

**Marcio Vinicius Bertacine Dias**

**Estudo Estrutural de Proteínas Alvo  
de *Mycobacterium tuberculosis***



Tese apresentada ao Departamento de Física do Instituto de Biociências, Letras e Ciências Exatas (IBILCE) da Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP) para obtenção de título de Doutor em Biofísica Molecular.

**Orientador: Prof. Dr. João Ruggiero Neto**  
**Co-orientador: Prof. Dr. Walter Filgueira de Azevedo Jr.**

O presente trabalho foi realizado no Laboratório de Sistemas Biomoleculares (BMSys), Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas/IBILCE de São José do Rio Preto, SP, Universidade Estadual Paulista/UNESP, sob orientação do Prof. Dr. João Ruggiero Neto e co-orientação do Prof. Dr. Walter Filgueira de Azevedo Júnior, com auxílio da Fundação de Amparo à Pesquisa do Estado de São Paulo/FAPESP (processo 03/12472-2).

*À minha amada e querida mãe, Catarina,  
E em memória ao meu inestimável pai, David,*

*Dedico este trabalho aos amigos*

*Aos que se tornaram familiares,  
Aos que nasceram familiares  
E aos que conheci antes de ontem.*

*Dedico tanto aos que me deixam louco,  
Quanto aos que enlouqueço.*

*Aos que me criticam em tudo,  
E a um ou outro que atura  
Minha "chatura".*

*Aos amigos que correm,  
Aos amigos que contemplam.*

*Aos que me consideram muito,  
E aos muitos que, com razão, fazem pouco.*

*Aos que conhecem o que penso,  
E aos que só conhecem o que faço.*

*Aos que passam o dia todo comigo,  
E aos que estão todo tempo em mim.*

*Este trabalho é a soma de todos vocês.  
E se ele não é melhor,  
É por falta de memória,  
Mas não por falta de amigos.*

*(autor desconhecido)*

## *Agradecimentos:*

*Ao final deste trabalho são inúmeras as pessoas que eu devo o meu sincero obrigado, mas entre estas as que me veio à mente no momento que eu escrevi esta parte do trabalho, sem serem mais ou menos importantes as que infelizmente, pelo meu egoísmo ou esquecimento, deixei de colocar neste local, são:*

*Ao meu orientador Prof. Dr. João Ruggiero Neto, por ter me socorrido no momento que mais precisei; Pela amizade e confiança depositada em mim;*

*Ao meu co-orientador, Prof. Dr. Walter F. de Azevedo Jr., pela amizade, orientação, ensinamentos, oportunidades e confiança depositada em mim;*

*Ao Prof. Dr. Valmir Fadel, pela amizade, apoio, conselhos e confiança depositada em mim e pelo auxílio nas coletas realizadas no LNS;*

*À Fernanda Canduri, pela amizade, apoio, conselhos, sendo que mesmo distante sempre esteve perto nos momentos que eu mais precisei;*

*Às minhas ex-orientadoras, Profas. Dras. Marlene K. H. Kobayashi, Hermione Elly Melara de Campos Bicudo, Ana Elizabete Silva e Fernanda Canduri, por a cada uma de sua maneira, por todos os conhecimentos transmitidos e por me despertarem o gosto pela ciência;*

*Às minhas queridas “alunas”, Brianna, Lúvia, Michele e Samantha, que estiveram sempre comigo me ensinando a ensinar e sempre trabalhando com dedicação e afinco no Laboratório. E que sem essas maravilhosas pessoas grande parte deste trabalho talvez não teria sido efetuada;*

*Ao meu grande amigo José Henrique Pereira, por toda amizade, conselhos, e ajuda tanto na vida profissional como pessoal. A sua esposa, Adriana e sua família, que se tornou uma extensão da minha própria família;*

*Ao Prof. Dr. Jorge Chahine e a Profa. Dra. Fernanda Canduri, por terem participado da minha banca de qualificação;*

*Aos Prof(a)s Dr(a)s. Beatriz Guimarães, João Alexandre Barbosa, Maria Célia Bertolini e Valmir Fadel por terem aceitado o convite em participar da minha banca de doutorado; E também aos Prof(a)s Dr(a)s Marcos Fontes, Maria Cristina Nonato e Jorge Chahine por terem aceitado o convite de serem suplentes da minha banca examinadora;*

*À Marisa, minha leal suportadora, que sempre me agüentou na sala de estudos. Pelos momentos de risadas, tristezas e desabaços; Pela amizade e sinceridade que sempre teve comigo;*

*Aos Profs. Drs. Diógenes Santiago Santos e Luiz Augusto Basso, pelo fornecimento de proteínas;*

*Ao pessoal da PUC-Rio Grande do Sul, pela amizade e pela expressão das proteínas fornecidas, em especial, Fernanda Ely, Igor, Jaím, Ana Luíza, Patrícia, e Jordana;*

*À grande amiga Márcia, que sempre esteve disposta a me ajudar e me escutar. E que fez também o ambiente de trabalho ser mais gostoso e profissional;*

*À outra grande amiga Denise Mello, que também sempre fez a minha vida se tornar mais alegre;*

*À Daiana e Angélica por todo o carinho, simpatia e confiança;*

*A todos os alunos do Grupo BMSys tanto os atuais como os que hoje não freqüentam mais (Pepeu, Diego, Denis, Janaina, Nathália, Nelsinho, Nelsão, Henrique, Marisa, Galo, Paulo, Alexandre, Hugo, Helen, Alessandra, Bona, Joane, Guilherme, Daiana, Ana Helena, Karisa, José Renato, Cristiane, Mariane, Lisandra, Marcos Michel, Brianna, Angélica, Lívia, Michele e Samantha), pelas gargalhadas, risadas, choros, brigas e tudo mais que faz parte das novas vidas;*

*Aos grandes amigos da pós-graduação, por tornar a vida cotidiana menos monótona e solitária (Priscila, Gisele, Magno, Flávio, Sabrina, Jorge, Eduardo, Sidney, Leandro, Ronaldo, Ricardo, José Ézio, Aperreio, Cristina, Fernanda, André, Renato);*

*Ao Prof. Dr. Jorge Chahine, pelo empenho com a pós-graduação e pelo exemplo de pessoa que é;*

*Aos demais professores do Departamento pelos ensinamentos transmitidos;*

*Aos funcionários técnico-administrativos, que sempre estão correndo para nos ajudar com os problemas menores, mas que fazem toda a diferença;*

*À Grande Rosemar, por toda a simpatia que é, e sempre correndo para agilizar as coisas para os alunos na seção de pós-graduação. Aos demais funcionários da Seção de pós-graduação por sempre terem contribuído aos alunos;*

*Aos professores da graduação, por ter me feito um biólogo e me dar à visão de mundo que tenho hoje, em especial a Cláudia Carareto e Eleni, por quem tenho grande admiração;*

*Aos funcionários do Laboratório Síncrotron, pela ajuda nos equipamentos de raios X, em especial ao Lucas,;*

*Aos amigos eternos da graduação, que mesmo distante no momento, eu sei que estarão torcendo por mim, principalmente a Ana Paula, Marcela, Leandra, Aline, Amanda e Renata e Gustavo;*

*A todos os demais amigos conquistados durante todos estes anos de IBJLCE;*

*Aos amigos conquistados em Fortaleza, em especial ao Gustavo, Jaianá e Vitor, por me ajudarem quando precisei;*

*Aos amigos conquistados em São Paulo, pelos ensinamentos transmitidos em “virtual screen”, em especial, a profa. Dra Antônia Favares e ao Dr. Humberto;*

*À Andréia, pela amizade e carinho que tem por mim;*

*Aos meus amigos de Fernandópolis, que com certeza estarão torcendo por mim, não importa onde eu ou eles estejam, principalmente ao Ricardo, Rosa, Vagner e Adriana, Fernando e Miriam, Rogério e Elisângela;*

*À minha família, em especial a minha querida e amada mãe, do qual sem ela nada sou. Que é o suporte para toda a minha vida;*

*À Fapesp pelo auxílio financeiro e concessão da bolsa de doutorado;*

*A todas as pessoas que de alguma forma contribuiriam para a minha formação ou para a realização deste trabalho, inclusive aquelas que por força do destino não tinham simpatia por mim ou eu por elas, mas que contribuiriam para o meu crescimento pessoal e profissional;*

*E finalmente a Deus por minha existência;*

*“A maior recompensa do nosso trabalho  
não é o que nos pagam por ele,  
mas aquilo que ele nos transforma”  
John Ruskin*

*“-Podes dizer-me, por favor, que caminho devo seguir para sair daqui?  
-Isso depende muito de para onde queres ir – respondeu o gato.  
-Preocupa-me pouco aonde ir – disse Alice.  
-Nesse caso, pouco importa o caminho que sigas – replicou o gato.”  
(Lewis Carroll – Alice no país das Maravilhas)*



## Lista de Abreviações

$\mu\text{M}$	micro molar
ACP	<i>Acyl carrier protein</i> - Proteína carreadora de acila
ADP	Adenosina difosfato
AIDS	<i>Acquired immunodeficiency syndrome</i> - Síndrome da Imunodeficiência Adquirida
ATP	Adenosina trifosfato
BCG	Bacilo Calmette Ghérin
C	Carbono
CQ	Chiquimato quinase
DAHP	Deoxi-D-arabino-heptulose-7-fosfato
DHFR	Enzima Dihidrofolato redutase
E4P	Eritrose-4-fosfato
EPSP	5-enoil-piruvil-chiquimato-3-fosfato
EPSPS	5-enoil-piruvil-chiquimato-3-fosfato sintase
ESB	<i>Extended shikimate binding</i> -domínio de ligação do chiquimato estendido
FAS I	Sistema de síntese de ácidos graxos tipo I
FAS II	Sistema de síntese de ácidos graxos tipo II
FMN	Flavina mononucleotídeo
G3P	D-gliceraldeído-3-fosfato
HCl	Ácido clorídrico
Hepes	Ácido N-(2-hidroxietil)piperazina-N'-(2-etanosulfônico)
HIV	<i>Human Immunodeficiency virus</i> (Vírus da Imunodeficiência Humana)
IGP	Indol glicerol-3-fosfato
INH	Isoniazida
InhA	Proteína Enoil ACP redutase
IPP	Indol propanol-3-fosfato
KatG	Proteína catalase peroxidase
$K_i$	Constante de inibição
$\text{Mg}^{2+}$	Íon Magnésio
MtCQ	Chiquimato quinase de <i>Mycobacterium tuberculosis</i>
N	Nitrogênio
Na	Sódio
NAD	Nicotinamida adenina difosfato

## ÍNDICE DE FIGURAS

Figura 1	<i>Mycobacterium tuberculosis</i> corado pelo método ácido-resistente .....	3
Figura 2	Representação esquemática da via metabólica do ácido chiquímico .....	5
Figura 3	Representação esquemática das ramificações da via do ácido chiquímico .....	7
Figura 4	Reação catalisada pela chiquimato quinase .....	9
Figura 5	Representação esquemática da chiquimato quinase .....	10
Figura 6	Domínios estruturais para a chiquimato quinase .....	11
Figura 7	Alterações estruturais e efeito sinérgico observado na chiquimato quinase .....	13
Figura 8	Reação catalisada pela corismato sintase .....	15
Figura 9	Representação esquemática da corismato sintase .....	15
Figura 10	Completo mecanismo catalítico para a corismato sintase .....	17
Figura 11	Representação esquemática da via de síntese de triptofano .....	20
Figura 12	Reação catalisada pela triptofano sintase .....	21
Figura 13	Estrutura tridimensional da triptofano sintase .....	22
Figura 14	Formula molecular dos ligantes para triptofano sintase .....	24
Figura 15	Representação esquemática do envelope celular de micobactérias .....	25
Figura 16	Reação catalisada pela enzima InhA .....	26
Figura 17	Representação esquemática da InhA .....	27
Figura 18	Mecanismo de ativação da isoniazida .....	28
Figura 19	Contatos moleculares entre o isonicotínico acil-NADH e o sítio ativo da InhA .....	29

<b>NADH-INH</b>	<b>Complexo de nicotinamida adenina dinucleotídeo e isoniazida</b>
<b>NADP</b>	<b>Nicotinamida adenina dinucleotídeo fosfato</b>
<b>NB</b>	<i>Nucleotide binding</i> – referente ao domínio de ligação do nucleotídeo
<b>NMP</b>	<b>Mononucleotídeo fosfato</b>
<b>OMS</b>	<b>Organização Mundial de Saúde</b>
<b>PDB</b>	<i>Protein Data Bank</i> – Banco de dados de proteína
<b>PEG</b>	<b>Polietileno glicol</b>
<b>PEP</b>	<b>Fosfoenol piruvato</b>
<b>Pi</b>	<b>Fosfato inorgânico</b>
<b>PLP</b>	<b>Piridoxal fosfato</b>
<b>RC</b>	<i>Reduced core</i> – referente ao domínio central reduzido
<b>SB</b>	<i>Shikimate binding</i> – referente ao domínio de ligação do chiquimato
<b>SDR</b>	<i>Short dehydrogenase/reductase</i> – desidrogenase/reductase de cadeia curta
<b>SpCS</b>	<b>Corismato sintase de <i>Streptococcus pneumoniae</i></b>
<b>TIM</b>	<b>Triose fosfato isomerase</b>
<b>Tris</b>	<b>Tris(hidroximetil)aminometano</b>
<b>TRPS</b>	<b>Triptofano sintase</b>

## RESUMO

A tuberculose representa hoje um dos principais problemas de saúde pública mundial. É estimado que cerca de cem milhões de pessoas sejam infectadas anualmente. Para este grande número de casos, dois fatores têm contribuído significativamente, a resistência da *Mycobacterium tuberculosis* aos antibióticos existentes e o aumento de casos de co-infecção com HIV. Por isso, é importante o desenvolvimento de novas drogas contra o seu agente causador. Há duas abordagens principais para o desenvolvimento de drogas. Primeiramente pode-se identificar e inibir proteínas que estejam presentes em uma determinada classe de microorganismos, mas não sejam conservadas em humanos, levando a antibióticos de largo espectro, ou ainda pode-se identificar e inibir proteínas de um determinado microorganismo, levando a drogas específicas. Exemplos da primeira abordagem são as enzimas da via do ácido chiquímico e suas ramificações, como a via de síntese de triptofano. A via do ácido chiquímico é responsável pela biossíntese de corismato, que é o precursor comum utilizado na geração de vários compostos aromáticos importantes para sobrevivência da bactéria, entre eles os aminoácidos e co-enzimas; e exemplos da segunda abordagem são enzimas relacionadas com a síntese de ácidos micólicos, que são importantes constituintes da parede celular de micobactérias. O objetivo do presente trabalho foi realizar estudos estruturais de quatro proteínas que são alvos em potencial para o desenvolvimento de drogas contra tuberculose, sendo elas: a Chiquimato quinase e Corisamto sintase, ambas da via do ácido chiquímico; Triptofano sintase, última enzima da via de síntese de triptofano; e a enzima dependente de NADH enoil ACP redutase (InhA), relacionada com a síntese de ácidos micólicos. Neste trabalho são apresentadas as estruturas da Chiquimato quinase de *M. tuberculosis* em complexo com ADP e magnésio, na ausência do substrato ácido chiquímico, e da Chiquimato quinase em

complexo com ADP e ácido chiquímico, na ausência do íon magnésio. Nestas estruturas pôde-se observar o efeito de íons e a influência do ácido chiquímico sobre a estrutura da chiquimato quinase. Apresentamos também, a primeira estrutura da corismato sintase de *Mycobacterium tuberculosis*, evidenciando suas características estruturais. Além disso, foram realizados estudos estruturais com a proteína InhA na sua forma nativa, mostrando o seu estado conformacional na ausência de ligantes, e três estruturas para esta proteína em complexo com isoniazida ativa, sendo a proteína selvagem e dois mutantes identificados clinicamente como resistentes a este medicamento (S94A e I21V). Além desses estudos, foi realizada a modelagem molecular da proteína triptofano sintase em complexo com seis inibidores, e feita análise da especificidade destes com relação a esta proteína. Com a realização deste trabalho pretendemos estar contribuindo para a solução do grande problema que a tuberculose representa hoje, por afligir milhões de pessoas em todo o mundo.

## ABSTRACT

Tuberculosis represents today a large problem of world public health. It is estimated that approximately a hundred million people are infected annually. For this large number of cases, two factors have contributed significantly, the resistance of *Mycobacterium tuberculosis* to existent antibiotics and the increase of the cases of the co-infection with HIV. Therefore, the development of new drugs against its etiologic agent is important. There are two main approaches for the development of drugs. Firstly, proteins of a determined class of microorganism, that are not present in humans, can be identified and inhibited leading to large-spectra antibiotics; or it can identify and inhibit proteins of a determined organism leading to specific drugs. Examples of the first approach are enzymes of the shikimate pathway and its branches, such as the tryptophan synthesis pathway. The shikimate pathway is responsible for biosynthesis of chorismate, it is a common precursor utilized in the production of various important aromatic compounds used in the survival of the bacteria, such as amino acids and co-enzymes; an example of the second approach are enzymes related to the synthesis of mycolic acids, they are important constituents of the cellular wall of mycobacteria. The objective of the present work was to perform structural studies of four enzymes that are potential targets to the development of drugs against tuberculosis, such as: shikimate kinase and chorismate synthase, both of the shikimate pathway; tryptophan synthase, last enzyme of the tryptophan pathway; and the dependent of NADH enzyme enoyl ACP reductase (InhA) related to the synthesis of mycolic acids. Here, the shikimate kinase structures from *M. tuberculosis* in complex with ADP and magnesium ion, in the absence of shikimate and the shikimate kinase in complex with ADP and shikimate, in the absence of the magnesium ion, are presented. The effects of ions and shikimate on the structure of shikimate kinase were observed. Here, we also present the first structure

of chorismate synthase from *M. tuberculosis* showing its structural characteristics. Furthermore, we perform structural studies of the InhA protein in its native form showing its conformational state in the absence of ligands and three structures for this protein in complex with active isoniazid (wild type and two resistant clinical isolated mutants to this medicine, S94A and I47T). We also performed the molecular modeling of the tryptophan synthase protein in complex with six inhibitors, and the analysis of the specificity of these in relation to the protein was performed. With the completion of this work, we intend to contribute to the solution of the big program that tuberculosis represents today causing suffering to the millions of people worldwide.

## **SUMÁRIO**

<b>1. INTRODUÇÃO</b>	<b>1</b>
<b>1.1 A tuberculose e a necessidade do desenvolvimento de novas drogas</b>	<b>1</b>
<b>1.2 A via do ácido chiquímico</b>	<b>5</b>
<b>1.3 A chiquimato quinase</b>	<b>8</b>
<b>1.4 A corismato sintase</b>	<b>14</b>
<b>1.5 A triptofano sintase</b>	<b>18</b>
<b>1.6 InhA e o mecanismo de resistência ao antibiótico isoniazida</b>	<b>24</b>
<b>2. OBJETIVOS</b>	<b>30</b>
<b>3. RESULTADOS</b>	<b>31</b>
<b>3.1 Crystallization and preliminary X-ray crystallographic analysis of chorismate synthase from <i>Mycobacterium tuberculosis</i></b>	<b>32</b>
<b>3.2 Structure of chorismate synthase from <i>Mycobacterium tuberculosis</i></b>	<b>36</b>
<b>3.3 Molecular models of tryptophan synthase from <i>Mycobacterium tuberculosis</i> complexed with inhibitors</b>	<b>52</b>
<b>3.4 Effects of the magnesium and chloride ions and shikimate on the structure of shikimate kinase from <i>Mycobacterium tuberculosis</i></b>	<b>64</b>



<b>3.5 Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-trans- Enoyl-ACP (CoA) Reductase from <i>Mycobacterium tuberculosis</i></b>	<b>72</b>
<b>4. CONCLUSÕES</b>	<b>118</b>
<b>5. BIBLIOGRAFIA</b>	<b>121</b>
<b>6. ANEXO – Artigo de revisão publicado</b>	<b>126</b>
<b>6.1 Chorismate Synthase: an attractive target for drug development against neglected disease</b>	<b>127</b>

# 1. INTRODUÇÃO

## 1.1 A tuberculose e a necessidade do desenvolvimento de novas drogas

As doenças infecciosas é a quarta maior causa de morte da população humana (KING et al., 2006), sendo responsável pelo sofrimento e morte de centenas de milhões de pessoas, especialmente em áreas tropicais e subtropicais do mundo, local onde ocorre cerca de 90% das mortes causadas por esse tipo de enfermidade (TROUILLER et al., 2001). Segundo a Organização Mundial de Saúde (OMS) cerca de 14 milhões de pessoas morrem a cada ano devido a desse tipo de doença, destacando-se entre elas AIDS/HIV, infecções respiratórias, malária e tuberculose, que lideram a causa de morte e de morbidade, principalmente nos países da África, Ásia e América do Sul – regiões nas quais estão presentes oitenta por cento de toda a população mundial (TROUILLER et al., 2001).

A organização mundial de saúde identifica três fatores chaves que podem contribuir coletivamente para a pandemia associada às doenças infecciosas. São eles: falência no uso de ferramentas existentes no controle de tais doenças, ferramentas inadequadas ou inexistentes, e conhecimento insuficiente sobre a doença (TROUILLER et al., 2002).

Dentre essas doenças, a tuberculose, hoje, é considerada a causa infecciosa líder de mortes. É estimado que cem milhões de pessoas são infectadas anualmente. De cem milhões, aproximadamente dez milhões desenvolvem a doença e três milhões morrem (MORGAN; HARITAKUL; KELLER, 2003).

A incidência de tuberculose teve um declínio rápido no início do século vinte nos países em desenvolvimento, devido à melhora nas condições sanitárias e de moradias. Essa tendência foi, inicialmente, acelerada pela introdução da vacinação BCG (1927) e da descoberta de antibióticos como a estreptomicina (1944) e, posteriormente, com a descoberta do ácido p-aminosalicílico (1946), isoniazida (1952) e rifampicina (1965) (DUNCAN, 2003). Mas está ocorrendo uma ressurgência dos casos de tuberculose nos

países em desenvolvimento. Muitos são os fatores envolvidos, porém, o mais notável é a falta de recursos e o descaso governamental que impedem a implementação adequada de medidas de controle. Além disso, a incidência aumentou também nos países desenvolvidos, principalmente devido à imigração e à negligência direta do mundo todo com relação ao problema (RAVIGLIONE et al., 1997). Todavia, dois recentes eventos têm contribuído significativamente para a pandemia de tuberculose: a) o aumento de bactérias multi-resistentes devido a terapias inadequadas e o uso indiscriminado de antibióticos, principalmente a isoniazida e rifampicina, as duas drogas mais utilizadas no tratamento quimioterápico (BAPTISTA et al., 2002; DUNCAN, 2003); e b) a AIDS, pois pessoas portadoras de ambos os agentes causadores da tuberculose e AIDS têm o sistema imune debilitado, e, portanto, maiores chances de desenvolver a doença (BAPTISTA et al., 2002; DYE et al., 2002). Em 1993, a Organização Mundial de Saúde declarou a tuberculose como uma emergência global. Entretanto, a incidência exata desta doença no panorama mundial não é bem conhecida (RAVIGLIONE et al., 1997).

O principal responsável pela tuberculose em humanos é a bactéria *Mycobacterium tuberculosis*. Este organismo é um bastonete fino, com comprimento médio de 4 µm, como pode ser visto pela coloração por métodos ácido-resistentes (Fiehl-Neelsen) (Figura 1) ou fluorescentes (COTRAN et al., 1991). Apresenta crescimento lento, dormência e complexo envelope celular. É um organismo obrigatoriamente aeróbico e é um parasita intracelular facultativo, usualmente de macrófagos. O tempo de geração da *M. tuberculosis* em animais infectados é de aproximadamente vinte e quatro horas. Isto contribui para a característica crônica da doença (apesar da doença em si não apresentar uma fase crônica), impossibilitando fortemente o regime de tratamento e representando um formidável obstáculo para pesquisas (COLE et al., 1998). *M. tuberculosis* é quase exclusivamente um parasita humano e é transmitido pelo ar. Usualmente causa tuberculose pulmonar, uma

infecção dos pulmões, embora, seja capaz de infectar outros tecidos ou órgãos do corpo (PELCZAR et al., 1996).



**Figura 1. *Mycobacterium tuberculosis* corado pelo método ácido-resistente (Fiehl-Neelsen).**

Clinicamente a tuberculose pode ser classificada como primária e secundária. Sendo que a tuberculose primária é caracterizada por uma forma de doença que se desenvolve previamente numa pessoa não-exposta e, portanto, não-sensibilizada. A tuberculose primária inicia-se pela inalação do *M. tuberculosis* e culmina com o desenvolvimento da imunidade mediada por células no organismo. Na maioria das pessoas, a tuberculose primária é assintomática, ainda que possa causar febre e derrame pleural. Das pessoas recém-infectadas somente 5% desenvolvem a doença em uma forma clinicamente significativa. A tuberculose secundária é um padrão de doença que observa em um hospedeiro previamente sensibilizado. Pode surgir pouco depois da tuberculose primária, porém mais frequentemente, ela surge da reativação das lesões primárias latentes muitas décadas após a infecção inicial, particularmente quando a resistência do hospedeiro estiver enfraquecida. No caso da tuberculose pulmonar secundária, ela é geralmente localizada no ápice dos lóbulos superiores de um ou de ambos os pulmões. A tuberculose secundária localizada pode ser assintomática. Quando as manifestações aparecem, elas são, geralmente, insidiosas no início. Os sintomas sistêmicos aparecem, com freqüência no início do curso e incluem indisposição, anorexia, perda de peso e febre. Com o

envolvimento pulmonar progressivo, surgem quantidades cada vez maiores de escarro, primeiramente mucóide e posteriormente purulento. As manifestações extra pulmonares da tuberculose são em grande número e dependem do sistema orgânico envolvido (KUMAR, ABBAS, FASTO, 1999)

Com a determinação completa do genoma da *M. tuberculosis* ocorreu um profundo avanço em pesquisas sobre tuberculose (COLE et al., 1998). Novas abordagens e técnicas têm sido utilizadas para avaliar as informações obtidas pelo projeto genoma deste organismo. Entre elas, inclui a utilização de ferramentas de bioinformática, proteômica, transcriptoma e genômica estrutural combinadas com ferramentas de manipulação de genoma. Com a utilização destas técnicas tem-se permitido um melhor entendimento da complexa biologia deste patógeno (DUNCAN, 2003). Além disso, a determinação do genoma e o melhor esclarecimento dos mecanismos de virulência deste organismo, permitem a possibilidade da descoberta de novos alvos para o desenvolvimento de drogas (COLE et al., 1998; DUNCAN, 2003; KHASNOBIS; ESCUYER; CHATTERJEE, 2002).

Para a identificação de alvos com potencial de gerar novas drogas contra *M. tuberculosis* pode se utilizar duas abordagens significativas. A primeira, em que se podem isolar proteínas que são conservadas entre diferentes espécies de bactérias, mas que não conservadas em humanos, conduzindo aos antibióticos de largo espectro; e a segunda na qual se podem identificar alvos que são únicos em *M. tuberculosis*, levando a obtenção de drogas anti-tuberculose específicas (KHASNOBIS; ESCUYER; CHATTERJEE, 2002).

Exemplos da primeira abordagem são as enzimas que compõe a via de síntese de ácido corísmico (via do ácido chiquímico) e da segunda abordagem são as enzimas responsáveis pela síntese de ácidos micólicos.

## 1.2 A Via do Ácido Chiquímico

A via do ácido chiquímico (Figura 2) é essencial para algas, plantas, bactérias, fungos (BENTLEY, 1990) e parasitas do filo apicomplexa, como *Plasmodium*, *Toxoplasma* e *Cryptosporidium* (ROBERTS et al., 1998), e não está presente em mamíferos.

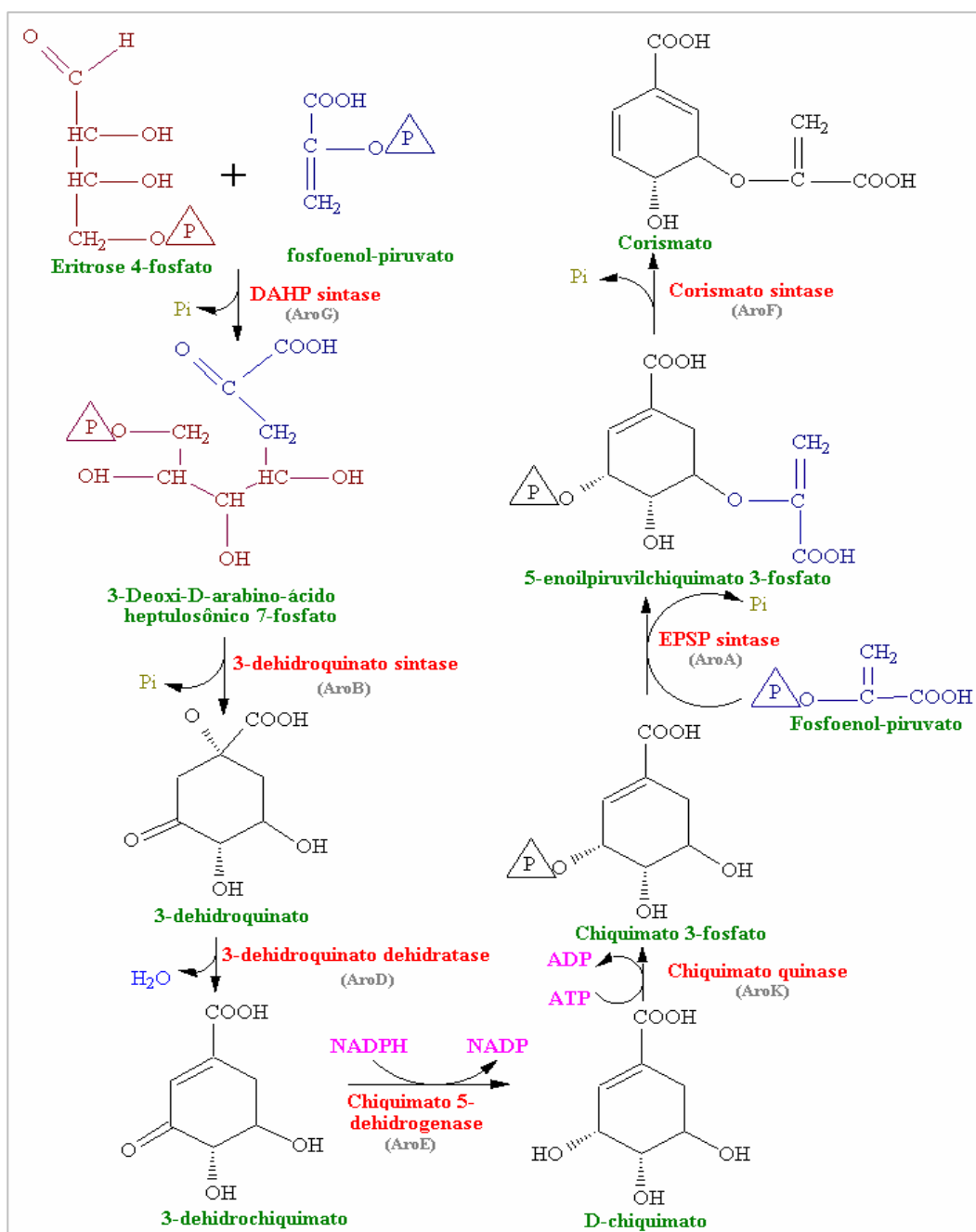
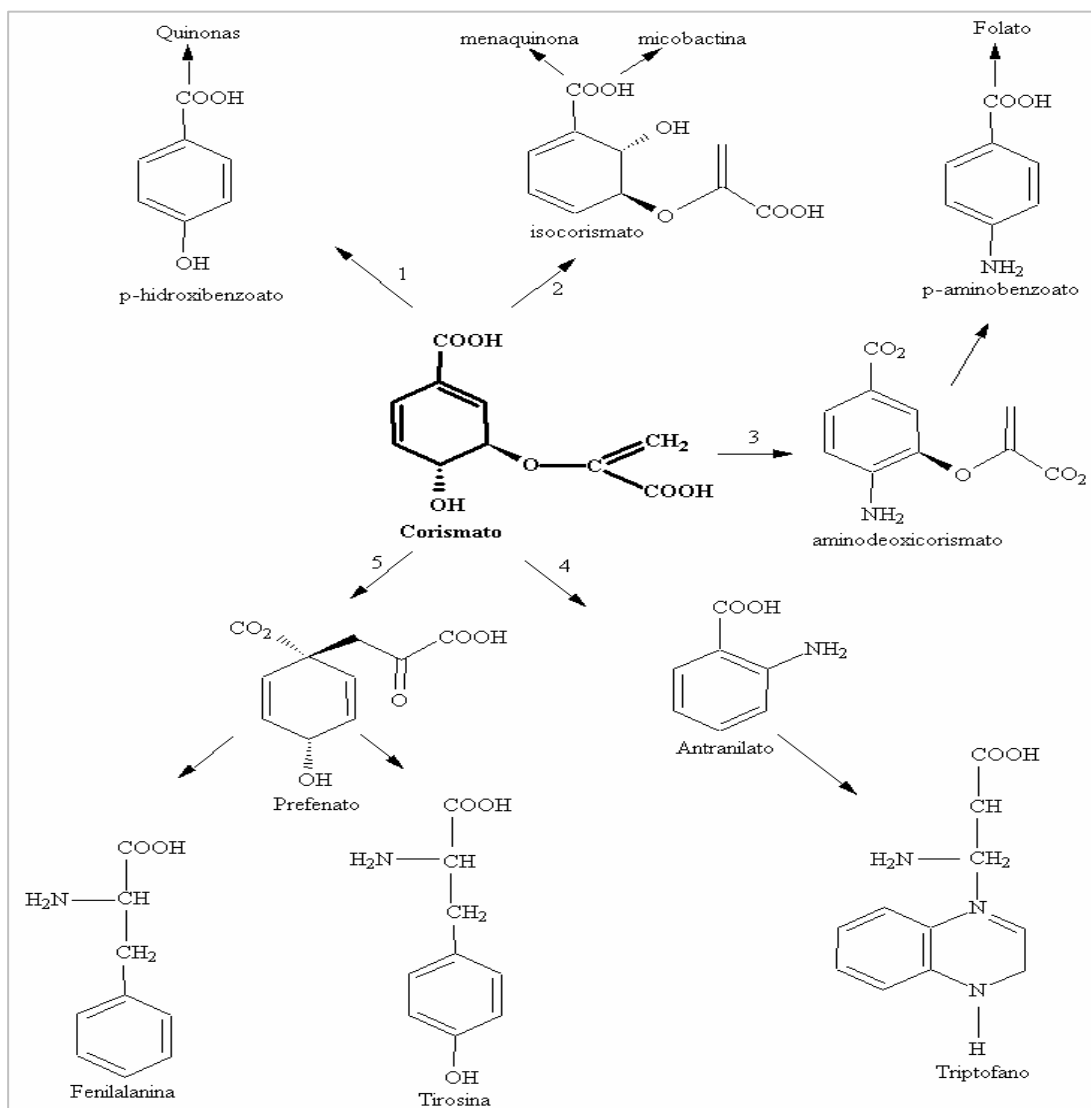


Figura 2. Representação esquemática da via metabólica do ácido chiquímico.

A via do ácido chiquímico é definida como os sete passos metabólicos iniciais da condensação de fosfoenol piruvato (PEP) e eritrose-4-fosfato (E4P) finalizando com a síntese de corismato. Os sete passos desta via foram originalmente descobertos através de estudos em bactéria, principalmente *Escherichia coli* e *Salmonella typhimurium* (HERRMANN; WEARVER, 1999).

A via do ácido chiquímico foi inicialmente descrita como a rota biossintética para a produção de aminoácidos aromáticos (Fenilalanina, Tirosina e Triptofano) através dos estudos clássicos de Bernhard Daves e David Sprinson e seus colaboradores por volta de 1950. A identificação de seus vários intermediários foi completada no início de 1960, sendo que o primeiro deles a ser identificado foi o ácido chiquímico (PITTARD, 1987; HERMANN, 1995). Em microorganismos, esta via, além de ser utilizada para síntese de aminoácidos aromáticos, também faz parte da síntese de folato, quinonas e alguns metabólitos secundários, como sideróforos e micobactinas. A via do ácido chiquímico ramifica-se em muitos pontos, porém, corismato é o último ramo comum da via para os compostos citados acima. Corismato é convertido por cinco enzimas distintas para pefenato, antranilato, aminodeoxicorismato, isocorismato e *p*-hidroxibenzoato. Esses metabólitos compreendem os primeiros intermediários na biossíntese de fenilalanina e tirosina, triptofano, folato, menaquinonas e sideróforos, e ubiquinonas, respectivamente (Figura 3) (DOSSELAERE; VANDERLEYDEN, 2001). Sendo que destes intermediários, folato é um composto que participa de várias formas na síntese de purinas, timidilato, metionina, glicina e pantotenato; Sideróforos são moléculas responsáveis pela captação de ferro por bactérias; Quinonas são moléculas lipofílicas, componentes não protéicos da cadeia transportadora de elétrons ligados à membrana que podem ser divididas em dois grandes grupos estruturais: benzoquinonas e naftoquinonas. As benzoquinonas são denominadas de ubiquinonas e as naftoquinonas são chamadas de menaquinonas (DOSSELAERRE; VANDERLEYDEN, 2001).



**Figura 3. Representação esquemática das ramificações da via do ácido chiquímico. (1) corismato piruvato-liase; (2) isocorismato sintase; (3) aminodeoxicorismato sintase; (4) antranilato sintase; (5) corismato mutase.**

A análise completa da seqüência do genoma da *M. tuberculosis* revelou a presença dos sete genes *aro* envolvidos na biossíntese de corismato. São eles, *aro D, B, K, F, G, E* e *A*. Estes genes são responsáveis pela síntese das enzimas 3-dehidroquinato desidratase, 3-dehidroquinato sintase, chiquimato quinase, corismato sintase, DAHP sintase, chiquimato-5-dehidrogenase e EPSP sintase, respectivamente (Figura 2) (COLE et al., 1998; PARISH; STOKER, 2002).

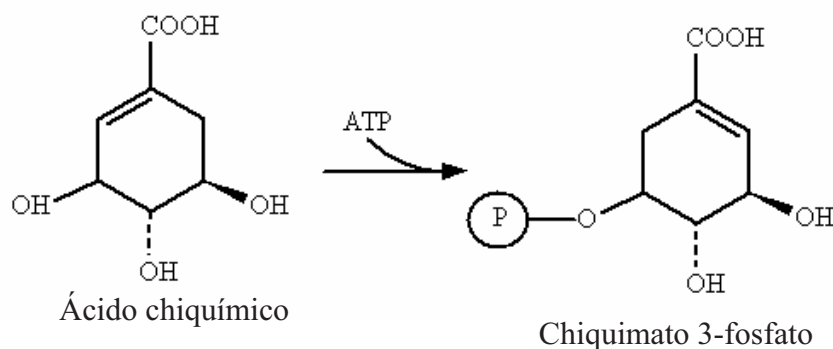


A importância desta via, tanto para *M. tuberculosis* como para outros patógenos foi justificada por diversos trabalhos: Parish e Stoker (2002) demonstraram que a via do ácido chiquímico é essencial para *M. tuberculosis*, pois, eles foram incapazes de isolar cepas contendo o gene *aroK* (responsável pela produção da chiquimato quinase) inativado. A inativação do gene *aroD* (responsável pela produção de 3-dehidroquinato desidratase) também é usada com sucesso para gerar cepas atenuadas para a vacina oral contra *Salmonella typhi* e outras bactérias (TACKET et al., 1997). Os estudos de inibição utilizando glifosato demonstram que este herbicida inibe a EPSP sintase, e apresenta uma significativa inibição no crescimento de *T. gondii*, *P. falciparum* e *C. parvum* (ROBERTS et al., 2002).

### 1.3 A Chiquimato quinase

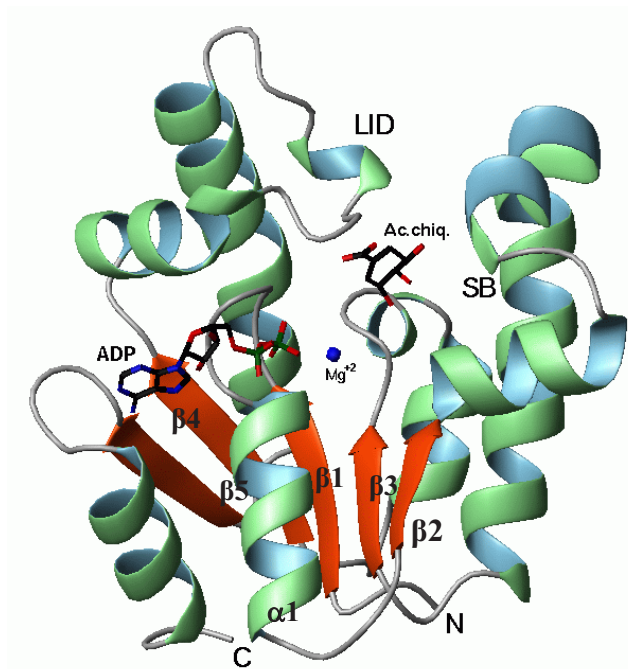
A Chiquimato quinase, principalmente de *Mycobacterium tuberculosis* (peso molecular de 18583,3 Da), é uma enzima que teve sua estrutura intensivamente estudada a partir de 2002 (GU et al., 2002, DHALIWAL et al., 2004, PEREIRA et al., 2004, GAN et al., 2006, HARTMANN et al., 2006). Esta proteína, que é, a quinta enzima da via do ácido chiquímico, catalisa a fosforilação do grupo 3-hidroxil do ácido chiquímico, usando ATP como um co-substrato (Figura 4) (HERRMANN, 1995) e os íons magnésio e cloreto como co-fatores. Em *Escherichia coli*, esta reação é catalisada por duas diferentes isoformas, sendo elas a CQ I, codificada pelo gene *aroK* e a CQ II codificada pelo gene *aroL*. Estas proteínas são monoméricas, possuem uma massa molecular de aproximadamente 19.500 Dáltons e apresentam uma identidade de 30% entre si. Ambas as enzimas possuem atividade *in vitro* na biossíntese de aminoácidos aromáticos. Mas *in vivo*, a isoenzima II não tem uma atividade exclusiva da via do ácido chiquímico, pois parece apresentar também uma função na divisão celular (HERRMANN; WEAVER, 1999; 1995; KRELL; COGGINS; LAPHORN, 1998). Diferente de *E. coli*, o genoma

completo de algumas bactérias, como *Mycobacterium tuberculosis* e *Haemophilus influenzae*, tem revelado a presença de apenas um gene para chiquimato quinase. Muitas destas chiquimato quinases parecem ser codificadas pelo gene *aroK*, pois sua seqüência de aminoácidos tem alta identidade com a CQ I de *E. coli* (GU et al., 2002).



**Figura 4. Reação catalisada pela chiquimato quinase.**

A primeira estrutura cristalográfica da chiquimato quinase de *M. tuberculosis* (*MtCQ*) foi resolvida na presença de Mg-ADP (GU et al., 2002). A *MtCQ* é uma proteína da classe  $\alpha/\beta$  que apresenta na sua estrutura fitas  $\beta$  rodeadas por hélices  $\alpha$ . No centro da proteína estão localizadas cinco fitas  $\beta$  paralelas ( $\beta 1$ - $\beta 5$ ) formando uma folha  $\beta$  na ordem 23145, que é flanqueada por oito hélices (Figura 5) (GU et al., 2002).

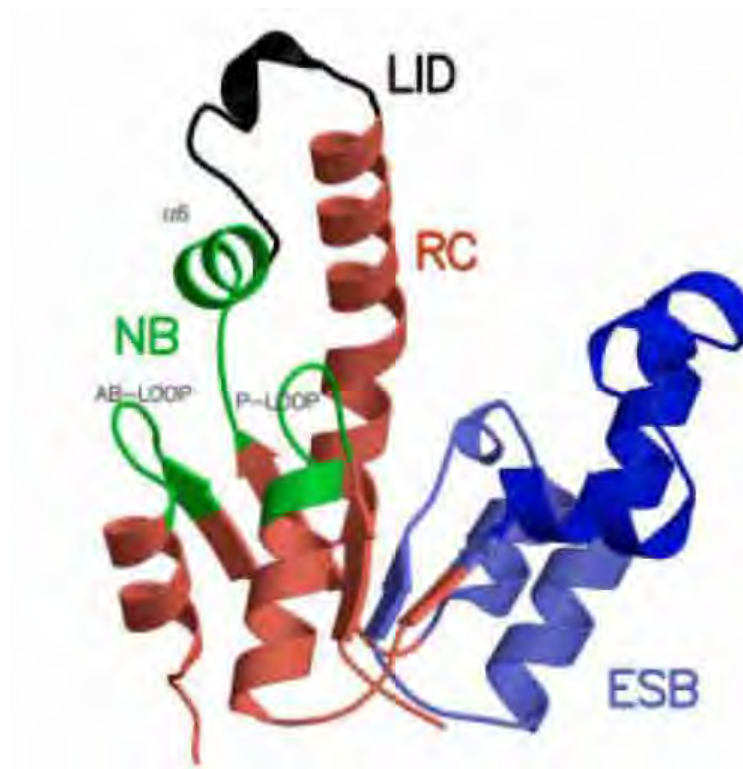


**Figura 5. Representação de elementos de estrutura secundária da *MtCQ* em complexo com Magnésio ( $Mg^{2+}$ ), Adenosina difosfato (ADP) e ácido chiquímico (PEREIRA et al., 2004) gerada pelo programa MolMol (KURADI; BILLETER; WÜTHRICH, 1996).**

A ordem da folha  $\beta$  23145 observada na estrutura da *MtCQ* classifica esta proteína como pertencente à família das monofosfato nucleosídeo (NMP) quinases (YAN; YSAI, 1999). As proteínas desta família são compostas por três domínios: central (local de ligação do fosfato), domínio da tampa (*Lid domain*) e o domínio de ligação do NMP ou SB (*shikimate binding*). A *MtCQ* possui uma seqüência de aminoácidos conservada, GXXXXGKT/S, conhecida como motivo Walker A, que forma o *P-loop* (SCHULZ, 1992). Esta região liga-se ao  $\beta$ -fosfato presente no ADP e está localizada entre a  $\beta 1$  e  $\alpha 1$ . As NMP quinases sofrem grande mudança conformacional durante a sua atividade catalítica. As regiões responsáveis por este movimento estão localizadas no sítio de ligação do NMP e no *Lid domain* (MÜLLER et al., 1996; GERSTEIN; SCHULZ; CHOTHIA, 1993).

Hartmann et al. (2006) baseado na análise de movimentos globais acompanhados no processo de ligação de substratos no sítio ativo da chiquimato quinase, reclassificou os

domínios da chiquimato quinase. Segundo sua classificação, temos agora quatro domínios, que são: o ESB (*extended SB*) que inclui o subdomínio SB; o NB (*nucleotide-binding*) que inclui o *P-loop* e o *AB-loop* (*adenine binding loop*) e um curto seguimento da hélice  $\alpha 6$ ; o LID; e a parte remanescente que forma o RC (*reduced core*) (Figura 6). O antigo domínio central compreende os presentes domínios RC e NB e um curto seguimento do domínio ESB.

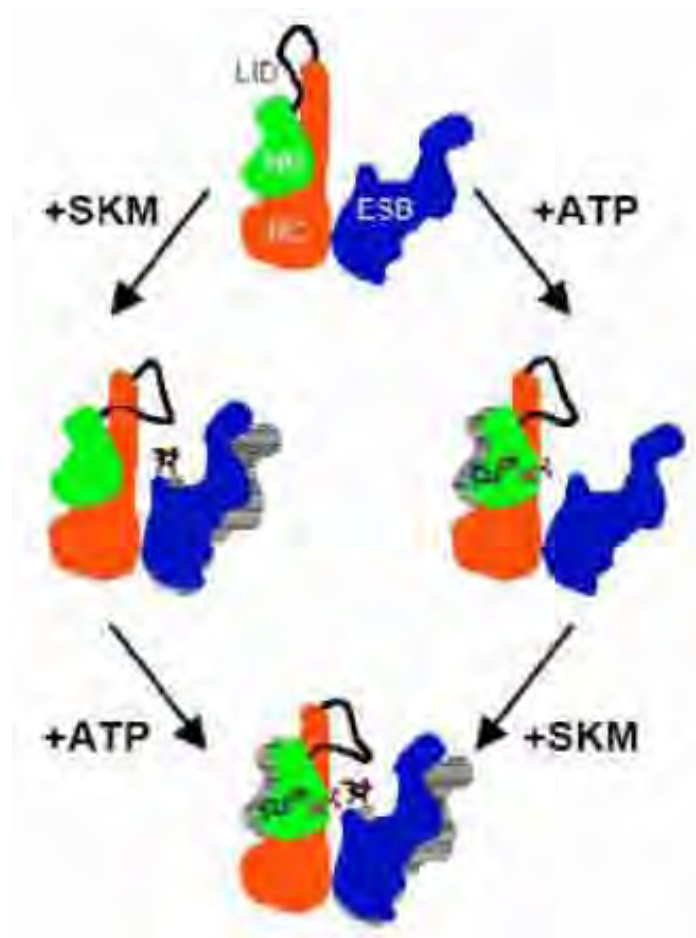


**Figura 6. Domínios estruturais para a Chiquimato quinase proposto por Hartmann et al. (2006). NB: nucleotide-binding, RC: reduced core, ESB: extended Shikimate binding, LID: Lid domain**

Apesar dos estudos cristalográficos feitos com as chiquimato quinases de *M. tuberculosis* (GU et al., 2002), *E. coli* (ROMANOWSKI; BURLEY, 2002) e *Erwinia chrysanthemi* (KRELL; COGGINS; LAPTHORN, 1998) a posição precisa do ácido chiquímico só foi determinada em 2004 por Pereira et al. (2004) e Dhaliwal et al. (2004). A determinação da estrutura da chiquimato quinase em complexo com ácido chiquímico

revelou os resíduos do sítio ativo que participam das interações entre este substrato e a enzima. Pela análise da estrutura deste complexo puderam-se observar também as mudanças conformacionais ocasionadas pela presença do substrato em seu sítio ativo. As principais alterações conformacionais ocorrem no domínio ESB e no *Lid domain*, na qual ambos movem-se em direção ao centro da enzima causando um fechamento parcial do sítio ativo da enzima (PEREIRA et al., 2004; DHALIWAL et al., 2004).

Entretanto, muito recentemente Gan et al. (2006) e Hartmann et al. (2006) mostraram através de dados estruturais que a molécula de nucleotídeo é o principal responsável pelo fechamento do *Lid domain* sobre o sítio ativo da *MtCQ*. Hartmann et al. (2006) mostrou que ocorre um sinergismo de ligação dos substratos no sítio ativo desta enzima. Segundo estes autores, ocorrem um efeito sinérgico aleatório de ligação tanto para o ácido chiquímico como para o nucleotídeo. Esse efeito sinérgico pode ser consistente com o possível aumento da afinidade para ambos os substratos, independente da ordem de ligação. A Figura 7 mostra o efeito sinérgico de ambos os substratos e suas influências na estrutura da chiquimato quinase de *Mycobacterium tuberculosis* proposto por Hartmann et al. (2006).



**Figura 7. Alterações conformacionais observadas na Chiquimato quinase de *M. tuberculosis* na presença de seus ligantes e o possível efeito sinérgico desses ligantes segundo Hartmann et al. (2006). SKM: ácido chiquímico; ATP: adenosina trifosfato**

Até muito recentemente, não se conhecia a conformação da estrutura da chiquimato quinase na ausência de qualquer ligante. Entretanto Gan et al. (2006) e Hartmann et al. (2006) determinaram a estrutura desta enzima no seu estado nativo que possibilita intuir os movimentos causados pela molécula de ADP. Esses autores também determinaram a estrutura da chiquimato quinase em complexo com análogo de ATP (GAN et al., 2006) e ATP (HARTMANN et al., 2006) e, desta forma, pode-se acrescentar informações sobre o mecanismo de transferência do fosfato da molécula de ATP para a molécula de ácido chiquímico.

Com relação ao íon cloreto que interage com a chiquimato quinase, foi observada a influência deste sobre a proteína apenas por técnicas de baixa resolução (CERASOLI et al., 2003). Assim, por meio de dicróismo circular e de cinética enzimática foi observado que a presença do íon cloreto enfraquece a interação entre o ácido chiquímico e a chiquimato quinase de *Erwinia Chrysanthemi*, mas, por outro lado, fortalece a afinidade da enzima por ADP ou ATP (CERASOLI et al., 2003). Entretanto, apesar deste estudo, pouco se sabe sobre o papel deste íon e do íon magnésio sobre a estrutura tridimensional desta proteína.

#### 1.4 A Corismato Sintase

Outra enzima da via do ácido chiquímico, que a partir de 2004 se tornou alvo de muitas pesquisas na área de biologia estrutural, é a corismato sintase (AHN et al., 2004, VIOLA; SARIDAKIS; CHRISTENDAT et al., 2004, MACLEAN; ALI et al., 2003, QUEVILLON-CHERUEL et al., 2004, DIAS et al., 2006). Esta proteína é a última enzima da via do ácido chiquímico e realiza a conversão de 5-enoilpiruvil-3-chiquimato fosfato (EPSP) em corismato, através de uma reação de eliminação de um grupo fosfato da molécula de EPSP (Figura 8) (KITZING; MACHEROUX; AMRHEIN, 2001). A eliminação desse grupo fosfato é responsável pela adição da segunda dupla ligação na molécula de corismato (PITTARD, 1987). A corismato sintase requer para sua reação a presença de flavina mononucleotídeo (FMN), embora a reação não envolva um estado de mudança redox (HERRMANN, 1995). A eliminação do grupo fosfato é realizada por um mecanismo não usual, através de uma 1,4-anti-eliminação. Até agora não foi descrita nenhuma outra enzima que catalisa uma reação deste tipo. Sabe-se também, que na reação está envolvido um intermediário instável de vida curta, formado depois da perda do fosfato da molécula de EPSP. Estudos recentes têm dado evidência de que esse intermediário é um radical, formado pela transferência de um elétron da molécula de FMN para o EPSP (MACLEAN; ALI, 2004; KITZING, 2004).

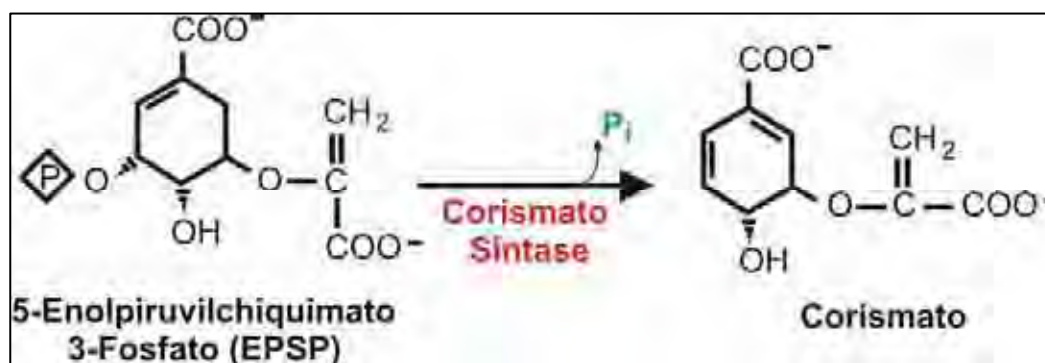


Figura 8. Reação catalisada pela corismato sintase.

A reação catalisada pela corismato sintase ocorre em duas etapas, em que, primeiramente ocorre a clivagem da ligação C3-O3 e o lançamento do fosfato, produzindo o intermediário instável, que colapsa para formar o corismato no segundo passo devido à clivagem da ligação *6proR*-C-H. A orientação da molécula de FMN na estrutura da corismato sintase sugere que ela pode promover a estabilização de qualquer intermediário deficiente em elétron da molécula do EPSP (Figura 9) (MACLEAN; ALI, 2004).

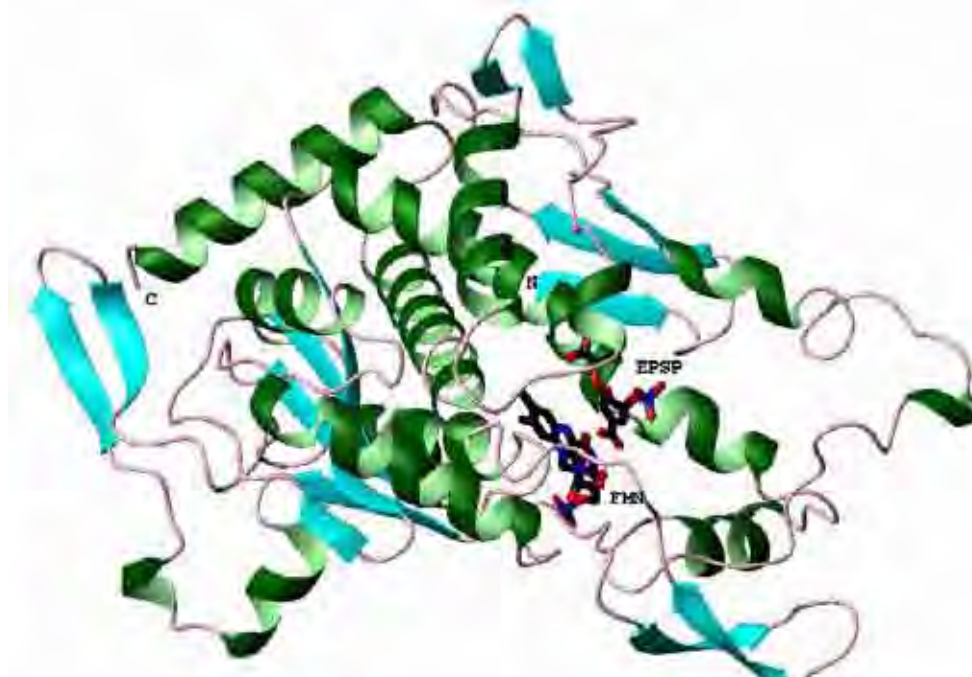


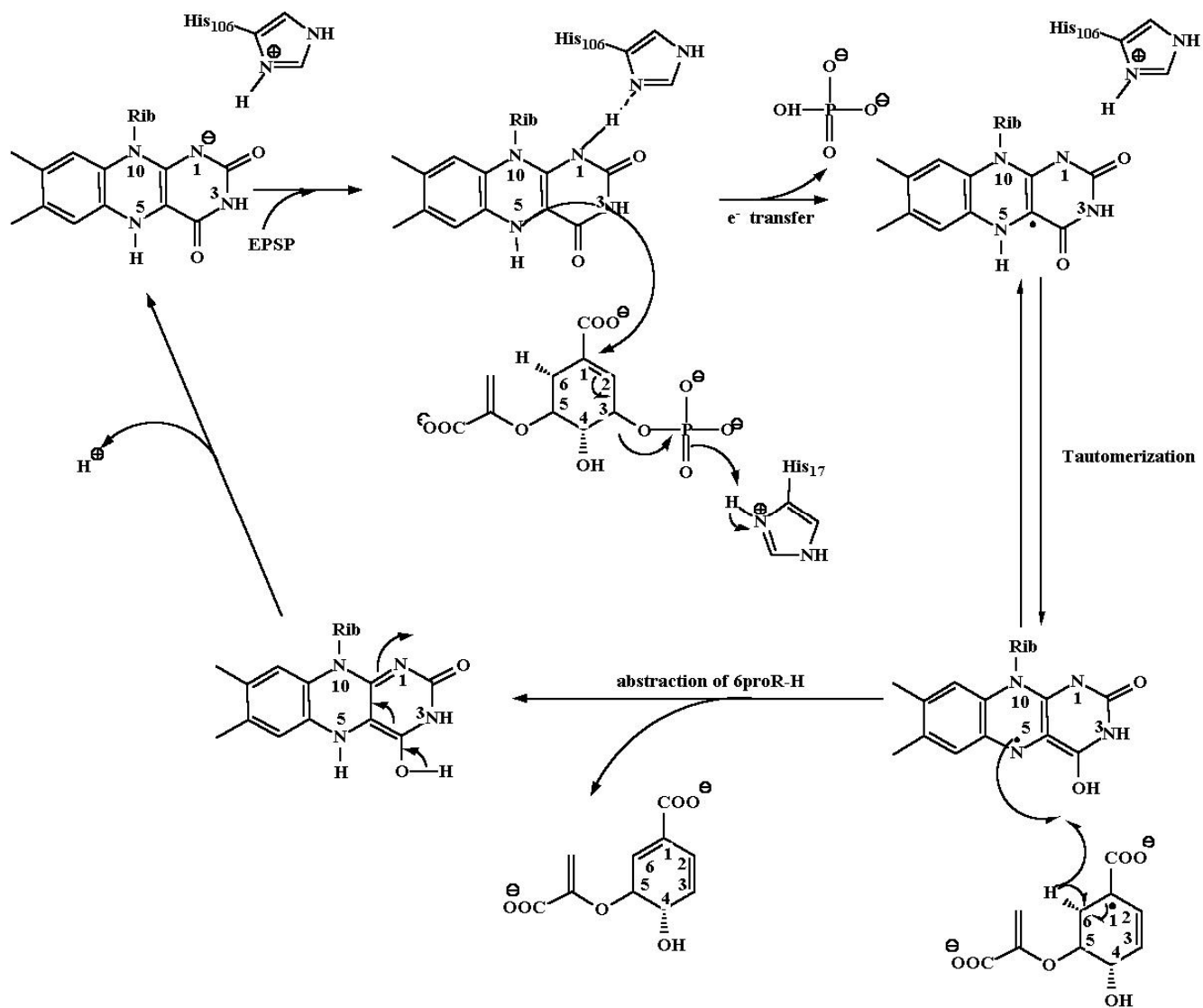
Figura 9. Representação de elementos de estrutura secundária da corismato sintase de *Streptococcus pneumoniae* complexada com 5-enoilpiruvil-chiquimato 3-fosfato (EPSP) e flavina mononucleotídeo (FMN) (MACLEAN; ALI 2004), gerada pelo programa MolMol (KURADI; BILLETER; WÜTHRICH, 1996).



Corismato sintases de diferentes organismos podem ser classificadas de acordo com sua capacidade em reduzir a molécula de flavina mononucleotídeo. Desta forma, temos as corismato sintases monofuncionais e as bifuncionais. Corismato sintases de fungos apresentam uma segunda função, além da conversão de EPSP em corismato. Elas são também capazes de realizar a redução da molécula de FMN utilizando NADH. Assim, estas corismato sintases são classificadas como bifuncionais, por apresentarem também uma função intrínseca de flavina redutase (MACHEROUX et al., 1999). Por outro lado, as corismato sintases de plantas, bactérias e parasitas do filo apicomplexa não possuem essa capacidade de redução de flavina mononucleotídeo, tendo que obter esta molécula na sua forma reduzida do meio e, portanto, são classificadas como monofuncionais (MACHEROUX et al., 1999).

Apesar da estrutura cristalográfica da corismato sintase de vários organismos terem sido resolvidas (AHN et al., 2004, VIOLA; SARIDAKIS; CHRISTENDAT et al., 2004, MACLEAN; ALI et al., 2003, QUEVILLON-CHERUEL et al., 2004), somente a estrutura de *Streptococcus pneumoniae* (SpCS) apresenta ambas moléculas de flavina mononucleotídeo (FMN) e o substrato 5-enoilpiruvil-3-chiquimato fosfato (EPSP) ligadas no seu sítio ativo (MACLEAN, ALI, 2003). Desta forma, esta estrutura revelou muitas informações sobre o complexo mecanismo catalítico da corismato sintase. Baseado nesta estrutura e na análise de proteínas mutantes de *Neurospora Crasa*, nas quais dois importantes resíduos de histidina do sítio ativo foram mutados para alanina, pôde-se propor um mecanismo mais completo para a reação catalisada por esta enzima. Segundo este mecanismo, após a ligação da molécula de EPSP no sítio ativo da enzima com a molécula de FMN reduzida já ligada, um elétron é transferido para a dupla ligação do substrato, iniciando desta maneira, a clivagem e o lançamento do fosfato, com um resíduo de histidina atuando como um neutralizador da carga incipiente do átomo de oxigênio (Figura 10). O resultante carbono neutro C(4a) da flavina semiquinona tautomeriza-se para uma

espécie radical na qual o elétron não pareado reside sobre o átomo N(5), que concomitantemente causa a abstração do hidrogênio ligado ao C(6). E finalmente, ocorre a deprotonação da flavina reduzida restaurando o seu estado inicial (Figura 10) (KITZING et al., 2004).



**Figura 10. Mecanismo catalítico completo para a Corismato Sintase proposto por Kitzing et al. (2004). Rib: Ribose**

Apesar das estruturas cristalográficas da corismato sintase de *Aquiflex aeolicus* (VIOLA; SARIDAKIS; CHRISTENDAT et al., 2004) e *S. cerevisiae* (QUEVILLON-CHERUEL et al., 2004) estarem na forma nativa, grande parte de suas estruturas não foi resolvida, devido à ausência de densidade eletrônica, principalmente alguns importantes resíduos do sítio ativo. Assim, a conformação desses no sítio ativo da corismato sintase na ausência de qualquer ligante é desconhecida.

A estrutura da corismato sintase apresenta-se como um tetrâmero, onde cada monômero possui um único domínio central, que é rodeado por *loops* e trechos discretos de hélices  $\alpha$  e folhas  $\beta$ . O centro do monômero consiste de uma camada de quatro longas hélices  $\alpha$ , sanduíche entre um par de 4 fitas de folhas  $\beta$  antiparalelas (MACLEAN; ALI, 2003) (Figura 9). O bolsão de ligação apresenta uma parte hidrofóbica, na qual está localizado o sistema de anéis isoaloxazino da molécula de FMN e uma outra parte hidrofílica carregada positivamente devido a grande quantidade de argininas e histidinas. Nesta segunda parte do sítio, está localizado a porção remanescente da molécula de FMN e o EPSP, na qual estas duas moléculas apresentam-se justapostas a uma distância aproximada de 3,0 Å, favorecendo, assim, a transferência do elétron entre a molécula de EPSP e FMN (MACLEAN; ALI, 2004).

### 1.5 A Triptofano Sintase

As ramificações da via do ácido chiquímico também representam alvos interessantes para o desenvolvimento de drogas de amplo espectro e vacinas para doenças causadas por microrganismos, pelo fato de também estarem ausentes em mamíferos. Exemplos clássicos, é a inibição da dihidrofolato redutase (DHFR), terceira enzima da via de síntese do ácido *p*-aminobenzóico, por trimetoprim (HAWSER; LOCIURO; ISLAM, 2006) e a inibição da dihidropteroato sintase, quarta enzima da mesma via, pelos antibióticos da série das sulfonamidas (SKÖLD, 2001). Outros exemplos de alvo são as

enzimas da via de síntese de triptofano, como demonstrado por Smith (2001), em que a inativação do gene *trpD*, responsável pela síntese da proteína antranilato fosforibosiltransferase, produz cepas atenuadas de *M. tuberculosis*. A inativação do gene *trpA* (corresponde a subunidade I da antranilato sintase) também foi demonstrado gerar cepas atenuadas de *Bordetella bronchiseptica* e *Leptospira meyeri* (McARTHUR et al., 2003; BAUBY; GIRON; PICARDEAU, 2003). A enzima triptofano sintase demonstrou ser alvo para o desenvolvimento de herbicidas, pois testes *in vitro* com inibidores análogos de substrato demonstraram ter atividade inibitória para esta enzima (FINN et al., 1999).

Desta forma a via de síntese de triptofano é bastante estudada no desenvolvimento de terapias contra doenças infecciosas. Esta via é composta por cinco enzimas, sendo elas, a antranilato sintase, antranilato fosforibosiltransferase, fosforibosil-antranilato isomerase, indol-glicerol-3-fosfato sintase e a triptofano sintase (Figura 11).

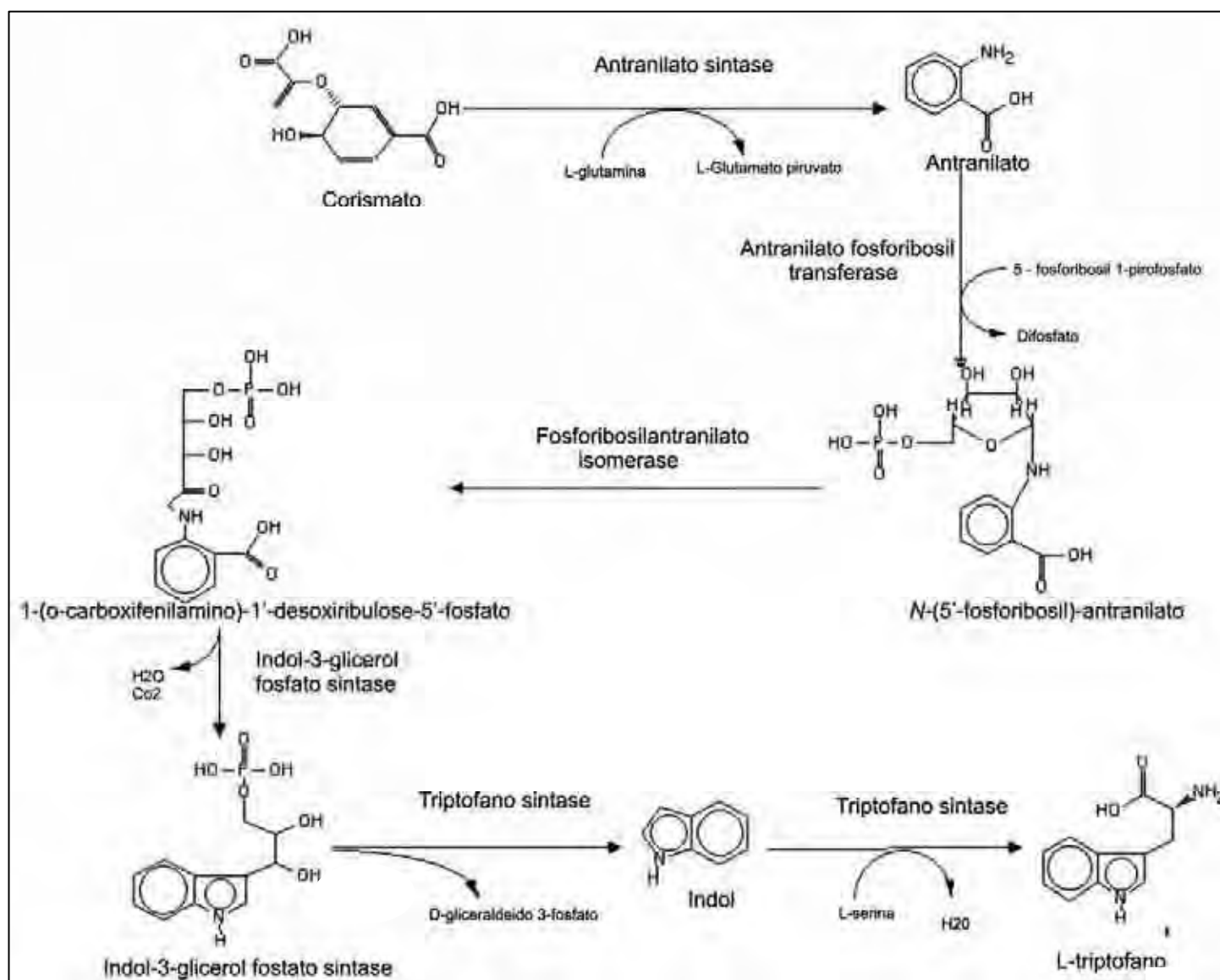
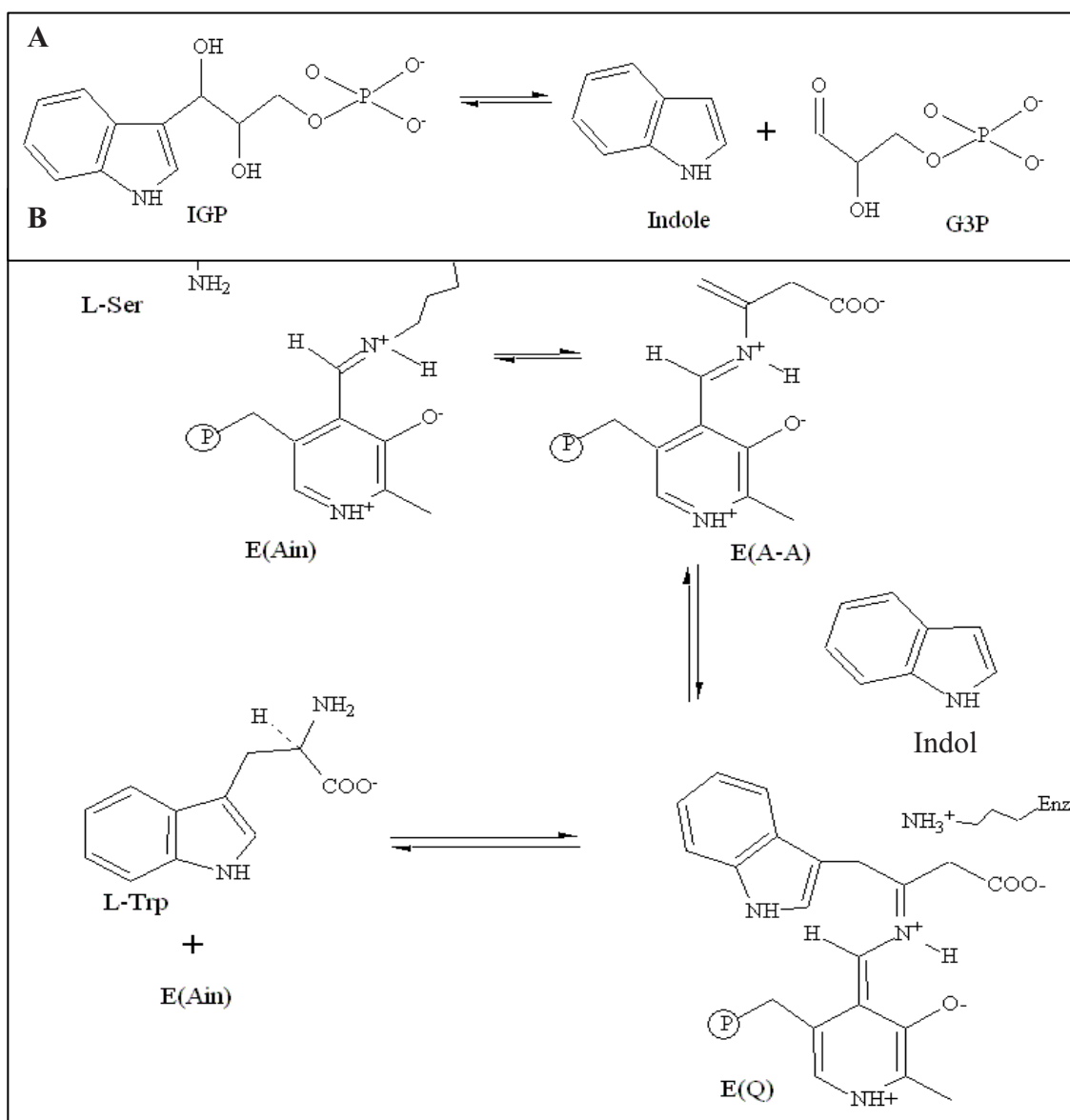


Figura 11. Via de síntese do aminoácido triptofano

Dentre estas enzimas, uma das mais bem estudadas é a triptofano sintase, sendo muito bem conhecido os detalhes de sua purificação, bioquímica, inibição e modificações estruturais de sua atividade (PAN; WOEHL; DUNN, 1997).

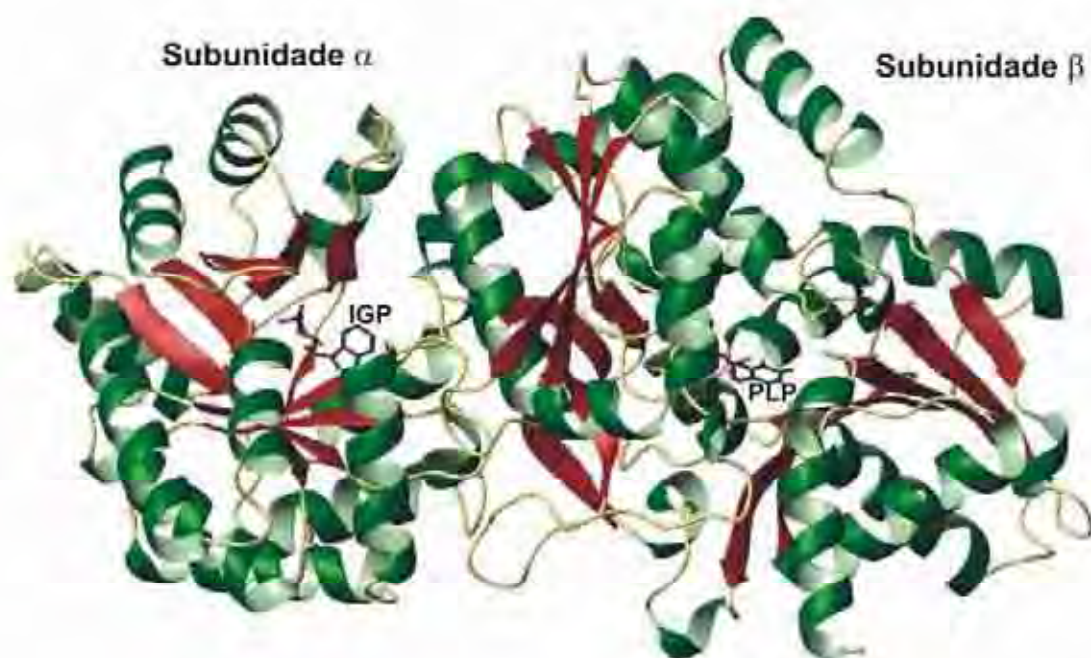
A triptofano sintase é responsável pela catálise das duas últimas reações da biossíntese de L-triptofano. Esta proteína é uma enzima bifuncional com duas subunidades, denominadas  $\alpha$  e  $\beta$ . A subunidade  $\alpha$  catalisa uma reação retroaldol em que o indol-3-

glicerol fosfato (IGP) é clivado em indol e D-gliceraldeído-3-fosfato (G3P). O indol produzido pela subunidade  $\alpha$  reage com L-serina ativada por uma reação dependente de piridoxal fosfato (PLP) no sítio ativo da subunidade  $\beta$  para formar L-triptofano e água (Figura 12A, B).



**Figura 12.** Reação catalisada pela triptofano sintase. A) Reação catalisada pela subunidade  $\alpha$ . B) Reação catalisada pela subunidade  $\beta$ . IGP: Indol glicerol 3-fosfato; G3P: Glicerol fosfato; L-Ser: L-Serina; Enz: Enzima; E(Ain): Almidina interna; E(A-A): Aminoacrilato; EQ: Espécie quinóide; L-trp: L-triptofano.

A subunidade  $\alpha$  da triptofano sintase apresenta um enovelamento similar a da proteína triose fosfato isomerase (TIM) (BANNER et al., 1975), apresentando uma estrutura  $\alpha/\beta$  barril (MILES, 1995). Entretanto, a subunidade  $\beta$  apresenta dois domínios de aproximadamente mesmo tamanho, sendo eles denominados N-domínio e C-domínio (HYDE et al, 1988). O sítio ativo da subunidade  $\alpha$  é localizado próximo à interface com a subunidade  $\beta$  em uma depressão que permite interações de hidrogênio com o grupo fosfato do substrato IGP e um ambiente hidrofóbico para a sua porção indol (WEYAND, M.; SCHILICHTING, 1999). O sítio ativo da subunidade  $\beta$  apresenta-se enterrado no centro desta subunidade entre os N- e C- domínios. A molécula de PLP ligada ao sítio ativo desta subunidade forma uma ligação covalente tipo C=N (base de Schiff) com um grupo amino de um resíduo de lisina (Figura 12) (HYDE et al, 1988).

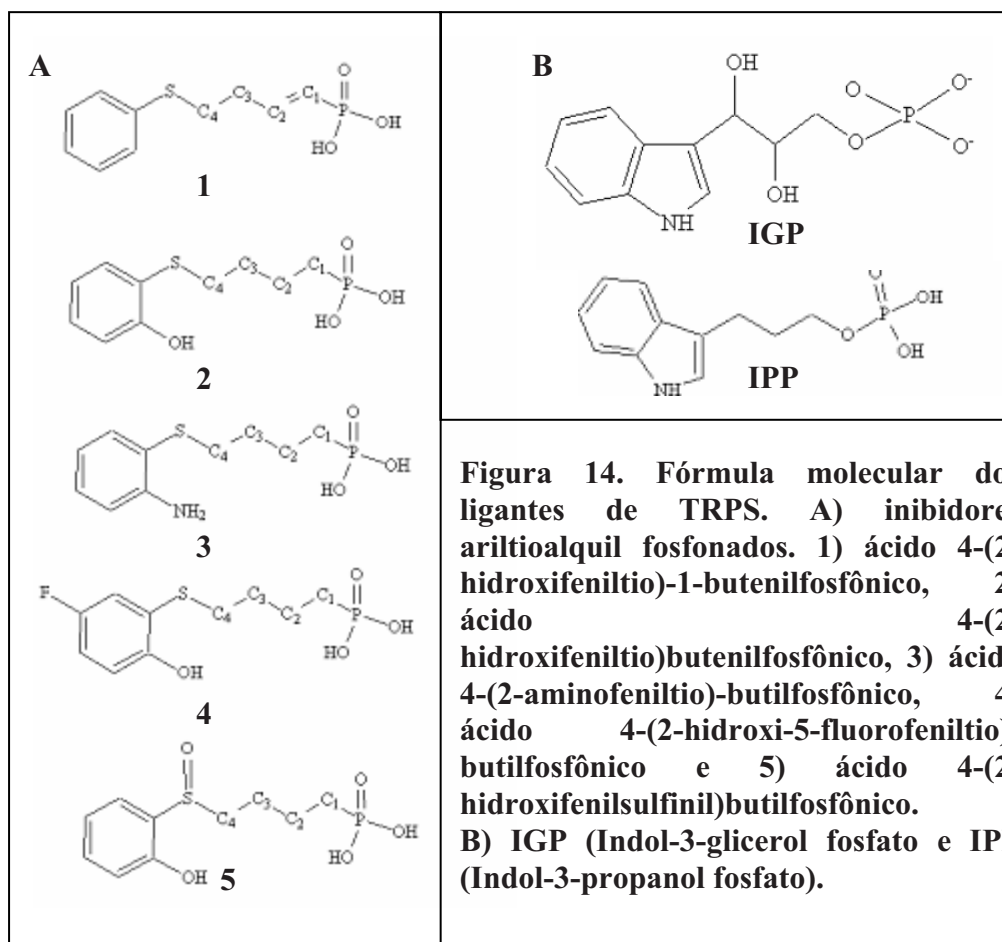


**Figura 13. Estrutura tridimensional da Triptofano sintase de *Salmonella typhimurium* em complexo com Indol glicerol 3-fosfato (IGP) na subunidade  $\alpha$  e piridoxal fosfato (PLP) na subunidade  $\beta$  (HYDE et. al, 1988).**

A estrutura tridimensional da triptofano sintase de *Salmonella typhimurium* foi a primeira a ser resolvida e revelou a presença de um longo canal hidrofóbico que conecta os sítios ativos da subunidade  $\alpha$  e da subunidade  $\beta$ . Este canal teria o papel de impedir a difusão do substrato e do produto da reação catalisada pela proteína, e também de estar envolvido na regulação alostérica que sincroniza as reações nas duas subunidades (SACHPATZIDIS et al., 1999). Após a determinação da estrutura nativa e de outras estruturas com diferentes ligantes para *S. typhimurium*, pode-se observar que ligantes induzem mudanças conformacionais que são importantes para o posicionamento do indol dentro do canal e também para a comunicação alostérica entre as duas subunidades. Drásticas mudanças conformacionais ocorrem com a adição de ligantes, tanto na subunidade  $\alpha$  como na subunidade  $\beta$ . Estas mudanças incluem ordenação de *loops* próximos ao sítio ativo da subunidade  $\alpha$  e movimentos de subdomínios da subunidade  $\beta$  (MILES, RHEE, DAVIES, 1999).

Entre os ligantes estudados que se ligam na triptofano sintase, está o indol-3-propanol fosfato (IPP) (WEYAND; SCHLICHTING, 1999) e cinco análogos do estado de transição ariltioalquil fosfonados, que se ligam e inibem a subunidade  $\alpha$  da triptofano sintase (Figura 14) (SACHPATZIDIS et al., 1999). O IPP é um análogo de IGP que inibe a subunidade  $\alpha$  com um  $K_i$  de 15  $\mu\text{M}$  e foi o ponto inicial para a estratégia de desenho dos inibidores ariltioalquil fosfonados (FINN et al., 1999). Estes inibidores foram desenhados de forma a mimetizar o estado de transição formado durante a reação  $\alpha$  da enzima e possuem uma afinidade maior do que o substrato natural IGP e o análogo IPP. Estes inibidores são derivados ariltioalquilfosfonados orto-substituídos que apresenta um átomo de enxofre substituindo C2 (FINN et al., 1999; SACHPATZIDIS et al., 1999). Entretanto, somente foram realizados estudos sobre o efeito inibitório dos inibidores ariltioalquil fosfonados sobre a triptofano sintase de *Arabidopsis thaliana* e *S. typhimurium*.

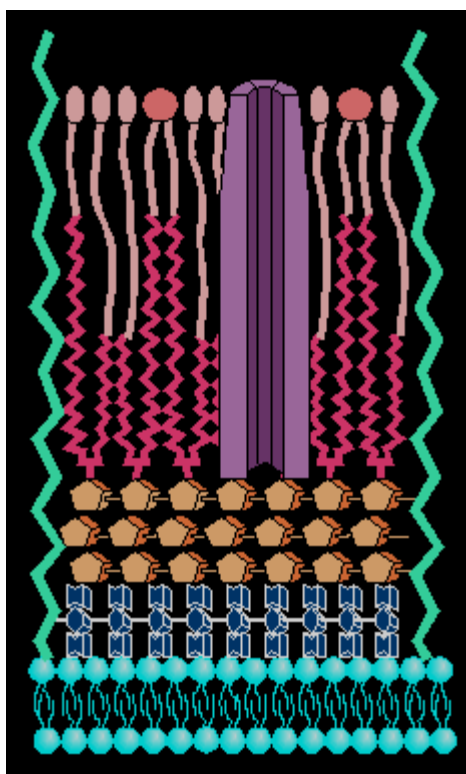











### 1.6 InhA e o mecanismo de resistência ao antibiótico isoniazida

Apesar das enzimas comentadas anteriormente serem alvos em potencial para o desenvolvimento de novos compostos que tenham atividade inibitória e, portanto, possivelmente medicamentosas, nenhum composto foi desenvolvido especificadamente para estas enzimas de *Mycobacterium tuberculosis*. Entretanto, uma das enzimas mais estudadas para o desenvolvimento de drogas em *M. tuberculosis*, e que é alvo da isoniazida, é a enzima dependente de NADH, 2-*trans* enoil-ACP (*acyl carrier protein*) redutase (InhA). Esta enzima exibe alta especificidade para ácidos graxos de cadeia longa ( $C_8 > C_{16}$ ) com grupo enoil tioéster, que é consistente com o envolvimento desta enzima na biossíntese de ácidos micólicos (QUÉRMAD et al. 1995). Ácidos micólicos, que por sua

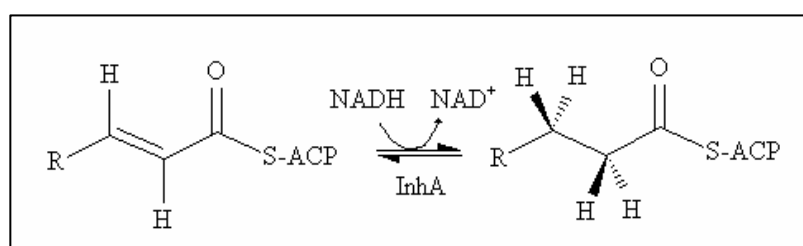
vez, são ácidos graxos  $\alpha$ -alquil- $\beta$ -hidroxil, são os principais componentes da parede celular de micobactérias (BRENNAN; NIKAIDO, 1995). Esta parede celular ou envelope celular é uma característica única das micobactérias e apresenta um alto conteúdo de lipídios constituindo uma barreira composta de ácidos micólicos ancorados a moléculas de arabinogalactano, ligados ao peptidoglicano da membrana celular da bactéria (Figura 15) (MOLLE et al., 2006). Essa parede é um grande obstáculo para a penetração de drogas na célula da micobactéria, representando assim um grande desafio para o desenvolvimento de novas drogas contra tuberculose.



**Figura 15. Representação esquemática do envelope celular de micobactérias.**  bicamada lipídica,  peptidoglicano,  arabinogalactano,  micolato,  acil lipídeos,  lipoarabinomanano (LAM),  porfirina

Dois tipos de sistemas de síntese de ácidos graxos (FAS) são conhecidos. FAS-I, encontrado em vertebrados e está relacionado com a síntese *de novo* de ácidos graxos; e

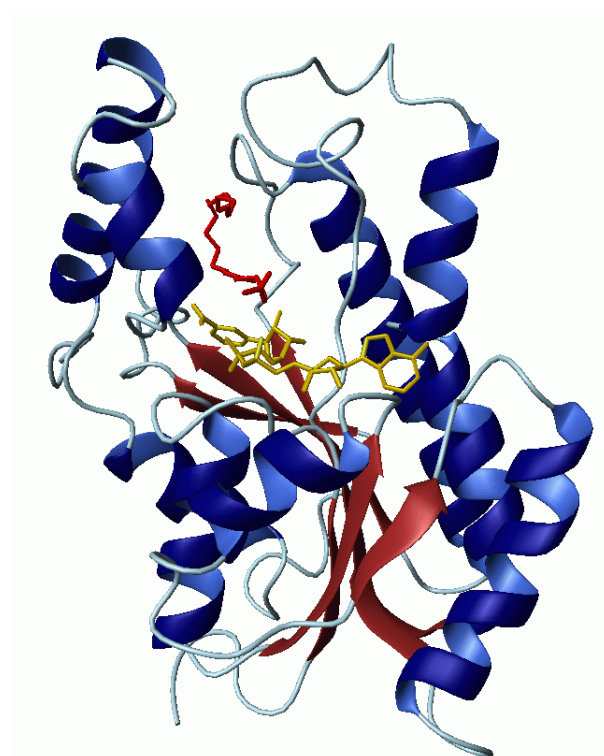
FAS-II, que é usualmente encontrado em bactérias, e está relacionado com a elongação de cadeias de ácidos graxos. Entretanto, *M. tuberculosis* apresenta ambos os sistemas, uma característica que não é muito usual (DOVER et al., 2004). InhA é uma proteína pertencente ao sistema de síntese micobacterial de ácidos graxos tipo II (FAS-II), que realiza a elongação de precursores acil ácidos graxos, provindos do sistema de síntese micobacterial de ácidos graxos tipo I (FAS-I) rendendo longas cadeias de carbono da parte meromicolato dos ácidos micólicos (SCHROEDER et al., 2002). A função da InhA dentro do sistema FAS-II é catalisar a redução dependente de NADH de duplas ligações *trans* entre a posição C2 e C3 de substratos acil ácidos graxos (Figura 16) (QUEMARD et al., 1995).



**Figura 16. Reação catalisada pela enzima InhA**

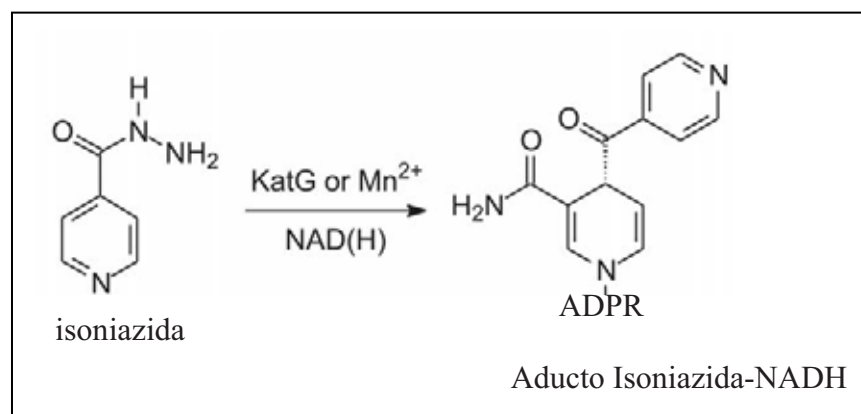
A primeira estrutura da InhA em complexo com NADH foi determinada por Dessen et al. (1995). A estrutura desta proteína apresenta uma estrutura terciária  $\alpha/\beta$  com um domínio apresentando um enovelamento Rossmann, que consiste em fitas  $\beta$  paralelas ligadas por hélices  $\alpha$  em uma ordem topológica  $\beta$ - $\alpha$ - $\beta$ - $\alpha$ - $\beta$ - $\alpha$ . A folha  $\beta$  formada apresenta 6 fitas na ordem 321456 (Figura 17). Este tipo de enovelamento classifica a InhA como pertencente a superfamília das dehidrogenase/reductase de cadeia curta (SDR) (*short dehydrogenase/reductase*) (DESSEN et al., 1995). Posteriormente à determinação desta estrutura, outras estruturas da InhA com diferentes ligantes foram obtidas, entre elas os complexos da InhA:NADH:isoniazida (ROSWARSKI et al., 1998), InhA:NADH:ácido graxo C16 (ROSWARSKI et al., 1999) e o complexo da InhA com diversos ligantes inibidores análogos de substrato ( KUO et al., 2003). Porém a conformação da InhA na

ausência de qualquer ligante ainda não foi elucidada, e assim, não se conhece a conformação desta proteína antes da ligação da molécula de NADH.



**Figura 17. Elementos de estrutura secundária da InhA de *Mycobacterium tuberculosis* em complexo com NADH (em amarelo) e o substrato C16 (em vermelho) (DESSEN et al., 1995)**

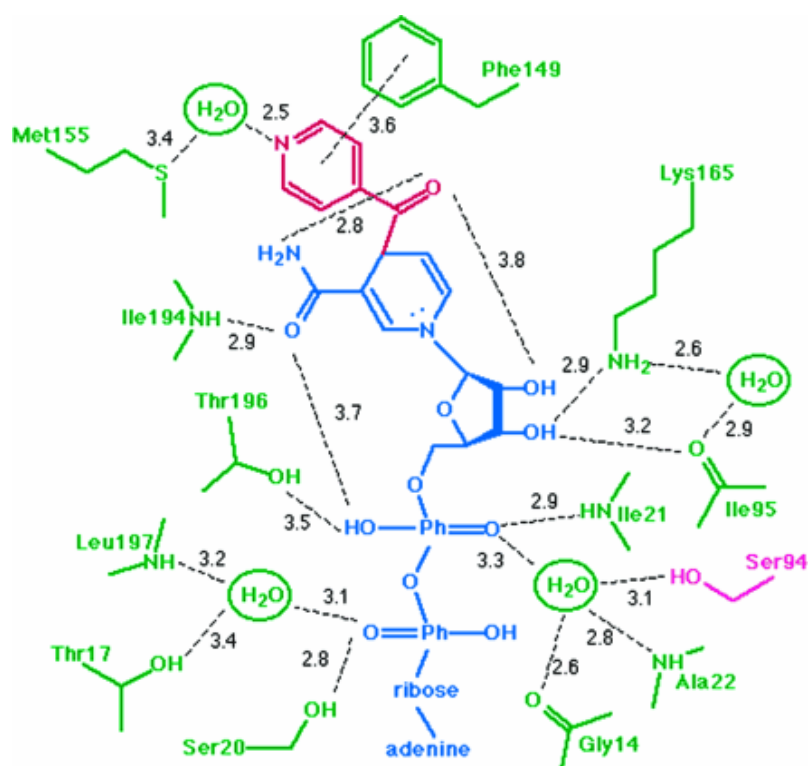
A relação entre a ação da isoniazida e a inibição da InhA é um pouco complexa, pois na verdade a isoniazida é uma pré-droga que deve ser primeiramente convertida para uma forma ativada através da enzima catalase-peroxidase (KatG) para um radical isonicotínico acil (Figura 18) (ROZWARSKI et al., 1999). Este radical torna-se covalentemente ligado ao anel nicotidamida do NADH ancorado no sítio ativo da InhA, criando, desta forma, um NADH aducto que é o potente inibidor da InhA (ROZWARSKI et al., 1998). A posição em que o radical isonicotínico acil liga-se à molécula de NADH é a mesma região responsável pela transferência híbrida do átomo de hidrogênio que ocorre durante a redução do substrato enoil-ACP (ROZWARSKI et al., 1998).



**Figura 18. Mecanismo de ativação da isoniazida pela enzima KatG. ADPR: ADP ribose.**

Porém, tem sido observado mutações na estrutura do gene *inhA* que confere resistência a isoniazida. Cepas apresentando mutações foram obtidas de isolados clínicos de pacientes que vivem nos Estados Unidos (BASSO et al., 1998). Os mutantes isolados apresentam a substituição de um único resíduo de aminoácido da estrutura primária da *InhA*. Entre as mutações obtidas podemos destacar as seguintes: S94A, I21V e I47T. Estas mutações estão bem caracterizadas cineticamente (RAWART; WHITTY; TONGE, 2003, OLIVEIRA et al., 2006). Recentemente as estruturas cristalográficas dos complexos de *InhA*:NADH com essas mutações foram determinadas (OLIVEIRA et al., 2006). Entretanto, a estrutura da *InhA* mutante S94A em complexo com NAD-isoniazida foi determinada por Rozwarski et al. (1998) a 2,8 Å de resolução, mostrando tanto o efeito desta mutação sobre a estrutura como a influência na conformação dos resíduos do sítio ativo (figura 19). Porém, os efeitos de outras mutações como a I21V e I47T sobre a estrutura do complexo da *InhA*:NAD-isoniazida ainda não foram revelados. A resistência à isoniazida por esses mutantes parece estar relacionada a uma baixa afinidade da proteína pela molécula de NADH, que desta maneira, poderia promover a ligação da molécula de acil substratos antes da molécula de NADH, desta forma dificultando a ligação da molécula NAD-isoniazida, devido a um impedimento estérico. Lembrando que análise de dados de cinética enzimática para *InhA* tem demonstrado que a seqüência de ligação do

NADH e ácidos graxos não é estritamente ordenada, mas ocorre uma preferência para o NADH ligar-se primeiro à enzima (QUEMARD et al., 1995). Entretanto quando ocorre a ligação do NAD-isoniazida primeiramente a baixa afinidade pela molécula de NADH poderia permitir a rápida liberação desta molécula, e assim promovendo posterior ligação da molécula de NADH permitindo a normal catálise da enzima (ROZWARSKI et al., 1998, OLIVEIRA et al., 2006).



**Figura 19.** Contatos moleculares entre o complexo isonicotínico acil-NADH e o sítio ativo da InhA. O grupo acil isonicotínico derivado da isoniazida está em vermelho, a porção NADH está em azul, a cadeia lateral da InhA está em verde e a Ser94, o resíduo que causa a resistência a isoniazida quando convertida em Ala está em rosa. Os números representam a distância (em Å) entre os átomos selecionados (ROZWARSKI et al., 1998).

## 2. OBJETIVOS

Os objetivos do presente trabalho foram:

Cristalizar, coletar e processar dados de difração de raios X e resolver as estruturas de duas enzimas da via do ácido chiquímico de *M. tuberculosis*, sendo elas a chiquimato quinase complexada com ADP-ácido chiquímico (*MtCQ*-ADP-ác.Chiquímico), da chiquimato quinase complexada com ADP e da corismato sintase na sua forma nativa (*MtCS*).

Realizar a modelagem molecular da proteína triptofano sintase de *M. tuberculosis* (*MtTRPS*) em complexo com inibidores indol propanol fosfato e uma série de 5 inibidores ariltalquil fosfonados.

Cristalizar, coletar e processar dados de difração de raios X e resolver as estruturas das proteínas InhA selvagem e dos mutantes I21V e S94A em complexo com isoniazida e a proteína mutante S94A na sua forma nativa.

### **3. RESULTADOS**



### 3.1 Crystallization and preliminary X-ray crystallographic analysis of chorismate synthase from *Mycobacterium tuberculosis*.

Marcio V. B. Dias; Fernanda Ely; Fernanda Canduri; José H. Pereira; Jeverson Frazzon; Luiz A. Basso; Mário S. Palma; Walter F. de Azevedo Jr; Diógenes S. Santos. **Acta Crystallographica Section D Biological Crystallography** (ISSN 0907-4449), v. D60, p. 2003-2005, 2004.

Neste trabalho é descrita a cristalização e a análise preliminar de dados de difração de raios X para a corismato sintase de *M. tuberculosis* (peso molecular de 42014 Da).

A proteína corismato sintase foi clonada, expressa e purificada pelo grupo de Pesquisa do Prof. Dr. Diógenes Santiago Santos da PUC – Porto Alegre – RS. A proteína inicialmente precipitada em sulfato de amônio foi dialisada contra tampão Tris HCl, 50 mM, pH 7.8. Para a cristalização foram utilizados os métodos de difusão de vapor, sistema *hanging drop* e matriz esparsa. Após a obtenção dos primeiros cristais, a condição inicial foi otimizada, variando-se a concentração de sal, precipitante e proteína. Foram obtidos cristais para a corismato sintase solubilizada em Tris-HCl, 50 mM, pH 7,8, com dimensões adequadas para a coleta de dados de difração de raios X, em uma condição composta por Hepes-Na, 0,1 M, pH 7,5, Cloreto de Magnésio hexahidratado, 0,6 M e PEG 400, 25%. A proteína estava em uma concentração de 60 mg.ml<sup>-1</sup> e a razão entre a solução de cristalização e a solução de proteína era de 2:1. A coleta de dados de difração de raios X foi realizada no LNLS. Os cristais obtidos são hexagonais, pertencente ao grupo espacial P6<sub>4</sub>22 com dimensões aproximadas de 0,3x0,25x0,25 mm. O conjunto de dados obtido para um cristal foi processado a 2.8 Å de resolução, apresentou uma completude de 97,9% e um R<sub>sym</sub> de 5,6%.

Acta Crystallographica Section D

Biological  
Crystallography

ISSN 0907-4449

**Crystallization and preliminary X-ray  
crystallographic analysis of chorismate synthase  
from *Mycobacterium tuberculosis*****Marcio Vinicius Bertacine Dias,<sup>a</sup>  
Fernanda Ely,<sup>b</sup> Fernanda  
Canduri,<sup>a,b</sup> José Henrique  
Pereira,<sup>a</sup> Jeverson Frazzon,<sup>b</sup>  
Luiz Augusto Basso,<sup>b</sup>  
Mário Sérgio Palma,<sup>c</sup>  
Walter Filgueira de  
Azevedo Jr,<sup>a,b\*</sup> and  
Diógenes Santiago Santos<sup>d\*</sup>**<sup>a</sup>Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, <sup>b</sup>Rede Brasileira de Pesquisas em Tuberculose Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil, <sup>c</sup>Laboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil, and <sup>d</sup>Centro de Pesquisas em Biologia Molecular e Funcional/ Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS 90619-900, BrazilCorrespondence e-mail:  
walterfa@df.ibilce.unesp.br, diogenes@puccrs.br

The enzymes of the shikimate pathway are potential targets for the development of new therapies because they are essential for bacteria but absent from mammals. The last step in this pathway is performed by chorismate synthase (CS), which catalyzes the conversion of 5-enolpyruvylshikimate-3-phosphate to chorismate. Optimization of crystallization trials allowed the crystallization of homogeneous recombinant CS from *Mycobacterium tuberculosis* (*MtCS*). The crystals of *MtCS* belong to space group *P6<sub>4</sub>22* (or *P6<sub>2</sub>22*) and diffract to 2.8 Å resolution, with unit-cell parameters  $a = b = 129.7$ ,  $c = 156.8$  Å. There are two molecules in the asymmetric unit. Molecular-replacement trials were not successful. Heavy-atom derivative screening is in progress.

Received 15 July 2004

Accepted 10 August 2004

**1. Introduction**

Tuberculosis is the second leading cause of deaths worldwide, killing nearly 2 million people each year. Most cases are in underdeveloped countries; over the past decade, tuberculosis incidence has increased in Africa, mainly as a result of the burden of HIV infection, and in the former Soviet Union, owing to socioeconomic change and decline of the health-care system (Frieden *et al.*, 2003). In 1993, the gravity of the situation led the World Health Organization (WHO) to declare tuberculosis a global emergency in an attempt to heighten public and political awareness (Cole *et al.*, 1998). Thus, newer and more efficient anti-tuberculosis drugs are needed. Potential targets for the development of new therapies are the enzymes of the shikimate pathway, because they are essential for bacteria, fungi and apicomplexan parasites, but absent from mammals (Bentley, 1990; Roberts *et al.*, 1998). In microorganisms, the shikimate pathway is used to synthesize the three proteinogenic aromatic amino acids phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp), the folate coenzymes benzoid and naphthoid quinones and a broad range of mostly aromatic secondary metabolites, including siderophores (Dosselaere & Vanderleyden, 2001). This pathway consists of seven enzymes that catalyse the sequential conversion of erythrose-4-phosphate and phosphoenolpyruvate to chorismate, the common precursor of aromatic compounds (Herrmann & Weaver, 1999). The last step in this pathway is performed by chorismate synthase, which catalyzes the conversion of 5-enolpyruvylshikimate 3-phosphate to chorismate via the 1,4-*anti*-elimination of phosphate and a proton, a reaction that is unique in nature. The enzyme

has an absolute requirement for reduced FMN as a cofactor, although the 1,4-*anti*-elimination of phosphate and the C(6-*pro-R*) H atom does not involve a net redox change. The role of the reduced FMN in catalysis has long been elusive. However, recent detailed kinetic and bioorganic approaches have fundamentally advanced our understanding of the mechanism of action of chorismate synthase (Macheroux *et al.*, 1999).

Structural information is necessary for structure-based design in drug discovery, for a better understanding of the catalytic mechanism and also for an understanding of protein–ligand interaction (de Azevedo *et al.*, 1996, 1997, 2002). Although structures of CS from some microorganisms have been determined (Ahn *et al.*, 2004; Maclean & Ali, 2003), the crystal structure of CS from *Mycobacterium tuberculosis* has not been described. The crystal structure of shikimate-pathway enzymes will help in the development of new drugs against tuberculosis and other infectious diseases (de Azevedo *et al.*, 2002; Pereira *et al.*, 2003). In this work, we initiated the structure determination of the *aroF*-encoded *MtCS*, which is composed of 401 residues with a molecular weight of 41 792 Da. We report here the crystallization and the preliminary X-ray crystallographic study of *MtCS*.

**2. Materials and methods****2.1. Cloning, protein expression and purification**

Synthetic oligonucleotide primers were designed based on the *aroF* structural gene sequence of *M. tuberculosis* H37Rv (Cole *et al.*, 1998), containing 5' *Nde*I and 3' *Bam*HI restriction sites. The PCR product was cloned

## crystallization papers

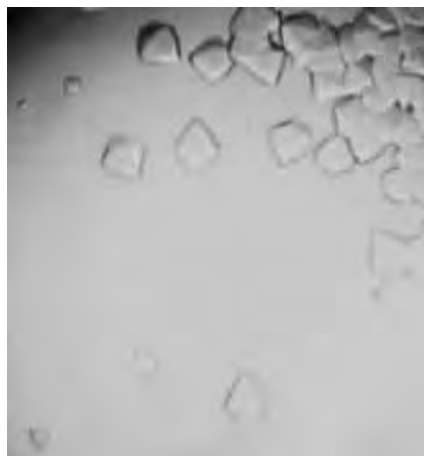
**Table 1**  
Summary of data-collection statistics for *MtCS*.

Values in parentheses are for the highest resolution shell.

X-ray wavelength (Å)	1.427
Unit-cell parameters	
<i>a</i> (Å)	129.74
<i>b</i> (Å)	129.74
<i>c</i> (Å)	156.77
Space group	<i>P</i> 6 <sub>2</sub> 22 or <i>P</i> 6 <sub>2</sub> 22
No. measurements with <i>I</i> > 2σ( <i>I</i> )	92610
No. independent reflections	19341 (2716)
Completeness in the resolution range 55.8–2.8 Å (%)	97.9 (97.9)
<i>R</i> <sub>sym</sub> † (%)	5.6 (16.5)
<i>I</i> /σ( <i>I</i> )	3.8 (4.0)
Highest resolution shell (Å)	2.94–2.80

†  $R_{\text{sym}} = \frac{\sum_h \sum_i |I(h)_i - \langle I(h) \rangle|}{\sum_h \sum_i I(h)_i}$ , where  $I(h)$  is the intensity of reflection  $h$ ,  $\sum_h$  is the sum over all reflections and  $\sum_i$  is the sum over  $i$  measurements of reflection  $h$ .

into pET-23a(+) (Novagen, USA) expression vector and transformed into *Escherichia coli* Rosetta(DE3) host cells. Transformed *E. coli* cells were grown in LB medium at 310 K for 18 h in the absence of IPTG, as used in previously reported expression protocols (Oliveira *et al.*, 2001; Grossman *et al.*, 1998), and harvested by centrifugation. Although it is often argued that the cost of IPTG limits the usefulness of the *lac* promoter to high-added-value products, it has previously been shown that high levels of expression could be obtained with pET vectors as cells entered the stationary phase without the addition of inducer in LB medium. It has been demonstrated that when DE3 hosts are grown to stationary phase in media lacking glucose, cyclic AMP mediated derepression of both the wild type and *lacUV5* promoters occurs (Oliveira *et al.*, 2001). These authors also proposed that cyclic AMP, acetate and low pH are required for high-level expression in the absence of IPTG induction when cells

**Figure 1**  
Hexagonal crystals of *MtCS*. Approximate dimensions are 0.30 × 0.25 × 0.25 mm.

approach the stationary phase in complex media and that derepression of the *lac* operon in the absence of IPTG may be part of a general cellular response to nutrient limitation.

The cell pellet was resuspended in 50 mM Tris–HCl pH 7.8 (Sigma Co., USA), sonicated and centrifuged to remove cell debris. The protein-purification protocol included the following steps: streptomycin sulfate precipitation, ammonium sulfate precipitation, Q Sepharose anion-exchange chromatography, phenyl Sepharose hydrophobic interaction chromatography and Mono Q anion-exchange column (Amersham Pharmacia Biotech, UK). A detailed description of the cloning and expression of recombinant *MtCS*, the protein-purification protocol, N-terminal sequencing, mass spectrometry and determination of the oligomeric state of homogeneous *M. tuberculosis* chorismate synthase will be given elsewhere (manuscript in preparation).

## 2.2. Crystallization

The purified *MtCS* was concentrated and dialyzed against 50 mM Tris–HCl buffer pH 7.8 (Hampton Research, USA). The final protein concentration was about 10 mg ml<sup>-1</sup>. Crystallization was performed by the hanging-drop vapour-diffusion and sparse-matrix methods (Jancarik & Kim, 1991) using tissue-culture multiwell plates with covers (Linbro, ICN Biomedicals, Inc, USA) at a temperature of 293 K. Each hanging drop was prepared by mixing 1 μl each of protein solution and reservoir solution and was placed over 700 μl reservoir solution. Initial conditions were screened using Crystal Screen I and II kits (Hampton Research, USA).

Crystal optimization was carried out by altering the concentration of the salt, precipitant and protein and also the ratio between the protein solution and reservoir solution.

## 2.3. X-ray data collection

A data set was collected at a wavelength of 1.427 Å using a synchrotron-radiation source (Station PCr, LNLS, Campinas-Brazil; Polikarpov *et al.*, 1998). This wavelength was previously set to optimize the overall data-collection statistics. The data set was collected from a single *MtCS* crystal using a MAR CCD image-plate system. The crystal was looped out from the drop and flash-cooled. The PEG 400 present in the crystallization conditions served as a cryoprotectant. X-ray diffraction data were collected at a temperature of 100 K under a

cold nitrogen stream generated and maintained with an Oxford Cryosystem. The crystal was rotated through a total of 160°, with a 1° oscillation range per frame, a crystal-to-detector distance of 130 mm and an exposure time of 60 s. Data were processed on a Silicon Graphics Octane2 computer using the programs *MOSFLM* (Leslie, 1990) and *SCALA* (Collaborative Computational Project, Number 4, 1994).

## 2.4. Molecular-replacement trials

Molecular-replacement trials were carried out with the program *AMoRe* (Navaza, 1994). The crystal structures of *SpCs* (PDB code 1qxo; Maclean & Ali, 2003) and *AaCs* (PDB code 1q1l; Viola *et al.*, 2004) were used as search models.

## 3. Results and discussion

The initial crystals were obtained with reservoir solution comprising 0.2 M magnesium chloride hexahydrate, 0.1 M Na HEPES pH 7.5 and 30% (w/v) PEG 400. The ratio of protein solution to well solution in the drop was 1:1. After optimization of these conditions, better crystals were obtained. These diffracting crystals grew from a reservoir solution containing 0.6 M magnesium chloride hexahydrate, 0.1 M Na HEPES pH 7.5 and 25% (w/v) PEG 400 and the protein solution was concentrated to 60 mg ml<sup>-1</sup> in 50 mM Tris–HCl buffer pH 7.8. The ratio of protein solution to well solution in the drop was 2:1. The crystals grew reproducibly to approximate dimensions of 0.3 × 0.25 × 0.25 mm within 2 d (Fig. 1). The crystal diffracted to 2.8 Å with relatively low mosaicity (0.35°). Fig. 2 shows a typical X-ray diffraction pattern. A total of 92 610 measured reflections were merged

**Figure 2**  
A typical diffraction pattern of the *MtCS* crystal with 1° oscillation range. The crystal diffracts to 2.8 Å resolution.

into 19 341 unique reflections. The overall  $R_{\text{sym}}$  was 5.6% and the completeness was 97.9%. The crystal belongs to the hexagonal space group  $P6_422$  (or  $P6_222$ ), with unit-cell parameters  $a = b = 129.7$ ,  $c = 156.8$  Å. Assuming the asymmetric unit content to be two monomers of molecular weight 41 792 Da, the  $V_M$  value is  $2.28$  Å<sup>3</sup> Da<sup>-1</sup>. Assuming a value of  $0.74$  cm<sup>3</sup> g<sup>-1</sup> for the protein partial specific volume, the calculated solvent content in the crystal is 45.96% (Matthews, 1968). Table 1 summarizes the data-collection statistics.

With the native data of *MtCS* to  $2.8$  Å resolution, molecular replacement was used to attempt to solve the structure using *AMoRe* (Navaza, 1994). Various search models, including complete and modified structures of *SpCS* (PDB code 1qxo) and *AaCS* (1q11), did not yield any meaningful results. This was most likely to be because of conformational differences between the search models and the structure under study. Heavy-atom screening is in progress.

This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0, 02/04383-7, 04/00217-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT) to DSS

and LAB. WFA (CNPq, 300851/98-7), MSP (CNPq, 500079/90-0), DSS (CNPq, 304051/1975-06), LAB (CNPq, 520182/99-5) and JF (CNPq, 301131/2003-01) are researchers for the National Research Council.

## References

- Ahn, H. J., Yoon, H. J., Lee, B. & Suh, S. W. (2004). *J. Mol. Biol.* **27**, 903–915.
- Azevedo, W. F. de Jr, Canduri, F. & da Silveira N. J. F. (2002). *Biochem. Biophys. Res. Commun.* **293**, 566–571.
- Azevedo, W. F. de Jr, de Oliveira, J. S., Basso, L. A., Palma, M. S., Pereira, J. H., Canduri, F. & Santos, D. S. (2002). *Biochem. Biophys. Res. Commun.* **295**, 142–148.
- Azevedo, W. F. de Jr, Leclerc, S., Meijer, L., Havlicek, L., Strnad, M. & Kim, S.-H. (1997). *Eur. J. Biochem.* **243**, 518–526.
- Azevedo, W. F. de Jr, Mueller-Dieckmann, H. J., Schulze-Gahmen, U., Worland, P. J., Sausville, E. & Kim, S.-H. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 2735–2740.
- Bentley, R. (1990). *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384.
- Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Dosselaere, F. & Vanderleyden, J. (2001). *Crit. Rev. Microbiol.* **27**, 75–131.
- Frieden, T. R., Sterling, T. R., Munsiff, S. S., Watt, C. J. & Dye, C. (2003). *Lancet*, **362**, 887–899.
- Grossman, T. H., Kawasaki, E. S., Punreddy, S. R. & Osburne, M. S. (1998). *Gene*, **209**, 95–103.
- Herrmann, K. M. & Weaver, L. M. (1999). *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* **50**, 473–503.
- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Leslie, A. G. W. (1990). *Crystallographic Computing*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 27–38. Oxford University Press.
- Macheroux, P., Schmid, J., Amrhein, N. & Schaller, A. (1999). *Planta*, **207**, 325–334.
- Maclean, J. & Ali, S. (2003). *Structure*, **11**, 1499–1511.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Navaza, J. (1994). *Acta Cryst.* **A50**, 157–163.
- Oliveira, J. S., Pinto, C. A., Basso, L. A. & Santos, D. S. (2001). *Protein Expr. Purif.* **22**, 430–435.
- Pereira, J. H., Canduri, F., de Oliveira, J. S., da Silveira, N. J., Basso, L. A., Palma, M. S., de Azevedo, W. F. Jr & Santos, D. S. (2003). *Biochem. Biophys. Res. Commun.* **312**, 608–614.
- Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. C. & Craievich, A. (1998). *J. Synchrotron Rad.* **5**, 72–76.
- Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. E., Krell, T., Coggins, J. R., Coombs, G. H., Milhous, W. K., Tzipori, S., Ferguson, D. J., Chakrabarti, D. & McLeod, R. (1998). *Nature (London)*, **395**, 801–805.
- Viola, C. M., Saridakis, V. & Christendat, D. (2004). *Proteins*, **54**, 166–169.

### 3.2 Structure of chorismate synthase from *Mycobacterium tuberculosis*.

Marcio V.B. Dias; Júlio C. Borges; Fernanda Ely; José H. Pereira; Fernanda Canduri; Carlos H.I. Ramos; Jeverson Frazzon; Mário S. Palma; Luis A. Basso; Diógenes S. Santos; Walter F. de Azevedo Jr. **Journal of Structure Biology** (ISSN 1047-8477), v. 154, p. 130-143, 2006.

Neste trabalho é descrita a estrutura cristalográfica da proteína corismato sintase de *M. tuberculosis* (peso molecular de 42014 Da).

A estrutura da corismato sintase de *M. tuberculosis* foi resolvida a 2,65 Å de resolução através da técnica de substituição molecular utilizando como modelo de busca a estrutura da corismato sintase de *S. pneumoniae* (MACLEAN; ALI, 2003). Os resíduos não conservados entre as duas estruturas foram trocados para alanina, com exceção de glicinas. A função translação obtida na substituição molecular apresentava um coeficiente de correlação de 37,6% e um R-factor de 53%. Após a confirmação do grupo espacial correto (P6<sub>4</sub>22), realizou-se refinamento cristalográfico. O modelo para a corismato sintase de *M. tuberculosis* foi iterativamente ajustado com base em mapas de densidade eletrônica com coeficientes  $2Fo - 2Fc$  e  $Fo - Fc$ . A estrutura final obtida apresenta um R-factor de 16,2% e um R-free de 22,1% e uma boa qualidade estereoquímica, apresentando 99,3% dos resíduos em regiões favoráveis do gráfico de Ramachandran e apenas 0,3% em regiões não permitidas.

A estrutura da corismato sintase de *M. tuberculosis* é similar às outras estruturas determinadas anteriormente para outros microorganismos. A unidade assimétrica do cristal apresenta um único monômero, que pertence à classe de proteínas  $\alpha/\beta$ . Este monômero apresenta somente um domínio, cuja topologia estrutural é um sanduíche  $\beta-\alpha-\beta$ . Este sanduíche apresenta um centro formado de hélices que se encontra entre duas camadas de

folhas  $\beta$  formadas por quatro fitas. Este domínio é rodeado por *loops* e trechos discretos de hélices  $\alpha$  e fitas  $\beta$ . Entretanto, a estrutura cristalográfica da corismato sintase de *M. tuberculosis* pode ser descrita como um tetrâmero formado por dímeros de dímeros, com simetria 222, sendo que a principal característica da interface de dimerização é uma folha  $\beta$  antiparalela intermonomérica, formada por oito fitas, sendo quatro de um monômero e quatro do outro monômero adjacente. A comparação do sítio ativo da corismato sintase de *M. tuberculosis* com a de outros organismos mostra que muitos resíduos são conservados entre estas estruturas. Entretanto a posição da cadeia lateral de vários resíduos apresenta-se em conformações diferentes, talvez devido à ausência de ligantes em nossa estrutura e a presença de ligantes nas outras estruturas que foram utilizadas para a análise. Portanto, devido à ausência de ligantes em nossa estrutura e a determinação da posição de quase todos os resíduos do sítio ativo, com exceção dos resíduos 49-52, pode-se inferir importantes movimentos ocorridos pela ligação do substrato EPSP e da co-enzima FMN no sítio ativo desta proteína.

Structure of chorismate synthase from *Mycobacterium tuberculosis*

Marcio V.B. Dias<sup>a</sup>, Júlio C. Borges<sup>a</sup>, Fernanda Ely<sup>b</sup>, José H. Pereira<sup>a</sup>,  
Fernanda Canduri<sup>a,c</sup>, Carlos H.I. Ramos<sup>d</sup>, Jeverson Frazzon<sup>b</sup>, Mário S. Palma<sup>e</sup>,  
Luis A. Basso<sup>f</sup>, Diógenes S. Santos<sup>f,\*</sup>, Walter F. de Azevedo Jr.<sup>a,g,\*</sup>

<sup>a</sup> Programa de Pós-Graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil

<sup>b</sup> Rede Brasileira de Pesquisa em Tuberculose, Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil

<sup>c</sup> Departamento de Morfofisiologia, Laboratório de Bioquímica, CCBS, Universidade Federal de Mato Grosso do Sul, Campo Grande, MS, CEP 79070-900, Brazil

<sup>d</sup> Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Síncrotron, Campinas SP, Brazil

<sup>e</sup> Departamento de Biologia, CEIS/IBRC, UNESP, Rio Claro, SP, CEP 13506-900, Brazil

<sup>f</sup> Pontifícia Universidade Católica do Rio Grande do Sul, Centro de Pesquisa em Biologia Molecular e Funcional, Porto Alegre, RS, Brazil

<sup>g</sup> Faculdade de Biociências—Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681. Porto Alegre-RS CEP 90619-900, Brazil

Received 17 July 2005; received in revised form 8 November 2005; accepted 9 December 2005

Available online 17 January 2006

## Abstract

In bacteria, fungi, plants, and apicomplexan parasites, the aromatic compounds, such as aromatic amino acids, are synthesized through seven enzymes from the shikimate pathway, which are absent in mammals. The absence of this pathway in mammals make them potential targets for development of new therapy against infectious diseases, such as tuberculosis, which is the world's second commonest cause of death from infectious disease. The last enzyme of shikimate pathway is the chorismate synthase (CS), which is responsible for conversion of the 5-enolpyruvylshikimate-3-phosphate to chorismate. Here, we report the crystallographic structure of CS from *Mycobacterium tuberculosis* (*MtCS*) at 2.65 Å resolution. The *MtCS* structure is similar to other CS structures, presenting  $\beta$ - $\alpha$ - $\beta$  sandwich structural topology, in which each monomer of *MtCS* consists of a central helical core. The *MtCS* can be described as a tetramer formed by a dimer of dimers. However, analytical ultracentrifugation studies suggest the *MtCS* is a dimer with a more asymmetric shape than observed on the crystallographic dimer and the existence of a low equilibrium between dimer and tetramer. Our results suggest that the *MtCS* oligomerization is concentration dependent and some conformational changes must be involved on that event.

© 2005 Elsevier Inc. All rights reserved.

**Keywords:** Chorismate synthase; Crystallography; Analytical ultracentrifugation; *Mycobacterium tuberculosis*; Shikimate pathway

## 1. Introduction

In mammals, diet has to provide the essential amino acids, such as, phenylalanine, tryptophan, and tyrosine. In bacteria, fungi, plants, and apicomplexan parasites, these aromatic amino acids are synthesized through the

complex shikimate pathway (Bentley, 1990; Macheroux et al., 1999). Furthermore, the shikimate pathway provides these organisms the basic building blocks for the synthesis of the other aromatic compounds required for different functions as UV protection, electron transport, signaling, iron uptake, etc. The shikimate pathway is responsible for conversion of D-erythrose-4-phosphate and phosphoenolpyruvate to shikimate and subsequently to the dihydroaromatic compound chorismate (Bentley, 1990; Dosselaere and Vanderleyden, 2001). The absence of shikimate pathway in mammals has rendered its enzymes as

\* Corresponding authors. Fax: +55 17 51 3220 3629.

E-mail addresses: diogenes@puccs.br (D.S. Santos), walter.junior@puccs.br (W.F. de Azevedo Jr.).

potentially targets for the development of new therapy against infectious disease, such as tuberculosis (TB).<sup>1</sup>

TB is considered one of the most serious global public health challenges of the 21st century. It is the world's second commonest cause of death from infectious disease, after acquired immune deficiency syndrome. According to a recent report compiled by the World Health Organization, the total number of new cases of TB worldwide in 2002 has increased to around 9 million (Duncan, 2004; WHO, 2004), from which 2–3 million people died, despite the chemotherapy available (WHO, 2004). These deaths are mostly of young adults but also include about 100 000 children under the age of 5 years (Gandy and Zumla, 2002). The key driver of the increase of TB is its synergy with the human immunodeficiency virus epidemic, which has a devastating impact in some parts of the world, such as, the African region (Duncan, 2004). Thus, there is an urgent need for new anti-mycobacterial inhibitors.

The enzymes of shikimate pathway are good candidates for development of new therapies against TB. Enzymes from this metabolic pathway have been submitted to intensive structural studies (Azevedo et al., 2002; Basso et al., 2005; Pereira et al., 2003, 2004; Silveira et al., 2005). The last enzyme from this pathway is the chorismate synthase (CS), which catalyzes the conversion of the 5-enolpyruvylshikimate-3-phosphate (EPSP) to chorismate. The CS reaction comprises an *anti*-1,4-elimination of the 3-phosphate group and the C(6*proR*)-hydrogen (Bornemann et al., 2003; Hill and Newkome, 1969; Macheroux et al., 1998). It is the only enzymatic reaction known of such transformation in biological systems, making the CS a unique enzyme in nature. The CS requires reduced flavin mononucleotide (FMN), an essential cofactor typically found in many biological redox reactions. Surprisingly, the reaction catalyzed by CS does not involve an overall change in redox state (Bornemann et al., 1996; Kitzing et al., 2004; Macheroux et al., 1996; Macheroux et al., 1999). According to Bornemann et al. (1995), the reduced FMN donates an electron to EPSP to facilitate the loss of the phosphate and receive it back after the reaction. So, only flavin in its reduced form is functional and it is not consumed during the reaction (Macheroux et al., 1998; Welch et al., 1974). Furthermore, two classes of CS are distinguished among the microorganisms that possess shikimate pathway. CS from yeasts have the ability to use  $\beta$ -nicotinamide adenine dinucleotide phosphate (NAD(H)P) for the reduction of oxidized FMN, having therefore, an additional catalytic activity and wherefore are called bifunctionals, while all the other CS lacking it are called monofunctional (Macheroux et al., 1999).

There are crystal structures available of CS from three bacteria: CS from *Streptococcus pneumoniae* complexed with FMN and EPSP (Maclean and Ali, 2003); CS from *Helicobacter pylori* complexed with FMN (Ahn et al., 2004); and native CS from *Aquiflex aelicus* (Viola et al., 2004) (PDB access codes: 1QXO, 1UMO, and 1Q1L, respectively) and one CS from yeast (native CS of *Saccharomyces cerevisiae*) (PDB access code: 1R53, Quevillon-Cheruel et al., 2004). In all these structures, the CS is presented as tetramer composed of two dimers, like a dimer of dimers. The monomer comprises a single large core domain, which is surrounded by loops and discrete stretches of  $\alpha$ -helix and  $\beta$ -sheet. Here, we present structural data of the CS from *M. tuberculosis* (*MtCS*) in its native form. The structure of the *MtCS* was determined at 2.65 Å resolution and it presents a  $\beta$ - $\alpha$ - $\beta$  architecture. The analytical ultracentrifugation data suggest that the *MtCS* is predominantly a dimer in solution presenting equilibrium with a tetrameric form. This structure can help us understand the action mechanism of the *MtCS*.

## 2. Experimental procedures

### 2.1. *MtCS* recombinant production, purification, and crystallization

The *MtCS* cloning, expression, purification, and crystallization were reported elsewhere (Dias et al., 2004). The protein concentration was determined spectrophotometrically as described by Edelhoff (1967), using a calculated extinction coefficient for denatured proteins (Gill and von Hippel, 1989).

### 2.2. Circular dichroism spectroscopy

Circular dichroism (CD) measurements were performed using a Jasco J-810 spectropolarimeter with the temperature controlled by a Peltier-type control system PFD 425S. The data were collected at a scanning rate of 100 nm/min with a spectral band width of 1 nm and using a 1 mm path length cell. The thermal-induced unfolding experiments followed by CD were performed at a scan rate of 1 °C/min. The average of three unfolding curves was used to build the *MtCS* thermal-unfold profile and the temperature at the midpoint of the unfolding transition was determined by fitting with Gaussians of the first derivative function.

### 2.3. Structure determination

The data set of *MtCS* was collected at a wavelength of 1.427 Å using the Synchrotron Radiation Source (Station PCr, LNLS, Campinas—Brazil) (Polikarpov et al., 1998) and a CCD detector (MARCCD). The crystal was flash-frozen at 104 K under cold nitrogen stream generated and maintained with an Oxford Cryosystem. The data set was processed up to 2.65 Å resolution using the program

<sup>1</sup> Abbreviations used: TB, tuberculosis; CS, chorismate synthase; *MtCS*, chorismate synthase from *Mycobacterium tuberculosis*; EPSP, 5-enolpyruvylshikimate-3-phosphate; FMN, flavin mononucleotide; CD, circular dichroism; RMSD, root-mean-square deviation; AUC, analytical ultracentrifugation.



MOSFLM (Leslie, 1992) and scaled with SCALA (CCP4, 1994).

The crystal structure of the *MtCS* was determined by standard molecular replacement methods implemented in the program AMoRe (Navaza, 2001), using as search model the monomer structure of *Streptococcus pneumoniae* CS (*SpCS*) (PDB access code: 1QXO) (Maclean and Ali, 2003), in which nonconserved residues were changed to alanines. The rotation search was calculated using data in the range of 12.0–3.0 Å resolution, which gave a solution with a correlation of 7.3%. The translation search yielded a unique solution with a correlation of 37.6% and an *R*-factor of 53.0% was obtained. At this stage, the correct space group was confirmed to be *P*<sub>6</sub><sub>4</sub><sub>2</sub><sub>2</sub> by visualization of crystal packing using the program O (Jones et al., 1991). Further crystallographic refinement was carried out using X-PLOR and simulated annealing (Brunger, 1992). The protein model obtained by molecular replacement was refined resulting *R*-factor and *R*-free of about 33.6 and 43.1%, respectively. Throughout the refinement, the composite electron-density maps with coefficients  $2F_o - F_c$  and  $F_o - F_c$  were calculated with X-PLOR and visualized using XtalView/Xfit (McRee, 1999), and the model was built and iteratively adjusted. After several refinement steps the *R*-factor and the *R*-free decreased to 25.7 and 32.5%, respectively. Further, the group *B* factors were adjusted and it was used the data from 8.0 to 2.65 Å. The model was subjected to further refinement using the maximum-likelihood based program REFMAC5 (Murs̄hudov et al., 1997) using data between 52.93 and 2.65 Å resolution. At this stage, water molecules were added to the model using the program XtalView and the ARP routine of REFMAC5. Refinement converged to *R*-factor and *R*-free values of 16.2 and 22.1%, respectively. The correctness of the stereochemistry of the model was checked using PROCHECK (Laskowski et al., 1994). Root-mean-square deviation (RMSD) differences from ideal geometries for bond lengths, angles, and dihedrals were calculated with X-PLOR (Brünger, 1992). Atomic models were superposed using the program LSQKAB from CCP4 (CCP4, 1994). The molecular surface areas have been calculated using the program AREAIMOL/RESAREA (CCP4, 1994). The PARMODEL (Uchôa, 2004) was used in the analysis of the final model.

#### 2.4. Gel filtration

Gel filtration was performed using a Superdex S-200 column (Amershan Pharmacia Biotech) in FPLC system. The column was equilibrated in 50 mmol/L of Tris–HCl buffer, pH 7.8, containing 200 mmol/L of NaCl. Protein standards used were ribonuclease A (13.7 kDa), chymotrypsinogen (25 kDa), ovalbumin (43 kDa), albumin (67 kDa), aldolase (158 kDa), catalase (232 kDa), ferritin (440 kDa), and thyroglobulin (669 kDa) (Amershan Pharmacia Biotech). Elution time of *MtCS* was recorded and molecular weight was calculated by estimating the elution

volumes of standards of known molecular weight. The *MtCS* were loaded on the gel filtration column at a concentration of 1 mg/mL.

#### 2.5. Analytical ultracentrifugation

Sedimentation velocity and sedimentation equilibrium experiments were performed using a Beckman Optima XL-A analytical ultracentrifuge. The sedimentation velocity experiments were carried out at 20 °C and 25000 rpm (AN-60Ti rotor) with the scan data acquisition at 232 and 236 nm. *MtCS* was tested in concentrations of 300–1000 µg/mL in 20 mmol/L Tris–HCl buffer, pH 8.0, containing 50 mmol/L NaCl and 1 mmol/L β-mercaptoethanol. The analysis involved fitting a model of absorbance versus cell radius data by nonlinear regression. The analysis was performed with the ORIGIN software package (MicroCal Software) supplied with the instrument. The second moment (Goldberg, 1953) and the sedimentation time derivative ( $g(s^*)$  integral distribution) (Stafford, 1994) methods were used to analyze the sedimentation velocity experiments. These methods allow the calculation of the apparent sedimentation coefficient ( $s^*$ ), the diffusion coefficient *D*, and the molecular mass MM. The ratio of the sedimentation coefficient to diffusion coefficient gives the molecular mass by use of the follow equation:

$$MM = \frac{sRT}{D(1 - Vbar\rho)}, \quad (1)$$

where *R* is the gas constant and *T* is the absolute temperature. The software Sednterp ([www.jphilo.mailway.com/download.htm](http://www.jphilo.mailway.com/download.htm)) was used to estimate protein partial specific volume at 20 °C ( $Vbar = 0.7329$  mL/g), buffer density ( $\rho = 1.00087$  g/mL), and viscosity ( $\eta = 1.0126 \times 10^{-2}$  P), and the  $s_{max}$  for globular proteins of about 41.8, 83.6, and 167.2 kDa. This software was also used to estimate the standard sedimentation coefficients ( $s_{20,w}$ ) for each protein concentration that were used to estimate the  $s_{20,w}$  at 0 mg/mL protein concentration by extrapolation ( $s_{20,w}^0$ ). That procedure minimizes interferences caused by temperature, viscosity solution, and molecular crowd (Laue, 2001).

The sedimentation equilibrium experiments were made at 20 °C at speeds of 8000, 10000, and 12000 rpm with the AN-60Ti rotor, scan data acquisition at 236 nm and at protein concentration of 300, 450, and 600 µg/mL. Briefly, the samples were accelerated to 8000 rpm and scans were collected until no changes in the scan were observed. Three scans were obtained at 8000 rpm. Similar procedures were made at 10000 and 12000 rpm. The self-association method was used to analyze the sedimentation equilibrium experiments using several models of association to fit the *MtCS* data. The distribution of the protein along the cell, obtained in the equilibrium sedimentation experiments, was fitted with the following equation (Johnson et al., 1981):

$$C = C_0 e^{\left[ \frac{M(1 - V\bar{\rho})\omega^2(r^2 - r_0^2)}{2RT} \right]}, \quad (2)$$

where  $C$  is the protein concentration at radial position  $r$ ,  $C_0$  is the protein concentration at radial position  $r_0$ , and  $\omega$  is the centrifugal angular velocity.

The HydroPro software (García de la Torre et al., 2000) was applied to estimate the standard sedimentation coefficient  $s_{20,w}$ , starting from the high resolution  $MtCS$  structure: as a monomer, dimer and tetramer. The HydroPro software was setup with the radius of the atomic elements of 2.5 Å, with sigma factors from 5 to 8 (as indicated by supplier) and minibeads radius (SIGMIN and SIGMAX) from 6 to 2 Å after initial evaluation of the two extremes. The parameters  $MtCS$   $V\bar{\rho}$ , and  $\rho$  and  $\eta$  (for standard conditions) were estimated using the software Sednterp as described above.

### 3. Results and discussion

#### 3.1. $MtCS$ spectropolarimetry analysis

The observed spectrum of CD data is shown in Fig. 1, and it is similar to the one obtained for CS from *Escherichia coli* ( $EcCS$ ) (Macheroux et al., 1998). The deconvolution of CD data predicted that the  $MtCS$  has nearly 30 and 18% of  $\alpha$ -helices and  $\beta$ -sheet, respectively. This result is in agreement with the crystallographic structure presented here (34 and 17% of  $\alpha$ -helices and  $\beta$ -sheet, respectively). The  $MtCS$  thermal stability is presented in Fig. 1 (inset). The data show that the  $MtCS$  has two well-characterized thermal unfolding transitions: one at 51.5 °C and another at 64 °C, which is disagreement with the data from thermal unfolding of the  $EcCS$ . The data for  $EcCS$  suggests only one unfolding transition at 54 °C (Fitzpatrick et al.,

2001), despite of the experimental conditions were different than that applied for the  $MtCS$ . Thus, the data obtained here, may suggest that the  $MtCS$  possesses two well-defined domains, each one presenting different thermal stabilities, despite  $MtCS$  crystal structure does not show the existence of these domains (see below). Then, other hypotheses is that the protein may loss its oligomeric status before undergoing the fully unfold. Furthermore, as the protein is a dimer in solution (see below), the secondary structures on the flanks of the dimerization interface core can also undergo unfolded at 51.5 °C and the core lost its secondary structure at 64 °C. The  $MtCS$  unfolding mechanism is currently under investigation.

#### 3.2. $MtCS$ presents a $\beta$ - $\alpha$ - $\beta$ architecture

The crystal structure of the  $MtCS$  was solved at 2.65 Å using the Synchrotron Radiation Source (Polikarpov et al., 1998), and the structure was determined by molecular replacement. The final statistics for data processing and refinement are in Table 1. There is one molecule of  $MtCS$  in the asymmetric unit, containing residues 1–48 and 53–392 and 243 water molecules (Fig. 2A). The residues 49–52 and 393–401 did not present electronic density, and they were omitted in the final model. The analysis of stereochemical quality,  $G$ -factor and 3D-profile are presented in Table 1.

The structure of  $MtCS$  is similar the other CS structures deposited in the Protein Data Bank (PDB) (PDB access code: 1QXO, 1Q1L, 1UMO, and 1R53; Ahn et al., 2004; Maclean and Ali, 2003; Quevillon-Cheruel et al., 2004; Viola et al., 2004, respectively). The structure of the  $MtCS$  monomer belongs to the  $\alpha/\beta$  class. Each monomer of  $MtCS$  contains 13  $\alpha$ -helices and 17  $\beta$ -sheets, which fold into an approximate dimension of 75 Å × 47 Å × 36 Å.  $MtCS$  presents only one domain and its dominant structural topology is a  $\beta$ - $\alpha$ - $\beta$  sandwich, in which each monomer of  $MtCS$  contains a central helical core (formed by helices  $\alpha 1$ ,  $\alpha 5$ ,  $\alpha 11$ , and  $\alpha 8$ ). This helical core is sandwiched between two four-stranded antiparallel  $\beta$ -sheets ( $\beta 1$ ,  $\beta 2$ ,  $\beta 7$ ,  $\beta 4$ , and  $\beta 8$ ,  $\beta 9$ ,  $\beta 15$ ,  $\beta 10$ ), which are surrounded by loops and discrete stretches of  $\alpha$ -helix and  $\beta$ -sheet. The topology of  $MtCS$  is shown in Fig. 2B and can be observed that the core of structure of  $MtCS$  has pseudo 2-fold symmetry, despite the molecule does not present such symmetry.

The crystal structure obtained for  $MtCS$  may be described as a tetramer formed by a dimer of dimers, presenting approximately 222 symmetry. The intermolecular interactions among the monomers A and D form the first dimer and those among the monomers B and C forming the second dimer. But each monomer of CS is in contact with three others, creating an intricate packing arrangement (Fig. 3). The contact areas between the monomers are 796.0, 1944.0, and 3977.0 Å<sup>2</sup> for the unit pairs AB/CD, AC/BD, and AD/BC, respectively, and many residues involved on the interface of tetramerization are conserved

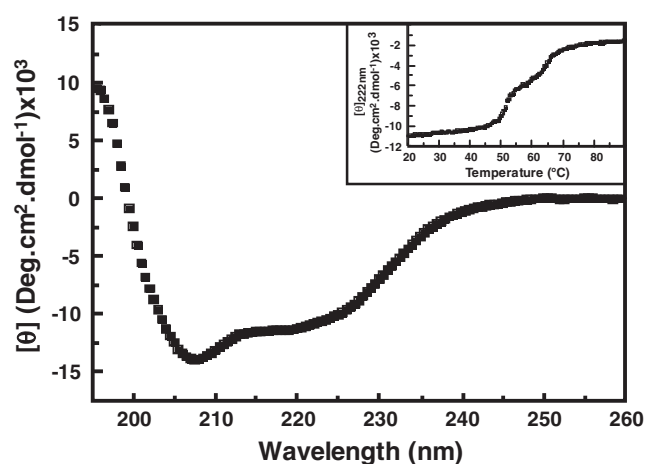


Fig. 1. Circular dichroism analyses. Residual molar ellipticity of the  $MtCS$  was measured from 195 to 260 nm in 20 mmol/L Tris-HCl buffer (pH 8.0) containing 50 mmol/L of NaCl, and 1 mmol/L  $\beta$ -mercaptoethanol at 20 °C. The amount of secondary structure was estimated by CDNN Deconvolution software (Böhm et al., 1992). Thermal stability of  $MtCS$  was analyzed by CD at 222 nm.

Table 1  
Crystallographic data, refinement statistics, and analysis of the quality from *MtCS*

1° Crystal data	
Unit-cell parameters	$a = b = 129.74$ , $c = 159.77$
Space group	$P6_422$
Number of measurements with $I > 2 \sigma(I)$	92610
Number of independent reflections	19341
Completeness (%) (outermost shell)	97.90 (97.9)
$R_{\text{sym}}^a$ (%) (outermost shell)	5.40 (16.50)
Highest resolution shell (Å)	2.94–2.65
2° Refinement data	
Resolution range (Å)	52.93–2.65
Reflections used for refinement	20093
Number of atoms	3083
Number of residues	388
Number of O atoms	243
Final $R$ -factor (%) <sup>b</sup>	16.2
Final $R$ -free final (%) <sup>c</sup>	22.1
Correlation coefficient $F_o - F_c$ (%)	95.6
$B$ values (Å <sup>2</sup> ) <sup>d</sup>	
Main chain	33.81
Side chain	36.05
3° Quality of structure	
Observed RMSD from ideal geometry <sup>e</sup>	
Bonds lengths (Å)	0.021
Bonds angles (°)	2.655
$G$ -factor <sup>e</sup>	
Torsion angles	−0.49
Covalent geometry	−0.15
Overall	−0.27
3D profile <sup>f</sup>	( $S$ ) = 170.39; IS = 177.45; S/IS = 0.96 IS
Ramachandran plot (%)	
Favorable	92.2
Additional allowed	7.5
Generously allowed	0.0
Disallowed	0.3

<sup>a</sup>  $R_{\text{sym}} = 100 \sum |I(h) - \overline{EI(h)}| / \sum I(h)$  with  $I(h)$ , observed intensity and  $\overline{EI(h)}$ , mean intensity of reflection  $h$  overall measurement of  $I(h)$ .

<sup>b</sup>  $R$ -factor =  $100 \times \sum (|F_{\text{obs}} - F_{\text{calc}}|) / \sum (F_{\text{obs}})$ , the sums being taken over all reflections with  $F/\sigma(F) > 2\sigma(F)$ .

<sup>c</sup>  $R$ -free =  $R$ -factor for 10% of the data that were not included during crystallographic refinement.

<sup>d</sup>  $B$  values = average  $B$  values for all non-H atoms.

<sup>e</sup> Ideally, scores should be above −0.5. Values below −1.0 may need investigation.

<sup>f</sup> Total score is the sum of the 3D–1D scores (statistical preferences) of each residue present in protein. Ideal score  $S_{\text{ideal}} = \exp(-0.83 + 1.008 \times \ln(L))$ ; where  $L$  is the number of amino acids.  $S_{\text{ideal}}$  score is compatibility of the sequence with their 3D structure. It is obtained from total score/ideal score.  $S_{\text{ideal}}$  score above  $0.45S_{\text{ideal}}$ .

in several sequences from bacteria (Fig. 4). Fig. 5 introduces the interface contact area between the monomers pairs AB, AC, and AD, emphasizing the molecular surface.

The dimers formed with monomers A/D or B/C of CS are characterized by a large antiparallel eight-stranded  $\beta$ -sheet that is formed by combination of two sheets, one from each monomer (Fig. 5A). This large  $\beta$ -sheet is the major feature of the *MtCS* dimer interface and may be said that the two strands  $\beta$ 10 are the principal region for the

dimer stabilization. Another secondary structure element involved in stabilization of the dimer is the helix  $\alpha$ 8 that interacts with  $\alpha$ 8 from another monomer. These interactions contribute highly to dimer stabilization, presenting considerable amount of hydrophobic interactions on the surface at the dimer interface. The dimer also presents two additional four-strand antiparallel  $\beta$ -sheet, one of each side of the core formed by association between the  $\beta$ 11 and  $\beta$ 12 from one monomer and  $\beta$ 16 and  $\beta$ 17 of another monomer. Moreover, the dimer interface presents several interactions (44 hydrogen bonds), mainly involving loops and the  $\alpha$ 3 and  $\alpha$ 4.

The main residues involved in the *MtCS* dimerization are Lys123, Tyr124, Arg135, Gly171, Asp217, Gly220, Gly233, Gly235, Asp245, Ile255, Lys259, Gly274, and Pro316, which are conserved in almost all CS sequences from bacteria including the four-solved structures (Fig. 4, see below). These results suggest that probably all CS are at least dimeric due the high number of hydrogen bonds and conservation of the residues involved in this interaction. Fig. 4 shows the alignment of 16 sequences of CS from bacteria evidencing the conserved residues involved on the interface of two monomers A and D or B and C, which form the dimer.

The *MtCS* crystal structure clearly shows that the interface interactions between the dimer AD and BC, which form of tetrameric structure of CS involve both hydrophobic and polar residues, and some regions present complementary of charge.

On the other hand, the forces that stabilize the interaction between the monomers A and B are made mainly by two salt-bridges and six hydrogen bonds localized in the following regions:  $\beta$ 2,  $\alpha$ 7,  $\alpha$ 11,  $\alpha$ 12, and L29, mainly between  $\alpha$ 12 from one monomer with the adjacent from other monomer. The number of the interactions between the monomers A and C is higher than those observed for monomers A and B. The interaction between monomer A and C is made mainly by 18 hydrogen bonds localized principally in regions of loops. The protein regions responsible for these interactions are: L1,  $\alpha$ 1, L5,  $\beta$ 5, L6,  $\beta$ 7, L8, L9,  $\alpha$ 3, L10, and L29. The predominant region of interaction between monomer A and C is L8, which interact with the L8 from the adjacent monomer.

### 3.3. Comparison between *MtCS* structure and others CS structures

The monomeric structure of *MtCS* was compared with five other CS structures from four organisms: three from bacteria (*Aquifex aeolicus*—*AaCS*; *Streptococcus pneumoniae*—*SpCS* and *Helicobacter pylori*—*HpCS*) and one from yeast (*Saccharomyces cerevisiae*—*ScCS*) deposited in the PDB (PDB access code: 1Q1L, 1QXO, 1UMO, and 1R53, respectively) (Ahn et al., 2004; Maclean and Ali, 2003; Quevillon-Cheruel et al., 2004; Viola et al., 2004, respectively). Maclean and Ali (2003) reports the *SpCS* structure (which shares with *MtCS*, on their amino acid

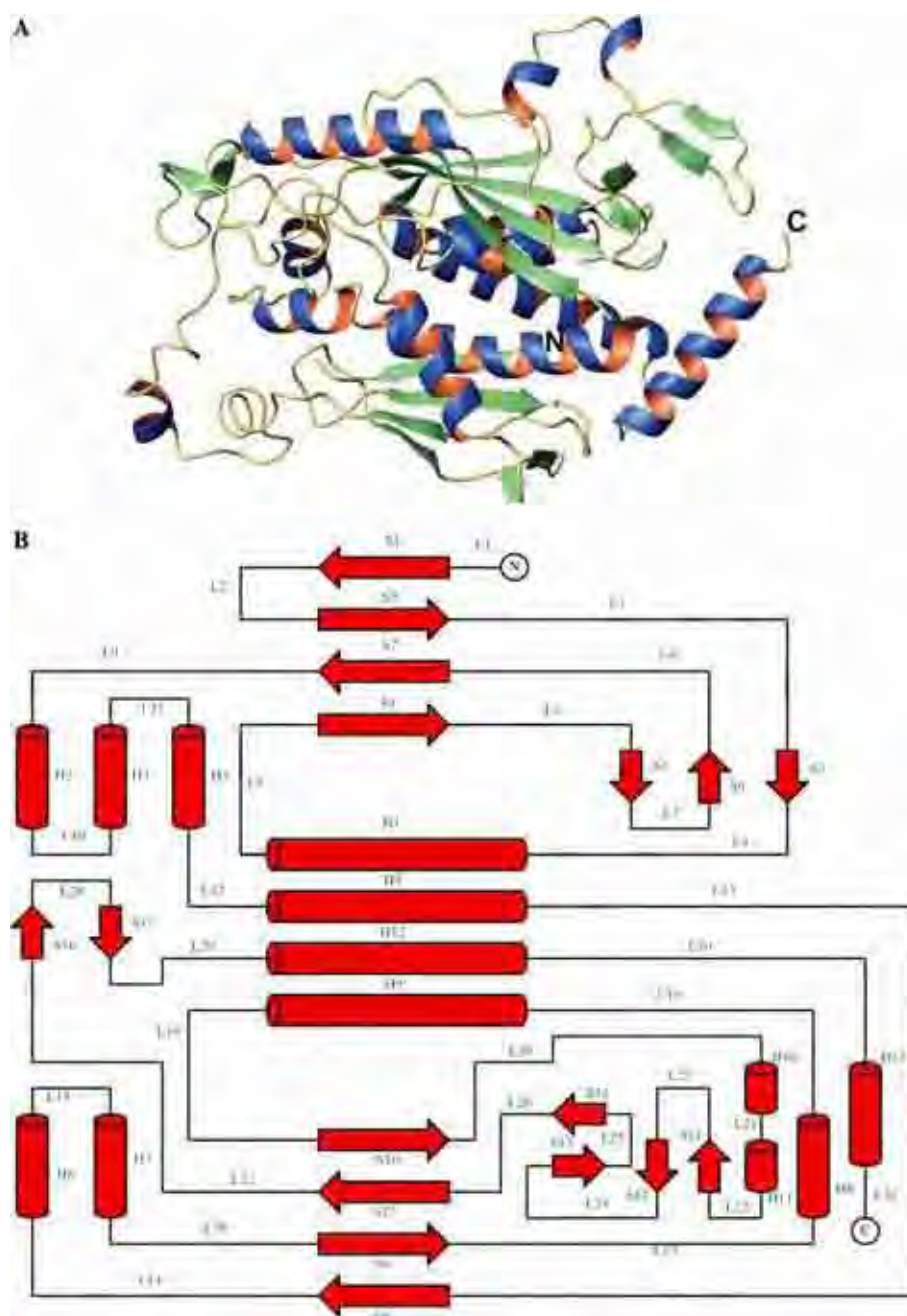


Fig. 2. *MtCS* monomer structure. (A) Representation of the monomer *MtCS* secondary structure elements. (B) *MtCS* secondary structure representation.

sequence, 46 and 52% of identity and similarity, respectively) in two conformational states. In the monomer A, the loop 22 is opened over the FMN binding site and in the other monomers (B, C, and D), the loop 22 is closed over FMN binding site. Fig. 6 (A to D) shows the superposition of C $\alpha$  atoms of the *MtCS* with others solved structures. Fig. 6 demonstrates that *MtCS* core is similar to other structures. However, some dislocated regions can be observed, mainly the H7 in *MtCS*. In the superposition of the *MtCS* with *HpCS*, it can be observed that the H13 in *MtCS* is in opposing direction. Furthermore, superposition of *MtCS* against other CS structures indicates highest RMSD for *HpCS* and *ScCS*, possibly due to low identity of

these structures with *MtCS* (31 and 33% of identity, respectively). On the other hand, the *AaCS* and *SpCS* structures present higher identity with *MtCS* (45 and 46% of identity, respectively) and therefore a major structural resemblance. Furthermore, the *MtCS* structure presents highest resemblance with the monomer B, C, and D from *SpCS* (which are in the closed conformation—Fig. 6D). The residues of FMN binding site of complex *HpCS*:FMN also present different conformations when compared with *MtCS*. The position of residues of active site of *HpCS* resemble with the structure of monomer A of *SpCS*. The FMN binding site from *AaCS* and *ScCS* structures are not defined and they can not be compared with that region from *MtCS*.

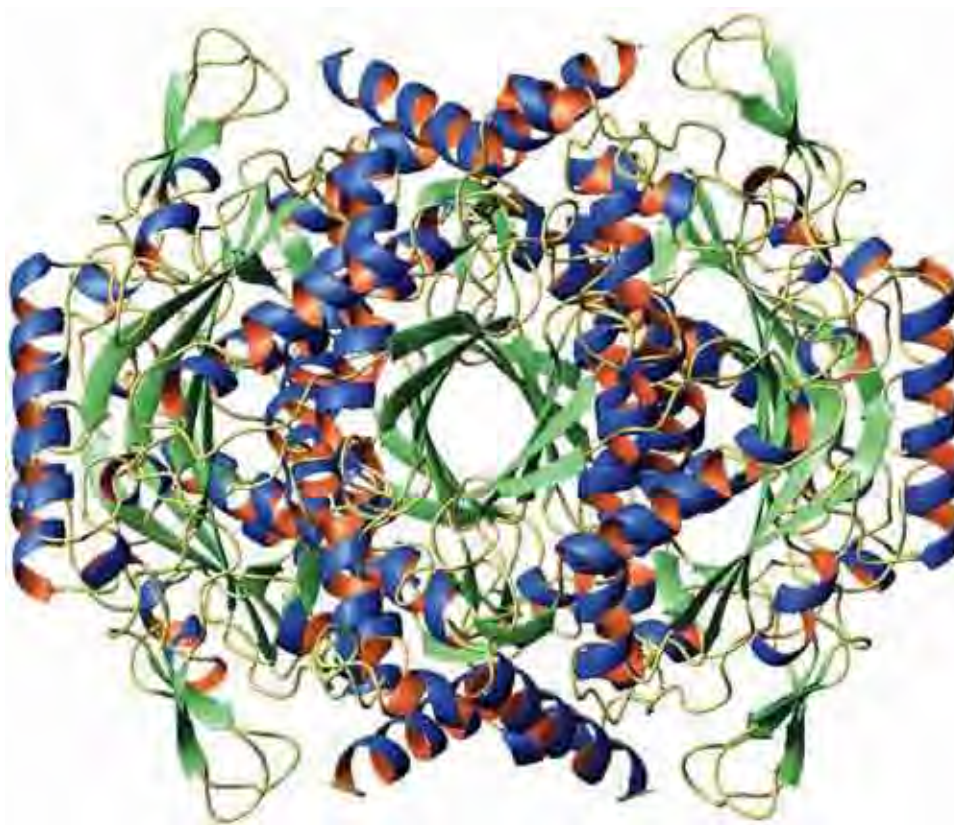


Fig. 3. *MtCS* tetrameric structure. The figure represents the *MtCS* tetrameric structure with its secondary structure content.

This information lead us conclude that in *MtCS*, the structure may be in its closed conformation, and in the presence of FMN, the loop 29 (in *MtCS*) moves away opening the EPSP binding site. In the EPSP bound state, the loop 29 from *MtCS* molecule must close over the ligand. The *SpCS* structure has confirmed that the binding of FMN and EPSP is ordered (Maclean and Ali, 2003), so as Macheroux et al. (1996, 1998) observed in experiments for *EcCS* that oxidized FMN which has a  $K_D$  of about 30  $\mu\text{M}$ , and decreases to around 20 nM in the presence of EPSP. Therefore, it is clear that the binding of EPSP blocks any possible exit of FMN from the active site, and the induced structural changes in the active site allow FMN to make further interactions with *SpCS* (Maclean and Ali, 2003).

The residues of FMN and EPSP binding active site were superposed with the same residues in the two conformational states of *SpCS* (Figs. 7A and B). The positions of many residues are conserved, but there are significant differences in the positions of some residues, mainly the

Arg46 and Arg341 at the EPSP binding site and in Arg110 and Ile317 at FMN binding site. The His11, an essential residue to CS catalysis (Kitzing et al., 2004) is conserved in several sequences (Fig. 4) and presents similar conformation in *MtCS* and *SpCS* (in the both *SpCS* conformational states). The His115 in the *MtCS* structure, another conserved and important residue to CS catalyzed reaction (Kitzing et al., 2004), also keeps its position, when compared with the *SpCS* open structure. This same residue, when compared to the closed structure of *SpCS* presents dislocated approximately 30° outward from the EPSP binding site (Fig. 7). In *SpCS*, the His110 (equivalent the His115 in *MtCS*) demonstrated therefore some conformational flexibility and interacts with FMN in the open form but with EPSP in the closed form (Maclean and Ali, 2003). The closing of the loop 22 on the closed structure of *SpCS* can be the cause of this displacement. The Arg341, located in the loop 29 in *MtCS* (relative to loop the 22 in *SpCS*) present the  $C\alpha$  closer to the closed

Fig. 4. Sequence alignment of several sequences of CS. Alignment was performed with MultAlin program (Corpet, 1988). Numbering is relative to *MtCS* primary sequence. Residues shown in red have a correlation of at least 90% with the consensus sequence and are therefore highly conserved. Residues shown in blue have 50% or more correlation with the consensus sequence and are therefore moderately conserved. Residues in (—) boxes are relative the active site. Residues in (- - - - -) boxes are relative the dimerization interface. Residues in (· · · · ·) boxes are relative the tetramerization interface. Myctub (*Mycobacterium tuberculosis*); Aquaeo (*Aquifex aeolicus*); Bacsub (*Bacillus subtilis*); Laclac (*Lactococcus lactis*); Strpne (*Streptococcus pneumoniae*); Lismon (*Listeria monocytogenes*); Staaar (*Staphylococcus aureus*); Brumel (*Brucella melitensis*); Esccol (*Escherichia coli*); Saltyp (*Salmonella typhimurium*); Yerpes (*Yersinia pestis*); Haeinf (*Haemophilus influenzae*); Xilfas (*Xylella fastidiosa*) Neimen (*Neisseria meningitidis*); Camjej (*Campylobacter jejuni*); Clamur (*Chlamydia muridarum*).



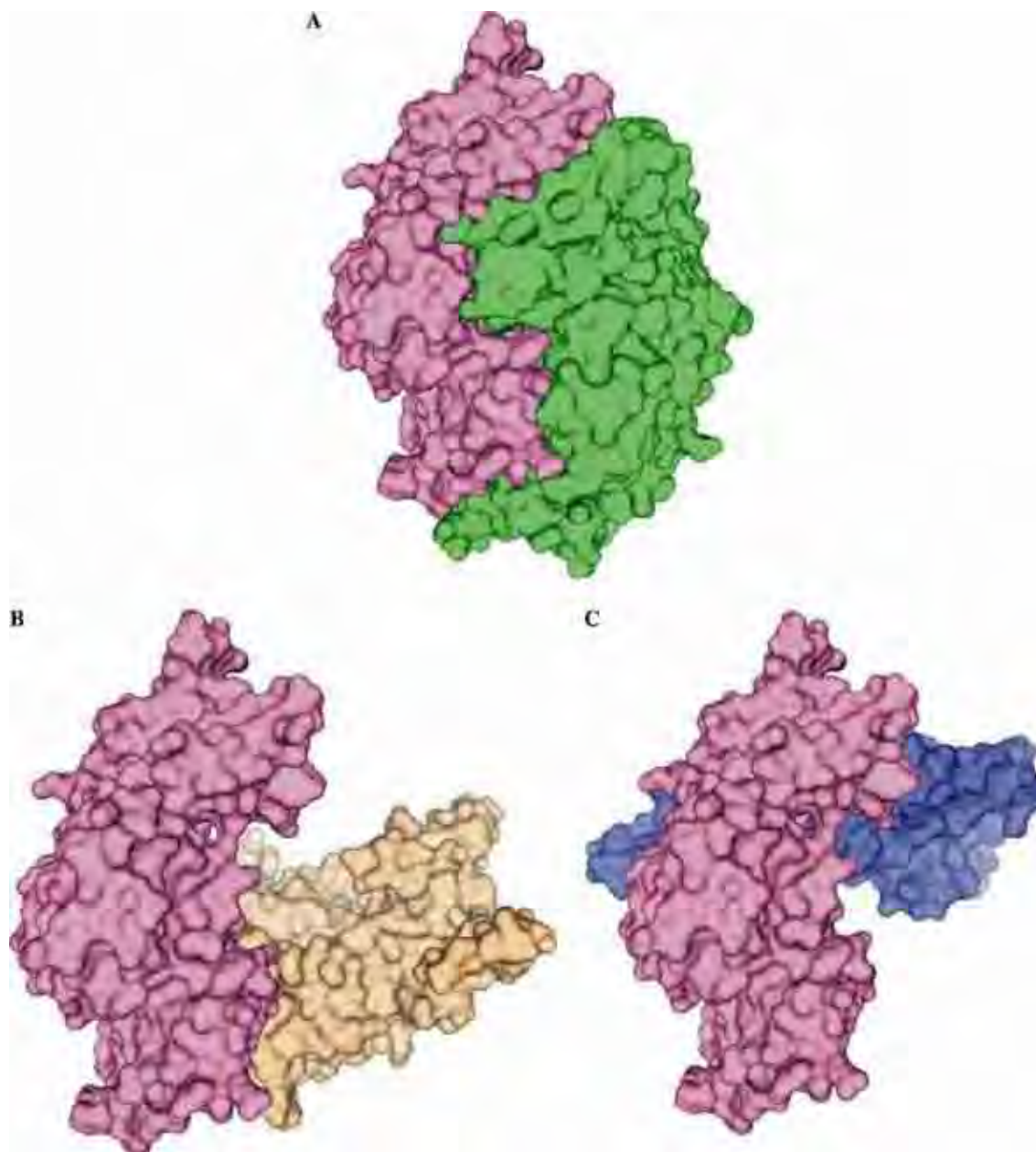


Fig. 5. *MtCS* dimeric structure and interface. (A) Figure shows the molecular surface of *MtCS* dimer formed by monomer A (pink) and D (lime). (B) Representation of the molecular surface on the interface between monomers A (pink) and B (wheat). (C) Representation of the molecular surface on the interface between monomers A (pink) and D (slate). The figures were generated for Pymol program (DeLano, 2002). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

structure of *SpCS*, but its side chain is coming from other direction, occupying the EPSP binding site while in closed structure of *SpCS*, it presents dislocated outward from the site. This effect occurs probably due to the EPSP binding at its site, therefore in *HpCS*, this arginine residue is moved away from the EPSP binding site (data not shown). The different orientations of important residues of active site of *MtCS*, as Arg341, Ser342, and Asp343, when compared with the structure of *SpCS*, likely occurs due to apo state of *MtCS*.

The arginine residues 112 and 139, which are involved in the formation of binding pocket of EPSP by the aliphatic portions from *MtCS* structure do not present significant

modifications when compared with *SpCS*. On the other hand, the Arg40 of the *MtCS* is more distended upon the EPSP binding site than the same arginine in the two *SpCS* conformational states, showing that the Arg40 presents some flexibility. This observation can also be justified by absence of the EPSP in the *MtCS* structure. The arginine residues 46 and 49 that are important to the EPSP binding (Maclean and Ali, 2003) are localized at the loop 5, which was not determined with precision probably due to the absence of interactions of protein with EPSP.

The largest differences in FMN binding site are observed in the position of Arg110 and Thr319, where

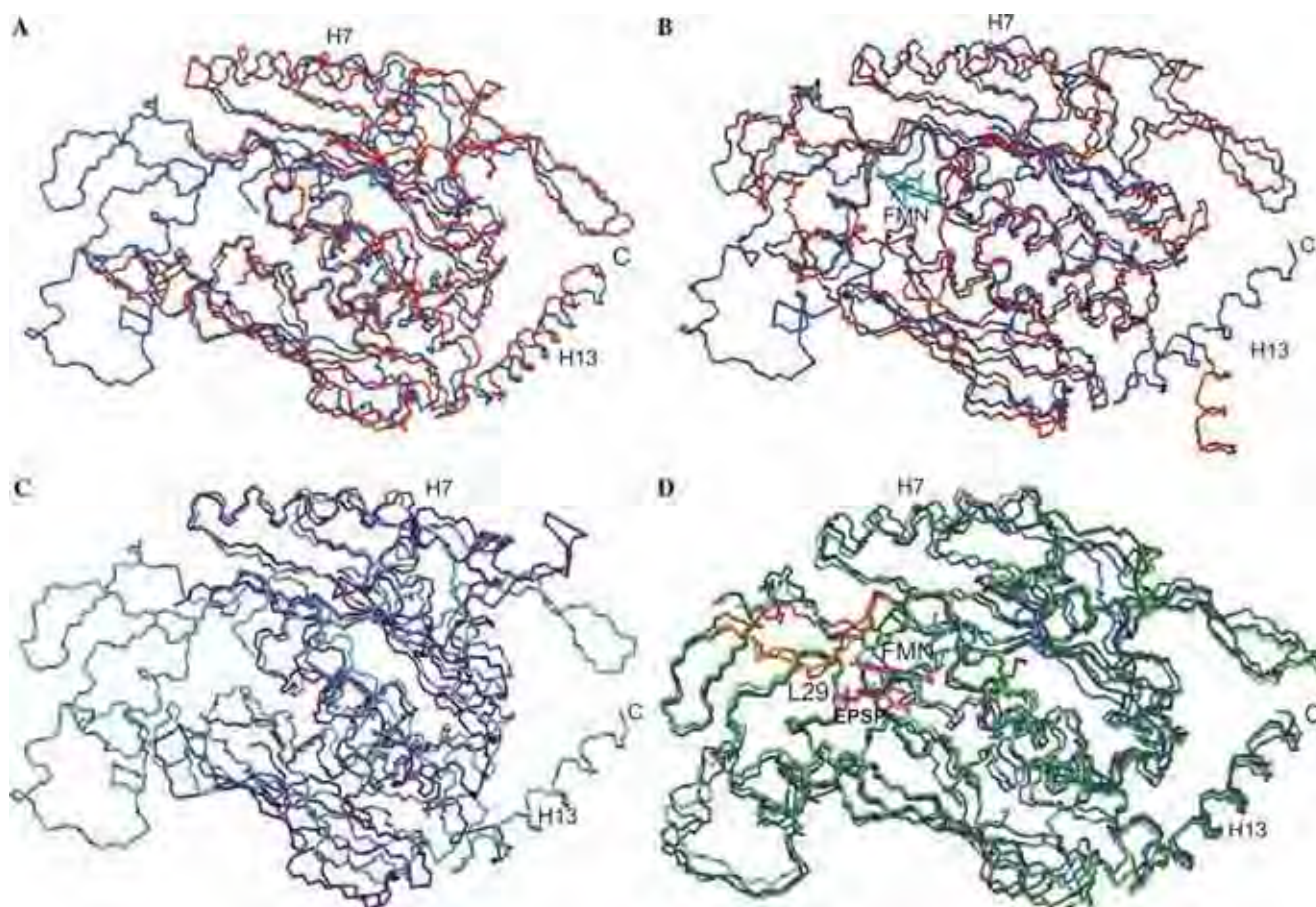


Fig. 6. Superposition of C $\alpha$  between MtCS and others CS. (A) from *A. aeolicus*; (B) *H. pylori*; (C) *S. cerevisiae*, and (D) *S. pneumoniae*. Structure in light blue is MtCS, and in light green is monomer B of SpCS. Structure in dark green is monomer A of SpCS; Region in yellow is loop 22 of SpCS of the monomer B; region in light red is loop 22 of SpCS of the monomer A; Structure in violet is loop 29 of MtCS. H7 and H13 is referent of MtCS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

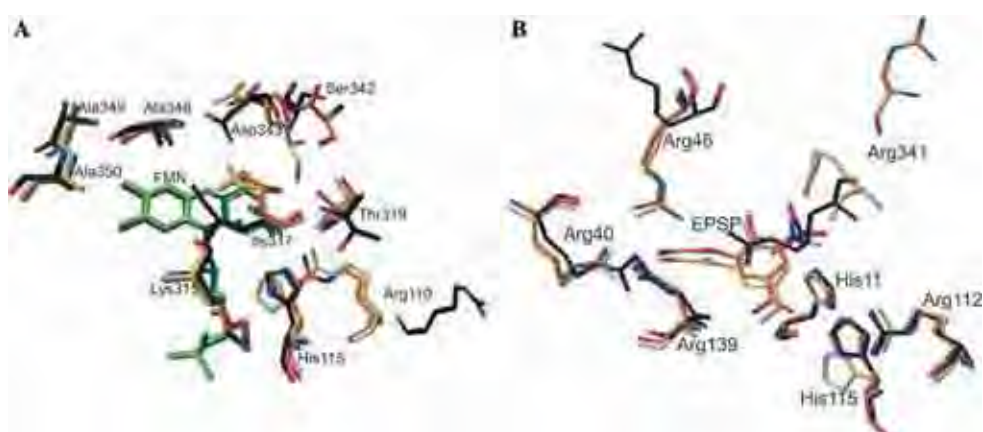


Fig. 7. Superposition of residues of the (A) FMN binding site and (B) EPSP binding site. Residues in black are from MtCS, yellow are from monomer B of the SpCS and in orange is from monomer A of the SpCS. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

in MtCS structure, their lateral chains are shifted outward from the FMN binding site. The side chains of Ile37 and Lys315 in MtCS are occupying the FMN bind-

ing site. The Ser342 and Asp343 also present different orientations when compared with SpCS, may be due to MtCS to be in apo state.



### 3.4. Quaternary structure in solution

Gel filtration column was performed to determine the native molecular weight of *MtCS*. The results showed that *MtCS* is a dimer in solution.

The hydrodynamic properties of the *MtCS* were established by both sedimentation velocity and sedimentation equilibrium. The samples did not present apparent aggregation and it shows monodisperse behavior. The sedimentation coefficient  $s$  normalized to the standard sedimentation coefficient  $s_{20,w}$  and extrapolated at 0 mg/mL ( $s_{20,w}^0$ ) is  $4.60 \pm 0.02$  S which is in accordance with a particle with molecular mass (MM) of 80–100 kDa, and of asymmetrical shape. Furthermore, considering the diffusion coefficient at standard conditions ( $3.7\text{--}3.8 \times 10^{-7}$  cm<sup>2</sup> seg<sup>-1</sup>) determined by dynamic light scattering and by the application of the Eq. (1), a MM of *MtCS* was  $80 \pm 2$  kDa, a value closer of the expected for its dimer (MM of monomer is 41.7 kDa). The MM derived from the sedimentation equilibrium was  $83 \pm 1$  kDa and it is also in accordance with the MM expected for the *MtCS* dimer.

However, the resultant plot (Fig. 8A, inset) suggests that  $s_{20,w}$  increases with the *MtCS* concentration, instead to reduce with the viscosity and molecular crowd effect (Laue, 2001). One hypothesis that can explain this behavior of *MtCS* is a rapid dimer–tetramer equilibrium (Rowe, 1977), which should be concentration dependent.

Furthermore, the results of sedimentation equilibrium also suggest the existence of dimer–tetramer equilibrium (Fig. 8B). The association constant determined for sedimentation equilibrium is approximately 20 mM<sup>-1</sup>, and therefore, the dissociation constant— $K_D$ —about 50 M.

Thus, due to very high  $K_D$ , probable the interactions that form the tetramer are few or/and very weak in these experimental conditions. *SpCS*, on the other hand, also shows dimer–tetramer equilibrium with dissociation constant of 0.8  $\mu$ M (Maclean and Ali, 2003). So, we can conclude that in solution, the *MtCS* is predominantly dimer while *SpCS* is tetramer, despite of their structural resemblance.

The comparison between the predicted  $s_{20,w}$  for *MtCS* crystal structures and the  $s_{max}$  for a globular protein relative MM correspondents (Table 2) shows that the *MtCS* in solution can have an asymmetric shape. Through of the sedimentation equilibrium and sedimentation velocity data of *MtCS*, we can considerate the predicted  $s_{0,w}$  for future analysis. But, the experimental  $s_{20,w}^0$  data for *MtCS* do not agree with the predicted data for crystallographic dimer formed for monomers A/D or B/C of this structure. These ambiguous results suggest that, in solution, the dimer of *MtCS* is slightly different from the crystallographic dimers (A/D or B/C) that form the tetramer, presenting a more asymmetrical shape.

These differences of the oligomeric states observed for *MtCS* when compared with other CS structures may not

Table 2  
Comparison between the predicted  $s_{20,w}$  for *MtCS* crystal structure and the  $s_{max}$  estimated for a globular protein

Oligomeric state	$s_{20,w}$ (S)	$s_{max}$ (S)
Monomer	$3.18 \pm 0.05$	4.20
Dimer	$5.42 \pm 0.06$	6.70
Tetramer	$9.02 \pm 0.06$	10.70

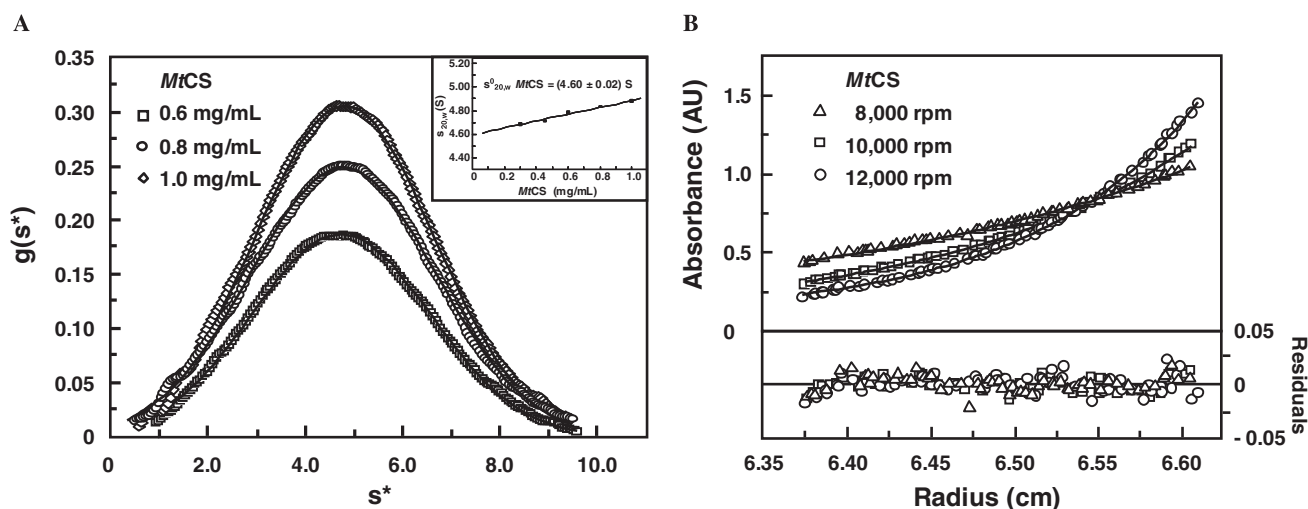


Fig. 8. Sedimentation velocity experiments. (A) Sedimentation velocity experiments: The  $g(s^*)$  distributions were fitted using the Origin (Microcal Software) with a Gaussian giving apparent sedimentation coefficients  $s^*$  for *MtCS* at different protein concentrations (see Section 2 for details). Inset: plots of  $s_{20,w}^0$  versus protein concentration fitted by linear regression to calculate the  $s_{20,w}^0$ :  $4.60 \pm 0.04$  Svedberg. Despite the data were well fitted by only one Gaussian suggesting the existence of only one *MtCS* specie in solution, the positive slop suggests that the *MtCS* undergoes an oligomerization process. (B) Sedimentation equilibrium experiments. The figure shows the best fits of experimental data for 450  $\mu$ g/mL of *MtCS* at 8000, 10000, and 12000 rpm with the self-association methods (see Section 2). The random distribution of the residuals (bottom panel) indicates that the fit is satisfactory. The fitness of the data obtained at other protein concentration presented similar results (data not shown). Sedimentation equilibrium data of *MtCS* agree with a dimer structure with  $83 \pm 1$  kDa in equilibrium with a tetrameric form with an affinity constant of  $20 \pm 5$  mM<sup>-1</sup>.

Table 3  
Contact analysis of interface area between the dimers that form the CS tetramer

Structure	<i>Mycobacterium tuberculosis</i>	<i>Streptococcus pneumoniae</i>	<i>Helicobacter pylori</i>	<i>Aquiflex aelicus</i>	<i>Saccharomyces cerevisiae</i>
Overall contact surface (Å <sup>2</sup> )	5027.0	12132.0	5305.0	4425.0	5745.0
Contact surface apolar (Å <sup>2</sup> ) (%)	2661.0 (53)	7303.0 (60)	3369.0 (63)	2658.0 (60)	3653.0 (63)
Contact surface hydrophilic polar (Å <sup>2</sup> ) (%)	2338.0 (47)	4825.0 (40)	1936 (37)	1737.0 (40)	2092.0 (37)
Saline bridges (number)	2	6	1	2	0
Hydrogen bonds (number)	24	40	26	27	22

The contact surface was calculated by AREA/AREAIMOL (CCP4, 1994) and saline bridges and hydrogen bonds were identified for protein–protein interaction (<http://www.biochem.ucl.ac.uk/bsm/PP/server>), on the basis of principles described by Jones and Thornton (1996).

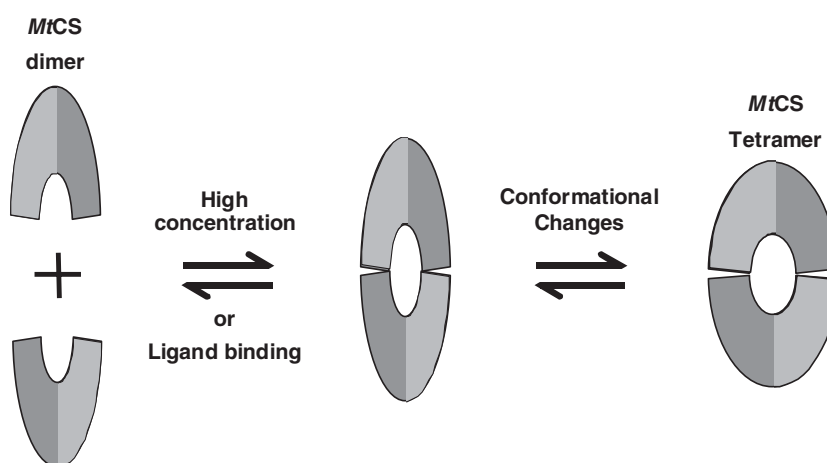


Fig. 9. MtCS oligomerization mechanism. Initial contacts on the MtCS dimer can be induced for protein concentration or ligand binding, would lead to some conformational changes increasing the dimers contact area and resulting on MtCS tetramerization.

be justified by number of saline bridges and hydrogen bonds in the dimer interface area, which form the tetramer. Table 3 shows the analysis of the tetramerization interface for five microorganisms. It can be observed that the number of saline bridges varies is between 0 and 6 and the number of hydrogen bonds is between 6 and 22 in the analyzed structures. Analysis of contact area between dimers reveals that all structures present contact interfaces above of 4000 Å<sup>2</sup>, but *SpCS* present a tetramerization interface of 12,132.0 Å<sup>2</sup>. Therefore contacts between dimers that form the tetramer of *SpCS* are larger than those observed in the other structures. The most striking difference between *MtCS* and other CS structures is the area percentage of polar atoms, which participate in the interface regions. *MtCS* present 53% of hydrophobic area (referent to the carbon atoms) while the other CS structures present a minimum 60%. Therefore, we can affirm that the interface between the dimers that forms tetramer in *MtCS* is more polar and hydrophilic when compared with other CS structures. However, this result can not justify the predominance of dimer in solution, but it shows that the hydrophobic effect has a minor participation in the formation of tetramer in *MtCS* structure. On the other hand, in some way, a minor hydrophobic surface can contribute to predominance of dimer in solution. Furthermore, a hypothesis to explain the low presence of tetramer in solution is that *MtCS* tetramerization may involve small initial

contacts, which lead to some conformational changes increasing the contact area between dimers, resulting on *MtCS* tetramerization. The initial contacts can be concentration-dependent and/or still due to the presence of some ligand, which may trigger the conformational changes resulting in the oligomerization (Fig. 9).

#### 4. Conclusion

Here, we report the first crystal structure and analysis in solution of the chorismate synthase from *M. tuberculosis*. The CD analyses show that *MtCS* secondary structures content agree with the CS structures from other organisms. Furthermore, *MtCS* thermal stability experiments show two well defined transitions, suggesting a complex unfolding mechanism. The monomeric and tetrameric structures of *MtCS* are similar to other CS structures. Many of conserved residues in the primary sequence of the active site of CS present conserved conformations in all structures solved, but some residues have large conformational changes when compared to other structures. All differences and resemblances observed in the *MtCS* structure when compared to the *SpCS* and *HpCS* structures may be due to absence of any ligand in the active site from *MtCS*. The determination of *MtCS* structure in the presence of FMN, EPSP, and chorismic acid may confirm these evidences. The intricate association between two monomers

of the *MtCS* leads to dimer organization, connecting to form a tetramer, which may be described as a dimer of dimers. In solution, the *MtCS* must be mainly in dimeric state that is slightly different from crystallographic dimer, which forms the tetramer. Analysis of all CS led us to conclude that all CS must present quaternary structure in solution in the dimeric or tetrameric forms or in equilibrium of them.

Despite of the CS structure presented here is in apo form, the structural data obtained are enough relevant. This is the first structure in the apo form that presents almost all residues of active site, excepting the residues 49–52. In this form, it is possible to observe several movements of the important residues of active site of CS. Furthermore, this is the first CS predominantly dimeric in structure, which can be due to the amount of interaction between two dimers, which form the tetrameric structure.

The PDB accession code for the apo crystal structure of Chorismate synthase from *Mycobacterium tuberculosis* is 1ZTB.

### Acknowledgments

We thank Fundação de Amparo à Pesquisa do Estado de São Paulo (SMOLBNet, Proc. 01/07532-0 and 03/12472-2), Conselho Nacional de Desenvolvimento Científico e Tecnológico and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior for fellowships and financial support. We thank the LNLS technical staff for assistance.

### References

- Ahn, H.J., Yoon, H.J., Lee, B., Suh, S.W., 2004. Crystal structure of chorismate synthase: a novel FMN-binding protein fold and functional insights. *J. Mol. Biol.* 336, 903–915.
- Azevedo Jr., W.F., Canduri, F., Oliveira, J.S., Basso, L.A., Palma, M.S., Pereira, J.H., Santos, D.S., 2002. Molecular model of shikimate kinase from *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* 295 (1), 142–148.
- Basso, L.A., Silva, L.H.P., Fett-neto, A.G., Azevedo Jr., W.F., Moreira, I.S., Calixto, J.B., Astolfi Filho, S., Santos, R.R., Soares, M.B.P., Santos, D.S., 2005. The use of biodiversity as source of new chemical entities against defined molecular targets for treatment of malaria, tuberculosis, and T-cell mediated diseases—A review. *Memórias do Instituto Oswaldo Cruz* 100 (6), 475–506.
- Bentley, R., 1990. The shikimate pathway—metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol.* 25, 307–384.
- Böhm, G., Muhr, R., Jaenicke, R., 1992. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* 5, 191–195.
- Bornemann, S., Ramjee, M.K., Balasubramanian, S., Abell, C., Coggins, J.R., Lowe, D.J., Thorneley, R.N., 1995. *Escherichia coli* chorismate synthase catalyzes the conversion of (6S)-6-fluoro-5-enolpyruvylshikimate-3-phosphate to 6-fluorochorismate. Implications for the enzyme mechanism and the antimicrobial action of (6S)-6-fluoroshikimate. *J. Biol. Chem.* 270, 22811–22815.
- Bornemann, S., Lowe, D.J., Thorneley, R.N., 1996. The transient kinetics of *Escherichia coli* chorismate synthase: substrate consumption, product formation, phosphate dissociation, and characterization of a flavin intermediate. *Biochemistry* 35 (30), 9907–9916.
- Bornemann, S., Lawson, D.M., Thorneley, R.N., 2003. A branch point in chorismate synthase research. *Structure* 11, 1463–1465.
- Brünger, A.T., 1992. X-PLOR, a System for Crystallography and NMR. Yale University Press, London.
- Collaborative Computational Project No. 4., 1994. The CCP4 suite: programs for proteins crystallography. *Acta Crystallogr.* D 50, 760–763.
- Corpet, T.F., 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acid Res.* 16 (22), 10881–10890.
- Dosselaere, F., Vanderleyden, J.A., 2001. A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Crit. Rev. Microbiol.* 27, 75–131.
- DeLano, W.L., 2002. The PyMOL Molecular Graphics System. CA DeLano Scientific.
- Dias, M.V.B., Ely, F., Canduri, F., Pereira, J.H., Frazzon, J., Basso, L.A., Palma, M.S., De Azevedo, W.F., Santos, D.S., 2004. Crystallization and preliminary X-ray diffraction analysis of chorismate synthase *Mycobacterium tuberculosis*. *Acta Crystallogr.* D. Biol. Crystallogr. 60, 2003–2005.
- Duncan, K., 2004. Identification and validation of novel drug targets in tuberculosis. *Curr. Pharm. Des.* 10 (26), 3185–3194.
- Edelhoc, H., 1967. Spectroscopic determination of tryptophan and tyrosine in protein. *Biochemistry* 6, 1948–1954.
- Fitzpatrick, T.B., Killer, P., Thomas, R.M., Jelesarov, I., Amrhein, N., Macheroux, P., 2001. Chorismate synthase from the hyperthermophile *thermotoga maritima* combines thermostability and increased rigidity with catalytic and spectral properties similar to mesophilic counterparts. *J. Biol. Chem.* 276, 18052–18059.
- Gandy, M., Zumla, A., 2002. The resurgence of disease: social and historical perspectives on the 'new' tuberculosis. *Soc. Sci. Med.* 55 (3), 385–396.
- García de la Torre, J., Huertas, M.L., Carrasco, B., 2000. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* 78, 719–730.
- Gill, S.C., von Hippel, P.H., 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319–326.
- Goldberg, R.J., 1953. Sedimentation in the ultracentrifuge. *J. Phys. Chem.* 57, 194–202.
- Hill, R.K., Newkome, G.R., 1969. Stereochemistry of chorismic acid biosynthesis. *J. Am. Chem. Soc.* 91, 5893–5894.
- Johnson, M.L., Correia, J.J., Yphantis, D.A., Halvorson, H.R., 1981. Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* 36, 575–588.
- Jones, S., Thornton, J.M., 1996. Principles of protein–protein interactions. *Proc. Natl. Acad. Sci. USA* 93, 13–20.
- Jones, T.A., Zou, J.Y., Cowan, S.W., 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* A 47, 110–119.
- Kitzing, K., Auweter, S., Amrhein, N., Macheroux, P., 2004. Mechanism of chorismate synthase. Role of the two invariant histidine residues in the active site. *J. Biol. Chem.* 279 (10), 9451–9461.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1994. PROCHECK v.3.0—Program to Check the Stereochemistry Quality of Protein structures—Operating Instructions.
- Laue, T.M., 2001. Biophysical studies by ultracentrifugation. *Curr. Opin. Struct. Biol.* 11, 579–583.
- Leslie, A.G.W., 1992. MOSFLM version 6.11 for processing image plate and CCD data.
- Macheroux, P., Bornemann, S., Ghisla, S., Thorneley, R.N., 1996. Studies with flavin analogs provide evidence that a protonated reduced FMN is the substrate-induced transient intermediate in the reaction of *Escherichia coli* chorismate synthase. *J. Biol. Chem.* 271 (42), 25850–25858.
- Macheroux, P., Schonbrunn, E., Svergun, D.I., Volkov, V.V., Koch, M.H., Bornemann, S., Thorneley, R.N., 1998. Evidence for a major structural change in *Escherichia coli* chorismate synthase induced by flavin and substrate binding. *Biochem. J.* 335, 319–327.
- Macheroux, P., Schmid, J., Amrhein, N., Schaller, A., 1999. A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway. *Planta* 207, 325–334.

- Maclean, J., Ali, S., 2003. The structure of chorismate synthase reveals a novel flavin binding site fundamental to a unique chemical reaction. *Structure* 11, 1499.
- McRee, D.E., 1999. XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* 125, 156–165.
- Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., D* 53, 240–255.
- Navaza, J., 2001. Implementation of molecular replacement in AMoRe. *Acta Crystallogr. D. Biol. Crystallogr.* 57 (Pt. 10), 1367–1372.
- Pereira, J.H., Canduri, F., Oliveira, J.S., Silveira, N.J.F., Basso, L.A., Palma, M.S., Azevedo Jr., W.F., Santos, D.S., 2003. Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* 312 (3), 608–614.
- Pereira, J.H., Oliveira, J.S., Canduri, F., Dias, M.V.B., Palma, M.S., Basso, L.A., Santos, D.S., Azevedo Jr., W.F., 2004. Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. *Acta Crystallogr. D. Biol. Crystallogr.* 60 (Pt. 12, Pt. 2), 2310–2319.
- Polikarpov, I., Perles, L.A., de Oliveira, R.T., Oliva, G., Castellano, E.E., Garratt, R.C., Craievich, A., 1998. Set-up and experimental parameters of the protein crystallography beam line at the Brazilian National Synchrotron Laboratory. *Nucl. Instrum. Methods Phys. Res. A* 405, 159–164.
- Quevillon-Cheruel, S., Leulliot, N., Meyer, P., Graille, M., Bre-mang, M., Blondeau, K., Sorel, I., Poupon, A., Janin, J., van Tilbeurgh, 2004. Crystal structure of the bifunctional chorismate synthase from *Saccharomyces cerevisiae*. *J. Biol. Chem.* 279, 619–625.
- Rowe, A.J., 1977. The concentration dependence of transport processes: a general description applicable to the sedimentation, translational diffusion, and viscosity coefficients of macromolecular. *Biopolymers* 16, 2595–2611.
- Silveira, N.J.F., Uchôa, H.B., Pereira, J.H., Canduri, F., Basso, L.A., Palma, M.S., Santos, D.S., Azevedo Jr., W.F., 2005. Molecular models of protein targets from *Mycobacterium tuberculosis*. *J. Mol. Model.* 11 (2), 160–166.
- Stafford, W.F., 1994. Boundary analysis in sedimentation velocity experiments. *Methods Enzymol.* 240, 478–501.
- Uchôa, H.B., Jorge, G.E., Silveira, N.J.F., Câmara, J.C., Canduri, F., Azevedo Jr., W.F., 2004. Parmodel: a web server for automated comparative modeling of proteins. *Biochem. Biophys. Res. Commun.* 325 (4), 1481–1486.
- Viola, C.M., Saridakis, V., Christendat, D., 2004. Crystal structure of chorismate synthase from *aquifex aeolicus* reveals a novel beta alpha beta sandwich topology. *Proteins: Struct. Funct. Genet.* 54, 166–169.
- Welch, G.R., Cole, K.W., Gaertner, F.H., 1974. Chorismate synthase of *Neurospora crassa*: a flavoprotein. *Arch. Biochem. Biophys.* 165, 505–518.
- WHO, 2004. World Health Organization. WHO Report, Geneva, Switzerland. WHO/CDS/TB/2004.331.

### 3.3 Molecular models of tryptophan synthase from *Mycobacterium tuberculosis* complexed with inhibitors

Marcio V. B. Dias; Fernanda Canduri; Nelson J. F. da Silveira; Clarissa M. Czekster; Luiz A. Basso; Mário S. Palma; Diógenes S. Santos; Walter F. de Azevedo Jr. **Cell Biochemistry and Biophysics** (ISSN 1085-9195), v. 44, p. 375-384, 2006.

Neste trabalho é apresentada a modelagem molecular da proteína triptofano sintase de *M. tuberculosis* em complexo com 6 inibidores análogos do estado de transição desta enzima.

Para a modelagem molecular da triptofano sintase de *M. tuberculosis*, foi utilizado como ponto de partida (*templates*), as estruturas cristalográficas da triptofano sintase de *S. typhimurium* em complexo com os respectivos inibidores. Para realizar a modelagem molecular foi utilizado o programa MODELLER (SALI; BLUNDELL, 1993). As duas subunidades da triptofano sintase foram modeladas separadamente e foram gerados 1000 modelos para cada subunidade e para cada complexo. Os modelos finais foram selecionados baseando-se em análise estereoquímica. A análise da afinidade de ligação entre o complexo proteína:ligante foi realizada baseando-se na contribuição individual de afinidade de ligação de cada átomo do ligante, utilizando o programa SCORE (WANG et al., 1998). Além disso foi levado em consideração o número de ligações de hidrogênio e a área de contato entre o ligante e a proteína. Os modelos gerados apresentam boa qualidade estereoquímica e são bastante similares a seus *templates*. A comparação dos sítios ativos dos complexos mostra que muitas interações de hidrogênio são conservadas entre os modelos obtidos e os *templates*, assim como muitas interações entre os próprios inibidores, apesar de estes apresentarem algumas interações únicas. Análise da área de contato entre a

proteína e seus inibidores mostra que a triptofano sintase de *M. tuberculosis* apresenta uma maior área de contato do que a de *S. typhimurium*, entretanto, análise das interações favoráveis para a ligação, mostra que a triptofano sintase *S. typhimurium* apresenta uma afinidade de ligação (*score*) maior do que a triptofano sintase de *M. tuberculosis*. Desta forma, algumas alterações nos átomos dos inibidores estudados podem aumentar a afinidade destes pela triptofano sintase de *M. tuberculosis*.

ORIGINAL ARTICLE

## Molecular Models of Tryptophan Synthase From *Mycobacterium tuberculosis* Complexed With Inhibitors

Marcio Vinicius Bertacine Dias,<sup>1</sup> Fernanda Canduri,<sup>2</sup>  
Nelson José Freitas da Silveira,<sup>1</sup> Clarissa Melo Czekster,<sup>3</sup>  
Luis Augusto Basso,<sup>4</sup> Mário Sérgio Palma,<sup>5</sup>  
Diógenes Santiago Santos,<sup>6,\*</sup> and Walter Filgueira de Azevedo, Jr.<sup>4,\*</sup>

<sup>1</sup>Programa de Pós-Graduação em Biofísica Molecular-Departamento de Física, UNESP, São José do Rio Preto, Brazil; <sup>2</sup>Departamento de Morfofisiologia-CCBS-UFMS, Campo Grande-MS, 79070-900, Brazil; <sup>3</sup>Rede Brasileira de Pesquisas em Tuberculose, Grupo de Microbiologia Molecular e Funcional, Centro de Biotecnologia, UFRGS, Porto Alegre-RS, 91501-970, Brazil; <sup>4</sup>Faculdade de Biociências-PUCRS, Porto Alegre-RS, 90619-900, Brazil; <sup>5</sup>Laboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil; and <sup>6</sup>Centro de Pesquisas em Biologia Molecular e Funcional/PUCRS, Avenida Ipiranga, 6681, Tecnopuc, Partenon 90619-900, Porto Alegre, RS, Brazil

### Abstract

The development of new therapies against infectious diseases is vital in developing countries. Among infectious diseases, tuberculosis is considered the leading cause of death. A target for development of new drugs is the tryptophan pathway. The last enzyme of this pathway, tryptophan synthase (TRPS), is responsible for conversion of the indole 3-glycerol phosphate into indol and the condensation of this molecule with serine-producing tryptophan. The present work describes the molecular models of TRPS from *Mycobacterium tuberculosis* (*Mt*TRPS) complexed with six inhibitors, the indole 3-propanol phosphate and five arylthioalkyl-phosphonated analogs of substrate of the  $\alpha$ -subunit. The molecular models of *Mt*TRPS present good stereochemistry, and the binding of the inhibitors is favorable. Thus, the generated models can be used in the design of more specific drugs against tuberculosis and other infectious diseases.

**Index Entries:** Tryptophan synthase; *Mycobacterium tuberculosis*; molecular modeling; drug design; structural bioinformatics.

### INTRODUCTION

Tryptophan synthase (TRPS) is a target for development of vaccines and drugs. The interaction between *Mycobacterium tuberculosis* TRPS (*Mt*TRPS) and inhibitors derived from the arylthioalkyl-phosphonated inhibitors (e.g., indole 3-propanol phosphate [IPP]) is not known. In this work, we show that IPP and their analogs (or derivatives) present strong interaction with *Mt*TRPS and that they can inhibit this protein efficiently.

\* Author to whom all correspondence and reprint requests should be addressed. E-mail: walter.junior@pucrs.br or diogenes@pucrs.br

The development of new therapies as vaccines and drugs is vital in developing country, where a majority of the population die because of infectious diseases, such as tuberculosis, AIDS, cholera, and malaria, among others (1). Among these diseases, tuberculosis deserves special attention because it is the main cause of death in these countries. *M. tuberculosis*, the main causative agent of tuberculosis, infects approx 8 million individuals and causes 2 million deaths annually (2).

Bacteria, fungi, apicomplexan parasites, and plants synthesize aromatic compounds, and this synthesis is essential for their survival (3,4). Therefore, the enzymes responsible for synthesis of aromatic compounds are considered targets

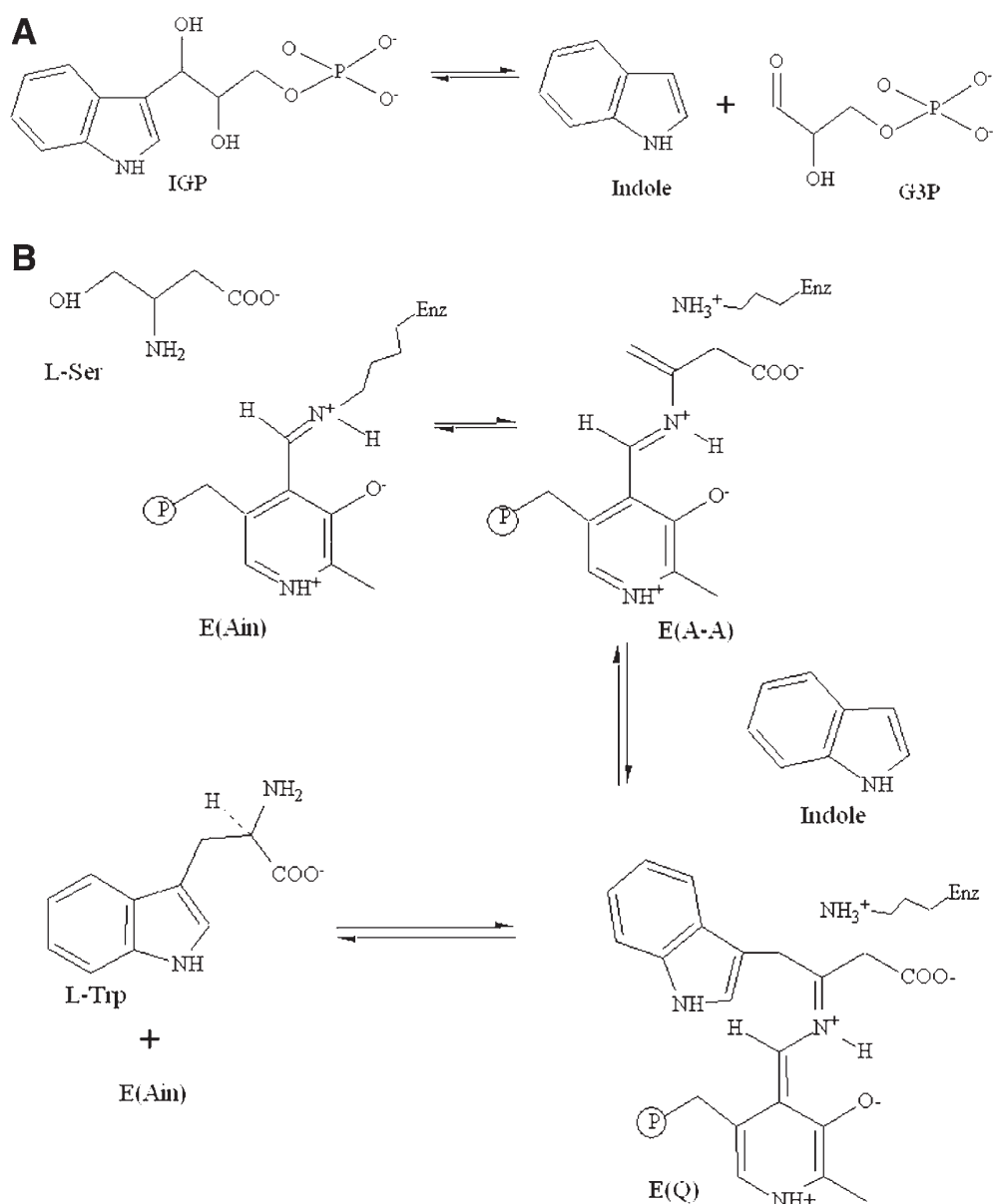


Fig. 1. Reaction scheme of the TRPS. (A) Reaction of the  $\alpha$ -subunit. (B) Reaction of the  $\beta$ -subunit.

for development of new therapies, because these compounds are absent in animals. Thus, the inhibition of amino acid biosynthesis is a demonstrated mode of action for important classes of inhibitors. Glyphosate and 6-fluoroshikimate, for example, inhibit biosynthesis of aromatic compounds in bacteria, fungi, apicomplexan parasites, and plants (5,6). Enzymes involved specifically in tryptophan biosynthesis have been proposed as potential herbicide and vaccine targets (7,8). The last enzyme of the pathway of tryptophan biosynthesis is TRPS, and the following reasons led to selection of this enzyme as a potential target. For one, all organisms that synthesize tryptophan are known to

do so by a single route, such that inhibition of this pathway should dramatically reduce tryptophan levels. TRPS is also one of the most studied enzymes of those involved in the biosynthesis of amino acids (9). It catalyzes two distinct reactions by separate polypeptide chains, referred to as  $\alpha$ - and  $\beta$ -subunits. The  $\alpha$ -subunit catalyzes the cleavage of indole 3-glycerol phosphate to yield an indole and glyceraldehyde-3-phosphate, whereas the  $\beta$ -subunit catalyzes the condensation of the indole with serine to produce tryptophan (Fig. 1) (10). The three-dimensional (3D) structure of the TRPS from *Salmonella typhimurium* (*St*TRPS) with several inhibitors has been deposited in the



Protein Data Bank (PDB). Among them are IPP (11) and five arylthioalkyl-phosphonated transition state analogs that inhibit the  $\alpha$ -subunit (12). IPP is an analog of IGP that inhibits the  $\alpha$ -subunit with a  $K_i$  value of  $15 \mu\text{M}$ , and it was the starting point for inhibitor design strategy of the arylthioalkyl-phosphonate (7). The arylthioalkyl-phosphonated inhibitors were designed to mimic the transition state formed during the  $\alpha$ -reaction of the enzyme, and as expected, they have affinities much greater than that of the natural substrate IGP or its nonhydrolyzable analog IPP. These inhibitors are ortho-substituted arylthioalkyl-phosphonate derivatives that have an  $\text{sp}^3$ -hybridized sulfur atom, designed to mimic the putative tetrahedral transition state at the C3 atom of the indole, and they lack the C2 atom to allow for higher conformational flexibility (7,12). These phosphonated inhibitors have been synthesized because of efforts to develop herbicidal and antimicrobial agents, and they have been tested for inhibition of enzyme activity in in vitro assay and for herbicidal activity in a biological assay (7). These phosphonated compounds possess enzyme inhibitory and herbicidal activities with micromolar  $\text{IC}_{50}$  values (7). The present work describes six molecular models of *Mt*TRPS, the complex with IPP, and the complexes with five arylthioalkyl-phosphonated analogs bound in  $\alpha$ -subunit and pyridoxal 5'-phosphate (PLP) in the  $\beta$ -subunit.

## MATERIALS AND METHODS

### Molecular Modeling

For modeling of the *Mt*TRPS, we used Parmodel, a Web server for automated comparative modeling of proteins (13) based on the program MODELLER (14). The atomic coordinates of the crystallographic structures of *Sf*TRPS were used as starting models. For the complex with IPP, we used PDB access code 1QOP (11) and for five complexes with arylthioalkyl-phosphonated derivatives we used the following PDB access codes: 1C29, 1C8V, 1CX9, 1C9D, and 1CW2 (12). The arylthioalkyl-phosphonated inhibitors are 4-(2-hydroxyphenylthio)-1-butenylphosphonic acid, 4-(2-hydroxyphenylthio)butylphosphonic acid, 4-(2-hydroxy-5-fluorophenylthio)butylphosphonic acid and 4-(2-hydroxyphenylsulfanyl)butylphosphonic acid. The complexes are designated by numbers, with 1 for TRPS:IPP and 2 to 6 for the five arylthioalkyl-phosphonates, respectively. The chemical structures of inhibitors are shown in Fig. 2. All the cited compounds bind to the  $\alpha$ -subunit. Moreover, the  $\beta$ -subunit was modeled in complex with PLP for all of the models. However, the models were generated separately for the  $\alpha$ - and  $\beta$ -subunit. The atomic coordinates of all water were removed from the TRPS templates. Several slightly different models can be calculated by varying the initial structure. In total, 1000 models were generated for each subunit of tryptophan synthase and

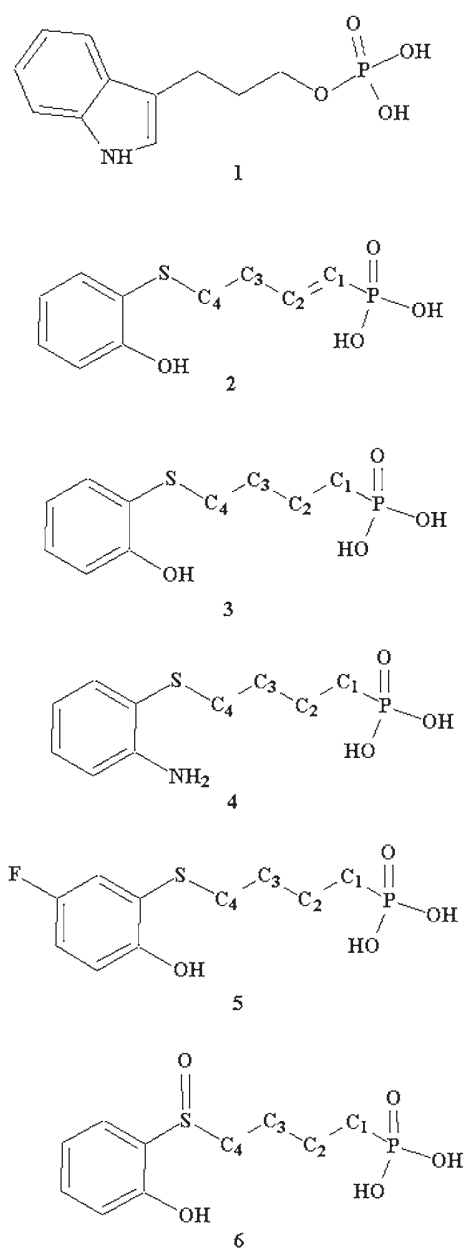


Fig. 2. Molecular formulae of TRPS inhibitors. (1) indolopropanol-phosphate; (2) 4-(2-hydroxyphenylthio)-1-butenylphosphonic acid; (3) 4-(2-hydroxyphenylthio)butylphosphonic acid; (4) 4-(2-aminophenylthio)butylphosphonic acid; (5) 4-(2-hydroxy-5-fluorophenylthio)butylphosphonic acid; and (6) 4-(2-hydroxyphenylsulfanyl)butylphosphonic acid.

each conformational state, and the final models were selected based on stereochemical quality. Subsequently, we performed the junction of the two subunits, and the complexes were optimized using MODELLER (14). The optimization of the complexes was carried out by the use of the variable target function method using methods of con-

jugate gradients and molecular dynamics with simulated annealing. All modeling processes were performed on a Beowulf cluster, with 16 nodes (B16/AMD Athlon 1800+; BioComp, São José do Rio Preto, SP, Brazil).

### Analysis of the Model

The overall stereochemical quality of the final models for the complexes of *Mt*TRPS were assessed by the program PROCHECK (15). Molecular models were superposed using the program LSQKAB from CCP4 (15). The cutoff for hydrogen bonds and salt bridges was 3.5 Å. The contact surfaces for the binary complexes were calculated using AREAIMOL and RESAREA (15). The root mean squared deviation (rmsd) differences from ideal geometries for bond lengths and bond angles were calculated with X-PLOR (16). The G-factor is essentially just log-odds score based on the observed distributions of the stereochemical parameters. It was computed for the following properties: torsion angles (the analyses provided the observed distributions of  $\phi$ - $\psi$ ,  $\chi^1$ - $\chi^2$ ,  $\chi$ -1,  $\chi$ -3,  $\chi$ -4, and  $\omega$  values for each of the 20 amino acid types) and covalent geometry (for the main-chain bond lengths and bond angles). These average values were calculated using PROCHECK (15). The 3D profile measures the compatibility of a protein model with its sequence, and it was calculated using Verify 3D program (17,18).

### Analysis of Binding Affinity of Protein-Ligand Complexes

For an estimative the absolute binding affinity of the protein-ligand complexes, we used the program SCORE (19). According to this method, the binding affinity of the ligand can be decomposed to the contribution of individual atoms. Each ligand atom obtains a score, called the atomic binding score, indicating its role in the binding process. The program reads the structure, assigns atom types and parameters, performs the calculation, and gives the dissociation constant of the given protein-ligand complex. The computational results are outputted into a text file in which the detailed information of each ligand atom, including the atomic binding score, is tabulated.

## RESULTS AND DISCUSSION

The sequence alignment of *St*TRPS (template) with *Mt*TRPS (target) is shown in Fig. 3. The sequence of  $\alpha$ -subunit of *Mt*TRPS shows 27% of identity with the sequence of  $\alpha$ -subunit of *St*TRPS, whereas the sequence of  $\beta$ -subunit of *Mt*TRPS shows 53% identity with the sequence of  $\beta$ -subunit of *St*TRPS.

### Quality of the Model

Ramachandran plots for the six *St*TRPS structures solved by crystallography were generated to compare the overall stereochemical quality of six *Mt*TRPS mod-

els. Analysis of the Ramachandran plots indicated that the models present a minimum of 93.4% of the residues in the most favored regions and a maximum of 0.2% of the residues in the disallowed regions, whereas structures solved by crystallography present a minimum of 91.6% of the residues in the most favored regions and also a maximum of 0.2% of the residues in the disallowed regions (Table 1). The Verify 3D values, the average G-factor, and rmsd values of bond lengths and bond angles are shown in Table 2. From analysis of the overall stereochemical quality of the molecular models, we feel that it is appropriate for structural studies.

### Overall Description

The structures of six molecular models were superposed with their templates. Only C $\alpha$  was considered in the superposition. The values of RMSD for the superpositions can be seen in Table 2. These results show that *Mt*TRPS models are most similar to the *St*TRPS structure, as expected. The models of *Mt*TRPS contain 270 residues in the  $\alpha$ -subunit (residues 1–8 were omitted) and 422 in the  $\beta$ -subunit (residues 1–23 were omitted). The model of *Mt*TRPS complexed with IPP is presented in Fig. 4.

The structure of the  $\alpha$ -subunit follows the 8-fold  $\alpha/\beta$  barrel motif first observed in triosephosphate isomerase (TIM) (20,21). The  $\alpha/\beta$  barrel structure can be described in terms of a canonical form in which the molecule is built up from eight repeating supersecondary structural units each comprised of a  $\beta$ -strand followed by an  $\alpha$ -helix. The eight helices, connected to both ends of the strands by loops, pack closely in parallel around the periphery of the barrel.

The  $\beta$ -subunit of the crystal structure of *St*TRPS and the molecular models of *Mt*TRPS reveal the presence of two domains of nearly equal size, denominated N-domain and C-domain (21). The core of the N-domain is formed by four strands, with three helices packed on one side of the sheet, and a fourth helix (helix 6) packed on the other side. The part of the N-domain constitutes a domain that has very few interactions with the rest of the protein and is named in the COMM-domain (21). The C-domain is comprised of a six-strand  $\beta$ -sheet surrounded by  $\alpha$ -helices. The core of the C-domain contains a central  $\beta$ -sheet sandwiched between layers of parallel packed helices, much like the core of the N-terminal domain. The  $\beta$ -sheet contains five parallel strands and one antiparallel.

The interface between  $\alpha$ - and  $\beta$ -subunits is mostly hydrophobic in character, and many residues of the interface between two subunits are conserved in almost all sequences. Approximately 1295 Å<sup>2</sup> of the surface area of the  $\alpha$ -subunit and 1285 Å<sup>2</sup> of the  $\beta$ -subunit are buried at the  $\alpha/\beta$  interface.

### Interaction Between Protein and Its Inhibitors

All inhibitors studied in this work bind to the active site of the  $\alpha$ -subunit. Potential hydrogen bonding inter-



Table 1  
Analysis of Ramachandram Plots

Model	Most favored (%)	Additional allowed (%)	Generously allowed (%)	Disallowed (%)
1	96.8 (94.6)	5.5 (5.1)	0.5 (0.2)	0.2 (0.2)
2	93.2 (94.7)	6.2 (5.1)	0.5 (0.2)	0.0 (0.0)
3	93.8 (92.7)	5.7 (7.1)	0.4 (0.2)	0.2 (0.0)
4	94.0 (91.6)	5.5 (8.2)	0.4 (0.0)	0.2 (0.2)
5	93.4 (91.6)	6.0 (8.0)	0.5 (0.4)	0.0 (0.0)
6	93.4 (93.6)	6.0 (6.2)	0.5 (0.2)	0.0 (0.0)

Values in parentheses are for templates.

plexes 2, 3, 5, and 6, or the amino group for complex 4 bonded to the phenyl ring of the inhibitors, or with the indole nitrogen atom of complex 1. Asp68 is thought to play a crucial role in catalysis of the  $\alpha$ -reaction and the intersubunit communication interface (22). Compounds 1, 4, 5, and 6 present a third interaction with main chain nitrogens of Ser240. In compound 1, the introduction of the double bond does not disturb the pattern of the hydrogen bond between the enzyme and inhibitor, although rigid conformation is imposed. The atom of the fluorine introduced in the ring of compound 5 does not interact by polar interaction with any residue of the protein. On the other hand, the sulfoxide oxygen of compound 6 forms one hydrogen bond interaction with the hydroxyl group of Tyr181. By comparison of the binding site of the six ligands in the models of *Mt*TRPS, we may conclude that it superimposes significantly well (figure not shown). The major differences occur in inhibitor 4 in overall positioning of its alkyl chain that is associated with an altered orientation of the phosphate group position relative to the other inhibitors. Owing to the orientation of the phosphate group in compound 4, it presents a weaker interaction between the inhibitor and the backbone amine of Gly190 compared with others inhibitors (3.4 Å).

The presence of different ligands in the  $\alpha$ -subunit does not cause any significant change in the binding site of PLP in the  $\beta$ -subunit, despite that TRPS is an allosteric protein and binding of the compounds of the  $\alpha$ -subunit is able to cause structural change in the  $\beta$ -subunit (11).

#### **Analysis of the Interaction Between the Molecular Models and Inhibitors**

Table 3 shows the analysis of the interaction between the molecular models of *Mt*TRPS and the molecular structure of *St*TRPS with the six compounds studied in the present work. Roughly, all ligands have high affinity

for *Mt*TRPS and *St*TRPS. They present from six to eight intermolecular hydrogen bonds, contact area above 179 Å<sup>2</sup>, and positive score value. Therefore, it can be said that the high affinity between the TRPS and IPP or arylthioalkyl-phosphonated inhibitors is related to directional hydrogen bonds and ionic interactions, as well as shape complementarity of the contact surface of the partners (23–25).

From analysis (Table 3), compound 5 presents greater contact area with *Mt*TRPS and seven hydrogen bonds, but does not present greater amount of favorable binding (score value). This suggests that compound 5 can be buried in the protein, but the binding is not completely favorable. Compound 6 presents eight hydrogen bonds, but the score value is the lowest. Therefore, this compound must present binding that is not favorable for interaction with protein. In contrast, compound 2 presents greater score value and thus a larger amount of favorable binding with protein and a large contact area (199 Å). Therefore, this compound presents higher affinity with the protein compared with other compounds. In biological activity assay, compound 2 presented greater herbicide activity against *Arabidopsis thaliana* (7). Accordingly, based on modeling results obtained and experimental results, compound 2 can also inhibit *Mt*TRPS. Inhibition studies of *Mt*TRPS against compound 2 and other inhibitors studied here may confirm these predictions.

#### **CONCLUSIONS**

The molecular models of *Mt*TRPS generated by MODELLER program possess good stereochemical quality. Furthermore, they show that the interaction with phosphonates and IPP is favorable as observed for *St*TRPS. Therefore, the molecular models can be used in the development of more specific drugs against tuberculosis and other infectious diseases, because the

Table 2  
Analysis of Quality of Models of *Mf*TRPS and Templates

Complex	3D profile <sup>a</sup>		G-factor <sup>b</sup>			rmsd from ideal geometry			Superposition
	Total score	Ideal score	S <sub>ideal</sub> score	Torsion angles	Covalent geometry	Global	Bond lengths (Å)	Bond angles (°)	
1	301.80 (281.35)	303.60 (298.97)	0.99 (0.94)	-0.24 (0.12)	-0.25 (-0.21)	-0.24 (0.17)	0.021 (0.008)	2.519 (0.02)	0.128
2	298.95 (338.03)	303.60 (298.97)	0.98 (1.13)	-0.27 (0.03)	-0.30 (0.43)	-0.28 (0.20)	0.021 (0.010)	2.515 (1.70)	0.232
3	302.55 (330.26)	303.60 (298.97)	1.00 (1.10)	-0.27 (0.05)	-0.28 (0.40)	-0.28 (0.20)	0.021 (0.010)	2.686 (1.80)	0.256
4	293.33 (324.04)	303.60 (298.97)	0.97 (1.08)	-0.28 (0.02)	-0.29 (0.45)	-0.28 (0.20)	0.021 (0.009)	2.686 (1.71)	0.248
5	298.52 (323.71)	303.60 (298.97)	0.98 (1.08)	-0.30 (0.05)	-0.33 (0.41)	-0.30 (0.20)	0.022 (0.010)	2.745 (1.83)	0.297
6	297.38 (327.72)	303.60 (298.97)	0.98 (1.10)	-0.14 (0.01)	-0.44 (0.42)	-0.25 (0.18)	0.022 (0.010)	2.785 (1.77)	0.249

Values obtained for the templates used for modeling are presented in parentheses. <sup>a</sup> Total score is the sum of the 3D-one-dimensional scores (statistical preferences) of each residue present in protein. Ideal score  $S_{ideal} = \exp(-0.83 + 1.008 \times \ln(L))$ , where  $L$  is the number of amino acids.  $S_{ideal}$  score is compatibility of the sequence with the 3D structure. It is obtained from total score/ideal score.  $S_{ideal}$  score above 0.455<sub>ideal</sub>.

<sup>b</sup> Ideally, scores should be above -0.5. Values below -1.0 may need investigation.

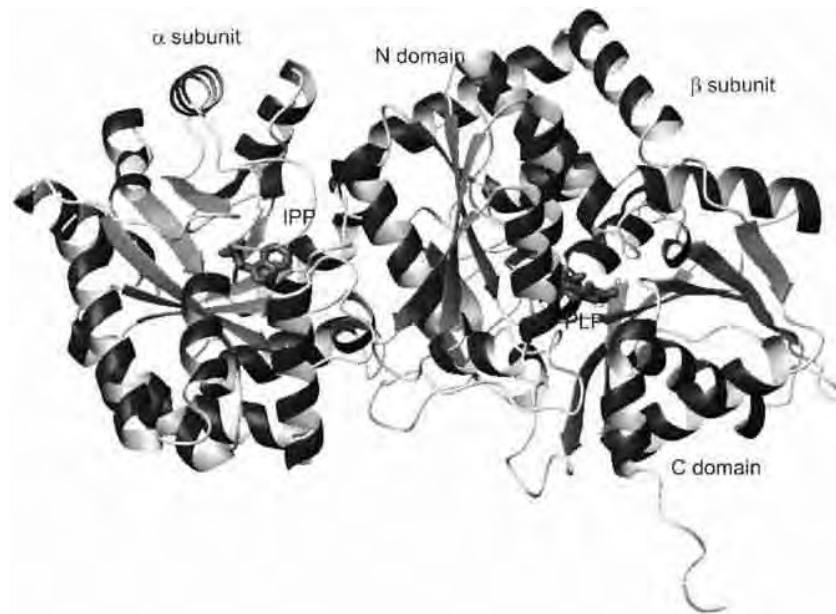


Fig. 4. Ribbon diagram of the  $\alpha$ - and  $\beta$ -subunits of *Mf*TRPS complexed with IPP and PLP generated by Molmol (27).

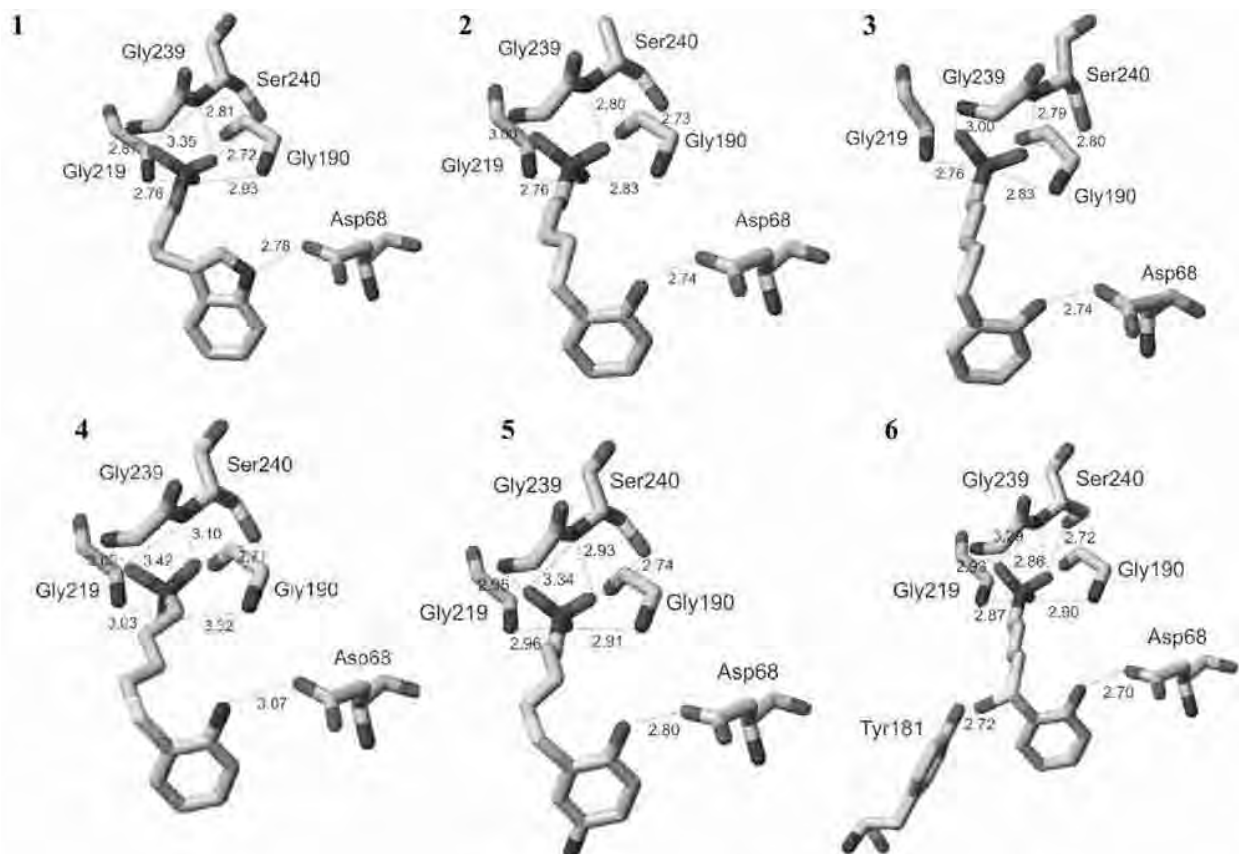


Fig. 5. Hydrogen bond pattern between *Mf*TRPS and inhibitors. (1) indol-propanol-phosphate; (2) 4-(2-hydroxyphenylthio)-1-butenylphosphonic acid; (3) 4-(2-hydroxyphenylthio)butylphosphonic acid; (4) 4-(2-aminophenylthio)butylphosphonic acid; (5) 4-(2-hydroxy-5-fluorophenylthio)butylphosphonic acid; and (6) 4-(2-hydroxyphenylsilynyl)butylphosphonic acid.

Table 3  
Interaction of Between Protein and Inhibitors<sup>a</sup>

Model	Number hydrogen bonds	Contact area (Å <sup>2</sup> )	Score
1	7 (7)	200.0 (167.0)	7.39 (7.76)
2	6 (6)	199.0 (176.0)	7.98 (8.29)
3	6 (6)	197.0 (191.0)	7.70 (7.94)
4	7 (7)	192.0 (172.0)	6.61 (6.95)
5	7 (7)	211.0 (185.0)	6.51 (7.60)
6	8 (7)	179.0 (175.0)	5.52 (7.04)

<sup>a</sup> Values in parenthesis are for templates.

examination of the binding site for arylthiophosphate inhibitor presents interactions that could be exploited to improve affinity with the protein. The replacement of hydrogen atoms along the alkyl chain with larger atoms or functional group can increase the affinity or the inhibition of the ligands.

## ACKNOWLEDGMENTS

This work was supported by grants from FAPESP (SMOLBNet 01/07532-0, 02/04383-7, 03/12472-2, 04/00217-0), CNPq, CAPES, and Instituto do Milênio (CNPq-MCT) PRONEX, FAPERGS. W.F.A., M.S.P., L.A.B., and D.S.S. are researchers for the Brazilian Council for Scientific and Technological Development from Brazil (CNPq).

## REFERENCES

- Trouiller, P., Torreele, E., Olliaro, P., et al. (2001) Drugs for neglected diseases: a failure of the market and a public health failure? *Trop. Med. Int. Health* **6**, 945–951.
- World Health Organization. Global Tuberculosis Control. WHO Report 2001. Geneva, Switzerland, WHO/CDS/TB/2001.287
- Dosselaere, F. and Vanderleyden, J. (2001) A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Crit. Rev. Microbiol.* **27**, 75–131.
- Roberts, F., Roberts, C. W., Johnson, J. J., et al. (1998) Evidence for the shikimate pathway in apicomplexan parasites. *Nature* **393**, 801–805.
- Schonbrunn, E., Eschenburg, S., Shuttleworth, W. A., et al. (2001) Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1376–1380.
- Daves, G. M, Barrett-Bee, K. J. Jude, D. A., et al. (1994) (6S)-6-Fluoroshikimic acid, an antibacterial agent acting on the aromatic biosynthetic pathway. *Antimicrob. Agents Chemother.* **38**, 403–406.
- Finn, J., Langevine, C., Birk, I., Nicherson, K., and Rodaway, S. (1999) Rational herbicide design by inhibition of tryptophan biosynthesis. *Bioorg. Med. Chem. Lett.* **9**, 2297–2302.
- Smith, D. A., Parish, T., Stoker, N. G., and Bancroft, G. J. (2001) Characterization of auxotrophic mutants of *Mycobacterium tuberculosis* and their potential as vaccine candidates. *Infect. Immun.* **69**, 1142–1150.
- Hyde, C. C. and Miles, E. W. (1990) The tryptophan synthase multienzyme complex: exploring structure-function relationships with X-ray crystallography and mutagenesis. *Biotechnology* **8**, 27–32.
- Pan, P., Woehl, E., and Dunn, M. F. (1997) Protein architecture, dynamics and allostery in tryptophan synthase channeling. *Trends Biochem. Sci.* **22**, 22–27.
- Weyand, M. and Schlichting, I. (1999) Crystal structure of wild-type tryptophan synthase complexed with the natural substrate indole-3-glycerol phosphate. *Biochemistry* **38**, 16,469–16,480.
- Sachpatzidis, A., Dealwis, C., Lubetsky, J. B., Liang, P. H., Anderson, K. S., and Lolis, E. (1999) Crystallographic studies of phosphonate-based alpha-reaction transition-state analogues complexed to tryptophan synthase. *Biochemistry* **38**, 12,665–12,674.
- Uchoa, H. B., Jorge, G. E., da Silveira, N. J. F., Camera, J. C., Jr., Canduri, F., and de Azevedo, W. F., Jr. (2004) Parmodel: a web server for automated comparative modeling of proteins. *Biochem. Biophys. Res. Commun.* **325**, 1481–1486.
- Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.
- Collaborative Computational Project No. 4 (1994) The CCP4 suite: program for protein crystallography. *Acta Crystallogr. D* **50**, 760–763.
- Brünger, A. T. (1992) X-PLOR version 3.1: a system for crystallography and NMR. Yale University Press, New Haven, CT.
- Bowie, J. U., Luthy, R., and Eisenberg, D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253**, 164–170.
- Luthy, R., Bowie, J., and Eisenberg D. (1992) Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85.
- Wang, R., Liu, L., Lai, L., and Tang, Y. (1998) SCORE: a new empirical method for estimating the binding affinity of a protein-ligand complex. *J. Mol. Model.* **4**, 379–394.
- Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) Three-dimensional structure of the tryptophan synthase  $\alpha_2\beta_2$  multienzyme complex from *Salmonella typhimurium*. *J. Biol. Chem.* **263**, 17,857–17,871.
- Banner, D. W., Bloomer, A. C., Petsko, G. A., et al. (1975) Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. *Nature* **255**, 609–614.
- Rhee, S., Miles, E. W., Mozzarelli, A., and Davies, D. R. (1998) Cryocrystallography and microspectrophotometry of a mutant (alpha D60N) tryptophan synthase alpha 2 beta 2 complex reveals allosteric roles of alpha Asp60. *Biochemistry* **37**, 10,653–10,659.

23. De Azevedo, W. F., Jr., Mueller-Dieckmann, H. J., Schulze-Gahmen, U., Worland, P. J., Sausville, E., and Kim, S. H. (1996) Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2735–2740.
24. De Azevedo, W. F., Jr., Canduri, F., Dos Santos, D. M., et al. (2003) Structural basis for inhibition of human PNP by immucillin-H. *Biochem. Biophys. Res. Commun.* **309**, 922–927.
25. Pereira, J. H., Canduri, F., de Oliveira, J. S., et al. (2003) Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* **312**, 608–614.
26. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
27. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOL-MOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.



### 3.4 Effects of the magnesium and chloride ions and shikimate on the structure of shikimate kinase from *Mycobacterium tuberculosis*.

Marcio V. B. Dias; Livia M. Faím; Igor B. Vasconcelos; Jaim S. de Oliveira; Luiz A. Basso; Diógenes S. Santos; Walter F. de Azevedo Jr. **Acta Crystallographica Section F Structural Biology and Crystallization Communications** (ISSN 1744-3091), 63, p. 1-6, 2007.

Neste trabalho é apresentado o efeito dos íons magnésio e cloro e do ácido chiquímico sobre a estrutura da chiquimato quinase de *M. tuberculosis* (peso molecular de 18583 Da).

A Chiquimato quinase de *M. tuberculosis* foi clonada, superexpressada e purificada pelo grupo de Pesquisas do Prof. Dr. Diógenes S. Santos da PUC – Porto Alegre – RS. Para a obtenção de cristais da chiquimato quinase de *M. tuberculosis* em complexo com ácido chiquímico e ADP, na ausência de  $Mg^{2+}$ , a proteína foi dialisada contra tampão Tris-HCl, 50 mM, pH 8,0 e concentrada a  $14 \text{ mg.mL}^{-1}$  e utilizada inicialmente a condição de cristalização estabelecida por Dhaliwal et al. (2004). Os melhores cristais foram obtidos em uma condição composta por Tris-HCl, 0,1 M, pH 8, 17% de PEG 1500 e de cloreto de lítio, 0,5-0,7 M. Por outro lado, a condição de cristalização na qual foram obtidos os melhores cristais para o complexo da chiquimato quinase de *M. tuberculosis* em complexo com ADP e magnésio era composta por Tris-HCl, 0,1 M, 20% PEG 3350 e 0,1 M de  $MgCl_2 \cdot 6H_2O$  e a concentração da proteína era de  $17 \text{ mg.mL}^{-1}$ . A coleta de dados de difração de raios X foi realizada no LNLS. O cristal obtido para a chiquimato quinase de *M. tuberculosis* em complexo com ADP e ácido chiquímico é similar aos descritos na literatura e é pertencente ao grupo espacial  $P3_221$ . Este cristal difratou a  $1,93 \text{ \AA}$  de resolução e apresenta um monômero na unidade assimétrica. Os cristais da chiquimato

quinase de *M. tuberculosis* em complexo com  $Mg^{2+}$  e ADP são pertencentes ao grupo espacial  $P2_12_12_1$  e difrataram a 2,8 Å de resolução, e apresentam na unidade assimétrica quatro monômeros, que formam um tetrâmero com simetria 222. As estruturas para os conjuntos de dados dos cristais obtidos foram determinadas por substituição molecular utilizando como modelo de busca a estrutura da chiquimato quinase de *M. tuberculosis* em complexo com MgADP e ácido chiquímico (PEREIRA et al., 2004). O refinamento cristalográfico das estruturas foi realizado pelo programa REFMAC 5 (MURSHUDOV; VAGIN; DODSON, 1997) e a inspeção visual e adição de moléculas de água foi efetuada pelo programa XtalView/Xfit (McREE et al., 1999). A estrutura da chiquimato quinase de *M. tuberculosis* em complexo com ADP e ácido chiquímico, na ausência do íon magnésio, foi comparada com estrutura da chiquimato quinase em complexo com MgADP e ácido chiquímico determinada por Pereira et al., 2004, que é a única estrutura que foi observado o íon magnésio quando o ácido chiquímico está presente. Pela comparação destas duas estruturas pôde-se observar o efeito que o íon magnésio causa sobre a estrutura da chiquimato quinase e sobre o ácido chiquímico. Nesta estrutura foram observadas alterações na posição da cadeia lateral de importantes resíduos do sítio ativo da enzima e alterações nos grupos hidroxilas da molécula de ácido chiquímico. Com relação à estrutura da chiquimato quinase de *M. tuberculosis* em complexo com magnésio e ADP pôde-se observar as possíveis interferências do ácido chiquímico sobre a conformação da chiquimato quinase, uma vez que esta molécula não foi adicionada na condição de cristalização. Além disso, em um dos monômeros do tetrâmero, a molécula de cloro, que representa um importante papel na catálise da chiquimato quinase, não foi observada (CERASOLI et al., 2003). Assim, este monômero apresentou alterações estruturais não observadas nos outros monômeros. Estas alterações conformacionais foram descritas e acreditamos que elas podem ser causadas pela ausência do íon cloreto na estrutura da chiquimato quinase de *M. tuberculosis*.

Marcio Vinicius Bertacine Dias,<sup>a</sup>  
Livia Maria Faím,<sup>a</sup> Igor Bordin  
Vasconcelos,<sup>b</sup> Jaim Simões de  
Oliveira,<sup>b</sup> Luiz Augusto Basso,<sup>b</sup>  
Diógenes Santiago Santos<sup>b\*</sup> and  
Walter Filgueira de Azevedo Jr<sup>c\*</sup>

<sup>a</sup>Programa de Pós-Graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, <sup>b</sup>Pontifícia Universidade Católica do Rio Grande do Sul, Centro de Pesquisa em Biologia Molecular e Funcional, Porto Alegre, RS, Brazil, and <sup>c</sup>Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681 Porto Alegre-RS, CEP 90619-900, Brazil

Correspondence e-mail: diogenes@pucls.br, walter.junior@pucls.br

Received 27 September 2006  
Accepted 6 November 2006

**PDB References:** shikimate kinase–ADP–shikimate complex, 2dfn, r2dfnsf; shikimate kinase–MgADP complex, 2dft, r2dftsf.



© 2007 International Union of Crystallography  
All rights reserved

## Effects of the magnesium and chloride ions and shikimate on the structure of shikimate kinase from *Mycobacterium tuberculosis*

Bacteria, fungi and plants can convert carbohydrate and phosphoenolpyruvate into chorismate, which is the precursor of various aromatic compounds. The seven enzymes of the shikimate pathway are responsible for this conversion. Shikimate kinase (SK) is the fifth enzyme in this pathway and converts shikimate to shikimate-3-phosphate. In this work, the conformational changes that occur on binding of shikimate, magnesium and chloride ions to SK from *Mycobacterium tuberculosis* (MtSK) are described. It was observed that both ions and shikimate influence the conformation of residues of the active site of MtSK. Magnesium influences the conformation of the shikimate hydroxyl groups and the position of the side chains of some of the residues of the active site. Chloride seems to influence the affinity of ADP and its position in the active site and the opening length of the LID domain. Shikimate binding causes a closing of the LID domain and also seems to influence the crystallographic packing of SK. The results shown here could be useful for understanding the catalytic mechanism of SK and the role of ions in the activity of this protein.

### 1. Introduction

*Mycobacterium tuberculosis*, the aetiological agent of tuberculosis (TB), infects one-third of the world's population. It is estimated that 1.7 million deaths resulted from TB in 2004, 95% of which occurred in developing countries (World Health Organization, 2006). The emergence of TB as a public health threat, the high susceptibility of HIV/TB co-infected patients and the proliferation of multi-drug-resistant strains have created a need for newer and better drugs for the treatment of TB.

The shikimate pathway is an attractive target for the development of herbicides (Coggins, 1998) and antibiotic agents (Davies *et al.*, 1994) because it is essential in algae, higher plants, bacteria, fungi and apicomplexan parasites but is absent from mammals (Bentley, 1990; Roberts *et al.*, 1999). The shikimate pathway is a seven-step biosynthetic route that links the metabolism of carbohydrates to the synthesis of aromatic amino acids. The shikimate pathway leads to the biosynthesis of chorismate, which is a precursor of aromatic amino acids and many other aromatic compounds (Ratledge, 1982). Shikimate kinase (SK; EC 2.7.1.71), the fifth enzyme of this pathway, catalyzes phosphate transfer from ATP to the carbon-3-hydroxyl group of shikimate, forming shikimate 3-phosphate (S3P).

SK belongs to the nucleoside monophosphate (NMP) kinase structural family. SK is a  $\alpha/\beta$  protein consisting of a central sheet of five parallel  $\beta$ -strands flanked by  $\alpha$ -helices, with overall topology similar to that of adenylate kinase (Pereira *et al.*, 2004; Krell *et al.*, 1998, 2001). A characteristic feature of the NMP kinases is that they undergo large conformational changes during catalysis (Vonnrhein *et al.*, 1995). The NMP kinases are composed of three domains: the CORE, which contains a highly conserved phosphate-binding loop (P-loop), the LID domain, which undergoes substantial conformational changes upon substrate binding, and the NMP-binding domain,

## protein structure communications

which is responsible for the recognition and binding of a specific substrate (Gu *et al.*, 2002).

MgADP induces concerted hinged movements of the shikimate-binding (SB) and LID domains such that the two domains move towards each other and towards the active centre of the enzyme in the presence of this ligand (Gu *et al.*, 2002).

The precise mode of binding of shikimate to MtSK and some conformational changes upon shikimate binding to MtSK have recently been proposed (Pereira *et al.*, 2004; Dhaliwal *et al.*, 2004). The binding of shikimate to MtSK causes a concerted conformational

change of the LID and SB domains towards each other and results in an additional closure of the active site.

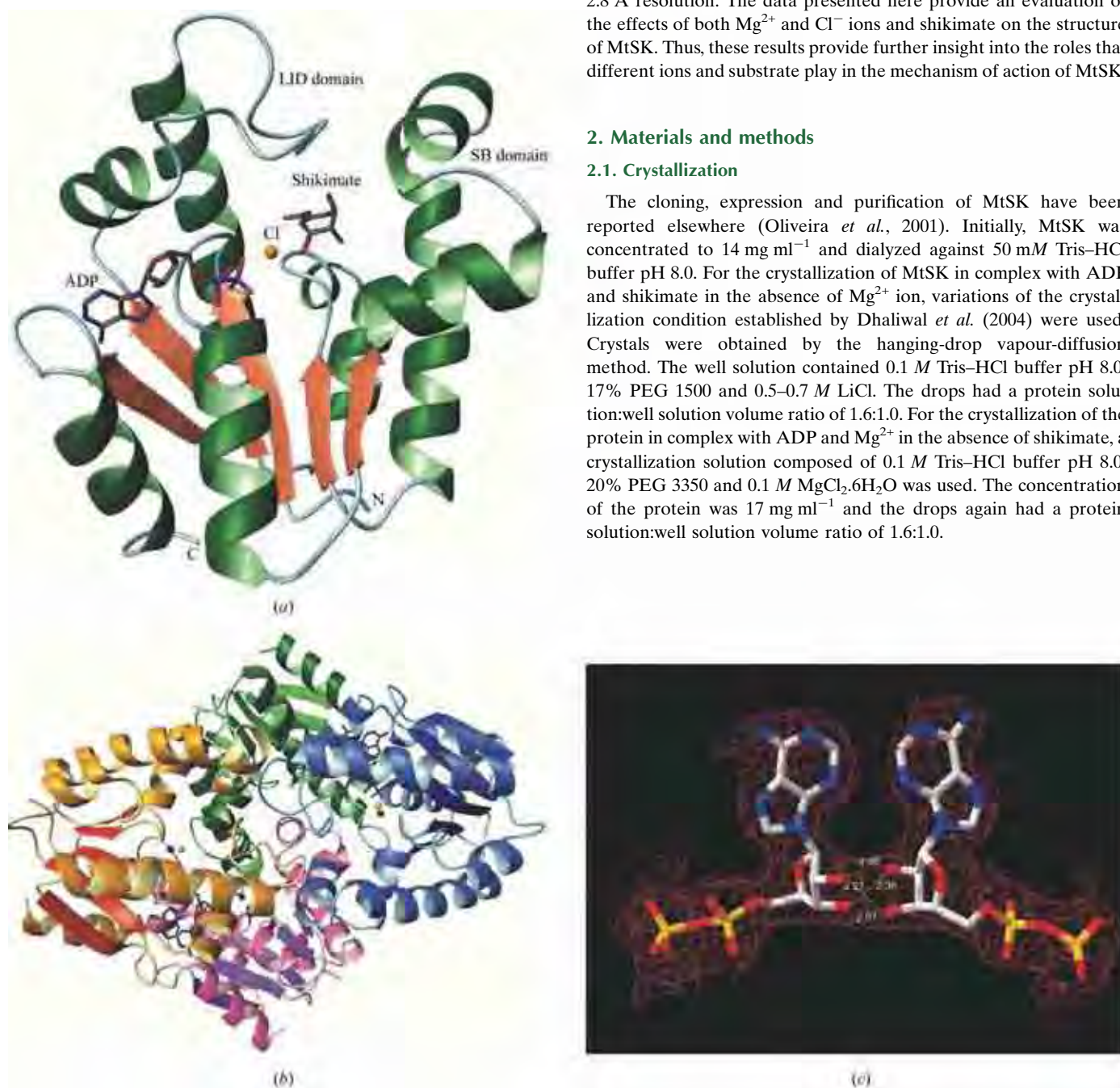
Chloride ions have been shown to weaken the interaction between shikimate and SK from *Erwinia chrysanthemi* and to strengthen the affinity of the enzyme for ADP and ATP (Cerasoli *et al.*, 2003). Thus, a chloride ion seems to occupy a site crucial for the binding of the nucleotide substrate in the correct orientation for catalysis.

In this work, we report two crystallographic structures: the MtSK–ADP–shikimate and MtSK–MgADP complexes. The MtSK–ADP–shikimate complex has been solved at 1.93 Å resolution. The crystal structure of MtSK in complex with ADP and Mg<sup>2+</sup> has been solved at 2.8 Å resolution. The data presented here provide an evaluation of the effects of both Mg<sup>2+</sup> and Cl<sup>-</sup> ions and shikimate on the structure of MtSK. Thus, these results provide further insight into the roles that different ions and substrate play in the mechanism of action of MtSK.

## 2. Materials and methods

### 2.1. Crystallization

The cloning, expression and purification of MtSK have been reported elsewhere (Oliveira *et al.*, 2001). Initially, MtSK was concentrated to 14 mg ml<sup>-1</sup> and dialyzed against 50 mM Tris–HCl buffer pH 8.0. For the crystallization of MtSK in complex with ADP and shikimate in the absence of Mg<sup>2+</sup> ion, variations of the crystallization condition established by Dhaliwal *et al.* (2004) were used. Crystals were obtained by the hanging-drop vapour-diffusion method. The well solution contained 0.1 M Tris–HCl buffer pH 8.0, 17% PEG 1500 and 0.5–0.7 M LiCl. The drops had a protein solution:well solution volume ratio of 1.6:1.0. For the crystallization of the protein in complex with ADP and Mg<sup>2+</sup> in the absence of shikimate, a crystallization solution composed of 0.1 M Tris–HCl buffer pH 8.0, 20% PEG 3350 and 0.1 M MgCl<sub>2</sub>·6H<sub>2</sub>O was used. The concentration of the protein was 17 mg ml<sup>-1</sup> and the drops again had a protein solution:well solution volume ratio of 1.6:1.0.



**Figure 1**

(a) Structure of MtSK in complex with ADP, shikimate and Cl<sup>-</sup>. (b) Tetrameric structure of MtSK in complex with ADP, Mg<sup>2+</sup> and Cl<sup>-</sup>. The monomers A, B, C and D are represented in green, blue, pink and yellow, respectively. The dark blue and yellow spheres represent the magnesium and chloride ions, respectively. The ADP molecules are represented as sticks. (c) Representation of the hydrogen-bonding interactions that occur between the ribose hydroxyl groups of ADP molecules bound to monomers A and B. The distances are shown in Å. Figures were generated with the program MolMol (Koradi *et al.*, 1996).

## 2.2. Data collection and processing

All data sets were collected at a wavelength of 1.427 Å using a synchrotron-radiation source (Station PCr, LNLS, Campinas, Brazil; Polikarpov *et al.*, 1998) and a CCD detector (MAR CCD). Data collection was performed under cryogenic conditions at a temperature of 100 K in a cold nitrogen stream generated and maintained with an Oxford Cryosystem. Prior to flash-cooling, glycerol was added to the crystallization drop to 20%(v/v). The data sets were processed using the program *MOSFLM* (Leslie, 1992) and scaled with *SCALA* (Collaborative Computational Project, Number 4, 1994).

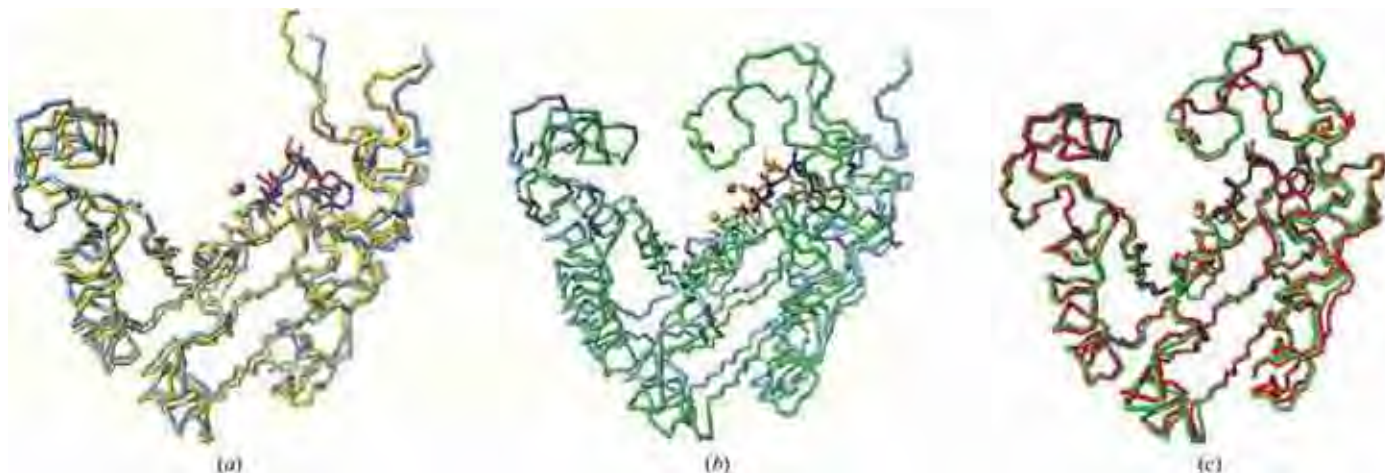
## 2.3. Structure determination

The crystal structures of both complexes were determined by standard molecular-replacement methods using the program *AMoRe*



**Figure 2**

Superposition of the structures of the MtSK-MgADP-shikimate and MtSK-ADP-shikimate ternary complexes. The C<sup>α</sup> trace of the MtSK-MgADP-shikimate complex is presented in green and that of the MtSK-ADP-shikimate trace is presented in blue. The C atoms of MtSK-MgADP-shikimate and MtSK-ADP-shikimate are coloured white and yellow, respectively. The Mg<sup>2+</sup> shown in yellow and the chloride ion shown in turquoise refer to the MtSK-MgADP-shikimate structure, while the chloride ion in dark blue refers to the MtSK-ADP-shikimate structure. The figure was generated with the program *MolMol* (Koradi *et al.*, 1996).



**Figure 3**

Superposition of the chains of the structure of the MtSK-MgADP binary complex obtained in the absence of shikimate. (a) Chains A (yellow) and C (light blue), (b) chains B (green) and C (light blue) and (c) chains B (green) and D (red).

(Navaza, 2001). For both complexes, we used as a search model the structure of MtSK-MgADP-shikimate (PDB code 1we2; Pereira *et al.*, 2004). Refinement of the structures was performed using *REFMAC5* implemented in the *CCP4* package (Murshudov *et al.*, 1997; Collaborative Computational Project, Number 4, 1994). *Xtal-View/Xfit* (McRee, 1999) was used for visual inspection and addition of water molecules. The stereochemical correctness of the models was checked using *PROCHECK* (Laskowski *et al.*, 1993). The final atomic models were superposed using *LSQKAB* from the *CCP4* package (Collaborative Computational Project, Number 4, 1994). *PAR-MODEL* (Uchôa *et al.*, 2004) was used in the final analysis of the model.

## 3. Results and discussion

MtSK crystallized in two different space groups depending on the complex. The crystals of the MtSK-ADP-shikimate ternary complex were trigonal, space group *P3<sub>2</sub>21*, and diffracted to 1.93 Å resolution. The asymmetric unit contains one molecule and the final values of *R* and *R<sub>free</sub>* were 20.2 and 27.0%, respectively. In contrast, the crystals of the MtSK-ADP-Mg<sup>2+</sup> ternary complex were orthorhombic, space group *P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>*, and diffracted to 2.8 Å resolution. The asymmetric unit contains four MtSK monomers that form a tetramer and the final values of *R* and *R<sub>free</sub>* were 18.3 and 28.0%, respectively. Table 1 details the data processing, refinement statistics and quality analysis of the two complexes.

The structures present good geometry, although some residues of the LID domain are located in disallowed regions of the Ramachandran plot. The N-terminal methionine residue is not observed since it was removed during the MtSK expression in *E. coli* (Oliveira *et al.*, 2001). The ten C-terminal residues are disordered in both structures and have not been included in the final structure.

The folding of the MtSK complexes presented here is similar to that of those reported previously (Pereira *et al.*, 2004; Krell *et al.*, 1998; Gu *et al.*, 2002; Dhaliwal *et al.*, 2004). SK is a  $\alpha/\beta$  protein and consists of five central parallel  $\beta$ -strands flanked by  $\alpha$ -helices. Fig. 1(a) shows a ribbon representation of the secondary-structure elements of MtSK-ADP-shikimate at 1.93 Å resolution.

## protein structure communications

### 3.1. Structure of the MtSK–ADP–Mg<sup>2+</sup> ternary complex

The asymmetric unit of the MtSK–ADP–Mg<sup>2+</sup> ternary complex structure contains four monomers forming a homotetramer with 222 symmetry. However, each subunit presents a different conformation state. In the *A* and *C* subunits the LID domain is disordered and residues 115–123 and 114–123 were thus not included in the final model.

The structure of the MtSK tetramer is shown in Fig. 1(*b*). Each monomer is in contact with the other three, creating an intricate packing arrangement. The ADP molecules appear to play an important role in the stabilization of the tetramer, since the hydroxyl groups of the ribose moiety of ADP of one monomer form hydrogen bonds with those of a neighbouring monomer (Fig. 1*c*).

### 3.2. Influence of Mg<sup>2+</sup> on the structure of MtSK

The absence of Mg<sup>2+</sup> ions seems to have a significant effect on the position of shikimate and some of the active-site residues, mainly Asp32 and Asp34 (these are conserved residues in all SKs and are involved in the binding of Mg<sup>2+</sup>; Fig. 2). In addition, the chloride ion and ADP molecule also undergo changes in position. Fig. 2 shows a

superposition of the MtSK–MgADP–shikimate (Pereira *et al.*, 2004) and MtSK–ADP–shikimate complexes, showing the structural changes caused by the absence of the Mg<sup>2+</sup> ion.

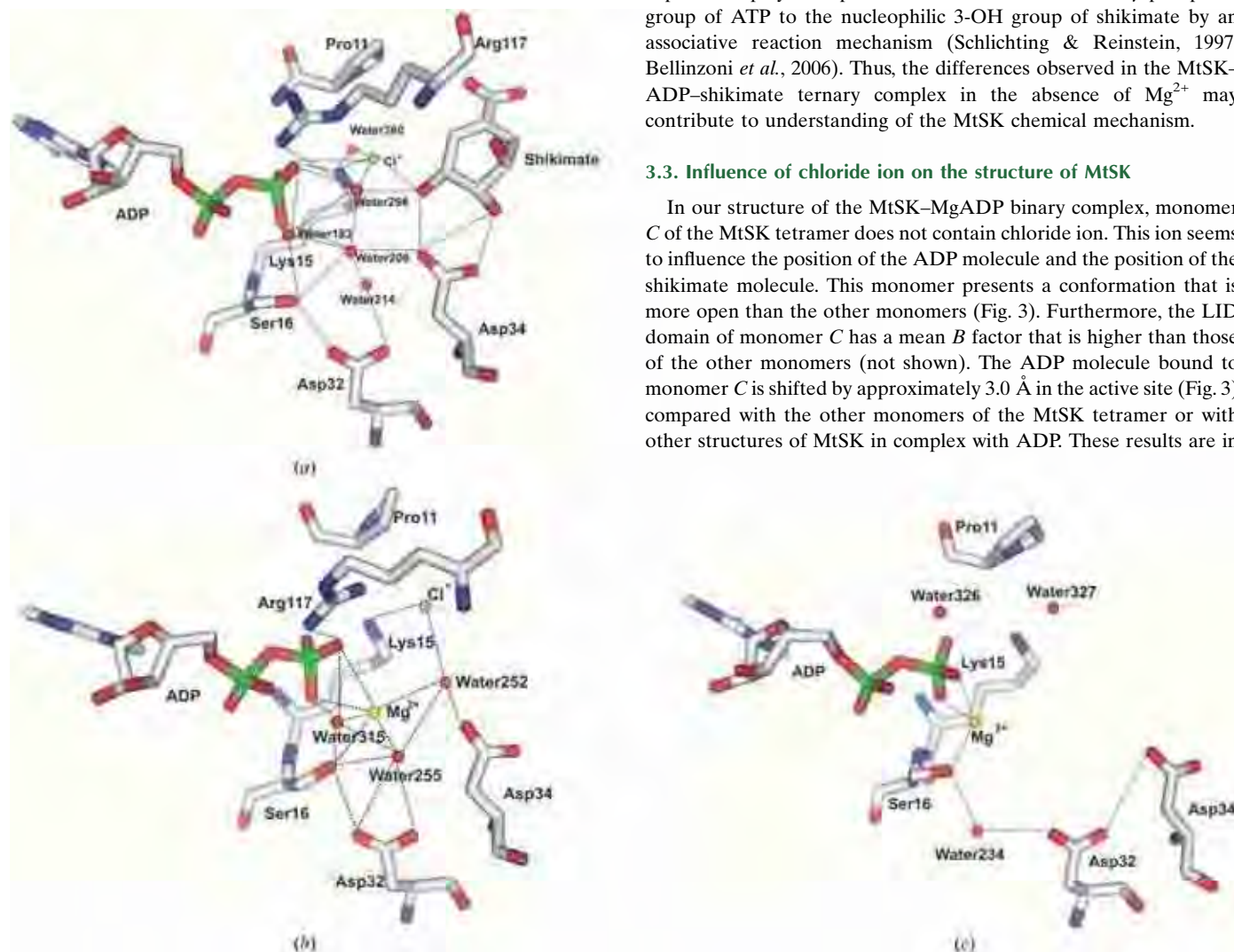
In the structure of the MtSK–ADP–shikimate ternary complex, the 3-, 4- and 5-hydroxyl groups of shikimate undergo a shift in their positions, leading to small differences in the hydrogen-bonding pattern between shikimate and MtSK. Although the Asp34 residue interacts with shikimate in both structures, a hydrogen bond between the OD2 atom of Asp34 and the O3 atom of the shikimate is not observed in the MtSK structure without Mg<sup>2+</sup> (not shown).

The position of the chloride ion is closer to shikimate in the MtSK–ADP–shikimate ternary complex than in the MtSK–MgADP–shikimate complex (Pereira *et al.*, 2004). In the structure of the MtSK–ADP–shikimate ternary complex, the chloride ion is 2.4 Å away from the O1 atom of the shikimate 3-hydroxyl group, while in the structure of MtSK–MgADP–shikimate the distance between these two atoms is 3.4 Å. The absence of the Mg<sup>2+</sup> ion can also cause alterations in the side chains of the hydrophobic residues Phe49 and Phe57 located in the SB domain and of Pro118, Ala46 and Ile45 located in the LID domain (Fig. 2).

As established previously for other NMP kinases, the Mg<sup>2+</sup> ion is expected to play an important role in the transfer of the  $\gamma$ -phosphate group of ATP to the nucleophilic 3-OH group of shikimate by an associative reaction mechanism (Schlichting & Reinstein, 1997; Bellinzoni *et al.*, 2006). Thus, the differences observed in the MtSK–ADP–shikimate ternary complex in the absence of Mg<sup>2+</sup> may contribute to understanding of the MtSK chemical mechanism.

### 3.3. Influence of chloride ion on the structure of MtSK

In our structure of the MtSK–MgADP binary complex, monomer *C* of the MtSK tetramer does not contain chloride ion. This ion seems to influence the position of the ADP molecule and the position of the shikimate molecule. This monomer presents a conformation that is more open than the other monomers (Fig. 3). Furthermore, the LID domain of monomer *C* has a mean *B* factor that is higher than those of the other monomers (not shown). The ADP molecule bound to monomer *C* is shifted by approximately 3.0 Å in the active site (Fig. 3) compared with the other monomers of the MtSK tetramer or with other structures of MtSK in complex with ADP. These results are in



**Figure 4**

Pattern of water bonding in the active site of MtSK. (*a*) Structure at 1.93 Å resolution (chloride ion and shikimate present, magnesium ion absent), (*b*) monomer *B* of the tetramer (chloride and magnesium ions present, shikimate absent), (*c*) monomer *C* of the tetramer (magnesium ion present, shikimate and chloride ion absent). The figures were generated using *PyMOL* (DeLano, 2004).

**Table 1**

Crystallographic data, refinement statistics and analysis of the quality of MtSK structures.

Values in parentheses are for the outermost shell.

	MtSK-ADP-shikimate	MtSK-MgADP
Crystallographic data		
Unit-cell parameters		
<i>a</i> (Å)	63.3	60.2
<i>b</i> (Å)	63.3	62.2
<i>c</i> (Å)	91.6	170.6
Space group	<i>P</i> <sub>3</sub> <sup>2</sup> <sub>2</sub> <sup>1</sup>	<i>P</i> <sub>2</sub> <sup>1</sup> <sub>2</sub> <sup>1</sup> <sub>2</sub> <sup>1</sup>
No. of measurements	76792	97059
No. of independent reflections	16017	17057
Completeness (%)	96.9 (91.5)	99.4 (96.9)
<i>R</i> <sub>sym</sub> † (%)	8.9 (58.8)	12.1 (58.0)
Redundancy	4.8	5.7
Refinement statistics		
Resolution range (Å)	35.16–1.93	57.17–2.80
Reflections used for refinement	15130	15670
Final <i>R</i> factor‡ (%)	20.2	18.3
Final <i>R</i> <sub>free</sub> § (%)	27.0	28.0
Correlation coefficient (%)	95.2	94.3
<i>B</i> values (Å <sup>2</sup> )		
Main chain	31	33
Side chain	34	36
ADP	21	27
Shikimate	32	—
Waters	39	30
Quality of structure		
Three-dimensional profile¶	S = 88.07, IS = 74.95, S/IS = 1.18IS	S = 343.59, IS = 294.34, S/IS = 1.17IS
Ramachandran plot		
Favoured	95.6	84.8
Additionally allowed	2.9	13.3
Generously allowed	0.7	0.8
Disallowed	0.7	1.1

†  $R_{\text{sym}} = 100 \sum I(h) - \langle I(h) \rangle / \sum I(h)$ , where  $I(h)$  is the observed intensity and  $\langle I(h) \rangle$  is the mean intensity of reflection  $h$  over all measurements of  $I(h)$ . ‡ *R* factor =  $100 \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$ , the sums being taken over all reflections with  $F/\sigma(F) > 2\sigma(F)$ . § *R*<sub>free</sub> is the *R* factor for 10% of the data that were not included during crystallographic refinement. ¶ The ideal score measures the compatibility of a protein model with its sequence, using a 3D profile. Each residue position in the 3D model is characterized by its environment and is represented by a row of 20 numbers in the profile. These numbers are the statistical preferences (called 3D-1D scores) of each of the 20 amino acids for this environment. Environments of residues are defined by three parameters: the area of the residue that is buried; the fraction of side chain area that is covered by polar atoms (O and N) and the local secondary structure. The 3D profile score *S* for the compatibility of the sequence with the model is the sum, over all residue positions, of the 3D-1D scores for the amino-acid sequence of the protein. For 3D protein models known to be correct, the 3D profile score *S* for the amino-acid sequence of the model is high, by contrast, the profile score *S* for the compatibility of a wrong 3D protein model with its sequence is often low. When this method is used to verify a structure, the raw compatibility score alone is difficult to interpret. In this case it is necessary to compare the score to those obtained using structures known to be correct, we use the Ideal Score (IS), that is calculated from the length of the protein. The IS is determined by  $\text{IS} = \exp[-0.83 + 1.008 \times \ln(L)]$ . Where *L* is the length of the sequence. Severely misfolded structures typically have scores less than 0.45 IS. A score near or above IS indicates a reliable structure.

agreement with the kinetic and spectroscopic data previously obtained by Cerasoli *et al.* (2003). This work shows that the chloride ion increases the stability of the *E. chrysanthemi* SK structure and that this same ion also influences the affinity of ADP for SK. Thus, the increase in the protein stability may be a consequence of chloride favouring the SK structure in its closed state.

The chloride ion bound to the MtSK active site seems to be part of an intricate network formed of water molecules, residues of the LID and SB domains, ADP and shikimate. This network of interactions appears to cause closure of the structure (Fig. 4). In all structures of MtSK reported so far and in the three monomers of the MtSK tetramer, an interaction network involving three water molecules is formed in the active site (Fig. 4). These water molecules bridge the interactions between ADP, chloride and magnesium ion and MtSK. In our MtSK-ADP-shikimate ternary complex, water 306 occupies a

similar position to the Mg<sup>2+</sup> ion. In the absence of chloride ion, this phenomenon does not occur and furthermore the formation of this intricate network that can induce the opening of the structure is avoided.

### 3.4. Influence of shikimate in the structure of MtSK

The crystals of the MtSK-MgADP binary complex in the absence of shikimate were obtained in space group *P*<sub>2</sub><sup>1</sup><sub>2</sub><sup>1</sup><sub>2</sub><sup>1</sup>, which has not been previously described for MtSK. The asymmetric unit presents four MtSK monomers that form a tetramer. The absence of shikimate from the crystallization conditions influences the crystal packing and also the conformation of the SB and LID domains of MtSK, as observed by Pereira *et al.* (2004) and Dhaliwal *et al.* (2004). In accordance with this, Gan and coworkers recently solved the apo-MtSK and MtSK-shikimate binary complex structures and suggested that shikimate binding defines the conformational change of the protein that arises when shikimate is bound: the LID domain is ordered and closes over the bound shikimate (Gan *et al.*, 2006).

## 4. Conclusion

Here, we report two structures of shikimate kinase from *M. tuberculosis*: the MtSK-ADP-shikimate and MtSK-MgADP complexes. In the former, we observe the effect of the Mg<sup>2+</sup> ion on the structure and in the latter we observe the effect of shikimate on the crystal packing and on the structure of MtSK. The Mg<sup>2+</sup> ion seems to influence the position of the hydroxyl groups of the shikimate molecule and some of the residues of the active site of MtSK. The crystal structure of the MtSK-MgADP complex was solved in space group *P*<sub>2</sub><sup>1</sup><sub>2</sub><sup>1</sup><sub>2</sub><sup>1</sup>, which has not previously been described for MtSK. In this space group, the MtSK presents a tetramer with 222 symmetry, in which the ribose moiety of the ADP molecule seems to play an important role in the stabilization of the tetramer and the contacts between the monomers, which occur mainly in the LID-domain region. However, one monomer of the tetramer does not contain chloride ion. The absence of this ion seems to cause large changes in the position of the ADP molecule and also causes a large opening of MtSK. This information is accordance with the results obtained by Cerasoli *et al.* (2003), which shows the importance of the chloride ion in the stability and the alignment of the substrates in the active site of SK.

We hope that the results described here will shed light on the structural changes of MtSK upon binding of substrate(s) that will be useful for the understanding of the catalytic mechanism and for structure-based drug design of novel inhibitors that may be potential anti-mycobacterial agents.

This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0, 03/12472-2, 04/00217-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT), DSS, WFA (CNPq, 300851/98-7) and LAB (CNPq, 520182/99-5) are researchers of the Brazilian Council for Scientific and Technological Development.

## References

- Bellinzoni, M., Haouz, A., Grana, M., Munier-Lehmann, H., Shepard, W. & Alzari, P. M. (2006). *Proteins*, **15**, 1–5.  
 Bentley, R. (1990). *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384.  
 Cerasoli, E., Kelly, S. M., Coggins, J. R., Lapthorn, A. J., Clarke, D. T. & Price, N. C. (2003). *Biochim. Biophys. Acta*, **1648**, 43–54.  
 Coggins, J. R. (1998). In *Herbicides and Plant Metabolism*, edited by A. Dodge. Cambridge University Press.

## protein structure communications

- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Davies, G. M., Barret-Bee, K. J., Jude, D. A., Lehan, M., Nichols, W. W. & Pinder, P. E. (1994). *Agents Chemother.* **38**, 403–406.
- DeLano, W. L. (2004). *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.
- Dhaliwal, B., Nichols, C. E., Ren, J., Lockyer, M., Charles, I., Hawkins, A. R. & Stammers, D. K. (2004). *FEBS Lett.* **574**, 49–54.
- Gan, J., Gu, Y., Li, Y., Yan, H. & Ji, X. (2006). *Biochemistry*, **45**, 8539–8545.
- Gu, Y., Reshetnikova, L., Li, Y., Wu, Y., Yan, H., Singh, S. & Ji, S. (2002). *J. Mol. Biol.* **319**, 779–789.
- Koradi, R., Billeter, M. & Wüthrich, K. (1996). *J. Mol. Graph.* **14**, 51–55.
- Krell, T., Coggins, J. R. & Laphorn, A. J. (1998). *J. Mol. Biol.* **278**, 983–997.
- Krell, T., Maclean, J., Boam, D. J., Cooper, A., Resmini, M., Brocklehurst, K., Kelly, S. M., Price, N. C., Laphorn, A. J. & Coggins, J. (2001). *Protein Sci.* **10**, 1137–1149.
- Laskowski, R. A., MacArthur, M., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Leslie, A. G. W. (1992). *Int CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* **D53**, 240–255.
- Navaza, J. (2001). *Acta Cryst.* **D57**, 1367–1372.
- Oliveira, J. S., Pinto, C. A., Basso, L. A. & Santos, D. S. (2001). *Protein Expr. Purif.* **22**, 430–435.
- Pereira, J. H., Oliveira, J. S., Canduri, F., Dias, M. V. B., Palma, M. S., Basso, L. A., Santos, D. S. & Azevedo, W. F. Jr (2004). *Acta Cryst.* **D60**, 2310–2319.
- Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. C. & Craievich, A. (1998). *Nucl. Instrum. Methods Phys. Res. A*, **405**, 159–164.
- Ratledge, C. (1982). *The Biology of the Mycobacteria*, Vol. 1, edited by C. Ratledge & J. L. Stanford, pp. 185–271. London: Academic Press.
- Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. E., Krell, T., Coggins, G. H., Milhous, W. K., Tzipoki, S., Ferguson, D. J., Chakrabarti, D. & McLeod, R. (1999). *Nature (London)*, **397**, 219–220.
- Schlichting, I. & Reinstein, J. (1997). *Biochemistry*, **36**, 9290–9296.
- Uchôa, H. B., Jorge, G. E., Silveira, N. J. F., Câmara, J. C., Canduri, F. & Azevedo, W. F. Jr (2004). *Biochem. Biophys. Res. Commun.* **325**, 1481–1486.
- Vonrhein, C., Schlauderer, G. J. & Schulz, G. F. (1995). *Structure*, **3**, 483–490.
- World Health Organization (2006). *Tuberculosis – Fact Sheet No. 104*. <http://www.who.int/mediacentre/factsheets>.



### **3.5 Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-trans- Enoyl-ACP (CoA) Reductase from *Mycobacterium tuberculosis***

Marcio V. B. Dias; Adriane M. X. Prado; Igor B. Vasconcelos; Valmir Fadel; Luis A. Basso; Diógenes S. Santos; Walter F. de Azevedo Jr. *Artigo aceito no Journal Structural Biology, 2007.*

Neste trabalho são apresentadas as estruturas cristalográficas da enzima Enoil (ACP) redutase (InhA) (peso molecular de 28528 Da) em complexo com INH-NAD na forma selvagem e para dois mutantes encontrados em isolados clínicos resistentes a isoniazida (S94A e I21V). É apresentada também a estrutura do mutante S94A na sua forma nativa.

As proteínas selvagens e os mutantes S94A e I21V foram expressadas e purificadas pelo grupo do Prof. Dr. Diógenes S. Santos da PUC – Porto Alegre –RS. Os cristais da InhA (selvagem e mutantes) em complexo com INH-NAD foram produzidos seguindo as condições estabelecidas por Dessen et al., 1995 e Rozwarski et al., 1998. Para a repetição destas condições as proteínas foram dialisadas contra HEPES, 50 mM, pH 7,5. A condição de cristalização na qual foram obtidos os melhores cristais era composta por HEPES, 50 mM, pH 7.2, citrato de sódio, 50 mM e 5-10% de MPD. Para a cristalização do mutante S94A na forma nativa, a proteína foi dialisada contra HEPES, 50 mM, pH 7,5, 10% de glicerol e KCl, 300 mM. A condição na qual foram obtidos os melhores cristais apresentava citrato de sódio, 100 mM, pH 5,6, acetato de amônio, 200 mM e 20-30% de PEG 4000. Os dados de difração de raios X foram coletados no LNLS. Os cristais obtidos para a InhA em complexo com INH-NAD são pertencentes

ao grupo espacial P6<sub>22</sub> e os cristais obtidos para a proteína na sua forma nativa são pertencentes ao grupo espacial P1. Todas as estruturas foram determinadas por substituição molecular e refinadas pelo programa REFMAC 5.2 (MURSHUDOV; VAGIN; DODSON, 1997).

Os dados obtidos para o complexo da InhA selvagem com INH-NAD apresenta uma melhor resolução (2.2 Å) do que os dados anteriormente descritos na literatura (2.8 Å) (ROZWARSKI et al., 1998) . Desta maneira, é possível realizar uma melhor análise do modo de interação desta droga com o sítio ativo da enzima, e evidenciar alterações ainda não descritas. Além disso, esta estrutura, devido a sua melhor resolução, é mais confiável para estudos *a posteriori* no desenvolvimento de novas drogas contra tuberculose. As estruturas determinadas para os mutantes S94A e I21V contribuem para a compreensão do mecanismo de resistência destes mutantes à isoniazida, além de observar a influência que estas mutações causam sobre o processo de ligação do aducto INH-NAD no sítio ativo da enzima. A estrutura do mutante S94A na sua forma nativa, é a primeira estrutura apresentada neste estado, e tem sido obtida em uma nova condição de cristalização e em um novo grupo espacial (P1). A estrutura apresenta um tetrâmero na unidade assimétrica, que pode corresponder ao tetrâmero encontrado em solução. Nesta estrutura, foi possível observar o estado da enzima antes da ligação de qualquer ligante e, desta forma, pôde dar evidências sobre os movimentos de *loops* e resíduos que estão envolvidos no processo de ligação de substratos acil ácidos graxos, da molécula de NADH ou do aducto INH-NAD.

**Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-*trans*- Enoyl-ACP (CoA) Reductase from *Mycobacterium tuberculosis***

Marcio Vinicius Bertacine Dias<sup>a</sup>, Igor Bordin Vasconcelos<sup>b</sup>, Adriane Michele Xavier Prado<sup>a</sup>, Valmir Fadel<sup>a</sup>, Luiz Augusto Basso<sup>b</sup>, Walter Filgueira de Azevedo Jr.<sup>c\*</sup>, Diógenes Santiago Santos<sup>b\*</sup>,

<sup>a</sup> Programa de Pós-Graduação em Biofísica Molecular – Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brasil;

<sup>b</sup> Centro de Pesquisa em Biologia Molecular e Funcional, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil;

<sup>c</sup> Faculdade de Biociências - Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681. Porto Alegre-RS CEP 90619-900 Brasil.

\* Corresponding authors: [diogenes@pucri.br](mailto:diogenes@pucri.br) (D.S. Santos) and [walter.junior@pucri.br](mailto:walter.junior@pucri.br) (W.F. de Azevedo Jr.).

telephone: +55 51 3320-3500 4529

fax: +55 17 51 3220-3629

## Abstract

The resumption of tuberculosis led to an increased need to understand the molecular mechanisms of drug action and drug resistance, which should provide significant insight into the development of newer compounds. Isoniazid (INH), the most prescribed drug to treat TB, inhibits an NADH-dependent enoyl-acyl carrier protein reductase (InhA) that provides precursors of mycolic acids, which are components of the mycobacterial cell wall. InhA is the major target of the mode of action of isoniazid. INH is a pro-drug that needs activation to form the inhibitory INH-NAD adduct. Missense mutations in the *inhA* structural gene have been identified in clinical isolates of *M. tuberculosis* resistant to INH. To understand the mechanism of resistance to INH, we have solved the structure of two InhA mutants (I21V and S94A), identified in INH-resistant clinical isolates, and compare them to INH-sensitive WT InhA structure in complex with the INH-NAD adduct. We also solved the structure of unliganded INH-resistant S94A protein, which is the first report on apo form of InhA. The salient features of these structures are discussed and should provide structural information to should improve our understanding of the mechanism of action of, and resistance to, INH in *M. tuberculosis*. The unliganded structure of InhA allows identification of conformational changes upon ligand binding and should help structure-based drug design of more potent antimycobacterial agents.

**Keywords:** *M. tuberculosis*; InhA; Crystal structure; Isoniazid; drug resistance

## Introduction

Tuberculosis (TB), which is caused mainly by *Mycobacterium tuberculosis*, is a global human health emergency that remains the leading cause of mortality among infectious diseases. It has been estimated that 8.2 million new TB cases occurred worldwide in the year 2000, with approximately 1.8 million deaths in the same year, which translates into more than 200 deaths per hour, and more than 95 % of these were in developing countries (Corbett et al., 2003). In the same year, 3.2 % of the world's new cases of TB were multidrug-resistant tuberculosis (MDR-TB), defined as strains resistant to at least isoniazid and rifampicin (Espinal, 2003; Ormerod, 2005). Treatment of MDR-TB strains requires the administration of second-line drugs that are more toxic and less effective, and are given for at least three times as long as, and 100 times as expensive as basic chemotherapeutic regimens (Pablos-Mendes et al., 2002). More recently, a survey of the frequency and distribution of extensively drug-resistant (XDR) TB cases, which are defined as cases in persons with TB whose isolates were resistant to isoniazid and rifampicin and at least three of the six main classes of second-line drugs, showed that during 2000-2004, of 17,690 TB isolates, 20% were MDR and 10% of these were XDR (CDC, 2006). XDR-TB has a wide geographic distribution, poses a public health threat, is an impediment to TB control, and opens up the possibility that epidemics of virtually untreatable TB may develop (No authors listed, 2006). New antimycobacterial agents are thus needed to improve the treatment of MDR- and XDR-TB, as well as to provide more effective treatment of drug-sensitive TB infection. An understanding of drug resistance mechanisms in this pathogen should contribute to the rational design of new chemotherapeutic agents to treat TB.

The modern, standard “short-course” therapy for tuberculosis is based on a four-drug regimen of isoniazid, rifampicin, pyrazinamide, and ethambutol or streptomycin

for two months, followed by treatment with a combination of isoniazid and rifampicin for four months (Mitchison, 1985). Isoniazid (INH, isonicotinic acid hydrazide) was first reported to be effective in the treatment of TB in 1952 (Bernstein et al., 1952) and, soon after, the first INH-resistant *M. tuberculosis* strains were isolated (Middlebrook, 1953). Genetic and biochemical studies have shown that the *inhA*-encoded protein is the primary target for isoniazid (Banerjee et al., 1994; Quémard et al., 1995; Larsen et al., 2002; Kremer et al., 2003). InhA was identified as an NADH-dependent 2-*trans* enoyl-ACP (acyl carrier protein) reductase enzyme that exhibits specificity for long-chain thioester substrates. InhA is a member of the mycobacterial type II fatty acid synthase system (FAS-II), which elongates acyl fatty acid precursors yielding the long carbon chain of the meromycolate branch of mycolic acids, the hallmark of mycobacteria (Schroeder et al., 2002). INH is a pro-drug that is activated by the mycobacterial catalase-peroxidase enzyme KatG in the presence of manganese ions, NAD(H) and oxygen (Johnsson and Schultz, 1994; Johnsson et al., 1995; Basso et al., 1996; Zabinski and Blanchard, 1997). The KatG-produced acylpyridine fragment of isoniazid is covalently attached to the C4 position of NADH forming an INH-NAD adduct, which, in turn, forms an inhibitory binary complex with the wild-type (WT) enoyl reductase of *M. tuberculosis* (Rozwarski et al., 1998) with an equilibrium dissociation constant value lower than 0.4 nM (Lei et al., 2000). The isonicotinyl-NAD adduct has been shown to be a slow, tight-binding competitive inhibitor of WT InhA with an overall inhibition constant value of 0.75 nM (Rawat et al., 2003).

The mechanism of action of isoniazid is complex, as mutations in at least five different genes (*katG*, *inhA*, *ahpC*, *kasA*, and *ndh*) have been found to correlate with isoniazid resistance (Schroeder et al., 2002; Blanchard, 1996; Basso and Blanchard, 1998; Glickman and Jacobs, 2001; Basso and Santos, 2005; Oliveira et al., 2007).

Consistent with InhA as the primary target of INH mode of action, INH-resistant clinical isolates of *M. tuberculosis* harboring *inhA*-structural gene missense mutations, but lacking mutations in the *inhA* promoter region, *katG* gene and *oxyR-ahpC* region, were shown to have higher dissociation constant ( $K_d$ ) values for NADH than INH-sensitive WT InhA, whereas there were only modest differences in the steady-state parameters (Blanchard, 1996). We have recently reported the crystal structures of binary complexes formed between NADH and INH-sensitive WT InhA, and INH-resistant S94A, I47T, and I21V InhA mutant enzymes (Oliveira et al., 2006). Even more recently, both specialized linkage transduction has been used to introduce S94A single point mutation within the *inhA* structural gene and X-ray crystallographic on INH-resistant S94A InhA protein has been reported (Vilchèze et al., 2006). However, even though there are several crystal structures of InhA in complex with a variety of ligands, there has been no report on unliganded InhA structure and, thus, no high resolution information on the InhA structure before ligand binding. In our efforts to understand the molecular basis for the reduced inhibition of the INH-NAD adduct to InhA mutants, here we report co-crystallization of INH-resistant I21V and S94A InhA mutant enzymes, which were identified in INH-resistant clinical isolates of *M. tuberculosis* (Blanchard, 1996; Morlock et al., 2003), with the INH-NAD adduct, and compare them to the INH-sensitive WT InhA structure. This is the first report on the crystal structure of the complex formed between INH-resistant I21V InhA and INH-NAD adduct refined to 2.2 Å. We also report the crystal structure of INH-resistant S94A InhA and INH-sensitive WT InhA both in complex with INH-NAD adduct to 2.0 and 2.2 Å of resolution, respectively. Moreover, we report, for the first time, the crystal structure of apo INH-resistant S94A InhA refined to 2.15 Å, which shows the protein conformational changes upon ligand binding. It is hoped that the data presented here

will provide structural insight into an understanding of the drug resistance mechanism, which, in turn, should aid the rational design of chemical compounds to efficiently inhibit both INH-resistant and -sensitive InhA enzymes with potential antimycobacterial activity.

## **Materials and methods**

### **Crystallization**

WT, I21V and S94A InhA enzymes were expressed and purified to homogeneity as described elsewhere (Quèrmard et al., 1995; Basso et al., 1998). INH-NAD synthesis was carried out as described elsewhere (Rozwarski et al., 1998). Crystals of binary complex InhA:INH-NAD were obtained by the hanging-drop vapor-diffusion method under similar conditions as described by Dessen et al., (1995) and Rozwarski et al., (1998). InhA enzymes were dialyzed against 50 mM Hepes, pH 7.5 and concentrated to 5-10mg/mL. Enzyme-inhibitor complexes were obtained at room temperature by incubating 5-10 mg/mL InhA proteins with NADH, MnCl<sub>2</sub> and isoniazid for 1 hour using molar ratios of 1:50, 1:10, and 1:100, respectively. The complexes InhA:INH-NAD were crystallized in hanging droplets containing 1 µL of inhibited-protein solution and 1 µL of crystallization solution containing 50 mM Hepes, pH 7.2; Sodium citrate buffer and 5-10% MPD.

The apoenzyme S94A InhA mutant was dialyzed against 50 mM Hepes, pH 7.5, 10% glycerol and 300 mM of KCl. The protein was concentrated to 10 mg/mL and crystallized in hanging droplets containing 1 µL of protein solution and 0.5 µL of crystallization solution containing 100 mM sodium citrate, pH 5.6, 200 mM ammonium acetate and 20-30% of PEG 4000.



### **Data collection and processing**

All data sets were collected at a wavelength of 1.427 Å using Synchrotron Radiation Source (Station PCr, LNLS, Campinas, Brazil; and a CCD detector (MARCCD) (Polikarpov et al., 1998). Data collection was performed in cryogenic conditions at 100 K in a cold nitrogen stream generated and maintained with an Oxford Cryosystem. Prior to flash-cooling, glycerol was added to the crystallization drop up to 20 % (v/v). The data sets were processed using the program MOSFLM (CCP4, 1994) and scaled with SCALA (CCP4, 1994).

### **Structure determination**

The crystal structures were determined by standard molecular replacement methods using the program AMoRe (Navaza, 1994). Initially, atomic coordinates for binary complex WTInhA:INH-NAD to 2.7Å resolution were used as search model for the structure of WT InhA:INH-NAD (PDB access code: 1ZID) (Rozwarski et al., 1998). Our atomic coordinates for WT InhA:INH-NAD at 2.2Å resolution were then used as search model to solve the other structures. The structure of apoenzyme S94A InhA mutant was solved using as search model the structure of binary complex InhA:Genz-10850 at 2.6Å resolution (PDB access code: 1P44) (Kuo et al., 2003). The atomic positions, which generated the higher correlation coefficient magnitude obtained from molecular replacement method, were used for the crystallographic refinement. The refinements of structures were performed using REFMAC 5.2 program (Murshudov et al., 1997). The XtalView/Xfit (McRee, 1999) was used for visual inspection and addition of water molecules. The water molecules were also checked based on B factor values. The stereochemistry correctness of the models was checked using PROCHECK

program (Laskowski et al., 1993) and PARMODEL (Uchoa et al., 2004). The final atomic models were superposed using the program LSQKAB program from CCP4i package (CCP4, 1994).

## Results and discussion

The crystals of apo S94A InhA are triclinic. These crystals diffracted to 2.15 Å resolution. The asymmetric unit presents four monomers forming the characteristic tetramer of InhA. This structure presents final R-factor and R-free values of 16.2% and 25.5%, respectively. The crystals of WT, I21V and S94A InhA in complex with the INH-NAD adduct are hexagonal and crystallized in the space group P6<sub>2</sub>2<sub>2</sub>, having one molecule in the asymmetric unit. Table 1 summarizes the data processing, crystallographic refinement statistics and structural quality for the four structures presented here. Analysis of the crystallographic refinement for the structures here described indicates that WtInhA:NADH:INH, I21V InhaA:NADH-INH and S94Inha:NADH-INH present difference between R-factor and R-free ranging from 3.9 to 5.5 %, which indicates fairly good overall refinement statistics, and the structure S94A InhA presents a difference of 9.3 %, below 10% , which makes this structure also acceptable for structural comparisons. These structures present good geometry, although some residues are located in regions not permitted in Ramachandran plots.

InhA belongs to the short chain dehydrogenase/reductase (SDR) family of enzymes. The main characteristic of this family is a polypeptide backbone topology in which each subunit consists of a single domain with a central core that contains a Rossmann fold supporting an NADH binding site. The structure displays a  $\alpha/\beta$  folding consisting of a central  $\beta$ -sheet composed of parallel strands and flanked by  $\alpha$ -helices

(Figure 1A). The structure presented here is in accordance with the homotetrameric quaternary structure in solution determined by analytical size-exclusion chromatography, and possesses an internal 222 symmetry (figure 1B).

### **Influence of INH on the structure of MtInhA**

The activated form of INH consists of an isonicotinic-acyl group attached through its carbonyl group to the C4 of the nicotinamide ring, replacing the 4S hydrogen of NADH, which is the same position involved in the hydride transfer that occurs during reduction of enoyl-ACP substrates (Quémard et al., 1995).

We compared the structure of INH-sensitive WT InhA presented here and the structure previously solved by Roswarski et al., (1998) at 2.7 Å of resolution. Although Vilchèze et al., (2006) have solved two structures of InhA in complex with NAD-INH (WT InhA and S94A mutant at 2.0 Å and 1.9 Å at resolution respectively) it was only observed one structure deposited in the Protein Data Bank (S94A InhA mutant). This way, the S94A InhA solved by Vilchèze et al., (2006) is argued in related to effect of this mutation on the binding process of the INH-NAD at a later stage.

The structure of the INH-sensitive WT InhA presented here shows differences when compared with the structure determined by Roswarski et al., (1998), mainly in important residues located in the binding site of INH-NAD (Tyr158, Phe149, Trp222, Leu218 and Phe41) and in the INH-NAD molecule. There is a difference of approximately 15° in the position of Tyr158 side chain of WT InhA in comparison with the structure solved by Roswarski et al., (1998), and the Tyr158 side chain moves closer to INH. Interestingly, a conformational change involving rotation of Tyr158 side chain upon binding of the enoyl substrate to InhA has been invoked to account for the observed inverse solvent isotope effect (Parikh et al., 1999). These authors have

proposed that Tyr158 functions as an electrophilic catalyst, stabilizing the transition state for hydride transfer by hydrogen bonding to the substrate carbonyl. The Phe41 side chain of WT InhA:INH-NAD structure presented here underwent 30° torsion as compared to the structure solved by Roswarski et al., (1998). This residue appears to be important in anchoring the adenine moiety of NADH. The side chain of Phe149 and Leu218 residues move closer to INH molecule (Figure 2A). Although it could be argued that these differences are due to the lower resolution of the previously published structure, these residues appear to play an importance role in the enzyme-ligand interaction and should be considered in structure-based drug design.

The main effect of INH moiety of INH-NAD adduct on WT InhA is a 90° rotation of Phe149 side chain (Figure 2B), which provides room to accommodate the INH moiety. Owing to this rotation, a water molecule, which has been identified in the InhA structures without the isonicotinic-acyl group (Oliveira et al., 2006), was removed from the InhA active site (not shown in figure 2). Furthermore, it is observed alterations in the Met155, Leu218 and Ile215, Met199, probably due to alteration in the side chain position of Phe149. The Ile215 moves away from the active site due to rotation of the side chain of the Tyr158 (Figure 2B). The side chain of Met199 underwent a rotation of approximately 30° away from the active site. Two water molecules are present in the structure of InhA:INH-NAD complex, but absent from the structures between WT and S94A, I21V and I47T mutants InhA proteins in complex with NADH (Oliveira et al., 2006). The isonicotinic group also causes the expulsion of one water molecule next to C4 of NADH molecule. This causes a rotation of approximately 20° of nicotinamide group of NADH molecule, when compared with other InhA:NADH structures, but the phosphates, ribose and adenine are highly superposed among the analyzed structures.

This change results in a more snugly fit of the INH moiety of INH-NAD in the active site of InhA.

There is one intermolecular hydrogen bond that is conserved in the present structures and in the structure reported by Roswarski et al., (1998) involving oxygen atom from Gly14 and O2 from the pentose of NADH, which strongly indicates the importance of this hydrogen bond for the interaction of NADH and the InhA.

### **Influence of mutations I21V and S94A on INH-NAD binding to InhAs**

Van der Waals contacts between Val21 side chain and the nicotinamide ring, nicotinamide ribose, and phosphate oxygen atoms are missing in the INH-resistant I21V InhA structure as compared to WT-sensitive InhA. These interactions do not occur due to the absence of the CD1 atom in the mutated residue Val21 (Oliveira et al., 2006), which appear to play an important role in stabilizing bound NADH or INH-NAD in the InhA active site (Oliveira et al., 2006; Basso et al., 1998).

The crystal structure of S94A InhA showed that a conserved hydrogen bond to a water molecule is lost owing to mutation to Ala94 (Oliveira et al., 2006), which has been proposed to account for the reduction in affinity for NADH observed for the INH-resistant S94A InhA enzyme (Basso et al., 1998). Vilchère et al. (2006) have solved the structure of INH-resistant S94A InhA enzyme (PDB code 2NV6) and INH sensitive InhA in complex with INH-NAD (data do not deposited in the PDB). These authors propose the loss of the serine causes a shift in position of water molecule that promotes a disruption in hydrogen binding network. Analysis of the structure of S94A InhA enzyme present here does not present the disruption between this water molecule and the atom O9 of the molecule of INH-NAD. Analysis of other structures of S94A InhA, such a ones solved by Oliveira et al. (2006), shows a binding between this conserved

water molecule and the O9 of the molecule of INH-NAD. Thus, we believed that the mutation S94A can cause increase flexibility or decrease affinity of this molecule due the loss binding between this water molecule and the OG atom of the Ser94, which is observed in the structure wild type. The influence of the mutation on this water molecule could cause a reduction in the affinity of INH-NAD or NADH by InhA S94A. Thus, the movement observed in the water molecule by Vilchèze et al. (2006) can be due an increase of flexibility of this molecule carrying a false impression of an ordered movement.

The effect of the mutations I21V and S94A on binding process of the NADH on protein is summarized in the table 2. Analysis of contact area between the InhAs and INH-NAD reveals that the INH-sensitive WT enzyme presents larger contact area than the INH-resistant mutants. The value for WT enzyme is  $467.4\text{\AA}^2$  while for S94A and I21V mutants the values are  $464.2\text{\AA}^2$  and  $461.9\text{\AA}^2$ , respectively. The smaller value observed for I21V mutant can be due the absence of the contact between the CD1 of Ile21 that is missing in the I21V mutant. The I21V and S94A mutations do not seem to alter the position of the isonicotinic-acyl group because there are no significant changes when the mutant structures are compared with WT type structure (figure 3A-C). These results are in agreement with the proposal that INH resistance is due to reduction in NADH affinity for mutant InhAs (S94A and I21V) thereby hampering binding of INH-NAD adduct.

#### **Effect of INH-NAD on the binding of substrate in the InhA**

The isonicotinic-acyl group of INH-NAD adduct replaces the 4S hydrogen of C4 of NADH, which is the same position involved in the hydride transfer catalyzed by InhA. As a member of the family of SDR enzymes, the catalytic triad formed by

Phe149, Tyr158 and Lys165 (Rozwarski et al., 1999) is present in InhA (Figure 4). The Phe149 may play a role in anchoring the nicotinamide moiety of NADH for hydride transfer to the fatty acyl substrate. Lys165 interacts with the 3'-hydroxyl oxygen of the nicotinamide ribose of NADH, in agreement with a previous proposal that Lys165 plays a role in cofactor binding (Parikh et al., 1999).

The proposed mechanism of how InhA catalyzes the reduction of the 2-*trans* double bond of the fatty acyl substrate consists of the formation of an enolate intermediate through the direct transfer of a hydride ion from NADH to position C3 of the substrate, followed by protonation of position C2 (Rozwarski et al., 1999). It is noteworthy that the Tyr158 occupies approximately the same position for S94A:INH-NAD and apo S94A, whereas there is a conformational change upon binding of NADH to S94A, and NAD<sup>+</sup> and a C16 enoyl substrate to WT InhA resulting in position of Tyr158 to similar positions in these latter structures (figure 4). A comparison between the WT InhA-NADH structure and an inactive ternary complex formed by WT InhA, NAD<sup>+</sup>, and a C16 enoyl substrate revealed that upon binding the enoyl substrate there is a 60° rotation about the Tyr158 C $\alpha$ -C $\beta$  bond that enables the Tyr158 to hydrogen bond to the substrate carbonyl group (Parikh et al., 1999). It has thus been suggested that Tyr158 provides electrophilic stabilization of the transition state(s) for the reaction by hydrogen bonding to the carbonyl of the substrate. However, the InhA:NAD:C16 enoyl substrate is a non-productive ternary complex and is likely not part of the reaction course, which is in agreement with the structural results presented here for apo S94A InhA. The isonicotyl-acyl group causes a large change in the position of the Phe149 side chain (figure 4), and appears to be involved in stacking interactions with the INH-NAD adduct.

**INH-resistant S94A mutant in an uncomplexed form**

The structure of unliganded S94A InhA reveals that the protein undergoes three noticeable conformational changes in its main chain upon NADH binding. The main alterations occur between the residues 99 and 112, between the residues 197 and 213 (figure 5A), and in the Phe41 (figure 5B). It appears that NADH binding is sufficient to cause a closure of the active site of InhA. Analysis of B-factor for three different structures of InhA (figure 6) shows that in the absence of both substrate and cofactor or in the presence of ones, the protein presents the substrate binding loop (residues 196-219) more disordered than the structure in the presence of cofactor. By analysis of these three structure, we can observe that in the uncomplexed structure, part of the substrate binding loop is closed (197-203) and part is opened (205-214) for InhA:NAD<sup>+</sup>:substrate structure (figure 5A). In the S94A:NADH structure, part of the substrate binding loop is more open (197-203) whereas part is more closed (205-214). The conformational change of InhA:NAD<sup>+</sup>:substrate ternary complex is due to the geometry of the  $\alpha$ -helix of the longer substrate binding loops that could sterically hinder the loop from folding downward into the substrate binding cavity, leaving a larger opening for the fatty acyl substrate between residues 205-214 (Rozwarski et al., 1999). The residues 205-214 of the substrate binding loop moves approximately 6 Å away from the active site in the unliganded InhA, while the substrate of InhA:NAD<sup>+</sup>:substrate presents the same loops away approximately 4Å. The opening of the substrate binding loop for these structures are consistent with apoenzyme providing access to substrates to enter the enzyme active site and non-productive ternary complex providing access to solvent for product release. Analysis of molecular surface shows the effect of the substrate binding loop on the active site of InhA (figure 7).



In this manner, the binding of NADH in the active site of the protein can cause the partial closure of substrate binding loop however arranging a conformation for the binding of the substrate. In the presence of the co-factor and substrate the protein presents the active site newly opened for liberation of the product of the reaction. Thus, in according with this hypothesis the molecule of NADH is the main responsible by closure of the active site of InhA.

The van der Waals interactions between the Phe41 and NADH molecule may play in NADH binding to the enzyme active site, since Phe41 moves closer to the adenine moiety upon cofactor binding (figure 5B). Asp42 and Arg43 side chains also undergo conformational changes upon binding of NADH or both NADH and substrate (figure 5B). However, the absence of substrate does not seem to alter the oligomeric state of InhA because the tetrameric form in the asymmetric unit and the contact area among the monomers in the unliganded structure are similar to the other structures (InhA:NAD<sup>+</sup>:substrate ternary complex and InhA:NADH binary complex) (data not show). Thus the movement of substrate binding loop appears not to interfere with the protein oligomeric state.

## Conclusion

Here we present four structures of InhA from *M. tuberculosis*: three structures for complexes of InhA:INH-NAD and one in the unliganded form. The structures in complex with INH-NAD are: INH-sensitive WT and two INH-resistant mutants (I21V and S94A). Comparison between our WT InhA:INH-NAD structure and the structure previously determined by Roswarski et al., (1998) reveals that there are changes in important residues in the active site that were not previously observed. Moreover, the comparison of our structures in complex with INH-NAD shows that there is no large

influence of mutation in the binding of INH-NAD. The INH-NAD adduct has been shown to be a slow, tight-binding competitive inhibitor of WT and INH-resistant InhA mutant enzymes (Rawat et al., 2003). Interestingly, the kinetic and thermodynamic parameters for the interaction of isonicotinyl-NAD<sup>+</sup> adduct with INH-resistant I21V, I47T, and S94A InhA mutant enzymes were found to be similar to those of the WT enzyme (Rawat et al., 2003). These results prompted the authors to suggest an alternative hypothesis to explain for INH resistance mechanism in strains harboring *inhA*-structural gene mutations, in which InhA may interact directly with other components of the FAS-II system. Accordingly, the resistance-associated mutations in the *inhA*-structural gene would affect the susceptibility of InhA to INH inhibition only in the context of the multienzyme complex, and not when InhA is tested in isolation as in *in vitro* assays. Several protein-protein interactions between FAS-II enzymes have been detected by yeast two-hybrid and co-immunoprecipitation studies and proposed that either these complexes might coexist or the quaternary structure of a “unique” FAS-II might change from one composition to another during the time and according to the degree of elongation of the substrate (Veyron-Churlet et al., 2004). In particular, *M. tuberculosis* InhA was shown to interact with KasA ( $\beta$ -ketoacyl synthase A) and this protein-protein interaction has been suggested as a probable explanation to occurrence of INH-resistant mutant in KasA, even if InhA is indeed the only primary target of INH. However, it remains to be shown whether *inhA* structural gene mutations identified in INH-clinical isolates of *M. tuberculosis* will affect the inhibition of InhA by INH in the context of, for instance, InhA-KasA multienzyme complex. In agreement with InhA as the primary target for INH mode of action, recessive mutations in *M. smegmatis* and *M. bovis* BCG *ndh* gene, which codes for a type II NADH dehydrogenase (NdhII), have been found to increase intracellular NADH/NAD ratios (Vilchèze et al., 2005).

Increasing NADH levels protected InhA against inhibition by the INH-NAD adduct formed upon KatG activation of INH. Hence, mutations in mycobacterial *ndh* gene resulted in increased intracellular NADH concentrations, which competitively inhibits the binding of INH-NAD adduct to InhA, in agreement with the higher dissociation constant values for NADH found for INH-resistant clinical isolates harbouring *inhA*-structural gene mutations as compared to WT InhA (Basso et al., 1998). Moreover, it has been shown that subtle structural changes result in increased values for the limiting dissociation rate constant for NADH from INH-resistant mutants (Oliveira et al., 2006). These observations are in agreement with the results described here.

The unliganded structure presented here shows conformational changes that occur upon binding of substrates (NADH and/or fatty acyl substrate). The structure presents movements in the substrate binding loop and a striking change in the position of Phe41, which may play an important role in NADH stabilization in the enzyme active site. NADH or substrate binding causes no apparent change in the oligomeric state of InhA.

The results presented here provide structural information that should better our understanding of the mechanism of action of, and resistance to, isoniazid in *M. tuberculosis*. The unliganded structure of InhA allows identification of conformational changes upon ligand binding that are important for substrate anchoring. We hope that these results will help structure-based drug design of more potent antimycobacterial agents.

### **PDB Accession codes:**

**Protein Data Bank:** Atomic coordinates and structure factors have been deposited with accession codes: 2idz, 2ie0, 2ieb, 2ied.

**Acknowledgments:** Financial support for this work was provided by Millennium Initiative Program MCT-CNPq, Ministry of Health-Department of Science and Technology (DECIT)-UNESCO (Brazil), and PRONEX/CNPq/FAPERGS (Brazil) to D.S.S. and L.A.B. This work was also supported by grants from FAPESP (SMOLBNet, proc. 01/07532-0, 03/12472-2, 04/00217-0) to WFA. D.S.S. (304051/1975-06), L.A.B. (520182/99-5), and WFA (CNPq, 300851/98-7) are research career awardees from National Research Council of Brazil (CNPq).

## References

- Banerjee, A., Dubnau, E., Quémard, A., Balasubramanian, V., Um, K.S., Wilson, T., Collins, D., de Lisle, G., Jacobs, W.R. Jr., 1994. *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* 263, 227-230.
- Basso, L.A., Blanchard, J.S., 1998. Resistance to antitubercular drugs. *Adv. Exp. Med. Biol.* 456, 115-144.
- Basso, L.A., Santos, D.S., 2005. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis* – an update. *Med. Chem. Ver. – online* 2, 393-413.
- Basso, L.A., Zheng, R., Blanchard, J.S., 1996. Kinetics of inactivation of WT and C243S mutant of *Mycobacterium tuberculosis* enoyl reductase by activated isoniazid. *J. Am. Chem. Soc.* 118, 11301-11302.

Basso, L.A., Zheng, R., Musser, J.M., Jacobs, W.R. Jr., Blanchard, J.S., 1998. Mechanisms of isoniazid resistance in *Mycobacterium tuberculosis*: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. *J. Infect. Dis.* 178, 769–775.

Bernstein, J.W., Lott, A., Steinberg, B.A., Yale, H.L., 1952 Chemotherapy of experimental tuberculosis. V. Isonicotinic acid hydrazide (isoniazid) and related compounds. *Am. Rev. Tuberc.* 65, 357-374.

Blanchard, J.S., 1996. Molecular mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Ann. Rev. Biochem.* 65, 215-239.

CDC (Centers for Disease Control and Prevention), 2006. Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs – worldwide, 2000-2004. *Morb. Mortal Wkly. Rep.* 55, 301-305.

Collaborative Computational Project, Number 4., 1994. The CCP4 suite: programs for protein. *Acta Crystallog. Sect. D* 50, 760–763.

Corbett, E.L., Watt, C.J., Walker, N., Maher, D., Williams, B.G., Raviglione, M.C., Dye, C., 2003. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch. Intern. Med.* 163, 1009-1021.

DeLano, W.L., 2004. The PyMOL Molecular Graphics System. CA DeLano Scientific.

- Dessen, A., Quémard, A., Blanchard, J.S., Jacobs, W.R. Jr., Sacchettini, J.C., 1995  
Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*.  
*Science* 267, 1638–1641.
- Espinal, M.A., 2003. The global situation of MDR-TB. *Tuberculosis*: 83, 44-51.
- Glickman, M.S., Jacobs, W. R. Jr., 2001 Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. *Cell*, 104, 477-485.
- Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18, 2714-2723.
- Johnsson, K., King, D.S., Schultz, P.G., 1995. Studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis. *J. Am. Chem. Soc.* 117, 5009-5010.
- Johnsson, K., Schultz, P.G., 1994. Mechanistic studies of the oxidation of isoniazid by the catalase peroxidase from *Mycobacterium tuberculosis*. *J. Am. Chem. Soc.* 116, 7425-7426.
- Koradi, R., Billeter, M., Wüthrich, K., 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14 (1), 51-55.
- Kremer, L., Dover, L.G., Morbidoni, H.R., Vilchèze, C., Maughan, W.N., Baulard, A., Tu, S.C., Honore, N., Deretic, V., Sacchettini, J.C., Loch, C., Jacobs, W.R. Jr., Besra, G.S., 2003. Inhibition of InhA activity, but not KasA activity, induces

formation of a KasA-containing complex in mycobacteria. *J. Biol. Chem.* 278, 20547-20554.

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. Targeting tuberculosis and malaria through inhibition of enoyl reductase: Compound activity and structural data. *J. Biol. Chem.* 278(23), 20851-20859.

Larsen, M.H., Vilchèze, C., Kremer, L., Besra, G.S., Parsons, L., Salfinger, M., Heifets, L., Hazbon, M.H., Alland, D., Sacchettini, J.C., Jacobs, W.R. Jr., 2002. Overexpression of *inhA*, but not *kasA*, confers resistance to isoniazid and ethionamide in *Mycobacterium smegmatis*, *M. bovis* BCG and *M. tuberculosis*. *Mol. Microbiol.* 46, 453-466.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* 26, 283-291.

Lei, B., Wei, C.-J., Tu, S.-C., 2000. Action mechanism of antitubercular isoniazid. *J. Biol. Chem.* 275, 2520-2526.

McRee, D.E., 1999. XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* 125, 156-165.

- Middlebrook, G., Cohn, M.L., 1953. Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli. *Science* 118, 297-299.
- Mitchison, D.A., 1985. The action of antituberculosis drugs in short-course chemotherapy. *Tubercle* 66, 219-225.
- Morlock, G.P., Metchock, B., Sikes, D., Crawford, J.T., Cooksey, R.C., 2003. *ethA*, *inhA*, and *katG* loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. *Antimicrob. Agents Chemother.* 47, 3799–3805.
- Murshudov, G.N., Vagin. A.A., Dodson. E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D* 53, 240–255.
- Navaza, J., 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallog. Sect. A* 50. 157–163.
- No authors listed, 2006 Addressing the threat of tuberculosis caused by extensively drug-resistant *Mycobacterium tuberculosis*. *Wkly Epidemiol. Rec.* 21, 386-390.
- 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. 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), 646-66.



- Oliveira, J. S. Vasconcelos, I. B., Moreira, I. S., Santos, D. S., Basso, L. A., 2007. Enoyl reductases as targets for the development of anti-tubercular and antimalarial agents. *Curr. Drug Targets* 8(3), 399-411.
- Ormerod, L.P., 2005. Multidrug-resistant tuberculosis (MDR-TB): epidemiology, prevention and treatment. *British Medical Bulletin* 73-74(1), 17-24.
- Pablos-Mendez, A., Gowda, D.K., Frieden, T.R., 2002. Controlling multidrug-resistant tuberculosis and access to expensive drugs: a rational framework. *Bull World Health Organ* 80, 489-495.
- Parikh, S., Moynihan, D.P., Xiao, G., Tonge, P.J., 1999. Roles of tyrosine 158 and lysine 165 in the catalytic mechanism of InhA, the enoyl-ACP reductase from *Mycobacterium tuberculosis*. *Biochemistry* 38, 13623-13634.
- Polikarpov, I., Perles, L.A., de Oliveira, R.T., Oliva, G., Castellano, E.E., Garratt, R.C., Craievich, A., 1998. Set-up and experimental parameters of the protein crystallography beam line at the Brazilian National Synchrotron Laboratory. *J. Synchrotron Radiat.* 5, 72-76.
- Quémard, A., Sacchetti, J.C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W.R. Jr., Blanchard, J.S., 1995. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochemistry* 34, 8235-8241.

- Rawat, R., Whitty, A., Tonge, P., 2003. The isoniazid–NAD adduct is a slow, tight-binding inhibitor of InhA, the *Mycobacterium tuberculosis* enoyl reductase: adduct affinity and drug resistance. *Proc. Natl. Acad. Sci. USA* 100, 13881–13886.
- Rozwarski, D.A., Grant, G.A., Barton, D.H.R., Jacobs, W.R., Sacchettini, J.C., 1998. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. *Science* 279, 98-102.
- Rozwarski, D.A., Vilchèze, C., Sugantino, M., Bittman, R., Sacchettini, J.C., 1999. Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD<sup>+</sup> and a C16 fatty acyl substrate. *The Journal of Biological Chemistry* 274(22), 15582-15589.
- Schroeder, E.K., de Souza, O.N., Santos, D.S., Blanchard, J.S., Basso, L.A., 2002. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. *Curr. Pharm. Biotechnol.* 3, 197–225.
- Uchoa, H.B.; Jorge, G.E.; Freitas Da Silveira, N.J.; Camera, J.C., Jr.; Canduri, F.; De Azevedo, W.F., Jr. *Biochem. Biophys. Res. Commun.*, 2004, 325, 1481-1486.
- Veyron-Churlet, R., Guerrini, O., Mourey, L., Daffe, M., Zerbib, D., 2004. Protein-protein interactions within the fatty acid synthase-II system of *Mycobacterium tuberculosis* are essential for mycobacterial viability. *Mol. Microbiol.* 54, 1161-1172.

- Vilchèze, C., Morbidoni, H.R., Weisbrod, T.R., Iwamoto, H., Kuo, M., Sacchetti, J.C., Jacobs, W.R. Jr., 2000. Inactivation of the *inhA*-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASII end products and cell lysis of *Mycobacterium smegmatis*. *J. Bacteriol.* 182, 4059-4067.
- Vilchèze, C., Wang, F., Arai, M., Hazbon, M.H., Colangeli, R., Kremer, L., Weisbrod, T.R., Alland, D., Sacchetti, J.C., Jacobs, W.R. Jr., 2006. Transfer of a point mutation in *Mycobacterium tuberculosis inhA* resolves the target of isoniazid. *Nat Med*: 12, 1027-1029.
- Vilchèze, C., Weisbrod, T.R., Chen, B., Kremer, L., Hazbon, M.H., Wang, F., Alland, D., Sacchetti, J.C., Jacobs, W.R. Jr., 2005. Altered NADH/NAD<sup>+</sup> ratio mediates coresistance to isoniazid and ethionamide in mycobacteria. *Antimicrob. Agents Chemother.* 49, 708-720.
- Zabinski, R.F., Blanchard, J.S., 1997. The requirement for manganese and oxygen in the isoniazid-dependent inactivation of *Mycobacterium tuberculosis* enoyl reductase. *J. Am. Soc. Chem.* 119, 2331-2332.

## Figures

**Figure 1.** Ribbon diagram of InhA. A) Monomeric structure of the wild type InhA:INH-NAD binary complex, and B) Tetrameric structure of the unliganded InhA. The figures were drawn using program MolMol (Koradi et al., 1996).

**Figure 2.** Superposition of active site residues of InhA. A) It is shown the superposition of the active site structure of wild type InhA: INH-NAD solved by Roswarski et al., (1998) at 2.7Å resolution (in green) and wild type InhA: INH-NAD at 2.0 Å resolution presented here (in grey). B) It is shown the superposition of the structure of wild type InhA:NADH complex binary solved by Oliveira et al., (2006) (in blue) and the structure of wild type InhA: INH-NAD presented here (in grey). The figures were generated with MolMol (Koradi et al., 1996).

**Figure 3.** Ribbon diagrams for InhA structures, showing the electronic density for INH-NAD and mutated residue. A) Wild type InhA, B) I21V mutant and C) S94A mutant. The figures were generated using program SpdbView (Guex and Peitsch, 1997).

**Figure 4.** Catalytic triad for InhA. In green is shown the residues of S94A InhA in complex with INH-NAD, in grey is shown the residues of S94A InhA unliganded, in orange is shown the residues of S94A in complex with NADH, and in blue is shown the residues of WT InhA in complex with C16 enoyl substrate and  $\text{NAD}^+$  (Rozwarski et al., 1999). The Figure shows conformational changes in these residues due to different ligands. The figure was generated using program Molmol (Koradi et al., 1996).

**Figure 5.** Conformational changes observed in the structure of S94A InhA upon binding of NADH or both NADH and substrate. A)  $\text{C}\alpha$  traces of S94A InhA unliganded (purple), S94AInhA in complex with NADH (orange), and WT InhA in complex with both  $\text{NAD}^+$  and substrate (green). B) Conformational changes of residues 41-43 (coloring was the same as in A). The figures were generated using program MolMol (Koradi et al., 1996).

**Figure 6.** Residue-averaged B factor for S94A resistant mutant apoenzyme (red line), WT InhA in complex with  $\text{NAD}^+$  and C16 substrate (black line), and S94A resistant mutant in complex with NADH (purple line).

**Figure 7.** Molecular surface for A) Unliganded S94A InhA, B) S94A InhA in complex with NADH, and C) WT InhA in complex with both NADH and substrate. The figures were generated using program Pymol (DeLano, 2004).



Ramachandran plot				
Favorable	86.2	87.1	88.9	88
Additional allowed	9.8	9.8	9.3	10.3
Generously allowed	0.0	0.0	0.0	0.7
Disallowed	4.0	3.1	1.8	1.0

<sup>a</sup>  $R_{\text{sym}} = 100 \frac{\sum |I(h) - \langle I(h) \rangle|}{\sum I(h)}$ , observed intensity and  $\langle I(h) \rangle$ , mean intensity of reflection  $h$  overall measurement of  $I(h)$ .

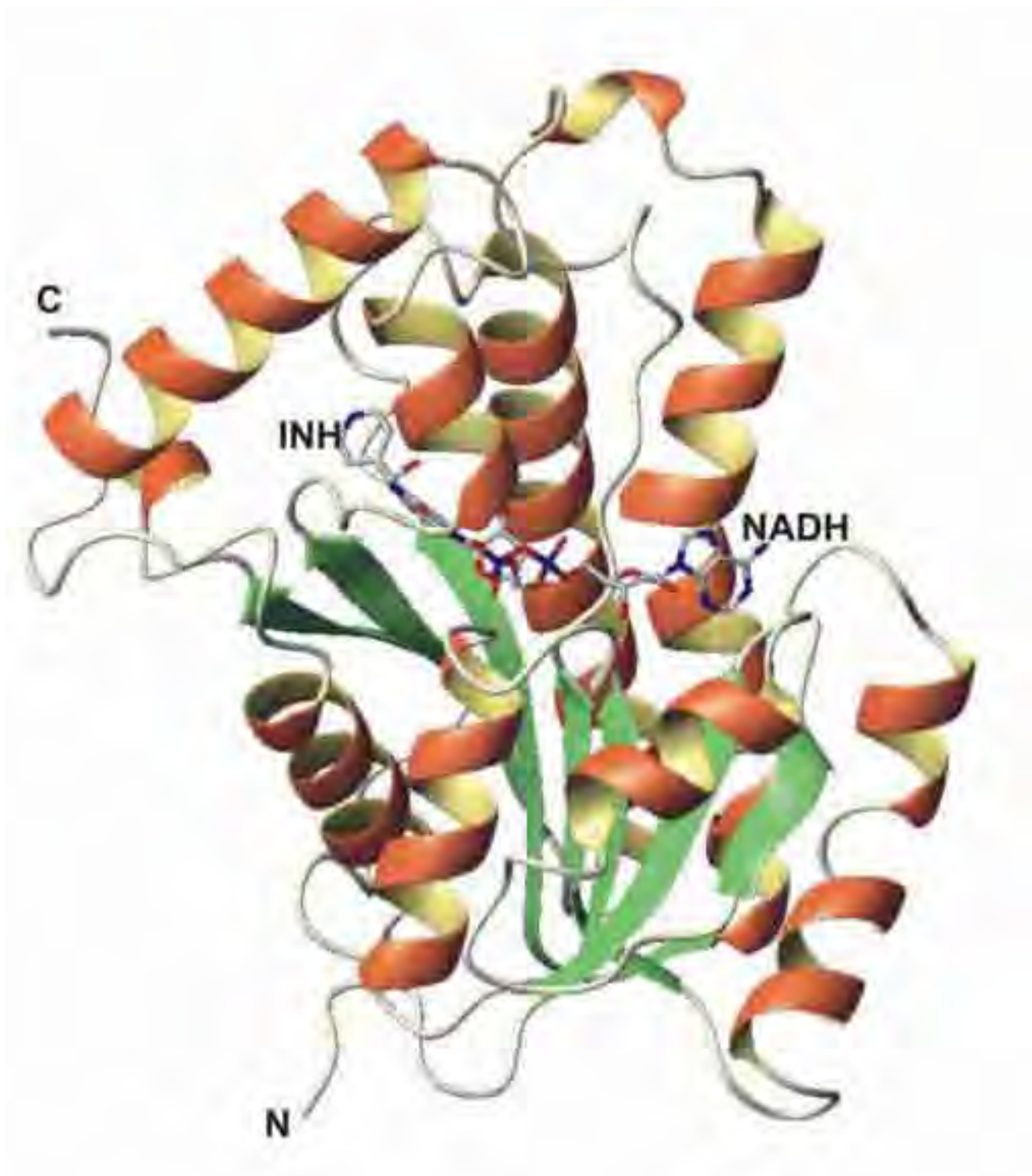
<sup>b</sup>  $R\text{-factor} = 100 \times \frac{\sum (F_{\text{obs}} - F_{\text{calc}})}{\sum F_{\text{obs}}}$ , the sums being taken over all reflections with  $F/\sigma(F) > 2\sigma(F)$ .

<sup>c</sup>  $R\text{-free} = R\text{-factor}$  for 10% of the data that were not included during crystallographic refinement.

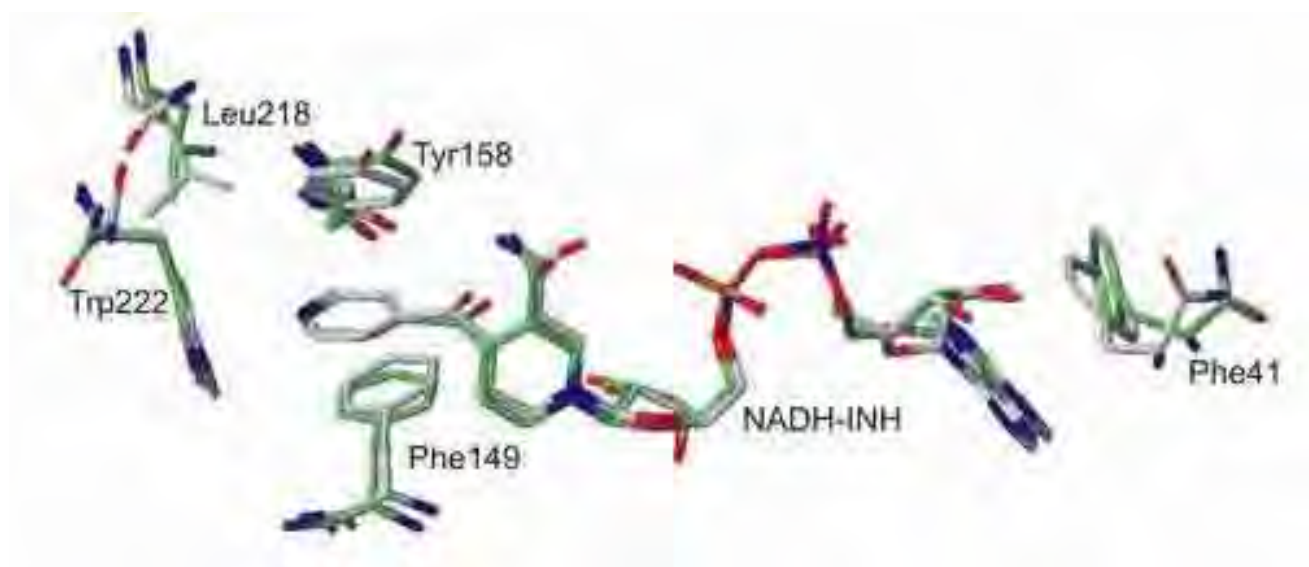
Table 2. Effect of the mutations I21V and S94A on the binding process of the NADH on the InhA.

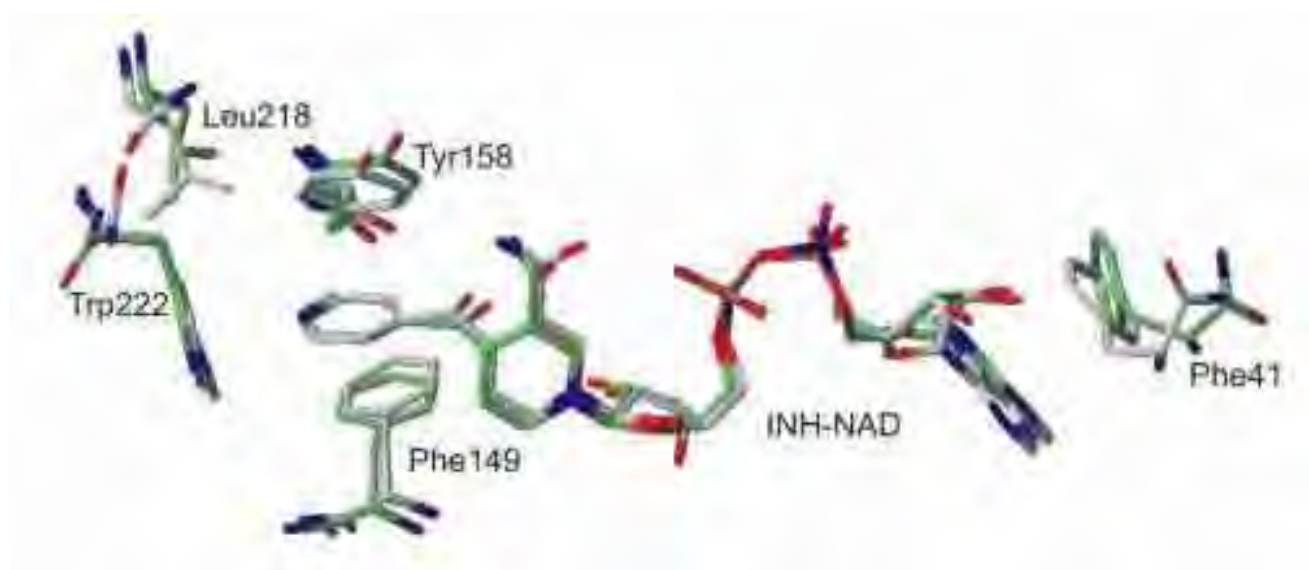
mutation	Mutation effect	consequence
I21V	Lose of Van der Waals interaction between NADH and the CD1 atom present in the valine residue	Decrease of the stability of the binding of the NADH in the active site of the protein
S94A	Alteration in the binding network involving a conserved water molecule and O9 atom of molecule of NADH	Increase of the flexibility of the conserved water molecule and decrease of the affinity of NADH by protein

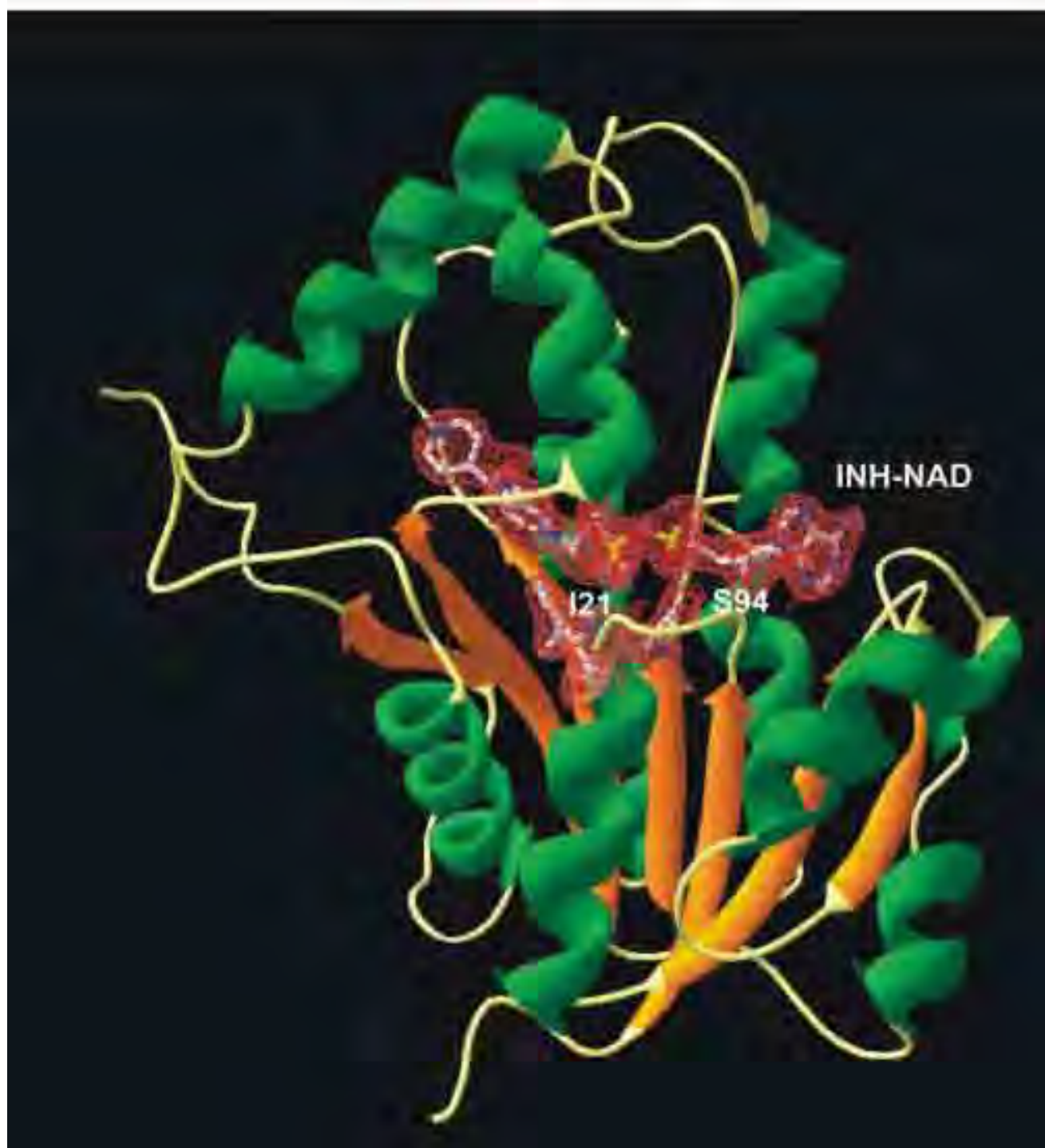




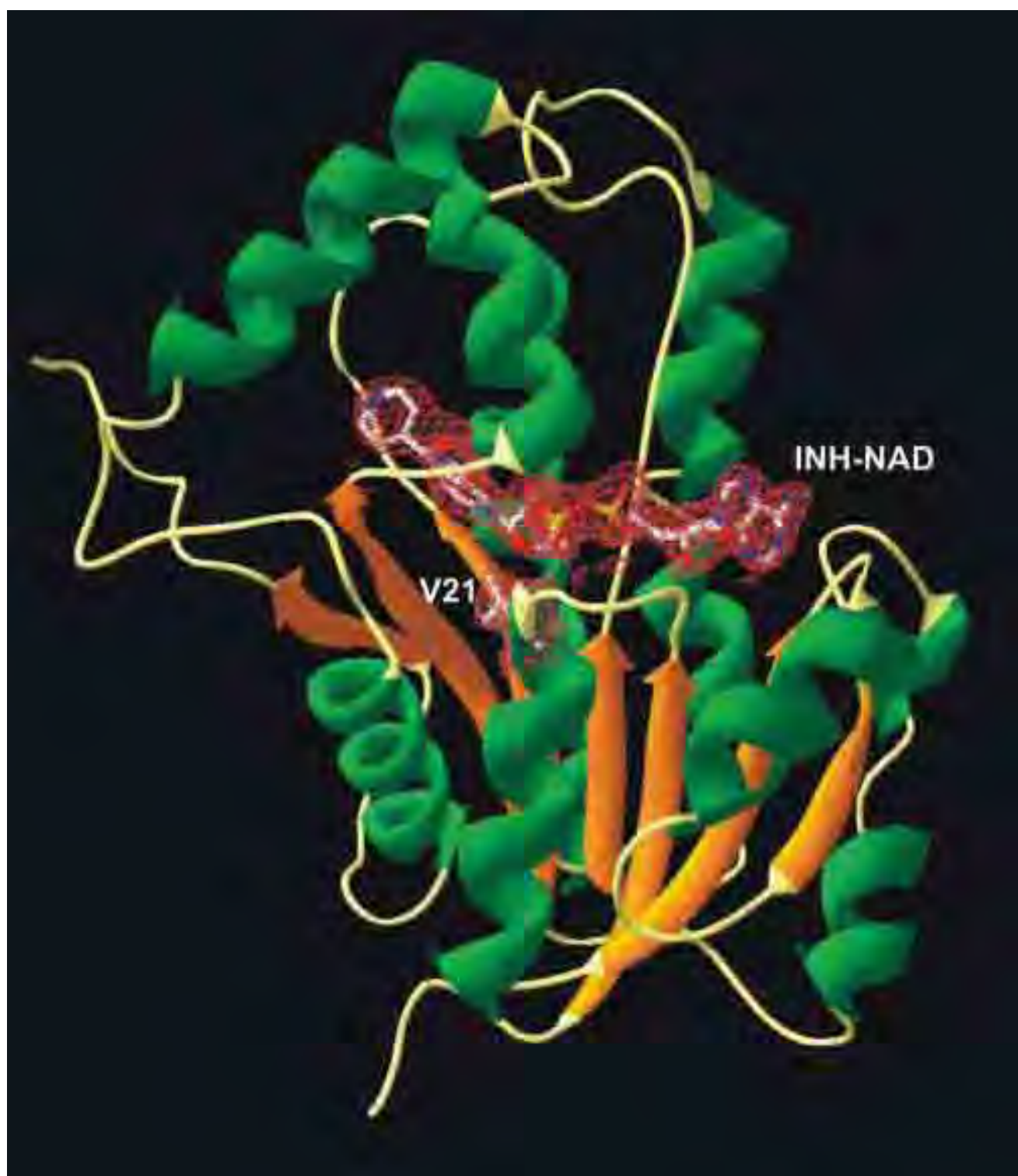


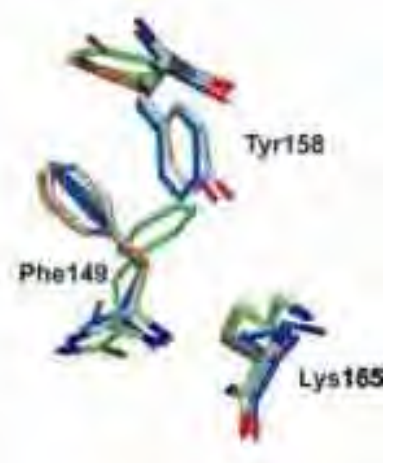




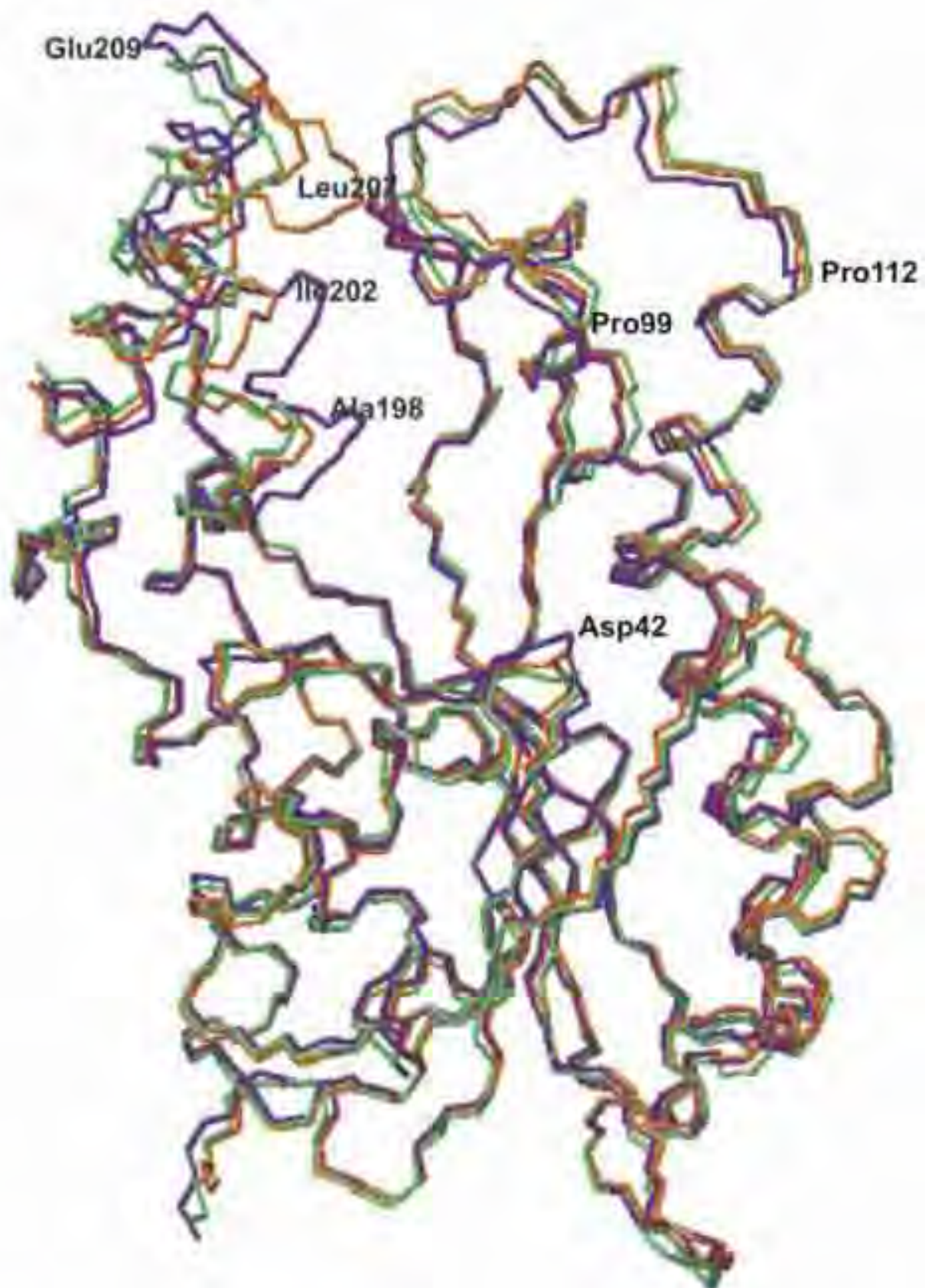


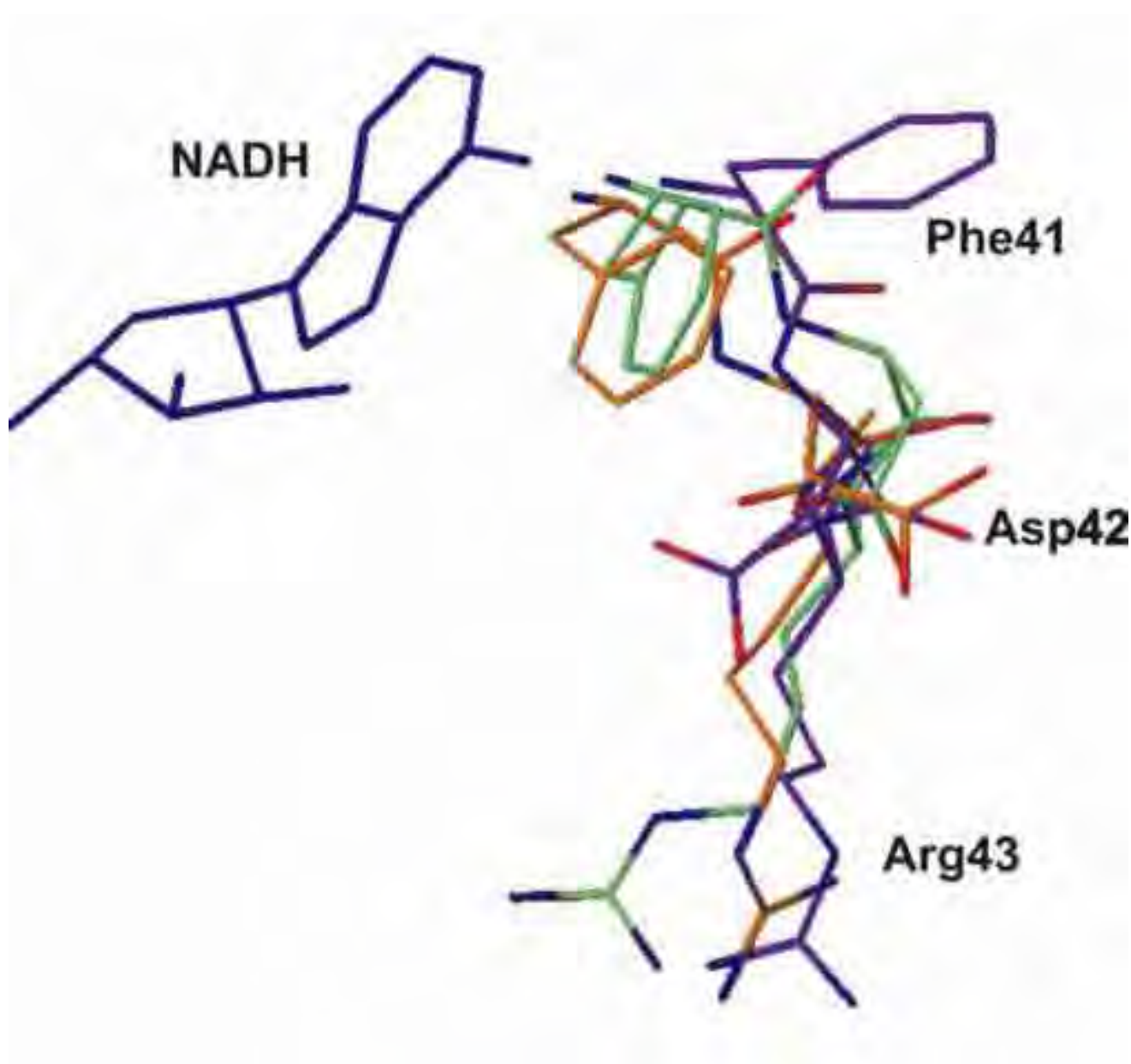




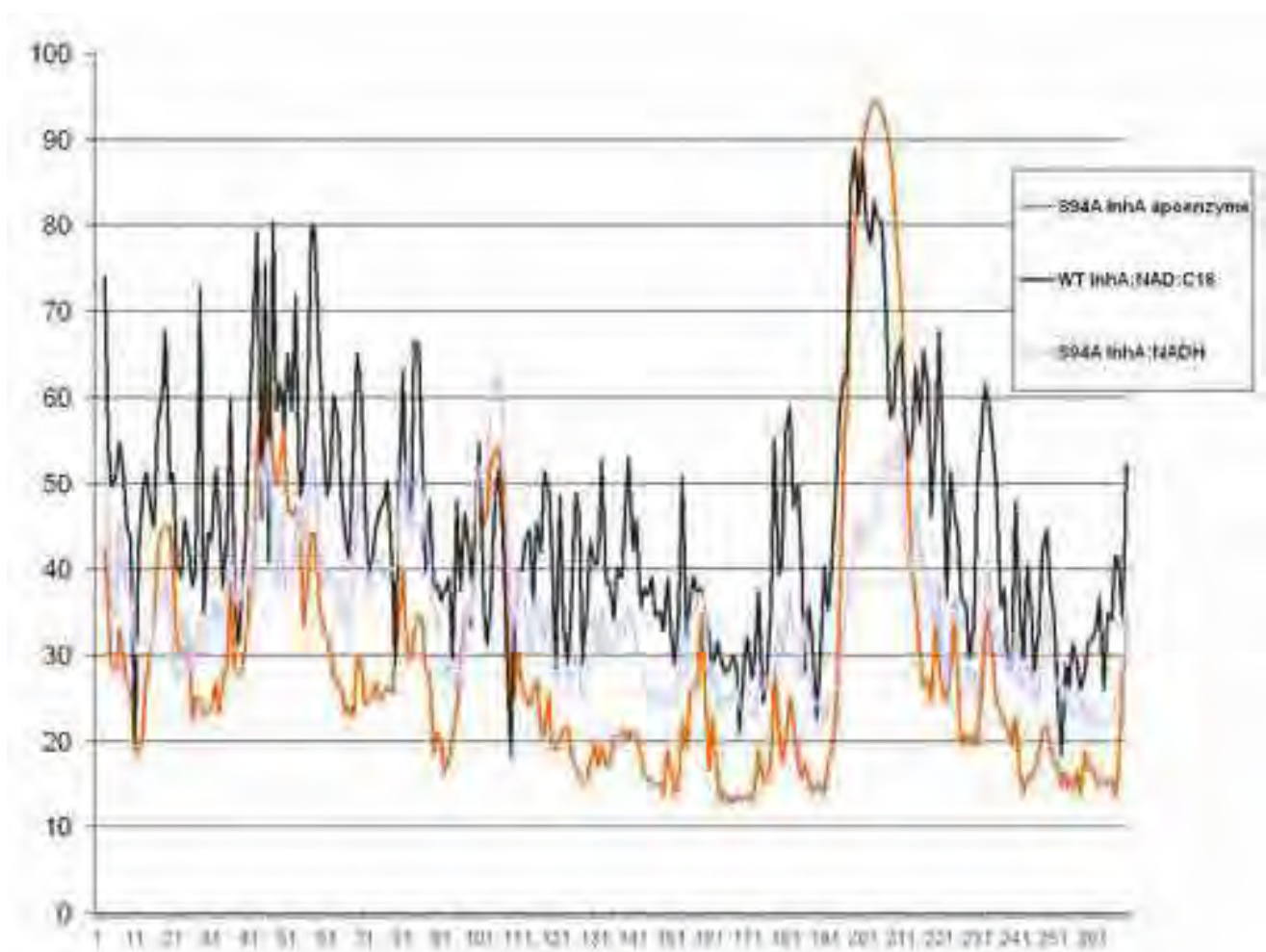


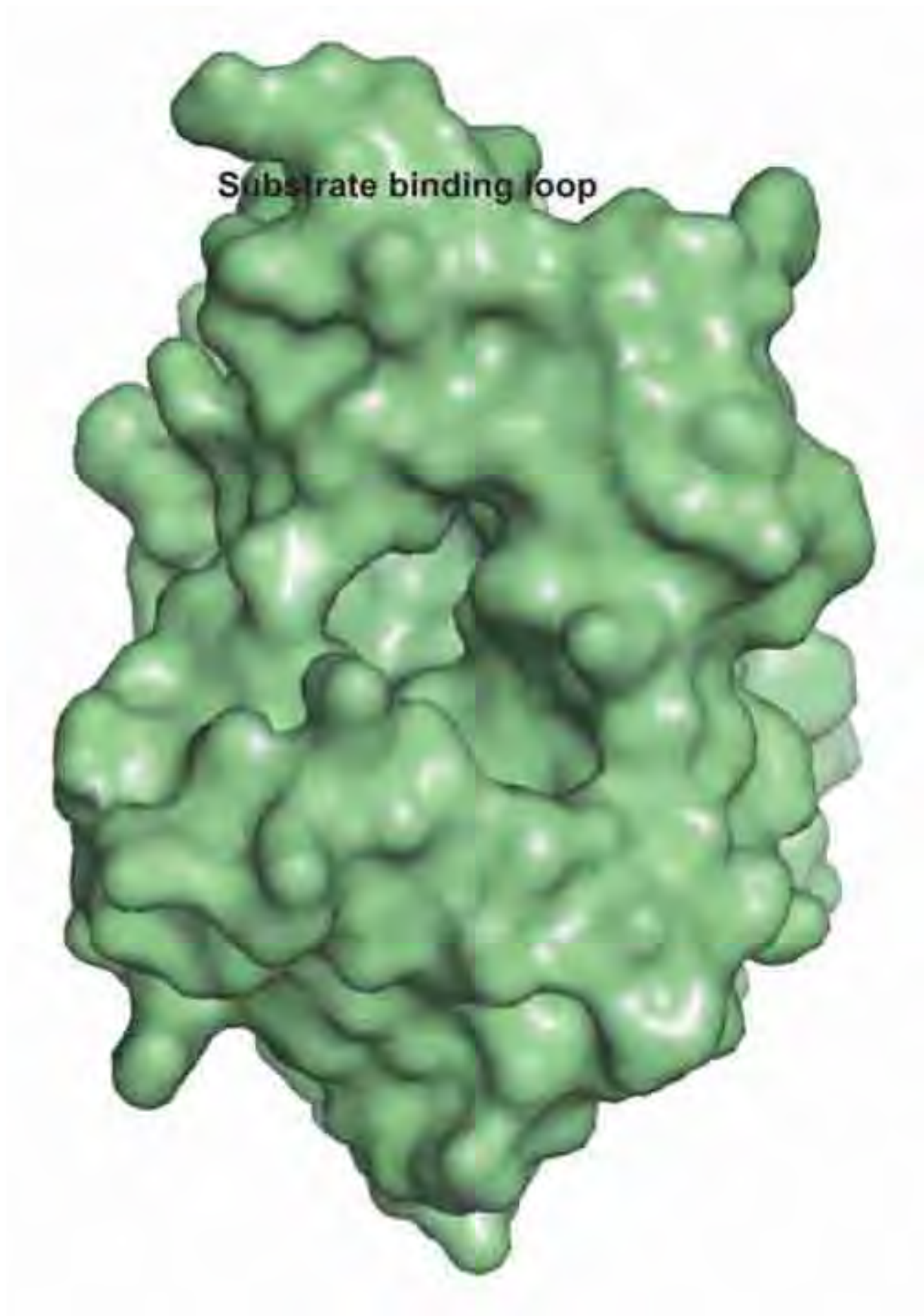


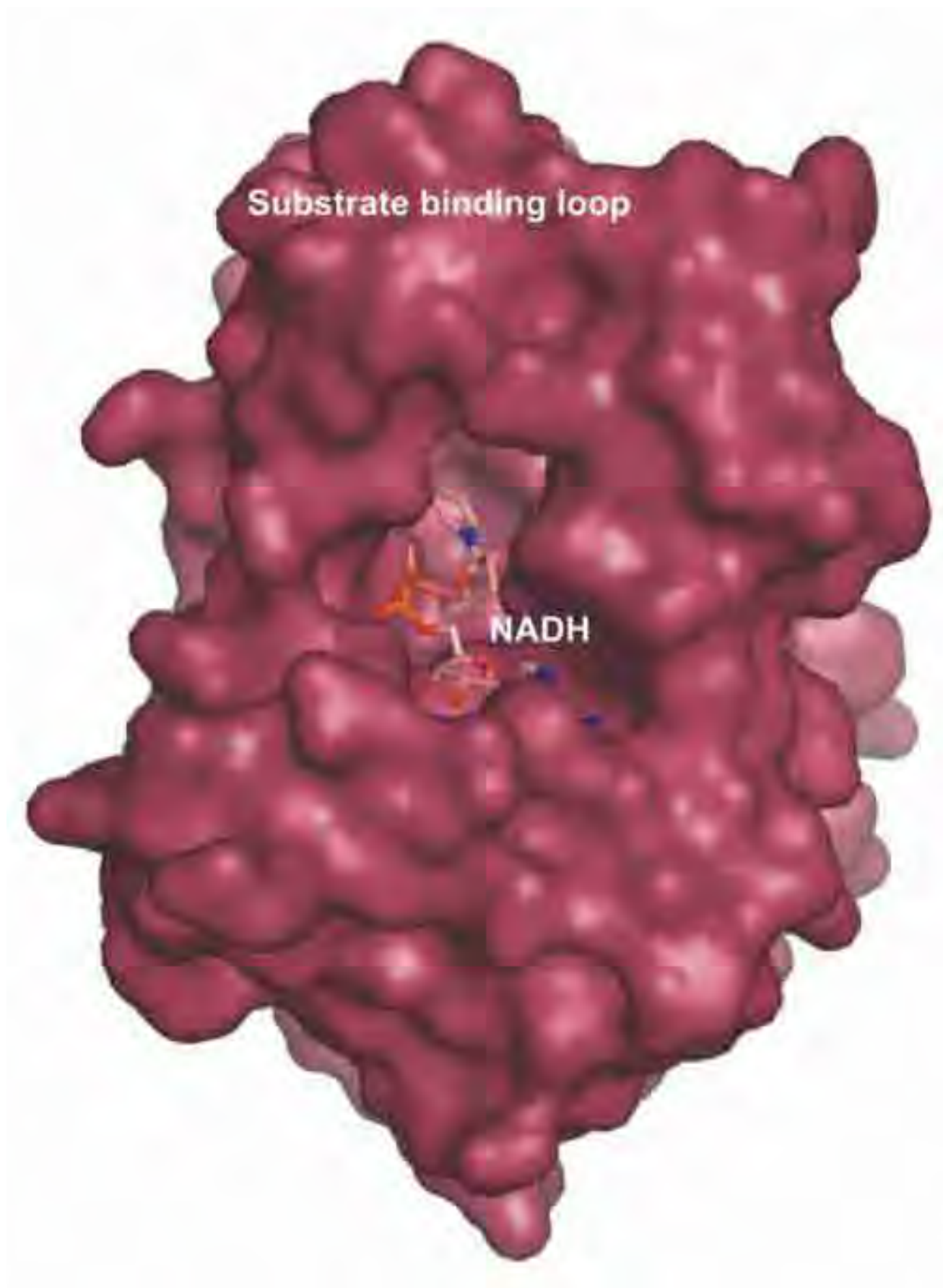


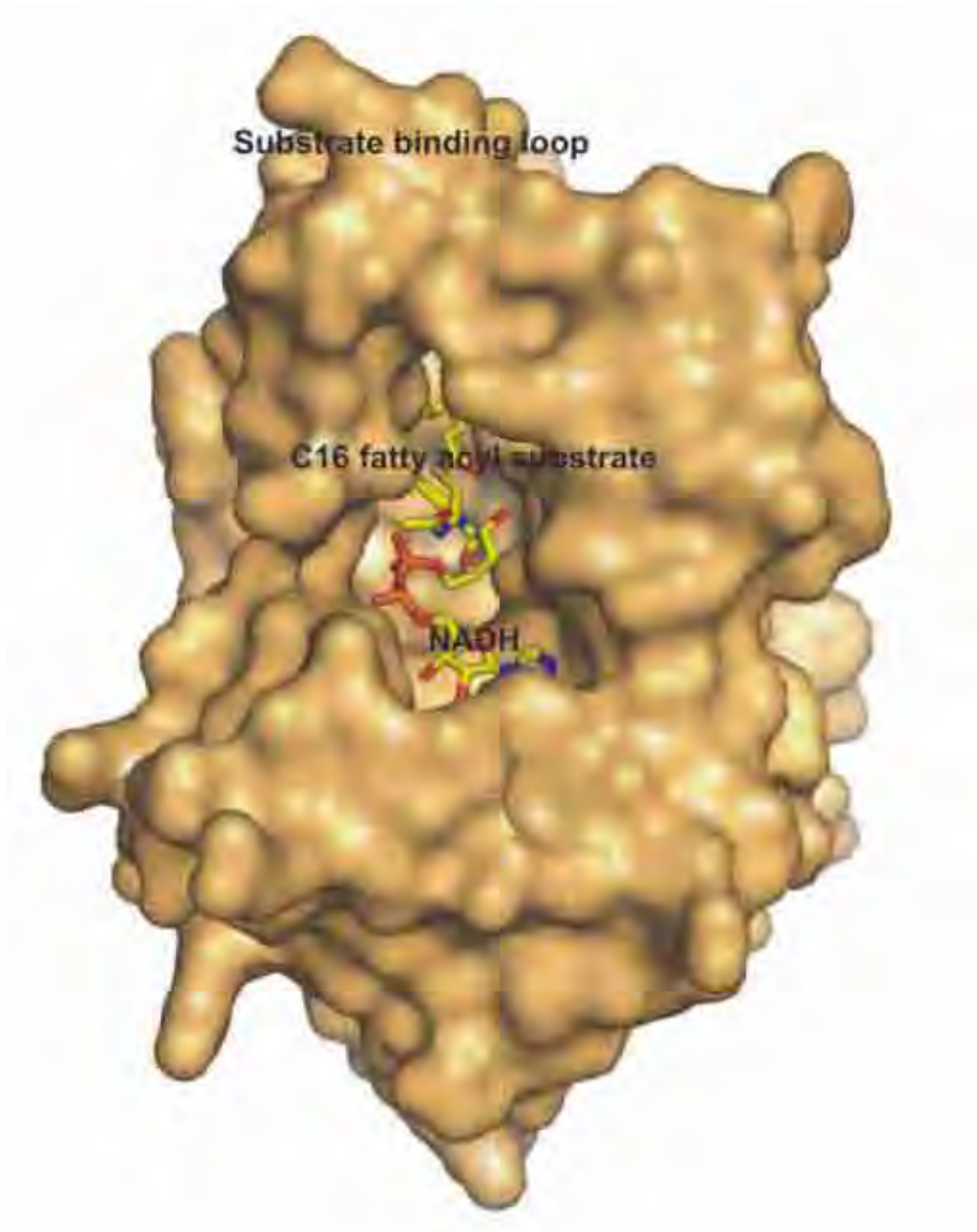


[Click here to download high resolution image](#)









## 4. CONCLUSÃO

Este trabalho teve como principal objetivo análise estrutural de enzimas que são importantes para a viabilidade de *M. tuberculosis*, mas que, entretanto tais enzimas estivessem ausentes em humanos. Assim essas enzimas são consideradas como alvos para o desenvolvimento de novas drogas contra a tuberculose e podem apresentar baixa toxicidade ou ausência de efeitos colaterais. Após os estudos realizados com cada uma das enzimas chegamos as seguintes conclusões para cada uma delas separadamente:

- Foi determinada a primeira condição de cristalização para a proteína corismato sintase de *M. tuberculosis*, na qual foram obtidos cristais hexagonais, grupo espacial P6<sub>4</sub>22. Estes cristais difrataram em torno de 2,65Å de resolução e os dados do processamento de difração de raios X foram utilizados para a resolução da estrutura cristalográfica desta proteína;
- Foi determinada a estrutura cristalográfica da corismato sintase de *Mycobacterium tuberculosis* em seu estado nativo a 2,65Å de resolução. A estrutura apresenta um *fold* bastante similar às outras estruturas de corismato sintase de outros microorganismos. Os resíduos do sítio ativo também são bastante conservados, entretanto a posição de alguns e resíduos apresenta diferentes conformações quando comparado com outras estruturas. Essas diferenças podem ser devido à ausência de ligantes. A determinação de estruturas da corismato sintase de *M. tuberculosis* em complexo com seus substratos podem esclarecer essas evidências;
- Foram construídos seis modelos estruturais para a proteína triptofano sintase de *M. tuberculosis* com diferentes inibidores análogos de substratos através da técnica de modelagem molecular comparativa. Os modelos obtidos apresentam qualidade estereoquímica satisfatória e mostram as interações entre estes ligantes e a enzima.

Pela análise da ligação podemos concluir que os inibidores estudados apresentam uma interação satisfatória com a enzima. Testes *in vitro* ou *in vivo* poderão confirmar essa hipótese. Entretanto os modelos obtidos podem ser utilizados na construção de outros inibidores mais potentes para esta enzima.

- Foram apresentadas também duas estruturas da proteína chiquimato quinase de *Mycobacterium tuberculosis* (uma em complexo com ADP e chiquimato e outra em complexo com MgADP). A estrutura da chiquimato quinase em complexo com ADP e chiquimato mostra que a ausência do íon magnésio pode influenciar a posição dos grupos hidroxilas da molécula de chiquimato e também em alguns importantes resíduos do sítio ativo da enzima, como o Asp34. A estrutura da chiquimato quinase em complexo com MgADP foi obtida em uma nova condição de cristalização e apresenta-se como um tetrâmero na unidade assimétrica, com simetria 222. Nesta estrutura pôde-se observar o possível efeito do chiquimato sobre o empacotamento cristalino da enzima. Além disso, devido à disposição do monômero na unidade assimétrica foi possível observar que a molécula ADP tem uma importante participação na estabilidade da estrutura quaternária da proteína, pelo menos na rede cristalina. Um dos monômeros do tetrâmero observado para a chiquimato quinase não apresentava o íon cloreto e assim é possível que ele tenha uma importante participação da orientação da molécula de ADP no sítio ativo da enzima, no fechamento do *LID domain* e também na formação de interações com moléculas de águas que podem levar o fechamento da enzima.. Desta forma, este íon pode ter importante influência na estabilidade e no mecanismo de catálise da enzima. Assim, as informações obtidas neste trabalho podem ser úteis para o entendimento do mecanismo catalítico desta enzima e também para o desenho de novos ligantes que tenham atividade inibitória.



- Foram apresentadas as estruturas cristalográficas da proteína Inha selvagem e de dois mutantes isolados clínicos resistentes a isoniazida (I21V e S94A) complexada com o aducto NAD-INH. A partir destas estruturas é possível determinar as alterações moleculares que a molécula de NAD-INH realiza no sítio ativo da enzima e também como ocorre o efeito inibitório desta droga. Foi observada nas estruturas dos mutantes em complexo com NAD-INH que, as mutações I21V e S94A parecem não influenciar na interação entre a proteína e o aducto. E assim, estas informações corroboram com os dados existentes na literatura de que a resistência pode estar relacionada com uma menor afinidade pela molécula de NADH. Entretanto a resistência a isoniazida poderia estar associada também a outros fatores nas quais estas mutações poderiam de forma indireta contribuir. É apresentado também neste trabalho a estrutura do mutante de InhA S94A na sua forma nativa, em um novo grupo espacial ainda não descrito na literatura para esta enzima (P1). Nesta estrutura puderam-se concluir quais são os possíveis movimentos que podem ocorrer durante o processo de ligação do NADH e do substrato acil C16. Além disso, foi possível evidenciar importantes resíduos que podem contribuir para a estabilidade da ligação entre o substrato e a enzima e importantes movimentos que ocorrem em alguns resíduos do sítio ativo durante o processo de catálise. Esperamos que estas informações contribuam para o entendimento do mecanismo de inibição e de resistência à isoniazida e que possam ser úteis na construção de novas drogas contra tuberculose.

## 5. BIBLIOGRAFIA

AHN, H. J. et al. Crystal structure of chorismate synthase: a novel FMN-binding protein fold and functional insights. **J. Mol. Biol.** v. 336(4), p 903-915, 2004.

BAPTISTA, I. M. et al. Drug resistance and genotypes of strains of *Mycobacterium tuberculosis* isolated from human immunodeficiency virus infected and non-infected tuberculosis patients in Bauru, São Paulo, Brazil. **Mem. Inst. Oswaldo Cruz**, v. 97(8), p. 1147-1152, 2002.

BASSO, L. A. et al. Mechanisms of isoniazid resistance in *Mycobacterium tuberculosis*: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. **J. Infect. Dis.**, v. 178, p. 769–775, 1998.

BAUBY, H.; GIRONS, I. S.; PICARDEAU, M. Construction and complementation of the first auxotrophic mutant in the spirochaete *Leptospira meyeri*. **Microbiology**, v. 149, p. 689–693, 2003.

BANNER, D. W. et al. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. **Nature**, v. 255, p. 609-614, 1975.

BENTLEY, R. The shikimate pathway - metabolic tree with many branches. **Crit. Rev. Biochem. Mol. Biol.**, v. 25, p. 307-384, 1990.

BRENNAN, P. J.; NIKAIDO, P. J. The envelope of mycobacteria. **Annu. Rev. Biochem.**, v. 64, p. 29-63, 1995.

CERASOLI, E. et al. Effects of salts on the function and conformational stability of shikimate kinase. **Biochem. Biophys. Acta**, v. 1648(1-2), p. 43-54, 2003.

COLE, S. T. et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. **Nature**, v. 393, p.537-544, 1998.

COTRAN, R. S.; KUMAR, V.; ROBBINS, S. L. **Robbins: Patologia estrutural e funcional**. Tradução de Amaury José da Cruz Júnior et al. 4.ed. Rio de Janeiro, Guanabara Koogan, 1991.

DESSEN, A. et al. Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. **Science**, v. 267, p. 1638–1641, 1995.

DIAS, M. V. B. et al. Crystal structure of chorismate synthase from *Mycobacterium tuberculosis*. **Journal of Structural Biology**, v. 154, p. 130-143, 2006.

DHALIWAL, B. et al. Crystallographic studies of shikimate binding and induced conformational changes in *Mycobacterium tuberculosis* shikimate kinase. **FEBS Lett.** V. 574(1-3), p. 49-54, 2004.

- DOSSELAERE, F.; VANDERLEYDEN, J.A. A metabolic node in action: chorismate-utilizing enzymes in microorganisms. **Crit. Rev. Microbiol.**, v.27, p.75-131, 2001.
- DOVER, L. G. et al. Comparative cell wall core biosynthesis in the mycolated pathogens, *Mycobacterium tuberculosis* and *Corynebacterium diphtheriae*. **FEMS Microbiology Reviews**, v. 28(2), p. 225-250, 2004.
- DUNCAN, K. Progress in TB drug development and what is still needed. **Tuberculosis**, v. 83, p. 201-207, 2003.
- DYE, C. et al. Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis **Science**, v. 295, p. 2042-2046, 2002.
- FINN, J. et al. Rational herbicide design by inhibition of tryptophan biosynthesis. **Bioorg. Med. Chem. Letters**, v. 9, p. 2297-2302, 1999.
- GERSTEIN, M.; SCHULZ, G.; CHOTHIA C. Domain closure in adenylate kinase. Joints on either side of two helices close like neighboring fingers. **J. Mol. Biol.** v. 229(2), p. 494-501, 1993.
- GAN, J. et al. Crystal structure of *Mycobacterium tuberculosis* shikimate kinase in complex with shikimate acid and an ATP analogue. **Biochemistry**, v. 45, p. 8539-8545, 2006.
- GU, Y. et al. Crystal structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the dynamic role of the LID domain in catalysis. **J. Mol. Biol.** v. 319, p. 779-789, 2002.
- HARTMANN, M. D. et al. Mechanism of phosphoryl transfer catalyzed by shikimate kinase from *Mycobacterium tuberculosis*. **J. Mol. Biol.**, 2006 (*in press*).
- HAWSER, S.; LOCIURO, S.; ISLAM, K. Dihydrofolate reductase inhibitors as antibacterial agents. **Biochemical Pharmacology**, v. 71(7), p. 941-948, 2006.
- HERRMANN, K. M. The shikimate pathway: early steps in the biosynthesis of aromatic compounds. **The Plant Cell**, v. 7, p. 907-919, 1995.
- HERRMANN, K. M.; WEAVER, L. M. The shikimate pathway. **Annu Rev Plant Physiol Plant Mol Biol.** v. 50, p. 473-503, 1999.
- HYDE, C. C. et al., Three-dimensional structure of the tryptophan synthase  $\alpha_2\beta_2$  multienzyme complex from *Salmonella typhimurium*. **J. Biol. Chem.** v. 263(33), p. 17857-17871, 1988.
- KHASNOBIS, S.; ESCUYER, V.; CHATTERJEE, D. Emerging therapeutic targets in tuberculosis: post-genomic era. **Expert. Opin. Ther. Targets**, v. 6, p. 21-40, 2002.
- KING, D. A. et al. Epidemiology. Infectious diseases: preparing for the future. **Science**, v. 313 (5792), p. 1392-1393, 2006.
- KITZING, K. et al., Mechanism of chorismate synthase. Role of the two invariant histidine residues in the active site. **J. Biol. Chem.** v. 279(10), p. 9451-9461, 2004.

- KITZING, K.; MACHEROUX, P.; AMRHEIN, N. Spectroscopic and kinetic characterization of the bifunctional chorismate synthase from *Neurospora crassa*: evidence for a common binding site for 5-enolpyruvylshikimate 3-phosphate and NADPH. **J. Biol. Chem.**, v. 276, p. 42658-42666, 2001.
- KUO, M. R. et al. Targeting tuberculosis and malaria through inhibition of enoyl reductase: Compound activity and structural data. **The journal of Biological Chemistry**, v. 278(23), p. 20851-20859, 2003.
- KRELL, T.; COGGINS, J. R.; LAPHORN, A. J. The three-dimensional structure of shikimate kinase. **J. Mol. Biol.** v. 278(5), p. 983-997, 1998.
- KUMAR, V.; ABBAS, A. K.; FASTO, N. **Robbins e Cotran: patologia – bases patológicas das doenças**. Tradução de Maria da Conceição Zacharias et al. Rio de Janeiro: Elsevier, 2005. 1592p.
- KURADI, R.; BILLETER, M.; WÜTHRICH, K. MOLMOL: a program for display and analysis of macromolecular structures. **J. Mol. Graphics** v. 14, p. 51-55, 1996.
- MACLEAN J.; ALI S. The structure of chorismate synthase reveals a novel flavin binding site fundamental to a unique chemical reaction. **Structure** v. 11(12), p. 1499-511, 2003.
- MACHEROUX, P. et al. A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway. **Planta**, v. 207(3), 325-334, 1999.
- McARTHUR J. D. et al. An aromatic amino acid auxotrophic mutant of *Bordetella bronchiseptica* is attenuated and immunogenic in a mouse model of infection. **FEMS Microbiology Letters**, v. 221, p. 7-16, 2003.
- McREE, D. E. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. **J. Struct. Biol.**, v. 125, p. 156-165, 1999.
- MILES, E. W.; RHEE, S.; DAVIES, D. V. The molecular basis of substrate channeling. **The journal of biological chemistry**, v. 274 (18), p. 12193-12196, 1999.
- MILES, E. W. Tryptophan synthase. Structure, function, and protein engineering. **Subcell Biochem.** v. 24, p. 207-254, 1995.
- MOLLE, V. et al. The condensing activities of the *Mycobacterium tuberculosis* type II fatty acid synthase are differentially regulated by phosphorylation. **J. Bio. Chem.**, v. 281(40), p. 30094-30103, 2006.
- MORGAN, J.; HARITAKUL, R.; KELLER, P. A. Anilinopyrimidines as novel antituberculosis agents. **Bioorg. Med. Chem. Lett.**, v. 13, p. 1755-1757, 2003.
- MÜLLER, C. W. et al. Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. **Structure** v. 4, p. 147-156, 1996.
- MURSHUDOV, G. N.; VAGIN A. A.; DODSON, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method **Acta Cryst.** D53, p. 240-255, 1997.

- 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.**, v. 359(3), p. 646-66 2006.
- PAN, P.; WOEHL, E.; DUNN, M. F. Protein architecture, dynamics and allostery in tryptophan synthase channeling. **TIBS**, v. 22, p. 22-27, 1997.
- PARISH, T.; STOKER, N. G. The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. **Microbiology**, v. 148, p. 3069-3077, 2002.
- PEREIRA, J. H. et al. Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. **Acta Crystallogr. D. Biol.**, v. 60 (pt 12 pt 2), p. 2310-2319, 2004.
- PELCZAR, M. J. Jr.; CHAN, E. C. S; KRIEG, N. R. **Microbiologia – conceitos e aplicações**. Tradução: Sueli F. Yamada et al. 2.ed. São Paulo: Makron Books, 1996. 2 v.
- PITTARD, A. J. Biosynthesis of the aromatic amino acids. In *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. Neidhardt, F. C., Washington, DC: American Society for microbiology, pp 368-394, 1987
- QUÉMARD, A. et al. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. **Biochemistry**, v. 34, p. 8235–8241, 1995.
- QUEVILLON-CHERUEL S. et al. Crystal structure of the bifunctional chorismate synthase from *Saccharomyces cerevisiae*. **J. Biol. Chem.** v. 279(1), p. 619-625, 2004.
- RAVIGLIONE, M. C. et al., Assessment of worldwide tuberculosis control. WHO global surveillance and monitoring project. **The lancet**, v. 350, p. 624-629, 1997.
- 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. USA**, v. 100, p. 13881–13886, 2003.
- ROBERTS, C. W. et al. The shikimate pathway and its branches in apicomplexan parasites. **J. Infect. Dis.**, v. 185, Suppl 1, p. S25-36, 2002.
- ROBERTS, F. et al. Evidence for the shikimate pathway in apicomplexan parasites. **Nature**, v. 393(6687), p. 801-805, 1998
- ROMANOWSKI, M. J.; BURLEY, S. K. Crystal structure of the *Escherichia coli* shikimate kinase I (AroK) that confers sensitivity to mecillinam. **Proteins**, v. 47(4), p. 558-62, 2002.
- ROSWARSKI, D. A. et al. Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD<sup>+</sup> and a C16 fatty acyl substrate. **The Journal of Biological Chemistry**, v. 274(22), p. 15582-15589, 1999.

- ROSWARSKI, D. A. et al. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. **Science**, v. 279, p. 98–102, 1998.
- SACHPATZIDIS, A. et al. Crystallographic studies of phosphonate-based  $\alpha$ -reaction transition-state analogues complexed to tryptophan synthase. **Biochemistry**, v. 38, p. 12665-12674, 1999.
- SALI, A.; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **J. Mol. Biol.**, v. 234, p. 779–815, 1993.
- SCHRUEDER, E. K. et al. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. **Curr. Pharm. Biotechnol.**, v. 3, p. 197–225, 2002.
- SCHULZ, G. E. Binding of nucleotides by proteins. **Curr. Opin. Struct. Biol.** v. 2, p. 61-67, 1992.
- SKÖLD, O. Resistance to trimethoprim and sulfonamides. **Veterinary Research**, v. 32, p. 261-273, 2001.
- SMITH, D. A. et al. Characterization of auxotrophic mutants of *Mycobacterium tuberculosis* and their potential as vaccine candidates. **Infection and Immunity**, v. 69 (2), p. 1142-1150, 2001.
- TACKET, C. O.; SZTEIN, M. B.; LOSONSKY, G. A.; WASSERMAN, S. S.; NATARO, J. P.; EDELMAN, R.; PICKARD, D.; DOUGAN, G.; CHATFIELD, S. N.; LEVINE, M. M. Safety of live oral *Salmonella typhi* vaccine strains with deletions in *htrA* and *aroC aroD* and immune response in humans. **Infect. Immun.**, v. 65, p. 452-456, 1997.
- TROUILLER, P. et al., Drug development for neglected diseases: a deficient market and a public-health policy failure. **Lancet**, v. 359(9324), p. 2188-2194, 2002.
- TROUILLER, P. et al., Drugs for neglected diseases: a failure of the market and a public health failure? **Tropical Medicine And International Health**, v. 6, p.945-951, 2001.
- VIOLA, C. M.; SARIDAKIS, V.; CHRISTENDAT, D. Crystal structure of chorismate synthase from *Aquifex aeolicus* reveals a novel beta alpha beta sandwich topology. **Proteins**, v. 54(1), p. 166-169, 2004.
- YAN, H.; YSAI, M. D. Nucleoside monophosphate kinases: structure, mechanism, and substrat specificity. **Advan. Enzymol. Relat. Areas Mol. Biol.** v. 73, p. 103-134, 1999.
- WANG, R et al. SCORE: a new empirical method for estimating the binding affinity of a protein-ligand complex. **J. Mol. Model.** v. 4, p. 379-394, 1998.
- WEYAND M.; SCHLICHTING I. Crystal structure of wild-type tryptophan synthase complexed with the natural substrate indole-3-glycerol phosphate. **Biochemistry**, v. 38(50), p.16469-16480, 1999.

## **6. Anexo**

**Artigo de Revisão publicado**

## **Chorismate Synthase: an attractive target for drug development against neglected disease.**

Marcio V. B. Dias; Fernanda Ely; Mário S. Palma; Luiz A. Basso; Diógenes S. Santos; Walter Filgueira de Azevedo Jr. **Current Drugs Targets**, v. 8, p. 48-55, 2007.

Neste artigo é apresentada uma revisão a respeito das informações sobre a corismato sintase publicadas em revistas indexadas nos últimos anos.

Nele é descrito a importância desta enzima e da via onde ela se encontra como alvo para o desenvolvimento de novas drogas. É comentado desde a descoberta do produto desta enzima, o corismato em 1962 até as informações publicadas em 2004. É realizado um estudo das características desta enzima em diferentes organismos que a possui, como fungos, bactérias, plantas e parasitas do filo apicomplexa. Portanto, é descrita as peculiaridades das corismato sintases monofuncionais encontradas em bactérias, plantas e parasitas do filo apicomplexa e das bifuncionais encontradas em fungos. São discutidos também, os dados existentes sobre o possível processo evolutivo da existência de enzimas monofuncionais e bifuncionais. São mostradas as peculiaridades das corismato sintases de organismos eucariotos (Plantas e parasitas de filo apicomplexa) e sua localização nestas células. É discutido detalhadamente o mecanismo catalítico desta enzima, com a importância da molécula de FMN para a transferência de elétrons para a molécula de EPSP e a importância da molécula de NADPH para as corismato sintases bifuncionais. É realizada uma descrição da estrutura desta enzima, os aspectos peculiares do seu sítio ativo e a disposição das moléculas de FMN e EPSP, e como isso pode auxiliar no processo de catálise desta enzima. E finalmente, são apresentados os estudos já realizados no desenvolvimento de novos ligantes e as estratégias utilizadas.



## Chorismate Synthase: An Attractive Target For Drug Development Against Orphan Diseases

Marcio V. B. Dias<sup>a\*</sup>, Fernanda Ely<sup>b</sup>, Mário S. Palma<sup>c</sup>, Walter F. de Azevedo Jr.<sup>d</sup>, Luiz A. Basso<sup>b</sup> and Diógenes S. Santos<sup>b\*\*</sup>

<sup>a</sup> Programa de Pós-Graduação em Biofísica Molecular – Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brasil; <sup>b</sup> Centro de Pesquisas em Biologia Molecular e Funcional / PUCRS. Avenida Ipiranga 6681, Tecnopuc, Partenon 90619-900, Porto Alegre, RS, Brazil; <sup>c</sup> Laboratório de Biologia Estrutural e Zooquímica-CEIS / Departamento de Biologia – Instituto de Biociências, UNESP, Rio Claro, SP 13506-900, Brasil and <sup>d</sup> Faculdade de Biociências-/ PUCRS. Avenida Ipiranga 6681, Partenon 90619-900, Porto Alegre, RS, Brazil

**Abstract:** The increase in incidence of infectious diseases worldwide, particularly in developing countries, is worrying. Each year, 14 million people are killed by infectious diseases, mainly HIV/AIDS, respiratory infections, malaria and tuberculosis. Despite the great burden in the poor countries, drug discovery to treat tropical diseases has come to a standstill. There is no interest by the pharmaceutical industry in drug development against the major diseases of the poor countries, since the financial return cannot be guaranteed. This has created an urgent need for new therapeutics to neglected diseases. A possible approach has been the exploitation of the inhibition of unique targets, vital to the pathogen such as the shikimate pathway enzymes, which are present in bacteria, fungi and apicomplexan parasites but are absent in mammals. The chorismate synthase (CS) catalyses the seventh step in this pathway, the conversion of 5-enolpyruvylshikimate-3-phosphate to chorismate. The strict requirement for a reduced flavin mononucleotide and the *anti* 1,4 elimination are both unusual aspects which make CS reaction unique among flavin-dependent enzymes, representing an important target for the chemotherapeutic agents development. In this review we present the main biochemical features of CS from bacterial and fungal sources and their difference from the apicomplexan CS. The CS mechanisms proposed are discussed and compared with structural data. The CS structures of some organisms are compared and their distinct features analyzed. Some known CS inhibitors are presented and the main characteristics are discussed. The structural and kinetics data reviewed here can be useful for the design of inhibitors.

**Key Words:** Infectious disease; neglected disease; chorismate synthase; shikimate pathway; flavin-dependent enzymes.

### INTRODUCTION

A worrying increase in incidence of infectious diseases was recognized in the late 1980's, causing the suffering of millions of people, especially in tropical and subtropical areas as Africa, Asia and South America – which account for four-fifth of the world's population. Each year, 14 million people are killed by infectious diseases, mainly diseases as HIV/AIDS, respiratory infections, malaria and tuberculosis [1]. The increase of immunodeficient population, mainly HIV-positive, the large homeless population and decline in health care structures are some problems responsible for the increased incidence of many infectious diseases in the world. Among the neglected infectious diseases are tuberculosis and malaria. The *Mycobacterium tuberculosis*, the aetiological agent of tuberculosis (TB), kills more than 3 million people each year and ninety percent of TB cases occur in developing countries [2]. In addition, malaria, caused by apicomplexan parasites of the genus *Plasmodium*, infects millions of people and account for the death of more than 2 million children annually [3]. In contrast, despite the enormous burden in the poor countries, drug discovery and development targeted at tropical diseases are at a standstill [1, 4]. Today, the pharmaceutical industry is reluctant to invest in drug development to treat the major diseases of the poor countries, because the financial return cannot be guaranteed. Thus, an urgent reorientation of priorities in drug development and health policy is needed, principally by national and international policies that need to direct the global economy and address the true health needs of society [1].

With the completion of the genome sequences of several pathogenic organisms is occurring an enormous impact on our under-

standing of the pathogenicity of these organisms. Thus, the information obtained by genome projects can be used by national and international research policy makers and scientific community for development of drugs against many of these diseases. The genome allows identification of metabolic pathways present in the pathogenic microorganisms, which can be target for the development of new drugs. A possible approach to selective antimicrobial chemotherapy has been to exploit the inhibition of unique targets, vital to the pathogen and absent in mammals [5]. The shikimate pathway is an attractive example of this kind of targets, since it is present in bacteria, fungi and apicomplexan parasites but absent from mammals [6]. In this pathway the glycolytic intermediate, phosphoenol pyruvate, and the pentose phosphate pathway intermediate, D-erythrose 4-phosphate, are converted to chorismate through seven metabolic steps [7]. The essentiality of shikimate pathway was observed in some microorganisms such as *Plasmodium falciparum* and *M. tuberculosis*. The disruption of *aroK* gene, which codes for the shikimate kinase, showed that this enzyme is essential for *M. tuberculosis* viability [8]. In *P. falciparum* the growth was inhibited by glyphosate, a well-characterized inhibitor of the shikimate pathway [9]. These reports provide strong evidence that shikimate pathway is essential for the survival of these pathogens; therefore, its enzymes are potential targets for drug development.

The product of shikimate pathway, the chorismate or chorismic acid, is a dihydroaromatic compound and it was first described by Frank and Margaret Gibson in 1962 [10]. This compound is the branch point in the biosynthesis of several important aromatic molecules. For this reason, it was named chorismate, which means, in Greek, separation, split, or divorce. The chorismate is the common precursor for the biosynthesis of a wide range of primary and secondary metabolites including aromatic amino acids (phenylalanine, tyrosine and tryptophan), folate, naphthoquinones, menaquinones and mycobactins [11]. Chorismate synthase (CS), which is responsible for the synthesis of chorismate, was first described in 1967 by Morell *et al.* [12]. This enzyme catalyses the seventh step

\*Address correspondence to this author at the Programa de Pós-Graduação em Biofísica Molecular – Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brasil; Tel: -----; Fax: -----; E-mail: diogenes@puers.br

in the shikimate pathway, the conversion of 5-enolpyruvylshikimate-3-phosphate (EPSP) to chorismate [13].

*Brucella suis*, a gram-negative bacteria, which is responsible for animal brucellosis in a variety of mammalian hosts, had the *aroC*-encoded CS gene disrupted in an array for oral vaccine against brucellosis. The *B. suis* mutant is highly attenuated in tissue culture and murine virulence models [14]. Other pathogenic bacteria as *Salmonella enterica* and *Yersinia enterocolitica* with *aroC* knockout [15-16] also grow slowly *in vivo*. The *in vivo* behavior of the bacterial *aroC* mutants (low-level residual virulence) suggests the importance of the CS for these microorganisms and shows that mutants can be candidate vaccines and the protein can be targeted for drug development.

The CS enzyme has been described in bacteria, fungi, plants and apicomplexan parasites. However, the requirement of reduced flavin cofactor is an intriguing factor, because the CS does not do net redox change when the substrate EPSP is converted to chorismate. The reaction catalyzed by CS involves an 1,4-anti-elimination of the 3-phosphate group and of the C-(6*proR*) hydrogen from EPSP, which is unusual and unique in nature [17-18]. The determination of CS structure from *Streptococcus pneumoniae* complexed with EPSP and FMN [19] could elucidate or confirm the role of the flavin in the chemical reaction catalyzed by CS. Kinetics studies could show that reduced flavin play a pivotal role in the reaction. In the structure of *S. pneumoniae*, EPSP molecule is stacked above the *si*-face of the isoalloxazine ring in an average distance of 3.3 Å. This close juxtaposition of FMN and EPSP supports earlier suggestions for a direct role of reduced FMN in the elimination reaction, such as a radical mechanism, in which electron transfer from the reduced cofactor to the substrate initiates C-O bond breakage of the phosphate group to yield a substrate-derived neutral radical [18, 20-24]. As there is not net redox change in the reaction, a reverse transfer of the transiently donated electron is required to complete the catalytic cycle [18]. This unusual reaction catalyzed by CS also makes it unique among flavin-dependent enzymes.

### CS FROM MICROORGANISMS (BACTERIA AND FUNGI)

CS was first described from three different microbial sources: *Escherichia coli* [12], *Bacillus subtilis* [25] and *Neurospora crassa* [26] (a gram-negative bacteria, a gram-positive bacteria and a fungi). Initially, it was believed that these identified proteins were three different enzymes because they were biochemically different. However, although they have many differences in their biochemical properties and molecular masses, these enzymes catalyze the same reaction, the conversion of EPSP to chorismate. Moreover, it had been observed that there was another feature common among the three enzymes: the reduced flavin requirement. It was observed that the *E. coli* CS is active only in anaerobic conditions in the presence of either chemically or enzymatically reduced flavin [12], while the enzyme of *N. crassa* and *B. subtilis* appeared to be associated with a second enzymatic activity, the NAD(P)H dependent-flavin reductase [25, 27-29]. However, these two organisms use the NAD(P)H differently. The *N. crassa* enzyme possesses an intrinsic flavin reductase activity, located in the same polypeptide chain of CS. On the other hand, the *B. subtilis* CS appears associated to two other enzymes: the 3-dehydroquinase synthase, the second enzyme of shikimate pathway, and a NAD(P)H dependent-flavin reductase, forming a heterotrimeric complex. Thus, the CS from *N. crassa*, owing to this additional capacity to reduce flavin is termed bifunctional CS, while the other two CS, which do not have additional capacity to reduce flavin, are termed monofunctional CS.

Initially, it was believed that the difference between molecular weight of monofunctional *E. coli* CS enzyme ( $M_r$  39,138 Da) and bifunctional *N. crassa* CS enzyme ( $M_r$  46,400 Da) was due to a domain responsible for reduction of flavin [30]. Based on sequence comparisons with monofunctional CSs, two regions of 18 internal

amino acids residues and 29 C-terminal amino acids residues unique to *N. crassa* were deleted. The presence of these two regions in CS polypeptide chain was found not to be essential by complementation with an *E. coli* strain lacking CS [31]. The further characterization of *Saccharomyces cerevisiae* CS has reported a smaller (40,800 Da) and yet bifunctional enzyme, which complies with the hypothesis that is not feasible to predict the mono or bifunctionality based on the molecular weight [32]. At present, all fungi CSs known are bifunctional, while all bacterial and plant CSs are monofunctional. However, it is not known exactly what residues or region of active site are responsible for the bifunctionality of fungi CS, despite the experimental efforts trying to determine these properties.

### CSS EVOLUTION

One question that remains without answer is how was the mono/bifunctional CSs evolution? A possible explanation is the existence of a common ancestor with flavin reductase activity [32]. It has been concluded from an earlier phylogenetic analysis that chorismate synthases are monophyletic [33], it is not known to date whether the ancestral chorismate synthase is mono- or bifunctional [34]. However, it has been suggested that the common ancestor was probably bifunctional given that it is difficult to image the evolution of the intrinsic reductase activity in a framework of monofunctional enzymes [33,34]. It was surmised that bifunctionality may have either been maintained only in organisms in which the availability of reduced flavin is limiting or perhaps there was positive selection of monofunctionality. *Thermotoga maritima* is thought to be one of the oldest eubacterium and appears to have undergone considerable lateral gene transfer from the archaea. A phylogenetic tree of all CSs presently known suggests that *T. maritima* CS diverged with the archaea and moreover, considerably before any of the CSs for which bifunctionality is known (fungi CSs). Thus the classification of its CS as monofunctional could be considered to be cognate to the ancestral CS and would therefore not lend support to bifunctionality being ascendant [34]. Furthermore, the plant and fungal CSs emerged from a common ancestral protein after diverging from the monofunctional bacteria proteins. In this case, if sufficient flavin is available in the cell, there may have been no selective pressure to maintain de flavin reductase activity in some species. By now, all bifunctional CSs known are from fungal origin, showing that among these organisms the bifunctionality persisted under selective pressure, *i.e.* in organisms where the availability of reduced flavin is limiting for growth [33]. The shikimate pathway in fungi has an interesting feature: the presence of *arom* complexes - a pentafunctional protein complex that catalyzes the second to sixth reaction of the pathway [35]. Whether there is a causal relationship in the presence of *arom* and the bifunctionality of fungal CSs remain to be solved.

Nowadays several microbial CSs were isolated from different sources such as *M. tuberculosis* [36], *Helicobacter pylori* [37], *Salmonella typhimurium* [38], *S. cerevisiae* [39] and *Staphylococcus aureus* [40] and these could help to solve problem of CS evolution.

### APICOMPLEXAN PARASITES

The herbicide glyphosate, a potent inhibitor of the shikimate pathway enzyme EPSP synthase, has been shown to inhibit the *in vivo* growth of the apicomplexan parasites *Toxoplasma gondii*, *P. falciparum* and *Cryptosporidium parvum*. This effect has provided evidence for the presence of shikimate pathway in apicomplexan parasites. The CS was the first shikimate pathway enzyme identified in apicomplexan parasites by anaerobic assay [9].

The genome projects have identified the ORFs from *aroC*-encoded CS gene by sequence homology in apicomplexan parasites. Many introns were found in *aroC* gene from *T. gondii*, but none was identified in *P. falciparum* *aroC* gene. The lack of introns allows the direct expression of *P. falciparum* CS in bacterial systems.

Moreover, the *T. gondii aroC* is in a single copy, which allows gene deletion experiments. Other interesting feature in the *T. gondii aroC* gene is a potential GCN4 binding motif (TGACTC) in the 5' non-coding region [3]. This specific upstream promoter element is present in genes that are regulated by the control-activator protein GCN4. In *S. cerevisiae* the expression of genes encoding shikimate pathway enzymes DAHP synthase (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase) and the pentafunctional arom complexes (which catalyzes the steps second to sixth in the pathway) are regulated by GCN4 protein [41], which indicates a possible regulation of shikimate pathway by CS expression in apicomplexan parasites.

In plants, the shikimate pathway enzymes, like other nuclear-encoded plastid enzymes, are post-translationally targeted to the plastid by an amino-terminal leader sequence. The apicomplexan parasites have a plastid-like organelle called apicoplast [42]. Although the apicoplast proteins from apicomplexan parasites also have leader sequences, there is no evidence that shikimate pathway enzymes are localized in the apicoplast. Neither the *T. gondii* CS gene nor *P. falciparum* CS gene have an obvious amino-terminal leader sequence similar to that found on CSs from higher plants, which strongly suggests that CS apicomplexan enzymes are in the cytosol, and not in the apicoplast [43]. The shikimate pathway in fungi also occurs in cytosol and phylogenetic analysis has demonstrated that CSs from apicomplexan parasites are closer to fungal CSs than plants CSs, which is consistent with the idea of a non-apicoplast CS. Moreover, it also indicates a common ancestor to fungal and apicomplexan enzymes [44].

Studies with CS from *P. falciparum* have characterized the enzyme biochemically and immunologically. Cofactor immunofluorescence and cellular fractions followed by Western blot analysis has shown that CS is located in the cytosol of this organism in different stages of infection, which is in agreement with the phylogenetic analysis. However, despite the close relationship with fungal proteins, the enzymatic assay has demonstrated that *P. falciparum* CS is monofunctional as in plants and bacteria [45]. Therefore, the CS from *P. falciparum* has a combination of properties from plant and fungal enzymes and it seems to possess properties distinct from previously described CSs. It was found that both *T. gondii* and *P. falciparum* enzymes differ from other known CS in possessing a number of unique insertions [3]. Although the CSs in apicomplexan parasites are larger in length, they share all of the amino acids to be highly conserved in CSs from other species [9]. The implication of a cytosolic location for the development of new drugs for apicomplexan parasites therapy is clear, fewer membranes will have to be crossed.

The inhibition of *P. falciparum* growth with a dsRNA encoding a 900 bp fragment of *aroC* has shown that CS is required for the normal growth of this pathogen [46]. It suggests that CS may be a viable target for chemotherapy. Moreover, two fluorinated analogs of shikimate: (6*R*)-6-fluoroshikimate and (6*S*)-6-fluoroshikimate have been shown to be inhibitors of the *P. falciparum* growth *in vitro* [47]. The possible function of apicomplexan CS in shikimate pathway regulation makes this enzyme a promising target for drug development. Increasing our understanding of the apicomplexan shikimate pathway enzymes may expedite the discovery and design of new antiparasitic agents.

## CS MECHANISM

The CS-catalyzed chemical reaction was first described in *E. coli* extracts. It was reported that CS was inactivated under aerobic conditions and it could be activated most effectively by a reduced flavin cofactor [12]. Although many flavoproteins are involved in redox reactions, there is not an overall redox when EPSP is converted to chorismate and the reduced flavin mononucleotide is not consumed during the reaction. Another interesting feature in CS catalysis is an unusual 1,4 *anti* elimination of 3-phosphate group

and the C(6*proR*) hydrogen from EPSP, implied in a nonconcerted reaction, which is interesting because it involved the cleavage of a nonactivated C-H bond [17]. The strict requirement for a reduced flavin mononucleotide and the *anti* 1,4 elimination are both unusual aspects which make CS mechanism unique among flavin-dependent enzymes [33]. Although there are two different classes of CS, distinguished by their ability to use NADPH for the reduction of FMN, there is no apparent difference in the mechanism of elimination reaction [18].

A nonconcerted mechanism was demonstrated by studies that include a secondary tritium kinetic isotopic effect at C(3) [48], transient kinetic studies [20] and a secondary  $\beta$  deuterium kinetic isotopic effect at C(4) [49]. The kinetic isotopic experiments have revealed that C(6)-H and C(3)-O bonds are cleaved in distinct steps, and the phosphate group is eliminated before the C-H cleavage. Two intermediates of nonconcerted mechanism have been proposed to CS reaction: the cationic and the radical intermediate [50]. Although reported isotopic effects could not clearly distinguish between the radical and cationic mechanisms, they have provided evidence against alternative mechanisms. Moreover, the observation of spectral changes associated with the flavin during the catalytic reaction has provided the strong evidence that the reduced flavin has been actively involved in the reaction mechanism [51]. The role of flavin mononucleotide in catalysis derives from the lack of activity of the bifunctional enzyme from *N. crassa* with 5-deaza-FMN [52] and the formation of a stable flavin semiquinone radical with the monofunctional *E. coli* enzyme and the substrate analog (6*R*)-6-fluoro-EPSP, as a ternary complex [23, 52].

The characterization of the flavin intermediate has shown that it is not simply associated with the conversion of substrate to product since it must be formed before any chemical step involving EPSP. The formation of a flavin intermediate is associated with the formation of the ternary complex between enzyme, EPSP and reduced flavin, as described above. The flavin intermediate is generated very rapidly within a few milliseconds and disappears when all substrate has been consumed [20]. Spectral observations have shown that the reduced flavin is bound to CS in its monoanionic (deprotonated) form in the absence of substrate [22]. On the other hand, studies with flavin analogs have demonstrated that CS binds preferably the flavin in a neutral form, including the protonated reduced form, rather than anionic forms in the presence of EPSP or (6*R*)-6-F-EPSP [21]. The reported data led to the conclusion that CS binds flavin in its monoanionic form and the protonation of the N(1) position of FMN occurs upon binding of either EPSP or (6*R*)-6-F-EPSP to the binary complex, which indicates that protonated, reduced flavin is the observed reaction intermediate. The protonation could lower the reduced flavin redox potential, converting the reduced flavin into a better reductant, as required to promote the radical chemistry [20]. Moreover, chorismate was produced from (6*R*)-6-F-EPSP with the formation of a flavin semiquinone radical, indicating the involvement of radical chemistry in this reaction. The ternary complex is formed and a single electron is transferred from the low potential flavin to the analogue, which promotes the elimination of a phosphate group to generate the allylic radical [23]. This explanation is in agreement with the essential requirement of a flavin co-factor.

Studies with substrate analogs have shown that even though the (6*S*)-6-F-EPSP is a competitive inhibitor in *N. crassa* CS [54], it is converted to 6-fluorochorismate in *E. coli* CS at a rate between 270 and 370 times slower than EPSP [50]. Studies with fluorinated substrate analogues have shown that in *E. coli* CS the C(6)-H bond breaking occurs after the formation of the flavin intermediate. Although the C(6)-H breaking does not contribute significantly to rate limitation in *E. coli* enzyme, it is partially rate-limiting with *N. crassa* enzyme [52-54]. On the other hand, this result and deuterium kinetic isotopic effect studies [55] have provided evidence that steps after flavin intermediate decay, which may include product

release, are not significantly rate-limiting with EPSP or (6S)-6-F-EPSP.

Spectroscopic and physical studies have shown that *E. coli* CS undergoes a major structural change when the ternary complex is formed. When the EPSP and FMN are bound, the enzyme has a more compact shape, designated as closed form. On the other hand, the apoenzyme probably exhibits more conformational flexibility, designated as opened form [56]. The elucidation of *S. pneumoniae* CS structure provides evidence that the binding of EPSP induces changes in protein conformation, concurrent with a protonation of FMN to give the neutral reduced form [19]. It is observed mainly by the clear change in the position of His<sup>110</sup> when the opened and closed structures are compared. It was proposed that the His<sup>110</sup> was the probable residue that protonates the FMN and neutralizes the reduced FMN in the first step of the reaction. The His<sup>10</sup> also appears to play a role in the reaction by providing a proton to neutralize the charge on the departing phosphate group [19].

The fundamental role of these two histidine residues in CS active site, as observed in *S. pneumoniae*, was confirmed by site-directed mutagenesis of two histidine residues in the active site of *N. crassa* CS. These histidines, His<sup>106</sup> and His<sup>17</sup>, were replaced by alanine, reducing the activity 10- and 20-fold, respectively. The phosphate release occurs before of the C-(6*proR*) hydrogen cleavage [20] and probably the two histidines residues are involved in chemical steps leading directly to the initial C-O bond cleavage. It was proposed that His<sup>106</sup> is involved in the protonation of monoanionic reduced FMN, so it functions as the acid that donates the proton to N(1). This protonation may serve to direct the electron density so that charge transfer to the double bond of the substrate is maximized. The function of His<sup>17</sup> is protonation of the leaving phosphate group of EPSP. A detailed mechanism for CS reaction was proposed based on structure and site-directed mutagenesis. According to this model the EPSP binds to enzyme-bound reduced FMN, forming a closed form. Then, an electron is transferred to the EPSP double bond, initiating C-O bond cleavage and the release of phosphate. In this step the His<sup>17</sup> acts as a general acid to neutralize the incipient charge on the oxygen atom. The resulting C(4a)-neutral flavin semiquinone tautomerizes to a radical species where the unpaired electron resides on N(5) with concomitant abstraction of the (*proR*-H) from C(6) by the N(5) [18]. Following the cleavage of phosphate and possibly partial re-opening of the active site, the intermediate remains bound while the proton is abstracted. The atom N(5) of FMN is positioned directly below the C(6-*proR*) hydrogen of EPSP and is ideally positioned to remove this proton from the transient intermediate. In the last step, the deprotonation of the reduced flavin restores the initial state of the cofactor. The regeneration of reduced FMN occurs with the reopening of the active site, involving deprotonation by His<sup>110</sup> in the *S. pneumoniae* [19]. It is also proposed that neither of the two histidines residues is essential for binding of the oxidized cofactor, representing only a role in the stabilization of the change of the cofactor [18].

The future contributions as studies with substrate analogues, high-resolution structures information and site-directed mutagenesis in other invariant amino acids residues in the active site will provide further insights into this unprecedented flavin-dependent reaction mechanism.

#### THE ROLE OF NADPH IN THE BIFUNCTIONAL CHORISMATE SYNTHASE

CSs from fungal sources, *i.e.* *S. cerevisiae* and *N. crassa*, have and additional intrinsic catalytic activity, which utilizes NAD(P)H to reduce the flavin cofactor. (13, 56). The NAD(P)H is utilized in the bifunctional CS to generate the essential reduced FMN cofactor. NAD(P)H is consumed for the reduction of FMN followed by re-oxidation of the reduced FMN by molecular oxygen. The rate of flavin reduction probably is independent of the concentration and

presence of EPSP. The most intriguing question about the bifunctional CS is the nature and the location of the NAD(P)H binding site. All studies carried out to identify the sequence responsible for binding of NAD(P)H have failed (31), therefore, the ability of bifunctional CS to bind NAD(P)H is embedded in the protein structure revealing an unusual mode of interaction between these CS and NAD(P)H [13]. Furthermore, Kitzing *et al.* (2001) [13] have observed that EPSP and NAD(P)H compete for a common binding site because the rate of NAD(P)H oxidation strongly depends on the concentration of EPSP. It is conceivable because both NAD(P)H and EPSP possess a phosphate group, which in turn plays an important role in binding to the active site of CS. The site-directed mutagenesis studies of important histidine residues of the active site of *N. crassa* CS demonstrated that specific histidines to alanines replacements affected the intrinsic NAD(P)H:FMN oxidoreductase activity. It is mentioned that the lack of activity of the two histidine to alanine mutant proteins under aerobic assay conditions does not lead to conversion of EPSP to chorismate. Although both amino acid exchanges affect the intrinsic oxidoreductase activity of CS, they are apparently not decisive for secondary enzymatic activity, as they do not abolish the flavin reductase activity [18]. Thus, the amino acid residues that bring about bifunctionality remain obscure [18]. Therefore, only the determination of the three-dimensional structure of a bifunctional CS complexed with NAD(P)H will provide a description at molecular level of the essential residues responsible for bifunctionality.

#### THE STRUCTURE OF CHORISMATE SYNTHASE

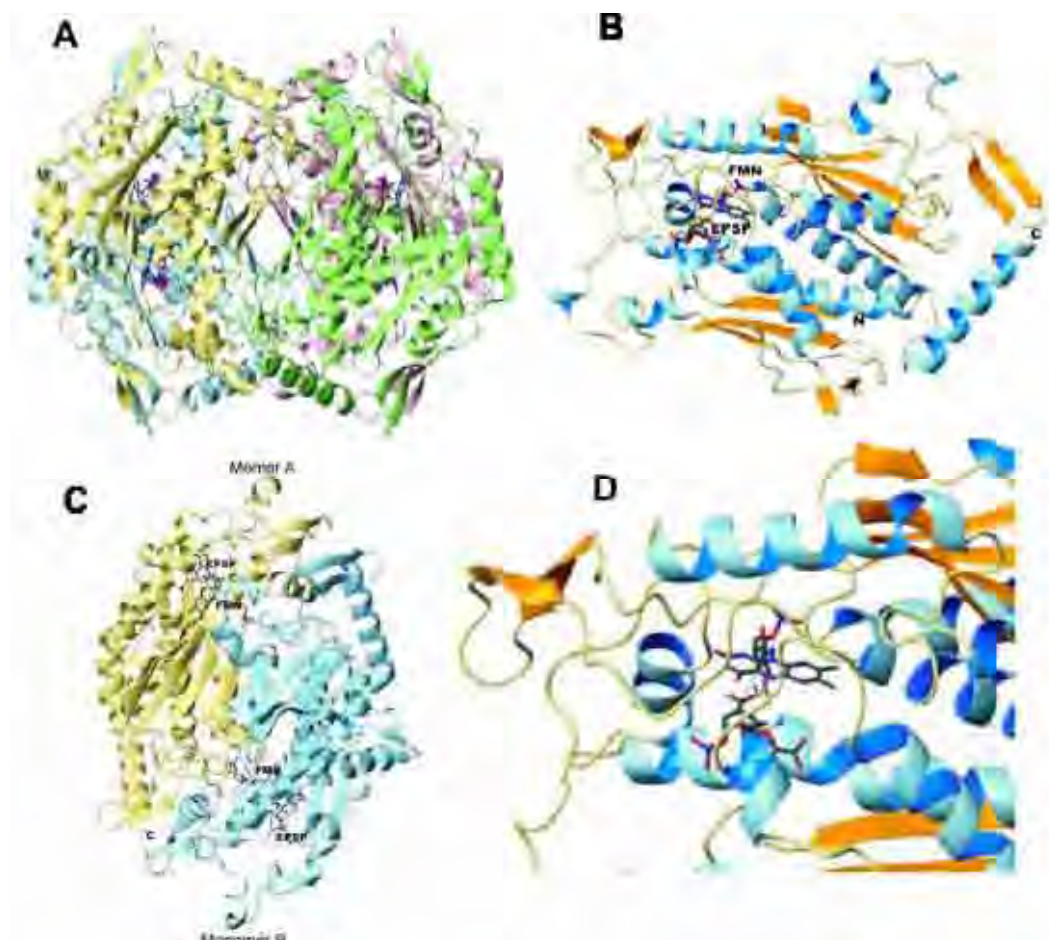
Macheroux *et al.* (1998) [56] have predicted the CS structure in a  $\alpha$ -barrel form as the result of secondary structure prediction efforts. However, the determination of high-resolution structures by crystallography has shown that they were not correct and the CS structure is very different from what was expected. The small-angle X-ray scattering data have indicated that the enzyme presents a more compact overall shape in the ternary complex when both oxidized FMN and EPSP are bound to the enzyme as compared to the apoenzyme. Probably, the presented apoenzyme exhibits more conformational flexibility, even though these ligands have no effect on the oligomerization state of the enzyme [5].

All crystal structures of CS solved to date show that the quaternary structure is tetrameric with approximately 222 symmetry (Fig. (1a)). Moreover all structures have high structural similarity representing other distinct class of a conserved enzyme family, since it lacks a primary sequence similarity to other classes of enzymes. Despite the CSs being tetrameric in its crystal form, in solution it can be present as a dimer, tetramer or dimer/tetramer equilibrium [12, 30, 39-40, 58].

The organisms whose CS crystal structures have been determined are: *S. pneumoniae* [19], *Aquifex aeolicus* [59], *H. pylori* [60], *S. cerevisiae* [61], *Campylobacter Jejuni* [PDB access code: 1SQ1]. Furthermore preliminary data of X-ray crystallography for CS from *M. tuberculosis* have also been reported [36].

The dominant structural topology of CS monomer is characterized by a  $\alpha$ -sandwich, in which each monomer consists of a central helical core formed for four long  $\alpha$ -helices, sandwiched

between two four-stranded antiparallel beta sheets which is surrounded by loops and discrete stretches of helix  $\alpha$  and sheet  $\beta$ . These layers are packed to form a compact structure [19, 59-61] (Fig. (1b)). The structure could also be described as two intimately contacted sub-domains, consisting of anti-parallel sheet covered by helices formed from  $-2\alpha$  halves that are related by pseudo 2-fold symmetry. However, the lack of identity between the sequences, which form two domains, shows that could not have occurred for gene duplication events [61-62].



**Fig (1).** Structure of CS; A) Tetramer structure; B) Monomer structure; C) Dimer; D) structure of active site.

Although CS is a tetramer in the crystals, it is better characterized as a dimer of dimers in its basic structural unit because two dimers compose the observed tetramer [19]. The dimer interface is extensive and includes antiparallel interactions between the ends of the sheet 2 from each subunit to give an extended eight-stranded  $\beta$  sheet spanning both subunits (Fig. (1c)) [62].

#### ACTIVE SITE CHARACTERISTICS

The interesting aspect of active site is the large number of loop regions that comprise the substrate binding pocket (fig. (1d)). Two of them are provided by the adjacent subunit across the dimer interface. In the ternary complex, FMN is almost completely buried, with part of its accessible surface area being occluded by the substrate, EPSP. Thus, the structure indicates an ordered binding of FMN followed by EPSP, which is consistent with previous rapid kinetic studies [56].

The FMN-binding site in CS is unique; no other flavoenzyme presents such characteristic. In the structure of *S. pneumoniae*, the FMN molecule makes few specific polar interactions with protein, most part of them are contacts between hydroxyl and phosphate oxygens of the ribityl chain. The flexible regions that surround the FMN site are rich in strictly or highly conserved residues. The FMN phosphate sits at the dimerization interface and it makes a number of contacts with residues from the adjacent monomer. Furthermore, there are a considerable number of solvent molecules close to both the phosphate and ribityl regions of FMN. The water molecules are discrete and ordered. They are responsible for medi-

ate interactions between FMN and the surrounding residues. The isoalloxazine ring system of FMN is hydrophobic and makes few specific interactions with the protein. However, it buries a considerable area of hydrophobic surface by packing the *re*-face against the complementary surface of the protein [19]. This environment contributes to the remarkable lowering of the flavin's reduction potential to a value comparable to the most reducing flavodoxins [62]. A notable feature of the bound FMN is a significant deviation from planarity displayed by isoalloxazine ring. The pyrimidine ring and the dimethylbenzene ring make an angle of approximately  $10^\circ$  and the *re*-face of the isoalloxazine ring is convex. It appears to assist the substrate EPSP-binding site. This characteristic is observed in very few other flavoenzymes [60, 63-65]. Furthermore the FMN binding site has a number of positive charges in the proximity of the isoalloxazine ring, which can increase the redox potential of FMN, whereas a negative charge or a hydrophobic environment is expected to lower it. The positive charges at this location can possibly stabilize negatively charged N1 in the reduced flavin, thus keeping the electron until the substrate EPSP is bound [60].

The EPSP molecule is oriented relative to the N (5) position of the isoalloxazine ring for electron transfer and hydrogen atom abstraction to and from the substrate, respectively. The EPSP molecule is attacked above the *si*-face of the isoalloxazine ring in an average distance of 3.3 Å as in the structure of CS from *S. pneumoniae* (Fig. (1d)). Owing to this fact, it can be observed that during the reaction there is no accumulation of flavin radical intermediates on the stopped-flow millisecond time scale. Hence, CS catalyzes a

unique reaction when occurs only the transfer of one electron, which is unusual among flavoenzymes. The localization of the FMN and EPSP in the CS has provided evidence that the flavin is well placed to abstract a hydrogen atom from the 6-*proR* position of the EPSP molecule, rather than just an electron with the proton being accepted by an amino acid side chain. This form of flavin can donate an electron and subsequently accept a hydrogen atom, stabilizing again [19, 62].

In the structure of CS from *S. pneumoniae*, the EPSP molecule makes a number of polar interactions with the protein, and few hydrophobic contacts. The site of EPSP is hydrophilic and has an environment extensively basic with many basic residues, mainly arginines around of the binding site, which are extremely conserved in many sequences, thereby showing its importance in the charge stabilization of EPSP molecule [19].

Maclean and Ali (2003) [19] observed that the quaternary structure of *S. pneumoniae* has two conformational states. One of the four monomers of the tetramer has differences when compared with the other three. These changes occur principally in two loops of active site comprised of highly conserved residues. The quaternary structure shows one monomer with the active site more accessible having therefore an open conformation, whereas the other three have closed conformations. These differences provide vital information about the mobility of these active site loops, which play an important mechanistic role [19]. These loops have important basic residues related with the catalytic mechanism of the protein. Differences between open and closed active sites demonstrate that the side chain of one histidine located in one of these loops has some conformational flexibility, and interacts with the FMN in the open form but with O12 of EPSP in the closed form. Furthermore other highly conserved residues in these loops also have different conformations in the two states showing that these residues play important role in the catalytic mechanism [19]. Ahn *et al.* (2004) [60] have observed also that the structure from *H. pylori* is surrounded by flexible regions that are also highly conserved. Ahn *et al.* (2004) [60] crystallized the apoenzyme and CS complexed with FMN. The authors have observed that FMN could cause conformational changes in these loops in CS from *H. pylori*, furthermore, they have observed that the movement of these regions results in a more apolar environment for the bound cofactor [60]. This structural information confirms the hypothesis, in which the CS undergoes a major change on binding oxidized flavin and EPSP and that the apoenzyme exhibits more conformational flexibility than ternary complex. Furthermore it was observed that alterations occurred in the visible CD spectrum of the enzyme-bound FMN on binding of EPSP showing that the environment of the flavin changes considerably [56].

#### STRATEGY FOR DEVELOPMENT DE DRUG BASED IN THE CHORISMATE SYNTHASE

Based on the structural knowledge of the enzymes of the shikimate pathway and their mechanisms of catalysis, it is possible to design potential inhibitors in a rational way against pathogenic bacteria, fungi and apicomplexan parasites. For amenable manner and for knowledge about the shikimate, it was chosen for evaluating of the design potential inhibitors. The strategy used was to incorporate fluorine at regio- and stereo-specific site of this molecule [66-68]. With the knowledge of loss of C(6-*proR*) hydrogen in the reaction catalyzed for CS [17, 66], two shikimate analogs were synthesized, the (6*R*)-6-fluoro- and (6*S*)-6-fluoro shikimate acids. The fluorine in (6*R*)-6-fluoro shikimate acid occupies the position of the hydrogen abstracted during conversion of EPSP to chorismate. The (6*R*)-6-fluoro- and (6*S*)-6-fluoro derivatives from shikimate acid are metabolized by shikimate kinase and then to EPSP synthase to give the corresponding diastereoisomeric 6-fluoro-EPSP. The (6*R*)-6-EPSP might inhibit the CS, while the (6*S*)-6-EPSP would be processed further and might lead to lethal synthesis [66]. The (6*R*)-6-

fluoro-EPSP binds to *E. coli* CS *in vitro* and causes the oxidation of enzyme-bound flavin, but it is not converted further, and neither chorismate nor any other product could be detected [53]. On the other hand, the (6*S*)-6-fluoro-EPSP is slowly converted to chorismate by CS from *E. coli* to a product identified as 6-fluoro-chorismate [66] that inhibits the 4-aminobenzoic acid synthesis [55, 66]. Because of its greater potency *in vitro*, the shikimic acid analogues were tested for antibacterial activity *in vivo*. The (6*S*)-6-fluoro-shikimate was more protective against bacterial intraperitoneal challenges in a mouse protection test. Mice were infected with *S. aureus* and *Salmonella dublin*, but none of these strains would grow when mice were treated with (6*S*)-6-fluoro-shikimate on the minimal chemically defined medium [66, 69].

The 6*R*- and 6*S*- isomers of 6-fluoro-shikimates have antibacterial activity displaying minimum inhibitory concentrations of 64 and 0.5  $\mu\text{g mL}^{-1}$  against *E. coli*, respectively [66]. Owing to the lack of (6-*proR*) hydrogen in (6*R*)-6-fluoro-EPSP, it cannot be considered a substrate of CS. Thus, this compound promotes one electron oxidation of the *E. coli* enzyme [53] and it is a competitive inhibitor of the *N. crassa* enzyme. The (6*S*)-6-fluoro-shikimate is also a competitive inhibitor of the *N. crassa* enzyme, because there was no conversion of this analog to 6-fluoro-chorismate by the enzyme of this organism [54]. On the other hand, in CS from *E. coli*, this analogue is not an inhibitor, since the conversion to 6-fluoro-chorismate occurs. It can occur due the *N. crassa* enzyme to be different, once, it is bifunctional [30, 57], differently of the *E. coli* enzyme. Despite of conversion of (6*S*)-6-fluoro-EPSP for *E. coli* CS, it is converted in a rate between 270 and 370 times slower than EPSP [50].

The decreased rate of reaction is consistent with the electron withdrawing fluoro constituent destabilizing an allylic cationic intermediate [70] that would be generated by the loss of phosphate from the substrate [50]. There is a strong evidence for a stepwise, rather than a concerted condensation reaction, involving an allylic cationic intermediate with the fluoro substituent adjacent to the allylic system. An allylic cationic intermediate in the CS reaction would also be destabilized, in part, by negative hyperconjugative effect of the adjacent fluoro substituent [71]. The negative hyperconjugative effect of the fluorine could destabilize an allylic radical intermediate that would be formed by a mechanism involving an additional one-electron reduction [50]. Once a ternary complex is formed between the reduced holoenzyme and the (6*R*)-fluoro-EPSP analogue, a single electron is transferred from the low potential flavin to the analogue. This promotes the elimination of the phosphate group to generate an allylic radical. This intermediate is incapable of transferring an electron and proton to the flavin from the 6*R* position due to its (6*R*)-fluoro group [23]. The (6*S*)-6-fluoro-shikimate does not inhibit the CS, but the product formed by CS, the 6-fluoro-chorismate inhibits the PabA and PabB. These enzymes are responsible for conversion of chorismate to 4-aminobenzoate precursor 4-amino-4-deoxychorismate [72].

The analogues of flavin also could be used to the design of inhibitors of CS, since it has been demonstrated that occurs lack of activity of the enzyme with reduced 5-deaza-FMN [23].

Other class of inhibitors of CS was developed by Thomas *et al.* (2003) [73]. Among these inhibitors are the benzofuran-3[2H]-one. This inhibitor is moderately potent, with  $\text{IC}_{50}$  of 8  $\mu\text{M}$ . Kinetic studies demonstrated that the benzofuran-3-[2H]-one is a competitive inhibitor with respect to EPSP and noncompetitive with FMN, suggesting that the compound binds in the active site of the enzyme. However, owing to the metabolic vulnerability and potential toxicity of the 6,7-dihydroxy functional group, and also in an effort to increase potency, further synthetic efforts have been undertaken [73].

#### CONCLUSION

CS is an important enzyme of the shikimate pathway, having a critical role in metabolism of microorganisms. Furthermore it is

fascinating, mainly for their unusual catalytic mechanism and for their unique folding in nature. The knowledge about the CSs has been evaluated day by day. The determination of various structures of this protein would contribute to an understanding of the details of the mechanism of catalysis and reveal the role played by important active site residues, such as the His<sup>10</sup> and His<sup>110</sup> in *S. pneumoniae*. Future research could also contribute to the understanding of the unprecedented flavin-dependence in the reaction mechanism, which can be through of site-directed mutagenesis and the utilization of substrate and cofactor analogues. The major challenge today, is to precisely determine the role of NAD(P)H in bifunctional CSs and the residues that are involved in their binding. On the other hand, today, major new opportunities in structure-based drug design are possible due the determination of CS structure complexed with both substrate and cofactor. This knowledge could contribute to the development of new drugs including the use of combinatorial chemistry and virtual screening methods that could render inhibitors more specific to compete with substrate and cofactor in binding to the enzyme. The development of new drugs or medicines against infectious disease such as tuberculosis, respiratory disease and malaria will help hundreds of millions of people worldwide [74-86], mainly in developing countries, where poor people will most benefit from these efforts.

#### ACKNOWLEDGMENTS

This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0, 03/12472-2, 04/00217-0) and Milkenium Institute, MCT/CNPq to DSS and LAB. MSP, DSS, WFA and LAB are Research Awardees from the National Research Council.

#### ABBREVIATION LIST

CS	=	Chorismate synthase
TB	=	Tuberculosis
EPSP	=	5-enolpyruvylshikimate-3-phosphate
FMN	=	Flavin mononucleotide
FMN <sub>red</sub>	=	Flavin mononucleotide reduced
FMN <sub>oxi</sub>	=	Flavin mononucleotide oxidized
FMN <sub>SQ</sub>	=	Flavin mononucleotide semiquinone
DAHP	=	3-deoxy-D-arabino-heptulosonate 7-phosphate
6(R)-6-F-EPSP	=	(6R)-6-fluoro-5-nolpyruvylshikimate-3-phosphate
6(S)-6-F-EPSP	=	(6S)-6-fluoro-5-nolpyruvylshikimate-3-phosphate
His <sup>10</sup>	=	Histidine 10
His <sup>110</sup>	=	Histidine 110
PDB	=	Protein data bank

#### REFERENCES

- Trouiller, P.; Torreele, E.; Oliario, P.; White, N.; Foster, S.; Wirth, D. and Pécoul, B. (2001) *Trop. Med. Internat. Health.*, **6**(11), 945-951.
- Corbett, E. L.; Watt, C. J.; Walker, N.; Maher, D.; Williams, B. G.; Raviglione, M. C. and Dye, C. (2003) *Arch. Intern. Med.*, **1639**, 1009-1021.
- Roberts, C. W.; Roberts, F.; Lyons, R. E.; Kirisits, M. J.; Mui, E. J.; Finnerty, J.; Johnson, J. J.; Ferguson, D. J. P.; Coggins, J. R.; Krell, T.; Coombs, G. H.; Milhous, W. K.; Kyle, D. E.; Tzipori, S.; Barnwell, J.; Dame, J. B.; Carlton, J. and McLeond, R. (2002) *J. Infect. Dis.*, **185**, S25-36.
- Pécoul, B.; Chirac, P.; Trouiller, P. and Pinel, J. (1999) *J. of the American Med. Association*, **281**, 361-367.
- Coates, A.; Hu, Y.; Bax, R. and Page, C. (2002) *Nat. Rev. Drug. Discov.*, **1**, 895 – 910.
- Bentley, R. (1990) *Crit. Rev. Biochem. Mol. Biol.*, **1990**, 307-384.
- Herrmann, K. M. (1995) *Plant Cell*, **7**, 907-919.
- Parish, T. and Stoker, N. G. (2002) *Microbiol.*, **148**, 3029-3077.
- Roberts, F.; Roberts, C. W.; Johnson, J. J.; Kyle, D. E.; Krell, T.; Coggins, J. R.; Coombs, G. H.; Milhous, W. K.; Tzipori, S.; Ferguson, D. J. P.; Chakrabarti, D. and McLeond, R. (1998) *Nature*, **393**, 801-805.
- Gibson, M.I. and Gibson, F. (1962) *Biochim. Biophys. Acta*, **65**, 160-163.
- Ratledge, C. (1982) Nutrition, growth and metabolism in: Ratledge, C., Stanford, J. L. (Eds.), *The biology of the Mycobacteria*, vol. 1, Academic Press, London, pp. 185-271.
- Morell, H.; Clark, M.J.; Knowles, P.F. and Sprinson, D.B. (1967) *J. Biol. Chem.*, **242**, 82-90.
- Kitzing, K.; Macheroux, P. and Amrhein, N. (2001) *J. Biol. Chem.*, **276**(46), 42658-42666.
- Foulongne, V.; Walravens, K.; Bourg, G.; Boschiroli, M.L.; Godfroid, J.; Ramuz, M. and O'Callaghan D. (2001) *Infect. Immun.*, **69**(1), 547-550.
- Bowe, F.; O'Gaora, P.; Maskell, D.; Cafferkey, M. and Dougan, G. (1989). *Infect. Immun.*, **57**, 3234-3236.
- Lowe, D.C.; Savidge, T.C.; Pickard, D.; Eckmann, L.; Kagnoff, M.F.; Dougan, G. and Chatfield, S.T. (1999) *Infect. Immun.*, **67**, 700-707.
- Floss, H. G.; Onderka, D. K. and Carroll, M. (1972) *J. Biol. Chem.*, **247**, 736-744.
- Kitzing, K.; Auweter, S.; Amrhein, N. and Macheroux, P. (2004) *J. Biol. Chem.*, **279**, 9451-9461.
- Maclean, J. and Ali, S. (2003) *Structure*, **11**(12), 1499-1511.
- Bornemann, S.; Lowe, D. J. and Thorneley, R. N. (1996) *Biochemistry*, **35**(30), 9907-9916.
- Macheroux, P., Bornemann, S.; Ghisla, S. and Thorneley, R. N. (1996) *J. Biol. Chem.*, **271**(42), 25850-25858.
- Macheroux, P.; Petersen, J.; Bornemann, S.; Lowe, D. J. and Thorneley, R. N. (1996) *Biochemistry*, **35**(5), 1643-1652.
- Osborne, A.; Thorneley, R. N.; Abell, C. and Bornemann, S. (2000) *J. Biol. Chem.*, **275**(46), 35825-35830
- Machoroux, P.; Bornemann, and Thorneley, R. N. F. (1997) in *Flavins and flavoproteins 1996* (Stevenson, K. J.; Massey, V. and Williams, C. H. J., eds) University of Calgary Press, Calgary, Canada, pp. 113-122.
- Hasan, N. and Nester, E. W. (1978) *J. Biol. Chem.*, **253**(14), 4987-4982.
- Gaertner, F. H. and Cole, K. W. (1973) *J. Biol. Chem.*, **248**(13), 4602-4609.
- Gaertner, F. H. (1987) *Methods Enzymol.*, **142**, 362-366.
- Hasan, N. and Nester, E. W. (1978) *J. Biol. Chem.*, **253**(14), 4993-4998.
- Hasan, N. and Nester, E. W. (1978) *J. Biol. Chem.*, **253**(14), 4999-5004.
- White P. J.; Millar, G. and Coggins, J. R. (1988) *Biochem. J.*, **251**(2), 313-322.
- Henstrand, J. M.; Amrhein, N. and Schmid, J. (1995) *J. Biol. Chem.*, **270**(35), 20447-20452.
- Henstrand, J. M.; Schaller, A.; Braun, M.; Amrhein, N. and Schmid, J. (1996) *Mol. Microbiol.*, **22**(5), 859-866.
- Macheroux, P.; Schmid, J.; Amrhein, N. and Schaller, A. (1999) *Planta*, **207**, 325-334.
- Fitzpatrick, T. B.; Killer, P.; Thomas, R.M.; Jelesarov, I.; Amrhein, N. and Macheroux P. (2001) *J. Biol. Chem.*, **276**(21), 18052-18059.
- Hawkins, A. R. and Smith, M. (1991) *Eur. J. Biochem.*, **196**, 717-724.
- Dias, M. V., Borges, J. C., Ely, F., Pereira, J. H., Canduri, F., Ramos, C. H., Frazzon, J., Palma, M. S., Basso, L. A., Santos, D. S., de Azevedo, W. F. Jr. (2006) *J. Struct. Biol.* **154**(2), 130-143.
- Ahn, H. J.; Yang J. K.; Lee, B. I.; Yoon, H. J.; Kim, H. W. and Suh, S. W. (2003) *Acta Crystallogr. D Biol. Crystallog.*, **59**, 569-571.
- Charles, I. G.; Lamb, H. K.; Pickard, D.; Dougan, G. and Hawkins, A. R. (1990) *J. Gen. Microbiol.*, **136**(2), 353-358.
- Jones, D. G.; Reusser, U. and Braus, G. H. (1991) *Mol. Microbiol.*, **5**(9), 2143-2152.
- Horsburgh, M. J.; Foster, T. J.; Barth, P. T. and Coggins, J. R. (1996) *Microbiology*, **142**, 2943-2950.
- Arndt, K. and Fink, G. R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8516-8620.

- [42] Köhler, S.; Delwiche, C. F.; Denny, P. W.; Tilney, L. G.; Webster, P. and Wilson, R. J. (1997) *Science*, **275**, 1485-1489.
- [43] McFadden, G. I.; Reith, M. E.; Munholland, J. and Lang-Unnasch, N. (1996) *Nature*, **381**, 482.
- [44] Keeling, P. J.; Jeffrey, D. P.; Donald, R. G. K.; Roos, D. S.; Waller, R. F. and McFadden, G. I. (1999) *Nature*, **397**, 219-220.
- [45] Fitzpatrick, T.; Ricken, S.; Lanzer, M.; Amrhein, N. Macheroux, P. and Kappes, B. (2001) *Mol. Biol.*, **40**(1), 65-75.
- [46] McRobert, L. and McConkey, G. A. (2002) *Mol. Biochem. Parasitol.*, **119**, 273-278.
- [47] McConkey, G. A. (1999) *Antimicrob. Agents Chemother.*, **43**, 175-177.
- [48] Balasubramanian, S.; Coggins, J. R. and Abell, C. (1995) *Biochemistry*, **34**(1), 341-348.
- [49] Bornemann, S.; Theoclitou, M. E.; Brune, M.; Webb, M. R.; Thorneley, R. N. and Abell, C. (2000) *Bioorg. Chem.*, **28**, 191-204.
- [50] Bornemann, S.; Ramjee, M. K.; Balasubramanian, S.; Abell, C.; Coggins, J. R.; Lowe, D. J. and Thorneley, R. N. F. (1995) *J. Biol. Chem.*, **270**, 22811-22815.
- [51] Ramjee, M. K.; Coggins, J. R.; Hawkes, T. R.; Lowe, D. J. and Thorneley, R. N. F. (1991) *J. Am. Chem. Soc.*, **113**, 8566-8567.
- [52] Lauhon, C. T. and Bartlett, P. A. (1994) *Biochemistry*, **33**, 14100-14108.
- [53] Ramjee M. N.; Balasubramanian, S.; Abell, C.; Coggins, J. R.; Davies, G. M.; Hawkes, T. R.; Lowe, D. J. and Thorneley, R. N. F. (1992) *J. Am. Chem. Soc.*, **114**, 3151-3153.
- [54] Balasubramanian, S. and Abell, C. (1991) *Tetrahedron Lett.*, **32**, 963-966.
- [55] Bornemann, S.; Balasubramanian, S.; Coggins, J. R.; Abell, C.; Lowe, D. J. and Thorneley, R. N. (1995) *Biochem. J.*, **305**, 707-710.
- [56] Macheroux P.; Schonbrunn, E.; Svergun, D. I.; Volkov, V. V.; Koch, M. H.; Bornemann, S. and Thorneley, R. N. (1998) *Biochem. J.*, **335**, 319-327.
- [57] Welch, G. R.; Cole, K. W.; and Gaertner, F. H. (1974) *Arch. Biochem. Biophys.*, **165**, 505-518.
- [58] Schaller, A.; Windhofer, V. and Amrhein, N. (1990) *Pers. Arch. Biochem. Biophys.*, **282**, 437-442.
- [59] Viola, C.M.; Saridakis, V. and Christendat, D. (2004) *Proteins*, **54**(1), 166-169.
- [60] Ahn, H.J.; Yoon, H.J.; Lee, B. 2nd and Suh, S.W. (2004) *J. Mol. Biol.*, **336**(4), 903-915.
- [61] Quevillon-Cheruel, S.; Leulliot, N.; Meyer, P.; Graille, M.; Bremang, M.; Blondeau, K.; Sorel, I.; Poupon, A.; Janin, J. and van Tilbeurgh, H. (2004) *J. Biol. Chem.*, **279**(1), 619-625.
- [62] Bornemann, S.; Lawson, D. M. and Thorneley, R. N. (2003) *Structure*, **11**(12), 1463-1465.
- [63] Binda, C.; Coda, A.; Angelini, R.; Federico, R.; Ascenzi, P. and Mattevi, A. (1999) *Structure*, **7**, 265-276.
- [64] Yue, Q. K.; Kass, I. J.; Sampson, N.S. and Vrielink, A. (1999) *Biochemistry*, **38**, 4277-4286.
- [65] Barber, M. J.; Neame, P. J.; Lim, L. W.; White, S. and Matthews, F. S. (1992) *J. Biol. Chem.*, **267**, 6611-6619.
- [66] Davies, G.M.; Barrett-Bee, K.J.; Jude, D.A.; Lehan, M.; Nichols, W.W.; Pinder, P.E.; Thain, J.L.; Watkins, W.J. and Wilson, R.G. (1994) *Antimicrob. Agents. Chemother.*, **38**(2), 403-406.
- [67] Abeles, R.H. and Alston, T.A. (1990) *J. Biol. Chem.*, **265**(28), 16705-16708.
- [68] Kollonitsch, J. (1978) *Isr. J. Chem.* **17**, 53-59.
- [69] Davis, B. D. and Mingioli, E. S. (1950) *J. Bacteriol.*, **60**, 17-28.
- [70] Hawkes, T. R.; Lewis, T.; Coggins, J. R.; Mousdale, D. M.; Lowe, D. J. and Thorneley, R. N. F. (1990) *Biochem. J.*, **265**, 899-902.
- [71] von Ragué Schleyer; P. and Kos, A. J. (1983) *Tetrahedron*, **39**, 1141-1150.
- [72] Green, J. M. and Nichols, B. P. (1991) *J. Biol. Chem.*, **266**, 12971-12975.
- [73] Thomas, M.G.; Lawson, C.; Allanson, N.M.; Leslie, B.W.; Bottomley, J.R.; McBride, A. and Olusanya, O.A. (2003) *Bioorg. Med. Chem. Lett.*, **13**(3), 423-426.
- [74] Basso, L. A., Pereira da Silva, L. H., Fett-Neto, A. G., de Azevedo Jr., W. F., Moreira, I. S., Palma, M. S., Calixto, J. B., Astolfi Filho, S., dos Santos, R. R., Soares, M. B. P., Santos, D. S. (2005) *Mem. Inst. Oswaldo Cruz*, **100**(6): 475-506.
- [75] da Silveira, N. J. F., Uchoa, H. B., Canduri, F., Pereira, J. H., Camera Jr., J. C., Basso, L. A., Palma, M. S., Santos, D. S., de Azevedo Jr., W. F. (2004) *Biochem Biophys. Res. Commun.*, **322**(1), 100-104.
- [76] Ducati, R. D., Basso, L. A., Santos, D. S. (2007) *Current Drug Targets*, **8**(1) In press.
- [77] Dias, M. V. B., Faim, L. M., Vasconcelos, I. B., Oliveira, J. S., Basso, L. A., Santos, D. S., De Azevedo, W. F. *Acta Crystallographica F*. (In press).
- [78] De Azevedo, W.F.Jr.; Mueller-Dieckmann, H.J.; Schulze-Gahmen, U.; Worland, P.J.; Sausville, E.; Kim S.-H. (1996) *Proc. Natl. Acad. Sci. USA.*, **93**(7), 2735-2740.
- [79] Canduri, F.; Uchoa, H.B.; de Azevedo, W.F.Jr. (2004) *Biochem. Biophys. Res. Commun.*, **324**(2), 661-666.
- [80] De Azevedo, W.F.Jr.; Canduri, F.; Silveira, N.J.F. (2002) *Biochem Biophys. Res. Commun.*, **293**(1), 566-571.
- [81] De Azevedo, W.F.Jr.; Leclerc, S.; Meijer, L.; Havlicek, L.; Strnad, M.; Kim, S.-H. (1997) *Eur. J. Biochem.*, **243**, 518-526.
- [82] De Azevedo, W. F. Jr., Canduri, F. (2005) *Current Computer-Aided Drug Design*, **1**, 53-64.
- [83] Uchoa, H. B., Jorge, G. E., da Silveira, N. J., Camera J. C., Canduri, F., De Azevedo, W.F. (2004) *Biochem. Biophys. Res. Commun.* **325**(4), 1481-1486.
- [84] De Azevedo, W.F.Jr., Gaspar, R.T., Canduri, F.; Camera, J.C.Jr., Silveira, N.J.F. (2002) *Biochem Biophys. Res. Commun.*, **297**(5), 1154-1158.
- [85] da Silveira, N. J. F., Bonalumi, C. A., Uchoa, H. B., Pereira, J. H., Canduri, F., Pereira, J. H., de Azevedo Jr., W. F. (2006) *Cell Biochem. Biophys.* **44**(3), 366-374.
- [86] Kim, S.-H.; Schulze-Gahmen, U.; Brandsen, J.; De Azevedo, W. F. Jr. (1996) *Prog. in Cell Cycle Res.*, **2**, 137-145.