binary complex are not available, we performed computational docking studies to predict the binding mode(s) of PIF in the InhA active site [27]. For that we used two crystal structures of InhA: the binary complex with the NADH coenzyme (PDB ID: 1ENY) [30] and the ternary complex with NAD+ and a substrate analog (PDB ID: 1BVR) [31]. We found that PIF preferentially occupies the pyrophosphate and nicotinamide sites in the NAD(H) binding pocket. However, we could not unambiguously assign a unique binding mode due to the distinct InhA active site conformations from the different PDB structures. We concluded that the flexibility of both, enzyme and inhibitor, should be taken into account to properly evaluate their interactions and to conform to the mechanism of slow-binding inhibition proposed for PIF based on WT InhA kinetic studies [27]. In a previous molecular dynamics (MD) simulation study of the NADH interaction with WT InhA and the mutants Ile21Val and Ile16Thr [31], both resistant to INH, we showed that InhA is a considerably flexible enzyme, capable of undergoing the conformational changes necessary to accommodate either substrate and inhibitor in an effective manner. Furthermore, we demonstrated that the mutations lead to conformational changes that reduced the affinity of the InhA-NADH complex. These results were soon after confirmed by X-ray crystallographic studies [32]. A recent characterization of PIF binding to MTB's InhA, using spectrofluorimetric techniques [28], hinted at the possibility that the quenching in protein fluorescence upon ligand binding, reported by the tryptophan amino acid (Trp) fluorescence, is due to conformational changes in the protein as previously suggested [27]. The identification of this enzyme conformational changes requires a considerable experimental effort, highlighting the practical value of computer simulations in their prediction. Consequently there is current interest in the prediction of the three-dimensional (3-D) nature of InhA-PIF specificity, how the enzyme binds to the inorganic complex, which conformational changes takes place upon binding, the effect of these changes on the Trp residues solvent accessibility, and which amino acid residues are responsible for PIF binding in the enzyme active site. We address these issues using computational methods, including automated molecular docking and MD simulations. Classical MD simulations make possible the detailed analysis of the individual movements of the atoms in the molecules as a function of time, resulting in an ideal model for understanding the atomic and molecular mechanisms involved in the formation of non-covalent enzyme-ligand complex [33-36].

2 Methods

2.1 The initial structure of the complex

We obtained the initial structure of WT InhA from the first crystal structure of the InhA-NADH binary complex (PDB ID: 1ENY) determined at 2.2 Å resolution [30]. The optimized structure of PIF and the Cartesian coordinates (x, y, z) of the InhA-PIF complex were taken from our previous molecular docking work, with the initial position of PIF in the active site of InhA taken from Oliveira *et al.* [27] (cluster 4 of Table 1 and figure 11e in reference 27). This cluster was top ranked and contained only one docked conformation. A simple translation and/or rotation could easily reproduce any of the other three top ranked clusters [27].

Although size exclusion chromatography analysis demonstrates that MTB's InhA biologically active structure is a homo-tetramer in solution, [37], each monomer binding cavity works independently. This is due to the fact that the active sites of the four monomers are about 40.0 Å apart from each other and are facing opposite sides in the quaternary structure. Hence, in this work, for all modeling and simulations, we use the InhA monomer (PDB ID: 1ENY). The all-atom model of the apo InhA enzyme contains 4,009 atoms with a net charge of -3 since His120 is protonated.

2.1.1 Force field and charges for PIF

Pentacyano(isoniazid)ferrate(II) is a new molecule [26] with 28 atoms and unknown charges and force field parameters. The partial atomic charges of PIF were determined as described by Oliveira *et al.* [27], and its force field parameters empirically derived by comparison with similar small molecules [38, 39]. To test these parameters, we run MD simulations of the PIF molecule fully solvated with TIP3P [39] water molecules at a temperature of 298.16 K. The initial cell dimensions containing the solute were 30.401 x 29.770 x 27.680 $Å^3$ with PIF solvated by a layer of water molecules of at least 10.0 Å in all orthogonal directions [39]. The test confirmed the PIF force field parameters we developed are adequate to be used in the MD simulation of the InhA-PIF complex. It is important to point out that the inorganic Fe^{+2} atom in the PIF molecule is covalently attached to its cyanide groups (Fig. 1a).

2.1.2 MD simulations

The main MD simulation started from the initial structure of InhA with the docked PIF inhibitor. InhA in complex with PIF (InhA-PIF) contains 4,037 atoms and a net molecular charge of -6, considering His120 protonated in the crystal structure. Hence, six sodium ions (Na^+) were added to neutralize the net negative charge density of the complex, which was then immersed in an orthorhombic box containing a total of 10,502 TIP3P water molecules [39]. The simulation cell dimension was 77.725 x 73.328 x 77.345 $Å^3$ and the complex was solvated by a layer of water molecules of at least 10.0 Å in all orthogonal directions [40]. The simulation cell contained a total of 35,549 atoms. Energy minimization, equilibration and production phases of the MD simulations were performed as described earlier [40]. The simulation was computed in a NPT ensemble at 298.16 K with the Berendsen temperature coupling [41] and constant pressure of 1.0 atm, with isotropic molecule-based scaling [42]. The SHAKE algorithm [43] was applied, with a tolerance of 10^{-5} Å, to fix all bonds that

contained a hydrogen atom, allowing the use of a 2.0 fs time step in the integration of the equations of motion. No extra restraints were applied after the equilibration phase. Periodic boundary conditions were applied, with electrostatic interactions between non-bonded atoms evaluated by the particle-mesh Ewald (PME) method. The Lennard-Jones interactions were evaluated using a 9.0 Å atom-based cutoff [44]. Four independent molecular systems' simulations were generated. The first one consisted of a 70.0 ns simulation of the PIF molecule alone in a neutral, with three Na+ ions, aqueous solution. The apo InhA enzyme consisted of the second one and the third was composed by the binary complex InhA-PIF. For these two simulations, data were collected for 25.0 ns. From the InhA-PIF MD simulation we built the fourth simulation, named InhA-PIF(-). For this simulation we removed the PIF molecule and three Na+ counter-ions "on-the-fly" from the third system (InhA-PIF) at 10.0 ns. The InhA-PIF(-) system was allowed to relax for another 15.0 ns. "On-the-fly" [31] means that we modified the InhA-PIF molecular systems and continued the MD simulation without reassignment of velocities.
Snapshots were collected at every 0.5 ps for analysis. All MD simulations were performed with the SANDER module of AMBER9 [42] using the ff99SB force field [45]. The stability of the simulations were analyzed in terms of the convergence of energy components, secondary structure content, solvent accessible surface area (SASA), radius of gyration, and the root-mean-squared deviation (RMSD) [46] from the initial, crystal structure (PDB ID: 1ENY). The tetramer structures for the InhA enzyme were generated using the symmetry operators available in the crystal structure (PDB ID: 1ENY) at the Protein Interfaces, Surfaces and Assemblies (PISA) web server of the EBI (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html) [47].

2.1.3 Analysis of the MD simulation trajectories

The MD simulation trajectories were visually monitored with the computer graphics software VMD [48]. Individual 3-D structures were further analyzed with Swiss-PdbViewer [49] and their illustrations prepared with the PyMOL Molecular Graphics System [50]. There are many different ways to evaluate the nature of intermolecular interactions or recognition, including making predictions of the estimated free energy of binding, involved in protein-ligand affinity. In this work intermolecular recognition is evaluated by analysing the total number of direct non-bonded interactions, i.e. hydrogen bonds (HB) and hydrophobic contacts (HC). Although waters play a major role in intermolecular recognition, here water-mediated H-bond interactions are not being considered (Rev 1, item 4; Rev 2, item 1). The total number of direct H-bonds in the InhA-PIF complex was calculated with the HBPLUS program [51] using a maximum donor-acceptor atoms distance of 3.4 Å and a minimum angle of 90.0º. We used the program LIGPLOT [52] for plotting H-bonds and HCs. PROMOTIF [53] was used to evaluate the InhA secondary structure pattern along the MD simulation trajectories. NACCESS [54] was used to calculate the SASA parameters of the Trp amino acid residues. The RMSD and the radius of gyration (Rgyr) were calculated with the Ptraj module of AMBER9 [42]. For all comparative analyses we used as reference the initial crystal structure (PDB ID: 1ENY). Graphics and statistical analyses were performed with Origin 7 Scientific Graphing and Analysis Software (Microcal Software, OriginLab, Northampton, MA). We also developed in house Python-based software to automate the analysis of the 270,000 snapshots generated by the four MD simulations described in this work. In the analyses, we adopted the PDB [55] numbering scheme of the reference structure with amino acid residues represented by a three-letter code.

3 Results and Discussion

3.1 Tests of the PIF parameters

To test the empirically derived PIF force field (FF) parameters, we first performed MD simulation in a water environment. No extraneous energy values and conformations were observed. Fig. 1B shows the RMSD of the coordinates of PIF in water with respect to the initial structure [27]. The all-atoms RMSD fluctuates between 0.3 Å and 2.5 Å along the 70.0 ns simulation. During this time the PIF molecule reversibly explores four different sets of conformers (Fig. 1b). These sets are populated with 39,591, 42,076, 28,677, and 29,656 molecules, respectively. Their RMSD averages 0.5 ± 0.3 Å, 1.3 ± 0.2 Å, 1.8 ± 0.2 Å, and 2.3 ± 0.2 Å, correspondingly. The observed fluctuations are due to PIF's intrinsic flexibility mainly due to the ability of its pyridine moiety to flip about the N6-Fe bond, and the torsions about the other two rotatable bonds of PIF (Fig. 1a). The rigid, pentacyanoferrate(II) moiety of PIF has stable RMSD values during the entire simulation, converging to an insignificant change of about 0.4 Å. Altogether, these results demonstrate the stability of the PIF molecule when free in an aqueous. They showed that its FF parameters are adequate for PIF to be further explored in other simulation studies. Hence, we simulated the aqueous InhA-PIF complex. In this simulation the PIF movements about the N6-Fe and the other two rotatable bonds are severely restricted by interactions with the amino acids residues in its InhA's binding site. As a result the RMSD of PIF in the complex fluctuates much less, ranging from 0.4 Å to 1.4 Å (Fig. 2) whilst the rather rigid pentacyanoferrate moiety RMSD values converged to 0.4 Å, as expected.

3.2 InhA conformational features

After the warm-up phase of the MD simulation, the first 120 ps, the enzyme backbone RMSD with respect to the initial structure of the apo InhA increases slowly and monotonically to a value close to 2.1 Å about which it remains the entire simulation time with very small fluctuations (Fig. 3). The RMSD for the InhA-PIF complex (Fig. 3) ranges from 0.7 to 2.9 Å with a median value of 2.1 Å. After the thermalization phase oscillates in the region of 1.7 Å remaining so for the next 2.0 ns. Then it rapidly decreases and stabilizes for another 1.0 ns. After that period the RMSD stably progresses to higher values reaching an average of 2.3 ± 0.1 Å in the last 5.0 ns (interval from 20.0 to 25.0 ns). Because this analysis involved InhA backbone atoms only, the observed drifting of the RMSD to higher values suggests that some degree of conformational change took place in the enzyme structure during the simulation of the complex. To test whether these changes were caused by the presence of the ligand, at 10.0 ns PIF and three Na^+ ions were removed from the complex. The remaining apoenzyme, named InhA-$PIF^{(-)}$, in the simulation box, was allowed to relax for another 15.0 ns. During this simulation the enzyme backbone RMSD reached a plateau around 2.5 ± 0.1 Å in the first 4.0 ns, followed by a gradual decrease over the remaining 11.0 ns, stabilizing at 2.4 ± 0.1 Å in the last 5.0 ns of the simulation. The apo InhA and the InhA-PIF simulations suggested that the conformational changes observed in the enzyme were in fact due to the formation of a tighter InhA-PIF complex during the first part of the simulation (0.0-10.0 ns of the InhA-PIF complex). It is interesting to notice that, even after 15.0 ns of simulation, the InhA-$PIF^{(-)}$ system did not recover from the changes caused by the interaction with PIF in the first 10.0 ns.

Inspection of other parameters such as the Rgyr and local secondary structure conservation, points to significant global and local changes in the InhA-PIF complex. The Rgyr is a measure of the protein dimension or compactness and, as the RMSD, is a measure of global structural changes in a protein. Fig.4 shows that, while for the apo InhA simulation Rgyr has an average of 17.9 ± 0.1 Å and fluctuates very close to the experimental, crystal structure value of 17.9 Å, in the InhA-PIF and InhA-PIF(-) Rgyr averages 18.3 ± 0.1 Å and 18.4 ± 0.1 Å, respectively, for the last 5.0 ns of their MD trajectories. While these values are statistically identical, they are different from that of the apo InhA. These differences become even clearer when we look at Fig. 4. From the InhA-PIF simulation, up to 10.0 ns, the Rgyr changes drastically.

In view of the fact that the RMSD as well as the Rgyr parameters measure global changes in the enzyme, it is possible that cooperation of local structural changes that led to the observed large deviations will take much longer than 15.0 ns to restore the RMSD values of the InhA-PIF(-) simulation to those of the apo InhA. Because Rgyr was calculated using the enzymes' backbone atoms, we hypothesize that these conformational changes could be the result of local disorder in regular (helices and sheets) and irregular (turns) secondary structure elements of the enzyme. These observations lead us to investigate the conservation of secondary structure in the InhA-PIF and InhA-PIF(-) systems.

To confirm this hypothesis and to detail the structural nature of the conformational changes, PROMOTIF [53] was used to measure the time dependence of InhA secondary structure (SS) content during the MD simulations (Fig. 5). These analyses suggest a considerable perturbation resulting in conversion of regular SS (α-helix mainly) in to irregular structures (coils) in the InhA-PIF complex. These alterations can be seen in the tertiary structure (Fig. 6). Comparing the MD snapshots at 10.0 ns (Fig. 6b) and 25.0 ns (Fig. 6c) with the initial enzyme structure (Fig. 6a) we notice that these structural changes mainly occur in the substrate binding loop [7], comprehending helices α6 and α7, and that the overall Rossmann fold [56-58] of the enzyme remained unchanged. This result is in agreement with the X-ray diffraction studies of Sullivan *et al.* [59]. These authors identified similar changes in MTB's InhA 3-D structure upon binding of alkyl diphenyl ethers or triclosan. Our observations are important as they support the hypothesis that PIF binding to InhA induces a conformational change [27,28] albeit, in our models, it is not directly interacting with the substrate binging site. Up to 20.0 ns PIF interacts with a site that overlaps with the NADH binding site (see details below). This observation is in agreement with the experimental results by Oliveira *et al.* [32] which imply that the inhibition mechanism of PIF may involve interaction to both the NADH coenzyme and the substrate binding sites [27]. However, after 20.0 ns and in the last 5,0 ns, PIF abruptly dissociates from the complex (see details below).

Experimental data on intrinsic InhA fluorescence indicated that PIF binding to the enzyme active site triggers a conformational change, inducing the formation of a more stable enzyme-inhibitor complex [28]. This observation is important since protein fluorescence usually is related to the solvent accessibility of Trp amino acid side-chains and this is a suitable method to study protein conformational changes and interactions with others molecules [60]. To understand and possibly correlate our data with the experimental fluorescence results of the InhA-PIF interaction [28], we identified all Trp residues in InhA and measured their solvent accessible surface area (SASA) in the initial structure and along the MD simulation.

3.3 Changes in SASA of tryptophan residues correlate with the fluorescence spectra of InhA-PIF complex

MTB's InhA enzyme has four Trp residues. Trp160 is located in the A-loop, Trp222 in helix α7, Trp230 is in the loop between helices α7 and α8, and Trp249 is located in a loop at the C-terminus. As we are simulating the InhA monomer, and the InhA biological unit is a tetramer [7], SASA for those Trp residues located in the tetramer interfaces were not realistic. In a single subunit these residues are overexposed to the solvent (Fig. 7). Examples are Trp249 and Trp160, located in the interfaces of the tetramer subunits. To partially overcome this problem, we built two InhA tetramers models as described in the Methods Section. One used the initial InhA structure (PDB ID: 1ENY) for reference purposes and the other the MD simulation snapshot at 25.0 ns. The total SASA for the Trp residues were calculated with NACCESS [54] and their values compared. Trp249 and Trp160, located internally in the tetramer, have a SASA remarkably reduced when compared to their values in the monomer ($Trp160_{monomer}$ = 53.02 Å^2 , $Trp160_{tetramer\ 0.0\ ns}$ = 4.62 Å^2, $Trp249_{monomer}$ = 189.91 Å^2, $Trp249_{tetramer\ 0.0\ ns}$ = 39.51 Å^2). The SASA of Trp222 and Trp230 are closer independently of the form of calculation, ($Trp222_{monomer}$ = 15.55 Å^2 , $Trp222_{tetramer\ 0.0\ ns}$ = 11.33 Å^2, $Trp230_{monomer}$ = 23.18 Å^2, $Trp230_{tetramer\ 0.0\ ns}$ = 17.72 Å^2). Hence, we conclude that Trp222 and Trp230 residues are the ones likely to be related to the changes in the InhA-PIF fluorescence [28]. Table 1 shows the Trp SASA in the tetramers in the initial structure (PDB ID: 1ENY) and at the end of the InhA-PIF simulation at 25.0 ns. The conformational changes affect the SASA of the Trp residues and consequently the fluorescence. After analyzing the data more closely we concluded that Trp222 is the one with the largest reduction in SASA when the tetramer's models are compared in the initial and final conformations of the InhA-PIF complex (Table 1). Fig. 8 shows the SASA variation of Trp222 and Trp230 along the 25.0 ns MD simulations. Although there is a tendency of the Trp230 SASA to increase during the dynamics simulation, it is evident that there is a reduction of the SASA for Trp222, which is located in helix α7 that, together with helix α6, forms the substrate binding loop. Most of the structural changes occur in this motif. Together, these data corroborate the hypothesis that PIF causes 3-D conformational changes in InhA, thus agreeing with the findings of Sullivan *et al.* [59], and explaining from a structural dynamics standpoint the experimental studies of Vasconcelos *et al.* [28].

3.4 InhA-PIF association

The initial structure of the InhA-PIF complex was taken from our previous automated molecular docking studies, with the initial position of PIF in the active site of InhA obtained from cluster 4 of table 1 and figure 11e from Oliveira *et al.* [27]. In this predicted complex the PIF ligand made 11 hydrophobic contacts (HCs) with InhA. To obtain further insight into the PIF-binding mechanism in InhA we explored their mode of interaction by atomistic MD simulations. Analysis of the InhA-PIF interactions along the simulation identified those residues important for PIF binding. We computed direct H-bonds as well as HCs with LIGPLOT [52]. Our analyses focused on the number and nature of the combined H-bonds and HCs interactions as described in the Methods Section. Estimations of free-energy of binding, and its corresponding enthalpic and entropic contribution, are out of the scope of this investigation. Tables 2 and 3 show the percentage of time H-bonds and HCs, respectively, lasted during the whole simulation. The values indicate that 13 InhA amino acids residues make H-bonds to PIF in at least one snapshot during the 25.0 ns simulation. Residues Ile15, Ser20, Phe41, Arg43 and Thr196 interact with PIF over 50% of the time. Hence, they are considered as the most important residues for InhA-PIF interaction (Table 2). For instance, at 10.0 ns (Fig. 6e) only H-bonds from Arg43 are missing. Furthermore, 25 residues make HCs with PIF, with Gly14, Ile15, Ile16, Ser20, Gly40, Phe41, Arg43, Ile47 and Thr196 being the

most important, interacting in more than 50% of the simulation time. These residues are located in the loop formed between helix α1 and strand β1 (Gly14, Ile15, Ile16, Ser20), in strand β2 (Gly40, Phe41), in the loop between strand β2 and helix α2 (Arg43), in helix α2 (Ile47), and in helix α6 (Thr196). Figs. 9 and 10 show the total number of InhA residues making H-bonds and HCs with PIF as a function of the simulation time. Approximately five residues make H-bonds interactions with PIF throughout almost the whole simulation time. Starting at about 20.0 ns this number drops sharply during the next 5.0 ns (Fig. 9), and at the end of the simulation there is only one residue (Thr196) making H-bond in the InhA-PIF complex (Fig. 6f).

The number of residues making HCs drops quickly from 17 to 7 in the first 10.0 ns (Fig. 10). Similarly to the dynamic behavior of H-bonding residues (Fig. 9), from 20.0 to 25.0 ns, the number HCs-making residues fell rapidly to a much lower value. At the end of the simulation time only three residues are making HCs to PIF in the InhA-PIF complex. This clearly indicates a dissociation of PIF from the initial InhA-PIF complex.

These analyses and visual inspection of the InhA-PIF MD simulation trajectory with VMD [48] showed that, after forming a tight binding InhA-PIF complex during the first 20.0 ns, PIF starts to dissociate from the complex.

All these data supports the notion that PIF first strongly interacts with the InhA enzyme, causing large conformational changes that mostly affect, by diminishing, the content of regular secondary structures . The fact that the residues involved in PIF interaction with InhA include those involved in InhA mutations related to resistance to INH and other anti-tuberculosis drugs that targets InhA might explain PIF's ability to inhibit both wild type and I21V INH-resistant InhA [27]. In addition to that, PIF dissociation observed at the end of the InhA-PIF complex simulation strongly suggests that, despite the strong InhA-PIF association at the beginning of the simulation, the initial position of PIF in the binding site adopted for the initial structure might not be the one expected for this ligand. This binding site overlaps with that of InhA innate NADH coenzyme. This is encouraging since because many important enzymes in humans contain the NAD(P)H coenzyme. If PIF does not compete with this site, it is likely to be a promising inhibitor candidate against MTB's InhA.

4. Conclusion

There is still no 3-D structure of the InhA-PIF binary complex to confirm the binding mode and to assess the molecular interactions PIF makes with the InhA active site residues which includes those that bind to the NADH coenzyme and those that make up de substrate binding cavity. While X-ray crystallography and Nuclear Magnetic Resonance (NMR) are powerful and the preferred techniques to determine the 3-D structure of enzymes and their complexes, it is not always possible to obtain single crystals or NMR solution of macromolecular complexes. Meanwhile, alternative, computational methods, such as molecular docking and fully solvated all-atom MD simulations of the wild-type InhA enzyme in complex with PIF, can provide comparable and useful information to help elucidate the molecular nature of InhA inhibition by the inorganic PIF ligand. We also performed MD simulations of apo InhA enzyme and the InhA enzyme ($InhA\text{-}PIF^{(-)}$) after PIF removal from the initial InhA-PIF complex. These two simulations served as control simulations. They showed that the results observed for the InhA-PIF complex simulation are not artifact of the method. Our approach for this work has provided important insights about the InhA-PIF interactions based on the analyses of the RMSD, Rgyr, SASA of Trp residues, H-bonds and HCs. We were capable of identifying all residues that interact with the inorganic PIF ligand along the simulation, the nature of these interactions, whether H-bonds and/or HCs, and details of the conformational

changes undergone by InhA upon PIF binding. The PIF compound appears to be a promising candidate to further antitubercular drug development and may represent a new class of lead compounds since it needs no activation by KatG or other enzyme, and furthermore, there is no need of the formation of any kind of adduct with the coenzyme NADH to bind to its molecular target, the *M. tuberculosis* InhA enzyme [26]. From this work, we can conclude that PIF strongly interacts with InhA and that these interactions leads to macromolecular instabilities reflected in the long time necessary for simulation convergence. The instability caused by the PIF interaction with InhA is mainly due to perturbation of the substrate binding loop, particularly the partial denaturation of helices α6 and α7. We also found that residues Gly14, Ile15, Ile16, Ser20, Gly40, Phe41, Arg43, Ile47, and Thr196 are responsible for the strong InhA-PIF association in the first 20.0 ns of the simulation. During this period, PIF directly competes for the NADH binding site. However, because at the end of the simulation PIF is almost completely dissociated from the InhA-PIF, we conclude that the mode of InhA-PIF interaction we simulated in this work may not reflect the actual InhA-PIF binding mode. It worth remembering that the initial position of PIF in the InhA-PIF complex investigated here overlapped with the NADH binding site. We have also been able to correlate the changes in the SASAs of the Trp residues with the recent spectrofluorimetric investigation of the InhA-PIF complex [27] and confirm their suggestion that the changes in fluorescence are due to InhA conformational changes upon PIF binding. Work underway in our laboratory is now being able to explore other possible InhA-PIF binding modes, as well as with other ligands. One possibility is to have PIF directly bound in the substrate binding cavity. As future work we will perform MD simulation of such complexes in order to provide knowledge and support to further the use of PIF as a lead compound to develop alternative treatment for tuberculosis using InhA as a target. Finally, we believe that this work constitutes a relevant contribution to the field of drug design and development with the use of molecular docking and MD simulations to help elucidate the binding mode of ligands, or prospective lead compounds, to their target protein receptor.

Acknowledgments
We would like to thank Luís Fernando Saraiva Macedo Timmers for his technical assistance. We thank the LAD-PUCRS for CPU time. This project was supported all or in part by grants from CAPES, FAPERGS and MCT-CNPq-DECIT, CNPq (Processes numbers 410505/2006-4 , 312027/2006-0, 559917/2010-4) to ONS and INCT-TB/CNPq to Prof. Diógenes Santiago Santos and PRONEX 2009/FAPERGS to Prof. Luiz Augusto Basso. ALPC was partially supported by a CAPES M.Sc. scholarship. IP was supported by a CAPES M.Sc scholarship. MD was supported by a CNPq M.Sc. scholarship. EKS was partially supported by a CAPES/PRODOC scholarship. ONS is a CNPq Research Fellow.

List of tables

Table 1. Trp residues' Solvent Accessible Surface Area (SASA) analysis. The Trp SASA (in $Å^2$) for each monomer in the tetramer built from the experimental initial protein structure (PDB ID: 1ENY) and from the snapshot at 25.0 ns.

Trp residue	Tetramer's SASA ($Å^2$) PDB ID: 1ENY	Tetramer's SASA ($Å^2$) for snapshot at 25.0 ns
160	4.62	12.42
222	11.33	0.02
230	17.72	28.43
249	39.51	93.7

Table 2. Hydrogen bonds analysis. Amino acids residues H-bonded to PIF along the 25.0 ns and during the last 5.0 ns of the MD simulation expressed as percentage (%) of time.

Amino acids that H-bonds to PIF	% of total time for 25.0 ns	% of total time for the last 5.0 ns
Gly14	47.97	19.64
Ile15	76.01	37.97
Ile16	0.05	0.24
Thr17	7.35	35.95
Ser19	0.85	4.24
Ser20	83.57	36.66
Phe41	76.48	45.10
Asp42	1.25	6.24
Arg43	56.84	35.82
Ser94	18.56	0.77
Gly96	0.24	0.09
Arg195	0.24	1.21
Thr196	73.58	87.21

Table 3. Hydrophobic contacts analysis. Amino acids residues making hydrophobic contacts to PIF along the 25.0 ns and during the last 5.0 ns of the MD simulation expressed as percentage (%) of time.

Amino acids making HCs to PIF	**% of total time for 25.0 ns**	**% of total time for the last 5.0 ns**
Gly14	83.94	40.71
Ile15	86.48	39.20
Ile16	98.53	94.39
Thr17	10.65	45.86
Ser19	2.56	12.73
Ser20	89.81	52.40
Ile21	14.73	0.06
Ala22	15.11	0.00
Gly40	72.47	30.81
Phe41	93.31	66.54
Asp42	10.69	24.94
Arg43	70.81	70.97
Leu46	0.02	0.12
Ile47	86.42	39.88
Ser94	48.69	20.93
Ile95	6.99	1.14
Gly96	4.75	3.40
Phe97	0.16	0.04
Gln100	0.01	0.00
Met103	0.11	0.53
Met147	1.16	0.00
Arg195	3.11	13.60
Thr196	82.84	92.89
Leu197	21.04	29.39
Ala198	13.83	0.43

Captions to figures

Fig.1. (a) Ball-and-stick model of the 3-D structure of pentacyano(isoniazid)ferrate(II) (PIF) molecule. The atoms at the ends of a rotatable bond are highlighted. The first one is between the nitrogen atom (N6) of the pyridine ring and the iron atom (Fe) of the pentacyanoferrate moiety. The second one is delimited by atoms C8 and C11 and the last one by C11 and N11. (b) The RMSD as a function of time for the entire PIF molecule (black line) and the pentacyanoferrate moiety (gray line). The four molecular structures at the right illustrates the each of the four sets of conformers adopted by PIF during the 70.0 ns MD simulation.

Fig. 2. Time dependence of the RMSD of (a) the entire PIF molecule in the binary InhA-PIF complex (black) and (b) the pentacyanoferrate moiety (gray). Comparing to Fig. 1b, PIF movement is severely restricted when in the complex with the InhA enzyme.

Fig. 3. RMSD of InhA backbone atoms (N, Cα, C, O) with respect to the initial, crystal structure (PDB ID: 1ENY) along the 25.0 ns MD simulation trajectory. The gray dashed line indicates the apo InhA simulation. The light gray represents the simulation of the InhA-PIF complex, and the black line indicates the InhA-$PIF^{(-)}$ simulation. The vertical dotted line at 10.0 ns indicates the time at which the PIF molecule was removed *on-the-fly* from the InhA-PIF complex. See the Methods Section for details.

Fig. 4. Radius of gyration (Rgyr) for the backbone atoms of InhA as a function of time. The vertical dashed line at 10.0 ns indicates the time at which the PIF molecule was removed *on-the-fly* from the InhA-PIF complex. The horizontal black line indicates the radius of gyration for the initial, crystal structure (PDB ID: 1ENY). The dark gray line indicates the apo InhA simulation, the light gray represents the simulation of the InhA-PIF complex while the black line indicates the InhA-$PIF^{(-)}$ simulation.

Fig. 5. InhA secondary structure content (in %) as a function of time along the MD simulations. The vertical dashed line at 10.0 ns indicates the time at which the PIF molecule was removed *on-the-fly* from the InhA-PIF complex. (a) Content of irregular structures and coils for the InhA-$PIF^{(-)}$ simulation. (b) The reference helical content based on the crystal structure (PDB ID: 1ENY). (c) Content of irregular structures and coils for the InhA-PIF simulation. (d) The reference coil and irregular structures' content based on the crystal structure (PDB ID: 1ENY). (e) Content of α-helices for the InhA-PIF complex. (f) Content of α-helices for the InhA-$PIF^{(-)}$ simulation. (g) Content of β-sheet for the InhA-PIF simulation. (h) The reference β-sheet content based on the crystal structure (PDB ID: 1ENY). (i) Content of β-sheet for InhA-$PIF^{(-)}$ simulation.

Fig. 6. Ribbon representation of the tertiary structure of the InhA-PIF complex (top) and their interactions calculated with HBPLUS and illustrated with LIGPLOT (bottom). (a) and (d) represent the initial simulation structure of the InhA-PIF complex; (b) and (e) the snapshot at 10.0 ns; and (c) and (f) the snapshot at 25.0 ns. The InhA motifs interacting with the inorganic PIF molecule (in dark blue stick model) are highlighted in different colors. The substrate binding loop (helices α6 and α7) is sand color, the A-loop is magenta, and the B-loop is cyan. At 10.0 ns we can notice a significant perturbation of the substrate binding loop and some disorder in helix α2, maintained until the end of the 25.0 ns simulation. In the initial structure PIF makes hydrophobic contacts and H-bond to nine and one residues of InhA, respectively. At the end of 10.0 ns PIF makes hydrophobic contacts and H-bonds to six and four residues of

InhA, respectively. Finally, at 25.0 ns, only three residues makes hydrophobic contacts and only one makes a H-bond.

Fig. 7. Molecular surface representation of the InhA tetramer built from the initial structure (PDB ID: 1ENY). The ribbon representation of the monomers is white colored. The Trp residues are represented by van der Waals spheres colored in red.

Fig. 8. The solvent accessible surface area (SASA) for Trp222 (black) and Trp230 (gray) as a function of simulation time for the three simulations: apo InhA, InhA-PIF complex and for InhA-PIF$^{(-)}$. The horizontal lines show the SASA values in reference tetramer for the initial structure. The vertical dashed line at 10.0 ns in the InhA-PIF$^{(-)}$ simulation indicates the time at which the PIF molecule was removed *on-the-fly* from the InhA-PIF complex.

Fig. 9. Number of residues making H-bonds to PIF in the InhA-PIF complex as a function of time. The solid line represents smoothing of the data to facilitate the visualization.

Fig. 10. Number of residues making HCs to InhA-PIF as a function of time. The solid line represents smoothing of the data to facilitate the visualization.

References

1. M.L. Schaeffer, G. Agnihotri, C. Volke, H. Kallender, B.J.Brennan, J.T.I. Lonsdale, Purification and biochemical characterization of the Mycobacterium tuberculosis beta - Ketoacyl-Acyl CarrierProtein Synthases KasA and KasB, J. Biol. Chem., 276, 47029-47037 (2001).
2. C. Ratledge, The biology of Mycobacteria. Academic Press, San Diego (1982).
3. K. Takayama, and N.Qureshi, The Mycobacteria: a Sourcebook, G.P. Kubica and L.G.Wayne Eds., Dekker, New York (1982).
4. E.K. Schroeder, O. Norberto de Souza, D.S. Santos, J.S. Blanchard, and L.A. Basso, Drugs that inhibit mycolic acid biosynthesis in Mycobacterium tuberculosis, Curr. Pharm. Biotech., 3, 197-225 (2002).
5. K. Mdluli, R.A. Slayden, Y. Zhu, S. Ramaswamy, X. Pan, D. Mead, D.D. Crane, J.M. Musser, and C.E. Barry III, Inhibition of a Mycobacterium tuberculosis beta-Ketoacyl ACP Synthase by Isoniazid, Science, 280, 1607-1610 (1998).
6. D. Chatterjee, The mycobacterial cell wall: structure, biosynthesis and sites of drug action, Curr. Opin. Chem. Biol., 4, 579-588 (1997).
7. A. Quémard, J.C. Sacchettini, A. Dessen, C. Vilcheze, l.R. Bittman, W.R. Jacobs Jr., and J.S. Blanchard, Enzymic characterization of the target for Isoniazid in Mycobacterium tuberculosis, Biochemistry, 34, 8235-8241 (1995).
8. A. Banerjee, E. Dubnau, A. Quemard, V. Balasubramanian, K.S. Um, T. Wilson, D. Collins, G. de Lisle, and W.R. Jacobs Jr., inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis, Science, 263, 227-230 (1994).
9. C. Lavender, M. Globan, A. Sievers, H.B. Jacobe, and J. Fyfe, Molecular characterization of Isoniazid-resistant Mycobacterium tuberculosis isolates collected in Australia, Antimicrob. Agents Chemother, 49, 4068-4074 (2005).
10. K. Johnsson, D.S. King, and P.G. Schultz, Studies on the mechanism of action of Isoniazid and Ethionamide in the chemotherapy of tuberculosis, J. Am. Chem. Soc., 117, 5009-5010 (1995).
11. Y. Zhang, B Heym, B. Allen, D. Young, and S. Cole, The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis, Nature, 358, 591-593 (1992).
12. L.V. Baker, T.J. Brown, O. Maxwell, A.L. Gibson, Z. Fang, M.D. Yates, and F.A. Drobniewski, Molecular analysis of Isoniazid-resistant Mycobacterium tuberculosis isolates from England and Wales reveals the phylogenetic significance of the ahpC-46A polymorphism, Antimicrob. Agents Chemother, 49, 1455-1464 (2005).
13. Y. Zhang, M.J. Garcia, R. Lathigra, B. Allen, C. Moreno, J.D. van Embden, and D. Young, Alterations in the superoxide dismutase gene of an isoniazid-resistant strain of Mycobacterium tuberculosis, Infect. Immun., 60, 2160-2165 (1992).
14. A. Scorpio, and Y. Zhang, Mutations in pncA, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus, Nat. Med., 2, 662-667 (1996).
15. A.R. Baulard, J.C. Betts, J. Engohang-Ndong, S. Quan, R.A. McAdam, P.J. Brennan, C. Locht, and G.S. Besra, Activation of the pro-drug ethionamide is regulated in mycobacteria, J. Biol. Chem., 275, 28326–28331 (2000).
16. A.E. DeBarber, K. Mdluli, M. Bosman, L.G. Bekker, and C.E. Barry 3rd, Ethionamide activation and sensitivity in multidrug-resistant Mycobacterium tuberculosis, Proc. Natl. Acad. Sci. USA, 97, 9677-9682 (2000).
17. T.A. Vannelli, A. Dykman, and P.R. Ortiz de Montellano, The antituberculosis drug ethionamide is activated by a flavoprotein monooxygenase, J. Biol. Chem., 277, 12824-12829 (2002).

18. C. Vilcheze, F. Wang, M. Arai, M.H. Hazbón, R. Colangeli, L. Kremer, T.R. Weisbrod, D. Alland, J.C. Sacchettini, and W.R. Jacobs Jr., Transfer of a point mutation in Mycobacterium tuberculosis inhA resolves the target of isoniazid, Nat. Med., 12, 1027-1029 (2006).
19. M. Nguyena, A. Quemard, H. Marrakchi, J. Bernadou, and B. Meunier, The nonenzymatic activation of isoniazid by MnIII-pyrophosphate in the presence of NADH produces the inhibitionof the enoyl-ACP reductase InhA from Mycobacterium tuberculosis, Comptes Rendus del'Académie des Sciences - Series IIC - Chemistry 4, 35-40 (2001).
20. L. Jia, J.E. Tomaszewski, C. Hanrahan, L. Coward, P. Noker, G. Gorman, B. Nikonenko, and M. Protopopova, Pharmacodynamics and pharmacokinetics of SQ109, a new diamine-based antitubercular drug, Br. J. Pharmacol., 144, 80-87 (2005).
21. T.J. Sullivan, J.J. Truglio, M.E. Boyne, P. Novichenok, X. Zhang, C.F. Stratton, H.J. Li, T. Kaur, A. Amin, F. Johnson, R.A. Slayden, C. Kisker, and P.J. Tonge, High affinity InhA inhibitors with activity against drug-resistant strains of Mycobacterium tuberculosis, ACS Chem. Biol., 1, 43-53 (2006).
22. X. He, A. Alian, and P.R.Ortiz de Montellano, Inhibition of the Mycobacterium tuberculosis enoyl acyl carrier protein reductase InhA by arylamides, Bioorg. Med. Chem., 15, 6649-6658 (2007).
23. M.R. Kuo, H.R. Morbidoni, D. Alland, S.F. Sneddon, B.B. Gourlie, M.M. Staveski, M. Leonard, J.S. Gregory, A.D. Janjigian, C. Yee, J.M. Musser, B. Kreiswirth, H. Iwamoto, R. Perozzo, W.R. Jacobs Jr., J.C. Sacchettini, and D.A. Fidock, Targeting tuberculosis and malaria through inhibition of Enoyl Reductase: compound activity and structural data, J. Biol. Chem., 278, 20851-20859 (2003).
24. Y. Zhang, K. Post-Martens, and S. Denkin, New drug candidates and therapeutic targets for tuberculosis therapy, Drug Discov. Today, 11, 21-27 (2006).
25. F. Wang, R. Langley, G. Gulten, G. Dover Lynn, G.S. Besra, W.R. Jacobs Jr., and J.C. Sacchettini, Mechanism of thioamide drug action against tuberculosis and leprosy, J. Exp. Med., 204, 73-78 (2007).
26. J.S. Oliveira, E.H.S Souza, L.A. Basso, M. Palaci, R. Dietze, D.S. Santos, and I.S. Moreira, An inorganic iron complex that inhibits wild-type and an isoniazid-resistant mutant 2-trans-enoyl-ACP (CoA) reductase from Mycobacterium tuberculosis, Chem. Commun., 3, 312-313 (2004).
27. J.S. Oliveira, E.H.S Souza, O. Norberto de Souza, I.S. Moreira, D.S. Santos, and L.A. Basso, Slow-Onset Inhibition of 2-trans-Enoyl-ACP (CoA) Reductase from Mycobacterium tuberculosis by an inorganic complex, Curr. Pharm. Design, 12, 2409-2424 (2006).
28. I. Vasconcelos, E. Meyer, F.A.M. Sales, I.S. Moreira, and D.S. Santos, The mode of inhibition of Mycobacterium tuberculosis wild-type and Isoniazid-resistant 2-trans-Enoyl-ACP(CoA) Reductase enzymes by an inorganic complex, Anti-Infective Agents in Medicinal Chemistry, 7, 50-62 (2008).
29. L.A. Basso, C. Z. Schneider, A. J. A. B. dos Santos, A. A. dos Santos Jr., M.M. Campos, A.A. Souto and D.S. Santos, An Inorganic Complex that Inhibits Mycobacterium tuberculosis Enoyl Reductase as a Prototype of a New Class of Chemotherapeutic Agents to Treat Tuberculosis, J. Braz. Chem. Soc., 00, 1-6 (2010).
30. A. Dessen, A. Quémard, J.S. Blanchard, Jr WR Jacobs, and J.C.Sacchettini, Crystal structure and function of the isoniazid target of Mycobacterium tuberculosis, Science, 267, 1638-1641 (1995).
31. E.K. Schroeder, L.A. Basso, D.S. Santos, and O. Norberto de Souza, 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 (2005).

32. J.S. Oliveira, J.H. Pereira, F. Canduri, N.C. Rodrigues, O. Norberto de Souza, W.F. de Azevedo, L.A. Basso, and D.S. Santos, Crystallographic and pre-steady-state kinetics studies on binding of NADH to wild-type and Isoniazid-resistant Enoyl-ACP(CoA) Reductase enzymes from Mycobacterium tuberculosis, J. Mol. Biol., 359, 646-666 (2006).
33. T.E. Cheatham, and B.R. Brooks, Recent advances in molecular dynamics simulation towards the realistic representation of biomolecules in solution, Theor. Chem. Acc., 99, 279-288 (1998).
34. M. Karplus, and J.A. McCammon, Molecular dynamics simulations of biomolecules Nat. Struct. Biol., 9, 646-652 (2002).
35. M. Karplus, and J. Kuriyan, Chemical Theory and computation special feature: Molecular dynamics and protein function, Proc. Natl. Acad. Sci. USA, 102, 6679-6685 (2005).
36. Z. Nyarady, T. Czompoly, S. Bosze, G. Nagy, A. Petrohai, J. Pál, F. Hudecz, T. Berki, and P. Németh, Validation of in silico prediction by in vitro immunoserological results of fine epitope mapping on citrate synthase specific autoantibodies, Mol.Immunol., 43, 830-838 (2006).
37. A, Quémard, J.C. Sacchettini, A. Dessen, C. Vilchèze, R. Bittman, W. R. Jr. Jacobs, J.S. Blanchard, Enzymatic characterization of the target for isoniazid in Mycobacterium tuberculosis, Biochemistry, 34, 8235-8241(1995).
38. W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman, A Second generation Force Field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc., 117, 5179-5197 (1995).
39. W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys., 79, 926-935 (1983).
19
40. O. Norberto de Souza, and R.L. Ornstein, Effect of periodic box size on aqueous molecular dynamics simulation of a DNA dodecamer with particle-mesh Ewald method, Biophys. J., 72, 2395-2397 (1997).
41. H.J.C. Berendsen, J.P.M. Postma, W.F. vanGunsteren, A. Dinola, and J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys., 81, 3684-3690 (1984).
42. D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M Merz , D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C. Schafmeister, W.S. Ross, and P.A. Kollman, AMBER 9, University of California, San Francisco (2006).
43. J.P. Ryckaert, G. Ciccotti, and H.J.C.Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes J. Comput. Phys., 23, 327-341 (1977).
44. O. Norberto de Souza, and R.L. Ornstein, Molecular dynamics simulations of a proteinprotein dimmer: particle-mesh Ewald electrostatic model yields far superior results to standard cutoff model, J. Biomol. Struct. Dyn., 16, 1205-1218 (1999).
45. D.R. Roe, A. Okur, L. Wickstrom, V. Hornak, and C. Simmerling, Secondary structure bias in Generalized Born Solvent Models: Comparison of conformational ensembles and free energy of solvent polarization from explicit and implicit solvation, J Phys Chem B, 111, 1846-1857 (2007).
46. V.N. Maiorov, and G.M. Crippen, Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins, J. .Mol. Biol., 235, 625-634 (1994).

47. E. Krissinel and K. Henrick, Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. **372**, 774-797(2007).
48. W. Humphrey, A. Dalke, and K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph., 14, 33-38 (1996).
49. W. Kaplan, and T.G. Littlejohn, Swiss-PDB Viewer (Deep View), Brief Bioinform., 2, 195-197 (2001).
50. W.L. DeLano. The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA (2002).
51. I.K. McDonald, and J.M. Thornton, Satisfying hydrogen bonding potential in proteins, J. Mol. Biol., 238, 777-793 (1994).
52. A.C. Wallace, R.A Laskowski, and J.M. Thornton, LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions, Protein Eng., 8, 127-134 (1995).
53. G. Hutchinson, and J.M. Thornton, PROMOTIF-A program to identify and analyze structural motifs in proteins, Protein Sci., 5, 212-220 (1996).
54. S.J. Hubbard, and J.M. Thornton, NACCESS, Computer Program, Department of Biochemistry and Molecular Biology. University College London (1993).
55. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, The Protein Data Bank, Nucleic Acids Res., 28, 235-242 (2000).
56. M.G. Rossmann, A. Liljas, C.I. Branden, and L.J. Banaszak, Evolutionary and structural relationships among dehydrogenases. The Enzymes, 3rd ed., Boyer, P. D., Ed. Academic Press, New York, 61-102 (1975).
57. H. Jornvall, B. Persson, M. Krook, S. Atrian, R.D. Gonzalez, J. Jeffery, and D. Ghosh, Shortchain dehydrogenases/reductases (SDR), Biochemistry, 18, 6003-6013 (1995).
58. U. Oppermann, C. Filling, M. Hult, N. Shafqat, X. Wu, M. Lindh, J. Shafqat, E. Nordling, YKallberg, B. Persson, and H. Jörnvall, Short-chain dehydrogenase/reductases (SDR): the 2002 update, Chem. Biol. Interact., 143, 247-253 (2003).
59. T.J. Sullivan, J.J. Truglio, M.E. Boyne, P. Novichenok, X. Zhang, C.F. Stratton, H.J.Li, T. Kaur, A. Amin, F. Johnson, R.A. Slayden, C. Kisker, and P.J. Tonge, High affinity InhA inhibitors with activity against drug-resistant strains of Mycobacterium tuberculosis, ACS Chem. Biol., 1, 43-53 (2006).
60. Y. Chen, and M.D. Barkley, Toward understanding Tryptophan fluorescence in proteins, Biochemistry, 37, 9976-9982 (1998).

Figure 1

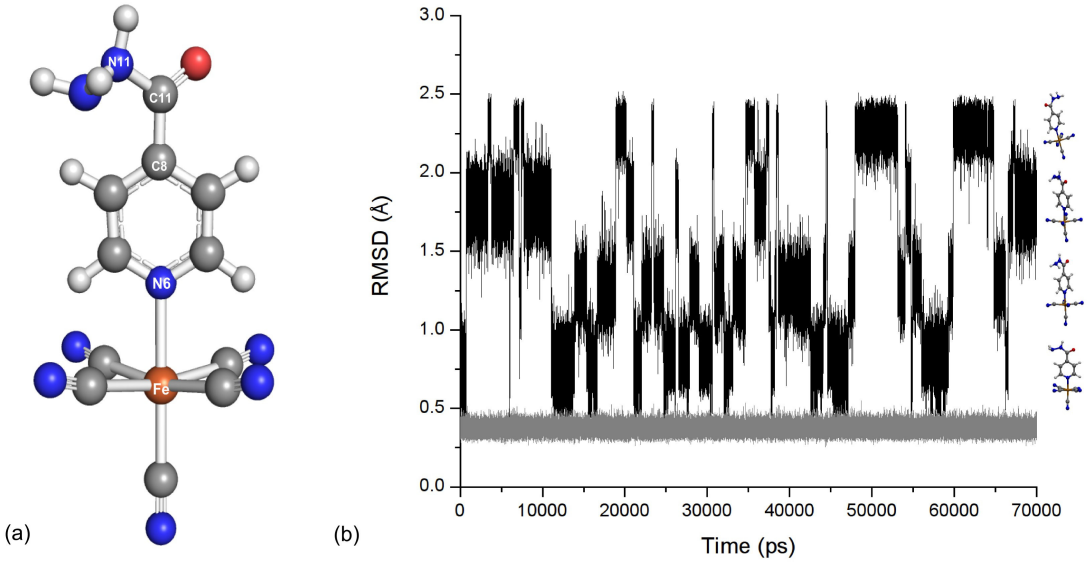


Figure 2

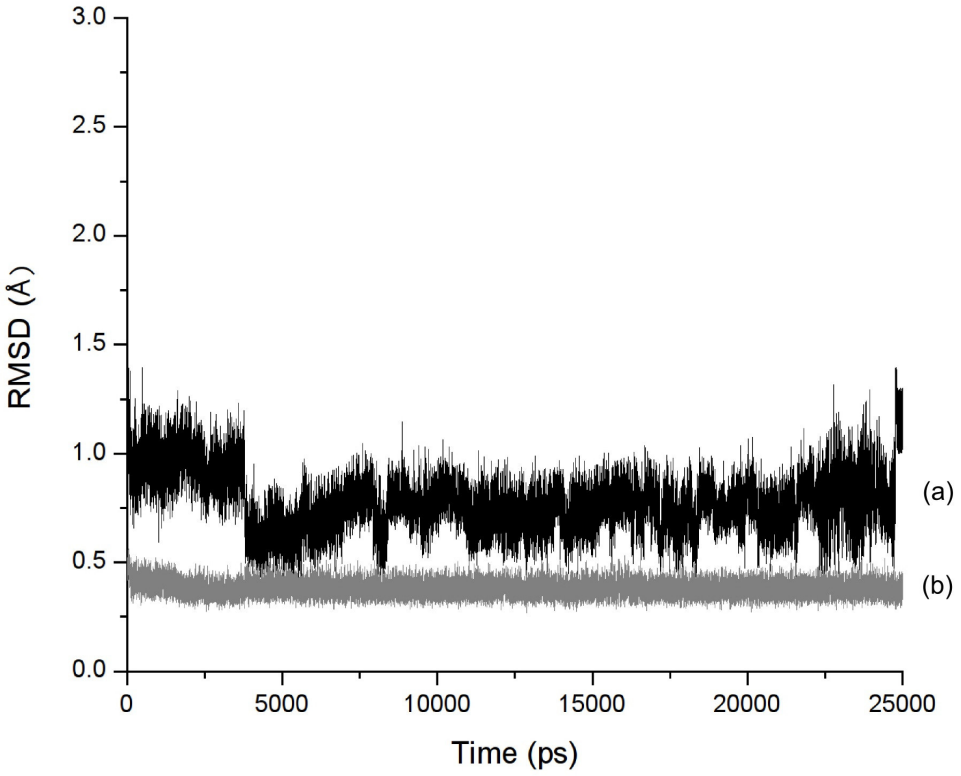


Figure 3

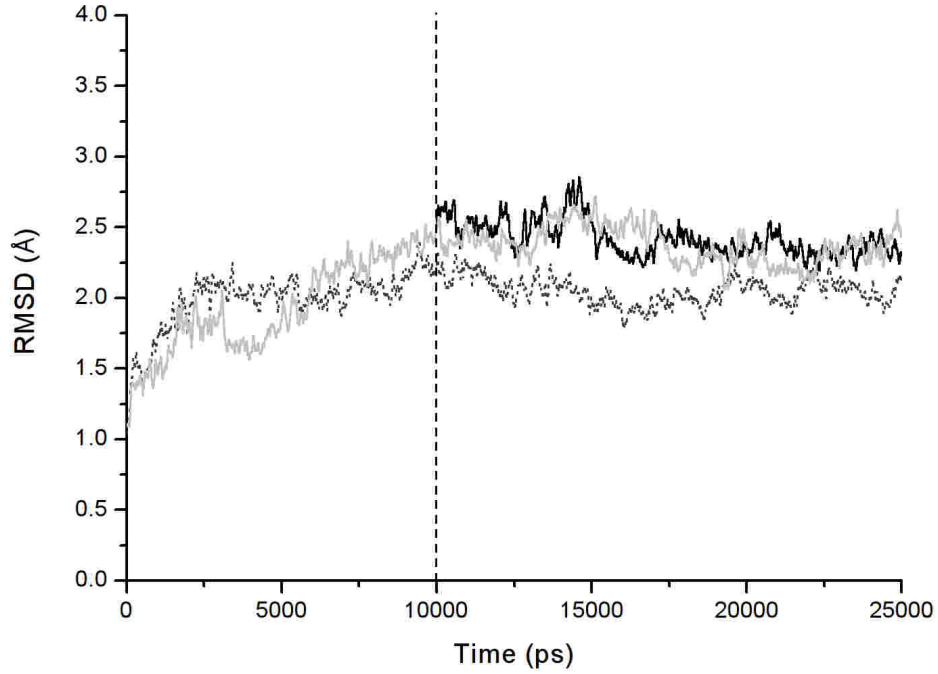


Figure 4

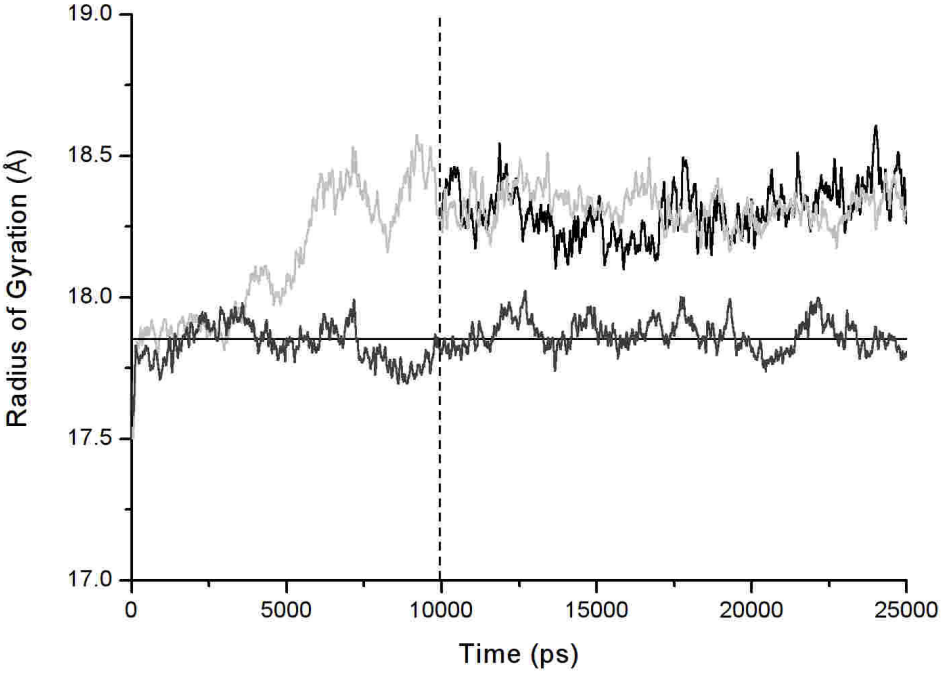


Figure 5

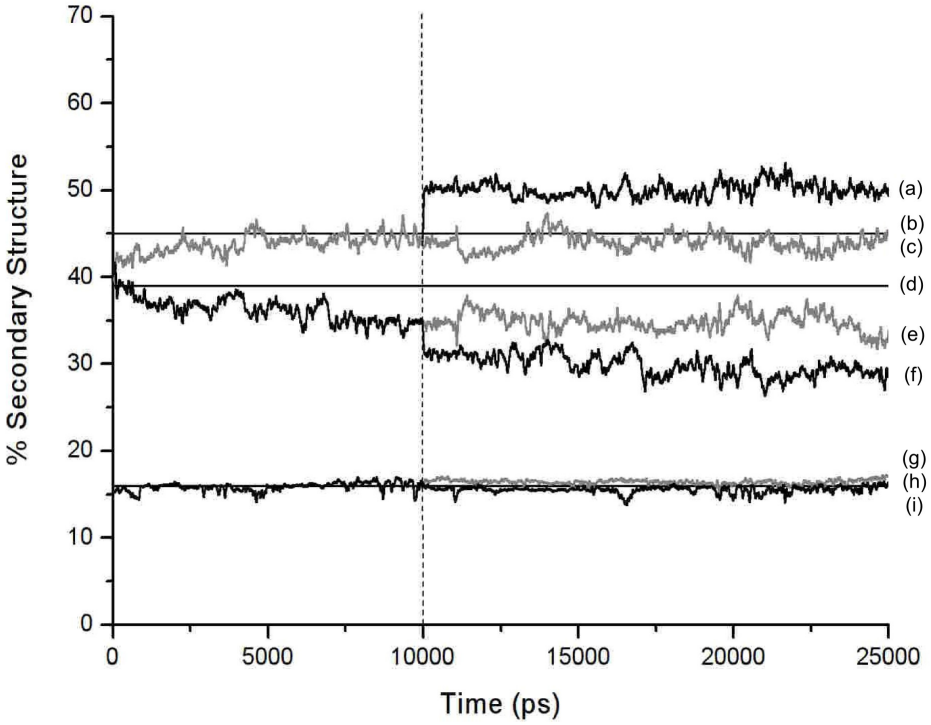

% Secondary Structure
Time (ps)
0
10
20
30
40
50
60
70
0
5000
10000
15000
20000
25000
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)

Figure 6

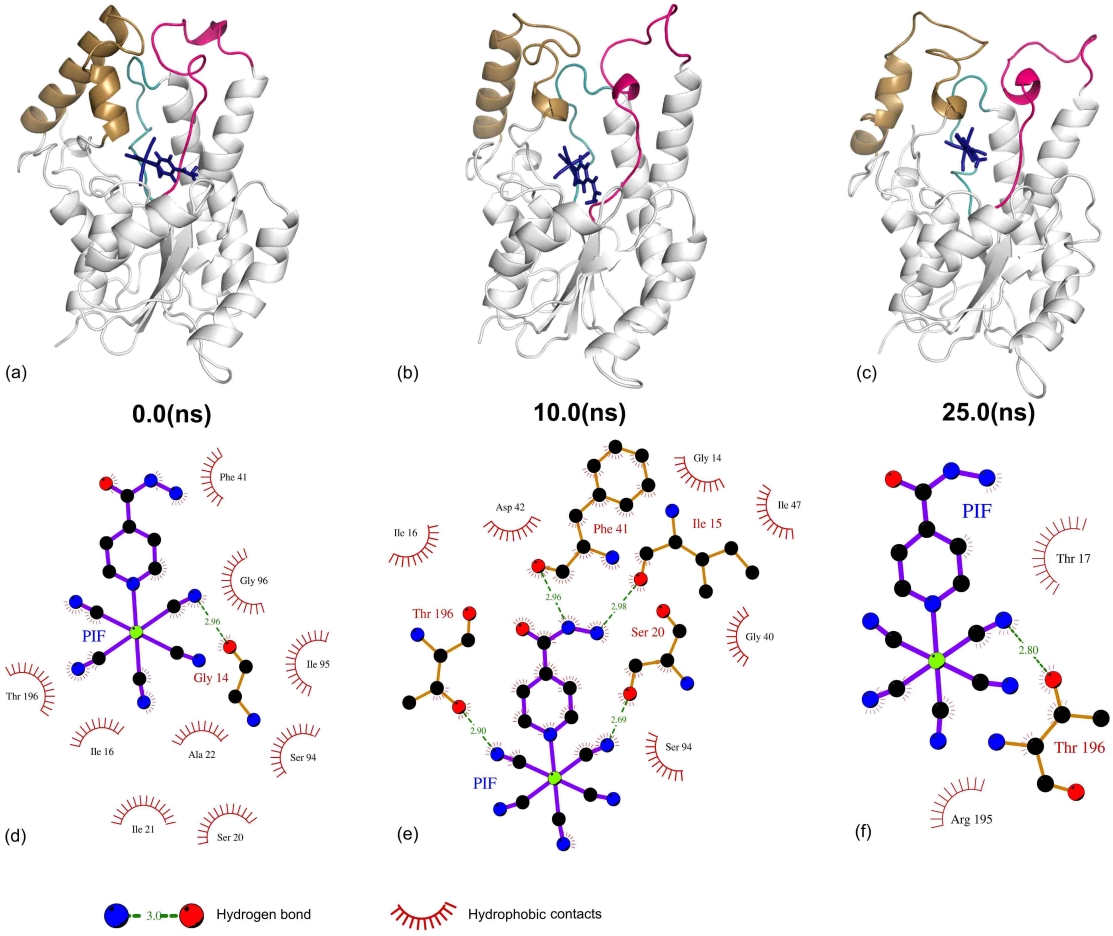

(a)
0.0(ns)
(b)
10.0(ns)
(c)
25.0(ns)
PIF
Phe 41
Gly 96
Gly 14
Ile 95
Thr 196
Ile 16
Ala 22
Ser 94
Ile 21
Ser 20
(d)
Asp 42
Gly 14
Ile 47
Ile 16
Phe 41
Ile 15
Thr 196
Ser 20
Gly 40
Ser 94
PIF
(e)
PIF
Thr 17
Thr 196
Arg 195
(f)
Hydrogen bond
Hydrophobic contacts

Figure 7

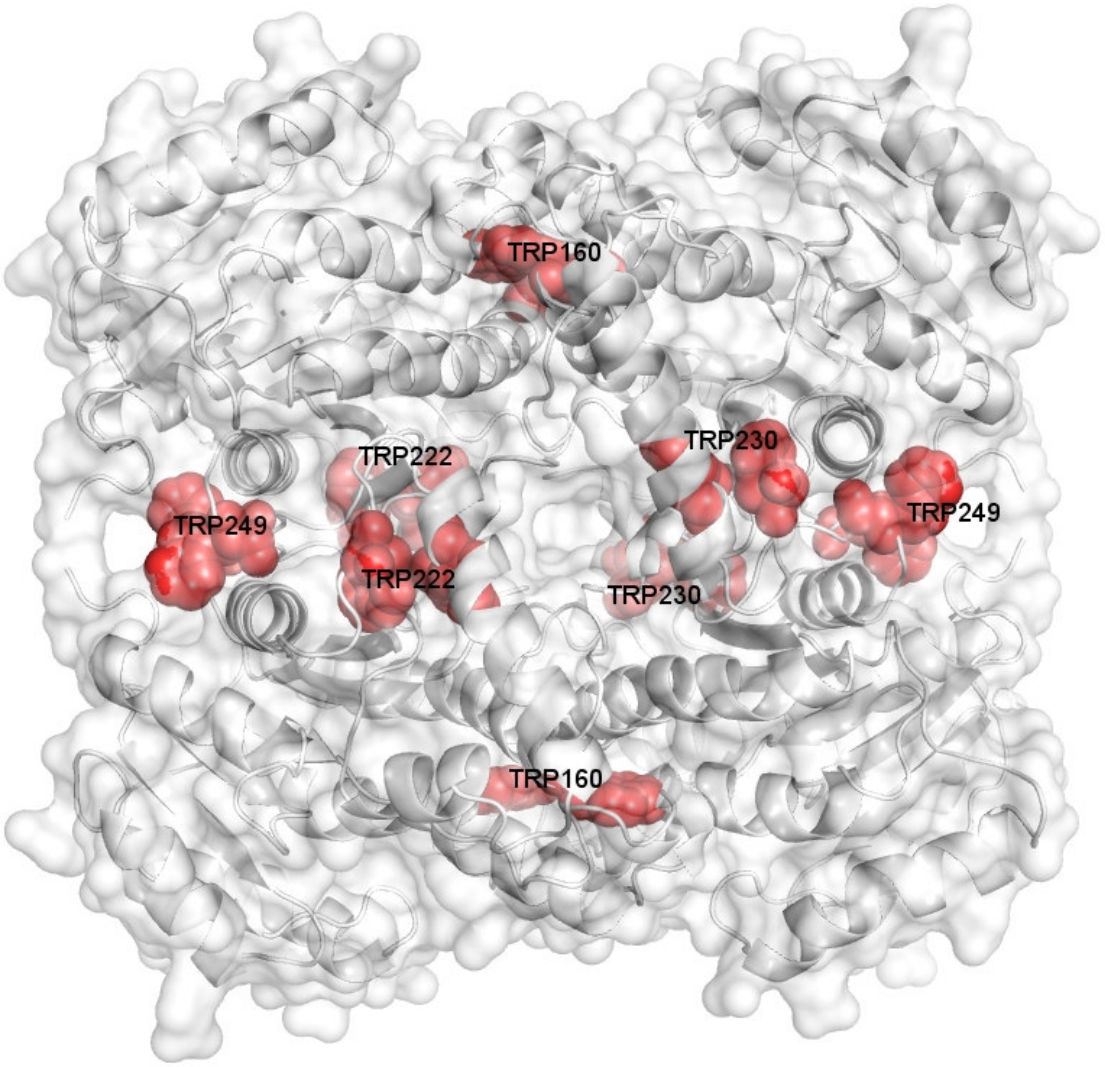


Figure 8

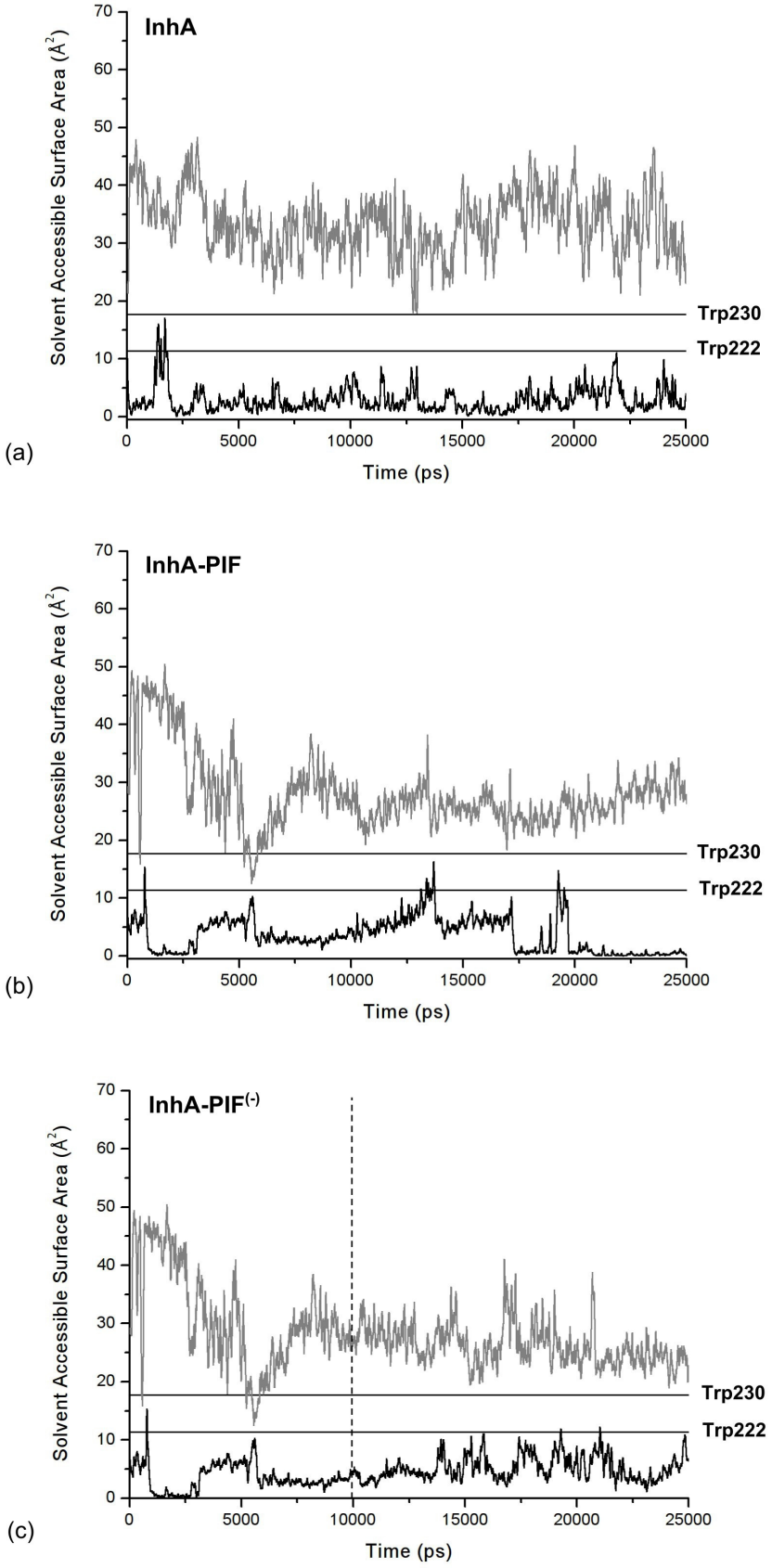

InhA
Solvent Accessible Surface Area (Å²)
Time (ps)
Trp230
Trp222
(a)
InhA-PIF
(b)
InhA-PIF(-)
(c)

Figure 9

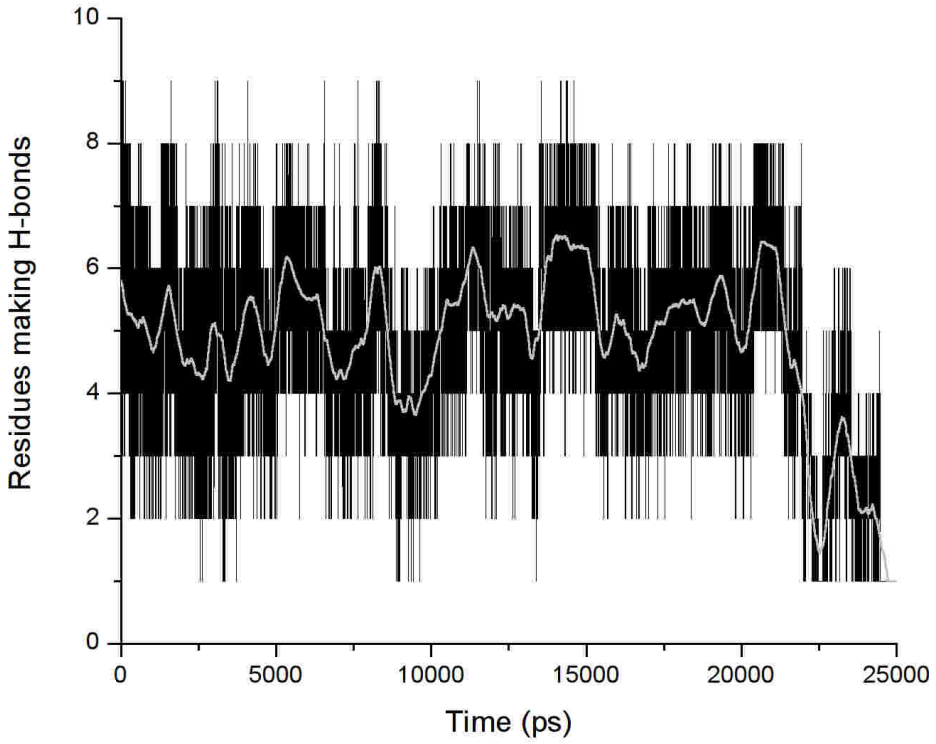
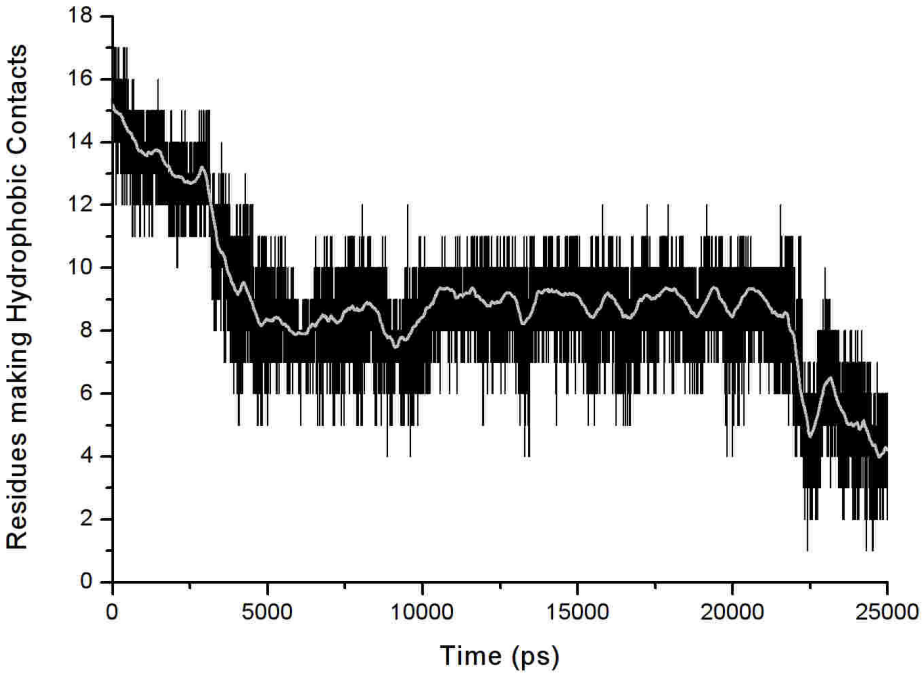

Residues making H-bonds
Time (ps)
0
2
4
6
8
10
0
5000
10000
15000
20000
25000

Figure 10


Residues making Hydrophobic Contacts
Time (ps)
0
2
4
6
8
10
12
14
16
18
0
5000
10000
15000
20000
25000