

ESCOLA DE BIOCÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR
MESTRADO EM BIOLOGIA CELULAR E MOLECULAR

RENATA FIORAVANTI TARABINI

**ESTUDO DA FLEXIBILIDADE DA ENZIMA INHA (EC 1.3.1.9) DE
MYCOBACTERIUM TUBERCULOSIS POR SIMULAÇÕES DE DINÂMICA
MOLECULAR**

Porto Alegre

2018

PÓS-GRADUAÇÃO - STRICTO SENSU



Pontifícia Universidade Católica
do Rio Grande do Sul



PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL

ESCOLA DE BIOCÊNCIAS

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

RENATA FIORAVANTI TARABINI

Estudo da flexibilidade da enzima InhA (EC 1.3.1.9) de *Mycobacterium tuberculosis* por simulações de dinâmica molecular

Porto Alegre

2018

RENATA FIORAVANTI TARABINI

Estudo da flexibilidade da enzima InhA (EC 1.3.1.9) de *Mycobacterium tuberculosis* por simulações de dinâmica molecular

Dissertação de Mestrado apresentada como requisito para a obtenção do grau de Mestre pelo Programa de Pós-Graduação em Biologia Celular e Molecular da Escola de Biociências da Pontifícia Universidade Católica do Rio Grande do Sul

Orientador: Osmar Norberto de Souza

Coorientador: Luís Fernando Saraiva Macedo Timmers

Porto Alegre

2018

Ficha Catalográfica

T176e Tarabini, Renata Fioravanti

Estudo da flexibilidade da enzima InhA (EC 1.3.1.9) de *Mycobacterium tuberculosis* por simulações de dinâmica molecular / Renata Fioravanti Tarabini . – 2018.

54 f.

Dissertação (Mestrado) – Programa de Pós-Graduação em Biologia Celular e Molecular, PUCRS.

Orientador: Prof. Dr. Osmar Norberto de Souza.

Co-orientador: Prof. Dr. Luís Fernando Saraiva Macedo Timmers.

1. *Mycobacterium tuberculosis*. 2. enzyme InhA. 3. flexibility. 4. Molecular Dynamics. I. Souza, Osmar Norberto de. II. Timmers, Luís Fernando Saraiva Macedo. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS
com os dados fornecidos pelo(a) autor(a).

Bibliotecária responsável: Salete Maria Sartori CRB-10/1363

Renata Fioravanti Tarabini

Estudo da flexibilidade da enzima InhA (EC 1.3.1.9) de *Mycobacterium tuberculosis* por simulações de dinâmica molecular

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Biologia Celular e Molecular, da Escola de Biociências da Pontifícia Universidade Católica do Rio Grande do Sul.

Aprovado em Março de 2018

BANCA EXAMINADORA

Prof. Dr. Luiz Augusto Basso

Prof. Dr. Rafael Andrade Caceres

Prof. Dr. Andre Arigony Souto

Porto Alegre

2018

RESUMO

A enzima InhA de *Mycobacterium tuberculosis*, ou 2-*trans*-enoil-ACP (CoA) redutase (EC 1.3.1.9) é codificada pelo gene *inhA*. Esta proteína realiza a quarta etapa da via de síntese de ácidos graxos tipo II (FAS II), a qual consiste de enzimas monofuncionais, responsáveis por alongar precursores de ácidos graxos produzindo longas cadeias de carbono do ramo meromicolato de ácidos micólicos. A InhA é responsável por converter 2-*trans*-enoil-ACP para acil-ACP por meio da transferência de um hidreto (H^-) do NADH para a posição C3 do substrato 2-*trans*-enoil-ACP, resultando na redução da ligação dupla α tio éster (entre C2 e C3) da porção lipídica do substrato. Esta reação pode ocorrer em cadeias carbônicas que variam de 16 a 56 carbonos, evidenciando a flexibilidade tanto da cavidade de ligação ao substrato quanto da estrutura desta enzima. A topologia desta enzima se assemelha a uma cadeira (chair-like) do tipo Rossmann, caracterizada por uma folha β contendo sete fitas e oito hélices α , numa arquitetura $\alpha\beta\alpha$ sanduíche, que forma o sítio de ligação da coenzima NADH e a cavidade de ligação do substrato. Essa cavidade, por sua vez, é formada por três alças estruturais: alças A e B e a alça de ligação ao substrato.

Essa composição estrutural é de fundamental importância para o funcionamento da InhA, pois, deve ser flexível o suficiente para permitir a adaptabilidade da InhA a substratos de tamanho variável. O estudo sobre a flexibilidade desta enzima é relevante para se ter o conhecimento de quais fatores podem estar envolvidos no processo de abertura e fechamento da sua cavidade de ligação.

Nesta dissertação é utilizado método computacional, por meio de simulações por Dinâmica Molecular, para avaliar a flexibilidade da cavidade de ligação do substrato da enzima InhA e analisar as respectivas mudanças conformacionais.

ABSTRACT

The enzyme InhA of *Mycobacterium tuberculosis*, or 2-*trans*-enoyl-ACP (CoA) reductase (EC 1.3.1.9) is encoded by the *inhA* gene. This protein performs the fourth step of the pathway of synthesis of type II fatty acid (FAS II), which consists of monofunctional enzymes responsible for fatty acid precursor's elongation producing long carbon chains of the meromycolate branch of mycolic acids. InhA converts 2-*trans*-enoyl-ACP to acyl-ACP by means of a hydride (H⁻) the transfer from NADH to the C3 position of the 2-*trans*-enoyl-ACP substrate, resulting in reduction of the double α -thioester bond (between C2 and C3) of the substrate's lipid portion. This reaction can occur in a carbon chain ranging from 16-56 carbons. This substrate sizes' range highlights the importance of the substrate binding cavity flexibility for InhA proper functioning. The topology of this enzyme is hair-like, characterized by a seven-strand β -sheet and eight α -helices, in a $\alpha\beta\alpha$ sandwich architecture, forming the NADH coenzyme-binding site and the substrate-binding cavity. This cavity, in turn, is formed by three structural loops: loops A and B, and the substrate-binding loop.

This structural composition is of fundamental importance for the functioning of InhA. It must be flexible enough to allow adaptation to substrates of varying sizes. Investigation of this flexibility is paramount for a better comprehension of which conformational features may be involved in the opening and closing events of its substrate-binding cavity.

In this work, a computational method is used, through Molecular Dynamics simulations, to evaluate the flexibility of the substrate-binding cavity of the InhA enzyme and analyze the respective conformational changes.

LISTA DE ILUSTRAÇÕES

Figura 1 - Estrutura monomérica da enzima InhA	14
Figura 2 - Região de reconhecimento do substrato da enzima InhA	15
Figura 3 - Mudanças conformacionais na cavidade de ligação da enzima InhA	17
Figura 4 - Descrição da alça de ligação ao substrato.....	18
Figura 5 - Flexibilidade da alça de ligação ao substrato.....	20

LISTA DE ABREVIATURAS E SIGLAS

ACP – *Acyl Carrier Protein* (Proteína Carregadora de Acilas)

BPTI – Inibidor da Tripsina Pancreática Bovina

CPK – Coloração para distinguir átomos de diferentes elementos químicos

CoA – Coenzima A

DM – Dinâmica Molecular

ENR – Enoil redutase

FabI – Enoil redutase de *E. coli*

FarmInf – Laboratório de FarmInformática

FAS II – *Fatty Acid Synthase Type II* (Síntese de Ácidos Graxos Tipo II)

InhA – Enzima 2-*trans*-enoil-ACP (CoA) redutase

inhA – Gene que codifica a enzima InhA

LABIO – Laboratório de Bioinformática, Modelagem e Simulação de Biosistemas

Mtb – *Mycobacterium tuberculosis*

NAD⁺ – Nicotinamida Adenina Dinucleotídeo, forma oxidada

NADH – Nicotinamida Adenina Dinucleotídeo, forma reduzida

NADPH – Nicotinamida Adenina Dinucleotídeo Fostato, forma reduzida

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

PUCRS – Pontifícia Universidade Católica do Rio Grande do Sul

PyMOL – Programa de visualização e modelagem molecular

SDR – *Short-Chain Dehydrogenase/ Reductase* (Família de enzimas desidrogenase/redutase de cadeia curta)

LISTA DE SÍMBOLOS

α – Alfa

β – Beta

H^- – Hidreto, forma aniônica do átomo de hidrogênio H

% – Porcentagem

kDa – Kilodaltons, unidade alternativa de massa molecular de proteínas multiplicada por mil

Å – Ångström, unidade de medida de comprimento

g/L – gramas por litro, unidade de medição

mg/mL – miligrama por mililitro, unidade de medição

ps – $1 \text{ ps} = 10^{-12}$ segundos ou um trilionésimo de segundo

SUMÁRIO

Capítulo 1

1	INTRODUÇÃO.....	12
	1.1 A importância da flexibilidade proteica.....	12
	1.2 A InhA de <i>Mycobacterium tuberculosis</i>	13
	1.3 A estrutura da InhA de <i>Mycobacterium tuberculosis</i>	13
	1.3.1 Cavidade de ligação do substrato.....	15
	1.3.2 Sítio de ligação do substrato.....	16
	1.3.3 As mudanças conformacionais observadas na InhA de <i>Mycobacterium tuberculosis</i>	18
2	JUSTIFICATIVA.....	21
3	OBEJTIVOS.....	22
	3.1 Objetivo Geral.....	22
	3.2 Objetivos Específicos.....	22

Capítulo 2

Artigo científico submetido ao periódico científico

Capítulo 3

	CONSIDERAÇÕES FINAIS.....	51
4	REFERÊNCIAS.....	53

Capítulo 1

Introdução

Justificativa

Objetivos

INTRODUÇÃO

A importância da flexibilidade proteica

O mecanismo para compreender a função de uma proteína, está em geral no conjunto de suas inúmeras conformações (Shinde *et al.*, 2013). A flexibilidade e/ou modificação estrutural de uma alça pode afetar a atividade de uma proteína. Desta forma, o estudo sobre como a flexibilidade afeta uma proteína pode proporcionar uma melhor compreensão da sua atividade biológica, auxiliando abordagens de desenvolvimento de novos fármacos.

A flexibilidade é a capacidade de uma proteína em adotar diferentes conformações. Esta dinâmica estrutural é uma etapa fundamental para o processo de reconhecimento biomolecular (Feixas *et al.*, 2014). Aspectos como a presença de outras moléculas e/ou mudanças no ambiente, podem influenciar a conformação de uma proteína em diferentes níveis. Estas variações estruturais podem ser pequenas como o rearranjo de algumas cadeias laterais de aminoácidos até influenciar no dobramento/desdobramento de proteínas (Teilum *et al.*, 2009). O fator comumente descrito como modulador da flexibilidade em proteínas é a associação de pequenas moléculas nas cavidades de ligação, o que favorece a uma maior estabilidade da estrutura proteica e, por consequência, uma redução na sua flexibilidade (Yesylevskyy *et al.*, 2006).

As técnicas mais difundidas na área de biologia estrutural são, a cristalografia de raios-X, a ressonância magnética nuclear e a microscopia crioeletrônica. Estas abordagens para a elucidação da estrutura tridimensional de proteínas apresentam metodologias distintas. A cristalografia de raios X é a metodologia mais utilizada para esta finalidade, entretanto, a proteína precisa estar no estado cristalino e a coleta dos dados de difração são realizadas em temperaturas criogênicas. Estas condições, podem influenciar diretamente na flexibilidade estrutural da proteína (Halle, 2004). Já a técnica de ressonância magnética nuclear, não necessita da formação de cristais para a sua resolução, os dados são coletados a partir de uma solução homogênea e concentrada da proteína de interesse. Esta condição permite que a flexibilidade estrutural seja analisada em detalhes. Contudo, devido a limitação no assinalamento dos espectros gerados, a utilização desta técnica para proteínas maiores que 30 kDa torna-se inviável. Por último, a técnica de microscopia crioeletrônica permite que grandes complexos proteicos sejam analisados,

entretanto a avaliação da flexibilidade também é limitada devido as baixas temperaturas utilizadas nos experimentos.

Com o intuito de auxiliar os estudos sobre flexibilidade estrutural em proteínas, diferentes abordagens computacionais foram desenvolvidas. Dentre estas técnicas, simulações computacionais por meio de dinâmica molecular (DM) clássica é a abordagem mais utilizada (Yesylevskyy *et al.*, 2006).

A InhA de *Mycobacterium tuberculosis*

A enzima InhA de *Mycobacterium tuberculosis* (Mtb), ou 2-*trans*-enoil-ACP (CoA) redutase (EC 1.3.1.9) é codificada pelo gene *inhA*. Esta proteína está envolvida na via de síntese de ácidos graxos tipo II (FAS II), a qual consiste de enzimas monofuncionais, responsáveis por alongar precursores de ácidos graxos produzindo longas cadeias de carbono do ramo meromicolato de ácidos micólicos (Takayama *et al.*, 2005). Estes são os componentes principais para a formação da parede celular do Mtb.

A quarta etapa da via FASII é catalisada pela enzima InhA, convertendo o 2-*trans*-enoil-ACP para acil-ACP por meio da transferência de um hidreto (H^-) do NADH para a posição C3 do substrato 2-*trans*-enoil-ACP (Rozwarski *et al.*, 1998), resultando na redução da dupla α tioéster (entre C2 e C3) da porção lipídica do substrato. Esta reação pode ocorrer em cadeias carbônicas que variam de 16 a 56 carbonos (Bloch, 1977; Qureshi *et al.*, 1984), evidenciando a flexibilidade tanto da cavidade de ligação ao substrato quanto da estrutura desta enzima.

A inibição ou inativação da InhA em Mtb bloqueia a biossíntese de ácidos micólicos, resultando no acúmulo de ácidos graxos saturados, dificultando na formação da parede celular de Mtb e, eventualmente, conduzindo à morte celular (Vilchère *et al.*, 2000).

A estrutura da InhA de *Mycobacterium tuberculosis*

A enzima InhA faz parte da família de enzimas desidrogenase/reductase de cadeia curta (do inglês, short-chain dehydrogenase/reductase) (SDR) (Persson *et al.*, 2003) e é composta por 269 aminoácidos com uma massa molecular de aproximadamente 29 kDa. As proteínas pertencentes a família SDR são constituídas por um núcleo central do tipo *Rossmann fold*, que compõe o sítio de ligação do NADH (Quemard *et al.*, 1995).

A InhA é um homotetrâmero em solução, sendo esta a sua estrutura biologicamente ativa (Quemard *et al.*, 1995). Cada subunidade da InhA é composta por uma cavidade de ligação independente, os quais distam aproximadamente 40 Å (Schroeder *et al.*, 2005). A topologia desta enzima se assemelha a uma cadeira (chair-like) do tipo *Rossmann fold*, caracterizada por uma folha- β contendo sete fitas- β e oito hélices- α , numa arquitetura $\alpha\beta\alpha$ sanduíche, que forma o sítio de ligação da coenzima NADH e a cavidade de ligação do substrato. A coenzima NADH associa-se no “assento” da estrutura, logo acima da folha β do domínio *Rossmann fold* e o sítio de ligação do substrato se localiza no “encosto” da estrutura, logo acima da cavidade de ligação do NADH (Dessen *et al.*, 1995). (**Figura 1**).

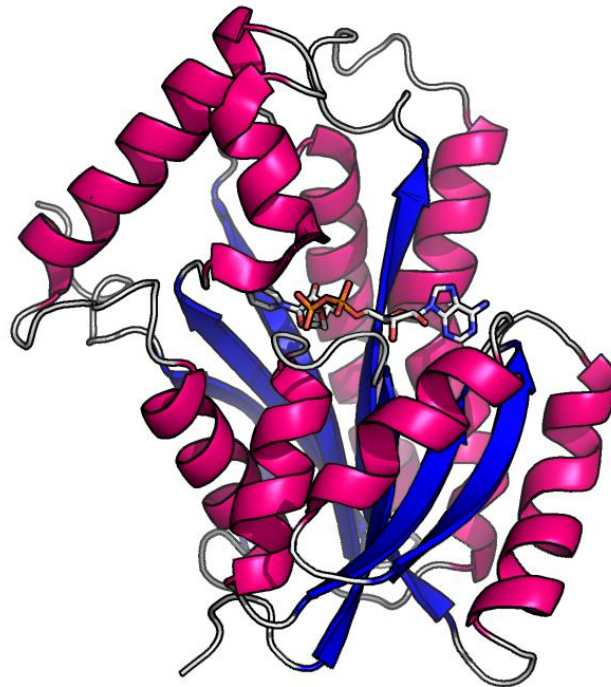


Figura 1. Estrutura monomérica da enzima 2-*trans*-enoil-ACP (CoA) redutase de *Mycobacterium tuberculosis* (código PDB: 1ENY). Representação do tipo *cartoon* da cadeia principal da enzima. As hélices α estão coloridas em rosa, as fitas- β da folha- β estão coloridas em azul. O NADH está na representação de palitos e colorido por tipo atômico, de acordo com o padrão CPK. Imagem gerada com PyMOL (Schrödinger, LLC, 2010).

A cavidade de ligação do substrato

A cavidade de ligação do substrato da InhA é composta por três regiões principais: a alça de ligação ao substrato, a alça A e a alça B (Dessen *et al.*, 1995). A alça de ligação ao substrato é formada por duas hélices- α perpendiculares as quais são mantidas por alças. Esta região se encontra no lado mais externo da cavidade de ligação, à esquerda do substrato. As outras duas grandes alças (A e B) encontram-se à direita do substrato (**Figura 2**). As duas hélices- α da alça de ligação ao substrato estão voltadas para fora do tetrâmero, em direção ao solvente, já as alças A e B estão do lado da enzima que faz contato com os demais monômeros (Rozwarski *et al.* 1999).

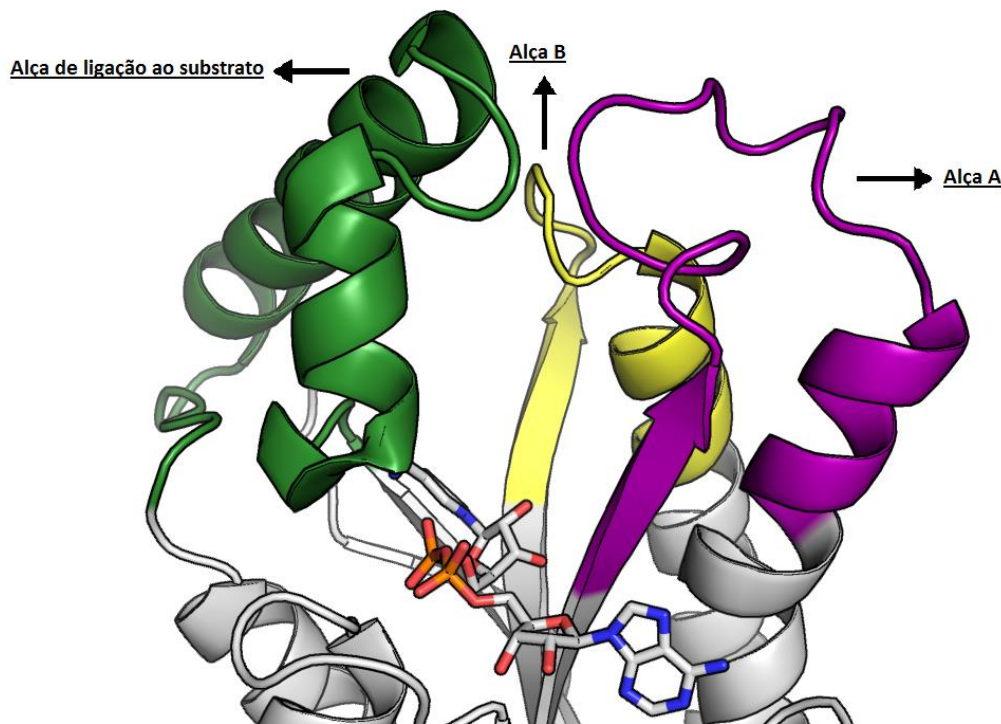


Figura 2. Região responsável pelo reconhecimento do substrato da enzima 2-*trans*-enil-ACP (CoA) redutase de *Mycobacterium tuberculosis* (código PDB: 1ENY). Representação do tipo *cartoon* da cadeia principal da enzima. A alça de ligação ao substrato (resíduos 192-233) está colorida em verde, a alça A (resíduos 96-120) em roxo e a alça B (resíduos 149-166) em amarelo. O NADH está na representação de palitos e colorido por tipo atômico, de acordo com o padrão CPK. Imagem gerada com PyMOL (Schrödinger, LLC, 2010).

Em 1999, Rozwarski e colaboradores depositaram a estrutura cristalográfica da InhA associada com o NADH e um análogo do substrato (código PDB: 1BVR). Esta foi a primeira estrutura apresentada em que o substrato poderia assumir conformação em “U” devido a cavidade de ligação ter uma forma oval, com dimensões de aproximadamente 16 Å x 13 Å x 7 Å. Já, a coenzima NADH assume

uma conformação estendida dentro de sua cavidade de ligação, em paralelo com a folha- β , onde a porção de adenina está direcionada para fora da cavidade e o anel nicotinamida está direcionado para o fundo da cavidade de ligação, próximo ao “encosto” da cadeira (Dessen *et al.*, 1995).

A cavidade de ligação da InhA é majoritariamente apolar, onde 69,82% dos resíduos são apolares (G96, F97, M98, P99, G102, M103, G104, I105, P107, F149, P151, A154, M155, P156, A157, W160, M161, G192, P193, I194, L197, A198, M199, A201, I202, V203, G204, G205, A206, L207, G208, A211, G212, A213, I215, L217 e L218) e apenas 30,18% são polares (Q100, N106, T101, D150, S152, R153, Y158, N159, R195, T196, S200, E209, E210, Q214, Q216 e E219). Dos resíduos apolares, a grande maioria está localizada na alça de ligação ao substrato (A198, M199, A201, I202, L207, I215, L218), o que permite que o substrato seja corretamente acomodado no interior da cavidade de ligação.

O sítio de ligação do substrato

A partir do alinhamento de sequências de aminoácidos das enoil-ACP redutases de FASII pode-se inferir uma correlação entre o comprimento da alça de ligação ao substrato e do substrato em si. As enoil-ACP redutases que apresentam preferência por substratos de cadeia curta (C2-C16) tem a alça de ligação ao substrato mais curta, composta por 13 resíduos (*E.coli* FabI e *B. napus* ENR), enquanto alças de substratos mais longas estão presentes em enzimas que apresentam preferência por substratos de cadeia longa (C16-C56) composta por 23 resíduos (*M. smegmatis* e *M. tuberculosis* InhA). Esta descoberta sugere que o tamanho da alça de ligação é determinante para distinguir entre substratos de cadeia longa e curta (Rozwarski *et al.*, 1999).

De acordo com Rozwarski e colaboradores, a cavidade de ligação do substrato pode ser dividida em duas partes, a parte mais exposta ao solvente representada pela alça de ligação ao substrato e as alças A e B, as quais estão voltadas para o centro do tetrâmero. Desta forma, a alça de ligação ao substrato pode apresentar uma maior liberdade conformacional, o que permite que o sítio ativo tenha uma certa adaptabilidade para acomodar substratos de comprimento variável (Rozwarski *et al.*, 1999). De acordo com as estruturas cristalográficas, a variação do volume da cavidade de ligação do substrato ($327,6 \text{ \AA}^3$ a $2,109.8 \text{ \AA}^3$) é o que confere a maior contribuição para a flexibilidade da enzima (Pauli *et al.*, 2013).

Esta flexibilidade é observada nas diferentes estruturas cristalográficas associadas a diferentes ligantes, onde a alça de ligação ao substrato se afasta 4 Å para a esquerda (**Figura 3**) enquanto que a cadeia lateral da Y158, resíduo essencial para a estabilização do substrato (Rozwarski *et al.* 1999) localizado no sítio ativo da enzima, rotaciona 60° para a direita.

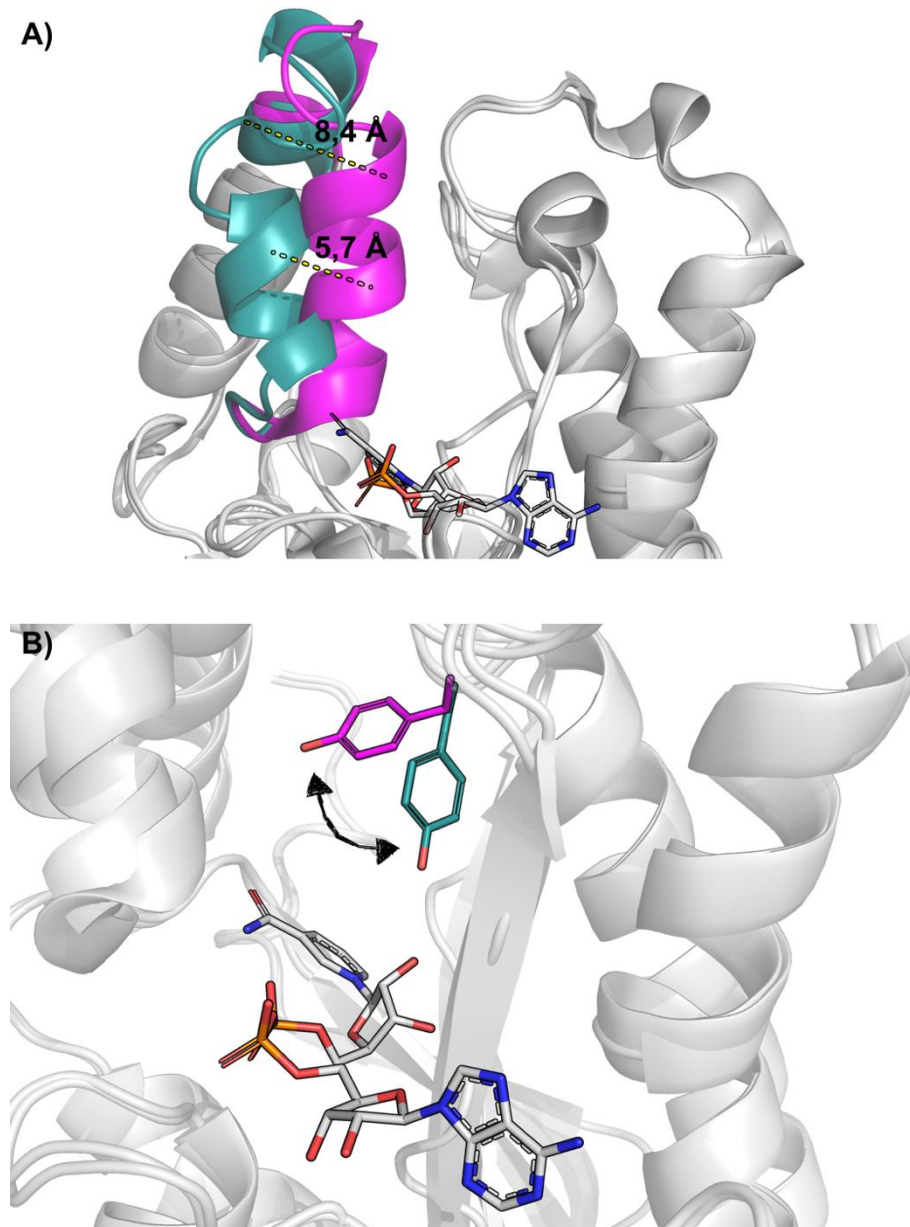


Figura 3. Mudanças conformacionais observadas na cavidade de ligação da enzima 2-*trans*-enoyl-ACP (CoA) redutase de *Mycobacterium tuberculosis*. Representação do tipo *cartoon* da cadeia principal da enzima. (A) Região da alça de ligação ao substrato, onde é possível observar a flexibilidade desta porção entre duas estruturas cristalográficas associada com o análogo do substrato em verde e com um inibidor em rosa (códigos de PDB: 1BVR e 3FNH, respectivamente). (B) Rotação da cadeia lateral do resíduo Y158 quando comparada a enzima associada apenas com o NADH e associada com o NADH e um análogo do substrato (códigos PDB: 1ENY e 1BVR, respectivamente). O NADH está na representação de palitos e colorido por tipo atômico, de acordo com o padrão CPK. Imagem gerada com PyMOL (Schrödinger, LLC, 2010).

As mudanças conformacionais observadas na InhA de *Mycobacterium tuberculosis*

De acordo com Qiu e colaboradores, a principal variação da enzima InhA é observada na região da alça de ligação ao substrato (resíduos: 192-233) (Quiu *et al.*, 1999). Esta é a alça mais longa encontrada na enzima, e também a mais longa quando comparada com a mesma região em enoi-ACP redutase de outras espécies (Kumar e Sobhia, 2013). Devido sua extensão, esta região é dividida em três porções: alça de ligação ao substrato superior (resíduos: 204-208), alça de ligação ao substrato médio (resíduos: 197-203) e alça de ligação ao substrato inferior (resíduos: 194-196) (**Figura 4**). A alça de ligação ao substrato superior interage com a alça B fechando a cavidade de ligação do substrato, fornecendo uma conformação adequada à cavidade do sítio ativo da InhA. Já a alça de ligação ao substrato inferior é uma volta que se comunica ao sítio de ligação do dinucleotídeo fazendo interações com a coenzima, NADH (Kumar e Sobhia, 2014).

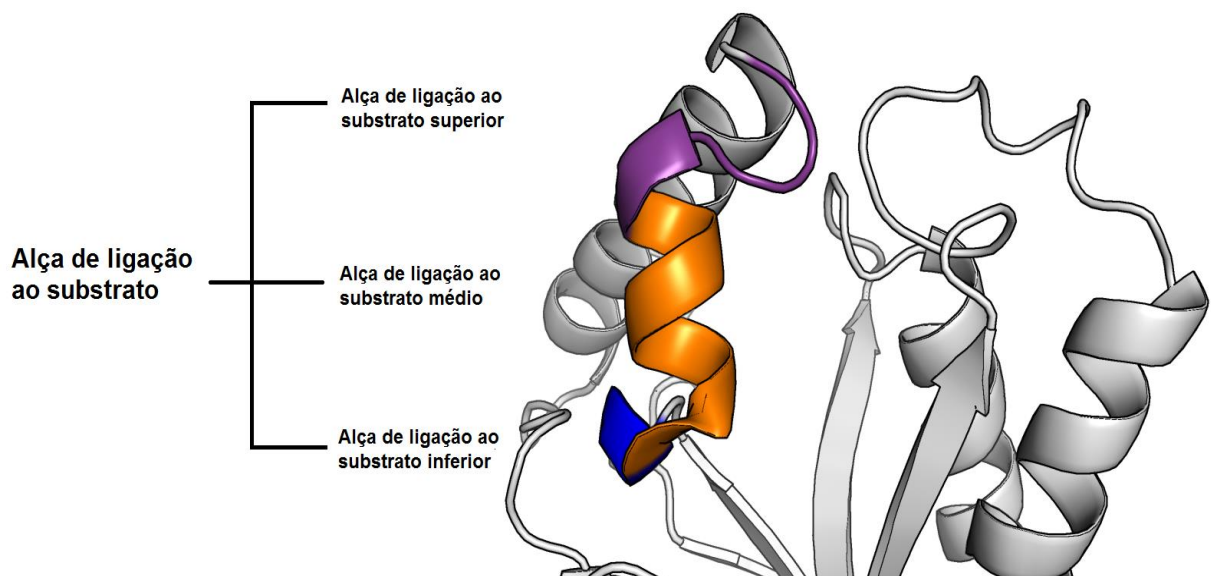


Figura 4. Descrição da alça de ligação ao substrato. As três partes da alça de ligação ao substrato estão destacadas em cores diferentes. Porção superior em roxo, porção média em laranja e porção inferior em azul. Imagem gerada com PyMOL (Schrödinger, LLC, 2010).

da Costa e colaboradores observaram uma desnaturação parcial das hélices α_6 e α_7 da InhA, quando associada ao composto IQG-607, enquanto que o domínio *Rossmann fold* permaneceu inalterado (**Figura 5**) (da Costa *et al.*, 2012). Estes resultados reforçam a capacidade de adaptação da cavidade de ligação da enzima InhA.

Kumar e colaboradores observaram mudanças conformacionais na hélice α_6 de acordo com a afinidade de interação da pequena molécula pela enzima.

Interações fracas entre proteína e pequenas moléculas do tipo fármaco, apresentam duas voltas na hélice $\alpha 6$, apresentando uma conformação aberta da alça de ligação ao substrato. Já os complexos com interações fortes, demonstraram uma conformação fechada e com três voltas na hélice $\alpha 6$. Desta forma, a estrutura da enzima InhA com ligantes menos potentes apresentam uma conformação mais aberta, quando comparada aos ligantes mais potentes que apresentam uma conformação mais fechada (Kumar e Sobhia, 2014).

Um estudo feito por Luckner e colaboradores afirmam que as interações entre os ligantes e os resíduos, A198, M199 e I202 são importantes para ordenar a alça de ligação ao substrato numa conformação fechada (Luckner *et al.*, 2010). Corroborando os dados de que interações intra e inter-moleculares são essenciais na dinâmica das proteínas e nos processos de reconhecimento molecular (Feixas *et al.*, 2014).

A fim de comprovar os diferentes estados conformacionais da enzima, como alça de ligação ao substrato “fechada” e “aberta”, Kumar e colaboradores demonstraram que nas estruturas cristalográficas, 1P44 e 1BVR (PDB IDs), a alça de ligação ao substrato encontra-se na conformação aberta, com o substrato exposto ao solvente (**Figura 5-A**). De acordo com estes resultados, é possível propor que ambas estruturas facilitam a entrada de substratos/inibidores. Entretanto, as estruturas cristalográficas de complexos ternários (PDB IDs 3FNH, 2NSD, 2H7M) e complexo binário (PDB ID 2IDZ) apresentam uma conformação fechada da alça de ligação ao substrato, ocasionando o fechamento da cavidade de ligação do substrato (**Figura 5-B**) (Kumar e Sobhia, 2013). Tendo em vista os trabalhos que descrevem a flexibilidade da InhA, é justificada a necessidade de estudar as mudanças conformacionais da enzima para que haja uma melhor compreensão sobre a afinidade de ligação a pequenas moléculas, assim como, a influência da estrutura quaternária na plasticidade da cavidade de ligação da enzima (Lima *et al.*, 2015).

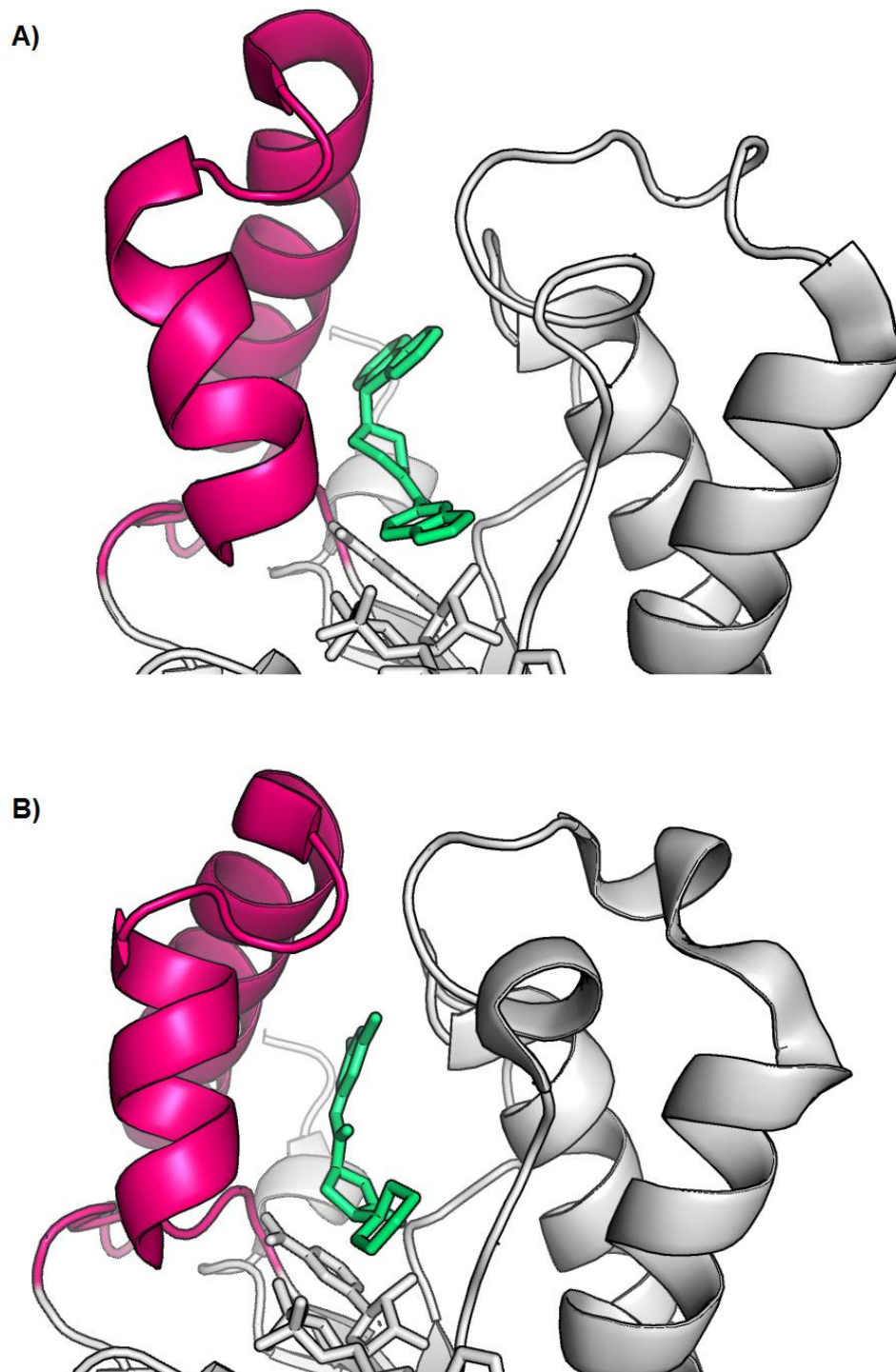


Figura 5. Flexibilidade da alça de ligação ao substrato. A) Conformação aberta (código PDB 1P44). B) Conformação fechada (código PDB 2H7M). Representação do tipo *cartoon* da cadeia principal da enzima, a alça de ligação ao substrato está colorida em rosa. Os inibidores e a coenzima NADH estão representados por palitos e coloridos em verde e cinza, respectivamente. Imagem gerada com PyMOL (Schrödinger, LLC, 2010).

2. JUSTIFICATIVA

A área de reconhecimento molecular é fundamental para a compreensão da bioquímica de processos biológicos. Saber como proteínas interagem entre si e/ou como um sítio ativo de uma enzima reconhece moléculas, é um procedimento fundamental para a compreensão de diversas doenças. Este ramo de estudos começou a ser desbravado em 1894 por Hermann Emil Fischer, quando o modelo de chave-fechadura foi proposto. Desde então, a compreensão de como proteínas se comportam começou a ser amplamente estudada e outros modelos foram descritos, como, por exemplo os modelos do encaixe induzido (Koshland Jr., 1958) e seleção conformacional (Weikl e Paul, 2014), evidenciando a importância da flexibilidade no processo de reconhecimento molecular.

A enzima InhA de *Mycobacterium tuberculosis* é capaz de reconhecer substratos que variam de 16 a 56 carbonos. A diversidade no tamanho dos seus reagentes requer que sua cavidade de ligação seja flexível o suficiente para que estes sejam devidamente posicionados durante a catálise da reação. Com isso, o estudo sobre a flexibilidade desta enzima torna-se interessante para a melhor compreensão de quais fatores podem estar envolvidos no processo de abertura e fechamento da sua cavidade de ligação e como esse processo pode influenciar na função catalítica da enzima.

3. OBJETIVOS

3.1. OBJETIVO GERAL

Analisar, por meio de simulações de dinâmica molecular clássica, a influência da estrutura quaternária da enzima InhA de *Mycobacterium tuberculosis* para o melhor entendimento da sua flexibilidade.

3.2. OBJETIVOS ESPECÍFICOS

- Realizar três simulações de 100ns da InhA na sua forma livre monomérica.
- Realizar três simulações de 100ns da InhA associada com a coenzima NADH na sua forma livre monomérica.
- Realizar três simulações de 100ns da InhA associada com a coenzima NADH e o análogo do substrato na sua forma livre monomérica.
- Realizar três simulações de 100ns da InhA na sua forma livre tetramérica.
- Realizar três simulações de 100ns da InhA associada com a coenzima NADH na sua forma livre tetramérica.
- Realizar três simulações de 100ns da InhA associada com a coenzima NADH e o análogo do substrato na sua forma livre tetramérica.

Capítulo 2

Artigo Científico

The importance of the quaternary structure to represent conformational ensembles of the major *Mycobacterium tuberculosis* drug target

Renata Fioravanti Tarabini^{a,b}, Luís Fernando Saraiva Macedo Timmers^{a,b,‡}, Carlos Eduardo Sequeiros-Borja^{a,b} and Osmar Norberto de Souza^{a,b*}.

^aLaboratório de Bioinformática, Modelagem e Simulação de Biosistemas (LABIO), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil.

^bPrograma de Pós-Graduação em Biologia Celular e Molecular, PUCRS, Porto Alegre, RS, Brazil.

[‡]Present address: Programa de Pós-Graduação em Biotecnologia (PPGBiotec), Universidade do Vale do Taquari - UNIVATES, Rua Avelino Talini, 171 - Bairro Universitário, Lajeado, RS, Brasil.

***Corresponding authors:** Dr. Osmar Norberto de Souza
e-mail: osmar.norberto@pucrs.br
Dr. Luis Fernando S.M. Timmers
e-mail luis.timmers@univates.br

Abstract

Flexibility is a feature intimately related to protein function, since conformational changes can be used to describe environmental changes, chemical modifications, protein-protein and protein-ligand interactions. In this study, we have investigated the influence of the quaternary structure of 2-*trans*-enoyl-ACP (CoA) reductase or InhA, from *Mycobacterium tuberculosis*, to its flexibility. We carried out classical molecular dynamics simulations using monomeric and tetrameric forms to elucidated the enzyme's flexibility. Overall, we observed statistically significant differences between conformational ensembles of tertiary and quaternary structures. In addition, the enzyme's binding site is the most affected region reinforcing the importance of the quaternary structure to evaluate the binding affinity of small molecules, as well as the effect of single point mutations to InhA protein dynamics.

Introduction

Flexibility is a feature intimately related to protein function since conformational changes in protein structures can be used to describe environmental changes, chemical modifications, in addition to protein-protein and protein-ligand interactions¹⁻³. Several research groups have sought to demonstrate the importance of protein flexibility at different levels. Weinkam and coworkers⁴ described mutations in a protein sequence that could be associated with changes in the allosteric conformational equilibrium. Plattner and Noé addressed the importance of conformational plasticity in receptors based on molecular dynamics simulations (MD) and Markov state models⁵. McCammon and coworkers⁶ also used MD simulations to provide insights into the function and mechanism of RNA polymerase II reactivation by the transcription factor IIS. In another work, Amaral and collaborators⁷ highlighted the importance of protein flexibility in structure-based drug discovery projects and more recently, Bringas and coworkers⁸ showed the importance of tertiary and quaternary structures flexibility to regulate oxygen affinity in human hemoglobin based on classical and hybrid MD simulations. Altogether, these works provide an overview of how the knowledge on protein flexibility achieved through computer simulations can be used to explain biological phenomena.

InhA or 2-*trans*-enoyl-ACP (CoA) reductase (EC 1.3.1.9) from *Mycobacterium tuberculosis* (Mt) is among the most widely studied enzymes for being the target of the first-line anti tubercular drug isoniazid⁹. The enzyme MtlNhA (InhA from Mt) has been investigated through different computational approaches to describe how inhibitor candidates bind to the active site¹⁰⁻¹², the hot-spot residues responsible for inhibitor binding¹³, as well as how point mutations, such as I21V, trigger isoniazid resistance¹⁴. This latter study was immediately confirmed using wet experiments¹⁵. Still, another work, by Tonge, Simmerling and coworkers, explored conformational changes in MtlNhA to explain experimental data on the modulation of slow-onset inhibitors¹⁶. Furthermore, MD simulations showed that the association of ligands in the MtlNhA substrate-binding pocket (SBP) can directly modulate the flexibility of its substrate-binding loop (SBL), A-loop, and B-loop^{10,13}. Schroeder and coworkers¹⁴ demonstrate a large flexibility of the SBL motif by using the monomeric form of the MtlNhA. Despite all computational studies on the MtlNhA flexibility, the impact of the quaternary structure was not taken into account since all information was retrieved

from tertiary structure simulation only. The use of the monomeric MtlNhA form in many computational studies can be supported on the fact that its binding sites occur about 40 Å apart from each other^{14,17}, suggesting that they are independent from each other. However, according to the tetramer organisation, the binding cavity is close to the monomers interface and the two major motifs of the substrate binding pocket (SBP), A-loop and B-loop, directly interact with the adjacent subunit, suggesting that the quaternary structure could be involved in the MtlNhA flexibility modulation. In addition, a detailed structural analysis of all available MtlNhA crystal structures corroborated the hypothesis in which A-loop and B-loop have lower values of B-factors in relation to the SBL, which is faced to the outside of the tetramer toward the solvent¹⁸⁻²¹.

Our work involves an extensive MD simulation study to evaluate the importance of the quaternary structure to MtlNhA flexibility by simulating three systems for the MtlNhA, (i) apo form, (ii) bound to NADH, and (iii) bound to NADH and the substrate analogue. Seeking to understand the role of the quaternary structure in the modulation of the MtlNhA flexibility, the tertiary and quaternary structures were also simulated to all systems. We subsequently applied probability density functions and a principal component analysis to assess the conformational ensembles obtained from the MD simulations in addition to using an ANOVA analysis to better describe the significance of the differences between the means of the simulated systems. The comparison of the flexibility between the tertiary and quaternary structures led to relevant insights into the conformational changes undertaken by this enzyme.

Material and Methods

MtlNhA crystal structures

We used three different systems to evaluate the impact of the quaternary structure on the MtlNhA flexibility: (i) apo form, (ii) bound to NADH, and (iii) bound to NADH and substrate analogue (THT). In addition, all systems were considered at two levels of structural organization: the tertiary (monomer) and quaternary (tetramer) forms. All crystal structures were obtained from the Protein Data Bank (PDB). Since the crystal structure of the MtlNhA apo form does not present the helix H6 of the SBL, we decided to use the PDB ID 1ENY without the NADH molecule as our starting

point to the apo form. For the systems associated with the NADH molecule and NADH:THT, we applied the PDB IDs 1ENY and 1BVR, respectively^{20,22}. Finally, we used the rotational and translational matrices deposited in the PDB to generate the quaternary structures, we used.

Molecular dynamics protocol

The molecular dynamics simulations of the MtlNhA enzyme were carried out using the Amber99SB force field²³ implemented in GROMACS 2016.3²⁴. Periodic bound conditions were applied and the numbers of particles, pressure, and temperature were maintained constant (NPT ensemble) during the whole production phase. The periodic boundary boxes of monomeric and tetrameric InhA systems were composed of 15116 and 47449 water molecules, respectively. V-rescale²⁵ thermostat was employed to maintain the system at constant temperature using a coupling time of 0.1 ps, and the Berendsen barostat²⁶ was applied to ensure that the system pressure was maintained at 1 bar. The LINCS algorithm²⁷ was implemented in order to constrain all of the covalent bonds involving hydrogen atoms and the system evolved in time. Steps of 2 fs. van der Waals interactions were computed using a 12 Å cutoff. The particle mesh Ewald method was applied to calculate electrostatic contributions in a grid with 1.2 Å spacing. The macromolecule was fully solvated using the TIP3P model²⁸ in a cubic box extending 10 Å from the macromolecule surfaces. The systems were submitted to steepest-descent energy minimization up to a tolerance of 1000 kJ mol⁻¹ nm⁻¹ in order to remove close contacts of van der Waals forces. The equilibration phase was performed in two steps: (i) a NVT equilibration for 500 ps followed by (ii) a NPT equilibration step lasting 1000 ps before the production phase. We performed three simulations to each MtlNhA system - (i) apo form, (ii) bound to NADH molecule, and (iii) bound to NADH:THT - either for the tertiary and quaternary structures by using different random seeds. The production phase of each MD simulation lasted for 100 ns and the total simulation time was 1.8 μs.

Structural analysis of the MtlNhA enzyme

To better describe the impact of the quaternary structure to the MtlNhA flexibility, the conformational ensembles were analysed according to the distances, angle, and area variations defined in the substrate-binding pocket (SBP). All

analyses were performed based on the three major motifs presented in the SBP, (i) A-loop (F96, M97, P98, Q99, T100, G101, M102, G103, I104, N105, P106, F107, F108, D109, A110, P111, Y112, A113, D114, V115, S116, K117, G118, I119, H120), (ii) B-loop (D149, P150, S151, R152, A153, M154, P155, A156, Y157, N158, W159, M160, T161, V162, A163, K164, S165, A166), and (iii) SBL (P192, I192, R194, T195, L196, A197, M198, S199, A200, I201, V202, G203, G204, A205, L206, G207, E208, E209, A210, G211, A212, Q213, I214, Q215, L216, L217, E218, E219, G220, W221, D222, Q223, R224, A225, P226, I227, G228, W229, N230, M231, K232, D233). To monitor the process of SBP open and closure, we defined a pincer angle (**Figure 1B**) based on the mass center of the motifs A-loop, B-loop, and SBL. In addition, we also assessed the triangle area formed by the pincer angle, in addition to monitoring the distances between SBL and A-loop (**Figure 1D**) as well as between A-loop and B-loop (**Figure 1C**) to reach further information on the influence of the quaternary structure. Furthermore, we assessed the root mean square displacement (RMSD) of the residues Y158 and K165 considering its importance to the catalytic mechanism²⁹. All analyses were carried out on GROMACS 2016.3 package in house scripts.

Principal Component Analysis

Molecular dynamics simulations produce large data amount to be used to describe protein function and dynamics, protein-ligand, and protein-protein interactions in microscopic details. However, it is important to reduce the dimensionality of these large data sets to extract functionally relevant information. Aiming at better understanding the differences among the systems, we applied a principal component analysis to analyse the internal collective motions of all systems simulated as the function of C α atoms by calculating a covariance matrix and its eigenvectors. Softwares Gromacs “gmx covar” and “gmx anaeig” were used to calculate the eigenvalues and eigenvectors of the MD trajectories and obtain a description of the most flexible regions of each system, respectively. We assumed that the tetramer monomers have a similar behaviour as well as the RMSF for the tetrameric systems correspond to a mean of all subunits in order to establish a comparison with the root mean square fluctuations (RMSF) between monomeric and tetrameric structures.

Statistical Analysis

Statistical analyses were performed on RStudio³⁰. One-way analysis of variance (ANOVA) was used to evaluate the occurrence of statistically significant differences between the means of the systems simulated. Since Levene's statistics test³¹ revealed an unequal variance of the system distributions, we applied the Games-Howell³² test at 95% of confidence. Bonferroni *post hoc* correction test was applied to ensure that relative false-positive rate applied to the data set does not exceed the specific value ($p_{\text{adjusted}} < 0.01$). Our null hypothesis (H0) was that the monomer structure has the same mean distribution as the tetramer monomers. The alternative hypothesis (HA) was that the mean distribution of the monomer structure was different from the tetramer monomers.

Results and Discussion

The MM/MD simulations were performed to better describe the impact of the quaternary structure on the flexibility of MtlNhA enzyme. The comparison allows us to highlight the influence of the tetrameric structure to the flexibility of A-loop, B-loop, and Substrate-binding loop (SBL). In addition, it was possible to demonstrate whether the conformational changes of these motifs are restricted due to tetramer contacts.

Comparison of monomeric and tetrameric ensembles of the MtlNhA apo form

The analysis of the relationship between area and pincer angle shows that the simulations of the apo structure were able to explore the conformational space considering the absence of interactions with small molecules into the active site. However, by comparing the tertiary and quaternary structures the monomeric system of apo MtlNhA proved to assess conformations which are not accessible to the tetrameric system; in addition, both systems have substantially different values for the most populated conformation. **Figure 2** illustrates the conformational ensembles of the tertiary and quaternary structures according to the probability density function. The monomeric system have a more spread plot (most representative conformation with an area of 167.6 \AA^2 and a pincer angle of 51.4°), whereas the conformation of the tetrameric system has an area of 148.8 \AA^2 and a pincer angle of 47.4° . Overall, the results suggest that the volume in the substrate-binding cavity is increased for the tertiary structure for not presenting any contacts with adjacent subunits. The

increased volume observed for monomeric structures is due to a higher flexibility of the A-loop, B-loop, and SBL. Furthermore, by analyzing the root mean square fluctuation (RMSF) of the first eigenvector, three flexible motifs (A-loop, B-loop, and SBL) are revealed in the monomeric form, while the tetrameric system has only SBL as a flexible motif (**Figure 3**). PCA analysis also demonstrated that total positional fluctuations described by the first 50 eigenvectors are 91.63%, 56.18%, and 85.78% for the simulations using the monomer and 73.33%, 76.95%, and 77.02% (**Table ST1**). In relation to crystallographic structures, the conformation obtained from the simulations of the quaternary structures may provide a more comprehensive description of the experimental results.

The representative conformation of the tetrameric ensemble presents a higher flexibility in the SBL, whereas A-loop and B-loop are more rigid than the monomeric ensemble. The rigidity of A-loop and B-loop results from protein-protein contacts between the adjacent subunits related to the quaternary structure. The analyses on the monomeric and tetrameric distributions means revealed statistically significant differences (p -value < 0.01). The analysis of distances between A-loop and B-loop illustrate pronounced differences when comparing monomeric and tetrameric systems with the most representative conformation presents values of 16.5 Å and 21.3 Å to monomeric and tetrameric systems, respectively. In addition, the comparisons of the distance between the SBL and A-loop also corroborate to different behaviour of protein dynamics, presenting values of 18.8 Å and 22.0 Å to monomeric and tetrameric systems. Therefore, the visual differences were confirmed through statistical analysis yielding p -values of < 0.01 to all comparisons between apo monomeric and tetrameric forms, indicating a statistically significant difference between these distributions (**Table ST2-ST5**).

Comparison of monomeric and tetrameric ensembles of the MtlNhA holo form

The dynamics behaviour of the MtlNhA holo form is also affected by the presence of the quaternary structure, mainly on the binding pocket. MtlNhA monomeric and tetrameric ensembles associated with NADH and NADH:THT show substantial differences regarding the conformational space sampled. **Figure 4** shows the probability density plots of monomeric and tetrameric ensembles of MtlNhA:NADH (**Fig. 4A** and **4B**) and MtlNhA:NADH:THT (**Fig. 4C** and **4D**). The most representative conformation of the monomeric MtlNhA:NADH system presents a

more compact conformation for the substrate binding cavity with a pincer angle of 42.1° and an area of 171.0 \AA^2 . It is interesting to emphasise that the value obtained for the pincer angle in the monomeric system does not occur in the tetrameric ensemble, in which the pincer angle of the most representative conformation is 48.0° and an area of 162.0 \AA^2 . Such a difference could be also associated with the quaternary structure since the monomeric system does not present adjacent subunits to restrict its flexibility. The monomeric ensemble is able to access conformations regarding the area parameter, which is also described by the tetrameric ensemble; however, these conformations represent just transient states and not local minimum structures. It is worth noticing that just the quaternary structure is able to stabilise the region Tyr158-X₆-Lys165, which is important to substrate recognition and enzyme catalysis. Furthermore, the analysis of the distances between the SBL and A-loop, as well as between A-loop and B-loop, also present a large difference in the conformational space sampled by the binary and ternary complexes in monomeric and tetrameric forms (**Figure SF1**). Our statistical analysis support the difference between the distributions of monomeric and tetrameric ensembles of the MtlNhA:NADH, yielding p-values of < 0.01 (**Tables ST1-ST4**).

The ternary complex (MtlNhA:NADH:THT) also have similar behaviour as apo form and binary complex (MtlNhA:NADH) (**Figure 4C** and **4D**) highlighting differences in protein dynamics comparing monomeric and tetrameric ensembles. **Figures 5** and **6** highlight the regions involved in an increased flexibility in relation to tertiary and quaternary structure. It is important to observe an intrinsic relation of A-loop, B-loop, and SBL to the dynamics behaviour of the MtlNhA binding pocket. All PCA analyses regarding the contributions of the eigenvectors are described in **Table ST1**.

Although the monomeric ensemble presents a more spread distribution regarding the pincer angle in addition to statistical analysis showing that the means of monomeric and tetrameric distributions are statistically different (p-value < 0.01), the most representative conformation have a similar tetrameric structure value, which indicates that the presence of a ligand in the substrate binding site can modulate the region flexibility. In addition, an open conformation of the substrate binding site occurred for the MtlNhA ternary complex, which is had appeared in a previous work on the conformational change possibly facilitating the entry and/or exit of external substrates¹⁶. Considering these results, monomeric ensemble diverge from all tetrameric subunits, demonstrating the importance of the quaternary structure to

study the MtlNhA flexibility. These findings reinforce the importance of the quaternary structure to evaluate the binding affinity of small molecules, as well as the effect of single point mutations to protein dynamics.

Conclusion

In previous works, the MtlNhA enzyme has been used as a monomer in MM/MD simulations^{9,33,34}, based on the assumption that the binding sites are independent from each other and thereby the tertiary structure could represent the behaviour of the biological structure in solution. However, based on our results there is a statistically significant difference between conformational ensembles of tertiary and quaternary structures. Our analyses show that the InhA active is much more flexible in the monomeric than the tetrameric form. Furthermore, by comparing the motions of A-loop, B-loop, and substrate-binding loop (SBL), we observed that in the quaternary structure, flexibility is restricted to the SBL motif, while in the monomeric, tertiary structure, flexibility varies among all these motifs. These differences could be associated with the tetrameric form in contact with an adjacent subunit due to the quaternary packing. In addition, statistical comparisons between the conformational ensembles of monomeric and tetrameric forms yielded p-values < 0.01, indicating a statistically significant difference between the means of these distributions. The MM/MD simulations allow us to observe that the tetramer subunits can exhibit conformational ensembles without significant differences. These findings suggest that the use of the MtlNhA quaternary structure may be mandatory in studies of its function and dynamics, as well as in virtual screening using multi-conformation or ensemble-docking approaches. Altogether, these results constitute a relevant contribution to the description of MtlNhA flexibility by highlighting the importance of the quaternary structure to the description of its conformational changes.

Acknowledgements

This work was supported in part by grants to ONS (CNPq, 304704/2018-0; FAPERGS, TO2054-2551/13-0). ONS is a CNPq Research Fellow. RFT was supported by CAPES scholarship. LFSMT was supported by CAPES/PNPD fellowship. C.E.S.B. was supported by a grant from the Graduate Agreement Student's Program - PEC-PG from CAPES/CNPq - Brazil (190602/2014-7).

Author contributions

R.F.T., L.F.S.M.T., C.E.S.B, and O.N.S. designed the study. R.F.T., L.F.S.M.T., C.E.S.B. performed the experiments and analyzed the data. R.F.T., L.F.S.M.T., C.E.S.B, and O.N.S. wrote the manuscript.

Competing interests

The authors declare no competing financial and/or non-financial interest.

References

- (1) Marsh, J. A., Teichmann, S. A. Protein flexibility facilitates quaternary structure assembly and evolution. *PLoS Biol.* **12**, e1001870 (2014).
- (2) Marsh, J. A., *et al.* Protein complexes are under evolutionary selection to assemble via ordered pathways. *Cell.* **153**(2), 461-470 (2013).
- (3) Teilum, K., Olsen, J. G., Kragelund, B. B. Protein stability, flexibility and function. *Biochim Biophys Acta Proteins Proteom.* **1814**, 969-976 (2011).
- (4) Weinkam, P., Chen, Y. C., Pons, J., Sali, A. Impact of mutations on the allosteric conformational equilibrium. *J. Mol. Biol.* **425**(3), 647-661 (2013).
- (5) Plattner, N., Noé, F. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun.* **6**, 7653 (2015).
- (6) Eun, C., Ortiz-Sánchez, J. M., Da, L., Wang, D., McCammon, J. A. Molecular dynamics simulation study of conformational changes of transcription factor *tffi*s during RNA polymerase II transcriptional arrest and reactivation. *PLoS One.* **9**(5), e97975 (2014).
- (7) Amaral, M., *et al.* Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat. Commun.* **8**(1), 2276 (2017).
- (8) Bringas, M., Petruk, A. A., Estrin, D. A., Capece, L., Martí, M. A. Tertiary and quaternary structural basis of oxygen affinity in human hemoglobin as revealed by multiscale simulations. *Sci. Rep.* **7**(1), 10926 (2017).
- (9) Schroeder, E. K., de Souza, O. N., Santos, D. S., Blanchard, J. S., Basso, L.A. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. *Curr Pharm Biotechnol.* **3**(3), 197-225 (2002).
- (10) da Costa, A. L., *et al.* Conformational changes in 2-trans-enoyl-ACP (CoA) reductase (InhA) from *M. tuberculosis* induced by an inorganic complex: a molecular dynamics simulation study. *J Mol Model.* **18**, 1779-1790 (2012).
- (11) Punkvang, A., *et al.* Investigating the structural basis of arylamides to improve potency against *M. tuberculosis* strain through molecular dynamics simulations. *Eur. J. Med. Chem.* **45**(12), 5585-5593 (2010).

- (12) Cohen, E. M., Machado, K. S., Cohen, M., de Souza, O. N. Effect of the explicit flexibility of the InhA enzyme from *Mycobacterium tuberculosis* in molecular docking simulations. *BMC Genomics*. **12**(4), S7 (2011).
- (13) Kumar, V., Sobhia, M. E. Molecular dynamics-based investigation of InhA substrate binding loop for diverse biological activity of direct InhA inhibitors. *J Biomol Struct Dyn*. **34**(11), 2434-2452 (2016).
- (14) Schroeder, E. K., Basso, L. A., Santos, D. S., de Souza, O. N. 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).
- (15) Oliveira, J. S., *et al.* 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**(3), 646-666 (2006).
- (16) Lai, C. T., *et al.* Rational modulation of the induced-fit conformational change for slow-onset inhibition in *Mycobacterium tuberculosis* InhA. *Biochem*. **54**(30), 4683-4691 (2015).
- (17) Pauli, I., *et al.* Discovery of new inhibitors of *Mycobacterium tuberculosis* InhA enzyme using virtual screening and a 3D-pharmacophore-based approach. *J. Chem. Inf. Model.* **53**, 2390-2401 (2013).
- (18) Chollet, A., Maveyraud, L., Lherbet, C., Bernardes-Genisson, V. An overview on crystal structures of InhA protein: Apo-form, in complex with its natural ligands and inhibitors. *Eur. J. Med. Chem.* **146**, 318-343 (2018).
- (19) Qiu, X., *et al.* Molecular basis for triclosan activity involves a flipping loop in the active site. *Protein Sci.* **8**, 2529-2532, (1999).
- (20) Dessen, A., Quemard, A., Blanchard, J. S., Jacobs Jr, W. R., Sacchettin, J. C. Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. *Science*. **267**, 1638 (1995).
- (21) Quemard, A., *et al.* Enzymic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochem*. **34**, 8235-8241 (1995).
- (22) Rozwarski, D. A., Vilchèze, C., Sugantino, M., Bittman, R., Sacchettini, J. C. Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate. *J Biol Chem*. **274**, 15582-15589 (1999).
- (23) Hornak, V., *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*. **65**, 712-725 (2006).
- (24) van der Spoel, D., *et al.* GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701-1718 (2005).
- (25) Bussi, G., Donadio, D., Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
- (26) Berendsen, H. J., Postma, J. V., van Gunsteren, W. F., DiNola, A. R. H. J., Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684-3690 (1984).

- (27) Hess, B., Bekker, H., Berendsen, H. J., Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463-1472 (1997).
- (28) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935 (1983).
- (29) Parikh, S., Moynihan, D. P., Xiao, G., Tonge, P. J. Roles of tyrosine 158 and lysine 165 in the catalytic mechanism of InhA, the enoyl-ACP reductase from *Mycobacterium tuberculosis*. *Biochem.* **38**(41), 13623-13634 (1999).
- (30) RStudio Team. RStudio: Integrated Development for R. RStudio Inc. 2017. URL <http://www.rstudio.com/>.
- (31) Levene H. Robust testes for equality of variances. Stanford Univ. Press. 1960; 278–292.
- (32) Games, P. A., Howell, J. F. Pair wise multiple comparison procedures with unequal n's and/or variances. *J Educ Stat.* **1**, 13-125 (1976).
- (33) Shaw, D. J., *et al.* Examining the role of protein structural dynamics in drug resistance in *Mycobacterium tuberculosis*. *Chem Sci.* **8**(12), 8384-8399 (2017).
- (34) Merget, B., Sotriffer, C. A. Slow-onset inhibition of *Mycobacterium tuberculosis* InhA: revealing molecular determinants of residence time by MD simulations. *PLoS One.* **10**(5), e0127009 (2015).
- (35) Schrödinger LLC . The PyMOL Molecular Graphics System. 2010; Version 1.3r1.
- (36) Hunter, J. D. Matplotlib: A 2D graphics environment. *Comp. Sci. Eng.* **9**(3), 90-95 (2007).
- (37) Turner, P. J. XMGRACE, Version 5.1. 19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology. 2005.

Figures

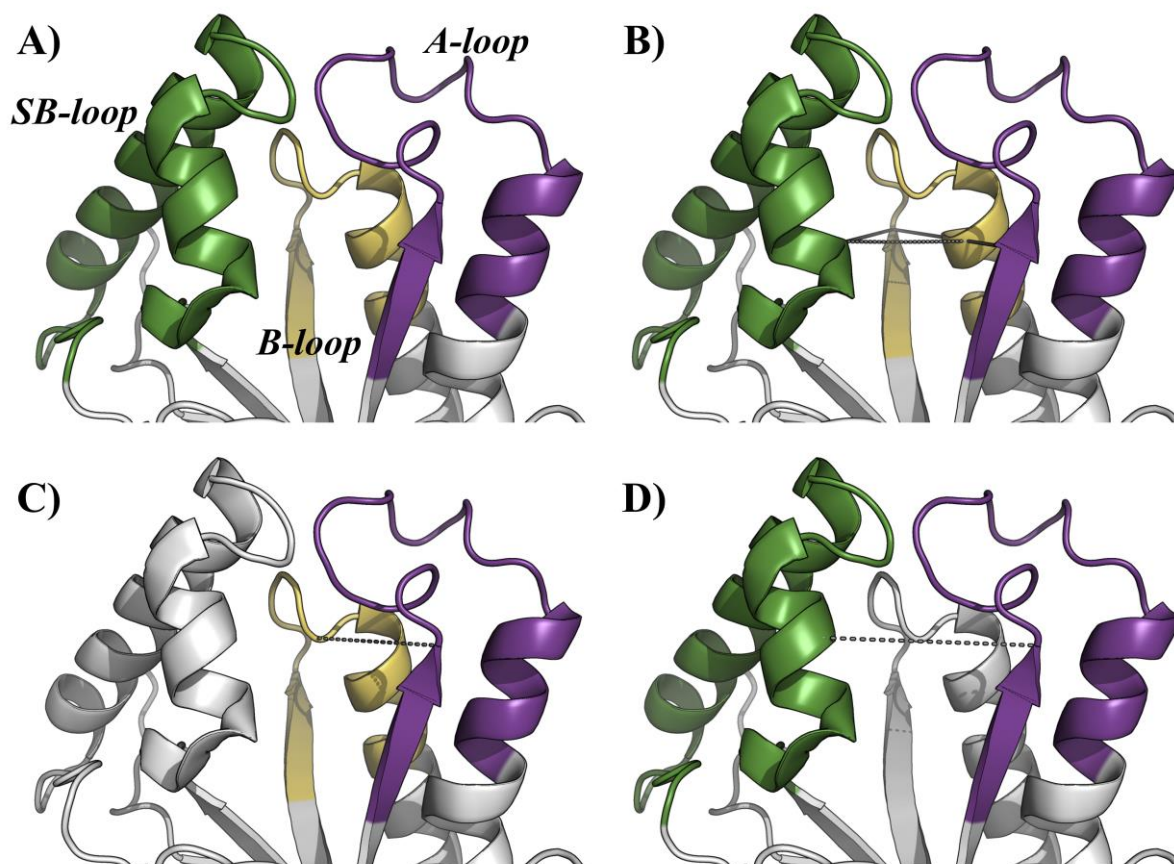


Figure 1. Definition of the structural analysis. The substrate binding cavity is defined by three motifs, the SB-loop (green), A-loop (purple) and B-loop (yellow). **A)** Description of the motifs involved in the analyses. **B)** Definition of the pincer angle and triangle area. **C)** A-loop and B-loop distance, and **D)** SB-loop and A-loop distance. The main chain of the enzyme is represented as *cartoon*. Image generated with PyMOL³⁵.

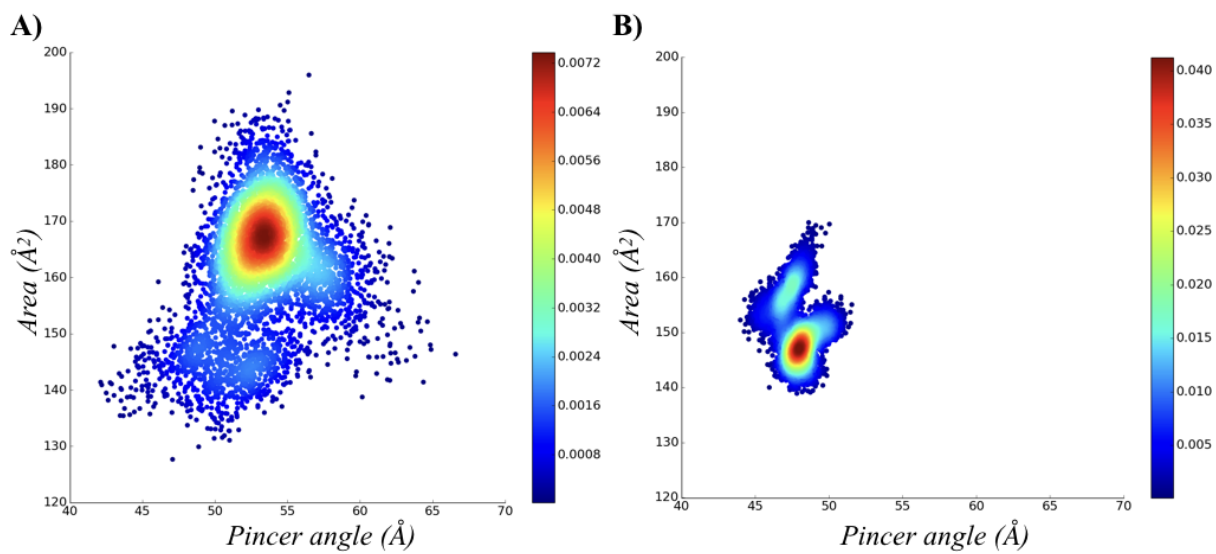


Figure 2. Description of the apo conformational ensembles of MtlNhA. The probability density plots of (A) monomeric and (B) tetrameric forms highlight the difference between the conformational ensembles. Image generated with Matplotlib³⁶.

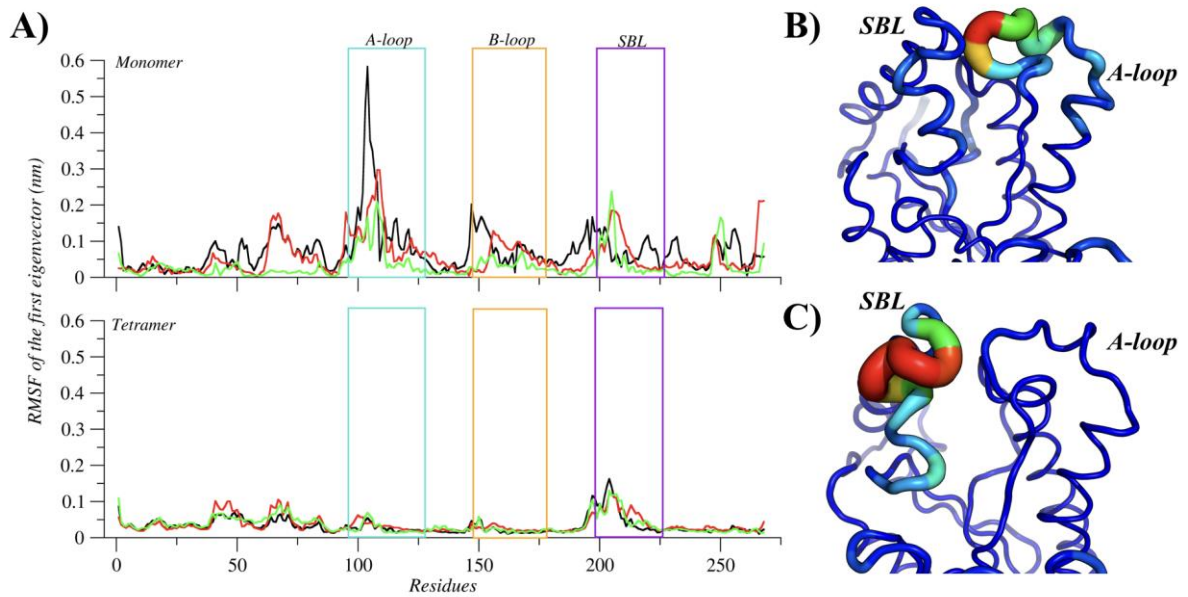


Figure 3. Flexibility of the MtlNhA apo form. (A) Root mean of the root mean square fluctuation (RMSF) for the monomeric and tetrameric systems, according to the first eigenvector. The regions A-loop, B-loop, and SBL are highlighted with boxes and coloured in light blue, orange, and purple, respectively. The mobility and conformation change of the most representative structure extracted from the MD trajectories for monomer (B) and tetramer (C). The thickness is proportional to B-factor values. Image generated with XMGrace³⁷ and PyMOL³⁵.

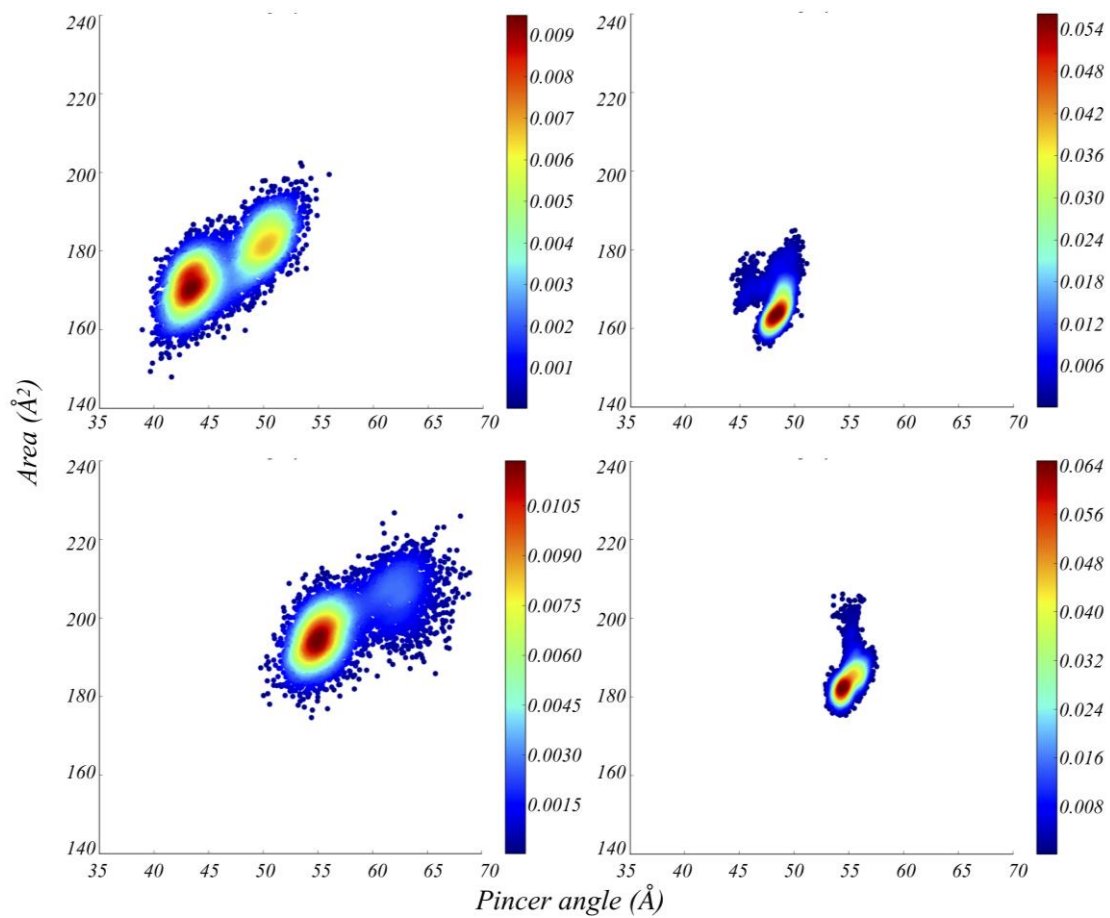


Figure 4. Description of the conformational ensembles of MtlNhA associated with NADH and NADH:THT. The probability density plots of (A) monomeric MtlNhA:NADH, (B) tetrameric MtlNhA:NADH, (C) monomeric MtlNhA:NADH:THT, and (D) tetrameric MtlNhA:NADH:THT highlights the differences of the conformational ensembles. Image generated with Matplotlib³⁶.

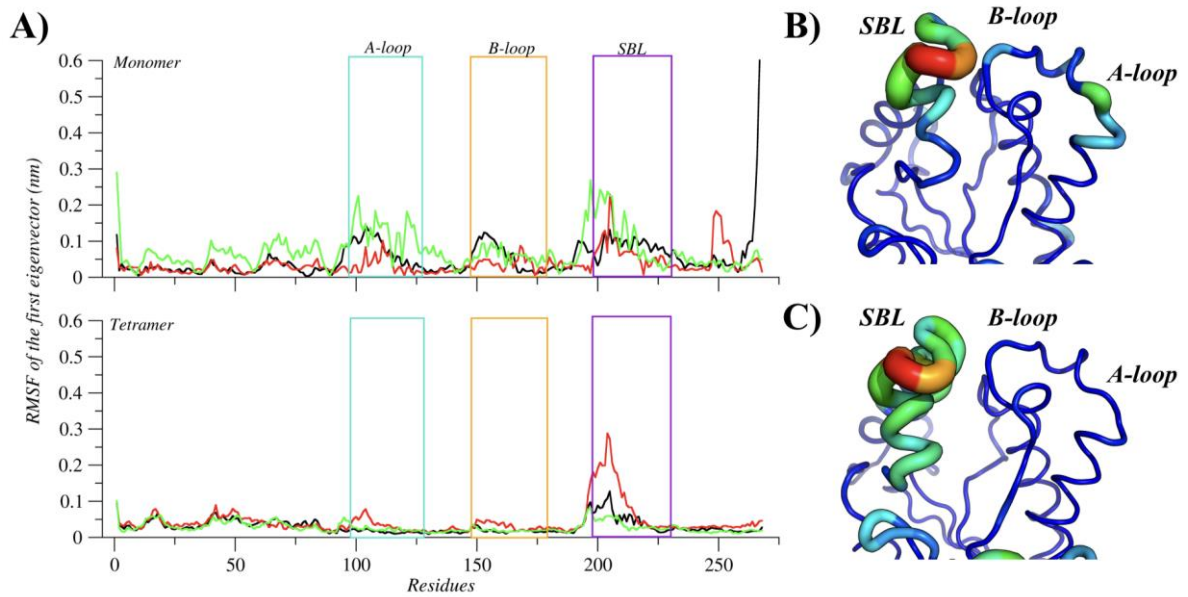


Figure 5. Flexibility of the MtlNhA associated with NADH. (A) Root mean of the root mean square fluctuation (RMSF) for the monomeric and tetrameric systems, according to the first eigenvector. The regions A-loop, B-loop, and SBL are highlighted with boxes and coloured in light blue, orange, and purple, respectively. The mobility and conformation change of the most representative structure extracted from the MD trajectories for monomer (B) and tetramer (C). The thickness is proportional to B-factor values. Image generated with XMGrace³⁷ and PyMOL³⁵.

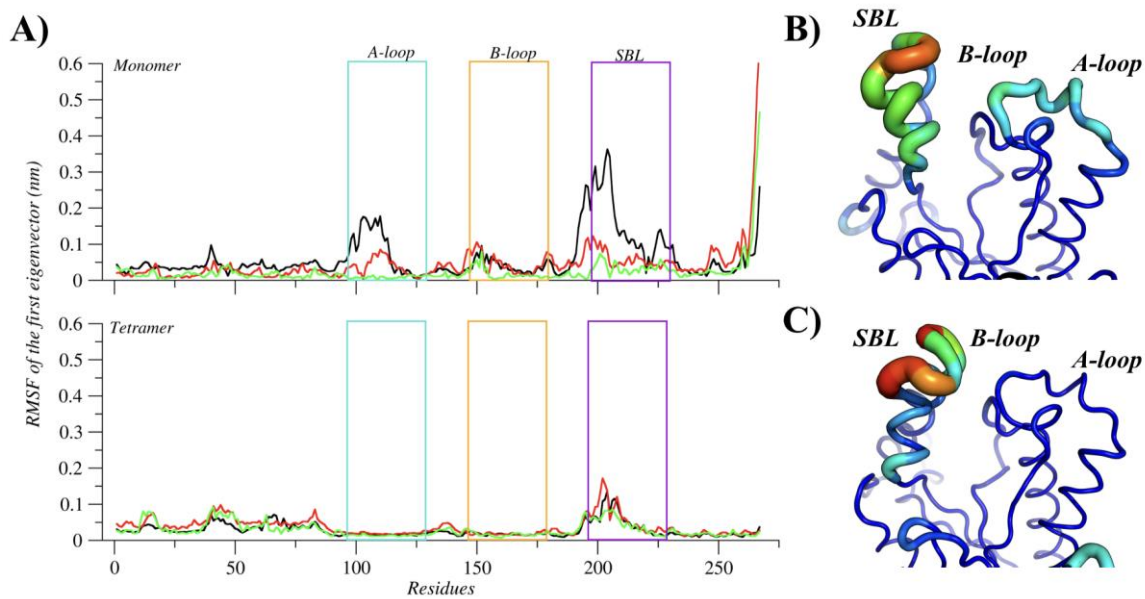


Figure 6. Flexibility of the MtlInhA associated with NADH and substrate analogue (THT). (A) Root mean of the root mean square fluctuation (RMSF) for the monomeric and tetrameric systems, according to the first eigenvector. The regions A-loop, B-loop, and SBL are highlighted with boxes and coloured in light blue, orange, and purple, respectively. The mobility and conformation change of the most representative structure extracted from the MD trajectories for monomer (B) and tetramer (C). The thickness is proportional to B-factor values. Image generated with XMGrace³⁷ and PyMOL³⁵.

SUPPLEMENTARY INFORMATION:

The importance of the quaternary structure to represent conformational ensembles of the major *Mycobacterium tuberculosis* drug target

Renata Fioravanti Tarabini^{a,b}, Luís Fernando Saraiva Macedo Timmers^{a,b*}, Carlos Eduardo Sequeiros-Borja^{a,b} and Osmar Norberto de Souza^{a,b*}.

^aLaboratório de Bioinformática, Modelagem e Simulação de Biosistemas (LABIO), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Av. Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil.

^bPrograma de Pós-Graduação em Biologia Celular e Molecular, PUCRS, Porto Alegre, RS, Brazil.

Present address:

Dr. Luis Fernando S.M. Timmers, Programa de Pós-Graduação em Biotecnologia (PPGBiotec), Universidade do Vale do Taquari -Univates, Rua Avelino Talini, 171 - Bairro Universitário, Lajeado, RS, Brazil.

Contents:

Figure SF1. Probability density plots of distance of SB-loop and A-loop by the distance of A-loop and B-loop. (A) MtlNhA apo form, (B) MtlNhA:NADH, (C) MtlNhA:NADH:THT monomeric ensembles. (D) MtlNhA apo form, (E) MtlNhA:NADH, and (F) MtlNhA:NADH:THT tetrameric ensembles

Table ST1. Contribution of the first and the sum of the first 50 eigenvectors to each system, according to the principal component analysis.

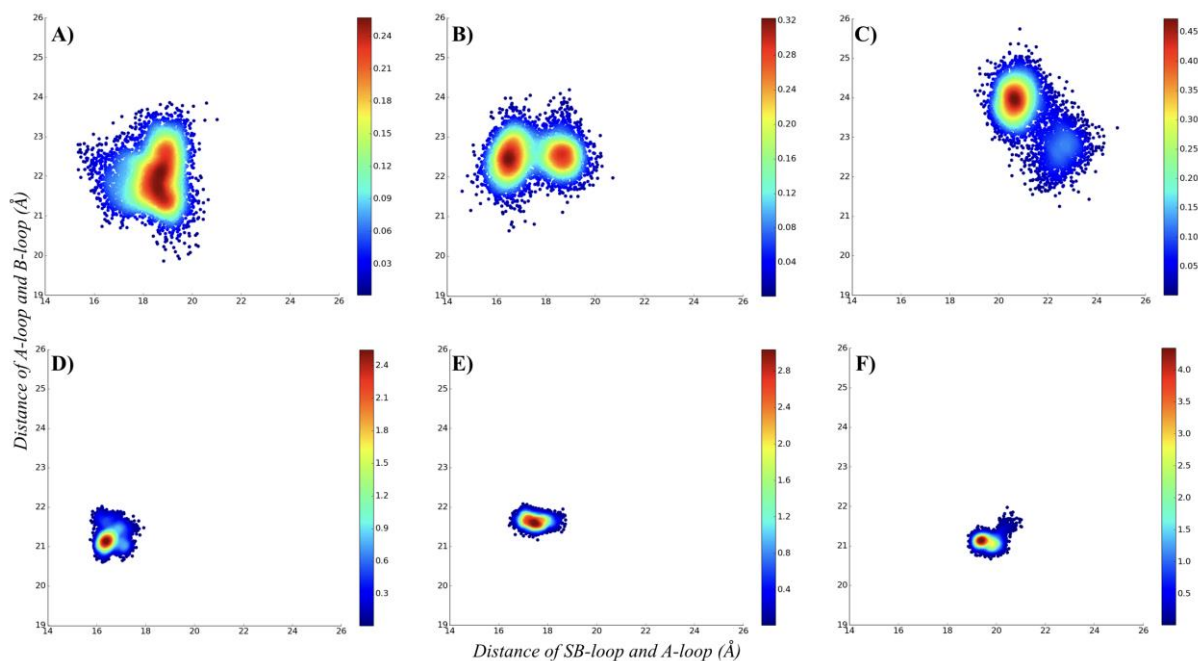
Table ST2. Statistical analysis of pincer angle distributions between monomeric and tetrameric structures.

Table ST3. Statistical analysis of area distribution between monomeric and tetrameric ensembles.

Table ST4. Statistical analysis of SB-loop and A-loop distances distributions between monomeric and tetrameric ensembles.

Table ST5. Statistical analysis of B-loop and A-loop distances distributions between monomeric and tetrameric ensembles.

Figure SF1. Probability density plots of distance of SB-loop and A-loop by the distance of A-loop and B-loop. (A) MtlNhA apo form, (B) MtlNhA:NADH, (C) MtlNhA:NADH:THT monomeric ensembles. (D) MtlNhA apo form, (E) MtlNhA:NADH, and (F) MtlNhA:NADH:THT tetrameric ensembles.



Supplementary Table ST1. Contribution of the first and the sum of the first 50 eigenvectors to each system, according to the principal component analysis.

<i>Ensemble/System</i>	<i>First eigenvector (%)</i>			<i>Sum of the 50 first eigenvectors (%)</i>		
	<i>Simulation 1</i>	<i>Simulation 2</i>	<i>Simulation 3</i>	<i>Simulation 1</i>	<i>Simulation 2</i>	<i>Simulation 3</i>
<i>Monomer:Apo</i>	41.63	26.29	22.05	91.98	56.18	85.78
<i>Monomer:NADH</i>	41.52	23.78	45.96	90.31	84.68	90.33
<i>Monomer:NADH:THT</i>	41.13	38.49	21.86	90.62	89.14	85.33
<i>Tetramer:Apo</i>	22.71	24.70	20.06	73.33	76.95	77.02
<i>Tetramer:NADH</i>	20.55	31.95	16.78	70.84	81.78	69.34
<i>Tetramer:NADH:THT</i>	19.87	25.86	17.26	72.36	74.50	73.11

Supplementary Table ST2. Statistical analysis of pincer angle distributions between monomeric and tetrameric structures.

Pincer angle

Ensembles	System	p.adjusted
Tetramer_A - Monomer	<i>Apo</i>	<0.01
Tetramer_B - Monomer	<i>Apo</i>	<0.01
Tetramer_C - Monomer	<i>Apo</i>	<0.01
Tetramer_D - Monomer	<i>Apo</i>	<0.01
Tetramer_A - Monomer	<i>NADH</i>	<0.01
Tetramer_B - Monomer	<i>NADH</i>	<0.01
Tetramer_C - Monomer	<i>NADH</i>	<0.01
Tetramer_D - Monomer	<i>NADH</i>	<0.01
Tetramer_A - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_B - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_C - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_D - Monomer	<i>NADH and substrate analogue</i>	<0.01

Supplementary Table ST3. Statistical analysis of area distributions between monomeric and tetrameric ensembles.

Area

Ensembles	System	p.adjusted
Tetramer_A - Monomer	<i>Apo</i>	<0.01
Tetramer_B - Monomer	<i>Apo</i>	<0.01
Tetramer_C - Monomer	<i>Apo</i>	<0.01
Tetramer_D - Monomer	<i>Apo</i>	<0.01
Tetramer_A - Monomer	<i>NADH</i>	<0.01
Tetramer_B - Monomer	<i>NADH</i>	<0.01
Tetramer_C - Monomer	<i>NADH</i>	<0.01
Tetramer_D - Monomer	<i>NADH</i>	<0.01
Tetramer_A - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_B - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_C - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_D - Monomer	<i>NADH and substrate analogue</i>	<0.01

Supplementary Table ST4. Statistical analysis of SB-loop and A-loop distances distributions between monomeric and tetrameric ensembles.

Distance of SB-loop and A-loop

Ensembles	System	p.adjusted
Tetramer_A - Monomer	<i>Apo</i>	<0.01
Tetramer_B - Monomer	<i>Apo</i>	<0.01
Tetramer_C - Monomer	<i>Apo</i>	<0.01
Tetramer_D - Monomer	<i>Apo</i>	<0.01
Tetramer_A - Monomer	<i>NADH</i>	<0.01
Tetramer_B - Monomer	<i>NADH</i>	<0.01
Tetramer_C - Monomer	<i>NADH</i>	<0.01
Tetramer_D - Monomer	<i>NADH</i>	<0.01
Tetramer_A - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_B - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_C - Monomer	<i>NADH and substrate analogue</i>	<0.01
Tetramer_D - Monomer	<i>NADH and substrate analogue</i>	<0.01

Supplementary Table ST5. Statistical analysis of B-loop and A-loop distances distributions between monomeric and tetrameric ensembles.

Distance of B-loop and A-loop

Ensembles	System	p.adjusted
Tetramer_A - Monomer	Apo	<0.01
Tetramer_B - Monomer	Apo	<0.01
Tetramer_C - Monomer	Apo	<0.01
Tetramer_D - Monomer	Apo	<0.01
Tetramer_A - Monomer	NADH	<0.01
Tetramer_B - Monomer	NADH	<0.01
Tetramer_C - Monomer	NADH	<0.01
Tetramer_D - Monomer	NADH	<0.01
Tetramer_A - Monomer	NADH and substrate analogue	<0.01
Tetramer_B - Monomer	NADH and substrate analogue	<0.01
Tetramer_C - Monomer	NADH and substrate analogue	<0.01
Tetramer_D - Monomer	NADH and substrate analogue	<0.01

Capítulo 3

Considerações finais

Perspectivas

4. CONSIDERAÇÕES FINAIS

A enzima InhA de *Mycobacterium tuberculosis* (Mtb), ou 2-trans-enoil-ACP (CoA) redutase (EC 1.3.1.9) codificada pelo gene *inhA*, tem como estrutura biologicamente ativa um homotetrâmero, sendo que cada subunidade é composta por 269 aminoácidos. A InhA tem por preferência substratos de cadeias carbônicas longas que variam de 16 a 56 carbonos, evidenciando a flexibilidade tanto da cavidade de ligação ao substrato quanto da estrutura desta enzima.

Estudar como as diferentes conformações estruturais que afetam esta enzima, proporciona uma melhor percepção de como esse processo pode influenciar na função catalítica da enzima e quais fatores podem estar envolvidos no processo de abertura e fechamento da sua cavidade de ligação. Esta dinâmica estrutural é fundamental para a técnica de reconhecimento biomolecular, ou seja, a compreensão da bioquímica de processos biológicos.

Este trabalho apresentou os resultados das simulações de dinâmica molecular clássica, a qual é utilizada para estudar a flexibilidade estrutural e função de macromoléculas biológicas em diferentes escalas de tempo e em nível atômico. A influência da estrutura quaternária da enzima InhA de Mtb foi analisada para uma melhor compreensão da sua flexibilidade. Foram realizadas três simulações para cada sistema: (i) InhA na forma livre (PDB ID 1ENY sem o NADH); (ii) InhA associada a coenzima NADH (PDB ID 1ENY); (iii) InhA associada ao NADH e análogo do substrato (PDB ID 1BVR). Todas as simulações foram realizadas para as formas monomérica e tetramérica. A construção da forma tetramérica da InhA foi feita por meio das matrizes translacionais e rotacionais depositadas no PDB. As estruturas conformacionais resultantes de cada simulação foram comparadas com estruturas consideradas referência, sendo estas, as estruturas cristalográficas 1ENY e 1BVR, também sendo comparadas estruturas monoméricas com estruturas tetraméricas.

Os objetivos propostos no trabalho foram realizados com sucesso. Os resultados demonstraram que há diferença estatisticamente significativa (p-values <0.01) entre os *ensembles* conformacionais da estrutura monomérica e tetramérica, para os três diferentes tratamentos. Com a utilização de diferentes parâmetros, foi visto que na forma apo, a estrutura monomérica se mostrou mais flexível em todas as três regiões (alça de ligação ao substrato, alça A e alça B), já a estrutura

tetramérica se mostrou mais rígida nas regiões de alça (A e B), devido estar em contato com subunidades adjacentes. Quando a InhA esta associada ao NADH, a estrutura monomérica apresenta uma conformação mais compacta da cavidade de ligação do substrato. No complexo ternário foi observada um conformação aberta do sítio de ligação do substrato, sugerindo que esta mudança conformacional poderia facilitar a entrada e/ou saída de substratos externos (Kumar e Sobhia, 2013). Todos os resultados foram obtidos por meio da técnica de dinâmica molecular clássica. Além disso, foi possível observar que as subunidades do tetrâmero também podem apresentar *ensembles* conformacionais distintos, exceto as subunidades C, B, D do sistema livre e C, B do sistema NADH:THT que não apresentam diferenças significativas.

Com este trabalho foi possível relatar as mudanças conformacionais que ocorrem na enzima InhA para melhor compreender a afinidade de ligação à pequenas moléculas e, principalmente, investigar a influência que a estrutura quaternária exerce em relação a flexibilidade da cavidade de ligação do substrato. O estudo sobre a flexibilidade desta enzima pode servir como um parâmetro para analisar a interação proteína-ligante, a fim de encontrar moléculas que apresentem uma maior afinidade pela cavidade do substrato da InhA, e com isso, desenvolver novas drogas capazes de combater a tuberculose.

REFERÊNCIAS

- da Costa, A. L., I. Pauli, M. Dorn, E. K. Schroeder, C. G. Zhan and O. N. de Souza (2012). "Conformational changes in 2-trans-enoyl-ACP (CoA) reductase (InhA) from *M. tuberculosis* induced by an inorganic complex: a molecular dynamics simulation study." J Mol Model **18**(5): 1779-1790.
- Dessen, A., A. Quemard, J. S. Blanchard, W. R. Jr. Jacobs and J. C. Sacchettini (1995). "Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*." Science **267**(5204): 1638.
- Feixas, F., S. Lindert, W. Sinko and J. A. McCammon (2014). "Exploring the role of receptor flexibility in structure-based drug discovery." Biophys Chem **186**: 31-45.
- Halle, B. (2004). "Biomolecular cryocrystallography: structural changes during flash-cooling." Proc Natl Acad Sci U S A **101**(14): 4793-4798.
- Koshland, D. E. (1958) "Application of a theory of enzyme specificity to protein synthesis." Proc Natl Acad Sci **44**: 98-104
- Kumar, V. and M. E. Sobhia (2013). "Characterisation of the flexibility of substrate binding loop in the binding of direct InhA inhibitors." Int J Comput Biol Drug Des **6**(4): 318-342.
- Kumar, V. and M. E. Sobhia (2014). "Insights into the bonding pattern for characterizing the open and closed state of the substrate-binding loop in *Mycobacterium tuberculosis* InhA." Future Med Chem **6**(6): 605-616.
- Lima, C. H. D. S., R. B. de Alencastro, C. R. Kaiser, M. V. N. de Souza, C. R. Rodrigues and M. G. Albuquerque (2015). "Aqueous molecular dynamics simulations of the *M. tuberculosis* enoyl-ACP reductase-NADH system and its complex with a substrate mimic or diphenyl ethers inhibitors." Int J Mol Sci **16**(10): 23695-23722.
- Luckner, S. R., N. Liu, C. W. Am Ende, P. J. Tonge and C. Kisker (2010). "A slow, tight binding inhibitor of InhA, the enoyl-acyl carrier protein reductase from *Mycobacterium tuberculosis*." J Biol Chem **285**(19): 14330-14337.
- Pauli, I., R. N. dos Santos, D. C. Rostirolla, L. K. Martinelli, R. G. Ducati, L. F. Timmers, L. A. Basso, D. S. Santos, R. V. Guido, A. D. Andricopulo and O. Norberto de Souza (2013). "Discovery of new inhibitors of *Mycobacterium tuberculosis* InhA enzyme using virtual screening and a 3D-pharmacophore-based approach." J Chem Inf Model **53**(9): 2390-2401.
- Persson, B., Y. Kallberg, U. Oppermann and H. Jörnvall (2003). "Coenzyme-based functional assignments of short-chain dehydrogenases/reductases (SDRs)." Chem Biol Interact **143**: 271-278.
- Qiu, X., S. S. Abdel-Meguid, C. A. Janson, R. I. Court, M. G. Smyth and D. J. Payne (1999). "Molecular basis for triclosan activity involves a flipping loop in the active site." Protein Sci **8**(11): 2529-2532.
- Quemard, A., J. C. Sacchettini, A. Dessen, C. Vilcheze, R. Bittman, W. R. Jacobs Jr and J. S. Blanchard (1995). "Enzymic characterization of the target for isoniazid in *Mycobacterium tuberculosis*." Biochemistry **34**(26): 8235-8241.

Rozwarski, D. A., G. A. Grant, D. H. Barton, W. R. Jacobs and J. C. Sacchettini (1998). "Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*." Science **279**(5347): 98-102.

Rozwarski, D. A., C. Vilcheze, M. Sugantino, R. Bittman and J. C. Sacchettini (1999). "Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate." J Biol Chem **274**(22): 15582-15589.

Schrödinger, LLC (2010). "The PyMOL Molecular Graphics System." Version 1.3r1.

Schroeder, E. K., L. A. Basso, D. S. Santos and O. Norberto de Souza (2005). "Molecular dynamics simulation studies of the wild-type, I21V, and I16T mutants of isoniazid-resistant *Mycobacterium tuberculosis* enoyl reductase (InhA) in complex with NADH: toward the understanding of NADH-InhA different affinities." Biophys J **89**(2): 876-884.

Shinde, R. N. and M. E. Sobhia (2013). "Binding and discerning interactions of PTP1B allosteric inhibitors: novel insights from molecular dynamics simulations." J Mol Graph Model **45**: 98-110.

Teilum, K., J. G. Olsen and B. B. Kragelund (2009). "Functional aspects of protein flexibility." Cell Mol Life Sci **66**(14): 2231.

Vilchèze, C., H. R. Morbidoni, T. R. Weisbrod, H. Iwamoto, M. Kuo, J. C. Sacchettini and W. R. Jacobs, (2000). "Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of *Mycobacterium smegmatis*." J. Bacteriol **182**(14): 4059-4067.

Weikl, T. R. and F. Paul (2014). "Conformational selection in protein binding and function." Protein Sci **23**(11): 1508-1518.

Yesylevskyy, S. O., V. N. Kharkyanen and A. P. Demchenko (2006). "The change of protein intradomain mobility on ligand binding: is it a commonly observed phenomenon?." Biophys J **91**(8): 3002-3013.