

José Henrique Pereira

Estudo estrutural de enzimas da via metabólica do ácido
chiquímico de *Mycobacterium tuberculosis*



1910001468



Estudo estrutural de enzimas da via metabólica do ácido
chiquímico de *Mycobacterium tuberculosis*

José Henrique Pereira

Orientador: Prof. Dr. Walter Filgueira de Azevedo Júnior



Tese apresentada para obtenção do grau de Doutor
em Biofísica Molecular, área de concentração em
Biofísica Molecular do Programa de Pós-Graduação
do Departamento de Física do Instituto de
Biociências, Letras e Ciências Exatas da
Universidade Estadual Paulista "Julio de Mesquita
Filho" – UNESP.

São José do Rio Preto / Outubro de 2005

577.112

P436c

13/03/2006

Doação

Autor(a)

20901003

Pereira, José Henrique.

Estudo estrutural de enzimas presentes na via metabólica do ácido chiquímico de *Mycobacterium tuberculosis* / José Henrique Pereira. – São José do Rio Preto : [s.n.], 2005
66 f. : il. ; 30 cm.

Orientador: Walter Filgueira de Azevedo Júnior

Tese (doutorado) – Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas

1. Cristalografia de raio X. 2. Modelagem molecular. 3. Bioinformática estrutural. 4. Chiquímico. 5. Biologia estrutural. 6. *Mycobacterium tuberculosis*. I. Azevedo Júnior, Walter Filgueira. II. Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU – 577.112



Agradecimentos

Ao Prof. Dr. Walter Filgueira de Azevedo Júnior pela orientação, confiança, incentivo e oportunidade de trabalho.

Aos professores Dr. Diego Luis Santiago Santos e Dr. Luis Augusto Basso responsáveis pela Coord. de Pesquisa em Biologia Molecular e Funcional (PUC-RS) que gentilmente acolheram as propostas envolvidas neste trabalho.

Aos Drs. Fernando Coutinho pelo apoio financeiro recebido durante todo o período de iniciação científica no presente momento.

O presente trabalho foi desenvolvido entre fevereiro/2002 e outubro/2005, junto ao Departamento de Física do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Julio de Mesquita Filho” – UNESP, sob a orientação do Professor Dr. Walter Filgueira de Azevedo Júnior.

Aos colegas do Laboratório de Sistemas Bimoleculares (Departamento de Física – IBFCEx): Alessandra, Bena, Bruno, Bruno, Danilo, Denize, Diogo, Fabiano, Joana, José, Rogério, João, Leandro, Lora, Marcos, Mariana, Marisa, Michelli, Nathalia, Nelson, Pepen.

José Henrique Pereira

Aos Prof. do Departamento de Física do IBFCEx – UNESP

A todos aqueles que ajudaram neste trabalho

A FAPESP, CAPES e CNPq pelo apoio financeiro e projetos



Agradecimentos

- Ao Prof. Dr. Walter Filgueira de Azevedo Júnior pela orientação, confiança, incentivo e oportunidade de trabalho.
- Aos professores Dr. Diógenes Santiago Santos e Dr. Luis Augusto Basso responsáveis pelo Centro de Pesquisas em Biologia Molecular e Funcional (PUC-RS) que gentilmente cederam as proteínas envolvidas neste trabalho.
- À Prof. Dra. Fernanda Canduri pela amizade e ensinamentos passados durante todo o período de iniciação científica até presente momento.
- Ao Dr. Jaim Simões de Oliveira pela fundamental contribuição para o desenvolvimento deste projeto.
- Aos colegas do Laboratório de Sistemas Biomoleculares (Departamento de Física - IBILCE): Alessandra, Bona, Briana, Daiane, Danilo, Denise, Diego, Helen, Joane, José Renato, Júlio, Lisandra, Livia, Márcio, Marcos, Marisa, Michelli, Nathalia, Nelson, Pepeu.
- Aos Prof. do Departamento de Física do IBILCE/UNESP.
- À minha Família pelo apoio incondicional.
- À FAPESP, CAPES e CNPq/MCT pela bolsa e apoio aos projetos.



Resumo

A tuberculose ressurgiu na metade dos anos 80 e atualmente, 2 milhões de pessoas morrem por ano devido a esta doença. O ressurgimento da tuberculose tornou-se uma ameaça à saúde pública. A alta susceptibilidade dos pacientes infectados com HIV e a proliferação de cepas resistentes a múltiplas drogas têm criado a necessidade de desenvolver novas terapias. No entanto, nenhuma nova classe de drogas contra a tuberculose foi desenvolvida nos últimos 30 anos. Deste modo, existe a necessidade de desenvolver rapidamente novos agentes anti-tuberculose. As enzimas da via metabólica do ácido chiquímico são alvos em potencial para o desenvolvimento de agentes antimicrobianos não-tóxicos e herbicidas, pois esta via é essencial para bactéria e plantas, enquanto que está ausente em mamíferos.

A via do ácido chiquímico é composta por sete passos metabólicos, catalisados pelas enzimas: DAHP sintase, 3-desidroquinato sintase, 3-desidroquinato desidratase, chiquimato desidrogenase, chiquimato quinase, EPSP sintase, e corismato sintase. Este trabalho apresenta o estudo através da cristalografia de raios X da chiquimato quinase, assim como a modelagem molecular da EPSP sintase e da corismato sintase.

Os resultados obtidos neste trabalho a partir do estudo estrutural das enzimas da via do ácido chiquímico presentes em *Mycobacterium tuberculosis*, usando modelagem molecular e cristalografia, devem contribuir para o entendimento do mecanismo catalítico destas enzimas e também para desenvolvimento de novos inibidores que futuramente poderão ser usados como agentes anti-tuberculose.



Abstract

Tuberculosis resurged in the mid-1980s and now kills approximately 2 million people a year. The reemergence of tuberculosis as a public health threat, the high susceptibility of HIV-infected persons, and the proliferation of multi-drug-resistant strains have created a need to develop new therapies. However, no new classes of drugs for TB have been developed in the past 30 years. There is an urgent need to developing faster acting and effective new anti-tubercular agents. The enzymes in the shikimate pathway are potential target for development of non-toxic antimicrobial agents, and herbicides because it is essential in bacteria, and plants, whereas it is absent from mammals.

The shikimate pathway is the seven-step biosynthetic route catalyzed by enzymes: 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHP synthase), 3-dehydroquinate synthase, 3-dehydroquinate dehydratase, shikimate dehydrogenase, shikimate kinase, 5-enolpyruvyl shikimate 3-phosphate synthase (EPSP synthase), and chorismate synthase. This work presents the X-ray crystallographic studies of shikimate kinase, as well as the molecular modeling of EPSP synthase and chorismate synthase.

The results obtain with this work through structural studies from enzymes in the shikimate pathway will provide crucial information for elucidation of the mechanism of catalyzed reaction and for the development of a new generation of drugs against tuberculosis.



Lista de Figuras

Figura 1.1 – Microscopia eletrônica de varredura do <i>Mycobacterium tuberculosis</i> .	01
Figura 1.2 – A via do ácido chiquímico. Fosfoenol-piruvato e eritrose 4-fosfato (precursores) são convertidos em ácido corísmico (corismato).	06
Figura 1.3 – Estrutura molecular do glifosato ($C_3H_8NO_5P$).	08
Figura 3.1 - Representação esquemática da gota pendurada (hanging drop).	15
Figura 3.2 A e B - Cristais de MtCQ-MgADP-Chiquimato com tamanho entre 0,35 mm e 0,4 mm.	17
Figura 3.3 – Gráficos de Ramachandran: A) MtCQ-MgADP-Chiquimato, B) MtCQ-MgADP (1L4Y), C) MtCQ-MgADP (1L4U).	20
Figura 3.4 – Elementos de estrutura secundária da MtCQ.	21
Figura 3.5 – Alinhamento das chiquimato quinases mostrando que os resíduos envolvidos na ligação do chiquimato são conservados.	22
Figura 3.6 – Coordenação do Mg^{+2} . A) 1L4Y, B) 1L4U, C) MtCQ-MgADP-Chiquimato.	25
Figura 3.7 – Mapa $F_{obs}-F_{calc}$ (3σ) gerado para posicionar o ácido chiquímico na estrutura da MtCQ. A figura mostra também os $C\alpha$ da MtCQ-mgADP-Ácido chiquímico, juntamente com as moléculas de ADP e Mg^{2+} .	26
Figura 3.8 – Estrutura molecular do chiquimato mostrando o grupo carboxila e os três grupos hidroxila.	27
Figura 3.9 – Interações envolvidas na ligação chiquimato no sítio ativo da MtCQ.	29
Figura 3.10 – Ligação do íon cloreto no sítio ativo das estruturas da MtCQ-MgADP-Chiquimato e MtCQ-MgADP (1L4U; Gu <i>et al.</i> , 2002).	31

Figura 3.11 – Superposição da MtCQ-MgADP-Chiquimato (vermelho) e o mutante K15M/P115L da EcCQ (azul) (Krell <i>et al.</i> , 2001) mostrando o grande deslocamento do domínio LID e do domínio SB devido a ligação do ADP e do ácido chiquímico.	33
Figura 3.12 – Superposição dos C α do domínio LID e do domínio SB das estruturas MtCQ-MgADP-Chiquimato (cinza) e MtCQ-MgADP (verde) (1L4U; Gu <i>et al.</i> , 2002).	35
Figura 3.13 – Superfície molecular das estruturas (A) MtCQ-MgADP (1L4U; Gu <i>et al.</i> , 2002) e (B) MtCQ-MgADP-Chiquimato.	36
Figura 3.14 – Reação catalisada pela EPSPS sintase.	39
Figura 3.15 – Passos para construção de um modelo utilizando a modelagem molecular comparativa.	41
Figura 3.16 – Alinhamento das seqüências de aminoácidos da MtEPSPS e da EcEPSPS.	41
Figura 3.17 – (A) MtEPSPS sem ligantes – estrutura aberta e (B) MtEPSPS complexada com chiquimato-3-fosfato e glifosato – estrutura fechada.	42
Figura 3.18 – Sobreposição da MtEPSPS aberta (linha grossa) e fechada (linha fina) mostrando a diferença conformacional entre as estruturas devido a ligação do chiquimato-3-fosfato e do glifosato.	44
Figura 3.19 – (A) Sítio de ligação do chiquimato-3-fosfato e do (B) glifosato da MtEPSPS.	45
Figura 3.20 – Reação catalisada pela corismato sintase.	48
Figura 3.21 – O alinhamento das seqüências da MtCS e SpCS. Os elementos de estrutura secundária para o modelo MtCS são indicados de acordo com a região onde são encontrados.	51
Figura 3.22 – Diagrama de Ribbon do modelo molecular da MtCS complexado com FMN e EPSP. A) Forma monomérica e B) Forma oligomérica (tetrâmero).	53
Figura 3.23A – Sítio de ligação da FMN.	54
Figura 3.23B – Sítio de ligação do EPSP.	55

Lista de Tabelas

Tabela 1.1 – Fármacos utilizados no tratamento de tuberculose.	04
Tabela 3.1. Estruturas cristalográficas da chiquimato quinase previamente depositadas.	13
Tabela 3.2 – Estatística do processamento de dados de difração de raios X para MtCQ-MgADP-Chiquimato.	17
Tabela 3.3 – Estatística do refinamento para MtCQ-MgADP-Chiquimato.	19
Tabela 3.4 – Ligação do Mg^{2+} nas MtCQs.	26
Tabela 3.5 – Ligações de hidrogênio direta ou mediada por água entre o chiquimato e MtCQ.	29
Tabela 3.6 – Ligação do íon cloreto nas estruturas da MtCQ-MgADP (1L4Y e 1L4U) e MtCQ-MgADP-Chiquimato.	32
Tabela 3.7 – Ligações de hidrogênios entre MtEPSP e chiquimato-3-fosfato.	45
Tabela 3.8 – Ligações de hidrogênio entre MtEPSP e glifosato.	46
Tabela 3.9 – Resultados das análises estruturais da MtCS e SpCS.	52
Tabela 3.10 – Porcentagem dos elementos de estrutura secundária calculado para os dados experimentais de CD e para o modelo molecular da MtCS.	57

Glossário

ADP – Adenosina difosfato

AK – Adenilato quinase

aroA – gene codificador da EPSP sintase

aroB – gene codificador da 3-desidroquinato sintase

aroD – gene codificador da 3-desidroquinato desidratase

aroE – gene codificador da chiquimato desidrogenase

aroF – gene codificador da corismato sintase

aroG – gene codificador da DAHP sintase

aroK – gene codificador da chiquimato quinase (isoforma I)

aroL – gene codificador da chiquimato quinase (isoforma II)

ATP – Adenosina trifosfato

CQ – Chiquimato quinase

CS – Corismato sintase

3D – Tridimensional

DAHP – 3-deoxi-D-arabino-heptulosonato 7-fosfato

DHFR – Disidrofolato redutase

DHQase – Desidroquinase

E. coli – *Escherichia coli*

EcEPSPS – EPSPS de *Escherichia coli*



EPSP – 5-enolpiruvilchiquimato-3-fosfato

EPSPS – 5-enolpiruvilchiquimato-3-fosfato sintase

ErcCQ – Chiquimato quinase de *Erwinia chrysanthemi*

FDA – *Food and Drug Administration*

FMN – Flavina mononucleotídeo

gaps – intervalos na sequência de aminoácidos

HIV – Vírus da imunodeficiência adquirida humana

K_m – Constante de Michaelis-Menten

MTB – *Mycobacterium tuberculosis*

MtCQ – Chiquimato quinase de *Mycobacterium tuberculosis*

MtCS – Corismato sintase de *Mycobacterium tuberculosis*

MtEPSPS – EPSPS de *Mycobacterium tuberculosis*

NMP – Nucleosídeo monofosfato

PAS – ácido *p*-aminossalicílico

PDB – *Protein Data Bank*

PEG – Polietilenoglicol

PEP – Fosfoenol-piruvato

P-loop – Alça de ligação do fosfato

RMSD – *Root mean square deviation* (Desvio médio quadrático)

SpCS – Corismato sintase de *Streptococcus pneumoniae*

TB – Tuberculose

TyrR – regulador de expressão da CQ II de *E. coli*



Índice

Agradecimentos	iii
Resumo	iv
Abstract	v
Lista de Figuras	vi
Lista de Tabelas	ix
Glossário	x
1. Introdução	01
1.1 A tuberculose	01
1.2 Via metabólica do ácido chiquímico	05
2. Objetivos	09
3. Resultados	10
3.1 Estrutura cristalográfica da chiquimato quinase	10
3.1.1 Materiais e Métodos	14
3.1.2 Resultados e discussão	16
3.1.3 Conclusão	36
3.2 Modelagem molecular da EPSP sintase	39
3.2.1 Materiais e Métodos	40



3.2.2 Resultados e discussão	41
3.3 Modelagem molecular e análise de CD da corismato sintase	48
3.3.1 Materiais e Métodos	49
3.3.2 Resultados e discussão	50
4. Referências	58

Anexos

- 1 - **Pereira, J.H., Oliveira, J.S., Canduri, F., Dias, M.B.V., Palma, M.S., Basso, L.A., Santos, D.S., De Azevedo, W. F. Jr. (2004).** Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. *Acta Crystallographica Section D*. **60**, 2310-2319.
- 2 - **Pereira, J.H., Vasconcelos, I.B., Oliveira, J.S., Basso, L.A., Santos, D.S.** Shikimate Kinase: A potential target for development of novel anti-tubercular agents (*Review*). *Current Drug Targets*. Submetido em Maio de 2005.
- 3 - **De Azevedo, W. F. Jr., 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*. *Biochemical and Biophysical Research Communications*. **295**, 142-148.
- 4 - **Pereira, J.H., Canduri, F., Oliveira, J.S., da Silveira, N.J.F., Basso, L.A., Palma, M.S., De Azevedo, W. F. Jr., Santos, D.S. (2003).** Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*. *Biochemical and Biophysical Research Communications*. **312** (3), 608-614.



- 5 - Pereira, J.H., Dias, M.V.B., Canduri, F., Ely, F., Ruggiero, J. N., Frazzon, J., Basso, L.A, Palma, M.S., Santos, D.S., De Azevedo, W. F. Jr. Molecular modeling and CD analysis of chorismate synthase.



Figura 1.1 - Microscopia eletrônica de varredura de *Mycobacterium tuberculosis*. Os bacilos aparecidos 1 e 2 são de 2,2 μ m e 2,2-2,9 μ m de comprimento (Bordas et al., 1994).

Adaptado por autorização da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) de 1992. Robert Koch descobriu o agente causador da tuberculose.

1. Introdução

1.1 A tuberculose

As doenças infecciosas ainda são responsáveis pelo sofrimento e morte de centenas de milhões de pessoas e 90% das mortes causadas por esse tipo de enfermidade ocorrem em áreas tropicais e subtropicais, regiões nas quais estão presentes 80% de toda população mundial (Trouiller *et al.*, 2001). Entre estas doenças destacam-se a AIDS/HIV, infecções respiratórias, malária e tuberculose. A infecção pelo *Mycobacterium tuberculosis*¹, agente causador da tuberculose, é a principal causa de morte em humanos devido a um único agente infeccioso. A tuberculose também pode ser causada por outras bactérias do gênero *Mycobacterium*, como por exemplo, o *Mycobacterium africanum*, o *Mycobacterium microti* e o *Mycobacterium bovis*.

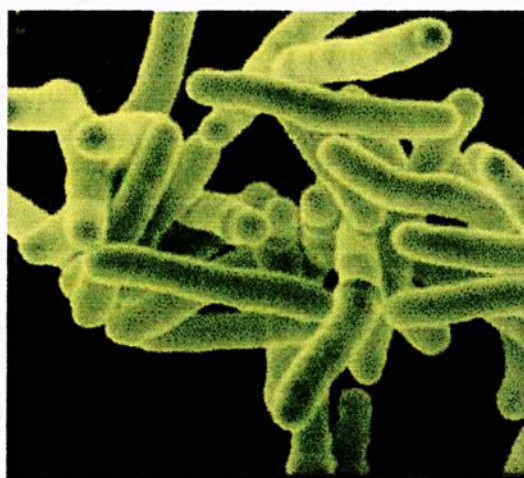


Figura 1.1 – Microscopia eletrônica de varredura do *Mycobacterium tuberculosis*. Os bacilos apresentam 1-4 μm de comprimento e 0,3-0,6 μm de diâmetro (Pelczar *et al.*, 1996).

¹ *Mycobacterium tuberculosis* também é conhecido como bacilo de Koch, pois em 24 de Março de 1882, Robert Koch identificou o agente causador da tuberculose.

De acordo com a Organização Mundial da Saúde, um terço da população mundial está infectada, e a estimativa é que 10% desta fração desenvolva a doença ao longo de sua vida (WHO, 2004a; Bloom & Murray, 1992). Anualmente, são mais de 8 milhões de pessoas que desenvolvem a tuberculose ativa, resultando em uma taxa de 2 milhões de mortes/ano (WHO, 2004a; 2004b). A cada dia, mais de 25 mil pessoas ficam doentes, e 5 mil pessoas acabam morrendo (WHO, 2003). No Brasil, de acordo com o relatório da Funasa², entre os anos de 1980 a 2000 a tuberculose apresentou uma ocorrência de 80 a 100 mil casos/ano, resultando em uma taxa de óbitos em torno de 6 mil mortes/ano³ (FUNASA, 2002).

Apesar da tuberculose ser considerada por muitos como uma doença do século XIX, na verdade ela nunca foi erradicada. Nos últimos anos tem-se verificado um aumento em sua incidência: entre 1997 e 2000, o número de novos casos de tuberculose aumentou 1,8% ao ano (Corbett *et al.*, 2003). Estima-se que entre 2002 e 2020, se não houver um esforço para seu controle, aproximadamente 1 bilhão de pessoas serão infectadas, mais de 150 milhões ficarão doentes, e a tuberculose levará à morte mais de 38 milhões de pessoas (WHO, 2001; 2002). Dada a gravidade da situação global da doença, em 1993 a Organização Mundial da Saúde declarou a tuberculose um sério problema de saúde pública mundial, como forma de alertar os setores de saúde pública e os governos para a necessidade de contenção desta doença.

A transmissão do bacilo ocorre via aérea, devido à inalação de pequenas partículas de aerossóis expelidas através da fala, espirro e, principalmente, pela tosse dos

² FUNASA – Fundação Nacional de Saúde.

³ Oficialmente, a doença mata a cada ano 5 a 6 mil pessoas no Brasil, mas especialistas acreditam que o número de óbitos deve chegar a 10 mil (Pivetta, 2004).



pacientes com a forma pulmonar da doença. Estas partículas são inaladas pelo indivíduo sadio, atingindo o espaço alveolar, onde iniciam sua multiplicação e são fagocitadas pelos macrófagos do hospedeiro (Hiriyanna & Ramakrishnan, 1986). Como o *Mycobacterium tuberculosis* (MTB) é capaz de resistir ao ataque, sobreviver e multiplicar-se no interior de células fagocitárias, como os macrófagos, o MTB é considerado um parasita intracelular (Clemmens, 1996).

À medida que macrófagos, monócitos e células T são recrutados para a área em torno do bacilo, ele diminui sua replicação e permanece em estado latente até um momento de deficiência do sistema imunológico. O MTB normalmente sobrevive durante anos neste estado de dormência dentro dos tecidos (Manabe & Bishai, 2000).

A incidência de tuberculose teve um declínio rápido no início do século vinte nos países em desenvolvimento, devido à melhora nas condições sanitárias e das moradias. Essa tendência foi acelerada inicialmente pela introdução da vacinação BCG⁴ (1921) e da descoberta de antibióticos como a estreptomicina (1944) e, posteriormente, com a descoberta do ácido *p*-aminossalicílico (PAS) (1946), isoniazida (1952) e rifampicina (1965) (Drobniewski *et al.*, 2003; Duncan, 2003).

Até o início dos anos 80, o tratamento com estes quimioterápicos promoveu o controle e o declínio da doença. Entretanto dois principais fatores têm contribuído para aumento no número de casos de tuberculose. O primeiro é a suscetibilidade de pacientes co-infectados com o vírus HIV, cujo sistema imunológico enfraquecido não pode controlar o crescimento do bacilo. Estes pacientes apresentam um risco cem

⁴ A Vacina BCG (Bacilo de Calmette-Guérin) contra a tuberculose foi desenvolvida pelo Dr. Albert Calmette, juntamente com seu assistente Camille Guérin. Os primeiros indivíduos (120 recém nascidos de mães tuberculosas) receberam a vacina em Julho de 1921 (Hawgood, 1998).



vezes maior de desenvolver a doença (El Syed *et al.*, 2000). O outro fator é a emergência de cepas resistentes aos antimicrobianos de primeira linha (isoniazida e rifampicina) utilizados no tratamento, devido a terapias inadequadas e o uso indiscriminado destes antibióticos (Baptista *et al.*, 2002; Duncan, 2003).

O tratamento contra tuberculose deve incluir uma ação bactericida contra o bacilo ativo, seguida da esterilização da população de bacilos dormentes (Mitchison, 1985). Atualmente, existem 10 medicamentos que estão aprovados pela *Food and Drug Administration* (FDA) dos E.U.A. (Villar, 2004). A isoniazida, rifampicina, pirazinamida e o etambutol são considerados como agentes terapêuticos de 1ª linha. A Cicloserina, Etionamida, o ácido *p*-aminosalicílico, capreomicina e a estreptomicina são considerados fármacos de 2ª linha. A estreptomicina, que pertenceu durante muitos anos aos antibacilares de 1ª linha, atualmente, devido ao cada vez maior número de resistências, viu muito diminuída a sua utilidade, tendo sido relegada para 2ª linha. A Rifabutina é aprovada na utilização para tratamento da tuberculose de pacientes infectados com vírus HIV.

Tabela 1.1 – Fármacos utilizados no tratamento da tuberculose.

1ª linha	2ª linha
Isoniazida	Cicloserina
Rifampicina	Etionamida
Pirazinamida	Levofloxacina [#]
Etambutol	Moxifloxacina [#]
Rifabutina*	Gatifloxacina [#]
	PAS
	Estreptomicina
	Amicacina / Canamicina [#]
	Capreomicina

* Utilizado em situações particulares

[#] Fármacos usados para tratamento de tuberculose multiresistente, apesar de não serem aprovados pelo FDA.



O tratamento recomendado pela Organização Mundial da Saúde apresenta uma duração mínima de seis meses. Nos dois primeiros meses, envolve a combinação de quatro antibióticos: isoniazida, rifampicina, pirazinamida e o etambutol, seguido da combinação de isoniazida e rifampicina por mais quatro meses (ou de isoniazida combinada com etionamida ou etambutol por mais seis meses).

Os mecanismos de resistência aos agentes anti-tuberculose são devido às alterações no DNA cromossomal, portanto estas cepas não estão sujeitas à seleção reversa, e permanecerão causando a falha de tratamentos padrões à tuberculose. Desta forma, o desenvolvimento de novas drogas para substituírem àquelas comprometidas pela resistência torna-se necessário para que se possa estabelecer um tratamento quimioterápico eficaz.

O principal objetivo da quimioterapia é atacar alvos peculiares aos microrganismos como, por exemplo, vias metabólicas ausentes no organismo humano. Sob este aspecto, as enzimas da via do ácido chiquímico representam bons exemplos de tal abordagem.

1.2 Via metabólica do Ácido Chiquímico

A via metabólica do ácido chiquímico foi descoberta através dos estudos de Bernhard Davis e David Sprinson (Davis & Mingioli, 1953; Sprinson, 1960). Esta rota biossintética faz a conexão entre o metabolismo de carboidratos e a síntese de compostos aromáticos através de sete passos metabólicos, onde o fosfoenol-piruvato e eritrose 4-fosfato são convertidos em ácido corísmico (Figura 1.2) (Pittard, 1987;



Haslam, 1993). O ácido corísmico é um precursor comum para a síntese de aminoácidos aromáticos, folato, ubiquinonas, menaquinonas e enterobactim.

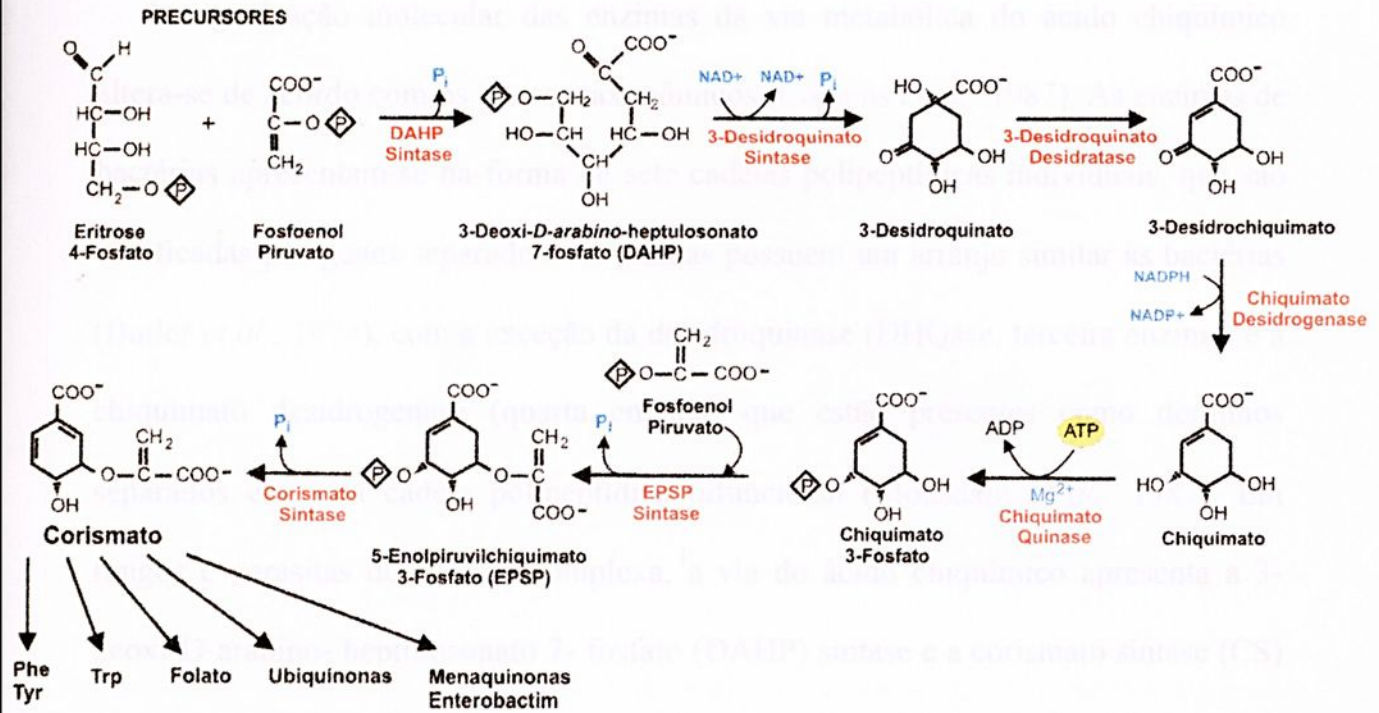


Figura 1.2 – A via do ácido chiquímico. Fosfoenol-piruvato e eritrose 4-fosfato (precursores) são convertidos em ácido corísmico (corismato). Corismato é essencial para síntese de aminoácido aromático (Phe, Tyr e Trp), folato, ubiquinonas, menaquinonas e enterobactim. Os sete passos são catalisados pelas enzimas: 3-deoxi-D-arabino-heptulosonato 7-fosfato sintase (DAHP sintase), 3-desidroquinato sintase, 3-desidroquinato desidratase, chiquimato desidrogenase, chiquimato quinase, 5-enolpiruvil chiquimato 3-fosfato sintase (EPSP sintase), e corismato sintase. Figura modificada a partir de Mathews & van Holde. (Mathews & van Holde, 1990).

As enzimas da via do ácido chiquímico são alvos potenciais para desenvolvimento de anti-microbianos (Davies *et al.*, 1994) e herbicidas (Coggins, 1989), pois esta via é essencial para algas, plantas superiores, bactérias, fungos, e inexistente em mamíferos (Bentley, 1990). Recentemente, a via do ácido chiquímico também foi identificada em parasitas do filo Apicomplexa (*Plasmodium falciparum*, *Toxoplasma gondii* e

Cryptosporidium parvum), fornecendo novos alvos para o desenvolvimento de drogas anti-parasitas (Roberts *et al.*, 1998).

A organização molecular das enzimas da via metabólica do ácido chiquímico altera-se de acordo com os grupos taxonômicos (Coggins *et al.*, 1987). As enzimas de bactérias apresentam-se na forma de sete cadeias polipeptídicas individuais, que são codificadas por genes separados. As plantas possuem um arranjo similar às bactérias (Butler *et al.*, 1974), com a exceção da desidroquinase (DHQase, terceira enzima) e a chiquimato desidrogenase (quarta enzima) que estão presentes como domínios separados em uma cadeia polipeptídica bifuncional (Mousdale *et al.*, 1987). Em fungos e parasitas do filo Apicomplexa, a via do ácido chiquímico apresenta a 3-deoxi-D-arabino- heptulose 7- fosfato (DAHP) sintase e a corismato sintase (CS) como enzimas monofuncionais e um polipeptídeo pentafuncional chamado AROM, que é responsável pelas cinco reações restantes da via do ácido chiquímico (Duncan *et al.*, 1987). Especificamente em *Mycobacterium tuberculosis*, com a publicação do genoma do bacilo (Cole *et al.*, 1998), verificou-se a presença dos genes codificadores das enzimas envolvidas na via do ácido chiquímico: *aroG* (DAHP sintase), *aroB* (3-desidroquinato sintase), *aroD* (3-desidroquinato desidratase), *aroE* (chiquimato desidrogenase), *aroK* (chiquimato quinase), e *aroA* (EPSP sintase), *aroF* (corismato sintase).

A utilização das enzimas da via biossintética do ácido chiquímico como alvos para o desenvolvimento de inibidores pode ser plenamente justificada e validada pela inibição da EPSP sintase através do glifosato (*N*-(fosfometil)glicina) (Kishore & Shah, 1988). Glifosato é um importante composto guia para a construção de novos



inibidores de EPSP sintase. Este composto atualmente é utilizado como herbicida, e está presente nos defensivos agrícola *Round Up* e *TouchDown*, usados no controle de ervas daninhas. Roberts e seus colaboradores demonstraram também que o glifosato é capaz de inibir *in vitro* o crescimento dos parasitas *Plasmodium falciparum*, *Toxoplasma gondii* e *Cryptosporidium parvum* (Roberts *et al.*, 1998).

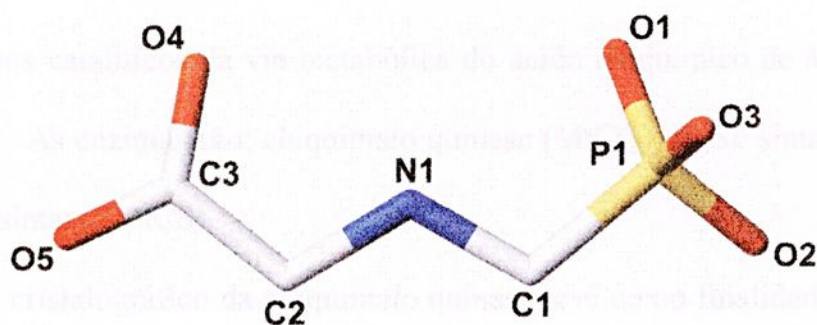


Figura 1.3 – Estrutura molecular do glifosato ($C_3H_8NO_5P$). Este composto exemplifica o grande potencial de desenvolvimento de novos inibidores baseados nas enzimas da via metabólica do ácido chiquímico. As coordenadas atômicas do glifosato em complexo com a EPSPS foram obtidas no PDB (código PDB:1G6S) (Schönbrunn *et al.*, 2001). Figura gerada utilizando o programa MolMol (Koradi *et al.*, 1996).

Com o objetivo de demonstrar a importância da via do ácido chiquímico em *Mycobacterium tuberculosis*, foi realizado a disrupção do gene *aroK*, que codifica a chiquimato quinase. Os resultados obtidos através destes experimentos confirmaram que a via do ácido chiquímico é essencial para viabilidade do bacilo (Parish & Stoker, 2002).

Uma vez validado a via como um potencial alvo para inibir o desenvolvimento do *Mycobacterium tuberculosis*, estudos estruturais das enzimas através de modelagem molecular e cristalografia foram realizados com o objetivo de fornecer informações que possam ser utilizadas para desenvolvimento de novos inibidores.

3. Resultados

2. Objetivos

Este trabalho teve como objetivo realizar o estudo estrutural através da cristalografia de raios X e modelagem molecular das enzimas responsáveis pelos três últimos passos catalíticos da via metabólica do ácido chiquímico de *Mycobacterium tuberculosis*. As enzimas são: chiquimato quinase (MtCQ), EPSP sintase (MtEPSPS) e corismato sintase (MtCS).

O estudo cristalográfico da chiquimato quinase teve como finalidade determinar o modo de interação do ácido chiquímico (chiquimato) com a enzima, pois nenhuma estrutura depositada de CQ apresentava seu substrato complexado (Krell *et al.*, 1998; Krell *et al.*, 2001; Gu *et al.*, 2002). Uma vez que pretendíamos auxiliar no desenvolvimento de novos inibidores para CQ, o conhecimento detalhado do sítio ativo da CQ era de fundamental importância para atingirmos nosso objetivo.

Experimentos de cristalização com MtEPSPS foram realizados, mas até o presente momento não foi obtido cristais desta enzima. Assim, a modelagem molecular da MtEPSPS apo enzima e MtEPSPS complexada com chiquimato-3-fosfato e glifosato foi realizada para extrair informações do modo que esses ligantes interagem com a enzima. Da mesma maneira, estudo por modelagem molecular da MtCS complexada com seu substrato (EPSP) e FMN (cofator) foi realizado, assim como experimentos de CD para reforçar o modelo que estava sendo proposto.

3. Resultados

3.1 Estrutura cristalográfica da chiquimato quinase complexada com MgADP e ácido chiquímico (chiquimato)

A chiquimato quinase (EC 2.7.1.71), quinta enzima da via, catalisa a fosforilação do grupo 3-hidroxila do chiquimato usando ATP como co-substrato (Figura 1.2). Em *Escherichia coli*, a reação da CQ é catalisada por duas isoformas: CQ I que é codificada pelo gene *aroK* (Whipp & Pittard, 1995) e CQ II que é codificada pelo gene *aroL* (Millar *et al.*, 1986). A principal diferença entre as isoenzimas são os valores de *K_m* encontrados para chiquimato, 20mM para a CQ I e 0,2mM para a CQ II. A isoforma CQ II apresenta um papel dominante na via do ácido chiquímico, onde a sua expressão é controlada pelo regulador *tyrR* e a sua repressão é realizada pela tirosina e triptofano (Ely & Pittard, 1979; De Feyter *et al.*, 1986). O papel fisiológico da CQ I em *E.coli* não está claro. Mutações em CQ I foram associadas com a sensibilidade ao antibiótico mecilinam (Vinella *et al.*, 1996), sugerindo que CQ I deve apresentar um papel biológico alternativo e que promove a fosforilação do chiquimato somente casualmente (De Feyter & Pittard, 1986).

Ao contrário de *E.coli*, seqüências de genomas completos de bactérias, como por exemplo *Haemophilus influenzae* e *Mycobacterium tuberculosis*, demonstraram a presença de somente um gene codificador de CQ. A maioria das CQs são codificadas pelo gene *aroK*, devido as seqüências de aminoácidos possuir um maior grau de identidade com a CQ I de *E. coli*. Os parâmetros cinéticos para MtCQ (codificada



pelo gene *aroK*) são mais similares a àqueles encontrados para CQ II *E. coli* do que aqueles de CQ I *E. coli*. A MtCQ apresenta um valor de *K_m* para chiquimato de 0,41mM indicando que nem todas as CQs codificadas pelo gene *aroK* possuem altos valores de *K_m* para substrato (Gu *et al.*, 2002).

A estrutura cristalográfica de CQ de *Erwinia chrysanthemi* (ErcCQ) (Krell *et al.*, 1998; Krell *et al.*, 2001), que é codificada pelo gene *aroL*, demonstrou que a chiquimato quinase é uma proteína da classe α/β que apresenta na sua estrutura fitas β rodeadas por hélices α . No centro da proteína estão localizadas as 5 fitas β paralelas (β 1- β 5) formando uma folha β na ordem 23145, que é flanqueada por oito hélices α . A ordem da folha β 23145 observada na estrutura da ErcCQ, classifica as CQs como pertencente à família das nucleosídeo monofosfato (NMP) quinases (Yan & Ysai, 1999). Esta família pode ser exemplificada pelas enzimas: hexoquinases (Bennett & Steitz, 1980), adenilato quinase (AK) (Dreusicke, *et al.*, 1988; Schlauderer & Schulz, 1996), guanilato quinase (Stehle & Schulz, 1990), uridilato quinase (Müller-Dieckmann & Schulz, 1994) e timidino quinase (Wild *et al.*, 1995).

As NMP quinases são compostas de três domínios: domínio central (*Core domain*), domínio da tampa (*LID domain*) e o domínio de ligação do NMP (*NMP-binding domain*). Uma característica que retrata as NMP quinases é que estas enzimas sofrem uma grande mudança conformacional durante a catálise (Vonnrhein *et al.*, 1995). Existem duas regiões flexíveis na estrutura que são responsáveis por este movimento: uma é o domínio de ligação do NMP que é formado por uma série de hélices entre as fitas- β 2 e 3, e a outra região é o domínio LID, uma região de



tamanho variado localizado após a quarta fita- β (β_4) (Müller *et al.*, 1996; Gerstein *et al.*, 1993). Em CQ, o domínio de ligação do chiquimato (Domínio SB) corresponde ao domínio de ligação do NMP de NMP quinases.

Três motivos funcionais presentes em enzimas de ligação de nucleotídeo são encontrados nas chiquimato quinases: motivo Walker A, motivo Walker B, e um motivo de ligação da adenina. O motivo Walker A está localizado entre a primeira fita- β (β_1) e a primeira hélice- α (α_1), contendo o trecho de sequência conservada GXXXXGKT/S, onde X representa qualquer resíduo (Walker *et al.*, 1982). Este motivo forma uma alça de ligação do fosfato (*P-loop*), onde diferentes amidas da cadeia principal fazem ligações de hidrogênio com β -fosfato do ADP (Smith & Rayment, 1996). Em MtCQ, este motivo compreende a região do resíduo 9 até 17, com a sequência GLPGSGKST. Em adição ao motivo Walker A, uma segunda sequência conservada ZZDXXG é chamada de motivo Walker B, onde Z representa resíduo hidrofóbico (Walker *et al.*, 1982). Um Asp $\rightarrow$ Ser é trocado em MtCQ. O consenso para o motivo Walker B em chiquimato quinases é ZZZTG β GGG, onde a segunda glicina (Gly80 em MtCQ) faz ligação de hidrogênio com o γ -fosfato do ATP. A alça de ligação da adenina, caracterizada pela sequência I/VDAXQ/NXP, foi descrita em Adenilato Quinase e ErcCQ. Em MtCQ, o motivo de ligação da adenina apresenta a sequência VDTNRRNP (resíduos 148 até 155). Portanto, o motivo de ligação da adenina poderia ser descrito mais precisamente pelo trecho I/VDXXX(X)XP (Gu *et al.*, 2002).

Os estudos cristalográficos realizados com ErcCQ (código de acesso PDB: 1SHK, 2SHK e 1E6C) (Krell *et al.*, 1998) não apresentaram densidade eletrônica suficiente



para incluir o chiquimato na estrutura. As duas estruturas cristalográficas de CQ de *Mycobacterium tuberculosis* resolvidas por Gu *et al* (Gu *et al.*, 2002) (código de acesso PDB: 1L4Y e 1L4U) revelou o movimento do domínio LID na catálise, mas a posição precisa do chiquimato e suas interações com a chiquimato quinase não foram determinadas. Embora as estruturas de ErcCQ e de MtCQ não apresentarem o chiquimato, o substrato foi adicionado em suas condições de cristalização (Tabela 3.1). A tabela 3.1 também mostra os resíduos e os ligantes que foram determinados nas estruturas de ErcCQ-MgADP e MtCQ-MgADP.

Tabela 3.1. Estruturas cristalográficas da chiquimato quinase previamente depositadas.

PDB	(Å)	Aditivos (mM)			Estr. cristalográficas das CQ com ligantes			
		Ác. Chiq.	ADP	MgCl ₂	Mol	Resid. depositados	Ác.Chiq	MgADP
1SHK ^a	1,9	5	5	10	A	1-112, 128-173	(1) ^b	
					B	1-112, 126-172		
2SHK ^a	2,6	5	5	10	A	1-112, 128-173	(1) ^b	
					B	1-112, 123-173		1
1E6C ^b	1,8	2,5	2,5	10	A	3-170		
					B	3-170		
1L4U ^d	1,8	4	4	4	--	2-166		1
1L4Y ^c	2,0	5	5	5	--	2-166		1

^aErcCQ – Chiquimato quinase de *Erwinia chrysanthemi* (Krell *et al.*, 1998)
^bMapas de densidade eletrônica *Fo-Fc* indicam uma possível posição do chiquimato (Krell *et al.*, 1998)
^capo-ErcCQ (K15M) (Krell *et al.*, 2001)
^dDerivativo-Pt da MtCQ (Gu *et al.*, 2002)
^eMtCQ nativa (Gu *et al.*, 2002)

Embora havendo estruturas de chiquimato depositadas no PDB, o modo de ligação do substrato com a enzima permanecia indeterminado. Deste modo, o estudo estrutural com a MtCQ teve como objetivo revelar detalhadamente o modo interação dos resíduos de aminoácidos com o chiquimato e as mudanças conformacionais que a



enzima sofre devido a ligação do substrato. A estrutura da MtCQ-MgADP-Chiquimato fornecerá informações importantes para o desenvolvimento de inibidores de CQ e também para entendimento do mecanismo catalítico.

3.1.1 Materiais e métodos

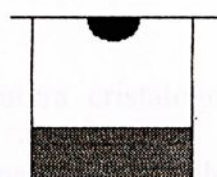
A MtCQ foi obtida a partir de métodos de clonagem, expressão e purificação pelo grupo do Prof. Dr. Diógenes Santiago Santos, da Rede Brasileira de Pesquisa em Tuberculose, do Departamento de Biologia Molecular e Biotecnologia (UFRGS), Porto Alegre, Brasil (Oliveira *et al.*, 2001).

Cristalização

A proteína, precipitada em sulfato de amônio, foi solubilizada e dialisada contra Hepes 50 mM (pH 7.5), NaCl 0,5M e MgCl₂ 5mM. A esta preparação foram adicionados chiquimato 13mM e ADP 13mM, e a mistura foi centrifugada por 5 minutos a 9.300 g. Os cristais foram obtidos usando uma variação do método de difusão de vapor, chamada de gota suspensa (*hanging drop*).

Para a cristalização do complexo MtCQ-MgADP-Chiquimato foi adicionado nos reservatórios de uma placa Linbro, 1 mL de uma solução composta de 0,1 M de Hepes (pH 7,5), 10 % de isopropanol e 35 % de PEG 3350. As gotas foram feitas sobre uma lamínula siliconizada de 22x22mm, na qual foi utilizado 1 µl da solução do reservatório e 1,5 µl da solução de proteína a uma concentração de 17 mg.ml⁻¹. Portanto, cada gota apresentava um volume final de 2,5 µl.





Hanging drop



Solução da macromolécula biológica



Reservatório com o agente de cristalização

Figura 3.1 - Representação esquemática da gota pendurada (*hanging drop*)

Coleta e processamento dos Dados

O conjunto de dados (160 imagens) da MtCQ-MgADP-Chiquimato foi coletado a um comprimento de onda de 1,431 Å, usando-se como fonte radiação síncrotron (Estação PCr, Laboratório Nacional de Luz Síncrotron, LNLS, Campinas, Brasil) (Polikarpov *et al.*, 1998a; Polikarpov *et al.*, 1998b), e um detector CCD (MARCCD) a uma distância de 90 mm do cristal. O protetor criogênico utilizado era composto glicerol 15 %, PEG 3350 12 %, Hepes 35mM, isopropanol 3,5 %, e a coleta foi mantida a uma temperatura de 104 K. O tempo de exposição foi de 50 segundos por imagem, usando um ângulo de oscilação de 1,0°. Os dados foram processados utilizando os programas MOSFLM (Leslie, 1992) e escalonados com o programa SCALA (Colaborative Computational Project nº 4, 1994).

Substituição Molecular e Refinamento cristalográfico

A estrutura cristalográfica da MtCQ-MgADP-Chiquimato foi determinada por métodos padrões de substituição molecular usando o programa AMoRe (Navaza, 1994), tendo sido usado como modelos de busca as estruturas da chiquimato quinase de *Mycobacterium tuberculosis* resolvida anteriormente (código de acesso PDB: 1L4Y e 1L4U) (Gu *et al.*, 2002). As posições atômicas obtidas por substituição molecular foram usadas para iniciar o refinamento cristalográfico.

Para o refinamento da estrutura foi utilizado o programa X-PLOR (Brünger, 1992). A qualidade estereoquímica do modelo final da MtCQ-MgADP-Chiquimato foi acessada pelo programa PROCHECK (Laskowski *et al.*, 1994). Modelos atômicos foram superpostos usando o programa LSQKAB presente no CCP4 (Colaborative Computational Project nº 4, 1994). A superfície molecular da MtCQ-MgADP-Chiquimato foi calculada usando o programa *Swiss PDB Viewer* v.3.7 (Guex & Peitsch, 1997).

3.1.2 Resultados e discussão

Os cristais de MtCQ-MgADP-Chiquimato foram obtidos na presença de 13mM chiquimato e 13mM ADP e apareceram após 1 dia a temperatura de 20° C (Figura 3.2). O conjunto de dados da MtCQ-MgADP-Chiquimato foi coletado a 2.3 Å usando fonte de radiação síncrotron (Tabela 3.2).

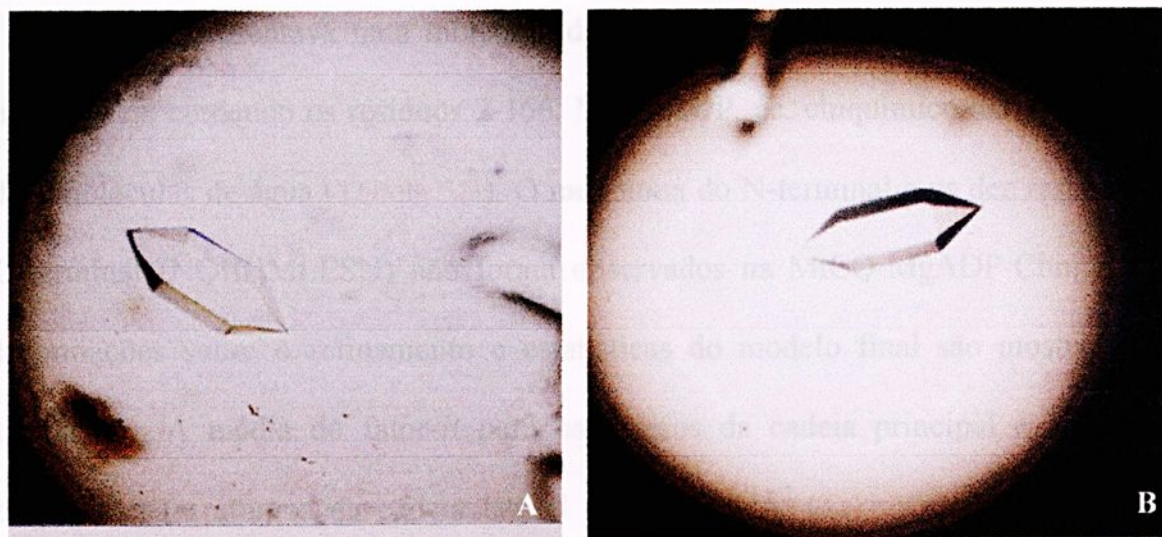


Figura 3.2 A e B - Cristais de MtCQ-MgADP-Chiquimato com tamanho entre 0,35 mm e 0,4 mm .

Tabela 3.2 – Estatística do processamento de dados de difração de raios X para MtCQ-MgADP-Chiquimato

a (Å)	62,91
b (Å)	62,91
c (Å)	90,92
$\alpha(^{\circ})$	90,00
$\beta(^{\circ})$	90,00
$\gamma(^{\circ})$	120,00
Grupo espacial	P3 ₂ 21
Número de medidas com $I > 2 \sigma(I)$	34.274
Número de reflexões independentes	9.563
Completeza (%)	98,7
R^*_{sym} (%)	3,0
Camada de mais alta resolução (Å)	2,41 – 2,30
Completeza à mais alta resolução (%)	98,7
R^*_{sym} à mais alta resolução (%)	7,20

* $R_{\text{sym}} = 100 \sum |I(h) - \langle I(h) \rangle| / \sum I(h)$ com $I(h)$, intensidade observada e $\langle I(h) \rangle$, intensidade de reflexões h sobre todas as medidas de $I(h)$.

O cristal apresentava uma molécula de MtCQ-MgADP-Chiquimato por unidade assimétrica, contendo os resíduos 2-166, Mg^{2+} , ADP, ác. chiquímico, dois íons Cl^- e 144 moléculas de água (Tabela 3.3). O metionina do N-terminal e os dez resíduos do C-terminal (NQIIHMLESN) não foram observados na MtCQ-MgADP-Chiquimato. Informações sobre o refinamento e estatísticas do modelo final são mostradas na tabela 3.3. A média do fator- B para os átomos da cadeia principal é $34,64 \text{ \AA}^2$, enquanto para átomos da cadeia lateral é de $35,68 \text{ \AA}^2$ (Tabela 3.3). Para obter a posição de ligação do chiquimato na MtCQ, uma maior concentração de substrato foi utilizada (13 mM) comparado com a usada por Gu et al. (Gu et al., 2002) para obter os cristais (5mM). A média do fator- B de $26,15 \text{ \AA}^2$ obtida para os átomos do chiquimato indica que este ligante está altamente ordenado na estrutura

Água	39,03
ADP	21,81
Mg^{2+}	34,20
Cl	17,65
Chiquimato	26,15
R.M.S.D. observado da geometria ideal	
Comprimento de ligação (Å)	0,017
Ângulo de ligação (graus)	1,895
Desvio (graus)	22,128
Gráfico de Ramachandran	
ϕ/ψ mais favorável (%)	93,4
ϕ/ψ não permitida (%)	0
$R_{\text{free}} = 100 \times (F_o - F_c) / \sum (F_o^2 + F_c^2)$ e $R_{\text{work}} = 100 \times (F_o - F_c) / \sum (F_o^2 + F_c^2)$ para 90% dos dados, os quais não foram incluídos no refinamento. Fator B = valor médio do fator de difração B médio para átomos de não-hidrogênio	

Tabela 3.3 – Estatística do refinamento para MtCQ-MgADP-Chiquimato

MtCQ-MgADP-Chiquimato	
Faixa de resol. usada (Å)	6,00 – 2,30
Reflexões usadas no ref.	8.885
Número de resíduos	165
Número de mol. de águas	144
Número de mol de ADP	1
Número de íons metal Mg ²⁺	1
Número de Cl ⁻	2
Número de moléculas de chiquimato	1
Final R_{factor} * (%)	20,7
Final R_{free} ** (%)	28,7
Fator B^+ (Å ²)	
Cadeia principal	34,64
Cadeias laterais	35,68
Águas	39,63
ADP	21,81
Mg ⁺²	34,70
Cl ⁻	37,65
Chiquimato	26,15
R.M.S.D observado da geometria ideal	
Comprimento de ligação (Å)	0,017
Ângulos de ligação (graus)	1,905
Diédros (graus)	22,125
Gráfico de Ramachandran	
Φ/Ψ mais favoráveis (%)	93,4
Φ/Ψ não permitidos (%)	0

* $R_{\text{factor}} = 100 \times \Sigma |F_{\text{obs}} - F_{\text{calc}}| / \Sigma (F_{\text{obs}})$, somas obtidas sobre todas as reflexões com corte de $F/\sigma(F) > 2$.

** $R_{\text{free}} = R_{\text{factor}}$ para 10% dos dados, os quais não foram incluídos no refinamento.

+ Fator B = valores médios do fator de vibração térmica para átomos de não-hidrogênios.

A qualidade estereoquímica da estrutura MtCQ-MgADP-Chiquimato foi analisada através do gráfico de Ramachandran (ângulos ϕ (ϕ) e ψ (ψ)) e apresentou 93,4% dos resíduos nas regiões mais favoráveis, e nenhum resíduo nas regiões não permitidas. As estruturas 1L4Y e 1L4U apresentaram, respectivamente, 90,5% e 94,2% dos resíduos nas regiões mais favoráveis e nenhum nas regiões não permitidas (Figura 3.3).

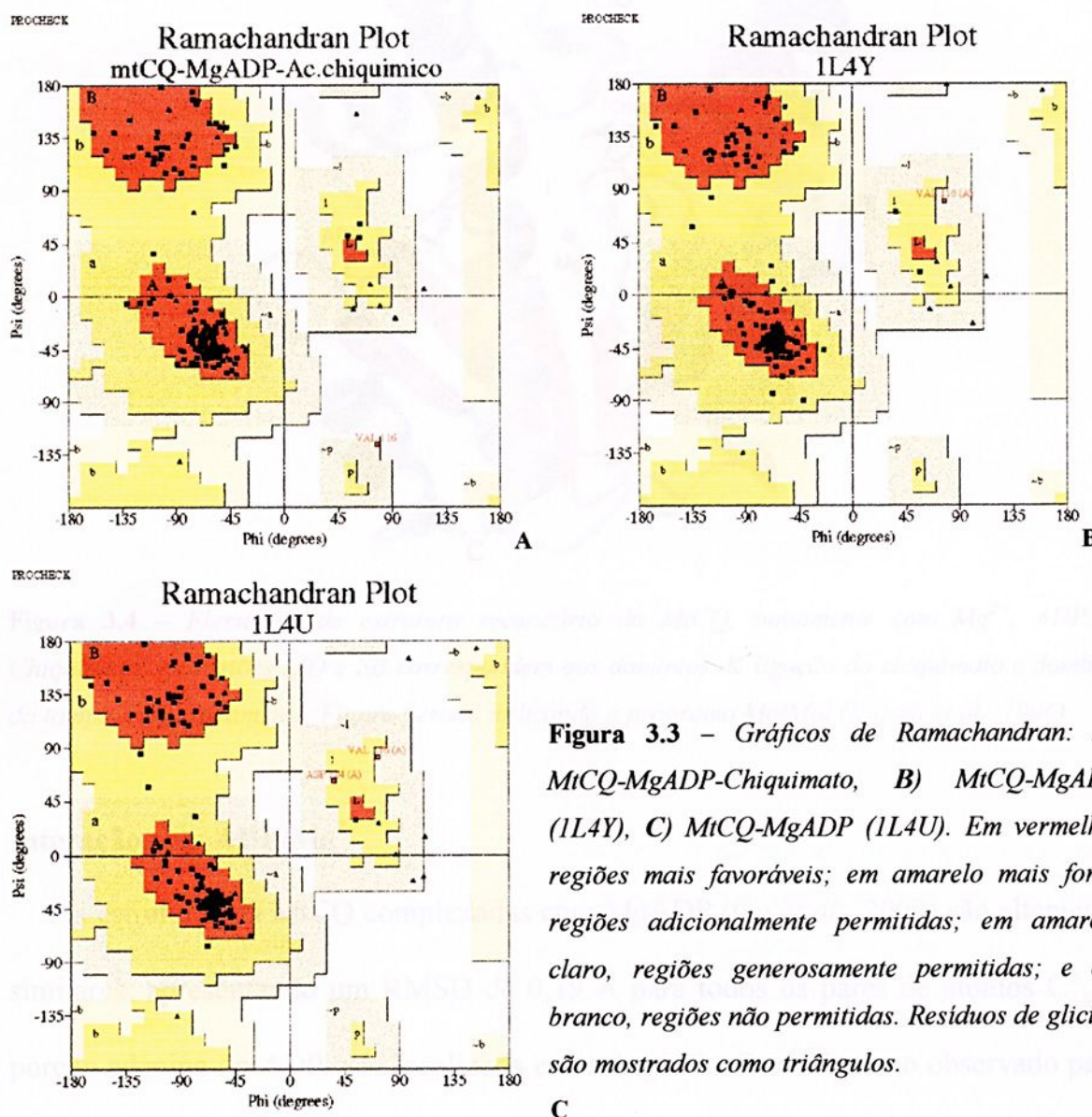


Figura 3.3 – Gráficos de Ramachandran: A) MtCQ-MgADP-Chiquimato, B) MtCQ-MgADP (1L4Y), C) MtCQ-MgADP (1L4U). Em vermelho, regiões mais favoráveis; em amarelo mais forte, regiões adicionalmente permitidas; em amarelo claro, regiões generosamente permitidas; e em branco, regiões não permitidas. Resíduos de glicina são mostrados como triângulos.

Como citado anteriormente, uma característica marcante da CQ é que esta enzima sofre uma grande mudança conformacional durante a catálise. As duas regiões responsáveis são o domínio de ligação do chiquimato (Domínio SB) e o domínio LID (Figura 3.4), que correspondem em MtCQ aos resíduos 33-61 e 112-124, respectivamente.

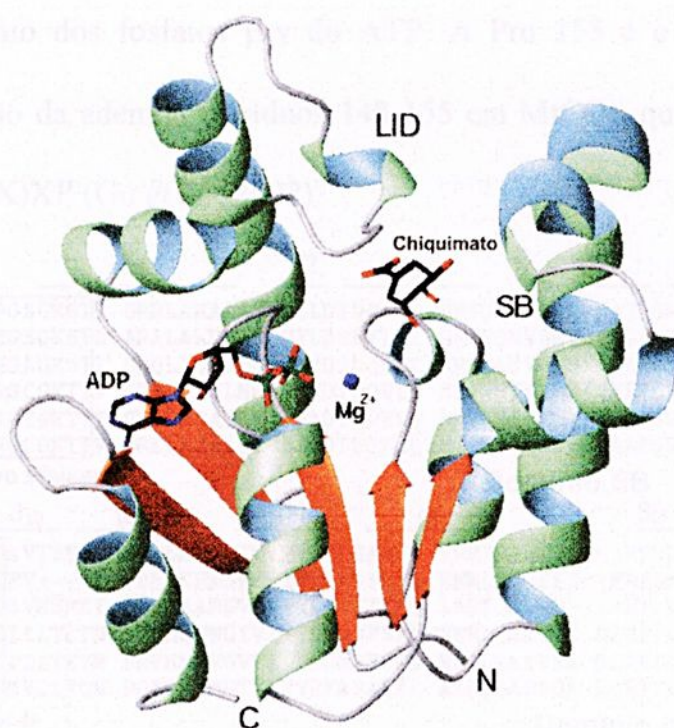


Figura 3.4 – Elementos de estrutura secundária da MtCQ, juntamente com Mg^{2+} , ADP, e Chiquimato. As regiões LID e SB correspondem aos domínios de ligação do chiquimato e domínio da tampa, respectivamente. Figura gerada utilizando o programa MolMol (Koradi et al., 1996).

Interação com ADP/ Mg^{2+}

As estruturas da MtCQ complexadas com MgADP (Gu et al., 2002) são altamente similares, apresentando um RMSD de 0,19 Å para todos os pares de átomos C^α . A porção adenina do ADP está localizada entre Arg110 e Pro155, como observado para as estruturas MtCQ-MgADP. A Arg110 em MtCQ representa o primeiro resíduo de

um motivo conservado do domínio LID observado para P-loop quinases (Leipe *et al.*, 2003). O segundo resíduo básico conservado deste motivo (Arg117) interage com os grupos α - e β -fosfato do ADP. Assim, este motivo conservado em CQ encontrado no domínio LID pode ser descrito como R(X)₆₋₉R (Figura 3.5). A Arg 117 deve estabilizar o estado de transição neutralizando a carga negativa formada entre os átomos de oxigênio dos fosfatos β - γ do ATP. A Pro 155 é o último resíduo do domínio de ligação da adenina (resíduos 148-155 em MtCQ), que pode ser descrito como I/VDXXX(X)XP (Gu *et al.*, 2002).

	$\beta 1$	$\alpha 1$	$\beta 2$	$\alpha 2$	$\alpha 3$	$\alpha 4$			
AK_MYCTU	--MAPKAVLV	GLPGSGKSTII	GRRLAKALGV	GLLDTDAVIE	QRTGRSIADI	FATDGEQEFR	RIREDVVRAA 68		
AK_CAMJE	MMKAKNIVFI	GFMGSGKSTL	APALAKDLDL	VFLDSDFLIE	QKFNQKVSEI	FEQKRENFRR	EQSQKMDFF 70		
AK_ECOLI	MAEKRNIPLV	GPMGAGKSTI	GRQLAQLNM	EFYDSQIEIE	KRTGADVGVV	FDLEGEEGFR	DRREKVINEL 70		
AL_ECOLI	--MTQPLFLI	GPRGCGKTTIV	GMALADSLNR	RFVDTDQWLQ	SQNMHTVAEI	VEREEWAGFR	ARSTAALEAV 68		
AL_SALTY	--MMQPLYLV	GPRGCGKTTI	GMALAQATGF	RFADTDRWLQ	SHVQMSVADI	VEKEGWGGFR	ARSTAALEAV 68		
AL_ERWCH	--MTEPIFMV	GARGCGKTTIV	GRELARALGY	EFVDTDIFMQ	HTSGMTVADV	VAAEGWPGFR	RRSEALQAV 68		
Motivo Walker A									
Domínio SB									
	$\alpha 4$	$\beta 3$	3_{10}	$\alpha 5$	$\beta 4$	$\alpha 6$	3_{10}	$\alpha 7$	
AK_MYCTU	LADHDGVLSL	GGGAVTSPGV	RAALAGH-TV	VYLEISAAEG	VRRITGGNTV-	--RPLL-AGP	DRAEKYRALM	133	
AK_CAMJE	SSCEKACIAT	GGGFV---N	VSNLEKAGFC	IYLKADFEYL	KKRLDKDEIS	-KRPLFYDEI	---KAKKLY	131	
AK_ECOLI	TEKQGIVLAT	GGGSVKSRER	RNRLSARGVV	VYLETTIEKQ	LARTQRDEK-	--RPLLHVET	PREVLEALA	137	
AL_ECOLI	TAP-STVIAT	GGGIILTEFN	RHFMQNNQIV	VYLCAPVSVL	VNRLQAAPPEE	DLRPTL-TGK	PLSEEVQEV	136	
AL_SALTY	SAP-STVVAT	GGGIILTEYN	RRYMHRVGVV	IYLCAPVSTL	VNRLQAAPPEE	DLRPTL-TGK	PLSEEVREVL	136	
AL_ERWCH	ATP-NRQVAT	GGGMVLLQ	RQFMRAHGT	VYLFAPAEEL	ALRLQASPOA	HQRPTL-TGR	PIAEEMEAVL	136	
Motivo Walker B									
Domínio LID									
	$\alpha 7$	$\beta 5$	$\alpha 8$						
AK_MYCTU	AKRAPLYRRV	ATMRVDTNRR	NNGAVVRHIL	SRLQVPSPE	AAT--	176			
AK_CAMJE	NERLSKYEQK	ANFILNIENK	NIDELLSEIK	KVIK-----	-----	165			
AK_ECOLI	NERNPLYEEI	ADVITRTDQ	SAKVVAQII	HMLESN----	-----	173			
AL_ECOLI	EERDALYREV	AHIIDATNE	PSQVISEIRS	ALAQTINC--	-----	174			
AL_SALTY	EQRDALYRET	AHYIIDATKA	PAQVSEIIA	ALPPSTQRLQ	GDVYT	181			
AL_ERWCH	REREALYQDV	AHYVVDATQP	PAIVCELMQ	TMLPAA---	-----	173			
Domínio de Ligação da Adenina									

Figura 3.5 – Alinhamento das chiquimato quinases mostrando que os resíduos envolvidos na ligação do chiquimato são conservados. Os elementos de estrutura secundária para MtCQ são mostrados acima da sequência. Motivos identificados em CQ são apresentados na figura, assim como o domínio SB e domínio LID em azul e verde, respectivamente. Regiões preenchidas em amarelo indicam os resíduos envolvidos na ligação com chiquimato, e a região preenchida em vermelho mostra a conservação do Glu61. AK - AroK; AL - AroL; MYCTU - *M.tuberculosis*; CAMJE - *C.jejuni*; ECOLI - *E.coli*; SALTY - *S.typhimurium*; ERWCH - *E.chrysanthemi*.

Apesar das estruturas da MtCQ-MgADP serem altamente similares, existem diferenças na posição do Mg^{2+} . Na MtCQ-MgADP (PDB 1L4Y) uma típica hexa-coordenação é observada para o Mg^{2+} (Figura 3.6A). Para a MtCQ-MgADP derivativo Pt (PDB 1L4U) hexa-coordenação (na verdade sete interações são observadas) apresenta-se desordenada (Figura 3.6B). A MtCQ-MgADP-Chiquimato (este trabalho) apresentou uma hexa-coordenação desordenada do Mg^{2+} (Figura 3.6C). Distâncias de ligação para a ligação do Mg^{2+} são mostradas na Tabela 3.4.

O Mg^{2+} da MtCQ-MgADP-Chiquimato interage com oxigênio β -fosfato do ADP, átomo OG da Ser16 do motivo Walker A, e quatro moléculas de água. Em ErcCQ, o Mg^{2+} é tetra-coordenado, interagindo com oxigênio β -fosfato do ADP, átomo OG1 da Thr16, átomo OD2 do Asp32, e uma molécula água (Krell *et al.*, 1998). A ligação do Mg^{2+} nas estruturas MtCQ-MgADP e MtCQ-MgADP-Chiquimato são de certa forma similares, onde a Água2 coordenada com Mg^{2+} faz ligação de hidrogênio com Asp32, que é conservado em todas as CQs (Figura 3.5). Superposição das estruturas da MtCQ-MgADP e MtCQ-MgADP-Chiquimato demonstrou que na última estrutura a Água1 e Água4 estão em posições equivalentes, mas um deslocamento de 1,32 Å e 2,75 Å são observados para Água2 e Água3, respectivamente (Figura 3.6). Em MtCQ-MgADP-Chiquimato o Asp32 forma ligação de hidrogênio com Ser16, enquanto que a interação entre Asp32 e Ser16 ocorre via ligação com uma molécula de água (Água6). A rotação dos ângulos diedros χ_1 e χ_2 do Asp32 promove expulsão da Água6 do sítio de ligação do Mg^{2+} e uma interação direta com átomo OG da Ser16 é observada para MtCQ-MgADP-Chiquimato (Figura 3.6C).



A Água1 coordenada pelo Mg^{2+} interage com o íon cloreto ao invés de fazer uma ligação com a Água5 que apresenta uma interação com Asp34 como observado para MtCQ-MgADP. Conseqüentemente, o modo diferente de interação observado para o Asp34 deve ser devido à presença do chiquimato que promove a expulsão da Água5 do sítio ativo.

O motivo Walker B para CQs (resíduos 75-81 em MtCQ) teve a sequência proposta como ZZZTG₇₅GG (Leipe *et al.*, 2003). A segunda glicina (Gly80 em MtCQ) faz ligação de hidrogênio com o γ -fosfato do ATP. O ataque nucleofílico ao γ -fosfato do ATP deve ser facilitado pela ligação do íon metálico (Mg^{2+}) aos grupos β - e γ -fosfato (Jencks, 1975a). Embora o mecanismo enzimático da MtCQ permaneça desconhecido, a interação entre a Gly80 e o íon cloreto, que está ocupando o local de ligação do γ -fosfato, indica que este resíduo deve participar do mecanismo catalítico das CQs.



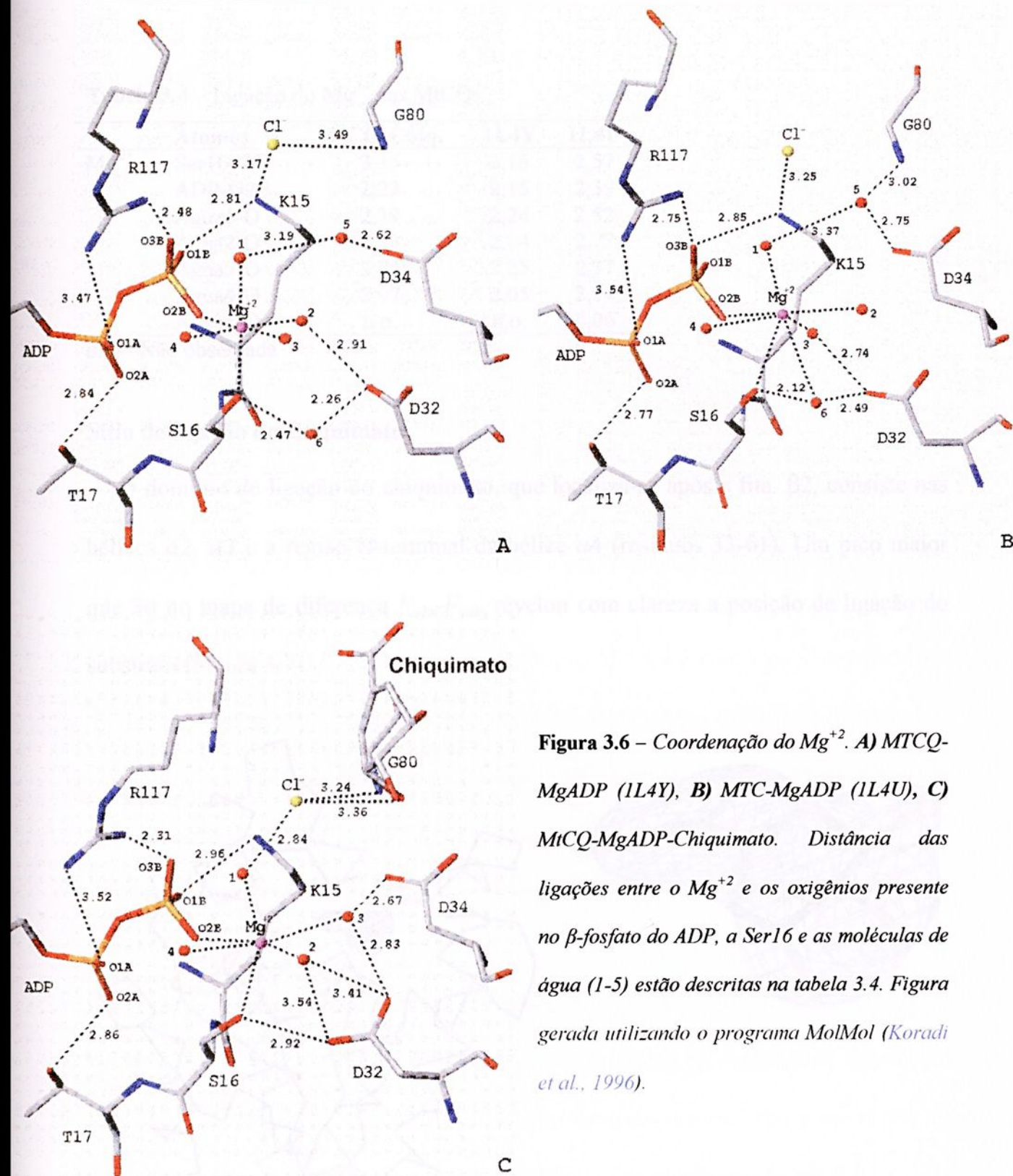


Figura 3.6 – Coordenação do Mg^{+2} . A) MTCQ-MgADP (1L4Y), B) MTC-MgADP (1L4U), C) MTCQ-MgADP-Chiquimato. Distância das ligações entre o Mg^{+2} e os oxigênios presente no β -fosfato do ADP, a Ser16 e as moléculas de água (1-5) estão descritas na tabela 3.4. Figura gerada utilizando o programa MolMol (Koradi et al., 1996).

Tabela 3.4 – Ligação do Mg^{2+} nas MtCQs.

	Átomos	MtCQ-Chiq.	1L4Y	1L4U
Mg^{2+}	Ser16-OG	2,36	2,15	2,57
	ADP-O2B	2,23	2,15	2,39
	Água1-O	2,39	2,24	2,52
	Água2-O	2,23	2,04	2,77
	Água3-O	2,75	2,25	2,37
	Água4-O	2,97	2,05	2,94
	Água5-O	n.o.	n.o.	3,06

n.o. – Não observada

Sítio de ligação do chiquimato

O domínio de ligação do chiquimato, que localiza-se após a fita $\beta 2$, consiste nas hélices $\alpha 2$, $\alpha 3$ e a região N-terminal da hélice $\alpha 4$ (resíduos 33-61). Um pico maior que 3σ no mapa de diferença $F_{obs}-F_{calc}$ revelou com clareza a posição de ligação do substrato (Figura 3.7).

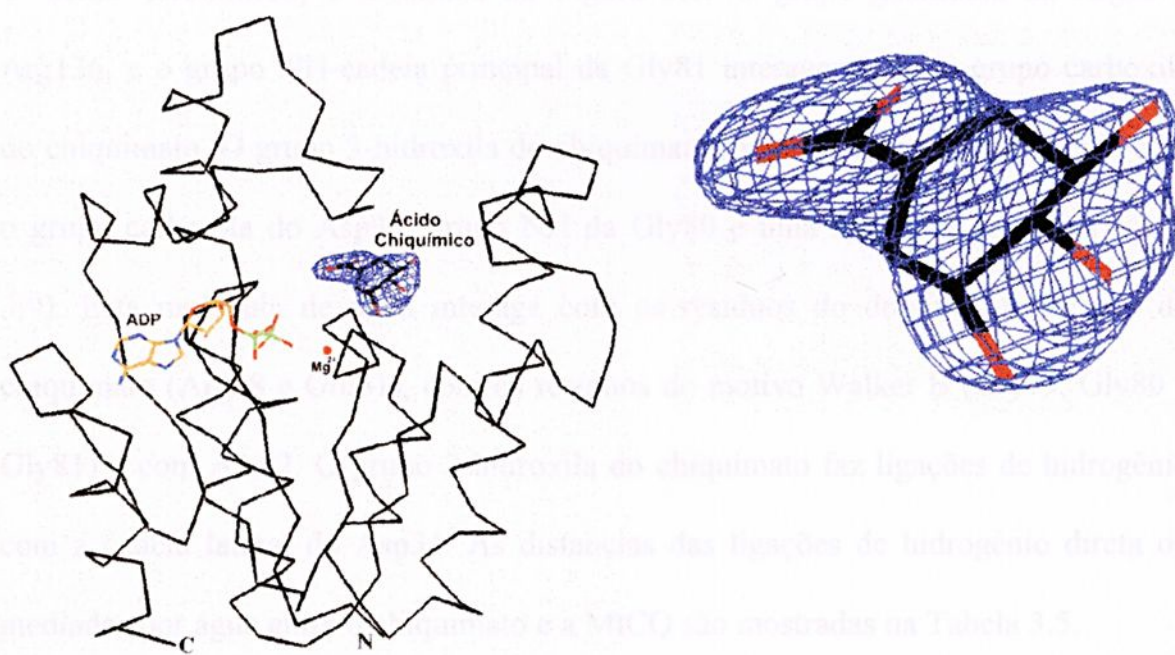


Figura 3.7 – Mapa $F_{obs}-F_{calc}$ (3σ) gerado para posicionar o chiquimato na estrutura da MtCQ. A figura mostra também os Ca da MtCQ-mgADP-Ác.chiq., juntamente com as moléculas de ADP e Mg^{2+} . A figura foi gerada utilizando o programa XtalView (McRee, 1999).

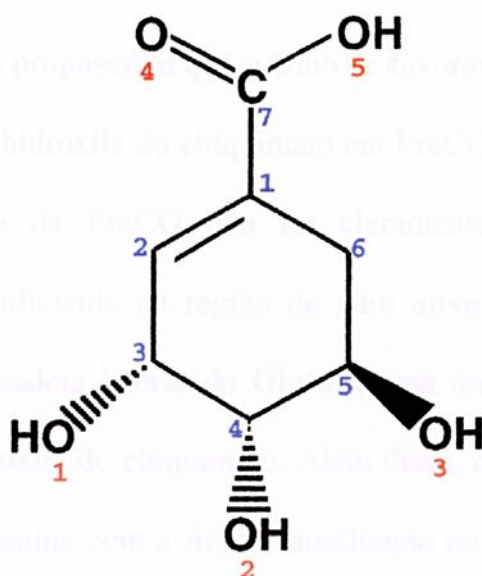


Figura 3.8 – Estrutura molecular do chiquimato evidenciando o grupo carboxila e os três grupos hidroxila. Os átomos de carbono estão numerados em azul e os átomos de oxigênio em vermelho.

A estrutura química do chiquimato [3R-(3 α ,4 α ,5 β)]3,4,5-trihidroxi-1-ciclohexeno-1-ácido carboxílico] é mostrada na Figura 3.8. O grupo guanidina da Arg58 e Arg136, e o grupo NH-cadeia principal da Gly81 interagem com o grupo carboxila do chiquimato. O grupo 3-hidroxila do chiquimato forma ligações de hidrogênio com o grupo carboxila do Asp34, grupo NH da Gly80 e uma molécula de água (Figura 3.9). Esta molécula de água interage com os resíduos do domínio de ligação do chiquimato (Arg58 e Glu61), com os resíduos do motivo Walker B (Gly79, Gly80 e Gly81) e com Ala82. O grupo 2-hidroxila do chiquimato faz ligações de hidrogênio com a cadeia lateral do Asp34. As distâncias das ligações de hidrogênio direta ou mediadas por água entre o chiquimato e a MtCQ são mostradas na Tabela 3.5.

O resíduo Glu61 é conservado em ambas chiquimato quinases (tanto as codificadas pelo gene *aroK* ou pelo gene *aroL*) e acreditava-se que este resíduo estava envolvido

na ligação do chiquimato (Krell *et al.*, 1998; Gu *et al.*, 2002; Romanowski & Burley, 2002). Krell e co-autores propuseram que o Glu61 é favoravelmente posicionado para a ligação com o grupo 5-hidroxila do chiquimato em ErcCQ. Entretanto, a posição do chiquimato na estrutura da ErcCQ não foi claramente demonstrado devido a densidade eletrônica insuficiente na região do sítio ativo. Na estrutura da MtCQ-MgADP-Chiquimato, a cadeia lateral do Glu61 forma uma interação mediada por água com o grupo 3-hidroxila do chiquimato. Além disso, o Glu61 forma ligações de hidrogênio e uma ponte salina com a Arg58, auxiliando no posicionamento do grupo guanidina deste resíduo, que é importante para ligação do substrato através de interações com o grupo carboxila. Portanto, o resíduo Glu61 demonstra um papel importante no posicionamento da molécula do chiquimato, embora não esteja diretamente envolvido na ligação. O Glu61 é conservado em todas as CQs seqüenciadas até o momento, o que reforça o papel de posicionar a Arg58 no sítio ativo. O Glu54 também apresenta o papel de posicionar a Arg58 para interação do chiquimato, como sugerido na estrutura da MtCQ-MgADP (Gu *et al.*, 2002). Entretanto, o Glu54 não é conservado em todas as CQs, com exceção das codificadas pelo gene *aroK* (Figura 3.5).



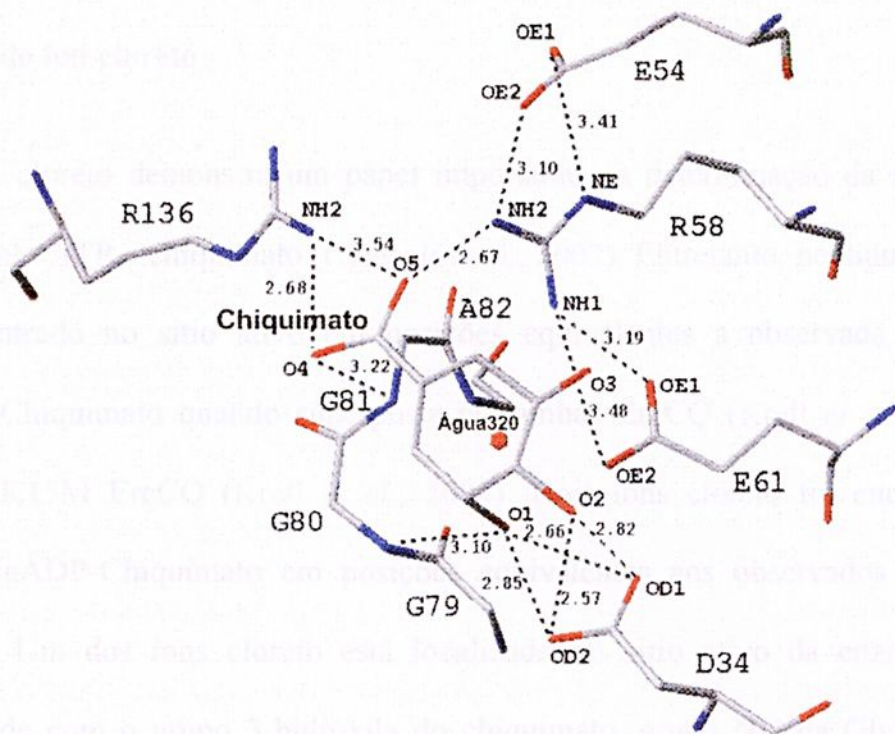


Figura 3.9 – Interações envolvidas na ligação chiquimato no sítio ativo da MtCQ. Ligações de hidrogênio entre chiquimato e grupos protéicos estão representadas por linhas tracejadas. A figura foi gerada utilizando o programa MolMol (Koradi et al., 1996).

Tabela 3.5 – Ligações de hidrogênio direta ou mediada por água entre o chiquimato e MtCQ

Ácido chiquímico	MtCQ	Distâncias (Å)	Ligações de hidrogênio mediadas por água	Distâncias (Å)
Grupos hidroxila				
O1	Gly80-N	3,10		
	Asp34-OD1	2,82		
	Asp34-OD2	2,57		
	Água320-O	2,46		
			Gly79-O	2,96
			Gly80-N	3,13
			Gly81-N	2,51
			Ala82-N	3,54
			Arg58-NH1	3,12
			Glu61-OE2	2,99
O2	Asp34-OD1	2,66		
	Asp34-OD2	2,85		
O3	n.o			
Grupo carboxila				
O4	Gly81-N	3,22		
	Arg136-NH2	2,68		
O5	Arg58-NH2	2,67		
	Gly 81-N	3,49		
	Arg136-NH2	2,34		

Todas as distâncias < 3,6 Å são mostradas.
n.o., Não observada

Ligação do íon cloreto

O íon cloreto demonstra um papel importante na determinação da afinidade da ErcCQ pelo ATP e chiquimato (Cerasoli *et al.*, 2003). Entretanto, nenhum íon cloreto foi encontrado no sítio ativo em posições equivalentes a observada em MtCQ-MgADP-Chiquimato quando superposta em ambas ErcCQ (Krell *et al.*, 1998) e o mutante K15M ErcCQ (Krell *et al.*, 2001). Dois íons cloreto foi encontrado em MtCQ-MgADP-Chiquimato em posições equivalentes aos observados em MtCQ-MgADP. Um dos íons cloreto está localizado no sítio ativo da enzima fazendo ligações de com o grupo 3-hidroxila do chiquimato, grupo NH da Gly80 e Água1 (Figura 3.10), enquanto que o outro está localizado na superfície da proteína, distante do sítio ativo. O resíduo conservado Arg117 do domínio LID está envolvido na ligação do ADP, promovendo o deslocamento observado na região do domínio LID. A Lys15 forma ligações de hidrogênio com um átomo de oxigênio do β -fosfato e com íon cloreto na estrutura MtCQ-MgADP (Figura 3.10). O NH da cadeia principal da Gly80 faz ligação de hidrogênio com íon cloreto em MtCQ-MgADP e MtCQ-MgADP-Chiquimato (Tabela 3.10). Estes resíduos (Lys15, Gly80 e Arg117) estão localizados próximos do local onde ocorre a reação química e devem apresentar um papel crucial na estabilização do estado de transição. Consistente com esta proposta, o mutante K15M da ErcCQ não possui atividade enzimática (Krell *et al.*, 2001). O resíduo Gly80 do motivo Walker-B está envolvido através de uma ligação de hidrogênio com γ -fosfato of ATP (Krell *et al.*, 1998). A distância entre o íon cloreto e o grupo 3-hidroxila do chiquimato (local onde o foforil é transferido) é de 3,36 Å em



31

Tabela 3.6 – Ligação do íon cloreto nas estruturas da MtCQ-MgADP (1L4Y e 1L4U) e MtCQ-MgADP-Chiquimato

Íon cloreto	Átomos	Distâncias (Å)		
		1L4Y	1L4U	MtCQ-MgADP-Chiquimato
Cl180	Grupo 3-hidroxila do chiquimato - O1	4,48 ^a	4,66 ^a	3,36
	Gly80-N	3,49	3,69	3,24
	Lys15-NZ	3,17	3,25	3,94
	Mg ²⁺ -Água1-O	3,87	3,75	2,84

^aAs distâncias foram medidas com a estrutura da MtCQ-MgADP superpostas na estrutura da MtCQ-MgADP-Chiquimato.

Mudanças conformacionais devido à ligação do chiquimato

Como citado anteriormente, MtCQ pertence a família das nucleosídeo monofosfato (NMP) quinases, que são compostas por três domínios: CORE, LID, e domínio de ligação NMP. As quinases apresentam um grande movimento durante a catálise para defender o sítio ativo da água evitando a hidrólise do ATP (Jencks, 1975b).

A MtCQ-MgADP-Chiquimato representa a estrutura mais fechada da MtCQ, mas não totalmente, desde que o total fechamento deve ser obtido com a ligação do ATP. O mutante K15M ErcCQ (Krell *et al.*, 2001) foi cristalizado em uma conformação aberta que presumidamente deve ser equivalente a conformação de uma estrutura apo-enzima (sem ADP e chiquimato) (Leipe *et al.*, 2003). Este mutante foi produzido para avaliar o papel da Lys15 do motivo Walker-A na catálise. Entretanto, uma mutação inesperada no domínio LID (Pro115Leu) foi detectada durante o refinamento do modelo. Além disso, a enzima foi cristalizada com duas moléculas na unidade assimétrica com contatos entre os domínios LID, o que não é observado na

estrutura cristalográfica nativa (Krell *et al.*, 1998; Krell *et al.*, 2001). Apesar do duplo mutante K15M/P115L ErcCQ como modelo para apo-enzima apresentarem estes “problemas”, consideramos como sendo a conformação da estrutura apo, uma que vez que não tínhamos um melhor modelo disponível no momento. A superposição da MtCQ-MgADP-Chiquimato e apo ErcCQ (K15M+P115L) mostra a grande mudança conformacional nos domínios LID e SB devido a ligação do ADP e chiquimato na MtCQ (Figura 3.11).

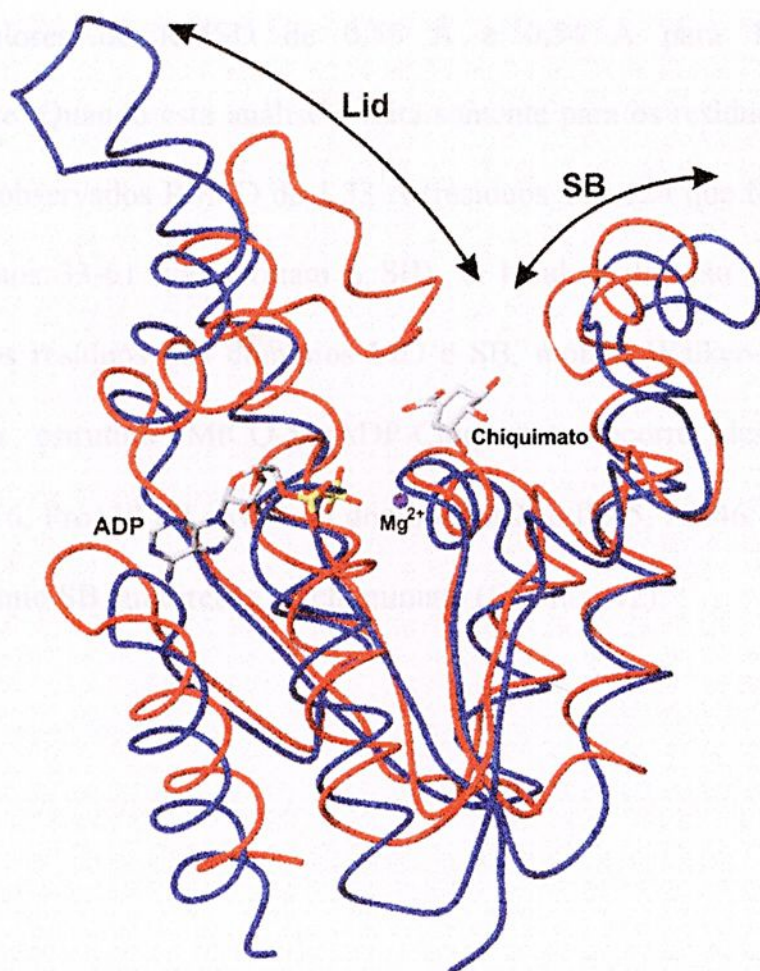


Figura 3.11 – Superposição da MtCQ-MgADP-Chiquimato (vermelho) e o mutante K15M/P115L da ErcCQ (azul) (Krell *et al.*, 2001) mostrando o grande deslocamento do domínio LID e do domínio SB devido a ligação do ADP e do chiquimato. A figura foi gerada utilizando o programa MolMol (Koradi *et al.*, 1996).

Mudanças nos elementos de estrutura secundária da ErcCQ foram observadas através dos espectros obtidos por dicroísmo circular da enzima livre (ErcCQ), do complexo binário (ErcCQ-Chiquimato), e do complexo ternário (ErcCQ-Chiquimato-Adenil-imidodifosfato) (Krell *et al.*, 1998). A superposição dos C α da MtCQ-MgADP-Chiquimato e MtCQ-MgADP demonstrou que os domínios LID e SB sofreram um notável deslocamento em direção um ao outro na estrutura MtCQ-MgADP-Chiquimato (Figura 3.12). A superposição incluindo todos os resíduos apresentou valores de RMSD de 0,56 Å e 0,54 Å para 1L4U e 1L4Y, respectivamente. Quando esta análise é feita somente para os resíduos dos domínios LID e SB são observados RMSD de 1,33 Å (resíduos 112-124 que formam o LID) e 0,74 Å (resíduos 33-61 que formam o SB). O local de ligação do chiquimato é composto pelos resíduos dos domínios LID e SB, motivo Walker-B, e Arg136 da hélice- α 7. Na estrutura MtCQ-MgADP-Chiquimato ocorre deslocamento dos resíduos Val116, Pro118 e Leu119 do domínio LID, e Ile45, Ala46, Glu54, Phe57 e Arg58 do domínio SB em direção ao chiquimato (Figura 3.12).

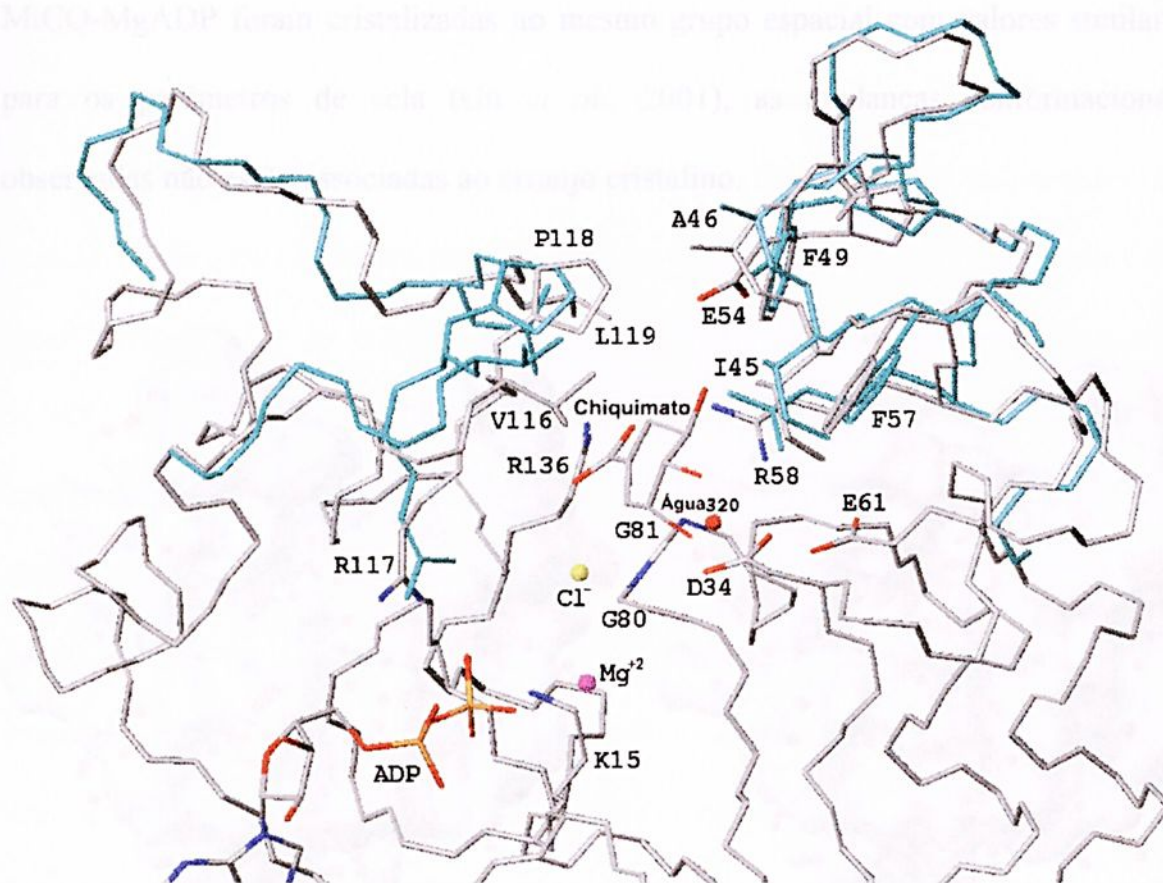


Figura 3.12 – Superposição dos C α do domínio LID e do domínio SB das estruturas MtCQ-MgADP-Chiquimato (cinza) e MtCQ-MgADP (verde) (1L4U; Gu et al., 2002). A figura mostra deslocamento em direção ao substrato dos resíduos Val116, Pro118 e Leu119 do domínio LID e Ile45, Ala46, Glu54, Phe57 e Arg58 do domínio SB devido a ligação do chiquimato. A figura foi gerada utilizando o programa MolMol (Koradi et al., 1996).

As mudanças conformacionais descritas acima podem ser demonstradas pela redução da superfície molecular da MtCQ-MgADP-Chiquimato quando comparada com MtCQ-MgADP (Figura 3.13). O ADP, Mg $^{2+}$, chiquimato, Cl $^{-}$ foram removidos para o cálculo. Os valores calculados são aproximadamente 7246 Å 2 para MtCQ complexada com MgADP (1L4U) e 6915 Å 2 para estrutura complexada com MgADP e chiquimato. Desta forma, aproximadamente 330 Å 2 na superfície molecular é reduzido devido à ligação do chiquimato. Uma vez que MtCQ-MgADP-Chiquimato e

MtCQ-MgADP foram cristalizadas no mesmo grupo espacial com valores similares para os parâmetros de cela (Gu *et al.*, 2001), as mudanças conformacionais observadas não estão associadas ao arranjo cristalino.

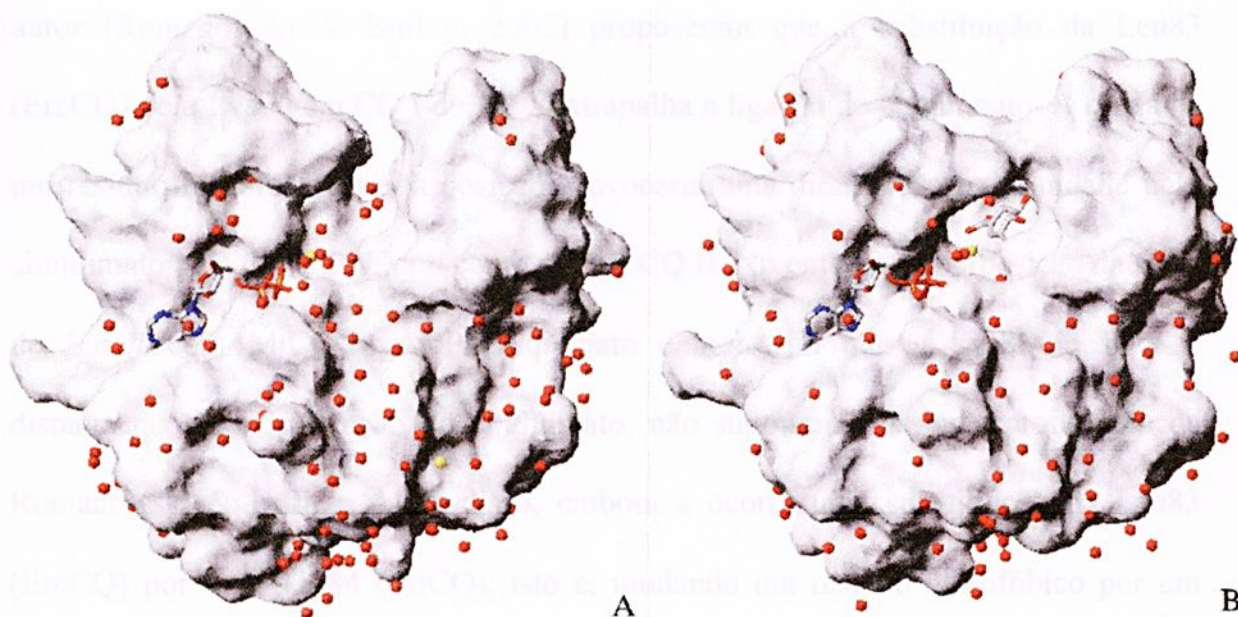


Figura 3.13 – Superfície molecular das estruturas (A) MtCQ-MgADP (1L4U; Gu *et al.*, 2002) e (B) MtCQ-MgADP-Chiquimato. As moléculas de água são coloridas em vermelho. A figura foi gerada utilizando o programa MolMol (Koradi *et al.*, 1996).

3.1.3 Conclusão

Os resíduos identificados na ligação do chiquimato diretamente ou indiretamente (Asp34, Arg58, Glu61, Gly79, Gly80, Gly81 e Arg136) são conservados em todas as CQs (codificadas pelos genes *aroK* e *aroL*) (Figura 3.5). Estruturas de CQs depositadas no PDB foram superpostas e as posições dos resíduos envolvidos na ligação entre o chiquimato e a CQ são altamente conservadas para as proteínas codificadas pelos genes *aroK* (Chiquimato quinase de *Mycobacterium tuberculosis* e

Campylobacter jejuni) e *aroL* (Chiquimato quinase de *Erwinia chrysanthemi*). A conservação do sítio ativo pode ser responsável pelos valores similares de *K_m* encontrados para MtCQ (*aroK*) e EcSK (*aroL*), entretanto esta informação não explica a diferença observada para valores *K_m* da CQ I de *E.coli*. Romanowski e co-autor (Romanowski & Burley, 2002) propuseram que a substituição da Leu83 (ErcCQ) pela Lys86 em CQ I de *E.coli* atrapalha a ligação do chiquimato. A perda de um resíduo hidrofóbico nesta posição provocaria uma diminuição na afinidade pelo chiquimato pela CQ I de *E.coli* comparada à CQ II. No entanto, superposição da CQ I de *E.coli* com MtCQ-MgADP-Chiquimato demonstrou que a Lys86 localiza-se distante do sítio de ligação do chiquimato, não suportando assim a proposição de Romanowski & Burley. Além disso, embora a ocorra uma substituição da Leu83 (ErcCQ) por uma Thr84 (MtCQ), isto é, mudando um resíduo hidrofóbico por um resíduo polar, os valores de *K_m* para a MtCQ e ErcCQ são similares.

A superposição da CQ I de *E.coli* na estrutura da MtCQ-MgADP-Chiquimato mostrou o resíduo Leu123 (*E.coli*) na posição ocupada pelo chiquimato, porém não podemos afirmar que os sítios ativos são diferentes, uma vez que a estrutura da CQ I de *E.coli* foi cristalizada na ausência de chiquimato, e a ligação deste composto pode promover futuras mudanças de conformação incluindo o deslocamento da Leu123.

Neste trabalho foram descritos os resíduos envolvidos na ligação do chiquimato e as mudanças conformacionais devido a ligação desta molécula. O fechamento completo do sítio ativo provavelmente será alcançado através do complexo MtCQ com chiquimato, Mg^{2+} , e um análogo de ATP não hidrolisável 5'-(β,γ -metileno) trifosfato (AMP-PCP). A disponibilidade da estrutura da chiquimato quinase de *M.*



tuberculosis complexada com chiquimato irá permitir o desenho molecular de inibidores específicos para CQ baseados no sítio de ligação do substrato. Além disso, as informações aqui apresentadas contribuirão para entendimento dos fatores responsáveis pelo fechamento do sítio ativo, que poderão ser utilizadas o desenvolvimento de inibidores que forcem a estrutura em uma conformação fechada que seja incapaz de catalisar a transferência do fosfato para o chiquimato.

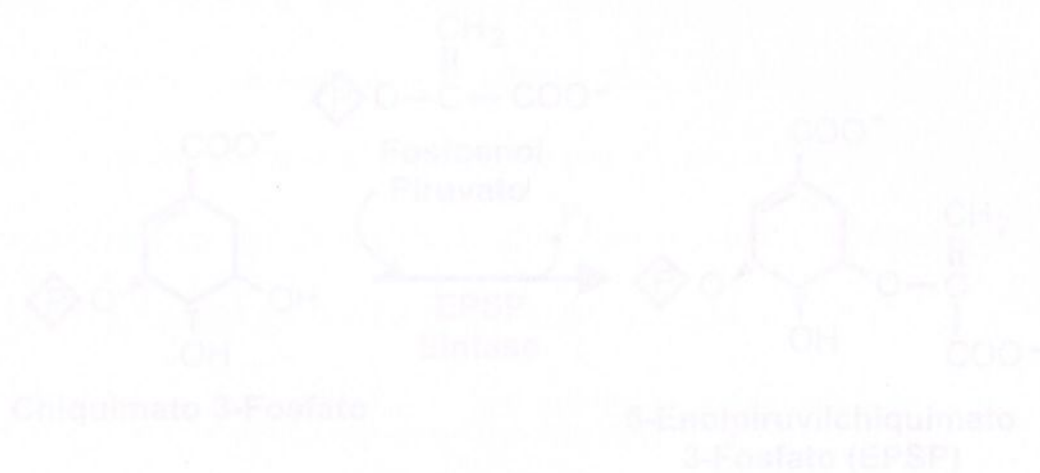


Figura 3.14 – Reação catalisada pela EPSPS. Figura adaptada de Muehle & van der Meulen, 2002 [18].

Como estado anti-conformacional, um inibidor é considerado ruim para a construção de novos inibidores de EPSPS é o glicolato. Devido a isso, modelos para EPSPS de *M. tuberculosis* (MEPSPS) são propostos em que o sítio de ligação do substrato é dividido em duas partes: a primeira para o glicolato e a segunda para o chiquimato. O glicolato é considerado um inibidor ruim porque ele não é capaz de se ligar ao sítio de ligação do substrato, mas a ligação do chiquimato é considerada boa porque ele é capaz de se ligar ao sítio de ligação do substrato.

3.2 Modelagem molecular da 5-enolpiruvilchiquimato-3-fosfato sintase (EPSP sintase)

A EPSP sintase (EC 2.5.1.19) catalisa o sexto passo da via do ácido chiquímico, que corresponde à transferência do enolpiruvil do fosfoenol piruvato (PEP) para o chiquimato-3-fosfato (produto da catálise da chiquimato quinase), dando origem ao 5-enolpiruvilchiquimato-3-fosfato e um fosfato inorgânico (Figura 3.14) (Bentley, 1990; Levin & Sprinson, 1964).

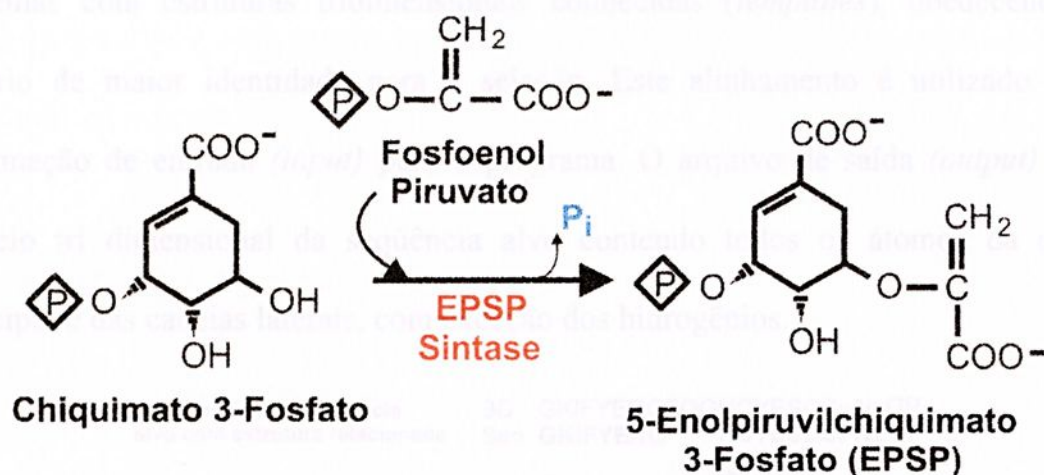


Figura 3.14 – Reação catalisada pela EPSPS sintase. Figura adaptada de Mathews & van Holde. (Mathews & van Holde, 1990).

Como citado anteriormente, um importante composto guia para a construção de novos inibidores de EPSPS é o glifosato. Deste modo, dois modelos para EPSPS de *M. tuberculosis* (MtEPSPS) são propostos, sendo um na ausência de ligantes (modelo aberto) e o outro na presença do chiquimato-3-fosfato e do glifosato (modelo fechado), uma vez que a literatura reporta uma mudança conformacional entre estes dois estados (Schönbrunn *et al.*, 2001)

3.2.1 Materiais e Métodos

Para construção do modelo da MtEPSPS foi usada modelagem molecular por homologia implementada no programa MODELLER (Šali & Blundell, 1993). Este programa aborda modelagem comparativa satisfazendo restrições espaciais. O procedimento para modelagem inicia-se com o alinhamento da sequência de aminoácidos a ser modelada (alvo), com uma ou várias sequências primárias de proteínas com estruturas tridimensionais conhecidas (*templates*), obedecendo ao critério de maior identidade para a seleção. Este alinhamento é utilizado como informação de entrada (*input*) para o programa. O arquivo de saída (*output*) é um modelo tri dimensional da sequência alvo contendo todos os átomos da cadeia principal e das cadeias laterais, com exceção dos hidrogênios.

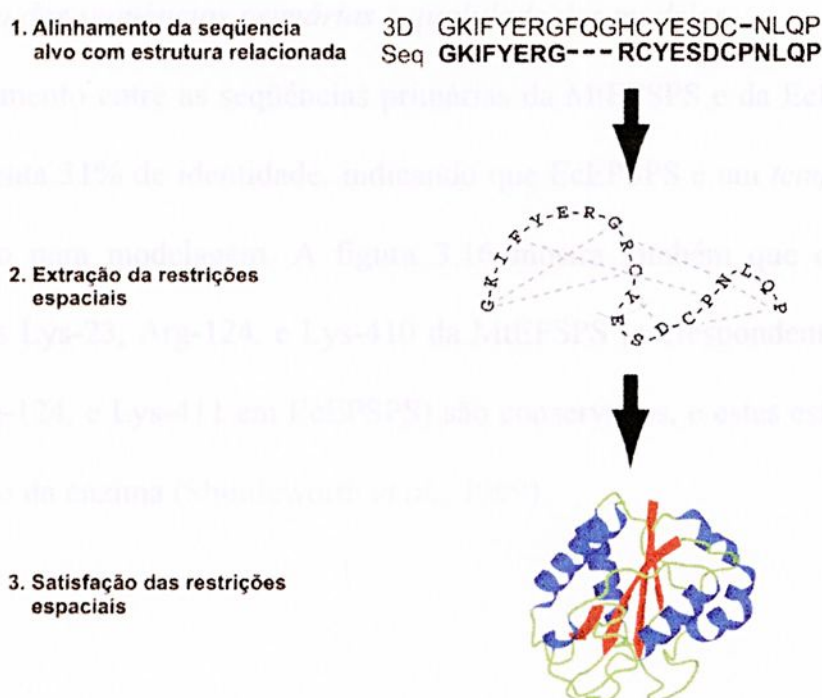


Figura 3.15 — Representação esquematizando as etapas de construção de um modelo utilizando a modelagem molecular comparativa.

As coordenadas atômicas da estrutura cristalográfica da EPSPS *Escherichia coli* (EcEPSPS), que apresenta duas estruturas cristalográficas, uma no estado aberto (sem ligante – PDB:1EPS) (Stallings *et al.*, 1991) e a outra no estado fechado (complexada com chiquimato-3-fosfato e glifosato – PDB:1G6S) (Schönbrunn *et al.*, 2001) foram utilizada como modelo inicial para modelagem da MtEPSPS apo-enzima e MtEPSPS complexada com chiquimato-3-fosfato e glifosato. As coordenadas atômicas de todas as águas foram removidas das estruturas da EcEPSPS. Um total de 1000 modelos foram gerados para cada estado, e o modelo final foi selecionado baseando-se na sua qualidade estereoquímica.

3.2.2 Resultados e discussão

Alinhamento das seqüências primárias e qualidade dos modelos

O alinhamento entre as seqüências primárias da MtEPSPS e da EcEPSPS (Figura 3.16) apresenta 31% de identidade, indicando que EcEPSPS é um *template* que pode ser utilizado para modelagem. A figura 3.16 mostra também que os resíduos de aminoácidos Lys-23, Arg-124, e Lys-410 da MtEPSPS (correspondente aos resíduos Lys-22, Arg-124, e Lys-411 em EcEPSPS) são conservados, e estes estão localizados no sítio ativo da enzima (Shuttleworth *et al.*, 1999).



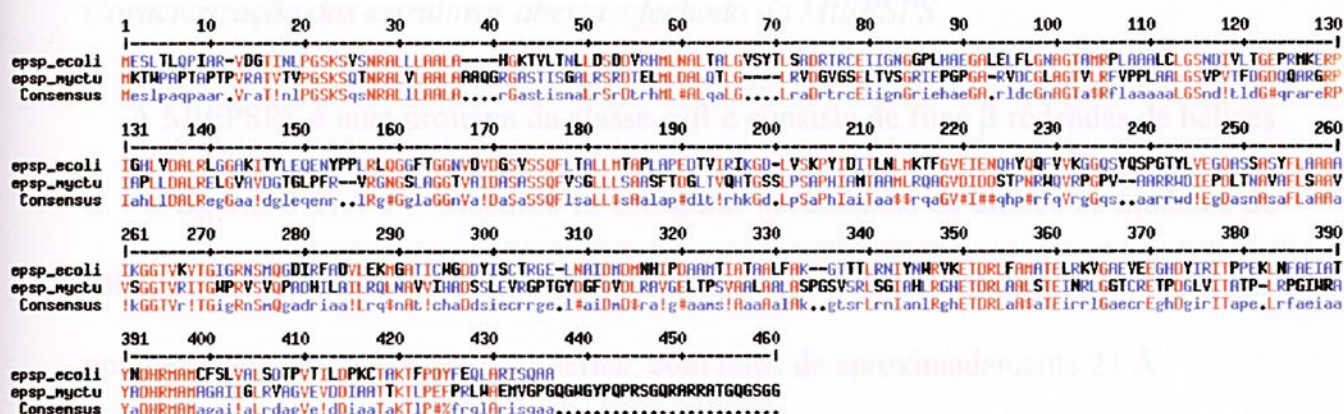


Figura 3.16 – Alinhamento das seqüências de aminoácidos da MtEPSPS e da EcEPSPS. Em azul e preto são mostrados resíduos que não são idênticos ou não foram alinhados, e em vermelho os resíduos idênticos ou semelhantes entre as seqüências.

A qualidade estereoquímica do modelo da MtEPSPS aberto e fechado foram analisados pelo diagrama ϕ - ψ . O modelo fechado apresentou 96% dos resíduos nas regiões permitidas, e 4% nas regiões não permitidas. O modelo aberto (sem ligante), apresentou 91,1% resíduos nas regiões permitidas, e 8,9% nas regiões não permitidas. A quantidade de resíduos na região não permitida do diagrama para o modelo aberto é devido ao *template* utilizado da EcEPSPS ter somente os carbonos α depositado no PDB (código: 1EPS), não fornecendo assim, informações suficientes para obtenção de um bom modelo. Não foi utilizado outro *template* para a modelagem da MtEPSPS, pois a única EPSPS depositada no PDB sem ligante é de *E. coli*.

Caracterização das estruturas aberta e fechada da MtEPSPS

A MtEPSPS é uma proteína da classe α/β e consiste de fitas β rodeadas de hélices α . As figuras 3.17A e B mostram as estruturas secundárias de ambos os modelos de MtEPSPS. O enovelamento das estruturas apresentou dois domínios similares, que possuem uma conformação hemisférica, com raios de aproximadamente 21 Å.

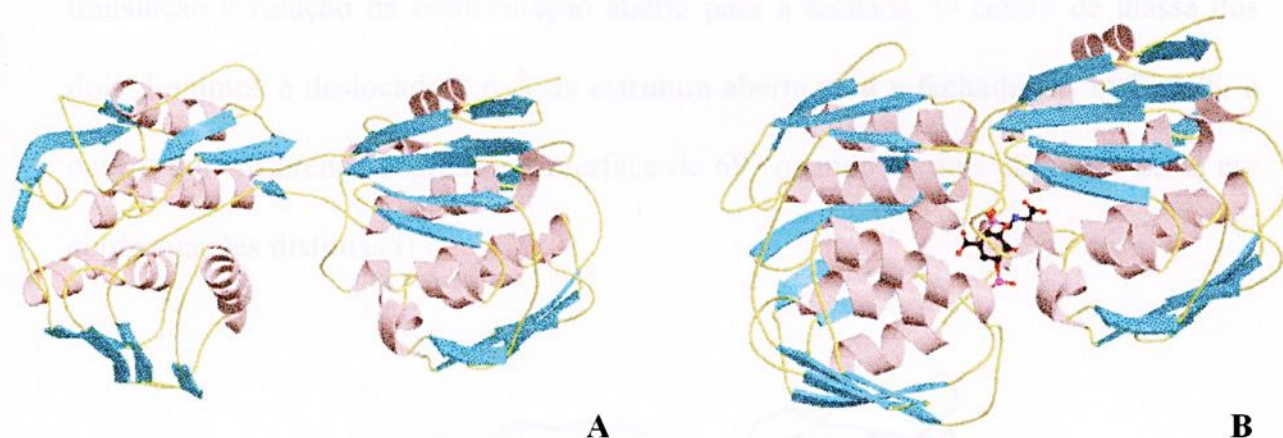


Figura 3.17 – (A) MtEPSPS sem ligantes – estrutura aberta e (B) MtEPSPS complexada com chiquimate-3-fosfato e glifosato – estrutura fechada.

Os dois domínios são ligados por dois segmentos da cadeia com ambos, amino e carboxi terminais da proteína localizados no mesmo domínio. Os efeitos macrodipolares observados nas hélices criam um potencial que facilita a ligação do chiquimate-3-fosfato e do glifosato (Stallings *et al.*, 1991). Os domínios são formados por três cópias de $\beta\alpha\beta\alpha\beta$ - unidade de enovelamento. As quatro fitas- β de cada unidade de enovelamento formam folhas- β paralela e antiparalela, e as duas hélices são paralelas a estas folhas. Na estrutura fechada, as moléculas do

chiquimato-3-fosfato e do glifosato estão ligadas na região interdomínio, promovendo o fechamento da estrutura.

“Dobradiça” molecular

As estruturas cristalográficas de EcEPSPS e dos modelos moleculares de MtEPSPS revelam a presença de dois domínios, designados A e B, que sofrem uma translação e rotação da conformação aberta para a fechada. O centro de massa dos dois domínios é deslocado 4,6 Å da estrutura aberta para a fechada em MtEPSPS, e ocorre uma diferença angular na interface de 69° quando os dois domínios estão em conformações distintas (Figura 3.18).

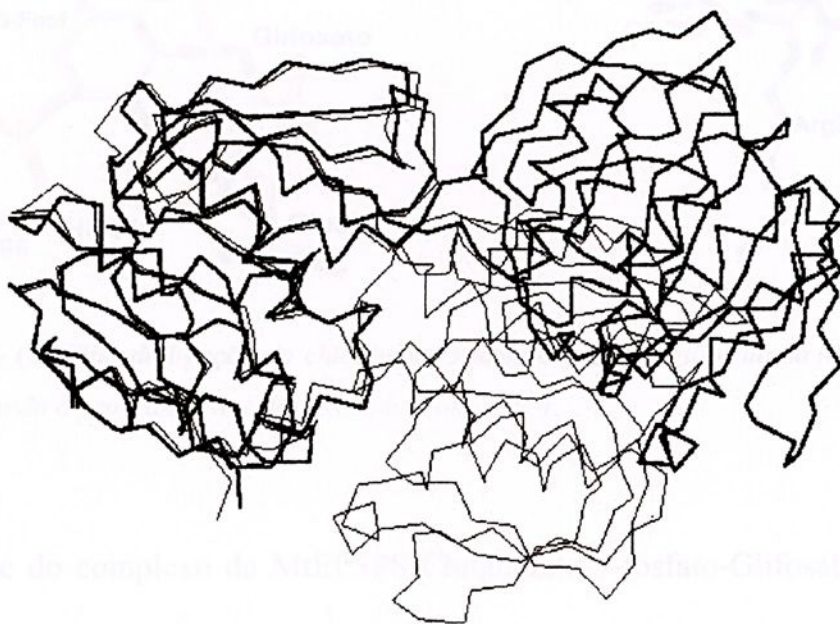


Figura 3.18 – Sobreposição da MtEPSPS aberta (linha grossa) e fechada (linha fina) mostrando a diferença conformacional entre as estruturas devido a ligação do chiquimato-3-fosfato e do glifosato. Figura gerada utilizando o programa XtalView (McRee, 1999).

Ligação do chiquimato-3-fosfato e do glifosato

Os sítios de ligação para o glifosato e o chiquimato-3-fosfato ficam próximos e o grupo 5-hidroxila do chiquimato-3-fosfato faz ligação de hidrogênio com o átomo de nitrogênio do glifosato a uma distância de 2,82 Å. A figura 3.19 mostra o sítio de ligação do chiquimato-3-fosfato e do glifosato da MtEPSPS.

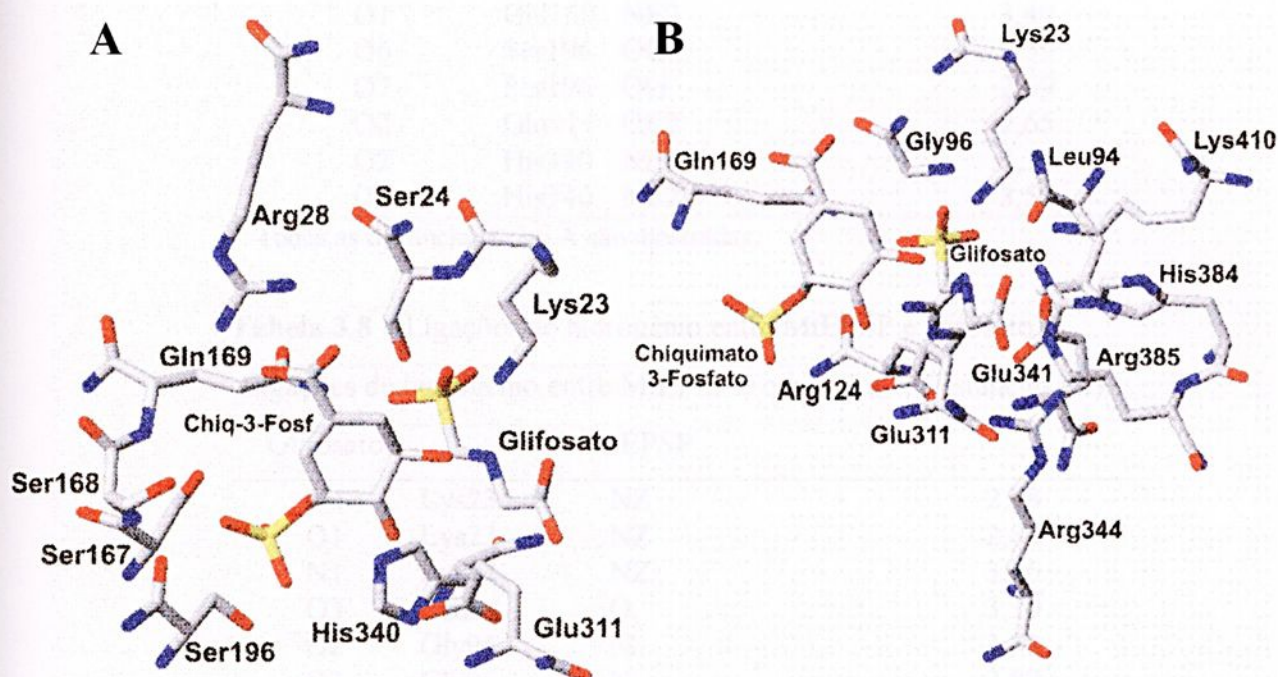


Figura 3.19 – (A) Sítio de ligação do chiquimato-3-fosfato e do (B) glifosato da MtEPSPS. Figura gerada utilizando o programa MolMol (Koradi et al., 1996).

A análise do complexo da MtEPSPS-Chiquimato-3-fosfato-Glifosato indicou um total de 15 e 22 ligações de hidrogênio da enzima com o chiquimato-3-fosfato e o glifosato, respectivamente (Tabelas 3.7 e 3.8).

Tabela 3.7 – Ligações de hidrogênios entre MtEPSP e chiquimato-3-fosfato

Ligações de hidrogênio entre MtEPSP e o substrato			Distâncias (Å)
Chiq-3-Fosfato	MtEPSP		
O3	Lys23	NZ	2,81
O5	Ser24	OG	2,70
O5	Arg28	NH1	2,86
O4	Arg28	NH2	2,78
O5	Arg28	NH2	3,60
O6	Ser167	OG	3,40
O8	Ser167	OG	2,70
O6	Ser168	OG	2,69
O6	Ser168	N	2,84
O1	Gln169	NE2	3,49
O6	Ser196	OG	3,59
O7	Ser196	OG	2,69
O2	Glu311	OE2	2,65
O2	His340	NE2	3,14
O1	His340	NE2	3,57

Todas as distâncias < 3,6 Å são mostradas.

Tabela 3.8 – Ligações de hidrogênio entre MtEPSP e glifosato.

Ligações de hidrogênio entre MtEPSP e o inibidor			Distâncias (Å)
Glifosato	EPSP		
O4	Lys23	NZ	2,98
O1	Lys23	NZ	2,84
N1	Lys23	NZ	3,46
O3	Leu94	O	3,10
O2	Gly96	N	3,53
O3	Gly96	N	2,87
O3	Arg124	NH1	2,87
O2	Arg124	NH2	2,81
O3	Arg124	NH2	3,60
O2	Gln169	NE1	3,58
O2	Gln169	NE2	2,88
O5	Glu311	OE1	2,96
N1	Glu341	OE1	3,35
N1	Glu341	OE2	2,88
O4	Glu341	OE2	3,59
O5	Arg344	NH1	2,81
O5	Arg344	NH2	3,02
O4	His384	NE2	3,49
O4	Arg385	NE	3,59
O5	Arg385	NE	2,75
O4	Arg385	NH2	3,12
O3	Lys410	NZ	2,95

Todas as distâncias < 3,6 Å são mostradas.



3.3 Modelagem molecular e análise de CD da chiquimato sintase

A maioria das ligações de hidrogênio entre glifosato e a MtEPSPS são formadas pelos resíduos Lys23, Arg124, Glu341, e Arg385, os mesmos resíduos identificados na ligação com a EcEPSPS. Além disso, o alinhamento de 69 seqüências primárias de EPSPS indicou que estes resíduos são conservados em todas as seqüências. Esta observação sugere que o glifosato deve ser capaz de inibir a maioria ou quase todas EPSPS, uma vez que a afinidade e especificidade de um inibidor estão relacionadas com as ligações de hidrogênio, interações iônicas, assim como complementariedade de cargas e superfície de contato entre enzima-inibidor (de Azevedo *et al.*, 1996; Kim *et al.*, 1996).

Dentre as sete enzimas presentes na via metabólica do ácido chiquímico, acredita-se que a EPSPS seja um dos principais alvos para inibir o *M. tuberculosis*, uma vez que a inibição desta proteína provocou a morte do organismo alvo. Portanto, o conhecimento das estruturas (na presença e na ausência de ligantes) da MtEPSPS e do modo de interação com chiquimato-3-fosfato (proteína-substrato) e com glifosato (proteína-inibidor) se tornam imprescindíveis para desenvolvimento de possíveis inibidores.



3.3 Modelagem molecular e análise de CD da corismato sintase

A corismato sintase (EC 4.2.3.5) catalisa a última reação desta via, a conversão do 5-enolpiruvil-3-chiquimato fosfato (EPSP) para corismato (Bornemann *et al.*, 2002; Macheroux *et al.*, 1998). A enzima utiliza flavina mononucleotídeo (FMN) reduzida como cofator, embora a reação catalisada não envolva uma mudança no estado de oxi-redução. O papel da FMN reduzida na catálise ainda não está totalmente elucidado. Entretanto, recentes detalhes cinéticos contribuíram para avançar no entendimento do mecanismo de ação da corismato sintase (Macheroux *et al.*, 1999).

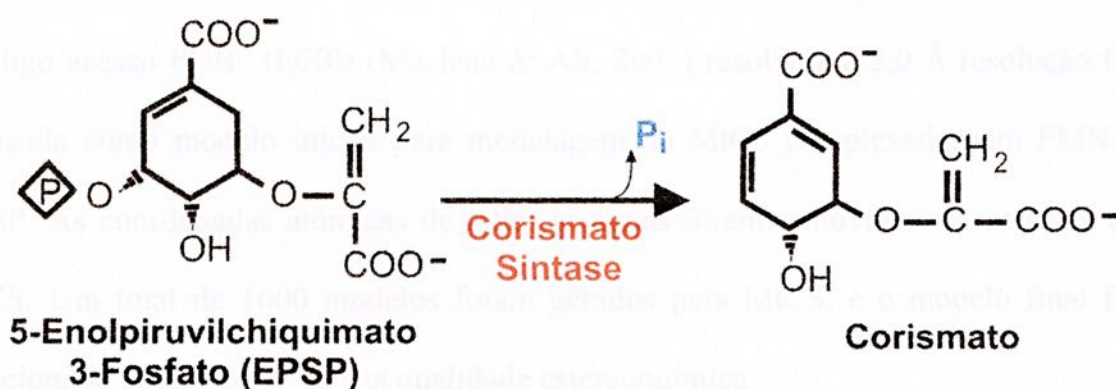


Figura 3.20 – Reação catalisada pela corismato sintase. Figura modificada a partir de Mathews & van Holde. (Mathews & van Holde, 1990).

O produto da corismato sintase é o precursor para cinco distintas vias metabólicas. O corismato é necessário para a produção de aminoácidos aromáticos, ácido para-aminobenzóico (PABA), folato, e outros metabólitos cíclicos como ubiquinonas e menaquinonas (Ratledge, 1982). A via do folato também é ausente em mamíferos, e suas enzimas estão sendo exploradas com sucesso como alvos antibacteriais, como por exemplo, a inibição da diidrofolato redutase (DHFR) pelo trimetoprim.

A modelagem molecular da corismato sintase de *M. tuberculosis* (MtCS) complexada com FMN e EPSP foi realizada, assim como análise de dicroísmo circular da proteína na presença e ausência de seu cofator.

3.3.1 Materiais e Métodos

Modelagem Molecular

Para construção do modelo da MtCS foi utilizada a mesma abordagem do modelo MtEPSPS, como descrito anteriormente. As coordenadas atômicas da estrutura cristalográfica da corismato sintase de *Streptococcus pneumoniae* (SpCS) (código acesso PDB: 1QX0) (Maclean & Ali, 2003) resolvida a 2,0 Å resolução foi utilizada como modelo inicial para modelagem da MtCS complexada com FMN e EPSP. As coordenadas atômicas de todas as águas foram removidas da estrutura da SpCS. Um total de 1000 modelos foram gerados para MtCS, e o modelo final foi selecionado baseando-se na sua qualidade estereoquímica.

Análise do modelo

A qualidade estereoquímica do modelo final do complexo MtCS-FMN-EPSP foi analisada utilizando o programa PROCHECK (Laskowski, 1994). Os modelos atômicos foram superpostos usando o programa LSQKAB (Colaborative Computational Project nº 4, 1994). As ligações de hidrogênio e as pontes salinas foram consideradas até 3,6 Å. As diferenças do desvio da raiz média quadrática (r.m.s.d.) da geometria ideal para comprimento de ligação e ângulos de ligação foram calculadas usando o programa X-PLOR (Brünger, 1992). O G-fator é essencialmente



uma análise obtida através de parâmetros estereoquímicos. Esta análise é calculada seguindo as propriedades: ângulos de torção (analisa a distribuição dos ângulos ϕ - ψ , χ^1 - χ^2 , χ^1 , χ^3 , χ^4 , e ω para cada um dos 20 tipos de aminoácidos) e geometria covalente (comprimentos e ângulos de ligação para cadeia principal). A compatibilidade do modelo tri dimensional da proteína com a sua seqüência primária foi analisada utilizando o programa VERIFY-3D (Bowie *et al.*, 1991; Luthy *et al.*, 1992).

Dicroísmo Circular

A clonagem, expressão e purificação da MtCS foi realizada como descrito em Dias *et al.* (Dias *et al.*, 2004). Espectros de dicroísmo circular foram obtidos entre os comprimentos de ondas de 195-260 nm, usando um espectropolarímetro Jasco-710 (Jasco, Tóquio-Japão) acoplado a um banho térmico Neslab RTE111. Os espectros foram obtidos a 25° C usando celas com caminho óptico de 0,2 cm. Os dados foram adquiridos com uma velocidade de leitura de 20 nm/min, largura de banda de 1,0 nm, 1,0s como tempo de resposta e 0,1 nm de resolução. Um total de cinco acumulações foram realizadas para obtenção de um espectro médio final. A deconvolução dos dados foi feita utilizando o programa DICROPROT (Deléage & Geourjon, 1993).

3.3.2 Resultados e discussão

Qualidade do modelo

O alinhamento das seqüências primárias da MtCS (alvo) e SpCS (template) apresenta 44% de identidade (Figura 3.21). Este grau de identidade entre as



seqüências demonstra que a estrutura cristalográfica da SpCS representa um bom “molde” para ser usado na modelagem da MtCS.

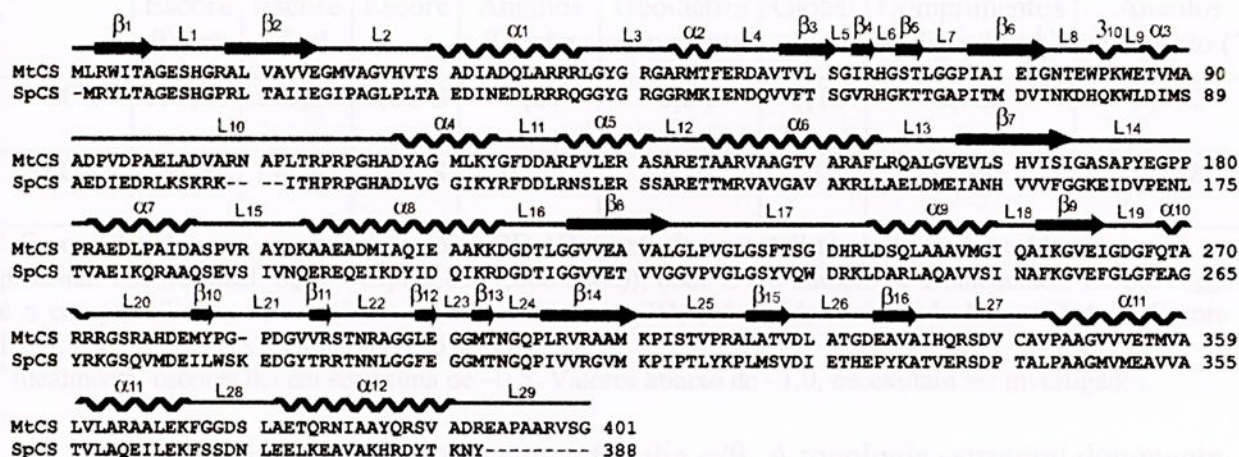


Figura 3.21 – O alinhamento das seqüências da MtCS e SpCS. Os elementos de estrutura secundária para o modelo MtCS são indicados de acordo com a região onde são encontrados. O alinhamento foi realizado com o programa CLUSTAL W (Thompson et al., 1994).

O diagrama *phi* (ϕ) e *psi* (ψ) (Diagrama de Ramachandran) foi utilizado para comparar a qualidade estereoquímica do modelo MtCS e da estrutura da SpCS resolvida por cristalografia. Análise do modelo da MtCS apresentou 94,6 % dos resíduos nas regiões mais favoráveis e 5,4% dos resíduos nas regiões permitidas. A mesma análise para a estrutura cristalográfica da SpCS foi realizada e verificou-se 93,1% dos resíduos nas regiões mais favoráveis e 6,9% dos resíduos nas regiões permitidas. A excelente qualidade do *template* permitiu que o modelo gerado também possuísse uma excelente qualidade estereoquímica. Os valores de r.m.s.d. para comprimento e ângulo de ligação com relação à geometria ideal, e valores correspondentes a compatibilidade da seqüência primária do modelo da MtCS com sua estrutura tri dimensional são mostrados na Tabela 3.9.

Tabela 3.9 – Resultados das análises estruturais da MtCS e SpCS.

Análise do modelo da MtCS e da estrutura da SpCS								
Proteínas	Verify-3D ^a			G-factor ^b			rmsd da geometria ideal	
	Escore Total	Escore Ideal	Escore S _{ideal}	Ângulos Torsão	Geometria Covalente	Global	Comprimentos de ligação (Å)	Ângulos de ligação (°)
MtCS	157,49	179,29	0,88 S	-0,07	-0,34	-0,16	0,022	3,172
SpCS	163,85	171,92	0,95 S	0,00	-0,16	-0,03	0,027	2,558

^a Escore Total: é a soma dos escores 3D-1D (preferência estatística) de cada resíduo presente na proteína. Escore Ideal: $S_{ideal} = \exp(-0,83+1,008 \times \ln(L))$; onde L é o número de aminoácidos. Escore S_{ideal}: é a compatibilidade da seqüência com sua estrutura 3D, e é obtido através do Escore Total / Escore Ideal. Escore S_{ideal} deve ser acima de 0,45S_{ideal}.

^b Idealmente, escores devem ser acima de -0,5. Valores abaixo de -1,0, necessitam ser investigados.

A estrutura da MtCS é pertencente a família α/β . A topologia estrutural dominante das CS é um sanduíche beta-alfa-beta, onde cada monômero de CS é formado por hélices na região central, na ordem $\alpha 1$, $\alpha 6$, $\alpha 12$ e $\alpha 9$, entre duas folhas betas compostas por quatro fitas antiparalelas, na ordem $\beta 1$, $\beta 2$, $\beta 6$ e $\beta 3$ para folha “um” e na ordem $\beta 7$, $\beta 8$, $\beta 14$ e $\beta 9$ para folha “dois” (Figura 3.21A). A topologia central da MtCS apresenta um pseudo eixo de ordem 2, embora esta simetria não seja observada para toda a molécula. O sítio ativo da MtCS está localizado entre as folhas 1 e 2 e este é formado pelas extremidades das hélices α centrais, e também pelos *loops* L1, L4, L10, L16, L25 e L27.

Análise dos alinhamentos das seqüências da corismato sintase de *Streptococcus pneumoniae*, *Helicobacter pylori*, *Aquifex aeolicus*, e *Saccharomyces cerevisiae* indicam que os resíduos envolvidos para a estabilização do tetrâmero são conservados na seqüência do *Mycobacterium tuberculosis*. Baseados nestas evidências, foi proposto que a estrutura quaternária da MtCS se apresente como um tetrâmero, como é mostrado na Figura 3.21B.



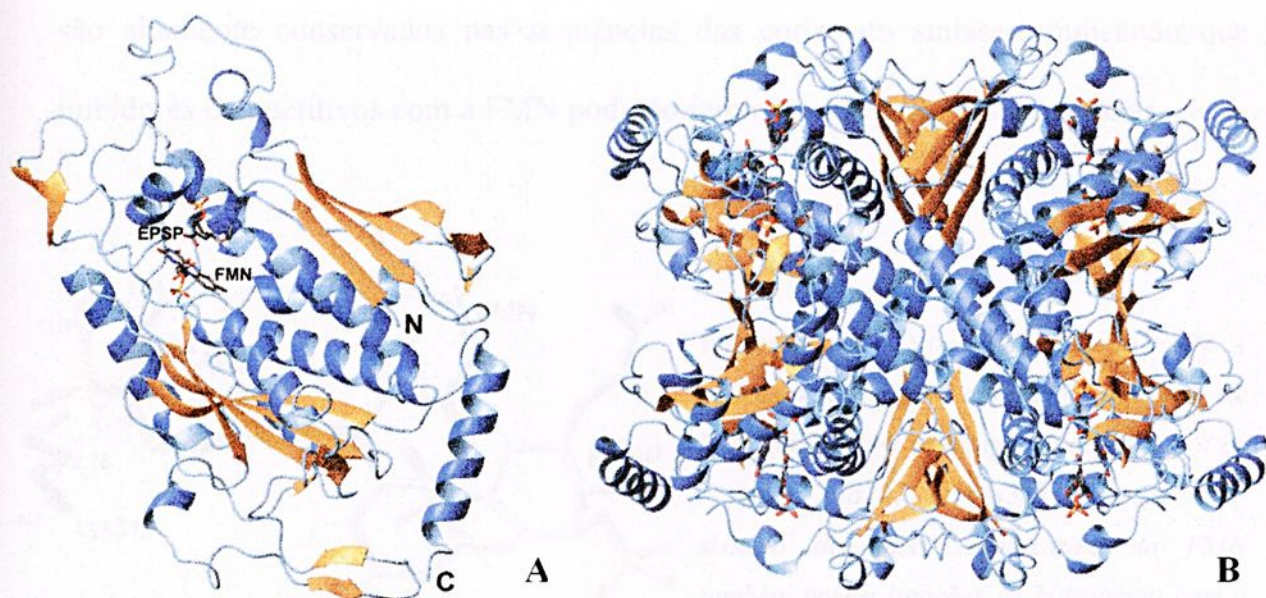


Figura 3.21 – Diagrama de Ribbon do modelo molecular da MtCS complexado com FMN e EPSP. A) Forma monomérica e B) Forma oligomérica (tetrâmero). A figura foi gerada utilizando o programa MolMol (Koradi *et al.*, 1996).

Sítio de ligação da FMN

O potencial eletrostático da superfície molecular da MtCS obtido pelo programa GRASP (Nicholls *et al.*, 1991), demonstrou uma grande concentração de resíduos positivos ao redor do sítio de ligação da FMN. Análise da superfície molecular do modelo da MtCS e da estrutura da SpCS indicaram que a FMN possui uma área de contato de $261 \text{ \AA}^2$ e $236 \text{ \AA}^2$, respectivamente.

A molécula de FMN apresenta poucas interações polares com a proteína, sendo a maioria contatos entre os grupos hidroxila e os oxigênios do fosfato com os resíduos presentes nos *loops* L10, L16, e L25. A figura 3.23A mostra as ligações de hidrogênio entre a molécula de FMN e MtCS. Foram observadas 6 ligações de hidrogênio, envolvendo os resíduos Ala257, Lys315, Thr319 e Ser342. Estes resíduos

são altamente conservados nas seqüências das corismato sintases, indicando que inibidores competitivos com a FMN poderão inibir grande parte destas enzimas.

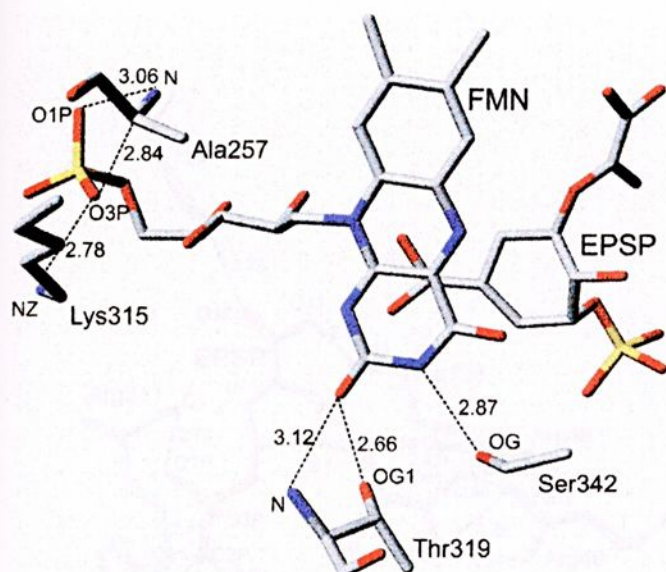


Figura 3.23A – Sítio de ligação da FMN. A molécula de FMN apresenta ligações de hidrogênio entre os oxigênios do fosfato e os resíduos Ala257 e Lys315. A região do sistema de anéis isoaloxazine da FMN também possui ligações de hidrogênio com a proteína, envolvendo os resíduos Thr319 e Ser342. Figura gerada utilizando o programa MolMol (Koradi et al., 1996).

Sítio de ligação do EPSP

A molécula do EPSP faz muitas interações polares com a MtCS. O sítio de ligação do EPSP é extremamente básico, com seis resíduos de arginina (R40, R41, R46, R49, R112, R139) e três histidinas (H11, H115, H339). Análise das ligações de hidrogênio entre a enzima e o EPSP indicaram um total de 14 interações. A molécula do EPSP pode ser dividida em três regiões: carboxila, enol-piruvil e do grupo fosfato. A região do enol-piruvil interage com três resíduos de argininas (R40, R46 e R139). O grupo fosfato faz ligações com a H11, R49, E81 e R341. O grupo carboxila forma ligações de hidrogênio com a H115 e A138 (Figura 3.23B).

Estudos espectroscópicos em corismato sintase de *Escherichia coli* demonstraram que molécula do EPSP pode ser responsável por induzir mudanças conformacionais

na estrutura da MtCS, diferentemente da FMN que pode ser acomodada no sítio ativo sem mudanças estruturais (Macheroux *et al.*, 1996a; Macheroux *et al.*, 1996b).

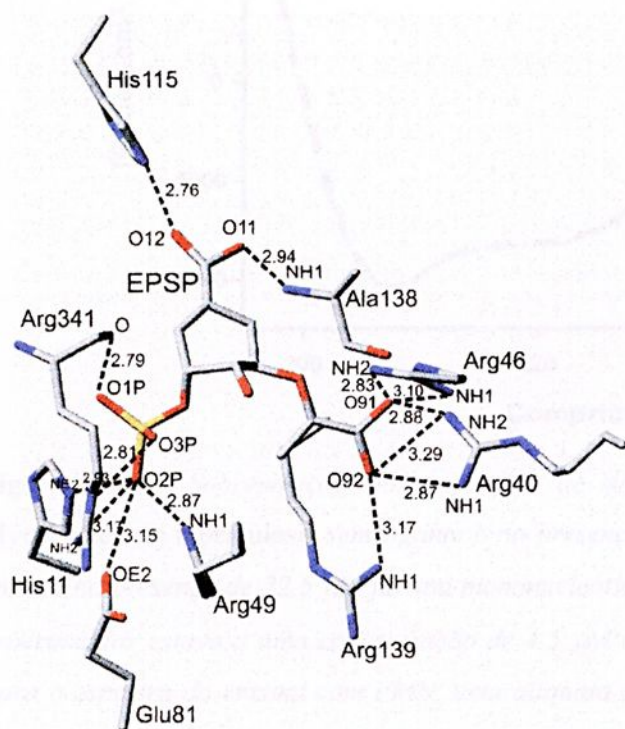


Figura 3.23B – Sítio de ligação do EPSP. A molécula de e H115 também faz ligação de hidrogênio com EPSP apresenta ligações de hidrogênio com as argininas R40, R46, R49, R139 e R341. Além das arginas, as histidinas H11 o EPSP, caracterizando o sítio como um local extremamente básico. Figura gerada utilizando o programa MolMol (Koradi *et al.*, 1996).

Dicroísmo Circular

Os espectros de CD da MtCS na presença e na ausência de FMN são mostrados na figura 3.24. As porcentagens obtidas dos elementos de estruturas secundárias para as duas medidas independentes e do modelo da MtCS pode ser observado na tabela 3.10.

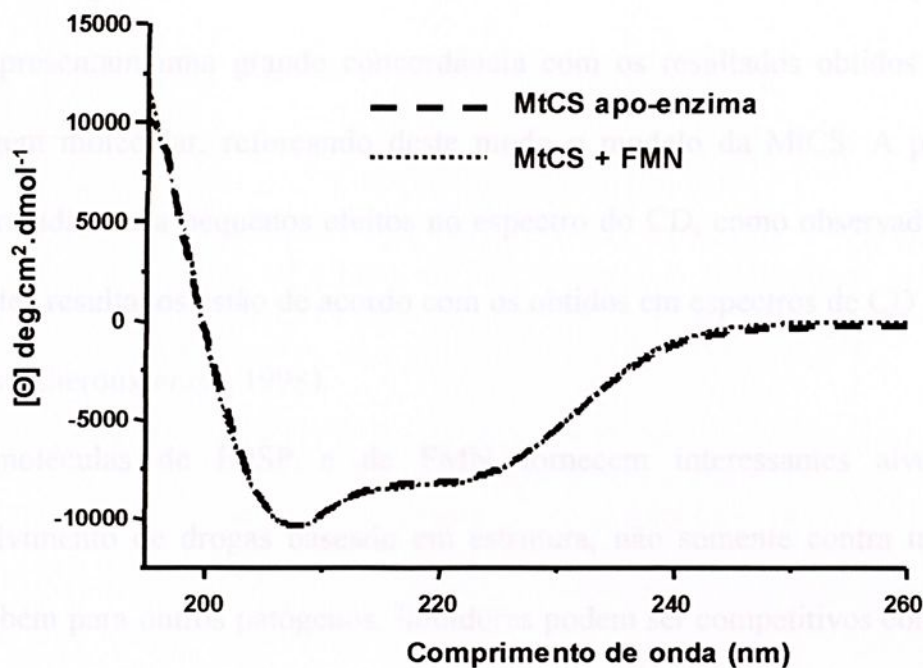


Figura 3.24 – Sobreposição dos espectros de dicroísmo circular da corismato sintase de *Mycobacterium tuberculosis* sem ligante e na presença de ligante (---) Enzima sem ligante; (.....) enzima na presença de 22,5 μ M flavina mononucleotideo (FMN). Solução de proteína utilizada nos experimentos estava a uma concentração de 4,5 μ M em tampão Fosfato de Sódio 5 mM (pH 7,8). Para o espectro da enzima com FMN, uma alíquota da solução do ligante a 10mM em Fosfato de Sódio 5 mM (pH 7,8) foi adicionado a solução de proteína, e no espectro resultante foi corrigido a diluição. O espectro de CD para enzima complexada com FMN também foi corrigido a contribuição do ligante.

Tabela 3.10 – Porcentagem dos elementos de estrutura secundária calculado para os dados experimentais de CD e para o modelo molecular da MtCS.

	Modelo Molecular da MtCS ^a	MtCS com FMN ^b	MtCS sem FMN ^b
Folhas β	18	18	16
Hélices α	40	44	45
Estr. aleatória	42	38	39

^a Elementos de estr. secundária foram calculados usando o programa MOLMOL (Koradi *et al.*, 1996).

^b Elementos de estr. secundária foram calculados usando o programa de análise de CD - DICROPROT (Deléage & Geourjon, 1993).



A porcentagem das estruturas secundárias obtidas através dos dados de CD da MtCS apresentam uma grande concordância com os resultados obtidos através da modelagem molecular, reforçando deste modo o modelo da MtCS. A presença de FMN oxidada causa pequenos efeitos no espectro do CD, como observado na figura 3.24. Estes resultados estão de acordo com os obtidos em espectros de CD para CS de *E.coli* (Macheroux *et al.*, 1998).

As moléculas de EPSP e de FMN fornecem interessantes alvos para o desenvolvimento de drogas baseado em estrutura, não somente contra tuberculose, mas também para outros patógenos. Inibidores podem ser competitivos com substrato e co-fatores, levando a inativação da enzima. A molécula de 5-deazaflavina (Lauhon & Bartlett, 1994), um análogo de FMN onde o nitrogênio da posição 5 é substituído por um carbono, e (6R)-6-fluoro-EPSP, onde o hidrogênio-6R do EPSP é substituído por um flúor, resulta na ausência de atividade da corismato sintase (Macheroux, 1998; Bornemann *et al.*, 2003). Estes são exemplos de moléculas que foram racionalmente modificadas e possuem a capacidade de inibição da enzima.



4. Referências Bibliográficas

- Baptista, I. M., Oelemann, M. C, Opromolla, D. V., Suffys, P. N. (2002).** 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*, **97**, 1147-1152.
- Bennett, W.S., Steitz, T.A. (1980).** Structure of a complex between yeast hexokinase A and glucose. II. Detailed comparisons of conformation and active site configuration with the native hexokinase B monomer and dimer. *J. Mol. Biol.* **140**, 211-230.
- Bentley, R. (1990).** The shikimate pathway - A metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol.* **25**, 307-384.
- Bloom, B.R., and Murray, C.J.L. (1992).** Tuberculosis: Commentary on a reemergent killer. *Science*, **257**, 1055-1064.
- Bornemann, S., Lawson, D.M., Thorneley, R.N. (2003).** A branch point in chorismate synthase research. *Structure*. **11**, 1463-1465.
- Bowie, J.U., Luthy, R., Eisenberg, D. (1991).** A method to identify protein sequences that fold into a known three-dimensional structure. *Science*. **253**, 164-170.
- Brünger, A. T. (1992).** *X-PLOR Version 3.1: A system for crystallography and NMR*. Yale University Press, New Haven.
- Butler, J.R., Alworth, W.L., Nugent, M.J. (1974).** *J. Am. Chem. Soc.* **96**, 1617-1618.
- Cerasoli, E., Kelly, S. M., Coggins, J. R., Laphorn, A. J., Clarke, D. T., Price, N. C. (2003).** Effects of salts on the function and conformational stability of shikimate kinase. *Biochim. Biophys. Acta* **1648**, 43-54.
- Clemmens, D.L. (1996).** Characterization of the *Mycobacterium tuberculosis* phagosome. *Trends Microbiol.*, **4**, 13-118.
- Coggins, J.R. (1989).** The shikimate pathway as a target for herbicides. In *Herbicides and Plant Metabolism* (Dodge, A., ed.), Cambridge University Press, Cambridge, UK, pp. 97-112.



Coggins, J.R., Duncan, K., Anton, I.A., Boocock, M.R., Chaudhuri, S., Lambert, J.M., Lewendon, A., Millar, G., Mousdale, D.M., Smith, D.D., (1987). The anatomy of a multifunctional enzyme. *Biochem. Soc. Trans.* **15**, 754–759.

Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M.A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J.E.; Taylor, K.; Whitehead, S., Barrell, B.G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, **393**, 537-544.

Collaborative computational project, No. 4, The CCP4 suite: programs for protein crystallography. (1994) *Acta Crystallog. sect. D.* **50**, 760–763.

Corbett, E.L., Watt, C.J., Walker, N., Mahier, D., Williams, B.G., Raviglione, M.C., Dye, C. (2003). The growing burden of tuberculosis. *Arch. Intern. Med.*, **163**, 1009-1021.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.*, **16**, 10881-10890.

Davies, G.M., Barret-Bee, K.J., Jude, D.A., Lehan, M., Nichols, W.W., Pinder, P.E. (1994). (6S)-6-fluoroshikimic acid, an antibacterial agent acting on the aromatic biosynthetic pathway. *Antimicrob. Agents Chemother.*, **38**, 403-406.

Davis, B.D., Mingioli, E. S. (1953). Aromatic biosynthesis. VII. Accumulation of two derivatives of shikimic acid by bacterial mutants. *J. Bacteriol.* **66**, 129–136.

de Azevedo, W.F. Jr., Mueller-Dieckmann, H.J., Schulze-Gahmen, U., Worland, P.J., Sausville, E., 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. USA.* **93**(7):2735-2740.

De Feyter, R. C., Pittard, J. (1986). Purification and properties of shikimate kinase II from *Escherichia coli* K-12. *J. Bacteriol.* **165**, 331-333.

De Feyter, R.C., Davidson, B.E., Pittard, J. (1986). Nucleotide sequences of the transcription unit containing the *aroL* and *aroM* genes from *Escherichia coli* K-12. *J. Bacteriol.* **165**: 233–239.



- Deléage, G., Geourjon, C. (1993).** An interactive graphic program for calculating the secondary structures content of proteins from circular dichroism spectrum. *Comp. Appl. Biosc.* **9**, 197-199.
- Dias, M. V. B., Ely, F., Canduri, F., Pereira, J. H., Frazzon, J., Basso, L. A., Palma, M. S., De Azevedo Jr., W. F., Santos, D. S. (2004).** Crystallization and preliminary X-ray crystallographic analysis of chorismate synthase from *Mycobacterium tuberculosis*. *Acta Crystal. Section D.* **60**, 2003-2005.
- Dreusicke, D., Karplus, A., Schulz, G. E. (1988).** Refined structure of porcine cytosolic adenylate kinase at 2.1 Å resolution. *J. Mol. Biol.* **199**, 359-371.
- Drobniewski, F. A., Caws, M., Gibson, A., Young, D. (2003).** Modern laboratory diagnosis of tuberculosis. *Lancet Infect. Dis.*, v. **3**, p.141-147.
- Duncan, K. (2003).** Progress in TB drug development and what is still needed. *Tuberculosis*, v. **83**, p. 201-207.
- Duncan, K., Edwards, R.M., Coggins, J.R. (1987).** The pentafunctional arom enzyme of *Saccharomyces cerevisiae* is a mosaic of monofunctional domains. *Biochem. J.* **246**, 375-386.
- El Sayed, K. A., Bartyzel, P., Shen, X. Y., Perry, T. L., Kjawiony, J. K., Hamann, M. T. (2000).** Marine natural products as antituberculosis agents. *Tetrahedron*, **56**, 949-953.
- Ely, B., Pittard, J. (1979).** Aromatic amino acid biosynthesis: Regulation of shikimate kinase in *Escherichia coli* K-12. *J. Bacteriol.* **138**: 933-943.
- Funasa, 2002.** Situação da Prevenção e Controle das Doenças Transmissíveis no Brasil. Setembro/2002 [http://www.funasa.gov.Br/guia_epi/]
- Gerstein, M., Schulz, G. E., Chothia, C. (1993).** Domain closure in adenylate kinase: Joints on either side of two helices close like neighbouring fingers. *J. Mol. Biol.* **229**, 494-501.
- Gu, Y., Reshetnikova, L., Li, Y., Wu, Y., Yan, H., Singh, S., Ji, X. (2002).** Crystal Structure of Shikimate Kinase from *Mycobacterium tuberculosis* Reveals the Dynamic Role of the LID Domain in Catalysis. *J. Mol. Biol.* **319**, 779-789.
- Gu, Y., Reshetnikova, L., Li, Y., Yan, H., Singh, S. V., Ji, X. (2001).** Crystallization and preliminary X-ray diffraction analysis of shikimate kinase from *Mycobacterium tuberculosis* in complex with MgADP. *Acta crystallographica Section D*, **D57**, 1870-1871.



- Guex, N., Peitsch, M.C. (1997).** SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. **18**, 2714-2723.
- Haslam, E. (1993).** Shikimic acid: metabolism and metabolites. John Wiley & Sons, Chichester, United Kingdom.
- Hawgood, B.J. (1999).** Doctor Albert Calmette 1863-1933: founder of antivenomous serotherapy and of antituberculous BCG vaccination. *Toxicon*, **37**, 1241-1258.
- Hiriyanna, K.T., Ramakrishnan, T. (1986).** Deoxiribonucleic-acid replication time in *Mycobacterium tuberculosis*. *Arch. Microbiol.*, **144**, 105-109.
- Jencks, W. P. (1975a).** In *Catalysis in Chemistry and Enzymology*. Dover Publications, Inc., New York, USA.
- Jencks, W. P. (1975b).** Binding energy, specificity and enzymatic catalysis: the circe effect. *Adv. Enzymol.* **43**, 219-410.
- Kim, S.-H., Schulze-Gahmen, U., Brandsen, J., de Azevedo, W. F. Jr. (1996).** Structural basis for chemical inhibitor of CDK2. *Prog. in Cell Cycle Res.* **2**, 137-145.
- Kishore, G.M., Shah, D.M. (1988).** Amino acid biosynthesis inhibitors as herbicides. *Annu. Rev. Biochem.* **57**, 627-663.
- Koradi, R., Billeter, M., Wüthrich, K. (1996).** MOLMOL: A program for display and analysis of macromolecular structures. *J. Molecular Graphics*, **14**(1), 51-55.
- Krell, T., Coggins, J. R., Laphorn, A. J. (1998).** The three-dimensional structure of shikimate kinase. *J. Mol. Biol.* **278**, 983-997.
- Krell, T., Maclean, J., Boam, D. J., Cooper, A., Resmini, M., Brocklehurst, K. (2001).** Biochemical and X-ray crystallographic studies on shikimate kinase: the important structural role of the P-loop lysine. *Protein Sci.* **10**, 1137-1149.
- Laskowski, R. A., MacArthur, M. W., Smith, D. K., Jones, D. T., Hutchinson, E. G., Morris, A. L., Naylor, D., Moss, D. S., Thornton, J. M. (1994).** PROCHECK v.3.0 - Program to check the stereochemistry quality of Protein structures - Operating instructions.
- Lauhon, C.T., Bartlett, P.A. (1994).** Substrate analogs as mechanistic probes for the bifunctional chorismate synthase from *Neurospora crassa*. *Biochemistry*. **33**, 14100-14108.
- Leipe, D. D., Koonin, E. V., Aravind, L. (2003).** Evolution and classification of P-loop kinases and related proteins. *J. Mol. Biol.* **333**, 781-815.



- Lennon, B.W., Williams Jr., C.H., Ludwig, M.L. (1999). Crystal structure of reduced thioredoxin reductase from *Escherichia coli*: structural flexibility in the isoalloxazine ring of the flavin adenine dinucleotide cofactor. *Protein Sci.* **8**, 2366–2379.
- Leslei, A.G.W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- Levin, J.G., Sprinson, D.B. (1964). The enzymatic formation and isolation os 3-EnolPyruylshikimate 5-Phosphate. *J.Biol.Chem.*, **239**, 1142-1150.
- Luthy, R., Bowie, J., Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, **356**, 83–85.
- Maclea, J., Ali, S. (2003). The structure of chorismate synthase reveals a novel flavin binding site fundamental to a unique chemical reaction. *Structure*. **11**, 1499-1511.
- Macheroux, P., Bornemann, S., Ghisla, S., Thorneley, R. N. F. (1996b). 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**, 25850-25858.
- Macheroux, P., Petersen, J., Bornemann, S., Lowe, D., Thorneley, R. N. F. (1996a). Binding of the oxidized, reduced, and radical flavin species to chorismate synthase. An investigation by spectrophotometry, fluorimetry, and electron paramagnetic resonance and electron nuclear double resonance spectroscopy. *Biochemistry*. **35**(5), 1643–1652.
- 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.
- Macheroux, P., Schönbrunn, E., Svergun, D.I., Volkov, V.V., Koch, M.H.J., Bornemann, S., Thorneley, R.N.F. (1998). Evidence for a major structural change in *Escherichia coli* chorismate synthase induced by flavin and substrate binding. *Biochem. J.* **335**, 319–327.
- Manabe, Y.C., Bishai, W.R. (2000). Latent *Mycobacterium tuberculosis* – persistence, patience, and winning by waiting. *Nat. Med.*, **6**, 1327-1329.
- Mathews, C.K., van Holde, K.E. (1990). *Biochemistry*, Benjamin Cummings, 1130p., California, USA.



- McRee, D. E. (1999).** XtalView/Xfit-A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.*, **125**, 156-165.
- Millar, G., Lewendon, A., Hunter, M. G., Coggins, J. R. (1986).** The cloning and expression of the *aroL* gene from *Escherichia coli* K-12. *Biochem. J.* **237**, 427-437.
- Mitchison, D.A. (1985).** The action of antitubercular drugs in short-course chemotherapy. *Tubercle*, **66**, 219-225.
- Mousdale, D.M., Campbell, M.S., Coggins, J.R. (1987).** Purification and characterization of bifunctional dehydroquinase-shikimate: NADP oxidoreductase from pea seedlings. *Phytochemistry*, **26**, 2665-2670.
- Müller, C. W., Schlauderer, G. J., Reinstein, J., Schulz, G. E. (1996).** Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure*, **4**, 147-156.
- Müller-Dieckmann, H.-J., Schulz, G. E. (1994).** The structure of uridylate kinase with its substrates, showing the transition state geometry. *J. Mol. Biol.*, **236**, 361-367.
- Navaza, J. (1994).** AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157-163.
- Nicholls, A., Sharp, K.A., Honig, B. (1991).** Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**(4), 281-296.
- Oliveira, J.S., Pinto, C.A., Basso, L.A., Santos, D.S. (2001).** Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*. *Protein Expression and Purification*, **22**, 430-435.
- Parish, T., Stoker, N. G. (2002).** The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. *Microbiology*, **148**, 3069-3077.
- Pelczar M. J. Jr., Chan, E.C.S., Krieg, N. R. (1996).** *Microbiologia – Conceitos e Aplicações*, Makron Books, vol. 2, 2ed., 517p., São Paulo.
- Pittard, A.J. (1987)** In F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. vol.1, pp. 368-394. American Society for Microbiology, Washington, D.C.
- Pivetta, M. (2004).** *Pesquisa Fapesp* **97**, 32-37.

Polikarpov, I., Oliva, G., Castellano, E.E., Garratt, R., Arruda, P., Leite, A., Craievich, A. (1998a). The protein crystallography beamline at LNLS, the Brazilian National Synchrotron Light Source. *Nuclear Instruments & Methods In Physics Research A*, **405**, 159-164.

Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E.E., Garratt, R., Craievich, A. (1998b). Set-up and experimental parameters of the protein crystallography beamline at the Brazilian National Synchrotron Laboratory. *J. Synchrotron Rad*, **5**, 72-76.

Ratledge, C. (1982). Nutrition, growth and metabolism, in: Ratledge, C., Stanford, J. L. (Eds.), *The biology of the mycobacteria*, Academic Press, London, Vol 1, 185-271.

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., Chakrebari, D., McLeod, R. (1998). Evidence for the shikimate pathway in apicomplexan parasites. *Nature*, **393**, 801-805.

Romanowski, M.J., Burley, S.K. (2002). Crystal structure of the *Escherichia coli* Shikimate Kinase I (AroK) That confers sensitivity to Mecillinam. *Proteins: Structure, Function, and Genetics*, **47**, 558-562.

Sali, A., Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, **234**, 779-815.

Schlauderer, G. J., Schulz, G. E. (1996). The structure of bovine mitochondrial adenylate kinase: Comparison with isoenzymes in other compartments. *Protein Sci.* **5**, 434-441.

Schönbrunn, E., Eschenburg, S., Shuttleworth, W. A., Schloss, J. V., Amrhein, N., Evans, J. N. S., Kabsch, W. (2001). Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc. Natl. Acad. Sci. USA*, **98**, 1376-1380.

Shuttleworth, W. A., Pohl, M. E., Helms, G. L., Jakeman, D. L., Evans, J. N. S. (1999). Site-directed mutagenesis of putative active site residues of 5-enolpyruvylshikimate-3-phosphate synthase. *Biochemistry*, **38**, 296-302.

Smith, C. A., Rayment, I. (1996). Active site comparisons highlight structural similarities between myosin and other P-loop proteins. *Biophys. J.*, **70**, 1590-1602.

Sprinson, D.B. (1960). The biosynthesis of aromatic compounds from D-glucose. *Adv. Carbohydrate Chem.* **15**, 235-270.



Stallings, W. C., Abdel-Meguid, S. S., Lim, L. W., Shieh, H.-S., Dayringer, H. E., Leimgruber, N. K., Stegeman, R. A., Anderson, K. S., Sikorski, J. A., Padgett, S. R., Kishore, G. M. (1991). Structure and topological symmetry of the glyphosate target 5-enol-pyruvylshikimate-3-phosphate synthase: a distinctive protein fold. *Proc. Natl. Acad. Sci. USA* **88**, 5046-5050.

Stehle, T., Schulz, G. E. (1990). Three-dimensional structure of the complex of guanylate kinase from yeast with its substrate GMP. *J. Mol. Biol.* **211**, 249-254.

Thompson, J.D., Higgins, D.G., 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.

Trouiller, P., Torreele, E., Olliaro, P., White, N., Foster, S., Wirth D., Pecoul, B. (2001). Drugs for neglected diseases: a failure of the market and a public health failure? *Tropical Medicine And International Health*, v. 6, p.945-951.

Villar, M. (2004). Toxicidade aos antibacilares.
[<http://www.sppneumologia.pt/docspdfs/Toxicidade.pdf>]. Acessado em Agosto/2005.

Vinella, D., Gagny, B., Joseleau-Petit, D., D'Ardi, R., Cashel, M. (1996). Mecillinam resistance in *Escherichia coli* is conferred by loss of a second activity of the *aroK* protein. *J. Bacteriol.* **178**, 3818-3828.

Vonrhein, C., Schlauderer, G. J., Schulz, G. E. (1995). Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure*, **3**, 483-490.

Walker, J. E., Saraste, M., Runswick, M. J., Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.

Whipp, M. J., Pittard, A. J. (1995). A reassessment of the relationship between *aroK*- and *aroL*-encoded shikimate kinase enzymes of *Escherichia coli*. *J. Bacteriol.* **177**, 1627-1629.

WHO (2001). Stop TB Annual Report 2001. World Health Organization, Geneva, Switzerland, WHO/CDS/STB/2002.17. Reprint:
[<http://www.who.int/gtb/publications/TBCatalogue.htm#2002>]. Acessado em Fevereiro/2004.

WHO (2002). Tuberculosis. Fact Sheet N° 104. World Health Organization, Geneva, Switzerland. August/2002.



[<http://www.who.int/mediacentre/factsheets/who104/en/print.html>]. Acessado em Março/2004.

WHO (2003). The global plan to stop tuberculosis. World Health Organization, Geneva, Switzerland, WHO/CDS/STB/2003.23. Reprint [<http://www.who.int/gtb/publications/TBCatalogue.htm#2003>]. Acessado em Fevereiro/2004.

WHO (2004a). Tuberculosis. Fact Sheet N° 104. World Health Organization, Geneva, Switzerland. March/2004. [<http://www.who.int/mediacentre/factsheets/fs104/en/print.html>]. Acessado em Março/2004.

WHO (2004b). Global tuberculosis control. In *WHO Report 2004*. World Health Organization, Geneva, Switzerland, WHO/HTM/TB/2004.331. Reprint [<http://www.who.int/gtb/publications/TBCatalogue.htm#2004>]. Acessado em Março/2004.

Wild, K., Bohner, T., Aubry, A., Folkers, G., Schulz, G. E. (1995). The three-dimensional structure of thymidine kinase from Herpes simplex virus type 1. *FEBS Letters*, **368**, 289-292.

Yan, H., Ysai, M.-D. (1999). Nucleoside monophosphate kinases: structure, mechanism, and substrat specificity. *Advan. Enzymol. Relat. Areas Mol. Biol.* **73**, 103-134.

Anexos



Anexos:

Anexo 1 - "Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid"

Anexo 2 - "Shikimate Kinase: A potential target for development of novel anti-tubercular agents" (Review)

Anexo 3 - "Molecular model of shikimate kinase from *Mycobacterium tuberculosis*"

Anexos

Anexo 4 - "Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*"

Anexo 5 - "Molecular modeling and CD analysis of chorismate synthase"

Anexos:

Anexo 1 - "Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid"

Anexo 2 - "Shikimate Kinase: A potential target for development of novel anti-tubercular agents" (Review)

Anexo 3 - "Molecular model of shikimate kinase from *Mycobacterium tuberculosis*"

Anexo 4 - "Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*"

Anexo 5 - "Molecular modeling and CD analysis of chorismate synthase"

José Henrique Pereira,^a
Jânio Simões de Oliveira,^b
Fernanda Cantúci,^{a,c}
Garcia Vinícius Bertolino Dias,^a
Mário Sérgio Fajina,^{a,d}
Luiz Augusto Siqueira,^{a,e}
Diógenes Santiago Santos^{a,*} and
Walter Hilgendorf de Azevedo^{a,f,g}

^aInstituto de Física de Caruaru, Universidade Federal de Pernambuco, Departamento de Física, 55087-900, Caruaru, PE, Brazil; ^bInstituto de Física de Caruaru, Universidade Federal de Pernambuco, Departamento de Física, 55087-900, Caruaru, PE, Brazil; ^cInstituto de Física de Caruaru, Universidade Federal de Pernambuco, Departamento de Física, 55087-900, Caruaru, PE, Brazil; ^dInstituto de Física de Caruaru, Universidade Federal de Pernambuco, Departamento de Física, 55087-900, Caruaru, PE, Brazil; ^eInstituto de Física de Caruaru, Universidade Federal de Pernambuco, Departamento de Física, 55087-900, Caruaru, PE, Brazil; ^fInstituto de Física de Caruaru, Universidade Federal de Pernambuco, Departamento de Física, 55087-900, Caruaru, PE, Brazil; ^gInstituto de Física de Caruaru, Universidade Federal de Pernambuco, Departamento de Física, 55087-900, Caruaru, PE, Brazil

*Correspondence: e-mail: dsantos@fisica.ufpe.br

© 2010 International Union of Crystallography
Printed in Great Britain — all rights reserved

Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid

Received 28 August 2009
Accepted 28 August 2009

DOI: 10.1107/S0107696009012574
Acta Cryst. B, 36, 1210–1218 (2010)

Tuberculosis made a resurgence in the mid 1990s and now kills approximately 3 million people a year. The re-emergence of tuberculosis as a public health threat, the high mortality rate of HIV-infected persons and the proliferation of multidrug-resistant strains have created a need to develop new drugs. Shikimate kinase and other enzymes in the shikimate pathway are attractive targets for development of new drugs against *M. tuberculosis*, but the development of new drugs, because the pathway is essential in these species, requires a detailed firm knowledge. The crystal structure of shikimate kinase from *Mycobacterium tuberculosis* (MSK) complexed with MgATP and shikimate acid (shikimate) has been determined at 2.0 Å resolution, clearly revealing the amino acid residues involved in shikimate binding. This is the first three-dimensional structure of shikimate kinase complexed with shikimate. In addition, the 3D structure that is strictly conserved in shikimate kinases forms a hydrogen bond and salt bridge with arginine and assists in positioning the gamma-phosphate group of arginine for shikimate binding. The carboxyl group of shikimate interacts with Arg227, Gly91 and Arg134 and the hydroxyl groups interact with Asp140 and Gly93. The crystal structure of MSK–MgATP–shikimate will provide crucial information for the elucidation of the mechanism of the shikimate kinase catalyzed reaction and for the development of a new generation of drugs against tuberculosis.

1. Introduction

The shikimate pathway is a seven-step biosynthetic route that generates chorismate and leads to phenylalanine, tyrosine and erythrose 4-phosphate. The shikimate pathway is an attractive target for the development of antimicrobial agents (Hosler *et al.*, 1998) and herbicides (Coggins, 1999) because it is essential in algae, higher plants, bacteria and fungi, whilst being absent from mammals (Beadles, 1995). Several enzymes of this pathway have been identified in structural study (Arora *et al.*, 2004; Pereira *et al.*, 2005) in order to propose inhibitors for these enzymes. More recently, by disruption of the *aroK* gene, which codes for the shikimate kinase enzyme, the shikimate pathway has been shown to be essential for the virulence of *Mycobacterium tuberculosis* (Hosler & Sander, 2002). Shikimate kinase (SK; EC 2.7.1.74), the fifth enzyme of the pathway, catalyzes the specific phosphorylation of the 3-hydroxyl group of shikimate with using ATP as a co-substrate.

Three previously solved structures of SK from *Erwinia caroliniana* (ECSK; Kroll *et al.*, 1998, 2001) show that SK belongs to the same structural family as nucleoside monophosphate (NMP) kinases. The NMP kinases are composed of three domains: the GMP, LID and NMP-binding (NMPB) domains. A characteristic feature of the NMP kinases is that

Anexo 1

José Henrique Pereira,^a
Jaim Simões de Oliveira,^b
Fernanda Canduri,^{a,c}
Marcio Vinicius Bertacine Dias,^a
Mário Sérgio Palma,^{c,d}
Luiz Augusto Basso,^{b,e}
Diógenes Santiago Santos^{e,*} and
Walter Filgueira de Azevedo
Jr^{a,d,*}

^aPrograma de Pós-Graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, ^bRede Brasileira de Pesquisa em Tuberculose Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil, ^cCenter for Applied Toxinology, Institute Butantan, São Paulo, SP 05503-900, Brazil, ^dLaboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil, and ^eCentro de Pesquisa e Desenvolvimento em Biologia Molecular e Funcional, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS 90619-900, Brazil

Correspondence e-mail: diogenes@puccs.br,
walterfa@df.ibilce.unesp.br

Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid

Tuberculosis made a resurgence in the mid-1980s and now kills approximately 3 million people a year. The re-emergence of tuberculosis as a public health threat, the high susceptibility of HIV-infected persons and the proliferation of multi-drug-resistant strains have created a need to develop new drugs. Shikimate kinase and other enzymes in the shikimate pathway are attractive targets for development of non-toxic antimicrobial agents, herbicides and anti-parasitic drugs, because the pathway is essential in these species whereas it is absent from mammals. The crystal structure of shikimate kinase from *Mycobacterium tuberculosis* (MtSK) complexed with MgADP and shikimic acid (shikimate) has been determined at 2.3 Å resolution, clearly revealing the amino-acid residues involved in shikimate binding. This is the first three-dimensional structure of shikimate kinase complexed with shikimate. In MtSK, the Glu61 residue that is strictly conserved in shikimate kinases forms a hydrogen bond and salt bridge with Arg58 and assists in positioning the guanidinium group of Arg58 for shikimate binding. The carboxyl group of shikimate interacts with Arg58, Gly81 and Arg136 and the hydroxyl groups interact with Asp34 and Gly80. The crystal structure of MtSK–MgADP–shikimate will provide crucial information for the elucidation of the mechanism of the shikimate kinase-catalyzed reaction and for the development of a new generation of drugs against tuberculosis.

Received 24 August 2004
Accepted 5 October 2004

PDB Reference: shikimate kinase–MgADP–shikimate complex, 1we2, r1we2sf.

1. Introduction

The shikimate pathway is a seven-step biosynthetic route that generates chorismic acid from phosphoenol pyruvate and erythrose 4-phosphate. The shikimate pathway is an attractive target for the development of antimicrobial agents (Davies *et al.*, 1994) and herbicides (Coggins, 1989) because it is essential in algae, higher plants, bacteria and fungi, whilst being absent from mammals (Bentley, 1990). Several enzymes of this pathway have been submitted to structural study (Arcuri *et al.*, 2004; Pereira *et al.*, 2003) in order to propose inhibitors for these enzymes. More recently, by disruption of the *aroK* gene, which codes for the shikimate kinase enzyme, the shikimate pathway has been shown to be essential for the viability of *Mycobacterium tuberculosis* (Parish & Stoker, 2002). Shikimate kinase (SK; EC 2.7.1.71), the fifth enzyme of the pathway, catalyses the specific phosphorylation of the 3-hydroxyl group of shikimic acid using ATP as a co-substrate.

Three previously solved structures of SK from *Erwinia chrysanthemi* (EcSK; Krell *et al.*, 1998, 2001) show that SK belongs to the same structural family as nucleoside monophosphate (NMP) kinases. The NMP kinases are composed of three domains: the CORE, LID and NMP-binding (NMPB) domains. A characteristic feature of the NMP kinases is that

they undergo large conformational changes during catalysis (Vonnrhein *et al.*, 1995).

Three functional motifs of nucleotide-binding enzymes are recognizable in *M. tuberculosis* SK, including a Walker A motif, a Walker B motif and an adenine-binding loop. The Walker A motif is located between the first β -strand ($\beta 1$) and first α -helix ($\alpha 1$), containing the conserved sequence GXXXXGKT/S (Walker *et al.*, 1982), where *X* represents any residue. This motif forms the phosphate-binding loop (P-loop; Smith & Rayment, 1996). In addition to the Walker A motif, a second conserved sequence ZZDXXG called the Walker B motif (Walker *et al.*, 1982) is observed, where *Z* represents a hydrophobic residue. An Asp→Ser replacement exists in MtSK. The Walker B motif consensus in SKs is ZZZTGGG and the second glycine (Gly80 in MtSK) has been implicated in hydrogen bonding to the γ -phosphate of ATP. The adenine-binding loop motif may be described as an IVDXXX(X)XP sequence stretch (Gu *et al.*, 2002).

The shikimate-binding domain has previously been assigned on the basis of the difference Fourier map of EcSK and structural comparison with NMP kinases. Shikimate was included in the crystallization conditions (co-crystallization); however, the electron density was not clear enough to include shikimate in the previously solved crystallographic structure (Krell *et al.*, 1998).

Two crystal structures of SK from *M. tuberculosis* (MtSK; Gu *et al.*, 2002; PDB codes 1l4y and 1l4u) have revealed the dynamic role of the LID domain in catalysis, but the precise position and interactions between shikimate and MtSK were not unequivocally demonstrated as the shikimate-binding site was not occupied by the substrate. A previously reported molecular-modelling study of the complex between MtSK and shikimate also failed to predict all the intermolecular hydrogen bonds (de Azevedo *et al.*, 2002). Here, we describe the crystal structure of the MtSK–MgADP–shikimate ternary complex at 2.3 Å resolution, unequivocally revealing in detail the interactions of amino-acid residues with bound shikimate and the conformational changes upon substrate binding. The crystal structure of MtSK–MgADP–shikimate will provide crucial information for the design of non-promiscuous SK inhibitors that target both the shikimate- and ATP-binding pockets or uniquely the shikimate-binding site.

2. Materials and methods

2.1. Crystallization

Cloning, expression and purification have been reported elsewhere (Oliveira *et al.*, 2001). MtSK was concentrated and dialyzed against 50 mM Na HEPES buffer pH 7.5 containing 0.5 M NaCl and 5.0 mM MgCl₂. This protein solution was brought to 13.0 mM in shikimate and ADP by addition of the pure solids and centrifuged prior to crystallization. The protein concentration was about 17.0 mg ml⁻¹. Crystals were obtained by the hanging-drop vapour-diffusion method. The well solution contained 0.1 M Na HEPES buffer pH 7.5, 10%

2-propanol, 35% PEG 3350 and the drops consisted of a mixture of 1.0 µl well solution and 1.5 µl protein solution.

2.2. Data collection and processing

The data set for MtSK–MgADP–shikimate was collected at a wavelength of 1.431 Å using the Synchrotron Radiation Source (Station PCr, LNLS, Campinas, Brazil; Polikarpov *et al.*, 1998) and a CCD detector (MAR CCD). The cryo-protectant contained 15% glycerol, 12% PEG 3350 and 3.5% propanol. The crystal was flash-frozen at 104 K in a cold nitrogen stream generated and maintained with an Oxford Cryosystem. The oscillation range used was 1.0°, the crystal-to-detector distance was 90 mm and the exposure time was 50 s. A data set containing 160 frames was collected and processed to 2.3 Å resolution using the program *MOSFLM* (Leslie, 1992) and scaled with *SCALA* (Collaborative Computational Project, Number 4, 1994).

2.3. Molecular replacement and crystallographic refinement

The crystal structure of MtSK–MgADP–shikimate was determined by standard molecular-replacement methods using the program *AMoRe* (Navaza, 1994), using as a search model the structure of MtSK–MgADP (PDB code 1l4y; Gu *et al.*, 2002). After translation-function computation the correlation was 65% and the *R* factor was 38.3%. The highest magnitude of the correlation coefficient function was obtained for the Euler angles $\alpha = 54.85^\circ$, $\beta = 85.32^\circ$, $\gamma = 91.90^\circ$. The fractional coordinates are $T_x = 0.7336$, $T_y = 0.5326$, $T_z = 0.2768$. The atomic positions obtained from molecular replacement were used to initiate the crystallographic refinement. Structure refinement was performed using *X-PLOR* (Brünger, 1992). During rigid-body refinement, the *R* factor decreased from 38.3 to 35.8%. Further refinement continued with simulated annealing using the slow-cooling protocol, followed by alternate cycles of positional refinement and manual rebuilding using *XtalView* (McRee, 1999). Finally, the positions of waters, MgADP and shikimate were checked and corrected in $F_{\text{obs}} - F_{\text{calc}}$ maps. The final model has an *R* factor of 20.7% and an R_{free} of 28.7%.

Root-mean-square deviation differences from ideal geometries for bond lengths, angles and dihedrals were calculated with *X-PLOR* (Brünger, 1992). The overall stereochemical quality of the final model for MtSK–MgADP–shikimate was assessed by the program *PROCHECK* (Laskowski *et al.*, 1994). Atomic models were superposed using the program *LSQKAB* from *CCP4* (Collaborative Computational Project, Number 4, 1994).

The molecular-surface areas were calculated using the program *Swiss PDB Viewer* v3.7 (<http://www.expasy.org/spdbv>), a probe radius of 1.4 Å and a fixed radius for all atoms.

3. Results and discussion

The crystals of MtSK–MgADP–shikimate were grown in the presence of 5.0 mM MgCl₂, 13.0 mM ADP and 13.0 mM shikimate. The data set of MtSK–MgADP–shikimate was

Table 1

Summary of data-collection statistics for MtSK-MgADP-shikimate.

Values in parentheses are for the highest resolution shell.

Unit-cell parameters	
$a = b$ (Å)	62.91
c (Å)	90.92
Resolution (Å)	29.74–2.30 (2.41–2.30)
Space group	$P3_221$
No. of measurements with $I > 2\sigma(I)$	34274
No. of independent reflections	9563
Overall redundancy	2.4
Completeness (%)	98.7 (98.7)
R_{sym} (%)	3.0 (7.2)

collected at 2.3 Å (Table 1) using the Synchrotron Radiation Source (Polikarpov *et al.*, 1998) and the structure was solved by molecular replacement. There is one molecule of MtSK-MgADP-shikimate in the asymmetric unit, containing residues 2–166, Mg^{2+} , ADP, shikimate, two Cl^- ions and 144 water molecules (Fig. 1). Therefore, the enzyme crystallized as a MtSK-MgADP-shikimate dead-end ternary complex. The N-terminal methionine residue is not observed and the ten C-terminal residues (NQIHHMLESN) are disordered. Details of the refinement and the final model statistics are presented in Table 2. Analysis of the Ramachandran diagram ϕ - ψ plot shows that 93.4% of non-glycine residues lie in most favoured regions and there are no residues in the disallowed region (Fig. 2). The average B factor for main-chain atoms is 34.64 Å², whereas that for side-chain atoms is 35.68 Å² (Table 2). In order to accurately determine the position of shikimate binding in the active site of MtSK, larger final

Table 2

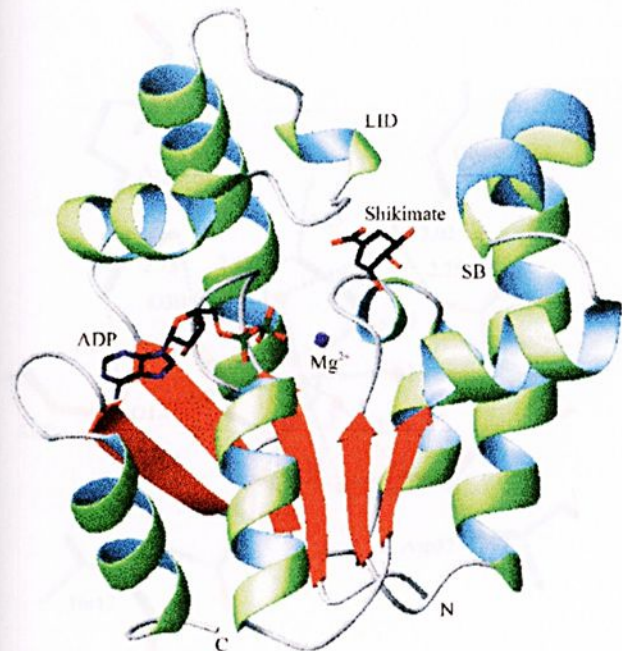
Summary of refinement statistics for MtSK-MgADP-shikimate.

Values in parentheses are for the highest resolution shell.

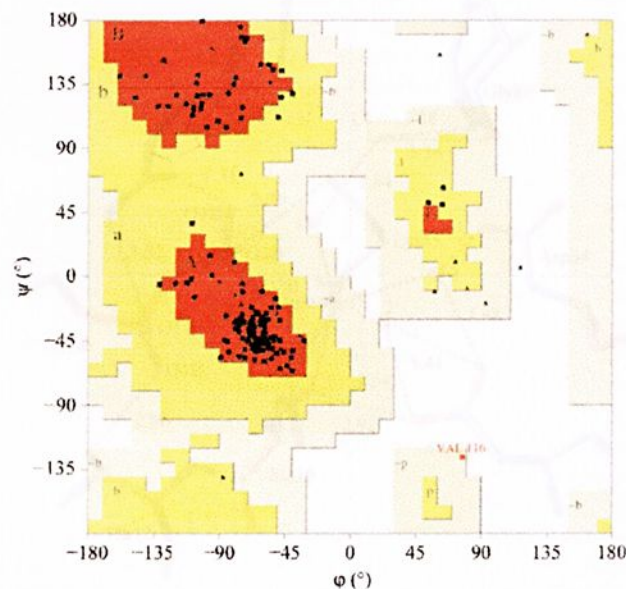
Resolution range (Å)	6.00–2.30
Reflections used for refinement	8885
No. of residues	165
No. of water O atoms	144
No. of ADP molecules	1
No. of metal ions (Mg^{2+})	1
No. of Cl^-	2
No. of shikimate molecules	1
Final R factor [†] (%)	20.7 (31.8)
Final R_{free} [‡] (%)	28.7 (36.9)
B factors§ (Å ²)	
Main chain	34.64
Side chains	35.68
Waters	39.63
ADP	21.81
Mg^{2+}	34.70
Cl^-	37.65
Shikimate	26.15
Observed r.m.s.d. from ideal geometry	
Bond lengths (Å)	0.017
Bond angles (°)	1.905
Dihedrals (°)	22.125
Ramachandran plot	
Most favoured ϕ/ψ angles (%)	93.4
Disallowed ϕ/ψ angles (%)	0

[†] R factor = $100 \times \sum (I_{\text{obs}} - I_{\text{calc}}) / \sum I_{\text{obs}}$, the sums being taken over all reflections with $I/\sigma(I) > 2\sigma(I)$. [‡] R_{free} = R factor for 10% of the data that were not included during crystallographic refinement. [§] B values = average B values for all non-H atoms.

concentrations of ADP and shikimate in the drop (8 mM) were used than those used by Gu *et al.* (2002) to obtain the crystals (4 mM). The average B -factor value of 26.15 Å² obtained for shikimate indicates a higher order and occupancy of this substrate in the present structure than in the previously solved SK structures, where either shikimate was not

**Figure 1**

MtSK complexed with Mg^{2+} , ADP and shikimate. The LID (residues 112–124) and SB (residues 33–61) domains are responsible for large conformational changes during catalysis and are shown. The figure was prepared with MOLMOL (Koradi *et al.*, 1996).

**Figure 2**

Ramachandran diagram for MtSK-MgADP-shikimate. 93.4% of non-glycine residues lie in the most favoured regions and no residues are in the disallowed region.

complexed or its electron density was too poor to accurately locate the substrate (Table 2) (Gu *et al.*, 2002; Krell *et al.*, 1998).

SK displays an α/β -fold and consists of five central parallel β -strands flanked by α -helices. As pointed out above, a characteristic feature of SKs is that they undergo large conformational changes during catalysis. There are two flexible

regions of the structure that are responsible for movement: the SB and LID domains (Fig. 1), which in MtSK correspond to residues 33–61 and 112–124, respectively.

3.1. Interaction with ADP/Mg²⁺

The structures of MtSK complexed with MgADP (Gu *et al.*, 2002) are highly similar, with a root-mean-square deviation of 0.19 Å for all pairs of C α atoms. The adenine moiety of ADP is sandwiched between Arg110 and Pro155 as observed for the MtSK–MgADP structure (Gu *et al.*, 2002). The Arg110 in MtSK represents the first residue in a conserved motif, typically RXX(X)R, of the LID domain observed for P-loop kinases (Leipe *et al.*, 2003). The second conserved basic residue of this motif interacts with the γ -phosphate of ATP. In MtSK, this residue would be Arg117, which interacts with the α - and β -phosphate groups (Fig. 3), and thus the conserved motif of the LID domain for P-loop shikimate kinases would be R(X)_{6–9}R (Fig. 4). The Arg117 residue may stabilize the transition state by neutralizing the developing negative charge on the β – γ bridge O atom (Hasemann *et al.*, 1996). The Pro155 is the last residue of the adenine-binding loop motif (residues 148–155 in MtSK), which was first recognized in AK and EcSK (Krell *et al.*, 1998) and has been described as an I/VDXXX(X)XP sequence stretch (Gu *et al.*, 2002). This motif forms a loop that wraps around the adenine moiety of ATP, connecting the β 5-strand with the C-terminal α 8-helix. The second (aspartate) residue and the last (proline) residue are not conserved in the adenine-binding loops of *aroK*-encoded SKs from *Escherichia coli* (Romanowski & Burley, 2002) and *Campylobacter jejuni* (PDB code 1via). However, they are

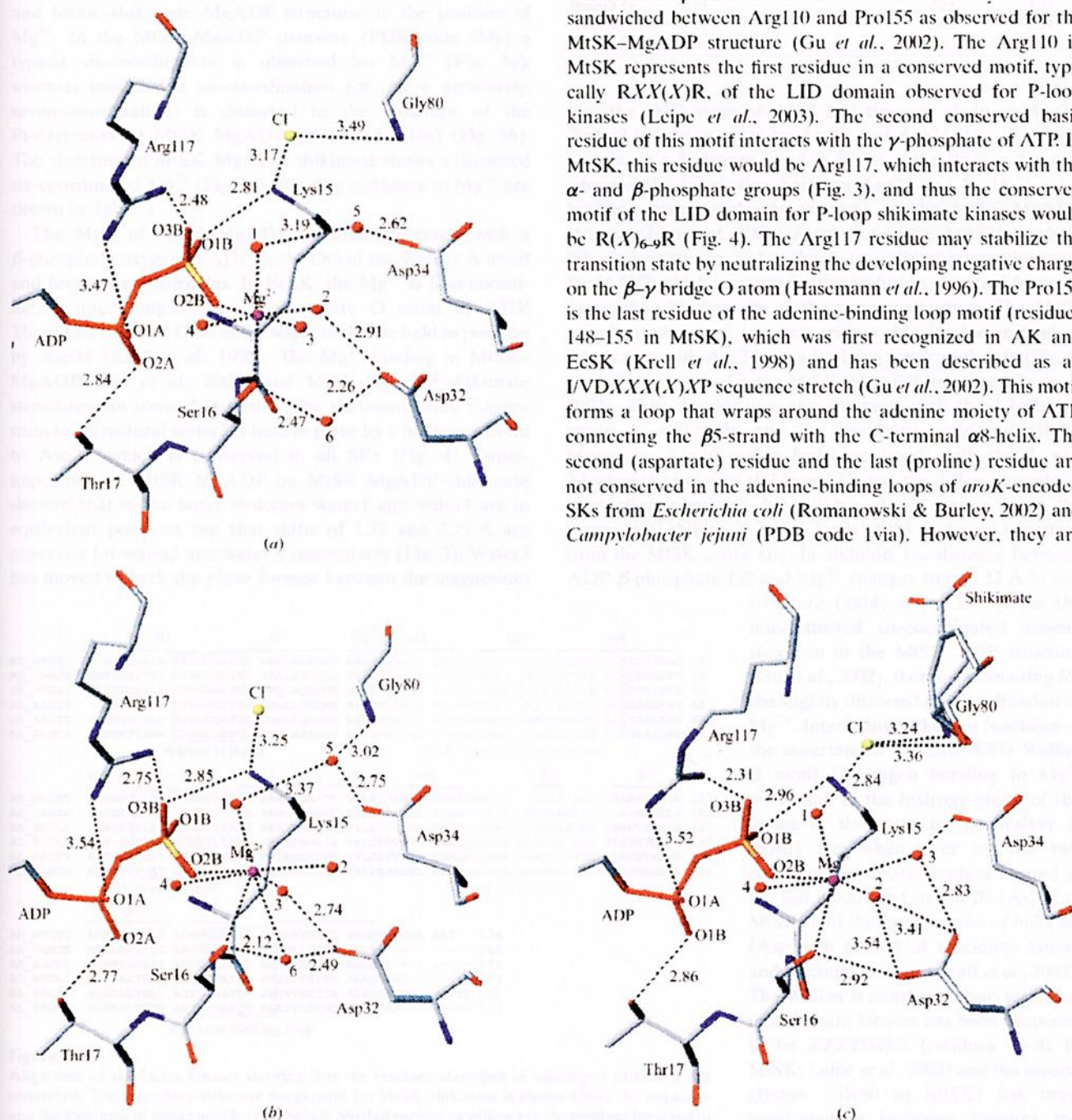


Figure 3
Mg²⁺ coordination. (a) 114y, (b) 114u (Gu *et al.*, 2002), (c) MtSK–MgADP–shikimate. For simplicity, only protein residues, atoms involved in binding of Mg²⁺, water and ADP are shown. Broken black lines represent hydrogen bonds or Mg²⁺ coordination. The distances are in Å. The following colour scheme was adopted: grey for carbon, red for oxygen, blue for nitrogen, yellow for chlorine, purple for magnesium and orange for phosphorus.

recognizable from structural alignment analysis using MtSK–MgADP as a reference structure. Accordingly, the I/VD-(X)XP primary sequence cannot be used to identify the adenine-binding loop sequences of all shikimate kinases. The carbonyl group of Arg153 residue forms hydrogen bonds to both the N6 atom of adenine and to a water molecule, which in turn hydrogen bonds to the N1 atom of adenine.

There are minor differences between the MtSK–MgADP and MtSK–shikimate–MgADP structures in the position of Mg²⁺. In the MtSK–MgADP structure (PDB code 1l4y) a typical six-coordination is observed for Mg²⁺ (Fig. 3a), whereas magnesium six-coordination (or, more accurately, seven-coordination) is distorted in the structure of the Pt-derivative of MtSK–MgADP (PDB code 1l4u) (Fig. 3b). The structure of MtSK–MgADP–shikimate shows a distorted six-coordinated Mg²⁺ (Fig. 3c). Binding distances to Mg²⁺ are shown in Table 3.

The Mg²⁺ of MtSK–MgADP–shikimate interacts with a β -phosphate oxygen of ADP, Ser16 OG of the Walker A motif and four water molecules. In EcSK, the Mg²⁺ is four-coordinated, interacting with a β -phosphate O atom of ADP, Thr16 OG1, Asp32 OD2 and a water molecule held in position by Asp34 (Krell *et al.*, 1998). The Mg²⁺ binding in MtSK–MgADP (Gu *et al.*, 2002) and MtSK–MgADP–shikimate structures are somewhat similar: the six-coordinated magnesium ion structural water2 is held in place by a hydrogen bond to Asp32, which is conserved in all SKs (Fig. 4). Superimposition of MtSK–MgADP on MtSK–MgADP–shikimate showed that in the latter structure water1 and water4 are in equivalent positions but that shifts of 1.32 and 2.75 Å are observed for water2 and water3, respectively (Fig. 3). Water3 has moved to back the plane formed between the magnesium

Table 3
Binding of Mg²⁺ in MtSKs.

n.o., not observed.			
Atom	MtSK–MgADP–shikimate (Å)	1l4y (Å)	1l4u (Å)
Ser16 OG	2.36	2.15	2.57
ADP O2B	2.23	2.15	2.39
Water1 O	2.39	2.24	2.52
Water2 O	2.23	2.04	2.77
Water3 O	2.75	2.25	2.37
Water4 O	2.97	2.05	2.94
Water5 O	n.o.	n.o.	3.06

ion, the O2B atom of ADP and the side chain of Asp34. Water3 is held in place by Asp32 and Asp34 (Fig. 3). Asp32 also forms a hydrogen bond to Ser16 of the Walker A motif, whereas the interaction between Asp32 and Ser16 is *via* a bridging water molecule (water6) in the MtSK–MgADP structure (Gu *et al.*, 2002). A rotation of the Asp32 side-chain dihedral angles χ_1 and χ_2 leads to a direct interaction with Ser16 OG, which accounts for the exclusion of water6 from the magnesium-binding site in the ternary structure. The Mg²⁺-coordinated water1 interacts with a chloride ion instead of interacting with Asp34 *via* a bridging water molecule (Fig. 3, water5) as observed in the MtSK–MgADP structure (Gu *et al.*, 2002). This chloride ion also interacts with the 3-hydroxyl group of shikimate and the backbone amide of Gly80. Moreover, Asp34 makes hydrogen bonds with the 2- and 3-hydroxyl groups of shikimate. Hence, the different mode of interaction observed for residue Asp34 arises from the presence of shikimate, which leads to the exclusion of water5 from the MtSK active site. In addition, the distance between ADP β -phosphate O2 and Mg²⁺ changes from 2.23 Å in our

structure (Table 3) to 2.15 Å for the non-distorted six-coordinated magnesium ion in the MtSK–ADP structure (Gu *et al.*, 2002), thereby accounting for the slightly distorted six-coordination of Mg²⁺. Interestingly, the two functions of the aspartate of the ZZZDXXG Walker B motif (hydrogen bonding to Mg²⁺ water and to the hydroxy group of the serine or threonine of the Walker A motif) are taken over by the two conserved aspartate residues located at the last position of strand $\beta 2$ (Asp32 in MtSK) and the first position of helix $\alpha 2$ (Asp34 in MtSK) of shikimate kinase and gluconate kinase (Kraft *et al.*, 2002). The Walker B motif consensus sequence in shikimate kinases has been proposed to be ZZZTGGG (residues 75–81 in MtSK; Leipe *et al.*, 2003) and the second glycine (Gly80 in MtSK) has been implicated in hydrogen bonding the γ -phosphate of ATP. Nucleophilic attack on the γ -phosphate of ATP will be most facilitated by metal-ion binding

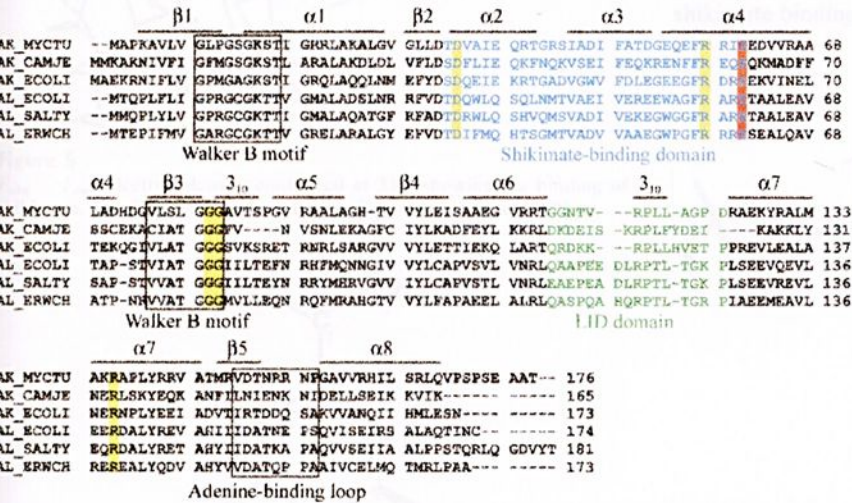


Figure 4
Alignment of shikimate kinases showing that the residues identified in binding of shikimate are conserved. The secondary-structure assignment for MtSK–shikimate is shown above the sequence and the locations of linear motifs are labelled. Shaded regions in yellow are the residues involved in binding of shikimate and the shaded region in red is the residue Glu61 (MtSK). Stretches in blue and green are the shikimate-binding domain and the LID domain, respectively. AK_MYCTU, *M. tuberculosis* SK I; AK_CAMJE, *C. jejuni* SK I; AK_ECOLI, *E. coli* SK I; AL_ECOLI, *E. coli* SK II; AL_SALTY, *S. typhimurium* SK II; AL_ERWCH, *E. chrysanthemi*.

All distances <3.6 Å are shown, n.o., not observed.

Shikimate	Atom	Distances (Å)	MtSK atom hydrogen bonded to water molecule	Distances (Å)
Hydroxyl groups				
O1	Gly80 N	3.10		
	Asp34 OD1	2.82		
	Asp34 OD2	2.57		
	Water320 O	2.46	Gly79 O	2.96
			Gly80 N	3.13
			Gly81 N	2.51
			Ala82 N	3.54
			Arg58 NH1	3.12
			Glu61 OE2	2.99
O2	Asp34 OD1	2.66		
	Asp34 OD2	2.85		
O3	n.o			
Carboxyl group				
O4	Gly81 N	3.22		
	Arg136 NH2	2.68		
O5	Arg58 NH2	2.67		
	Gly81 N	3.49		
	Arg136 NH2	2.34		

3.2. Shikimate binding

The shikimate-binding domain, which follows strand $\beta 2$, consists of helices $\alpha 2$ and $\alpha 3$ and the N-terminal region of helix $\alpha 4$ (residues 33–61). A peak of more than 3σ in the final $F_{\text{obs}} - F_{\text{calc}}$ difference Fourier map clearly indicates the position of bound shikimate in the electron density (Fig. 5).

The chemical structure of shikimic acid [3*R*-(3*α*,4*α*,5*β*)]-3,4,5-trihydroxy-1-cyclohexene-1-carboxylic acid] is shown in Fig. 6. The guanidinium groups of Arg58 and Arg136 and the NH backbone group of Gly81 interact with the carboxyl group of shikimate. The 3-hydroxyl group of shikimate forms hydrogen bonds with the carboxyl group of Asp34, the main-chain NH group of Gly80 and a water molecule (Fig. 7). This water molecule in turn mediates interactions with the side chains of SB residues Arg58 and Glu61, Walker B residues Gly79, Gly80 and Gly81, and Ala82. The 2-hydroxyl group of shikimate hydrogen bonds to the side chain of Asp34. The distances of direct and water-mediated hydrogen bonds in the shikimate binding of MtSK are shown in Table 4.

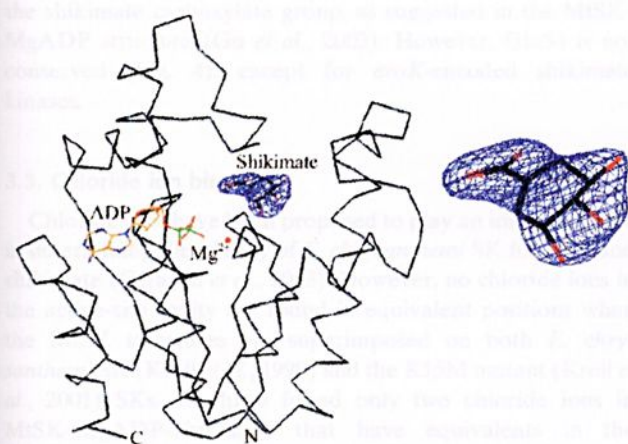


Figure 5
 $F_{\text{obs}} - F_{\text{calc}}$ electron density contoured at 3.0σ showing the binding of shikimate. The figure was prepared with *XtalView* (McRee, 1999).

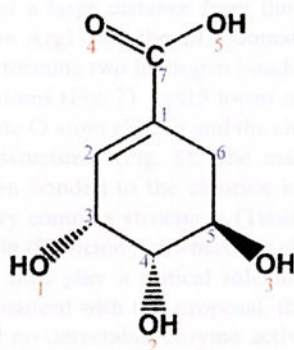


Figure 6
The molecular structure of shikimate shows the carboxyl group and three hydroxyl groups. C atoms are numbered in blue and O atoms in red.

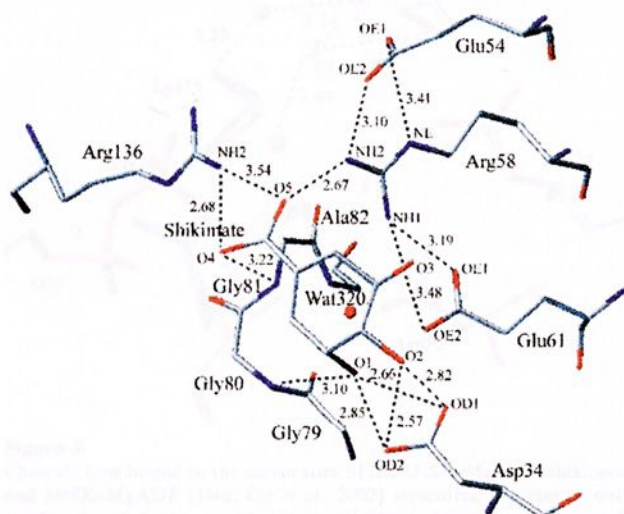


Figure 7
Polar interactions involved in the binding of shikimate to the MtSK active site. Hydrogen bonds between shikimate and protein groups and interactions between protein groups are represented as broken black lines. The interaction distances are in Å.

Glu61 is conserved in both *aroK*- and *aroL*-encoded shikimate kinase enzymes and has been implicated in shikimate binding (Gu *et al.*, 2002; Krell *et al.*, 1998; Romanowski & Burley, 2002). Krell and coworkers proposed that Glu61 is suitably positioned to bind the 5-hydroxyl group of shikimate in EcSK (Krell *et al.*, 1998). However, the amino-acid residues comprising the shikimate-binding domain were not clearly demonstrated because the electron density was not sufficient to position the shikimate molecule in the structure. In our structure, the Glu61 side chain forms a water-mediated interaction with the 3-hydroxyl group of shikimate. In addition, Glu61 forms a hydrogen bond and a salt bridge with the conserved Arg58 and assists in positioning the guanidinium group of Arg58 for substrate binding *via* interactions with the carboxylate group of shikimate. Glu61 OE2 makes a hydrogen bond with the amide N atom of Ala82. Moreover, Arg58 NH1 forms a hydrogen bond with Gly80 O. Therefore, Glu61 plays an important role in positioning a shikimate molecule, even though it is not directly involved in substrate binding. Furthermore, Glu61 is conserved in all SKs sequenced so far, which corroborates its role in anchoring Arg58 in the shikimate-binding site. The Glu54 carboxylate group also appears to anchor the guanidinium group of Arg58 for interaction with the shikimate carboxylate group, as suggested in the MtSK-MgADP structure (Gu *et al.*, 2002). However, Glu54 is not conserved (Fig. 4), except for *aroK*-encoded shikimate kinases.

3.3. Chloride ion binding

Chloride ions have been proposed to play an important role in determining the affinity of *E. chrysanthemi* SK for ATP and shikimate (Cerasoli *et al.*, 2003). However, no chloride ions in the active-site cavity are found in equivalent positions when the MtSK structures are superimposed on both *E. chrysanthemi* SK (Krell *et al.*, 1998) and the K15M mutant (Krell *et al.*, 2001) SKs. We have found only two chloride ions in MtSK-MgADP-shikimate that have equivalents in the MtSK-MgADP complex structures (114u and 114y; Gu *et al.*, 2002). One of them is bound to the active-site cavity and makes hydrogen bonds with 3-hydroxyl group of shikimate, Gly80 NH and water1 (Fig. 8), whereas the other is on the enzyme surface at a large distance from the active site. The conserved residue Arg117 in the LID domain is involved in ADP binding by forming two hydrogen bonds with the α - and β -phosphate O atoms (Fig. 7). Lys15 forms a hydrogen bond with a β -phosphate O atom (Fig. 3) and the chloride ion in the MtSK-MgADP structures (Fig. 8). The main-chain NH of Gly80 is hydrogen bonded to the chloride ion in the binary (114y) and ternary complex structures (Table 5). These residues are located in the vicinity of where the chemical reaction occurs and may thus play a critical role in transition-state stabilization. Consistent with this proposal, the K15M mutant of EcSK showed no detectable enzyme activity (Krell *et al.*, 2001), although it was in fact a K15M/P115L double mutant. Gly80 of the Walker B motif has been implicated in hydrogen bonding to the γ -phosphate of ATP (Krell *et al.*, 1998). The

Table 5

Binding of chloride ion in MtSK-MgADP and MtSK-MgADP-shikimate structures.

Chloride ion	Atom	Distance (Å)		
		114y	114u	MtSK-MgADP-shikimate
Cl180	Shikimate hydroxyl group O1	4.48†	4.66†	3.36
	Gly80 N	3.49	3.69	3.24
	Lys15 NZ	3.17	3.25	3.94
	Mg ²⁺ Water1 O	3.87	3.75	2.84

† The distance was measured with the MtSK-MgADP structures (114y and 114u) superimposed on the MtSK-MgADP-shikimate structure.

distance between the chloride ion and the hydroxyl group bound to carbon 3 of shikimate (to which a phosphoryl group would be transferred) is 3.36 Å in the MtSK-MgADP-shikimate structure (Fig. 8). Lys15 and Gly80 are in slightly different positions in the ternary complex compared with the MtSK-MgADP binary complex. The distance between O1B of the β -phosphate of ATP (2.81 Å) and Lys15 is 2.96 Å in the ternary complex (Fig. 3). However, the distance between Lys15 and the chloride ion is greater in the ternary complex (3.94 Å, Table 5). The distance between Cl⁻ and the main-chain NH of Gly80 (3.49–3.69 Å) in the binary complex is shortened to 3.24 Å in the ternary complex. There appears to be a concerted movement of Lys15, Gly80 and the Cl⁻ ion (Fig. 8). This chloride ion appears to be shifted by 1.76 and

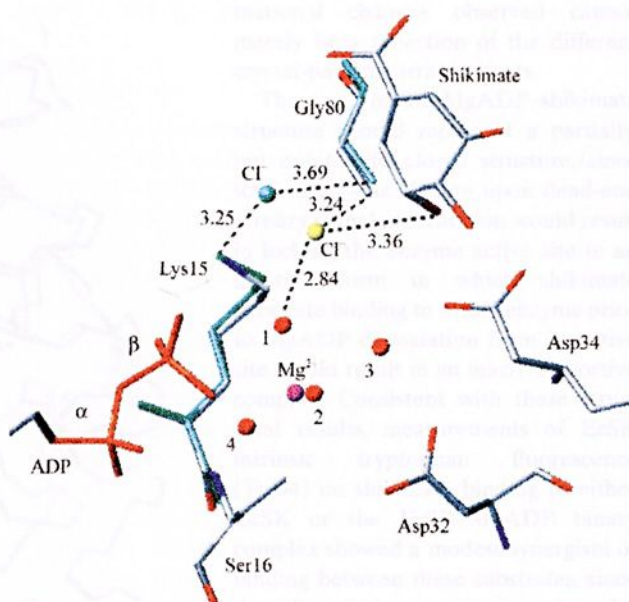


Figure 8

Chloride ions bound to the active sites of the MtSK-MgADP-shikimate and MtSK-MgADP (114u; Gu *et al.*, 2002) structures. For clarity, only residues Lys15, Ser16, Asp32, Asp34 and Gly80, shikimate, the α - and β -phosphates of ADP, Mg²⁺ and Cl⁻ ions and Mg²⁺-coordinated waters are shown. The chloride ion hydrogen bonds are represented as broken black lines and distances are in Å. The Cl⁻ ion and the Lys15 and Gly80 residues of the superimposed MtSK-MgADP structure 114u are coloured cyan.

1.52 Å in MtSK-MgADP-shikimate compared with the MtSK-MgADP 114u and 114y complexes, respectively. It is thus tempting to suggest that the chloride ion in the ternary complex occupies the phosphoryl position in the reaction coordinate to shikimate 3-phosphate formation.

3.4. Conformational changes upon substrate binding

As pointed out above, MtSK belongs to the family of nucleoside monophosphate (NMP) kinases, which are composed of three domains: the CORE, LID and NMP-binding domains. Kinases should undergo large movements during catalysis to shield their active centre from water in order to avoid ATP hydrolysis (Jencks, 1975*b*). The NMP and LID domains of NMP kinases have been shown to undergo large motions that are independent, in agreement with the observed random bi-bi kinetics (Vornrhein *et al.*, 1995). Ligand-induced changes in the secondary structure of EcSK have been detected by comparing the circular-dichroism spectra of free enzyme, EcSK-shikimate binary complex and a ternary complex of EcSK, shikimate and adenylyl imidodiphosphate, an ATP analogue (Krell *et al.*, 1998). Alignment of C α positions of the MtSK-MgADP-shikimate dead-end ternary complex and the MtSK-MgADP binary complex structures shows that the LID and SB domains undergo noticeable concerted movements towards each other (Fig. 9). The structural alignment included all residues and yielded r.m.s. deviation values of 0.56 and 0.54 Å for 114u and 114y, respectively. There is a shift of the LID domain, with an r.m.s.

deviation of 1.33 Å for residues 112–124. The SB domain shift is somewhat smaller, with a calculated r.m.s. deviation of 0.74 Å for residues 33–61. The shikimate-binding cavity is delineated mainly by residues from the LID and SB domains, the Walker B motif and Arg136 from the α 7 helix. A close inspection of the residues involved in these movements shows that the side chains of Val116, Pro118 and Leu119 from the LID domain and Ile45, Ala46, Glu54, Phe57 and Arg58 from the SB domain shifted upon shikimate binding to the MtSK-MgADP binary complex (Fig. 9). In MtSK with bound MgADP and shikimate, a cluster of hydrophobic contacts is formed between LID residues Val116, Pro118 and Leu119 and SB residues Ala46 and Phe49, which account for the stabilization of the partially closed shikimate-binding site cavity.

The conformational changes described above result in closure of the MtSK binding site, as shown by the reduction in molecular-surface area of MtSK-shikimate-MgADP compared with MtSK-MgADP (Fig. 10). The ADP, Mg²⁺, shikimate, Cl⁻ ions and water molecules were removed prior to calculation. The calculated values are approximately 7246 Å² for MtSK complexed with MgADP (114u) and 6915 Å² for MtSK complexed with MgADP and shikimate. Thus, approximately 330 Å² of molecular surface is buried on shikimate binding. Since both MtSK-MgADP and MtSK-MgADP-shikimate have been crystallized in the same space group with similar values for the unit-cell parameters (Gu *et al.*, 2001) (Table 1) and residues 114–124 from the LID domain and residues 33–61 from the SB domain form no crystal contacts with symmetry-related MtSK molecules, the conformational changes observed cannot merely be a reflection of the different crystal-packing arrangements.

The MtSK-MgADP-shikimate structure should represent a partially, but not totally, closed structure, since total active-site closure upon dead-end ternary complex formation would result in locking the enzyme active site in an inactive form in which shikimate substrate binding to MtSK enzyme prior to MgADP dissociation from its active site would result in an inactive abortive complex. Consistent with these structural results, measurements of EcSK intrinsic tryptophan fluorescence (Trp54) on shikimate binding to either EcSK or the EcSK-MgADP binary complex showed a modest synergism of binding between these substrates, since the dissociation constant value for shikimate ($K_d = 0.72$ mM) decreased to 0.3 mM in the presence of 1.5 mM ADP (Idziak *et al.*, 1997). Measurements of the quenching of protein fluorescence of the *aroL*-encoded EcSK upon nucleotide binding demonstrated the dissociation-constant values for ADP

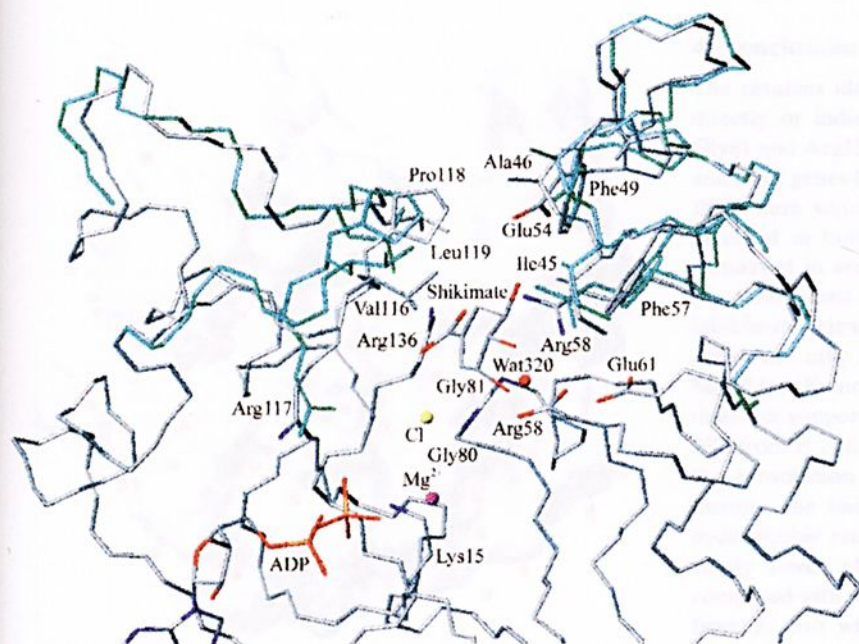


Figure 9
Shikimate binding in the MtSK-MgADP-shikimate structure. The C α traces of MtSK-MgADP-shikimate (grey) and MtSK-MgADP (cyan) (114u; Gu *et al.*, 2002) were superimposed. The side chains with large shifts owing to shikimate binding are shown for both structures: Val116, Pro118 and Leu119 from LID domain and Ile45, Ala46, Glu54, Phe57 and Arg58 from the SB domain. For clarity, only residues from the LID and SB domains of 114u (cyan) are shown.

and ATP to be 1.7 and 2.6 mM, respectively (Krell *et al.*, 2001). The K_m for ATP (620 μ M) was found to be approximately four times lower than the dissociation constant in the absence of shikimate (Krell *et al.*, 2001). These results prompted the proposal that the conformational changes in the enzyme associated with the binding of the first substrate lead to an increase in the affinity for the second substrate. However, even if this holds for EcSK, it does not appear to hold for MtSK since the apparent dissociation-constant values for ATP

(89 μ M) and shikimate (44 μ M) are similar to their K_m values, 83 μ M for ATP and 41 μ M for shikimate considering a random-order bi-bi enzyme mechanism (Gu *et al.*, 2002). Moreover, no evidence for synergism between shikimate and ATP could be observed in substrate binding to EcSK in a chloride-free buffer system (Cerasoli *et al.*, 2003). The K15M EcSK mutant has been crystallized in an open conformation that is proposed to presumably be equivalent to an apo-enzyme structure in which neither shikimate nor ADP (or ATP) would be bound (Krell *et al.*, 2001). This mutant was produced to evaluate the role of the conserved Lys15 of the Walker A motif (Krell *et al.*, 2001). However, an unwanted point mutation in the LID domain (Pro115Leu) was detected during refinement of the model and extensive contacts between LID domains of neighbouring EcSK enzymes were observed. It therefore appears unwarranted to consider the double K15M/P115L EcSK mutant a model for the apo enzyme. The incomplete LID-domain closure observed in the crystal structure presented here may suggest that the γ -phosphate of ATP plays a crucial role in the completion of the domain movement, as has also been proposed by others (Krell *et al.*, 1998). The equilibrium constant for the intramolecular hydrolysis of bound ATP to bound ADP and phosphate at enzyme active sites is considerably larger than the equilibrium constant for ATP hydrolysis in solution (Jencks, 1975a). Accordingly, the loss of two water molecules (water5 and water6) described in §3.1 is consistent with the exclusion of water molecules from the active site owing to the partial closure of MtSK upon shikimate binding in order to minimize ATP hydrolysis.

4. Conclusions

The residues identified in the binding of shikimate whether directly or indirectly (Asp34, Arg58, Glu61, Gly79, Gly80, Gly81 and Arg136) are conserved in all SKs encoded by *aroK* and *aroL* genes (Fig. 4). The structures of SKs deposited in the PDB were superimposed and the positions of the residues involved in binding between shikimate and SK are highly conserved in *aroK*-encoded proteins (shikimate kinase from *M. tuberculosis* and *C. jejuni*) and *aroL*-encoded proteins (shikimate kinase from *E. chrysanthemi*). The conserved active site may account for the similar K_m values found for MtSK (*aroK*-encoded) and EcSK (*aroL*-encoded); however, it does not support the difference observed in the K_m value of SKI from *E. coli*. Romanowski & Burley (2002) proposed that the substitution of Leu83 (EcSK) for Lys86 in *E. coli* SKI disrupts the binding to shikimate. The loss of a stabilizing hydrophobic residue at this position may explain the significantly lower affinity of SKI from *E. coli* for shikimate compared with that of SKII. However, superimposition of SKI from *E. coli* with MtSK–shikimate indicates that Lys86 is distant from shikimate, which does not support Romanowski and Burley's proposition. Moreover, even though a substitution of Leu83 (EcSK) for Thr84 occurs in MtSK, which would result in the loss of a stabilizing hydrophobic residue, the K_m values for MtSK and EcSK are similar.

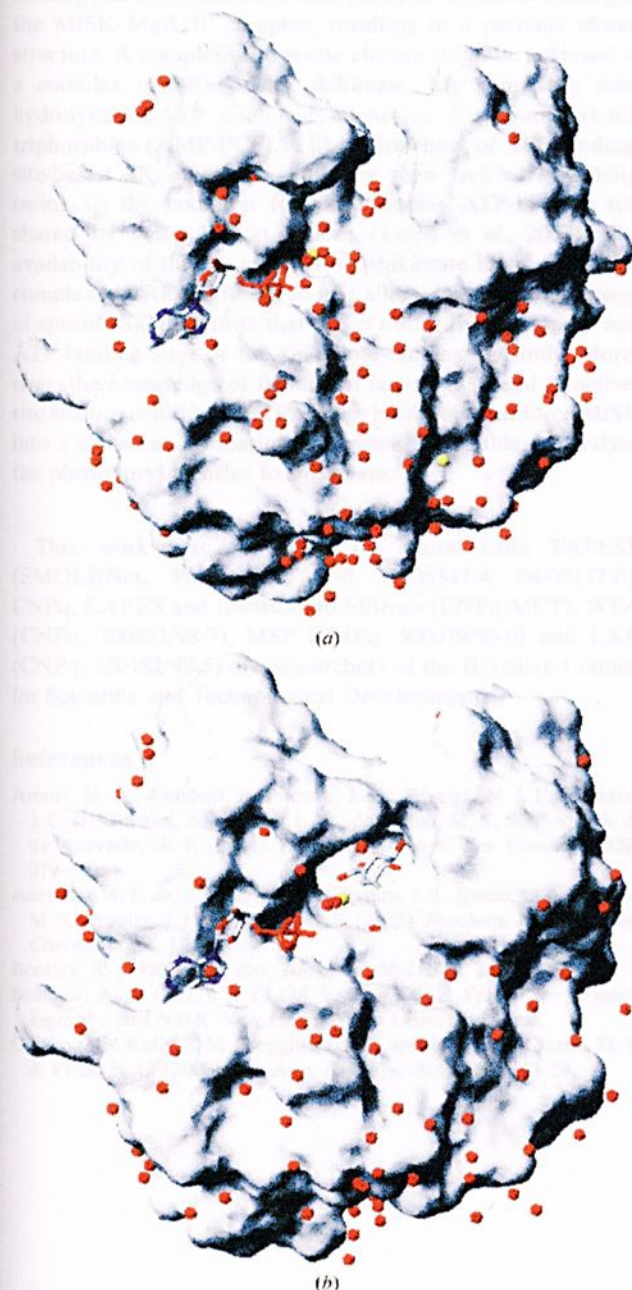


Figure 10
Molecular surface of (a) MtSK–MgADP (114u; Gu *et al.*, 2002) and (b) MtSK–MgADP–shikimate structures. Chloride ion is coloured yellow in MtSK–MgADP and MtSK–MgADP–shikimate. For water molecules only O atoms are shown.

Superimposition of SKI from *E. coli* on the MtSK–MgADP–shikimate structure shows the overlap of Leu123 onto the shikimate structure; however, we cannot affirm that the binding sites are different, since the structure of SKI from *E. coli* was crystallized without shikimate and shikimate binding in the SKI structure may promote further conformation changes that may include a displacement of the Leu123 side chain.

Here, we describe the residues involved in shikimate binding and conformational changes upon substrate binding to the MtSK–MgADP complex, resulting in a partially closed structure. A complete active-site closure could be achieved in a complex of MtSK with shikimate, Mg²⁺ and the non-hydrolysable ATP analogue adenosine 5'-(β,γ -methylene) triphosphate (AMP-PCP). A likely drawback of ATP-binding-site-based SK inhibitors would be their lack of specificity, owing to the common fold and similar ATP-binding site shared by many P-loop kinases (Leipe *et al.*, 2003). The availability of the *M. tuberculosis* shikimate kinase structure complexed with shikimate should allow the molecular design of specific SK inhibitors that target either the shikimate- and ATP-binding sites or the shikimate-binding site only. Moreover, the knowledge of functional factors that lead to active-site closure could be used to design inhibitors that force MtSK into a closed conformation that would be unable to catalyze the phosphoryl transfer to shikimate.

This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0, 02/05347-4, 04/00217-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT). WFA (CNPq, 300851/98-7), MSP (CNPq, 500079/90-0) and LAB (CNPq, 520182/99-5) are researchers of the Brazilian Council for Scientific and Technological Development.

References

- Arcuri, H. A., Canduri, F., Pereira, J. H., Silveira, N. J. F., Camara, J. C. Jr, Oliveira, J. S., Basso, L. A., de Palma, M. S., Santos, D. S. & de Azevedo, W. F. (2004). *Biochem. Biophys. Res. Commun.* **320**, 979–991.
- Azevedo, W. F. de, Canduri, F., de Oliveira, J. S., Basso, L. A., Palma, M. S., Pereira, J. H. & Santos, D. S. (2002). *Biochem. Biophys. Res. Commun.* **295**, 142–148.
- Bentley, R. (1990). *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384.
- Brünger, A. T. (1992). *X-PLOR Version 3.1: A System for Crystallography and NMR*. New Haven: Yale University Press.
- 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. (1989). *Herbicides and Plant Metabolism*, edited by A. Dodge, pp. 97–112. Cambridge University Press.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst. D50*, 760–763.
- Davies, G. M., Barrett-Bee, K. J., Jude, D. A., Lehan, M., Nichols, W. W., Pinder, P. E., Thain, J. L., Watkins, W. J. & Wilson, R. G. (1994). *Antimicrob. Agents Chemother.* **38**, 403–406.
- Gu, Y., Reshetnikova, L., Li, Y., Wu, Y., Yan, H., Singh, S. & Ji, X. (2002). *J. Mol. Biol.* **319**, 779–789.
- Gu, Y., Reshetnikova, L., Li, Y., Yan, H., Singh, S. V. & Ji, X. (2001). *Acta Cryst. D57*, 1870–1871.
- Hasemann, C. A., Istvan, E. S., Uyeda, K. & Deisenhofer, J. (1996). *Structure*, **4**, 1017–1029.
- Idziak, C., Price, N. C., Kelly, S. M., Krell, T., Boam, D. J., Lapthorn, A. J. & Coggins, J. R. (1997). *Biochem. Soc. Trans.* **25**, S627.
- Jencks, W. P. (1975a). In *Catalysis in Chemistry and Enzymology*. New York: Dover.
- Jencks, W. P. (1975b). *Adv. Enzymol.* **43**, 219–410.
- Koradi, R., Billeter, M. & Wüthrich, K. (1996). *J. Mol. Graph.* **14**, 51–55.
- Kraft, L., Sprenger, G. A. & Lindqvist, Y. (2002). *J. Mol. Biol.* **318**, 1057–1069.
- Krell, T., Coggins, J. R. & Lapthorn, 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., Lapthorn, A. J. & Coggins, J. R. (2001). *Protein Sci.* **10**, 1137–1149.
- Laskowski, R. A., MacArthur, M. W., Smith, D. K., Jones, D. T., Hutchinson, E. G., Morris, A. L., Naylor, D., Moss, D. S. & Thornton, J. M. (1994). *PROCHECK v3.0. Program to Check the Stereochemistry Quality of Protein Structures. Operating Instructions*. http://www.ccp14.ac.uk/ccp/ccp14/ftp-mirror/lnlrupp/Procheck_NT/manual.txt.
- Leipe, D. D., Koonin, E. V. & Aravind, L. (2003). *J. Mol. Biol.* **333**, 781–815.
- Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
- 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.
- Parish, T. & Stoker, N. G. (2002). *Microbiology*, **148**, 3069–3077.
- Pereira, J. H., Canduri, F., de Oliveira, J. S., da Silveira, N. J. F., Basso, L. A., Palma, M. S., de Azevedo, W. F. & 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.
- Romanowski, M. J. & Burley, S. K. (2002). *Proteins Struct. Funct. Genet.* **47**, 558–562.
- Smith, C. A. & Rayment, I. (1996). *Biophys. J.* **70**, 1590–1602.
- Vonrhein, C., Schläuderer, G. J. & Schulz, G. E. (1995). *Structure*, **3**, 483–490.
- Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). *EMBO J.* **1**, 945–951.

Shikimate Kinase: A potential target for development of novel anti-tubercular agents.

J.H. Pereira¹, I.B. Vasconcelos², J.S. Almeida³, L. A. Basso² and O.S. Santos^{1*}

¹ Department of Physics, Universidade São Paulo State, São José do Rio Preto, SP 15034-000, Brazil.

² Centro de Pesquisas em Biologia Molecular e Funções, Faculdade de Farmácia, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil.

³ Centro de Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.

* to whom correspondence may be addressed

LAB: phone, +55-51-33166234, fax, +55-51-33166239, e-mail: labasso@biut.ufpr.br

DSS: phone, +55-51-33203629, fax, +55-51-3320-3629, e-mail: ologos@uol.com.br

Anexo 2

Key Words: tuberculosis, sulfoxide, shikimate kinase, shikimate pathway, drug design, drug target, *Mycobacterium tuberculosis*, *Pharmaceuticals*

Shikimate Kinase: A potential target for development

of novel anti-tubercular agents.

J.H Pereira[†], I.B Vasconcelos[§], J.S.Oliveira[§], L.A. Basso^{‡*} and D.S. Santos^{§*}

[†] Department of Physics, University of São Paulo State, São José do Rio Preto, SP 15054-000, Brazil.

[§] Centro de Pesquisas em Biologia Molecular e Funcional, Faculdade de Farmácia, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil.

[‡] Centro de Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.

* to whom correspondence may be addressed

LAB: phone, +55-51-33166234; fax, +55-51-3316-7309; e-mail: labasso@cbiot.ufrgs.br

DSS: phone, +55-51-33203629; fax, +55-51-3320-3629; e-mail: diogenes@puers.br

Key Words: tuberculosis, malaria, shikimate kinase, shikimate pathway, drug design, drug target, *Mycobacterium tuberculosis*, *Plasmodium* sp



ABSTRACT

Tuberculosis (TB) remains the leading cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. However, no new classes of drugs for TB have been developed in the past 30 years. There is an urgent need to developing faster acting and effective new anti-tubercular agents, preferably belonging to new structural classes, to better combat TB, including MDR-TB, to shorten the duration of current treatment to improve patient compliance, and to provide effective treatment of latent tuberculosis infection. The enzymes in the shikimate pathway are potential targets for development of a new generation of anti-tubercular drugs. The shikimate pathway has been shown by disruption of *aroK* gene, which codes for the shikimate kinase enzyme, to be essential for the *Mycobacterium tuberculosis*. The shikimate kinase (SK) catalyses the phosphorylation of the 3-hydroxyl group of shikimic acid (shikimate) using ATP as a co-substrate. SK belongs to family of nucleoside monophosphate (NMP) kinases. The enzyme is an α/β protein consisting of a central sheet of five parallel β -strands flanked by α -helices. The shikimate kinases are composed of three domains: Core domain, Lid domain and Shikimate-binding domain. The Lid and Shikimate-binding domains are responsible for large conformational changes during catalysis. More recently, the precise interactions between SK and substrate have been elucidated, showing the binding of shikimate with three charged residues conserved among the SK sequences. The elucidation of interactions between *MtSK* and their substrates is crucial for the development of a new generation of drugs against tuberculosis through rational drug design.



ABBREVIATIONS LIST

TB, tuberculosis; *Mt*, *Mycobacterium tuberculosis*; SK, shikimate kinase; DAHPS, 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; CS, chorismate synthase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; PABA, p-aminobenzoate; GTP, guanosine triphosphate; ATP, adenosine triphosphate; ADP, adenosine diphosphate; Km, Michaelis-Menten constant; NMP, nucleoside monophosphate; AK, adenylate kinase; WHO, World Health Organization; HIV, human immunodeficiency virus; MDR, multidrug-resistant; DHQ, dehydroquinase; PCR, polymerase chain reaction; IPTG, isopropyl-beta-D-thiogalactoside; Mg, magnesium; r.m.s., root mean square; *Erc*, *Erwinia chrysantemi*; AMP-PCP, adenosine - β,γ -methylenadenosine 5'-triphosphate; AMP-PCP, β,γ -imidoadenosine 5'-triphosphate.



1. THE SHIKIMATE PATHWAY

The shikimate pathway was discovered as a biosynthetic route through the studies of Bernhard Davis and David Sprinson and their collaborators [1,2]. The shikimate pathway links metabolism of carbohydrates to biosynthesis of aromatic compounds through seven metabolic steps, where phosphoenolpyruvate and erythrose 4-phosphate are converted to chorismic acid (Fig. 1) [3,4]. Chorismic acid is a common precursor for the synthesis of aromatic compounds, such as aromatic amino acids, folate, ubiquinone and menaquinones. Amongst them, the only ones that can be synthesized by humans are tyrosine, which is synthesized from phenylalanine through a reaction catalyzed by phenylalanine hydroxylase enzyme, and ubiquinone from tyrosine through a cascade of eight aromatic precursors [5].

Fig. 1

The molecular organization of the shikimate pathway enzymes varies between taxonomic groups [6]. Bacteria have seven individual polypeptides, which are encoded by separate genes. Plants have a molecular arrangement similar to bacteria [7], with the exception of dehydroquinase (DHQase, third enzyme) and shikimate dehydrogenase (fourth enzyme) which have been shown to be present as separate domains on a bifunctional polypeptide [8]. In fungi and apicomplexan parasites (*Toxoplasma gondii*) the shikimate pathway has been shown to include monofunctional 3-deoxy-D-arabino- heptulosonate 7-phosphate (DAHP) synthase and chorismate synthase (CS) enzymes and a pentafunctional polypeptide termed AROM, which accounts for the remaining five shikimate pathway reactions [9].

The shikimate pathway enzymes are attractive targets for development of non-toxic anti-bacterial [10] and herbicides [11], because this pathway is essential for algae, higher plants, bacteria, fungi, whereas it is absent from mammals [12]. Thus, in the case of bacterial diseases, inhibition of any of shikimate pathway enzymes is unlikely to cause toxic side

effects on the host. In addition, the importance of shikimate pathway can be indicated by the finding that deletion of the *aroA* gene, which codes EPSPS, causes *Streptomyces pneumoniae* and *Bordetella bronchiseptica* strains to be attenuated for virulence [13,14].

The shikimate pathway has also been discovered in apicomplexan parasites providing several targets for the development of new anti-parasite drugs [15]. *In vitro* growth of apicomplexan parasites such as *Plasmodium falciparum* (malaria), *Toxoplasma gondii* (toxoplasmosis), and *Cryptosporidium parvum* (cryptosporidiosis) was inhibited by the herbicide glyphosate, a well-characterized inhibitor [16] of the shikimate pathway enzyme 5-enolpyruvyl shikimate 3-phosphate synthase (EPSPS), at concentration of 1-6 mM [15]. In *P. falciparum* and *T. gondii*, this inhibitory effect was reversed by co-addition of p-aminobenzoate (PABA) or folate suggesting that this pathway is essential for the biosynthesis of folate precursors. Moreover, several shikimate pathway enzymes have been detected in *T. gondii* and *P. falciparum* extracts [15,17] and a gene encoding a pentafunctional polypeptide AROM has been described in *T. gondii* indicating that a complete shikimate pathway is present in this parasite [18]. Whereas the shikimate pathway gives an array of compounds in bacteria, fungi and plants, only its folate biosynthesis function has been established in apicomplexan parasites. All folate pathway enzymes, which convert GTP to derivatives of tetrahydrofolate, were found in Apicomplexa when the complete genome was sequenced [19].

Two shikimic acid analogs, 6-*S*-fluorshikimate and 6-*R*-fluorshikimate, have been shown to inhibit *P. falciparum* growth and inhibition shown to be specific to the shikimate pathway [20]. Despite the completion of *P. falciparum* genome sequence [21], only a single gene encoding the chorismate synthase enzyme and a potential bifunctional SK/EPSP synthase protein has been identified in the genome annotation [22]. The coding DNA sequence of *P. falciparum* chorismate synthase has been cloned and the protein has been



shown to be located on cytosol by immunological studies [23]. Moreover, chorismate synthase, which catalyzes the last step of shikimate pathway, has been shown to be required for *Plasmodium falciparum* growth, as disruption of expression by RNA interference decreased parasite growth [19]. It has been proposed that the missing shikimate pathway enzymes are either substituted by non-homologous enzymes that catalyze the same reaction or that the enzymes are homologous but too divergent to be identified readily [22].

In mycobacteria, the chorismic acid intermediate is a precursor for the synthesis of naphthoquinones, menaquinones, and mycobactins, besides aromatic amino acids [24]. The salicylate-derived mycobactins siderophores have been shown to be essential for *M. tuberculosis* growth in macrophages [25]. Particularly, in *Mycobacterium tuberculosis*, the shikimate pathway has been shown to be essential for the bacterial viability. The disruption of *aroK* gene, which codes for the shikimate kinase enzyme (SK), was only possible when the second functional copy of *aroK* was integrated into the chromosome. Moreover, excision of the second integrated copy of *aroK* by the L5 excisionase could not be achieved in a *M. tuberculosis* strain carrying the disrupted copy of *aroK* gene, but was possible in a strain carrying a wild-type copy [26].

2.1 THE ENZYME FOLD

2. SHIKIMATE KINASE

The shikimate kinase (SK; EC 2.7.1.71), the fifth enzyme of the pathway, catalyses the specific phosphorylation of the 3-hydroxyl group of shikimic acid (shikimate) using ATP as a co-substrate (Fig. 1). In *Escherichia coli*, the SK reaction is catalyzed by two isoforms: SK I encoded by the *aroK* gene [27] and SK II encoded by the *aroL* gene [28]. The major difference between the isoenzymes is their K_m for shikimate, 20 mM for the SK I and 0.2 mM for the SK II enzyme [29]. The SK II isoform appears to play a dominant role in the shikimate pathway, its expression is controlled by the *tyrR* regulator, and it is repressed by

tyrosine and tryptophan [30,31]. The physiological role of SK I in *E.coli* is not clear. Since mutations in SK I are associated with sensitivity to the antibiotic mecillinam [32] it has been suggested that SK I may have an alternative biological role that is distinct and unrelated to its shikimate kinase activity [29]. As pointed out by Parish and Stoker [26], if *M. tuberculosis* *aroK*-encoded SK I possess a similar activity it is possible that disruption of this activity can account for the observed inability of *M. tuberculosis* to grow in the absence of a functional copy of *aroK* gene. However, the actual nature of second activity of *aroK* gene product remains to be established. Contrary to the presence of isoenzymes in *E. coli*, complete genome sequences of a number of bacteria, for example, *Haemophilus influenzae* and *Mycobacterium tuberculosis*, have revealed the presence of only one SK-coding gene. Most of these SKs appear to be encoded by *aroK* rather than *aroL* because their amino acid sequences have higher degree of identity with *E.coli* SK I. The kinetic parameters for *aroK*-encoded *M.tuberculosis* SK (*MtSK*) are more similar to those of *aroL*-encoded *E.coli* SK II than to those of *aroK*-encoded *E.coli* SK I. Thus, the *MtSK* K_m value (0.41mM) for shikimate suggests that not all *aroK*-encoded SKs have high K_m values for shikimate [33].

2.1 THE ENZYME FOLD

Shikimate kinase displays an α/β -fold and consists of five central parallel β -sheet with the strand order 23145, flanked by α -helices [34](Fig. 2). The three crystal structures of SK from *Erwinia chrysanthemi* (*ErcSK*) [34,35] showed that SK belongs to the same structural family of nucleoside monophosphate (NMP) kinases for which structures are known for adenylate kinase (AK) [36,37], guanylate kinase [38], uridylate kinase [39] and thymidine kinase [40].

The NMP kinases are composed of three domains: CORE, LID and NMP-binding (NMPB) domains [41]. A characteristic feature of the NMP kinases is that they undergo



large conformational changes during catalysis, for which AK is the most extensively studied [41]. There are two flexible regions of the structures that are responsible for movement: one is the NMP-binding site which is formed by a series helices between strands 2 and 3 of parallel β -sheet and the other is the LID domain, a region of varied size and structure following the fourth β -strand of the sheet [42,43]. In SK, the shikimate-binding (SB) domain corresponds to the NMPB domain of NMP kinases.

g.2 *Tuberculosis (TB) remains the leading cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. The interruption of centuries of decline in case rates of TB occurred in most cases in the late 19th and involved the USA and some European countries due to increased poverty in urban settings and the immigration from TB high-burden.*

2.2 FUNCTIONAL MOTIFS

Three functional motifs of nucleotide-binding enzymes are recognizable in shikimate kinase, including a Walker A-motif (A-motif), a Walker B-motif (B-motif), and an adenine-binding loop. The Walker A-motif is located between the first β -strand (β_1) and first α -helix (α_1), containing the GXXXXGKT/S conserved sequence [44], where X represents any residue. This motif forms the phosphate-binding loop (P-loop), a giant anion hole which accommodates the β -phosphate of the ADP by donating hydrogen bonds from several backbone amides [45]. The side chain of P-loop lysine have a catalytic role of stabilizing the pentavalent transition state of the γ -phosphoryl group as has been shown for adenylate kinase [46] and p21^{ras} [47].

In addition to the Walker A-motif, it is observed a second conserved sequence ZZDXXG called Walker B-motif [44], where Z represents a hydrophobic residue. More recently, however, sequence and structural comparisons for all P-loop-fold proteins classified Shikimate Kinase in the DxD group of enzymes [48], which has a conserved DxD motif in strand 2. The Walker B motif consensus in shikimate kinases is ZZZTGGG and the second glycine has been implicated in hydrogen bonding to the γ phosphate of ATP. This motif is located on the C-terminal segment of the third strand (β_3) of the central β -sheet The adenine-

binding loop motif may be described as a sequence stretch of I/VDXXX(X)XP [33]. This motif forms a loop that wraps around the adenine moiety of ATP, connecting the β 5-strand with the C-terminal α -helix.

3. TUBERCULOSIS AND *Mycobacterium tuberculosis* SHIKIMATE KINASE

Tuberculosis (TB) remains the leading cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. The interruption of centuries of decline in case rates of TB occurred, in most cases, in the late 1980s and involved the USA and some European countries due to increased poverty in urban settings and the immigration from TB high-burden countries [49]. Thus, no sustainable control of TB epidemics can be reached in any country without properly addressing the global epidemic. It is estimated that 8.2 million new TB cases occurred worldwide in the year 2000, with approximately 1.8 million deaths in the same year, and more than 95 % of those were in developing countries [50]. Approximately 2 billion individuals are believed to harbor latent TB based on tuberculin skin test surveys [51], which represents a considerable reservoir of bacilli. According to a recent report compiled by the World Health Organization (WHO), the total number of new cases of tuberculosis worldwide in 2002 had risen to approximately 9 million [52]. A key driver of the increase is the synergy with the HIV epidemic, which is having a devastating impact on some parts of the world mostly in the African Region, where 31% of new TB cases were attributable to HIV co-infection [50]. Another problem is the proliferation of multi-drug resistant (MDR) strains, defined as resistant to at least isoniazid and rifampicin, which are the most effective first-line drugs [53]. According to the 2004 Global TB Control Report of the World Health Organization, there are 300,000 new cases per year of MDR-TB worldwide, and 79 % of MDR-TB cases are now “super strains”, resistant to at least three of the four main drugs used to treat TB [52]. The factors that most influence the emergence of



drug-resistant strains include inappropriate treatment regimens, and patient noncompliance in completing the prescribed courses of therapy due to the lengthy standard “short-course” treatment or when the side effects become unbearable [54]. No new classes of drugs for TB have been developed in the past 30 years, reflecting the inherent difficulties in discovery and clinical testing of new agents and the lack of pharmaceutical industry in investing money and manpower for research in the area [55]. Hence, there is an urgent need to developing faster acting and effective new anti-tubercular agents, preferably belonging to new structural classes, to better combat TB, including MDR-TB, to shorten the duration of current treatment to improve patient compliance, and to provide effective treatment of latent tuberculosis infection [53].

In *M.tuberculosis*, the presence of the genes involved in the shikimate pathway began to be elucidated in the early 90s, with the cloning and characterization of *aroA* gene product, which codifies the EPSPS synthase [56]. Nevertheless, the complete genome sequence from *M.tuberculosis*, strain H37Rv reported by Cole et al. in 1998 [57] allowed the identification by sequence homology of all genes coding for shikimate pathway enzymes.

Four homologues to the shikimate pathway enzymes were located in a cluster containing the *aroD*-encoded type II DHQ dehydratase (Rv2537c), *aroB*-encoded DHQ synthase (Rv2538c), *aroK*-encoded type I shikimate kinase (Rv2539c), and *aroF*-encoded chorismate synthase (Rv2540c). The remaining homologues to shikimate pathway enzymes were annotated as follows: *aroG*-encoded class II phenylalanine-regulated DAHPS (Rv2178c), *aroE*-encoded shikimate dehydrogenase (Rv2552c), and *aroA*-encoded EPSP synthase (Rv3227).

The *aroK* structural gene is composed by 531 pb and codes a protein of 176 amino acids. The theoretical molecular mass of *MtSK* enzyme subunit is 18.58 KDa. The first report of cloning and overexpression in soluble and functional form of *MtSK* occurred in



2001 [58], where has been reported the PCR amplification *aroK* gene from genomic DNA of *M. tuberculosis* H37Rv strain, cloning in the plasmid pET-23a(+), and overexpression of *aroK*-encoded *MtSK* protein in *Escherichia coli* BL21(DE3) host cells, without IPTG induction.

The crystal structure of a protein complexed with its substrates is of crucial importance for the rational design of inhibitors that target the enzyme. Although two crystal structures of SK from *M.tuberculosis* [33] have revealed the dynamic role of LID Domain in catalysis and position in the structure of ligands Mg^{2+} and ADP, the precise positions and interactions between shikimate and *MtSK* was not demonstrated because the shikimate-binding site was not occupied by the substrate or the electron density was not sufficient to position the shikimate molecule. In case of SK enzymes, a likely drawback of ATP-binding-site-based SK inhibitors would be their lack of specificity, owing to the common fold and similar ATP-binding site shared by many P-loop kinases [48]. Hence, trying to obtain a description of molecular interaction between *MtSK* and shikimate, molecular-modeling and docking studies has been carried out and its results reported [59]. Despite these studies failed to predict all molecular interactions between *MtSK* and shikimate [60], Asp34 and Arg136 residues and its hydrogen bonds implicated on shikimate binding are in agreement with the crystal structure of *MtSK*-MgADP-Shikimate ternary complex reported afterwards [61]. The crystal structure of *MtSK*-MgADP-Shikimate was the first crystallographic structure of SK with bound shikimate deposited in Protein Data Bank (PDB access code: 1WE2) [61]. Crystals were obtained by the hanging-drop vapour-diffusion method, and larger final concentration of ADP and shikimate in the drop (8.0mM) were used than those used by Gu et al. (2002) to obtain the crystals (4.0mM), in order to accurately determine the position of shikimate binding in the active site of *MtSK*. The structure has been determined at 2.3 Å resolution, clearly reveling the amino-acids residues involved in shikimate binding. The



molecular replacement method was used, using as a search model the structure of *MtSK*-MgADP [33]. Almost at the same time of publishing of our article describing the structure of *MtSK*-MgADP-Shikimate ternary complex, another article describing a similar structure of the ternary complex has been published (PDB access code: 1U8A) [60]. The crystal structures of *MtSK*-MgADP-Shikimate [61] and *MtSK*-ADP-Shikimate [60] ternary complexes have unequivocally revealed in detail the interactions of amino acid residues with bound shikimate and conformational changes upon substrate binding.

3.1 ADP/Mg²⁺ INTERACTION

Essentially all kinases require a Mg²⁺-nucleotide complex as one of the enzyme substrates [62], an exception being the first partial reaction catalyzed by nucleoside diphosphate kinase, which can proceed independently of Mg²⁺ [63]. Nucleophilic attack on the γ -phosphate group of ATP will be most facilitated by meta-ion binding (Mg²⁺) to the β - and γ -phosphoryl groups, whereas departure of the leaving group will be most favoured in a structure with metal binding to the α - and β -phosphoryl groups [64]. Presumably Mg²⁺ also assists in orienting the γ -phosphoryl group of 'inline' with respect to the second substrate, creating the correct geometry to complete phosphoryl transfer [62]. The binding of this cation in many P-loop proteins such as myosin [65], elongation factor EF-Tu [66], p21-Ras [67] and the heterotrimeric G-proteins [68] involves hexa-coordination of the Mg²⁺ by two oxygen atoms (from the β and γ -phosphates of the bound nucleotide), two water molecules, and two protein ligands. As in the *MtSK*-MgADP structure (PDB code 1L4Y) [33], a typical six-coordination has been observed for Mg²⁺ in the *MtSK*-MgADP-Shikimate structure, with some minor differences between the position of Mg²⁺ in the structures. In the *MtSK*-MgADP-Shikimate structure Mg²⁺ interacts with a β -phosphate oxygen of ADP, Ser16 OG

of the Walker A motif and four water molecules [61]. Superimposition of shikimate-free *MtSK*-MgADP structure on shikimate-complexed *MtSK*-MgADP structure showed that in the ternary complex water1 and water4 are in equivalent positions but that shifts of 1.32 and 2.75 Å are observed for water2 and water3, respectively. In the *MtSK*-MgADP structure, the interaction between Asp32 and Ser16 of the Walker A motif is via a bridging water molecule (water6), whereas in the ternary complex structure this interaction occurs directly via a hydrogen bond between the two residues [61], which accounts for the exclusion of water6 from the magnesium-binding site in the ternary structure. In the ternary complex, the water1 molecule coordinated to the Mg^{2+} interacts directly with the chloride ion instead of interacting with Asp34 via a bridging water molecule (water5) as observed in the *MtSK*-MgADP structure [33]. This chloride ion also interacts with the 3-hydroxyl group of shikimate and the backbone amide of Gly80 [61]. Thus, the different mode of interaction observed for residue Asp34 arises from the presence of shikimate, which leads to the exclusion of water5 from *MtSK* active site [61]. The Mg^{2+} cation was not included in the final structure of *MtSK*-ADP-Shikimate ternary complex reported by Dhaliwal et. al.(2004), since the best diffracting crystals grew in the absence of $MgCl_2$ [60]. Accordingly, the chloride ion observed in the active sites of both *MtSK* binary complex [33] and *MtSK*-MgADP-Shikimate ternary complex [61] is absent in the Mg^{2+} -free structure of *MtSK*-ADP-Shikimate ternary complex [60]. Thus, the absence of $MgCl_2$ in crystallization mixture probably have accounted for the small differences observed in conformation, position and molecular interactions of shikimate when the structure reported by Dhaliwal et al. is compared with *MtSK*-MgADP-Shikimate-Cl structure [61]. In Mg^{2+} -free *MtSK* structure the 3-hydroxyl group of shikimate is closer to β -phosphoryl group of ADP than it is in the *MtSK*-MgADP-Shikimate structure. Furthermore, two additional water-mediated interactions between protein residues and shikimate has been shown. The NZ atom of Lys15



residue forms a 2.6 Å hydrogen-bond with oxygen atom of another water molecule, which in turn interacts with the 3-hydroxyl group of shikimate, and the main-chain nitrogen of Leu119 residue forms a 2.9 Å hydrogen bond with the oxygen atom of a water molecule, which in turn interacts with 5-hydroxyl group of substrate [60]. In the *MtSK*-MgADP-Shikimate structure, Lys15 cannot form an interaction with 3-hydroxyl group of shikimate, since the chloride ion is bound to the active site cavity between them. The distance between NZ atom of Lys15 and the chloride ion is 3.94 Å, which in turn forms a 3.36-Å hydrogen bond with 3-hydroxyl group of shikimate [61].

In the *MtSK*-MgADP structures [33], Lys15 forms a hydrogen bond with a β -phosphate O1B atom and the chloride ion. The main-chain NH of Gly80 is hydrogen bonded to the chloride ion in the binary (1L4Y) [33] and ternary complex structures [60,61]. These residues are located in vicinity of where the chemical reaction occurs and may thus play a critical role in the transition state stabilization.

The molecular interactions that describe the ADP binding mode on enzyme are very similar in all available structures of *MtSK*. The adenine moiety of ADP is sandwiched between Arg110 and Pro155 [33] and this interaction has also been observed in *ErcSK* [34], and in Adenylate kinase [69] and isoenzyme II [70]. Arg110 is located at the C-terminus of $\alpha 6$ where the LID domain starts. In *MtSK*, Arg110 and Arg117 residues represent, respectively, the first and the last residue of a conserved motif of LID domain observed for P-loop kinases (typically RXX(X)R) [48]. In P-loop shikimate kinases the conserved motif of the LID domain has been proposed to be R(X)₆₋₉R [61]. The Arg117 has been shown to interact with the α - and β -phosphate groups of ADP [33,60,61], it generally interacts with the γ -phosphate of ATP bound to enzymes, and it may stabilize the transition state by neutralizing the developing negative charge on the β - γ bridge O atom [71]. The Pro155 is the last residue of the adenine-binding loop motif (residues 148-155 in *MtSK*), which was

first recognized in Adenilate kinase and *ErcSK* [34] and has been described as an I/VDXXX(X)XP sequence stretch [33]. However, the second (aspartate) residue and the last (proline) residue are not conserved in the adenine-binding loops of *aroK*-encoded SKs from *Escherichia coli* [72] and *Campylobacter jejuni* (PDB access code: 1VIA). However, they are recognizable from structural alignment analysis using *MtSK*-MgADP as a reference structure [61]. In fact, it has been pointed out that the main-chain contacts with the adenine base and the presence of a structural motif independent of sequence may be more important for adenine binding in kinases [62]. The adenine-binding loop motif (residues 148-155 in *MtSK*) forms a loop that wraps around the adenine moiety of ATP, connecting the β 5-strand with the C-terminal α 8-helix. In addition, Lys15 forms a hydrogen bond with a β -phosphate oxygen atom [33].

For catalysis, the three protein interacting residues Lys15, Arg117, and Arg136 has been proposed to be the most important [33]. Consistent with this proposal, the K15M mutant of *ErcSK* showed no detectable enzyme activity [35], although it was in fact a K15M/P115L double mutant. Contrary to the previously proposed, analysis of the *MtSK* ternary complexes have revealed that Lys15 and Arg117 are the only positively charged residues located in the vicinity of where the reaction may occur and may therefore play critical roles in the stabilization of the transition state [33]. The Arg136 residue, instead, appears to interact with the carboxyl group of shikimate and probably, it is not involved in catalysis [60,61].

3.2 INTERACTION WITH SHIKIMATE

The shikimate-binding domain, which follows strand β 2, consists of helices α 2 and α 3 and the N-terminal region of helix α 4 (Fig. 2). In particular for *MtSK*, the precise interactions between shikimate and SK have been elucidated, showing that the guanidinium



groups of Arg58 and Arg136 and the NH backbone group of Gly81 interact with the carboxyl group of shikimate. The 3-hydroxyl group of shikimate forms hydrogen bonds with the carboxyl group of Asp34 and the main-chain NH group of Gly80 and a water molecule. The 2-hydroxyl group of shikimate hydrogen bonds to the side chain of Asp34 (Fig. 3B). The Glu61 is conserved in both *aroK* and *aroL*-encoded shikimate kinase enzymes. This residue is not directly involved in substrate binding, but it forms a hydrogen bond and a salt bridge with the conserved Arg58 and assists in positioning the guanidinium group of Arg58 for shikimate binding. In the ternary structure, Glu61 side chain also forms a water-mediated interaction with the 3-hydroxyl group of shikimate. Therefore, Glu61 plays an important role in the substrate-binding site [61]. The Glu54 carboxylate group also appears to anchor the guanidinium group of Arg58 for interaction with the shikimate carboxylate group; however, Glu54 is not conserved, except for *aroK*-encoded shikimate kinases [61]. In the *MtSK*-ADP-shikimate structure [60], the main-chain nitrogen of Leu119 forms a 2.9 Å H-bond with the water3 oxygen atom, which in turns interacts with the 1-hydroxyl group of shikimate, and the NZ atom of Lys15 forms a 2.6 Å H-bond with the oxygen atom of water2, which also interacts with the 3-hydroxyl group of the substrate.

3.3 CONFORMATIONAL CHANGES UPON SUBSTRATE BINDING

As pointed out above, NMP Kinases undergoes large conformational changes during catalysis, because their LID domain and the NMP binding-site are very flexible and can make large movements upon substrate binding. These structural changes act to position enzyme sidechains appropriately around the substrates and to sequester the substrates so as to prevent the hydrolysis of bound ATP or other phosphoryl-containing substrates prior to catalysis [41,62,64]. In addition, it has also been shown that these two domains are capable

to make independently moves towards each other [41]. Previous studies have shown that hexokinases [73] and adenylate kinases [42,74], which are classified as NMP kinases, undergo a large conformational change during catalysis.

Several structural and spectroscopic studies have demonstrated that SKs undergo conformational changes on ligand binding. Circular dichroism spectra of unliganded and liganded *ErcSK* enzyme in the presence of 2mM shikimate or 2mM γ N-ATP (an non-hydrolysable ATP analogue) have shown that SK undergoes conformational changes upon ligand binding [34]. Moreover, fluorescence studies were performed using a single tryptophan residue (W54) as a report group, which is positioned close to the shikimate binding site. The addition of shikimate to protein solution caused a quenching in *ErcSK* protein fluorescence and a blue shift of 3 nm in the emission maximum, consistent with the loop containing the tryptophan residue becoming more deeply buried within the protein following ligand-binding [75]. In addition, the averaged *B*-factor for all residues in crystal structure of *ErcSK* showed clear evidence of the flexibility of the molecule, where the temperature factors for both *ErcSK* molecules in the asymmetric unit (one with bound shikimate and other with unbound shikimate) indicate two regions of high mobility, corresponding to the shikimate binding-site and the LID domain and its flanking regions [34]. A comparison of the residue-averaged *B* factors between the crystal structures of *ErcSK*-MgADP, *MtSK*-MgADP (1L4Y), and *MtSK*-MgADP Pt-derivative (1L4U) binary complexes [33] shows that the enzyme from *M. tuberculosis* follows the same pattern of flexibility in the LID and SB domains as previously observed for *ErcSK*.

Another method used to evaluate conformational changes in proteins is the superposition of structures in order to compare different complexes of the same protein. To demonstrate conformational changes upon shikimate binding in *MtSK*, alignment of C α positions of the *MtSK*-MgADP-Shikimate dead-end ternary complex and the *MtSK*-MgADP binary complex



structures were made [61], showing that the LID and SB domains undergoes noticeable concerted movements towards each other. In this alignment, which included all residues, were verified r.m.s. deviation values of 0.56 and 0.54 Å for 1L4U and 1L4Y, respectively. Residues 112-124, that comprises the LID domain, showed a r.m.s. deviation of 1.33 Å, and the SB domain shift was somewhat smaller, with an r.m.s. deviation value of 0.74 Å for residues 33-61. Similar values were found when the same procedure was applied to the binary complex and the *MtSK*-ADP-Shikimate structure [60], with a value of 0.7 Å for the overall structure (excluding residues 114-115), 1.5 Å for the LID domain and 0.4 Å for the SB domain. The residues directly involved in these movements are Val116, Pro118 and Leu119 (LID domain), Ile45, Ala46, Glu54, Phe57 and Arg58 (SB domain), where its side chains shifted upon shikimate binding [61]. A comparison of *MtSK*-ADP-Shikimate ternary complex to *MtSK*-MgADP binary complex showed, within of SB region, a shift of 0.9 Å of Arg58 residue towards the carboxylate group of shikimate. Phe49 residue moves approximately ~1.7 Å away from Phe57 and closer to the substrate, translating towards shikimate [60]. In addition, it has been proposed by Dhaliwal et al. (2004) that the observed changes in the orientation and position of Phe49 and Phe57 residues disrupt their strong ring stacking interactions is probably a result from the Van Der Waals contacts made between shikimate with both phenylalanines [60]. However, in *MtSK*-MgADP-Shikimate the strong ring stacking interaction between the phenylalanines is not disrupted in spite of the shift of Phe49 and Phe57 residues towards shikimate of, 0.49 Å and 0.79 Å, respectively. Moreover, there is a reduction in molecular surface area value the ternary complex compared to binary complex. The calculated value for *MtSK* complexed with MgADP (1L4U) was approximately 7246 Å² and for *MtSK* complexed with MgADP and shikimate a value of 6915 Å² was found. Based on this results, the difference between the complexes molecular surface areas is 330 Å², which are buried on shikimate binding (Fig. 4) [61]. An reduction in



molecular surface area value has also been observed in the Mg^{2+} -free crystal structure of *MtSK*-ADP-Shikimate ternary complex (Fig. 4). Nevertheless, in this complex a molecular surface of only $206 \text{ \AA}^2$ is buried on shikimate binding, indicating that Mg^{2+} binding has a role in the closure of *MtSK* active site. It is important to point out that both the *MtSK*-MgADP binary complexes [33] and *MtSK*-MgADP-Shikimate [61] and *MtSK*-ADP-Shikimate [60] ternary complexes have been crystallized in the same space group with similar unit-cell parameters and residues from LID and SB domains form no crystal contacts with symmetry-related *MtSK* molecules. Thus, the conformational changes that has been described for *MtSK* cannot merely be a reflection of the different crystal-packing arrangements.

Fig. 4

The incomplete LID-domain closure observed in crystal structures of both *MtSK*-MgADP-Shikimate [61] and *MtSK*-ADP-Shikimate [60] ternary complexes suggests that γ -phosphate of ATP is necessary for the completion of the domain movement [34]. This find is not unexpected, since total active-site closure upon dead-end ternary complex formation would result in locking the enzyme active site in an inactive form in which shikimate substrate binding to *MtSK* enzyme prior to MgADP dissociation from its active site would result in an inactive abortive complex. Several previously reported results are consistent with these structural. Measurements of *ErcSK* intrinsic tryptophan fluorescence (Trp54) on shikimate binding to either *ErcSK* or the *ErcSK*-MgADP binary complex showed a modest synergism of binding between these substrates, since the dissociation constant value for shikimate ($K_d = 0.72 \text{ mM}$) decreased to 0.3 mM in the presence of 1.5 mM ADP [75]. In addition, measurements of the quenching of protein fluorescence of the *aroL*-encoded *ErcSK* upon nucleotide binding demonstrated the dissociation-constant values for ADP and ATP to be 1.7 and 2.6 mM , respectively [35]. The K_m for ATP ($620 \text{ }\mu\text{M}$) was found to be

approximately four times lower than the dissociation constant in the absence of shikimate [35]. These results prompted the proposal that the conformational changes in the *ErcSK* enzyme associated with the binding of the first substrate led to an increase in the affinity for the second substrate. However, this synergism on substrate binding does not appear to hold for *MtSK* since the apparent dissociation-constant values for ATP (89 μ M) and shikimate (440 μ M) are similar to their K_m values, 83 μ M for ATP and 410 μ M for shikimate considering either a rapid equilibrium random-order bi-bi enzyme mechanism or a steady-state-ordered bi-bi enzyme mechanism [33]. Moreover, no evidence for synergism between shikimate and ATP could be observed in substrate binding to *ErcSK* in a chloride buffer system [76].

The K15M *ErcSK* mutant has been crystallized in an open conformation that is proposed to presumably be equivalent to an apo-enzyme structure in which neither ADP (or ATP) neither shikimate would be bound [35]. This mutant was produced to evaluate the role of the conserved Lys15 of the Walker A motif. However, an unwanted mutation point mutation in the LID domain (Pro115Leu) was detected during the refinement of the model. Furthermore, the enzyme was crystallized with two independent molecules in the asymmetric unit and extensive contacts of neighboring LID domains lead to a stabilization of this part of the molecule that is not visible in the native crystal structure [34,35]. It therefore appears unwarranted to consider the double K15M/P115L *ErcSK* mutant a model for the apo enzyme. Since a better model for the apo-enzyme is not currently available, here we considered the double K15M/P115L *ErcSK* mutant as a model for the apo-SK. The superposition of *MtSK*-MgADP-Shikimate and apo *ErcSK* (K15M+P115L) show the large conformational changes in the LID and SB domains associated with the binding of ADP and shikimate on *MtSK* (Fig. 5).

3. CONCLUDING REMARKS

The disruption of gene *arok*, which codes for the shikimate kinase, has been shown to be essential for the viability of *M. tuberculosis*. The evidence that shikimate pathway is essential for *M. tuberculosis* even in the presence of exogenous supplements as p-aminobenzoate, p-hydroxibenzoate and aromatic amino acids reinforces its attractiveness as a drug target [26]. It is interesting to note that shikimic acid can be used to supply the aromatic amino acid and aromatic vitamin requirements of *E. coli* blocked in any of the first three steps of the shikimate pathway [77]. This ability has been related to the presence of the shikimic acid transport system encoded by the gene *shiA* [78]. Thus, there is the possibility that shikimic acid could be used as a supplement to allow the growth *in vitro* of mutant *M. tuberculosis* strains disrupted on *aroG*, *aroB*, *aroD* or *aroE* genes, which code for first four shikimate pathway enzymes in *M. tuberculosis*. These *aro*-disrupted *M. tuberculosis* strains could be used in experiments to determine if shikimate pathway is essential to *in vivo* growth of *M. tuberculosis*.

A comprehensive structural picture of the interactions between *M. tuberculosis* SK enzyme and Mg^{2+} , ADP and shikimate substrates have been obtained from crystallographic studies [33,60,61]. In addition, the LID and SB domains conformational changes upon ADP and shikimate substrates binding, which result in a partial closure of and solvent expulsion from *MtSK* active site, has been described. Probably, a complete active-site closure would be achieved in a complex of *MtSK* with shikimate, Mg^{2+} and a non-hydrolysable ATP analogue such as adenosine - β,γ -methylenadenosine 5'-triphosphate (AMP-PCP) or β,γ -imidoadenosine 5'-triphosphate (AMP-PCP). These ATP analogues have been used in



crystallographic studies for obtaining closed complexes with *E. coli* Gluconate Kinase [79] and 3-Phosphoglycerate Kinase enzymes [80].

There has been considerable debate within the literature as to whether enzyme-catalyzed phosphoryl transfer reactions operate primarily with an associative (S_N2 -like) or dissociative (S_N1 -like) transition state [62]. A criterion that has been used for distinguishing between these mechanisms is based on the geometry and reaction coordinate distances between the terminal phosphoryl group of ATP and the acceptor substrate in ground state enzyme-substrate or enzyme inhibitor complexes [81]. For the transition state to have some degree of associative character, these distances are expected to be less than the sum of a P-O van der Waals contacts and a P-O single bond (i.e in the order of approximately 4.9 Å) [62]. In dissociative mechanisms these distances would be expected to be longer than this to allow space for the intermediate monomeric metaphosphate to exist between the leaving group (ADP) and the entering group [81]. A crystal structure of *MtSK* in complex with MgADP, Shikimate, and AlF_3 (a structural analog of γ -phosphoryl group transferred in the reaction) could permit a snapshot of the transition-state like structure of *MtSK*. As it has been shown to other kinases [82,83,84], this structural study could distinguish between associative and dissociative transition states in the reaction catalyzed by *MtSK*.

Structures of SKs deposited in the protein data bank (PDB) were superimposed and positions of residues involved in binding (Lys15, Asp34, Arg58, Glu61, Gly79, Gly80, Gly81, Arg136) between shikimate and SK are highly conserved [61]. The precise interactions between *MtSK* and shikimate have been elucidated, showing the residues involved in the bind of substrate and the conformational changes upon substrate binding. Accordingly, the availability of the *M. tuberculosis* shikimate kinase structures complexed with shikimate provide crucial information for the design of non-promiscuous SK inhibitors that target both the shikimate- and ATP-binding pockets or uniquely, the shikimate-binding



site. Moreover, the knowledge of functional factors that lead to active site closure could be used for designing inhibitors that force *MtSK* to a closed conformation that would be unable to catalyze the phosphoryl transfer to shikimate. These inhibitors could block the biosynthesis of aromatic aminoacids and other compounds (folate, micobactins, etc), which are essential for the growth and viability of the microorganism.



ACKNOWLEDGMENTS

Financial support for this work was provided by FAPESP (Proc. 02/05347-4) and Millennium Initiative Program MCT-CNPq, Ministry of Health - Secretary of Health Policy (Brazil) to DSS and LAB. DSS (304051/1975-06) and LAB (520182/99-5) are research career awarded from the National Research Council of Brazil (CNPq).

Washington, D.C.

[4] Haslam, E. (1993) Shikimic acid, metabolism and metabolites. John Wiley & Sons, Chichester, United Kingdom.

[5] Folkers, K. (1996) *Biochem Biophys Res Commun* 224, 353-361.

[6] Coggins, J.R., Duncan, E., Acton, I.A., Boscock, M.L., Chaudhuri, S., Lambert, J.M., Lewendon, A., Millar, G., Mousdale, D.M., Smith, D.D. (1987) *Biochem Soc Trans* 15, 734-739.

[7] Butler, J.R., Alworth, W.L., Nugent, M.J. (1974) *J. Am Chem Soc* 96, 1617-1618.

[8] Mousdale, D.M., Campbell, M.S., Coggins, J.R. (1987) *Phytochemistry* 26, 2653-2670.

[9] Duncan, E., Edwards, R.M., Coggins, J.R. (1987) *Biochem J* 246, 373-386.

[10] Davies, G.M., Barrett-Bee, K.J., Jude, D.A., Lehan, M., Nichols, W.W., Pinder, P.E. et al. (1994) *Antimicrob Agents Chemother*, 38, 403-406.

[11] Coggins, J.R. (1989). The shikimate pathway as a target for herbicides. In *Herbicides and Plant Metabolism* (Dodge, A., ed.), Cambridge University Press, Cambridge, UK, pp. 97-112.

[12] Bentley, R. (1990). *Crit. Revs Biochem Mol Biol*, 25, 307-334.

[13] McDevitt, D., Payne, D.J., Holmes, D.J., and Rosenberg, M. (2002) *J. Appl. Microbiol* 92, 288-345.



4. REFERENCES

- [1] Davis, B.D. and Mingioli, E. S. (1953) *J. Bacteriol.* **66**, 129–136.
- [2] Sprinson, D.B. (1960) *Adv. Carbohydrate Chem.* **15**, 235–270.
- [3] Pittard, A.J. (1987) In F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol.1, pp. 368–394. American Society for Microbiology, Washington, D.C.
- [4] Haslam, E. (1993) Shikimic acid: metabolism and metabolites. John Wiley & Sons, Chichester, United Kingdom.
- [5] Folkers, K.(1996). *Biochem. Biophys. Res. Commun.* **224**, 358-361.
- [6] Coggins, J.R., Duncan, K., Anton, I.A., Boocock, M.R., Chaudhuri, S., Lambert, J.M., Lewendon, A., Millar, G., Mousdale, D.M., Smith, D.D., (1987) *Biochem. Soc. Trans.* **15**, 754–759.
- [7] Butler, J.R., Alworth, W.L., Nugent, M.J., (1974) *J. Am. Chem. Soc.* **96**, 1617–1618.
- [8] Mousdale, D.M., Campbell, M.S., Coggins, J.R. (1987) *Phytochemistry* **26**, 2665–2670.
- [9] Duncan, K., Edwards, R.M., Coggins, J.R. (1987) *Biochem. J.* **246**, 375–386.
- [10] Davies, G.M., Barret-Bee, K.J., Jude, D.A., Lehan, M., Nichols, W.W., Pinder, P.E. *et al.* (1994). *Antimicrob. Agents Chemother.*, **38**, 403-406.
- [11] Coggins, J.R. (1989). The shikimate pathway as a target for herbicides. In *Herbicides and Plant Metabolism* (Dodge, A., ed.), Cambridge University Press, Cambridge, UK, pp. 97-112.
- [12] Bentley, R. (1990). *Crit. Rev. Biochem. Mol. Biol.*, **25**, 307-384.
- [13] McDevitt, D., Payne, D.J., Holmes, D.J., and Rosenberg, M. (2002) *J. Appl. Microbiol.* **92**, 28S-34S.



- [14] McArthur, J.D., West, N.P., Cole, J.N., Jungnitz, H., Guzman, C.A., Chin, J., Lehrbach, P.R., Djordjevic, S.P., and Walker, M.J. (2003) *FEMS Microbiol. Lett.* **221**, 7-16.
- [15] Roberts, F., Roberts, C.W., Johnson, J.J., Kyle, D.E., Krell, T., Coggins, J.R., Coombs, G.H., Milhous, W.K., Tzipori, S., Fergunson, D.J.P., Chakrebarti, D., and McLeod, R. (1998). *Nature*, **393**, 801-805.
- [16] Kishore, G.M. & Shah, D.M. (1988) *Annu. Rev. Biochem.* **57**, 627-663.
- [17] Dieckmann, A., and Jung, A. *Mol. Biochem. Parasitol.*, **19**, 143-147.
- [18] Campbell, S.A., Richards, T.A., Mui, E.J., Samuel, B.U., Coggins, J.R., McLeod, R., Roberts, C.W. (2004) *Int. J. Parasitol.*, **34**, 5-13.
- [19] McRobert, L., McConkey, G.A. (2002) *Mol. Biochem. Parasitol.*, **119**, 273-278.
- [20] McConkey G.A. (1999) *Antimicrob. Agents Chemother.*, **43**, 175-177.
- [21] Gardner, M. J., Shallom, S.J., Carlton, J.M., et al. (2002) *Nature*, **419**, 498-511.
- [22] McConkey, G.A., Pinney, J.W., Westhead, D.R. (2004) *Trends Parasitol.*, **20**, 60-65.
- [23] Fitzpatrick T, Ricken S, Lanzer M, Amrhein N, Macheroux P, Kappes B. (2001) *Mol. Microbiol.*, **4**, 65-75.
- [24] Ratledge, C. (1982). The Biology of the Mycobacteria. *Academic Press, London*. **1**, 185-271.
- [25] Voos, J.J., Rutter, K., Schroder, B.G., Su, H., Zhu, Y., and Barry, C.B. III. (2000) *Proc. Natl. Acad. Sci. USA*. **97**, 1252-1257.
- [26] Parish, T., and Stoker, N. G. (2002). *Microbiology*, **148**, 3069-3077.
- [27] Whipp, M.J. & Pittard, A.J. (1995) *J. Bacteriol.* **177**, 1627-1629.
- [28] Millar, G., Lewendon, A., Hunter, M.G. & Coggins, J.R. (1986). *Biochem. J.* **237**, 427-437.
- [29] De Feyter, R.C. & Pittard, J. (1986) *J. Bacteriol.* **165**, 331-333.
- [30] Ely, B. and Pittard, J. (1979) *J. Bacteriol.* **138**: 933-943.



- [31] De Feyter, R.C., Davidson, B.E., and Pittard, J. (1986) *J. Bacteriol.* **165**, 233–239.
- [32] Vinella, D., Gagny, B., Joseleau-Petit, D., D'Ardi, R. & Cashel, M. (1996) *J. Bacteriol.* **178**, 3818–3828.
- [33] Gu, Y., Reshetnikova, L., Li, Y., Wu, Y., Yan, H., Singh, S., Ji, X. (2002) *J. Mol. Biol.* **319**, 779–789.
- [34] Krell, T., Coggins, J.R. & Lapthorn, A.J. (1998) *J. Mol. Biol.* **278**, 983–997.
- [35] Krell, T., Maclean, J., Boam, D.J., Cooper, A., Resmini, M., Brocklehurst, K. et al. (2001) *Protein Sci.* **10**, 1137–1149.
- [36] Dreusicke, D., Karplus, A. & Schulz, G.E. (1988) *J. Mol. Biol.* **199**, 359–371.
- [37] Schlauderer, G.J. & Schulz, G.E. (1996) *Protein Sci.* **5**, 434–441.
- [38] Stehle, T. & Schulz, G.E. (1990) *J. Mol. Biol.* **211**, 249–254.
- [39] Müller-Dieckmann, H.-J. & Schulz, G.E. (1994) *J. Mol. Biol.* **236**, 361–367.
- [40] Wild, K., Böhner, T., Aubry, A., Folkers, G. & Schulz, G.E. (1995) *FEBS Letters*, **368**, 289–292.
- [41] Vonnrhein, C., Schlauderer, G. J. & Schulz, G. E. (1995) *Structure*, **3**, 483–490.
- [42] Müller, C.W., Schlauderer, G.J., Reinstein, J. & Schulz, G.E. (1996) *Structure*, **4**, 147–156.
- [43] Gerstein, M., Schulz, G.E. & Chothia, C. (1993) *J. Mol. Biol.* **229**, 494–501.
- [44] Walker, J.E., Saraste, M., Runswick, M.J. & Gay, N.J. (1982). *EMBO J.* **1**, 945–951.
- [45] Smith, C.A. & Rayment, I. (1996) *Biophys. J.* **70**, 1590–1602.
- [46] Reinstein, J., Schlichting, I., and Wittinghofer, A. (1990) *Biochemistry* **29**, 7451–7459.
- [47] Sigal, I.S., Gibbs, J.B., D'Alonzo, J.S., Temeles, G.L., Wolanski, B.S., Socher, S.H., and Scolnick, E.M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 952–956.
- [48] Leipe, D.D., Koonin, E.V., and Aravind, L. (2003) *J. Mol. Biol.* **333**, 781–815.
- [49] Raviglione, M.C. (2003) *Tuberculosis*, **83**, 4–14.



- [50] Corbett, E.L., Watt, C.J., Walker, N., Maher, D., Williams, B.G., Raviglione, M.C., Dye, C. (2003). *Arch. Intern. Med.*, **163**, 1009-1021.
- [51] Dye C, Scheele S, Dolin P, Pathania V, Raviglione, M.C. (1999) *JAMA*, **282**, 677-686.
- [52] World Health Organization: Global Tuberculosis Control. WHO Report 2004. ISBN 92 4 156264 1.
- [53] Basso, L.A., and Santos, D.S. (2005) *Med. Chem. Ver.*- Online. (accepted for publication; Ms. MCRO/03/INV).
- [54] Duncan, K. (2003) *Tuberculosis*, **83**, 201-207.
- [55] O'Brien, R.J., Nunn, P.P. (2001) *Am. J. Resp. Crit. Care Med.*, **163**, 1055-1058.
- [56] Garbe, T., Joens, C., Charles, I., Dougan, G., Young, D. *Journal of Bacteriology* **172**, 6774-6782.
- [57] Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M.A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J.E.; Taylor, K.; Whitehead, S. and Barrell, B.G. (1998) *Nature*, **393**, 537-544.
- [58] Oliveira, J.S., Pinto, C.A., Basso, L.A., Santos, D.S. (2001) *Protein Expr. and Purif.*, **22**, 430-435.
- [59] Azevedo, W. F. de, Canduri, F., de Oliveira, J. S., Basso, L. A., Palma, M.S., Pereira, J. H. & Santos, D. S. (2002). *Biochem. Biophys. Res. Commun.*, **295**, 142-148.



- [60] Dhaliwal, B., Nichols, C.E., Ren, J., Lockyer, M., Charles, I., Hawkins, A.R., Stammers, D.K. *FEBS Letter* 2004, **574**, 49-54.
- [61] Pereira, J.H., Oliveira, J.S, Canduri, F., Dias, M.V.B., Palma, M.S., Basso, L.A., Santos, D.S., and Azevedo Jr., W.F. (2004) *Acta Cryst. Section D* **60**, 2310-2319.
- [62] Matte, A., Tari, L.W., and Delbaere, T.J. (1998) *Structure*, **6**, 413-419.
- [63] Williams, R.L., Oren, D.A., Muñoz-Dorado, J., Inouye, S., Inouye, M., and Arnold, M. (1993) *J. Mol. Biol.*, **234**, 1230-1247.
- [64] Jencks, W.P. (1975) *Adv. Enzymol.* **43**, 219-410.
- [65] Smith, C.A., Reymont, I. (1995) *Biochemistry*, **34**, 8973-8981.
- [66] Berchthold, H., Reshetnikovca, L., Reiser, C.O.A., Schrimmer, N.K., Sprinzl, M., Hilgenfeld, R. *Nature* 1993, **365**, 126-132.
- [67] Pai, E.F., Kregel, U., Petski, G.A., Goody, R.S., Kabsch, W., Wittinghofer, A. *EMBO* 1990, **9**, 2351-2359.
- [68] Coleman, D.E., Berghuis A.M., Lee, E., Linder, M.E., Gilman, A.G., Sprang, S.R. *Science* (1994), **265**, 1405-1412.
- [69] Abele, U. and Schulz, G. E. (1995) *Protein Sci.*, **4**, 1262-1271.
- [70] Schlauderer, G.J. & Schulz, G.E. (1996) *Protein Sci.* **5**, 434-441.
- [71] Hasemann, C. A., Istvan, E. S., Uyeda, K. & Deisenhofer, J. (1996). *Structure*, **4**, 1017-1029.
- [72] Romanowski, M.J., Burley, S.K. (2002) *Proteins: Structure, Function, and Genetics*. **47**, 558-562.
- [73] Bennett, W.S. and Steitz, T.A. (1980) *J. Mol. Biol.* **140**, 211-230.
- [74] Schulz, G.E., Müller, C.W. and Diederichs, K. (1990) *J. Mol. Biol.* **213**, 627-630.
- [75] Idziak, C., Price, N.C., Kelly, S.M., Krell, T., Boam, D.J., Lapthorn, A.J., and Coggins, J.R. (1997) *Biochem. Soc. Trans.*, **25**, S627.



- [76] Cerasoli, E., Kelly, S.M., Coggins, J.R., Laphorn, A.J., Clarke D.T. (2003) *Biochim. Biophys. Acta*, **1648**, 43-54.
- [77] Pittard, A.J., and Wallace, B.J. (1966) *J. Bacteriol.*, **92**, 1070-1075.
- [78] Whipp, M.J., Camakaris, H., and Pittard, A.J. (1998) *Gene*, **209**, 185-192.
- [79] Kraft, L., Sprenger, G. A., and Lindqvist, Y. (2002) *J. Mol. Biol.*, **318**, 1057-1069.
- [80] Flachner, B., Kovári, Z, Varga, A., Gugolya, Z., Vonderviszt, F., Náray-Szabó, G., and Vas, M. (2004) *Biochemistry*, **43**, 3436-3449.
- [81] Mildvan, A.S. (1979) *Adv. Enzymol. Relat. Areas Mol. Biol.*, **49**, 103-126.
- [82] Schlichting, I., and Reinstein, J. (1997) *Biochemistry*, **36**, 9290-9296.
- [83] Xu, Y.-W., Moréra, S., Janin, J., and Cherfils, J. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 3579-3583.
- [84] Scheffzek, K., and Wittinghofer, A. (1997) *Science*, **277**, 333-338.

Fig. (1). The Shikimate pathway: D-erythrose 4-phosphate and phosphoenolpyruvate (PEP) are converted to chorismate acid (chorismate). Chorismate is a essential intermediate for the synthesis of aromatic amino acid (Phe, Tyr) and Trp, folate, ubiquinones, menaquinones, and cephalosporins. The seven steps are catalyzed by enzymes: 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHP synthase), 3-deoxyquinate synthase, 3-dehydroquinate dehydratase, 4-ketoquinate dehydratase, shikimate kinase, 5-enolpyruvate shikimate 3-phosphate synthase (EPSP synthase), and chorismate synthase. The reaction catalyzed by shikimate kinase is circled by a dotted line.

Figures

Fig. (1).

PRECURSORS

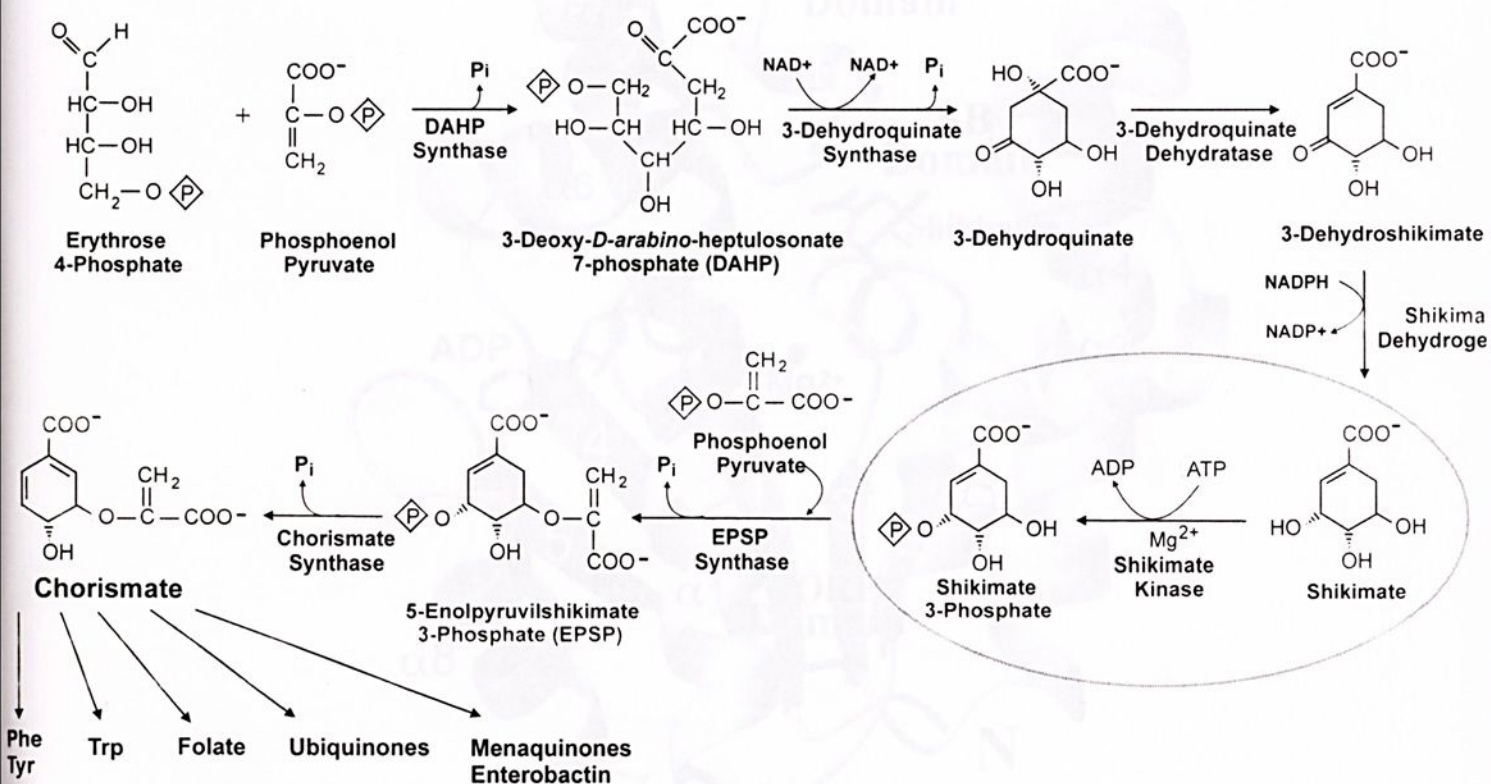


Fig. (1). The Shikimate pathway. Erythrose 4-phosphate and phosphoenolpyruvate (precursors) are converted to chorismic acid (chorismate). Chorismate is a essential intermediate for the synthesis of aromatic amino acid (Phe,Tyr and Trp), folate, ubiquinones, menaquinones, and enterobactin. The seven steps are catalyzed by enzymes: 3-deoxy-D-arabino-heptulosonate 7- phosphate synthase (DAHP synthase), 3-dehydroquinate synthase, 3-dehydroquinate dehydratase, shikimate dehydrogenase, shikimate kinase, 5-enolpyruvyl shikimate 3-phosphate synthase (EPSP synthase), and chorismate synthase. The reaction catalyzed by shikimate kinase is circled by a drawn line.

Fig. (2).

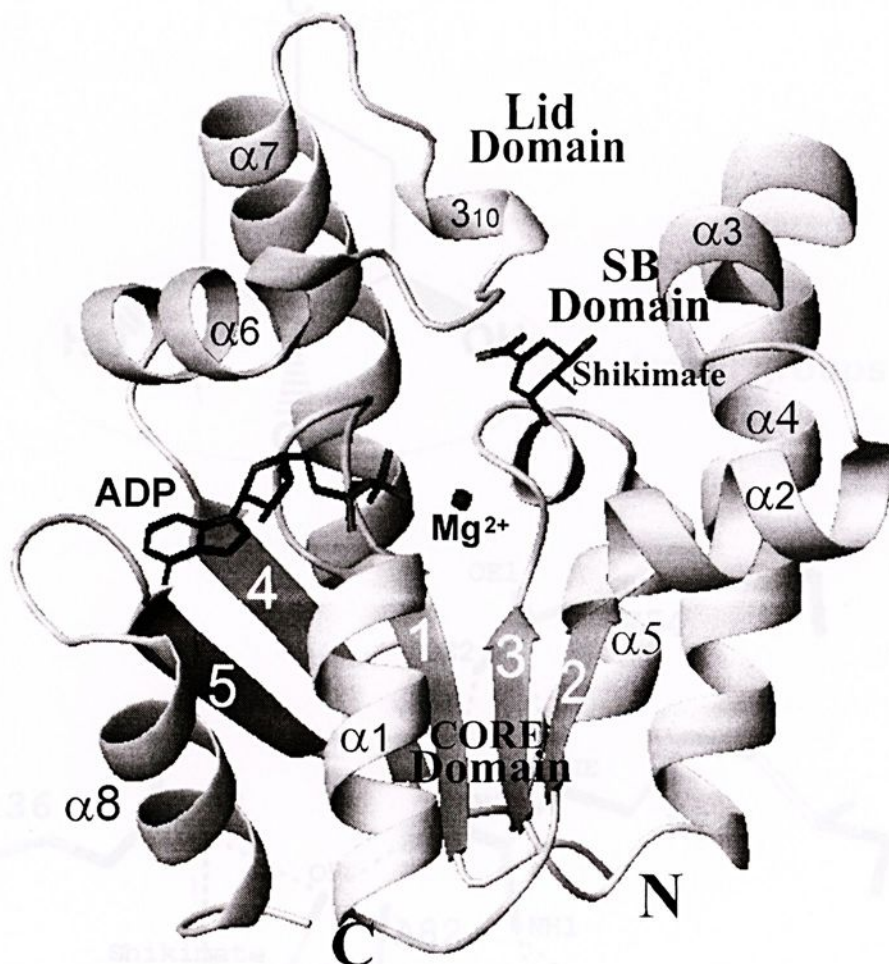


Fig. (2). Overall structure of Shikimate kinase from *M. tuberculosis* [61]. *MtSK* displays an α/β -fold and the precise ordering of the strands 23145 in the parallel β -sheet classifies *MtSK* as belonging to the family as the NMP kinases. The SKs are composed of three domains: CORE, LID and Shikimate-binding (SB) domains. The positions of the ADP, Mg^{2+} and shikimate are shown in the *MtSK* structure.

Fig. (3). A) The molecular structure of shikimate (3R-0 α ,4 α ,5 β)(3,4,5-trihydroxy-1-cyclohexene-1-carboxylic acid) shows the carboxyl group and three hydroxyl groups B)

Fig. (3).

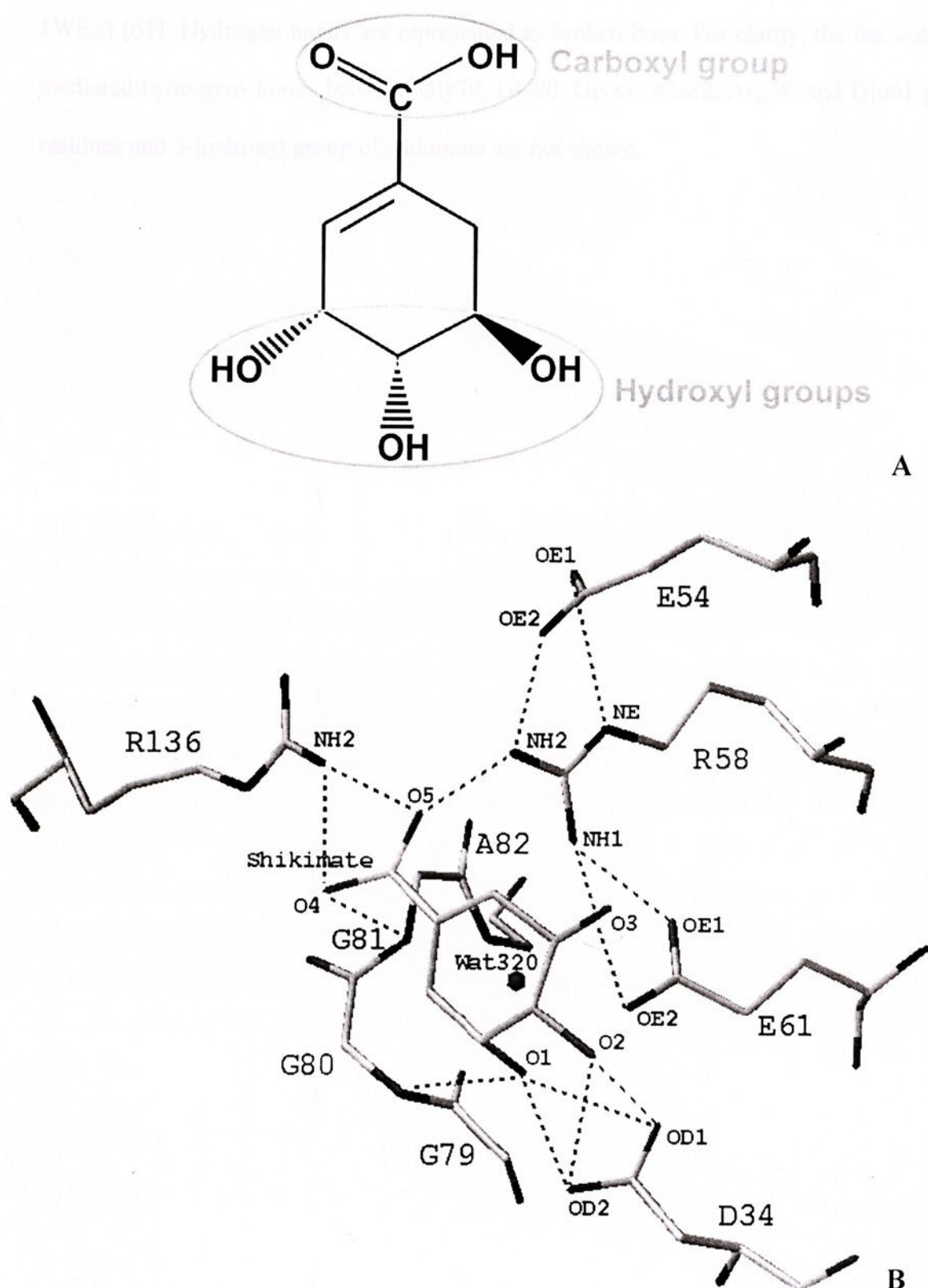


Fig. (3). A) The molecular structure of shikimate [3R-(3 α ,4 α ,5 β)]3,4,5-trihydroxy-1-cyclohexene-1-carboxylic acid] shows the carboxyl group and three hydroxyl groups B)

Interactions involved in the binding of shikimate to the *MtSK* active site (PDB access code: 1WE2) [61]. Hydrogen bonds are represented as broken lines. For clarity, the the water320-mediated hydrogens bonds between Gly79, Gly80, Gly81, Ala82, Arg58, and Glu61 protein residues and 3-hydroxyl group of shikimate are not shown.

Fig. (4). Molecular surface of (A) *MtSK*-MgADP (1L4U) [30], (B) *MtSK*-MgADP-Shikimate (1WE2) [61], and (C) *MtSK*-ADP-Shikimate (1U8A) [60] structures. The molecular surface areas have been calculated using the program Swiss-PD/Viewer v1.7 (www.expasy.org/spdov), probe radius of 1.4 Å, and a fixed radius for all atoms. Calculated values are approximately 7246 Å² for *MtSK* complexed with only MgADP (A), 6915 Å² for structure complexed with MgADP and Shikimate (B), and 7040 Å² for the Mg²⁺-free structure complexed with ADP and Shikimate. Thus, approximately 330 Å² and 200 Å² of solvent-accessible surface are buried on shikimate binding to, respectively, *MtSK*-MgADP-Shikimate and *MtSK*-ADP-Shikimate ternary complexes. The shikimate binding leads to a counter movement of LID and SB domains and, consequently, to a partial closure of shikimate-binding pocket. Probably, the Mg²⁺ cation has a role in the closure of *MtSK* active site. All atoms of shikimate and ADP, Mg²⁺ cation, and oxygen atoms of waters are colored grey.

Fig. (4).

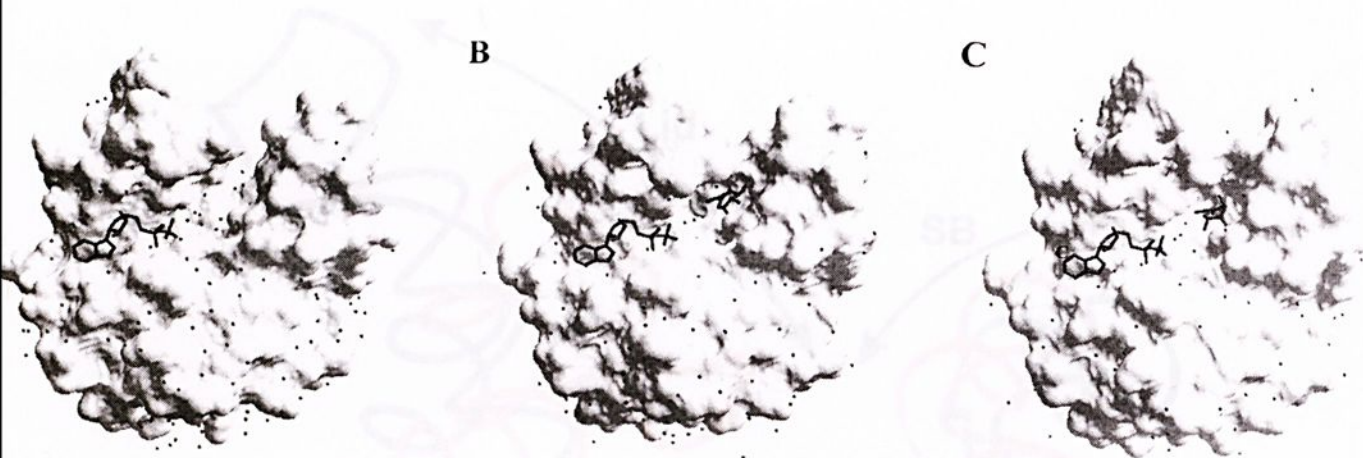


Fig. (4). Molecular surface of (A) *MtSK*-MgADP (1L4U) [33], (B) *MtSK*-MgADP-Shikimate (1WE2) [61], and (C) *MtSK*-ADP-Shikimate (1U8A) [60] structures. The molecular surface areas have been calculated using the program Swiss-PDBViewer v3.7 (www.expasy.org/spdbv), probe radius of 1.4 Å, and a fixed radius for all atoms. Calculated values are approximately 7246 Å² for *MtSK* complexed with only MgADP (A), 6915 Å² for structure complexed with MgADP and Shikimate (B), and 7040 Å² for the Mg²⁺-free structure complexed with ADP and Shikimate. Thus, approximately 330 Å² and 206 Å² of solvent-accessible surface are buried on shikimate binding to, respectively, *MtSK*-MgADP-Shikimate and *MtSK*-ADP-Shikimate ternary complexes. The shikimate binding leads to a counter movement of LID and SB domains and, consequently, to a partial closure of shikimate-binding pocket. Probably, the Mg²⁺ cation has a role in the closure of *MtSK* active site. All atoms of shikimate and ADP, Mg²⁺ cation, and oxygen atoms of waters are colored grey.

Fig. (5). Superposition of SK-MgADP-Shikimate ternary complex from *Aerobacterium* *aerobactans* [61] (red box) and apo-SK (1L4U) [33] from *Erwinia chrysanthemi* [34] (blue box).

Fig. (5).

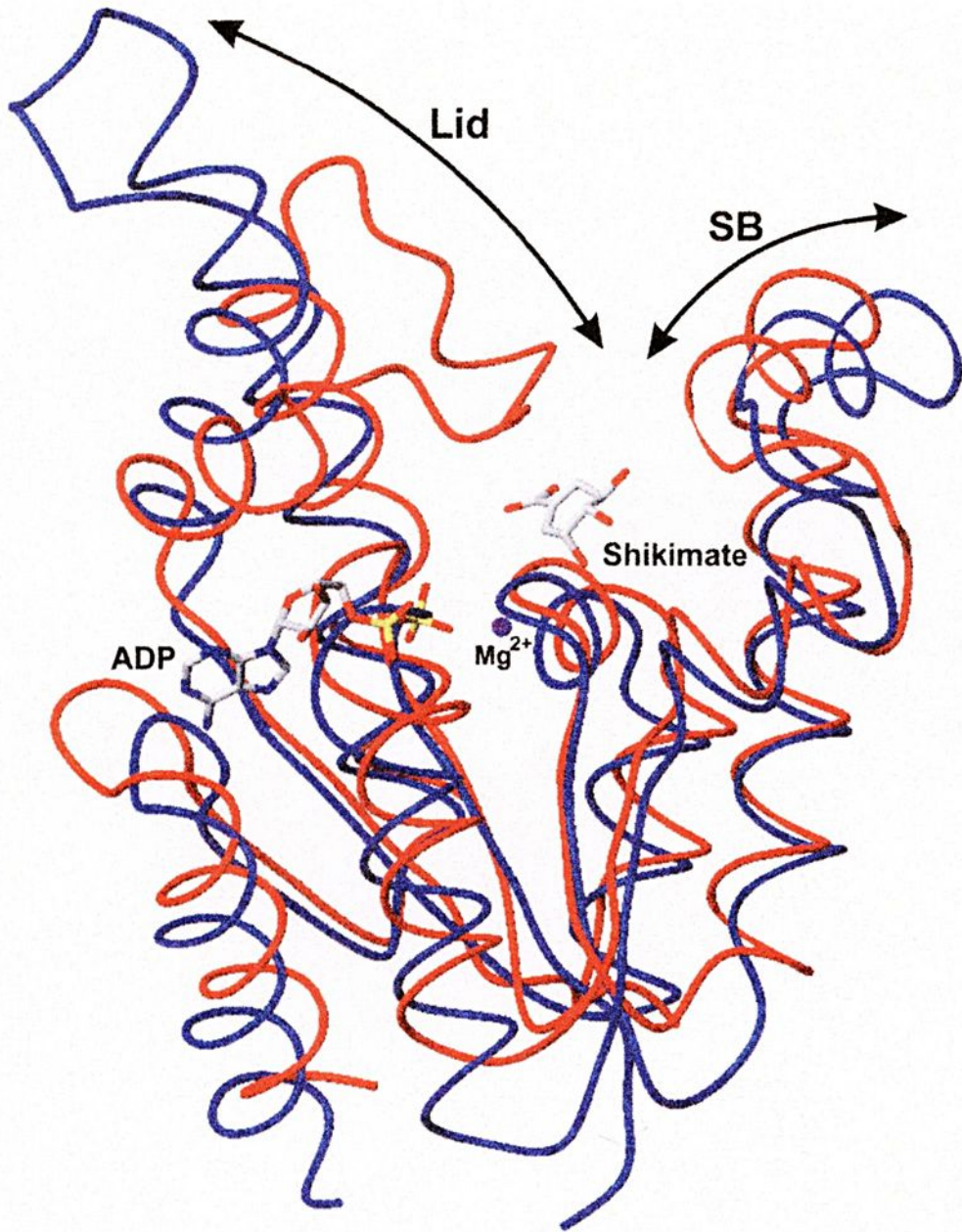


Fig. (5). Superposition of SK-MgADP-Shikimate ternary complex from *Mycobacterium tuberculosis* [61] (red line) and apo-SK (K15M/P115L) from *Erwinia chrysanthemi* [34] (blue line).

Molecular model of shikimate kinase from *Mycobacterium tuberculosis*

Walter Filgueira de Azevedo Jr.^{a,b,*}, Fernanda Candau,^{a,b} João Simões de Oliveira,^a
Luiz Augusto Basso,^a Maria Sérgio Palma,^{a,c} José Henrique Pereira,^a
and Douglas Santiago Santos^a

^a Departamento de Física, UNIFESP, Lab. Física, Av. Itália 33, 04032-000, Brazil

^b Centro de Estudos Teóricos, Instituto Brasileiro de Física, SP 05389-000, Brazil

^c Núcleo Brasileiro de Pesquisas em Tuberculose, Departamento de Biologia Molecular, Universidade Federal do Rio de Janeiro, RJ 21941-900, Brazil

* Corresponding author. Fax: +55 11 5082 1100; E-mail: wfilgueira@fma.unifesp.br

Received 25 May 2002

Abstract

Tuberculosis (TB) resurged in the late 1980s and now kills approximately 3 million people a year. The resurgence of tuberculosis as a public health threat has created a need to develop new antimicrobial agents. The shikimate pathway is an attractive target for herbicide and antimicrobial agents development because it is essential in plant, higher plants, bacteria, and fungi, but absent from mammals. Developments to interfere in this shikimate pathway have been observed in the genetic screens of *Mycobacterium tuberculosis*. Among them, the shikimate kinase (ShikK) gene (shk) was proposed to be present by sequence homology. Accordingly, to pave the way for structural and functional efforts for anti-mycobacterial agents development, here we describe the molecular modeling of *M. tuberculosis* shikimate kinase (ShikK) that shows potential structural framework on which the design of specific inhibitors may be based. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Shikimate kinase; Streptomycin; Neomycin; Drug design; Antimicrobial resistance

The 1998 annual report on global tuberculosis (7) by the control of the World Health Organization found that there were an estimated 8–10 million new cases in 1999, up from 8.0 million in 1997 (4). It is expected that there will be 10.7 million new cases in 2000 if the present trend continues. Approximately 3 million persons die from the disease each year (2). Twenty percent of tuberculosis is more severe in developing countries, where few resources are available to ensure proper treatment and where human immunodeficiency virus (HIV) infection may be present. The combination of deaths due to tuberculosis in socioeconomically developing nations and mortality rates in the range from 25 to 54 years, the most symptomatically fruitful years of life, causes substantial loss in productivity and contributes to the higher incidence of deaths of 1 case in 3 (1). The resurgence of

TB as a public health threat, the high susceptibility of HIV-infected persons to the disease, and the proliferation of multidrug (MDR) strains have created much scientific interest in developing new antimicrobial agents to both treat Mycobacterium tuberculosis and to develop drugs that can be used in combination with first-line drugs to increase patient compliance (3).

The shikimate pathway is an attractive target for the development of herbicide and antimicrobial agents because it is essential in plant, higher plants, bacteria, and fungi, but absent from mammals (2). In eukaryotes, the shikimate pathway leads to the biosynthesis of precursors for the synthesis of aromatic amino acids, poliphenolics, monoterpenes, and isoprenoids. In fungi, 10 enzymes in the shikimate pathway have been identified in the complete genomic sequence of *Mycobacterium tuberculosis* H37Rv strain (7). Among them, the shikimate kinase (ShikK, EC 2.7.1.31) catalyzing reaction, $R_2(5\%)$ was proposed to be present

Anexo 3

* Corresponding author. Fax: +55 11 5082 1100.

E-mail address: wfilgueira@fma.unifesp.br (W. Filgueira de Azevedo Jr.).



Molecular model of shikimate kinase from *Mycobacterium tuberculosis*

Walter Filgueira de Azevedo Jr.,^{a,b,*} Fernanda Canduri,^{a,b} Jaim Simões de Oliveira,^c
Luiz Augusto Basso,^c Mário Sérgio Palma,^{b,d} José Henrique Pereira,^a
and Diógenes Santiago Santos^c

^a Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil

^b Center for Applied Toxicology, Instituto Butantan, São Paulo, SP 05503-900, Brazil

^c Rede Brasileira de Pesquisa de Pesquisas em Tuberculose, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil

^d Laboratory of Structural Biology and Zoochemistry/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil

Received 26 May 2002

Abstract

Tuberculosis (TB) resurged in the late 1980s and now kills approximately 3 million people a year. The reemergence of tuberculosis as a public health threat has created a need to develop new anti-mycobacterial agents. The shikimate pathway is an attractive target for herbicides and anti-microbial agents development because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals. Homologs to enzymes in the shikimate pathway have been identified in the genome sequence of *Mycobacterium tuberculosis*. Among them, the shikimate kinase I encoding gene (*aroK*) was proposed to be present by sequence homology. Accordingly, to pave the way for structural and functional efforts towards anti-mycobacterial agents development, here we describe the molecular modeling of *M. tuberculosis* shikimate kinase that should provide a structural framework on which the design of specific inhibitors may be based. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Shikimate kinase; Bioinformatics; Structure; Drug design; *Mycobacterium tuberculosis*

The fifth annual report on global tuberculosis (TB) control of the World Health Organization found that there were an estimated 8.4 million new cases in 1999, up from 8.0 million in 1997 [1]. It is expected that there will be 10.2 million new cases in 2005 if the present trend continues. Approximately 3 million persons die from the disease each year [2]. Ninety percent of tuberculosis cases occur in developing countries, where few resources are available to ensure proper treatment and where human immunodeficiency virus (HIV) infection may be common. The concentration of deaths due to tuberculosis in demographically developing nations and mortality rate in the range from 25 to 54 years, the most economically fruitful years of life, causes substantial losses in productivity and contributes to the impoverishment of third-world countries [3]. The reemergence of

TB as a public health threat, the high susceptibility of HIV-infected persons to the disease, and the proliferation of multi-drug (MDR) strains have created much scientific interest in developing new anti-mycobacterial agents to both treat *Mycobacterium tuberculosis* strains resistant to existing drugs and shorten the duration of short-course treatment to improve patient compliance [4].

The shikimate pathway is an attractive target for the development of herbicides and anti-microbial agents because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals [5]. In mycobacteria, the shikimate pathway leads to the biosynthesis of precursors for the synthesis of aromatic amino acids, naphthoquinones, menaquinones, and mycobactin [6]. Homologs to enzymes in the shikimate pathway have been identified in the complete genome sequence of *Mycobacterium tuberculosis* H37Rv strain [7]. Among them, the shikimate kinase I (mtSK, EC 2.7.1.71) encoding gene (*aroK*, Rv2539c) was proposed to be present

* Corresponding author. Fax: +55-17-221-2247.

E-mail address: walterfa@df.ibilce.unesp.br (W. Filgueira de Azevedo Jr.).



by sequence homology. Shikimate kinase catalyzes a phosphate transfer from ATP to the carbon-3 hydroxyl group of shikimate resulting in the formation of shikimate-3-phosphate (S3P) and ADP.

The present paper describes the molecular model of *M. tuberculosis* shikimate kinase (mtSK) and analysis of SK and shikimate complex obtained by docking simulations. The homology modeling was performed using three crystallographic structures of SK from *Erwinia chrysanthemi*, solved to resolution better than 2.6 Å, as templates [8]. The mtSK has been cloned, sequenced, overexpressed in soluble and functional forms [9], thus allowing enzymological studies to be performed. The results presented here should provide a three-dimensional model of mtSK to both guide enzymological studies and aid the design of specific inhibitors.

Methods

Molecular modeling. For modeling of the mtSK we used restrained-based modeling implemented in the program MODELLER [10]. This program is an automated approach to comparative modeling by satisfaction of spatial restraints [11–13]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). This alignment is usually the input to the program. The output is a three-dimensional model for the target sequence containing all main-chain and side-chain non-hydrogen atoms.

The degree of primary sequence identity between mtSK and *Erwinia chrysanthemi* shikimate SK (ecSK) indicates that the crystallographic structures of ecSK are good models to be used as templates for mtSK. The atomic coordinates of three crystallographic ecSK structures (PDB access code: 1SHK, 2SHK, and 1E6C) [8,14], with two independent structures in each asymmetric unit, solved to resolution better than 2.6 Å were used to build up an ensemble of SK structures to be used as starting models for modeling of the mtSK. The atomic coordinates of all waters and ligands were removed from the ecSK structures. Next, the spatial restraints and CHARMM energy terms enforcing proper stereochemistry [15] were combined into an objective function. Finally, the model is obtained by optimizing the objective function in Cartesian space. The optimization is carried out by the use of the variable target function method [16] employing methods of conjugate gradients and molecular dynamics with simulated annealing.

Several slightly different models can be calculated by varying the initial structure. A total of 500 models were generated for mtSK, and the final model was selected based on stereochemical quality.

Docking simulations. To obtain information about the docking of shikimate to ecSK and mtSK, several rigid docking simulations were performed using the geometric recognition algorithm, which was developed to identify molecular surface complementarity. The geometric recognition algorithm was implemented in the program GRAMM [17].

The atomic coordinates of shikimate, used in the docking simulations, were obtained from structure of 5-enolpyruvylshikimate-3-phosphate synthase liganded with shikimate-3-phosphate and glyphosate (PDB access code: 1G6S) [18]. To generate the ternary complex mtSK–shikimate–ADP/Mg²⁺ we superposed the atomic coordinates of the ADP/Mg²⁺ to the binary complex of mtSK–shikimate. The optimization of the complexes was carried out by the use of the variable target function method [10] employing methods of conjugate gradients and molecular dynamics with simulated annealing. All docking simulations and optimization process were performed on SGI Octane, R12000.

Analysis of the model. The overall stereochemical quality of the final model for mtSK complex was assessed by the program PROCHECK [19]. The cutoff for hydrogen bonds and salt bridges was 3.6 Å.

Results and discussion

Primary sequence comparison

The sequence alignment of ecSK (template) and mtSK (target) is shown in Fig. 1. The secondary structural elements are indicated in the figure. The sequence mtSK shows 34% of identity with the sequence of ecSK.

Quality of the model

Figs. 2A and B show the Ramachandran diagram ϕ – ψ plots for the mtSK structure and for three crystallographic SK structures solved to resolution better than 2.6 Å. The Ramachandran plot for the three ecSK structures was generated to compare the overall stereochemical quality of mtSK model against SK structures solved by biocrystallography. Analysis of the Ramachandran plot of the mtSK model shows that 91.1% of the residues lie in the most favorable regions and the

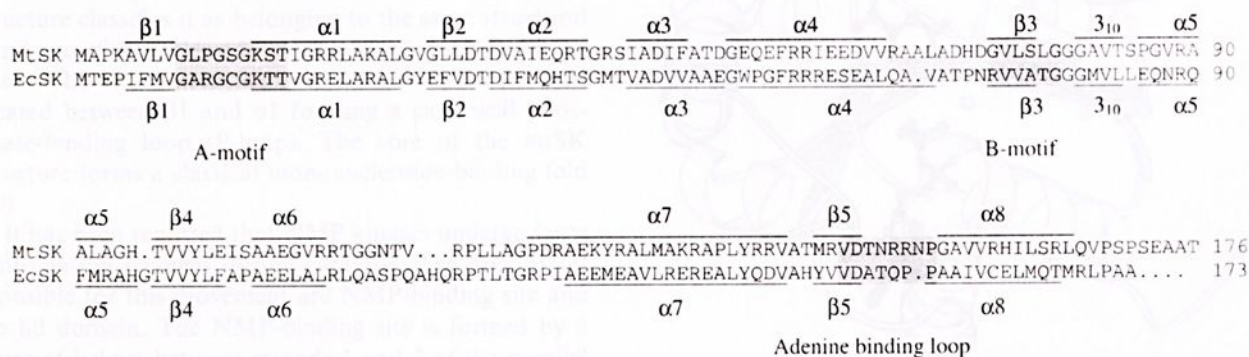


Fig. 1. The sequence alignment of ecSK and mtSK indicating the secondary structural elements. The sequence mtSK shows 34% of identity with the sequence of ecSK. The alignment was performed with the program CLUSTAL V [31].

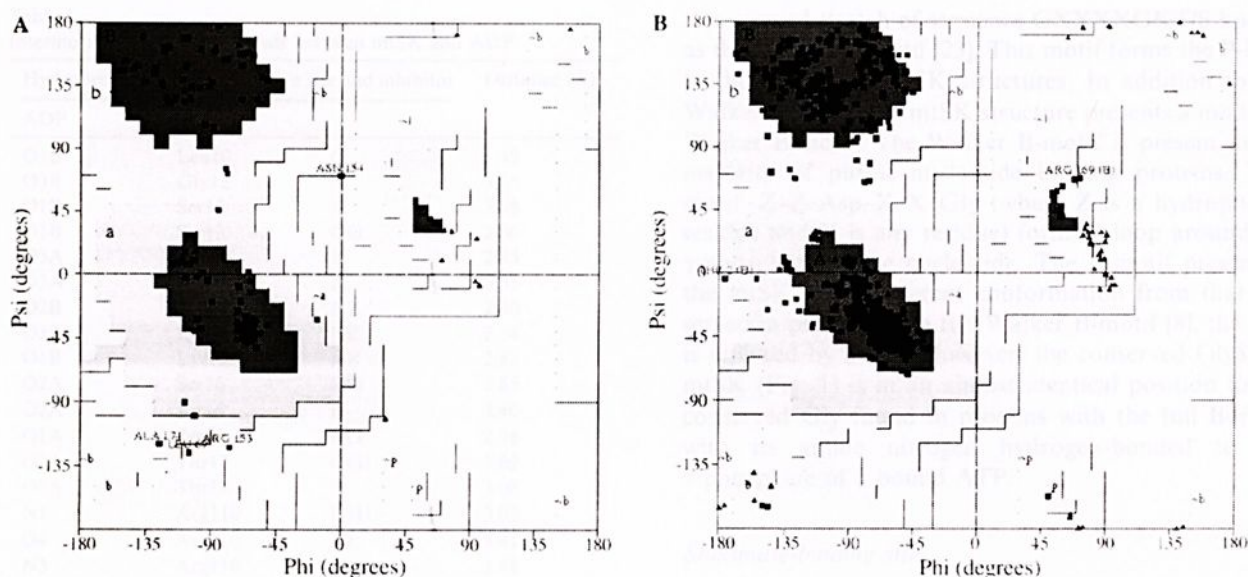


Fig. 2. (A) Ramachandran diagram ϕ - ψ plots for the mtSK structure and (B) for three crystallographic SK structures solved to resolution better than 2.6 Å.

remaining 8.9% in the additional allowed regions. The same analysis for three crystallographic ccSK structures (six chains) present 93.7% of residues in the most favorable, 6.1% additional allowed regions, and 0.7% generously allowed regions. The overall rating for the mtSK model is slightly poorer than the one obtained for the three structures of SK. However, it has over 90% of the residues in the most favorable regions.

Overall description

MtSK is an α/β protein consisting of a mixed β sheet surrounded by α helices. A central five stranded parallel β -sheet (β 1– β 5) presents the strand order 23145. The β -strands are flanked on either side by α helices (α 1 and α 8 on one side, α 4, α 5, and α 7 on the other). Fig. 3 shows a schematic diagram of the mtSK structure, with shikimate and ADP/Mg²⁺ bound to the structure.

The ordering of the strands 23145 observed the mtSK structure classifies it as belonging to the same structural family as the nucleoside monophosphate (NMP) kinases. The mtSK structure exhibits the Walker A-motif located between β 1 and α 1 forming a canonical phosphate-binding loop (P-loop). The core of the mtSK structure forms a classical mononucleoside-binding fold [20].

It has been reported that NMP kinases undergo large conformation changes during catalysis. The regions responsible for this movement are NMP-binding site and the lid domain. The NMP-binding site is formed by a series of helices between strands 1 and 2 of the parallel β -sheet. The lid domain is a region of variable size and structure following the forth β -strand of the sheet

[21,22]. The residues from 112 to 123 form the lid domain in the mtSK which has been reported to be highly dynamic and possibly flexible in solution in the ccSK structure [8].

ADP/Mg²⁺-binding site

The molecular model for ternary complex mtSK–shikimate–ADP/Mg²⁺ indicates that ADP/Mg²⁺ is

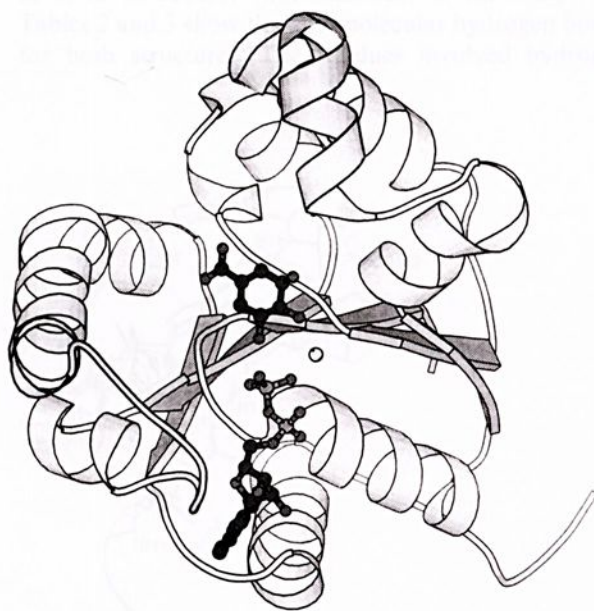


Fig. 3. Ribbon diagram of the mtSK structure with shikimate and ADP/Mg²⁺ bound to the structure generated by MOLSCRIPT [32].

Table 1
Intermolecular hydrogen bonds between mtSK and ADP

Hydrogen bonds between active site and inhibitor			Distance (Å)
ADP	mtSK		
O1B	Leu10	O	3.49
O3B	Gly12	N	3.15
O1B	Ser13	N	3.08
O1B	Ser13	OG	2.66
O3A	Gly14	N	2.75
O3A	Lys15	N	3.35
O2B	Lys15	NZ	2.80
O3B	Lys15	NZ	2.54
O1B	Lys15	NZ	2.82
O2A	Ser16	OG	2.65
O2A	Ser16	N	3.40
O1A	Ser16	OG	2.98
O2A	Thr17	OG1	3.02
O2A	Thr17	N	3.08
N1	Arg110	NH1	3.05
O4	Arg110	NE	3.41
N3	Arg110	NH1	2.88
N7	Arg110	NH1	3.29
N6	Arg152	O	2.52
N1	Arg152	O	3.23
N1	Arg153	O	3.37
N6	Arg153	O	2.66
N3	Arg153	NE	2.80
N3	Arg153	NH2	2.68
N1	Arg153	N	3.20

tightly bound to the mtSK structure. The intermolecular hydrogen bonds are described in Table 1. Most of the intermolecular hydrogen bonds observed in the ecSK structure is conserved in the ternary complex mtSK–shikimate–ADP/Mg²⁺. Fig. 4 shows the superposition of the ATP-binding site of mtSK and ecSK. As previously described a phosphate-binding loop (P-loop) accommodates the β -phosphate of ADP by donating hydrogen bonds from several backbone amides [8,14]. SKs contain

a conserved stretch of sequence GXXXXGKT/S known as the Walker A-motif [23]. This motif forms the P-loop in the ecSK and mtSK structures. In addition to the Walker A-motif, the mtSK structure presents a modified Walker B-motif. The Walker B-motif is present in the majority of purine–nucleotide binding proteins. This motif, Z–Z–Asp–X–X–Gly (where Z is a hydrophobic residue and X is any residue) forms a loop around the γ -phosphate of the nucleotide. The B-motif present in the mtSK has a different conformation from that observed in proteins with full Walker B-motif [8], the Asp is replaced by Ser77. However, the conserved Gly80 of mtSK (Fig. 1) is in an almost identical position to the conserved Gly found in proteins with the full B-motif with its amide nitrogen hydrogen-bonded to the γ -phosphate of a bound ATP.

Shikimate-binding site

The crystallographic structure of ecSK indicated the presence of a strong electron density peak attributed to shikimate. However, the electron density was not clear enough to include shikimate in the molecular structure [8]. The docking simulations of shikimate to ecSK and mtSK identified that shikimate binds in a position analogous to nucleotide monophosphate in NMP kinases [8]. The shikimate binding domain, which follows strand β 2, consists of helices α 2 and α 3 and the N-terminal region of helix α 4. A total of four hydrogen bonds between ecSK and shikimate, in the model, involving the residues Lys15, Asp34, and Arg 136 was observed. For the mtSK model the same pattern was observed. Fig. 5A and B show the main residues involved in contact with shikimate in the complexes. Tables 2 and 3 show the intermolecular hydrogen bonds for both structures. The residues involved hydrogen

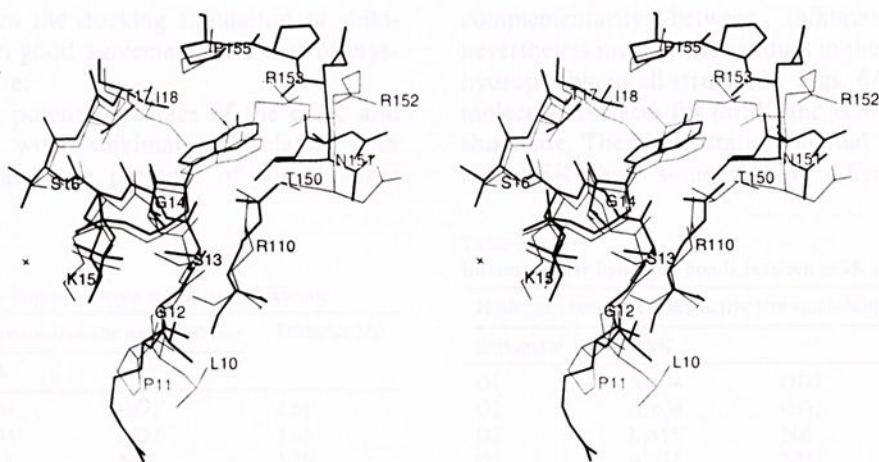


Fig. 4. Superimposed binding pockets of the ATP-binding site of mtSK (thick line) and ecSK (thin line).

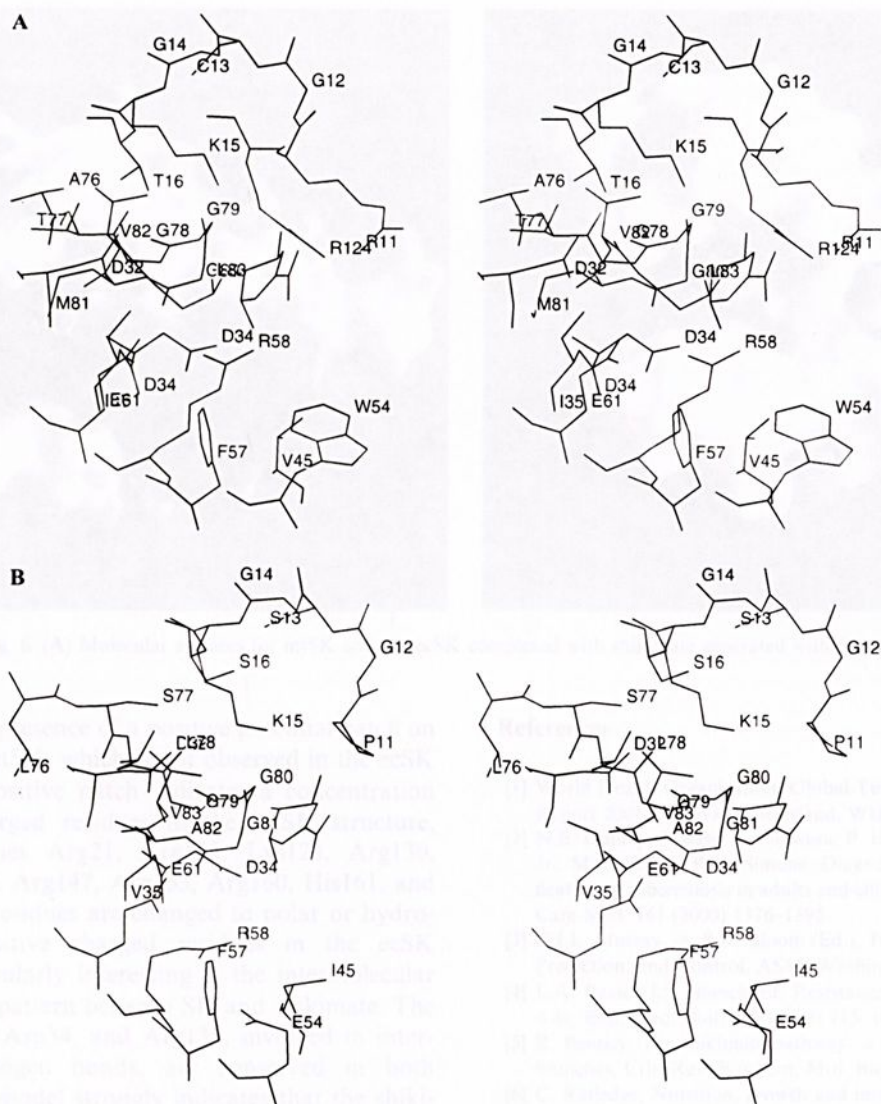


Fig. 5. Main residues involved in contact with shikimate in the complexes (A) mtSK:shikimate and (B) ecSK:shikimate.

bond identified from the docking simulation of shikimate to ecSK are in good agreement with that of crystallographic structure.

The electrostatic potential surface of the ecSK and mtSK complexed with shikimate calculated with GRASP [24] indicates the presence of some charge

complementarity between shikimate and enzyme, nevertheless most of the residues in the binding pocket are hydrophobic in all structures. Figs. 6A and B show the molecular surfaces for mtSK and ecSK complexed with shikimate. The electrostatic potential surfaces of mtSK and ecSK show some striking differences. The main

Table 2
Intermolecular hydrogen bonds between mtSK and shikimate

Hydrogen bonds between active site and inhibitor			Distance (Å)
Shikimate	mtSK		
O1	Asp34	OD2	2.58
O2	Asp34	OD2	3.55
O3	Lys15	NZ	3.36
O5	Arg136	NH1	3.47

Table 3
Intermolecular hydrogen bonds between ecSK and shikimate

Hydrogen bonds between active site and inhibitor			Distance (Å)
Shikimate	ecSK		
O1	Asp34	OD2	2.64
O2	Asp34	OD2	3.17
O3	Lys15	NZ	3.46
O5	Arg11	NH1	2.77
O5	Arg136	NH1	3.51

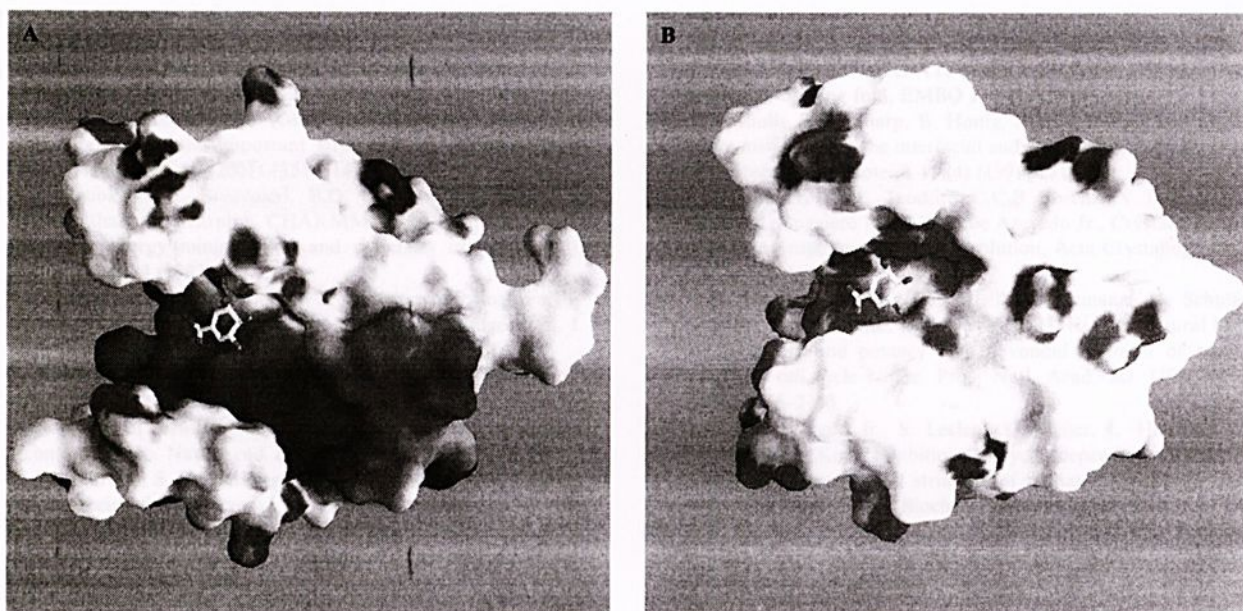


Fig. 6. (A) Molecular surfaces for mtSK and (B) ecSK complexed with shikimate generated with GRASP [24].

difference is the presence of a positive potential patch on the surface of mtSK, which is not observed in the ecSK surface. This positive patch indicates a concentration of positive charged residues in the mtSK structure, involving residues Arg21, Arg125, Lys128, Arg130, Lys135, Arg142, Arg147, Arg153, Arg160, His161, and Arg165. These residues are changed to polar or hydrophobic or negative charged residues in the ecSK sequence. Particularly interesting is the intermolecular hydrogen bond pattern between SK and shikimate. The residues Lys15, Asp34, and Arg136, involved in intermolecular hydrogen bonds, are conserved in both structures. The model strongly indicates that the shikimate binding domain is a well-conserved motif in SK structures. Furthermore, the alignment of 37 SK sequences, figure not shown, indicates that the main residues involved in intermolecular hydrogen bonds are conserved in all sequences. Such observation suggests that competitive inhibitors with shikimate will be able to inhibit most or even all SKs, since specificity and affinity between enzyme and its inhibitor depend on directional hydrogen bonds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners [25–30]. Further inhibition experiments may confirm this prediction.

Acknowledgments

This work was supported by grants from FAPESP (SMOLBNet), CNPq, CAPES, and Instituto do Milênio (CNPq-MCT). WFA (CNPq, 300851/98-7) and MSP (CNPq, 500079/90-0) are researchers for the Brazilian Council for Scientific and Technological Development.

References

- [1] World Health Organization, Global Tuberculosis Control, WHO Report 2001, Geneva, Switzerland, WHO/CDS/TB/2001.287.
- [2] N.E. Dunlap, J. Bass, P. Fujiwara, P. Hopewell, C.R. Horsburgh Jr., M. Salfinger, P.M. Simone, Diagnostic standards and classification of tuberculosis in adults and children, *Am. J. Respir. Crit. Care Med.* 161 (2000) 1376–1395.
- [3] C.J.L. Murray, in: B.R. Bloom (Ed.), *Tuberculosis: Pathogenesis, Protection, and Control*, ASM, Washington, 1994, pp. 583–622.
- [4] L.A. Basso, J.S. Blanchard, Resistance to antitubercular drugs, *Adv. Exp. Med. Biol.* 456 (1998) 115–144.
- [5] R. Bentley, The shikimate pathway—a metabolic tree with many branches, *Crit. Rev. Biochem. Mol. Biol.* 25 (1990) 307–384.
- [6] C. Ratledge, Nutrition, growth and metabolism, in: C. Ratledge, J.L. Stanford (Eds.), *The Biology of the Mycobacteria*, vol. 1, Academic Press, London, 1982, pp. 185–271.
- [7] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B.G. Barrell, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, *Nature* 393 (1998) 537–544.
- [8] T. Krell, J.R. Coggins, A.J. Lapthorn, The three-dimensional structure of shikimate kinase, *J. Mol. Biol.* 278 (1998) 983–997.
- [9] J.S. Oliveira, C.A. Pinto, L.A. Basso, D.S. Santos, Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate-3-phosphate synthase enzymes from *Mycobacterium tuberculosis*, *Protein Express. Purif.* 22 (2001) 430–435.
- [10] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [11] A. Sali, J.P. Overington, Derivation of rules for comparative protein modeling from a database of protein structure alignments, *Protein Sci.* 3 (9) (1994) 1582–1596.
- [12] A. Sali, L. Potterton, F. Yuan, H. van Vlijmen, M. Karplus, Evaluation of comparative protein modeling by MODELLER, *Proteins* 23 (3) (1995) 318–326.

- [13] A. Sali, Modeling mutations and homologous proteins, *Curr. Opin. Biotechnol.* 6 (4) (1995) 437–451.
- [14] T. Krell, J. Maclean, D.J. Boam, A. Cooper, M. Resmini, K. Brocklehurst, S.M. Kelly, N.C. Price, A.J. Lapthorn, J.R. Coggins, Biochemical and X-ray crystallographic studies on shikimate kinase: the important structural role of the P-loop lysine, *Protein Sci.* 10 (2001) 1137–1149.
- [15] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy minimization and dynamics calculations, *J. Comp. Chem.* 4 (1983) 187–217.
- [16] W. Braun, N. Go, Calculation of protein conformations by proton–proton distance constraints. A new efficient algorithm, *J. Mol. Biol.* 186 (3) (1985) 611–626.
- [17] E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A.A. Friesem, C. Aflalo, I.A. Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, *Proc. Natl. Acad. Sci. USA* 89 (1992) 2195–2199.
- [18] E. Schönbrunn, S. Eschenburg, W.A. Shuttleworth, J.V. Schloss, N. Amrhein, J.N.S. Evans, W. Kabsch, Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate-3-phosphate synthase in atomic detail, *Proc. Natl. Acad. Sci. USA* 98 (2001) 1376–1380.
- [19] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchinson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thornton, PROCHECK v.3.0—program to check the stereochemistry quality of protein structures—operating instructions, 1994.
- [20] G.E. Schulz, Binding of nucleotides by proteins, *Curr. Opin. Struct. Biol.* 2 (1992) 61–67.
- [21] C.W. Müller, G.J. Schlauender, J. Reinstein, G.E. Schulz, Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding, *Structure* 4 (1996) 147–156.
- [22] M. Gerstein, G. Schulz, C. Chothia, Domain closure in adenylate kinase Joints on either side of two helices close like neighboring fingers, *J. Mol. Biol.* 229 (2) (1993) 494–501.
- [23] J.E. Walker, M. Saraste, M.J. Runswick, N.J. Gay, Distantly related sequences in the α and β subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold, *EMBO J.* 1 (1982) 945–951.
- [24] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins* 11 (4) (1991) 281–296.
- [25] F. Canduri, L.G.V.L. Teodoro, C.C.B. Lorenzi, V. Hial, R.A.S. Gomes, J. Ruggiero Neto, W.F. de Azevedo Jr., Crystal structure of human uropepsin at 2.45 Å resolution, *Acta Crystallogr. D* 57 (2001) 1560–1570.
- [26] W.F. De Azevedo Jr., H.J. Mueller-Dieckmann, U. Schulze-Gahmen, P.J. Worland, E. Sausville, S.-H. Kim, Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase, *Proc. Natl. Acad. Sci. USA* 93 (7) (1996) 2735–2740.
- [27] W.F. De Azevedo Jr., S. Leclerc, L. Meijer, L. Havlicek, M. Strnad, S.-H. Kim, Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human CDK2 complexed with roscovitine, *Eur. J. Biochem.* 243 (1997) 518–526.
- [28] W.F. De Azevedo Jr., F. Canduri, V. Fadel, L.G.V.L. Teodoro, V. Hial, R.A.S. Gomes, Molecular model for the binary complex of uropepsin and pepstatin, *Biochem. Biophys. Res. Commun.* 287 (1) (2001) 277–281.
- [29] W.F. De Azevedo Jr., F. Canduri, N.J.F. da Silva, Structural basis for inhibition of cyclin-dependent kinase 9 by flavopiridol, *Biochem. Biophys. Res. Commun.* 293 (1) (2002) 566–571.
- [30] S.-H. Kim, U. Schulze-Gahmen, J. Brandsen, W.F. de Azevedo Jr., Structural basis for chemical inhibitor of CDK2, *Prog. Cell Cycle Res.* 2 (1996) 137–145.
- [31] D.G. Higgins, A.J. Bleasby, R. Fuchs, CLUSTAL V: improved software for multiple sequence alignment, *Comput. Appl. Biosci.* 8 (2) (1992) 189–191.
- [32] P. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, *J. Appl. Cryst.* 24 (1991) 946–950.

Anexo 4



Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*

José Henrique Pereira,^a Ferranda Canduri,^{a,b,*} Iam Soares de Oliveira,^c Nelson José Freitas de Silva,^a Luiz Augusto Basso,^c Mário Sérgio Palma,^{b,d} Walter Filgueira de Almeida Jr.,^{a,b,*} and Diógenes Santiago Santos^{a,c}

^a Departamento de Física, UFPA, Rua 48, s/n, Belém, PA, 66090-100, Brazil

^b Instituto de Física, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21945-970, Brazil

^c Laboratório de Física de Materiais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21945-970, Brazil

^d Instituto de Física, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21945-970, Brazil

^e Laboratório de Física de Materiais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21945-970, Brazil

^f Instituto de Física, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 21945-970, Brazil

Received 24 October 2011

Abstract

The glutamate pathway is an essential element for herbivores and microorganisms, being degraded because it is considered an essential amino acid, leucine, and fungi but absent from mammals. Herbivores, as in response to the herbivores pathway have been identified in the genome sequence of *Mycobacterium tuberculosis*. Among them, the EPSP synthase was proposed to be present in *Mycobacterium tuberculosis*. Accordingly, in order to pave the way for structural and functional effects both on antimicrobial agent development, here we describe a molecular modeling of *Mycobacterium tuberculosis* Glutamate Synthase (EPSP synthase) isolated from *M. tuberculosis* that also provides a structural model of the enzyme in the active site region. The design of active inhibitors may be based on the structural information of the enzyme in the active site region. The possible relevance of this structural information to the design of active inhibitors is discussed.

© 2012 Elsevier Inc. All rights reserved.

Keywords: EPSP synthase; *Mycobacterium tuberculosis*; Molecular modeling; Antimicrobial agent

Tuberculosis (TB) remains one of the most deadly infectious diseases in the world. It is estimated that approximately 1 billion individuals are infected with latent TB. *Mycobacterium tuberculosis* has caused nearly 2 billion deaths in the last 1000 years and more than 2 million people each year worldwide, according to the World Health Organization (WHO) which in 1993 declared tuberculosis to be a global emergency [1].

Mycobacterium tuberculosis is the most important opportunistic agent responsible for TB in all deaths and 1/3 of all people who die in developing world.

Several vaccines against TB have been used with no positive results. 70–90% of TB cases are in the immunocompetent population [2–10] and are not fatal [11].

Treatment of tuberculosis is a long process, and the drugs are used only for long-term treatment against TB. Treatment of the active disease includes isoniazid, rifampin, ethambutol, and pyrazinamide. Treatment of the active disease includes isoniazid, rifampin, ethambutol, and pyrazinamide. Treatment of the active disease includes isoniazid, rifampin, ethambutol, and pyrazinamide. Treatment of the active disease includes isoniazid, rifampin, ethambutol, and pyrazinamide.

Recent culture of tuberculosis caused by multi-drug-resistant (MDR) strains, mainly in individuals infected with HIV, have caused a warning against the standard and new drugs. A worldwide network is conducting various strategies of development of new drugs.

Anexo 4

Corresponding authors. Fax: +55 11 21251237.

E-mail addresses: canduri@fisica.uerj.br (F. Canduri), wfilgueira@fisica.uerj.br (W. Filgueira de Almeida Jr.), dsantos@fisica.uerj.br (D. Santiago Santos).

© 2012 Elsevier Inc. All rights reserved. This article is an open access article under the CC BY-NC-ND 4.0 International license.



Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*

José Henrique Pereira,^a Fernanda Canduri,^{a,b,*} Jaim Simões de Oliveira,^c
Nelson José Freitas da Silveira,^a Luiz Augusto Basso,^c Mário Sérgio Palma,^{b,d}
Walter Filgueira de Azevedo Jr.,^{a,b,*} and Diógenes Santiago Santos^{c,*}

^a Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil

^b Center for Applied Toxinology, Instituto Butantan, São Paulo, SP 05503-900, Brazil

^c Rede Brasileira de Pesquisa de Pesquisas em Tuberculose Grupo de Microbiologia Molecular e Funcional,
Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil

^d Laboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil

^e Faculdade de Farmácia/Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil

Received 23 October 2003

Abstract

The shikimate pathway is an attractive target for herbicides and antimicrobial agent development because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals. Homologues to enzymes in the shikimate pathway have been identified in the genome sequence of *Mycobacterium tuberculosis*. Among them, the EPSP synthase was proposed to be present by sequence homology. Accordingly, in order to pave the way for structural and functional efforts towards anti-mycobacterial agent development, here we describe the molecular modeling of 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase isolated from *M. tuberculosis* that should provide a structural framework on which the design of specific inhibitors may be based on. Significant differences in the relative orientation of the domains in the two models result in “open” and “closed” conformations. The possible relevance of this structural transition in the ligand binding is discussed.

© 2003 Elsevier Inc. All rights reserved.

Keywords: EPSP synthase; Bioinformatics; Molecular modeling; *Mycobacterium tuberculosis*

Tuberculosis (TB) remains one of the most deadly infectious diseases in the world. It is estimated that approximately 1 billion individuals are infected with latent TB. Tuberculosis has made a steady global comeback in the late 1980s and now kills more than 2 million people each year worldwide, according to the World Health Organization (WHO) which in 1993 declared tuberculosis to be a global emergence [1].

Mycobacterium tuberculosis is by far the biggest killer among infectious agents, accounting for 7% of all deaths and 26% of all avoidable deaths. In impoverished third

world countries where 95% of cases and 98% of deaths occur, 70–80% of TB cases are in the economically most productive 15–50 year age bracket [1,2].

The treatment of tuberculosis is a special problem in the field of chemotherapy. Many of the drugs employed to treat the disease are used only for treating infections caused by mycobacteria. Treatment of the active case of TB always includes simultaneous therapy with two or more of the frontline drugs: isoniazid, ethambutol, rifampicin, and streptomycin which are used to decrease the rate of emergence of resistant strains as well to increase the antibacterial effect [3–5].

Recent outbreaks of tuberculosis caused by multi-drug-resistant (MDR) strains, mainly in individuals infected with HIV, have created a scaring element to the scenario and also created a worldwide interest in expanding current programs of development of new drugs

* Corresponding authors. Fax: +55172212247.

E-mail addresses: canduri@df.ibilce.unesp.br (F. Canduri), walterfa@df.unesp.br (W.F. Azevedo Jr.), diogenes@puers.br (D.S. Santos).



different in kinds from existing ones and based on the principle of selective toxicity on enzymes or structure of the bacillus to both treat *M. tuberculosis* strains resistant to existing drugs and shorten the duration of short-course treatment to improve patient compliance [4,6].

The shikimate pathway is an attractive target for the development of herbicides and antimicrobial agents because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals [7]. In mycobacteria, the shikimate pathway leads to the biosynthesis of precursors for the synthesis of aromatic amino acids, naphthoquinones, menaquinones, and mycobactin [8]. Homologues to enzymes in the shikimate pathway have been identified in the complete genome sequence of *M. tuberculosis* H37Rv strain [9]. Among them, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase encoded by *aroA* gene, which catalyzes the transfer of the enolpyruvyl moiety from phosphoenolpyruvate (PEP) and inorganic phosphate. A valuable lead compound in the search of new inhibitors is glyphosate, which has proven to be potent and specific inhibitor of EPSP synthase [10]. Glyphosate is successfully used as a herbicide, being the active ingredient of the widely used weed control agent Roundup Ready, and was recently shown to inhibit the growth of the pathogenic parasites *Plasmodium falciparum*, *Toxoplasma gondii*, and *Cryptosporidium parvum* [11,12].

The present paper describes the two molecular models of *M. tuberculosis* EPSP synthase, one without any ligand and another in complex with 3-phosphoshikimate (S3P) and glyphosate. The homology model-

ing was performed using the crystallographic structures of EPSP synthase from *Escherichia coli* [12,13], as templates. The EPSP synthase has been cloned, sequenced, and overexpressed in soluble and functional form [14], thus allowing enzymological studies to be performed. The results presented here should provide a three-dimensional model of EPSP synthase to both guide enzymological studies and aid in the design of specific inhibitors.

Methods

Molecular modeling. For modeling of the EPSP synthase we used restrained-based modeling implemented in the program MODELLER [15]. This program is an automated approach to comparative modeling by satisfaction of spatial restraints [16–18]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). This alignment is usually the input to the program. The output is a three-dimensional model for the target sequence containing all main-chain and side-chain non-hydrogen atoms.

The degree of primary sequence identity between MtEPSP synthase and EPSP synthase from *E. coli* indicates that the crystallographic structures of EcEPSP are good models to be used as templates for MtEPSP synthase. The EPSP synthase isolated from *E. coli* showed two conformation states, open (unliganded structure) and closed (liganded structure) [12]. Models in the two conformations were generated, using the α -carbon atomic coordinates from 1EPSP [13] to generate the open structure and the atomic coordinates from 1G6S [12] to generate the closed structure. The atomic coordinates of all waters were removed from the EPSP structure. The 3-phosphoshikimate and glyphosate of the template were kept in the structure for the closed

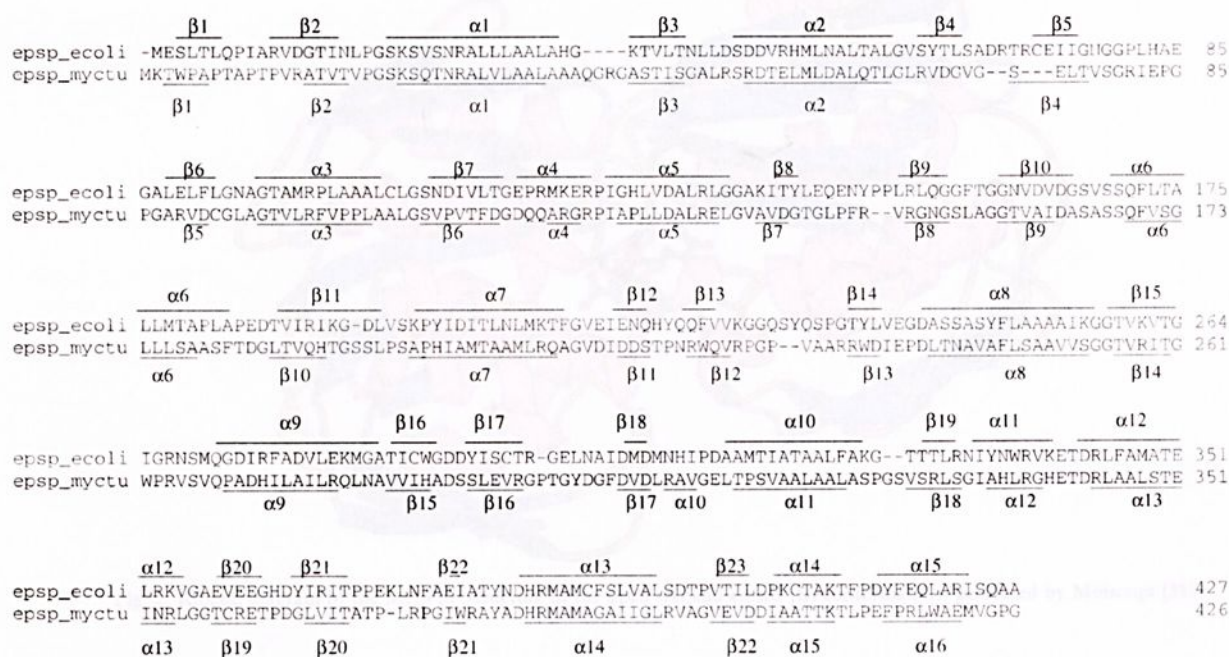


Fig. 1. The sequence alignment of EcEPSP and MtEPSP indicating the secondary structural elements. The sequence MtEPSP shows 31% of identity with the sequence of EcEPSP. The alignment was performed with the program CLUSTAL V [37].

conformation. Next, the spatial restraints and CHARMM energy terms enforcing proper stereochemistry [19] were combined into an objective function. Finally, the model is obtained by optimizing the objective function in Cartesian space. The optimization is carried out by the use of the variable target function method [20] employing methods of conjugate gradients. Several slightly different models can be calculated by varying the initial structure. A total of 1000 models were generated for each EPSP synthase conformation state, and the final models were selected based on stereochemical quality. The optimization of the complex was carried out by the use of the variable target function method [15] employing methods of conjugate gradients and molecular dynamics with simulated annealing. All modeling process was 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 model for EPSP complex was assessed by the program PROCHECK [21]. The cutoff for hydrogen bonds and salt bridges was 3.6 Å.

Results and discussion

Primary sequence comparison

The sequence alignment of EcEPSP (template) and MtEPSP (target) is shown in Fig. 1. The secondary structural elements are indicated in the figure. The sequence MtEPSP shows 31% identity with the sequence of EcEPSP.

Quality of the model

The Ramachandran plot for the 2 EcEPSP structures was generated in order to compare the overall stereochemical quality of MtEPSP model against EPSP

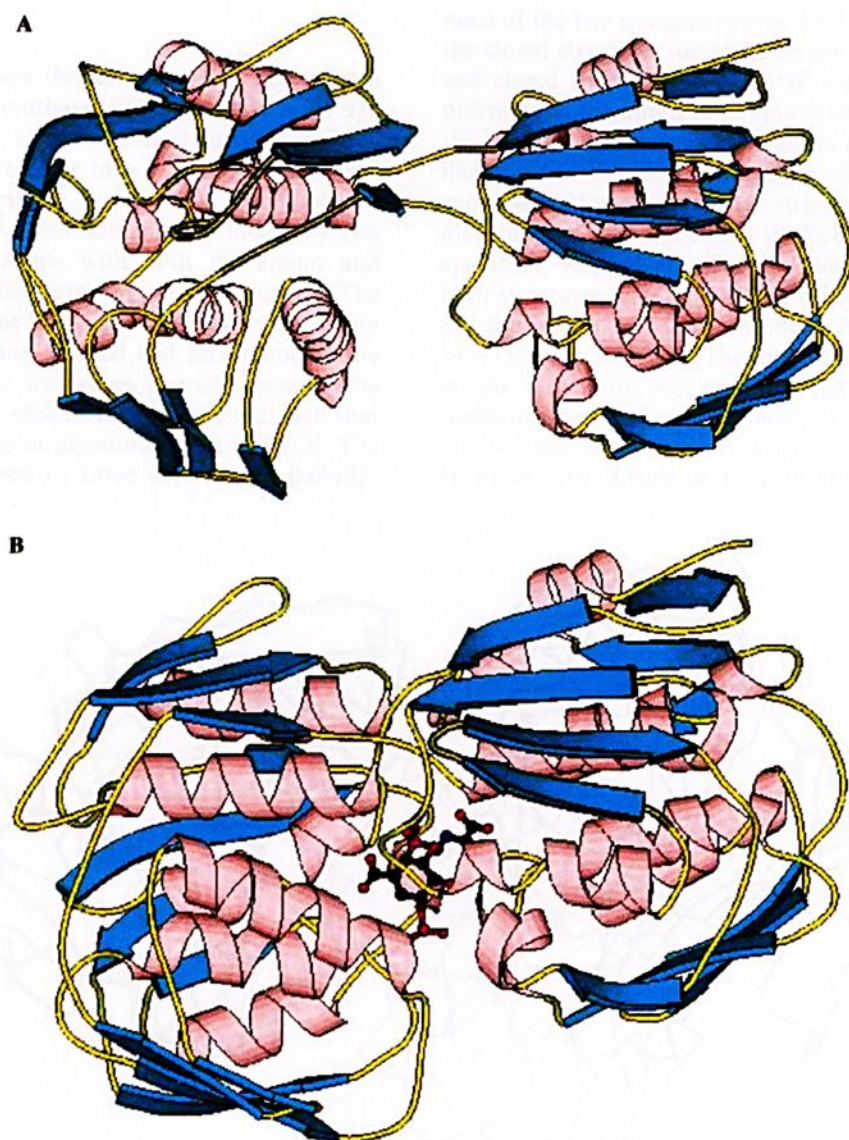


Fig. 2. Ribbon diagram of the MtEPSP structure in the open (A) and closed (B) conformations generated by Molscript [38].

structures solved by biocrystallography. Analysis of the Ramachandran plot of the MtEPSP models shows that 84.8% of the residues lie in the most favorable regions and the remaining 10.9% in the additional allowed regions for the closed structure, and 91.1% lie in the allowed regions for the open structure. The same analysis for two crystallographic EcEPSP structures present 91.2% of residues in the most favorable, 8.5% additional allowed regions, and 0.3% generously allowed regions. The overall rating for the MtEPSP model is slightly poorer than the one obtained for the two structures of EcEPSP. However, it has over 90% of the residues in the most favorable regions.

Overall description

Figs. 2A and B show the secondary structure of both models of MtEPSP synthase. EPSP synthase is an α/β protein consisting of a mixed β sheet surrounded by α helices. The structure folds into two similar domains, which are approximately hemispheric, each with a radius of about 21 Å. The domains are linked by two crossover chain segments with both the amino and carboxyl termini of the protein in the lower domain. The two flat surfaces of the hemispheres, which in projection form a "V," are almost normal and accommodate the amino termini of the six helices in each domain. The helical macrodipolar effects create a potential well that facilitates the binding of glyphosate and S3P [13]. The domains are composed of three copies of a $\beta\alpha\beta\alpha\beta$ -

folding unit. The four-stranded β -sheet structures contain both parallel and antiparallel strands and the helices are parallel. In the closed structure the glyphosate and S3P molecules are bound in the interdomain cleft, promoting the closure of the structure.

Molecular hinge

Both crystal structures of EcEPSP synthase and the molecular models of MtEPSP synthase reveal the presence of two domains in which the domains, designated as domains A and B, undergo a translation and rotation from the open to the closed conformation. The center of mass of the two domains moves 4.6 Å from the open to the closed structure, for MtEPSP synthase. In the open and closed structures of MtEPSP synthase an angular difference at the interface of approximately 69° between the two domains results in different conformations, as illustrated in Fig. 3. Superposition of C α atoms of the constituent domains of each structure results in rms differences of 0.20 and 0.39 Å for EcEPSP and MtEPSP synthases, respectively. The individual domains from both structures were superposed using the C α atoms in the four major α -helices (residues 240–252, 270–280, 343–355, and 385–395). The low rmsd values, observed in the superpositions, indicate that the alternative conformation of the two-domain structures is due to relative motions of structurally conserved domains. However, the difference in conformation suggests a

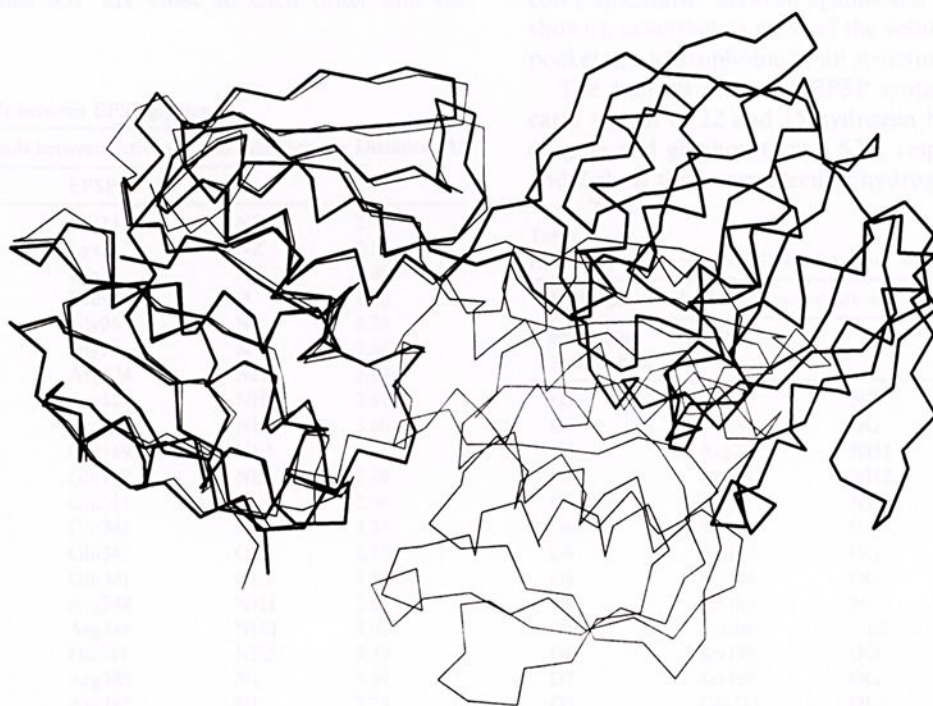


Fig. 3. Molecular hinge. α -Carbon traces of the superimposed structures of MtEPSP synthase from the open conformation (dark line) and closed conformation (light line).

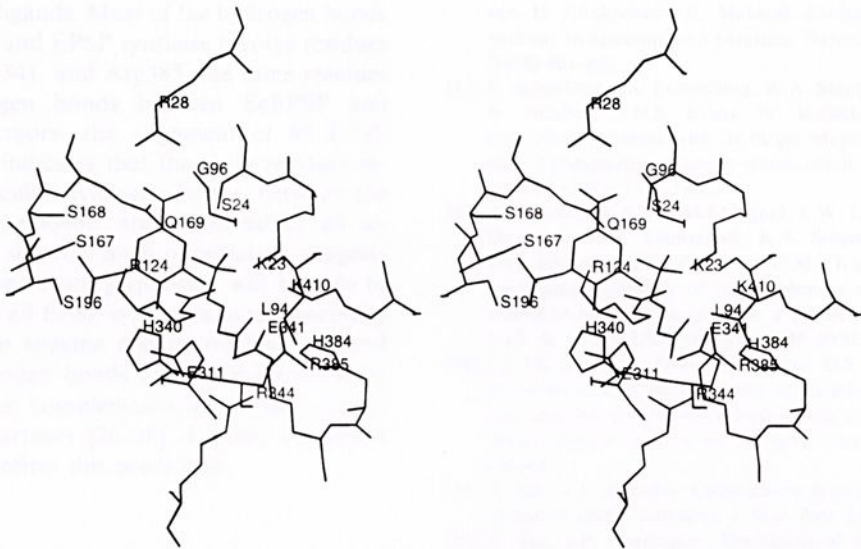


Fig. 4. Stereo view of the interdomain interface region illustrating the binding of glyphosate and S3P in the MtEPSP synthase structure.

degree of flexibility in the interdomain contact at the interface region, which acts as a hinge between the two monomers resulting in “open” and “closed” forms of the EPSP synthase structure. Similar molecular hinge is observed in PLA₂ structures [22–24].

3-Phosphoshikimate and glyphosate binding

Fig. 4 shows the active site of MtEPSP synthase. The glyphosate and S3P are close to each other and the

5-hydroxyl group of S3P is hydrogen bonded to the nitrogen atom of glyphosate with a distance of 2.82 Å. Analysis of the molecular surface of the MtEPSP synthase indicates that both ligands are buried in between the two domains, with a contact area of 29 and 134 Å² for glyphosate and 3-phosphoshikimate, respectively. The electrostatic potential surface of the MtEPSP synthase complexed with glyphosate and S3P calculated with GRASP [25] indicates the presence of some charge complementarity between ligands and enzyme (figure not shown), nevertheless most of the residues in the binding pocket are hydrophobic in all structures (see Fig. 4).

The analysis of the MtEPSP synthase complex indicates a total of 22 and 15 hydrogen bonds between the enzyme and glyphosate and S3P, respectively. Tables 1 and 2 show the intermolecular hydrogen bonds between

Table 1
Hydrogen bonds between EPSP:glyphosate

Hydrogen bonds between active site and inhibitor			Distance (Å)
Glyphosate	EPSP		
O4	Lys23	NZ	2.98
O1	Lys23	NZ	2.84
N1	Lys23	NZ	3.46
O3	Leu94	O	3.10
O2	Gly96	N	3.53
O3	Gly96	N	2.87
O3	Arg124	NH1	2.87
O2	Arg124	NH2	2.81
O3	Arg124	NH2	3.60
O2	Gln169	NE1	3.58
O2	Gln169	NE2	2.88
O5	Glu311	OE1	2.96
N1	Glu341	OE1	3.35
N1	Glu341	OE2	2.88
O4	Glu341	OE2	3.59
O5	Arg344	NH1	2.81
O5	Arg344	NH2	3.02
O4	His384	NE2	3.49
O4	Arg385	NE	3.59
O5	Arg385	NE	2.75
O4	Arg385	NH2	3.12
O3	Lys410	NZ	2.95

Table 2
Hydrogen bonds between EPSP:shikimate 3-phosphate

Hydrogen bonds between active site and ligand			Distance (Å)
Shikimate 3-phosphate	EPSP		
O3	Lys23	NZ	2.81
O5	Ser24	OG	2.70
O5	Arg28	NH1	2.86
O4	Arg28	NH2	2.78
O5	Arg28	NH2	3.60
O6	Ser167	OG	3.40
O8	Ser167	OG	2.70
O6	Ser168	OG	2.69
O6	Ser168	N	2.84
O1	Gln169	NE2	3.49
O6	Ser196	OG	3.59
O7	Ser196	OG	2.69
O2	Glu311	OE2	2.65
O2	His340	NE2	3.14
O1	His340	NE2	3.57

the enzyme and the ligands. Most of the hydrogen bonds between glyphosate and EPSP synthase involve residues Lys23, Arg124, Glu341, and Arg385, the same residues involved in hydrogen bonds between EcEPSP and glyphosate. Furthermore, the alignment of 69 EPSP synthase sequences indicates that the main residues involved in intermolecular hydrogen bonds, between the glyphosate and the enzyme, are conserved in all sequences (figure not shown). Such observation suggests that inhibitors derived from glyphosate will be able to inhibit most or even all EPSP synthases, since specificity and affinity between enzyme and its inhibitor depend on directional hydrogen bonds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners [26–36]. Further inhibition experiments may confirm this prediction.

Acknowledgments

This work was supported by Grants from FAPESP (SMOLBNet, Proc. 01/07532-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT) to DSS and LAB. WFA (CNPq, 300851/98-7) and MSP (CNPq, 500079/90-0) are researchers for the National Research Council.

References

- [1] World Health Organization. Global Tuberculosis Control. WHO Report 2001. Geneva, Switzerland, WHO/CDS/TB/2001.287.
- [2] N.E. Dunlap, J. Bass, P. Fujiwara, P. Hopewell, C.R. Horsburgh Jr., M. Salfinger, P.M. Simione, Diagnostic standards and classification of tuberculosis in adults and children, *Am. J. Respir. Crit. Care Med.* 161 (2000) 1376–1395.
- [3] J.D. McKinney, W.R. Jacobs Jr., B.R. Bloom, Persisting problems in tuberculosis, in: R.M. Krause (Ed.), *Emerging Infections*, Academic Press, New York, 1998, pp. 51–146.
- [4] C.J.L. Murray, Issues in operational, social, and economic research on tuberculosis, in: B.R. Bloom (Ed.), *Tuberculosis: Pathogenesis, Protection, and Control*, ASM Press, Washington, DC, 1994, pp. 583–622.
- [5] B.R. Bloom, C.J.L. Murray, Tuberculosis: commentary on a reemergent killer, *Science* 257 (1992) 1055–1064.
- [6] L.A. Basso, J.S. Blanchard, Resistance to antitubercular drugs, *Adv. Exp. Med. Biol.* 456 (1998) 115–144.
- [7] R. Bentley, The shikimate pathway—A metabolic tree with many branches, *Crit. Rev. Biochem. Mol. Biol.* 25 (1990) 307–384.
- [8] C. Ratledge, Nutrition, growth and metabolism, in: C. Ratledge, J.L. Stanford (Eds.), *The Biology of the Mycobacteria*, vol. 1, Academic Press, London, 1982, pp. 185–271.
- [9] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry III, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B.G. Barrell, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, *Nature* 393 (1998) 537–544.
- [10] H.C. Steinrücken, N. Amrhein, The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase, *Biochem. Biophys. Res. Commun.* 94 (4) (1980) 1207–1212.
- [11] F. Roberts, C.W. Roberts, J.J. Johnson, D.E. Kyle, T. Krell, J.R. Coggins, G.H. Coombs, W.K. Milhous, S. Tzipori, D.J. Ferguson, D. Chakrabarti, R. McLeod, Evidence for the shikimate pathway in apicomplexan parasites, *Nature (London)* 395 (6699) (1998) 801–805.
- [12] E. Schönbrunn, S. Eschenburg, W.A. Shuttleworth, J.V. Schloss, N. Amrhein, J.N.S. Evans, W. Kabsch, Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail, *Proc. Natl. Acad. Sci. USA* 98 (2001) 1376–1380.
- [13] W.C. Stallings, S.S. Abdel-Meguid, L.W. Lim, H.-S. Shieh, H.E. Dayringer, N.K. Leimgruber, R.A. Stegeman, K.S. Anderson, J.A. Sikorski, S.R. Padgett, G.M. Kishore, Structure and topological symmetry of the glyphosate target 5-enol-pyruvylshikimate-3-phosphate synthase: a distinctive protein fold, *Proc. Natl. Acad. Sci. USA* 88 (1991) 5046–5050.
- [14] J.S. Oliveira, C.A. Pinto, L.A. Basso, D.S. Santos, Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*, *Protein Express. Purif.* 22 (2001) 430–435.
- [15] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [16] A. Sali, J.P. Overington, Derivation of rules for comparative protein modeling from a database of protein structure alignments, *Protein Sci.* 3 (9) (1994) 1582–1596.
- [17] A. Sali, L. Potterton, F. Yuan, H. van Vlijmen, M. Karplus, Evaluation of comparative protein modeling by MODELLER, *Proteins* 23 (3) (1995) 318–326.
- [18] A. Sali, Modeling mutations and homologous proteins, *Curr. Opin. Biotechnol.* 6 (4) (1995) 437–451.
- [19] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy minimization and dynamics calculations, *J. Comp. Chem.* 4 (1983) 187–217.
- [20] W. Braun, N. Go, Calculation of protein conformations by proton-proton distance constraints. A new efficient algorithm, *J. Mol. Biol.* 186 (3) (1985) 611–626.
- [21] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchinson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thornton, PROCHECK v.3.0 Program to Check the Stereochemistry Quality of Protein structures—Operating Instructions, 1994.
- [22] W.F. De Azevedo Jr., R.J. Ward, F.R. Lombardi, J.R. Giglio, A.M. Soares, M.R.M. Fontes, R.K. Arni, Crystal structure of myotoxin-II: a myotoxic phospholipase A₂-homologue from *Bothrops moojeni* venom, *Protein Peptide Lett.* 4 (5) (1997) 329–334.
- [23] W.F. De Azevedo Jr., R.J. Ward, J.M. Gutierrez, R.K. Arni, Structure of a Lys49 phospholipase A₂ homologue isolated from the venom of *Bothrops nummifer* (jumping viper), *Toxicon* 37 (2) (1999) 371–384.
- [24] W.F. De Azevedo Jr., R.J. Ward, F. Canduri, A. Soares, J.R. Giglio, R.K. Arni, Crystal structure of piratoxin-I: a calcium-independent, myotoxic phospholipase A₂-homologue from *Bothrops pirajai* venom, *Toxicon* 36 (10) (1998) 1395–1406.
- [25] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins* 11 (4) (1991) 281–296.
- [26] F. Canduri, L.G.V.L. Teodoro, C.C.B. Lorenzi, V. Hial, R.A.S. Gomes, J. Ruggiero Neto, W.F. de Azevedo Jr., Crystal structure of human uropepsin at 2.45 Å resolution, *Acta Crystallogr. D* 57 (2001) 1560–1570.
- [27] W.F. De Azevedo Jr., H.J. Mueller-Dieckmann, U. Schulze-Gahmen, P.J. Worland, E. Sausville, S.-H. Kim, Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase, *Proc. Natl. Acad. Sci. USA* 93 (7) (1996) 2735–2740.
- [28] W.F. De Azevedo Jr., S. Leclerc, L. Meijer, L. Havlicek, M. Strnad, S.-H. Kim, Inhibition of cyclin-dependent kinases by

- purine analogues: crystal structure of human CDK2 complexed with roscovitine, *Eur. J. Biochem.* 243 (1997) 518–526.
- [29] W.F. De Azevedo Jr., F. Canduri, V. Fadel, L.G.V.L. Teodoro, V. Hial, R.A.S. Gomes, Molecular model for the binary complex of uropepsin and pepstatin, *Biochem. Biophys. Res. Commun.* 287 (1) (2001) 277–281.
- [30] W.F. De Azevedo Jr., F. Canduri, N.J.F. da Silveira, Structural basis for inhibition of cyclin-dependent kinase9 by Flavopiridol, *Biochem. Biophys. Res. Commun.* 293 (1) (2002) 566–571.
- [31] S.-H. Kim, U. Schulze-Gahmen, J. Brandsen, W.F. de Azevedo Jr., Structural basis for chemical inhibitor of CDK2, *Prog. Cell Cycle Res.* 2 (1996) 137–145.
- [32] W.F. De Azevedo, R.T. Gaspar, F. Canduri, J.C. Camera, N.J.F. da Silveira, Molecular model of cyclin-dependent kinase 5 complexed with roscovitine, *Biochem. Biophys. Res. Commun.* 297 (2002) 1154–1158.
- [33] W.F. De Azevedo Jr., J.S. de Oliveira, L.A. Basso, M.S. Palma, J.H. Pereira, F. Canduri, D.S. Santos, Molecular model of shikimate kinase from *Mycobacterium tuberculosis*, *Biochem. Biophys. Res. Commun.* 295 (1) (2002) 142–148.
- [34] W.F. De Azevedo, F. Canduri, D.M. Santos, R.G. Silva, J.S. Oliveira, L.P.S. Carvalho, L.A. Basso, M.A. Mendes, M.S. Palma, D.S. Santos, Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution, *Biochem. Biophys. Res. Commun.* 308 (3) (2003) 545–552.
- [35] D.M. Santos, F. Canduri, J.H. Pereira, M.V.B. Dias, R.G. Silva, M.A. Mendes, M.S. Palma, L.A. Basso, W.F. de Azevedo, D.S. Santos, Crystal structure of human purine nucleoside phosphorylase complexed with acyclovir, *Biochem. Biophys. Res. Commun.* 308 (3) (2003) 553–559.
- [36] W.F. De Azevedo, F. Canduri, D.M. Dos Santos, J.H. Pereira, M.V.B. Dias, R.G. Silva, M.A. Mendes, L.A. Basso, M.S. Palma, D.S. Santos, Structural basis for inhibition of human PNP by immucillin-H, *Biochem. Biophys. Res. Commun.* 309 (2003) 922–927.
- [37] D.G. Higgins, A.J. Bleasby, R. Fuchs, CLUSTAL V: improved software for multiple sequence alignment, *Comput. Appl. Biosci.* 8 (2) (1992) 189–191.
- [38] P. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, *J. Appl. Cryst.* 24 (1991) 946–950.

Anexo 5

Molecular modeling and CD analysis of chorismate synthase

João Henrique Pereira^{1*}, Márcio Vinícius Bergamaschi Dias^{2,3}, Fernanda Candori^{1,4}, Fernanda Ely⁵, João Ruggiero Neto⁶, Jeyaraj Varma⁷, Luis Augusto Basso⁸, Mário Sérgio Palma^{9,10}, Diogenes Santiago Santos¹¹, Walter Pinheiro de Azevedo Jr.¹²

¹ Programa de Pós-Graduação em Biofísica Molecular - Departamento de Física, UNESP, São José do Rio Preto, SP 13054-000, Brazil

² Centro de Toxicologia Aplicada, Instituto Botânico, São Paulo, SP 05503-900, Brazil

³ Rede Brasileira de Pesquisa em Tuberculose Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, ITRGS, Porto Alegre, RS 91501-970, Brazil

⁴ Laboratório de Biologia Estrutural e Zooquímica, CIJS/Departamento de Biologia, Instituto de Biociências, UNESP, Rio Claro, SP 13506-900, Brazil

⁵ Faculdade de Farmácia/Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil

⁶ J. H. Pereira and M. V. B. Dias Contributed equally to this work

⁷ Corresponding author

E-mail addresses of the corresponding authors: walteria@fot.unesp.br (W.F. de Azevedo Jr.) and diogenes@pors.br (D.S. Santos).

Anexo 5

Keywords: Chorismate synthase, Circular dichroism, Molecular modeling, Mycobacterium tuberculosis, Drug design.

Molecular modeling and CD analysis of chorismate synthase

José Henrique Pereira^{a,1}, Marcio Vinicius Bertacine Dias^{a,1}, Fernanda Canduri^{a,b}, Fernanda Ely^c, João Ruggiero Neto^a, Jeverson Frazzon^c, Luiz Augusto Basso^c, Mário Sérgio Palma^{b,d}, Diógenes Santiago Santos^{e*}, Walter Filgueira de Azevedo Jr^{a,b*}

^a Programa de Pós-Graduação em Biofísica Molecular - Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brasil

^b Centro de Toxicologia Aplicada, Instituto Butantan, São Paulo, SP 05503-900, Brasil

^c 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

^d Laboratório de Biologia Estrutural e Zooquímica, CEIS/Departamento de Biologia, Instituto de Biociências, UNESP, Rio Claro, SP 13506-900, Brasil

^e Faculdade de Farmácia/Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil

¹ J. H. Pereira and M. V. B. Dias contributed equally to this work

* Corresponding author

E-mail addresses of the corresponding authors: walterfa@df.unesp.br (W.F.de Azevedo Jr.) and diogenes@pucrs.br (D.S. Santos).

Keywords: Chorismate synthase, Circular dichroism, Molecular modeling, Mycobacterium tuberculosis, Drug design.

Abstract

Tuberculosis is considered a worldwide health problem mainly due to co-infection with HIV and proliferation of multi-drug-resistant strains. 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 (EPSP) to chorismate. Here, it is described molecular modeling of chorismate synthase from *Mycobacterium tuberculosis* complexed with EPSP and FMN (Flavin-mononucleotide), and analysis of circular dichroism (CD). The atomic coordinates of the structure of CS from *Streptococcus pneumoniae* were used as template for starting the modeling of MtCS. The molecular model presents good stereochemistry quality and CD analysis confirms the prediction of the secondary structure. The molecular modeling of MtCS should provide a structural framework on which the design of specific inhibitors may be based.

The catalyzed reaction involves no net redox change. The role of the reduced FMN in catalysis is still not clear. However, recent detailed kinetic and biorganic approaches have fundamentally advanced our understanding of the mechanism of action of chorismate synthase [6].

Chorismate synthase is particularly attractive as an anti-infective target as chorismate lies at a metabolic node, being the precursor for five distinct pathways. It is necessary for the production of aromatic amino acids, para-aminobenzoic acid (PABA), folate, and for other cyclic metabolites such as ubiquinone and menaquinone [7]. Much of the folate pathway is also absent in mammals, and enzymes within it have therefore been successfully explored as targets for antibacterial chemotherapy, as exemplified in the

Introduction

Tuberculosis resurged in the mid-1980s and now kills approximately 3 million people a year. The reemergence of tuberculosis as a public health threat, the high susceptibility of HIV-infected persons, and the proliferation of multi-drug-resistant strains have created a need to develop new drugs. Potential targets for the development of new therapies are the enzymes of shikimate pathway, because they are essential for bacteria, fungi and apicomplexan parasites, but absent from mammals [1,2]. This pathway consists of seven enzymes that catalyze the sequential conversion of erythrose-4-phosphate and phosphoenol pyruvate to chorismic acid (chorismate), the common precursor of aromatic compounds [3]. Chorismate synthase catalyses the most interesting and unusual reaction of the entire pathway, the conversion of 5-enolpyruvyl-3-shikimate phosphate (EPSP) to chorismate, via the 1,4-antielimination of phosphate and a proton, a reaction that is unique in nature [4,5]. The enzyme has an absolute requirement for reduced FMN as a cofactor, although the catalyzed reaction involves no net redox change. The role of the reduced FMN in catalysis still not is all clear. However, recent detailed kinetic and bioorganic approaches have fundamentally advanced our understanding of the mechanism of action of chorismate synthase [6].

Chorismate synthase is particularly attractive as an anti-infective target as chorismate lies at a metabolic node, being the precursor for five distinct pathways. It is necessary for the production of aromatic amino acids, para-aminobenzoic acid (PABA), folate, and for other cyclic metabolites such as ubiquinone and menaquinone [7]. Much of the folate pathway is also absent in mammals, and enzymes within it have therefore been successfully exploited as targets for antibacterial chemotherapy, as exemplified in the

inhibition of dihydrofolate reductase (DHFR) by trimethoprim. Despite of the crystallization of Chorismate Synthase from *Mycobacterium tuberculosis* has been reported [8], the structure was not determined, yet.

The present paper describes the molecular model of *M. tuberculosis* chorismate synthase (MtCS) complexed with FMN and EPSP, and analysis of circular dichroism spectra of protein in the presence and absence of co-factor.

Methods

Molecular modeling

For modeling of the MtCS was used restrained-based modeling implemented in the program MODELLER [9]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). This alignment is usually the input to the program. The output is a three-dimensional model for the target sequence containing all main-chain and side-chain non-hydrogen atoms. The atomic coordinates of the crystallographic structure of *Streptococcus pneumoniae* chorismate synthase (SpCS) (PDB access code: 1QXO) [10] solved at 2.0 Å resolution were used as starting model for modeling of the CS from *M. tuberculosis* complexed with FMN and EPSP. The atomic coordinates of all waters were removed from the SpCS structure. Several slightly different models can be calculated by varying the initial structure. A total of 1000 models were generated for MtCS, and the final model was selected based on stereochemical quality.

Analysis of the model

The overall stereochemical quality of the final model for the complex MtCS-FMN-EPSP was assessed by the program PROCHECK [11]. Molecular models were superposed using the program LSQKAB from CCP4 [12]. The cutoff for hydrogen bonds and salt bridges were 3.5 Å. The contact surfaces for the binary complexes were calculated using AREAIMOL and RESAREA [12]. The root mean square deviation (rmsd) differences from ideal geometries for bond lengths and bond angles were calculated with X-PLOR [13]. The G-factor is essentially just log-odds score based on the observed distributions of the stereochemical parameters. It is computed for the following properties: torsion angles (the analyses provided the observed distributions of ϕ - ψ , χ_1 - χ_2 , χ_1 , χ_3 , χ_4 , and ω values for each of the 20 amino acid types) and covalent geometry (for the main-chain bond lengths and bond angles). These values' averages were calculated using PROCHECK [14]. The Verify-3D measures the compatibility of a protein model with its sequence, and it was calculated using 3D profile program [14,15].

Circular dichroism

The cloning, protein expression and purification of the MtCS was performed as described in Dias *et al.* [8]. Circular dichroism (CD) was measured over the range 195-260 nm, using a Jasco-710 spectropolarimeter (Jasco, Tokyo, Japan) coupled to a Neslab RTE111 circulating water bath. Spectra were obtained at 25°C using cells with a path length of 0.2 cm. Data points were recorded at a scan speed of 20 nm/min, bandwidth 1.0 nm, 1 s response and 0.1 nm resolution. Five repeat scans were accumulated to obtain the

final averaged spectra. The deconvolution was performed using program of CD analysis by DICROPROT [16].

Results and discussion

Quality of the model

The sequence alignment of MtCS (target) and SpCS (template) shows 44.1% of identity (Figure 1). This degree of primary sequence identity indicates that the crystallographic structure of SpCS is good model to be used as template for modeling of MtCS.

The Ramachandran plot was used to compare the overall stereochemical quality of MtCS model against SpCS structure solved by crystallography. Analysis of the Ramachandran plot of the MtCS model shows that 94.6% of the residues lies in the most favorable regions and the remaining 5.4% in the allowed regions. The same analysis for crystallographic SpCS structure present 93.1% of residues in the most favorable and 6.9% allowed regions. The overall rating for the MtCS model is slightly better than the obtained for the structure of SpCS, however both structures shows excellent stereochemical quality. The r.m.s.d. of molecular modeling MtCS on crystallography SpCS, in the coordinates of $\text{C}\alpha$ is 0.13 Å. The r.m.s.d. value of bond lengths and bond angles, the average G-factor and Verify 3D values are shown in Table 1.

Overall description

The structure of MtCS belongs to the α/β family. The dominant structural topology of CS is a beta-alpha-beta sandwich, in which each monomer of CS consists of a central helical core, helix order $\alpha 1$, $\alpha 6$, $\alpha 12$ and $\alpha 9$, sandwiched between two four-stranded

antiparallel beta sheets, strand order $\beta 1$, $\beta 2$, $\beta 6$ and $\beta 3$ for sheet “one” and strand order $\beta 7$, $\beta 8$, $\beta 14$ and $\beta 9$ for sheet “two” (Figure 2). These layers are packed to form a compact structure. It is clear from the topology of MtCS that the protein core has a pseudo 2-fold symmetry, although the molecule as a whole does not.

The MtCS active site is at the interface between β sheet 2 and 1 end of the internal layer of α helices. At this point, close to the internal pseudo 2-fold axis, the two central helix diverges to leave a small hydrophobic pocket. The other secondary structure elements that participate of active site are the loops L1, L4, L10, L16, L25 and L27. FMN and EPSP are closely associated with each other and a considerable degree of the binding surface of each ligand is in contact with the other.

FMN Binding

The electrostatic potential at the molecular surface generated for program GRASP [17] shows that FMN binding site is highly positive. The six flexible loops are clustered around or near bound cofactor. These flexible regions are rich in strictly or highly conserved residues. Analysis of the molecular surface of the model of MtCS and the structure of SpCS indicates that FMN presents a contact area of the $261.0 \text{ \AA}^2$ and $236.0 \text{ \AA}^2$, respectively.

The molecule of FMN makes few specific polar interactions with the protein, the majority of which are contacts between the hydroxyl and phosphate oxygens of the ribityl chain with loops L10, L16, and L25. The ribityl-phosphate portion of FMN is buried deeply into the enzyme, while the entire *re* face and much of the ortho-xylyx ring of FMN contact the surface of the FMN binding site. These few residues that form the binding surface can provide additional flexibility necessary for the protein to accommodate reduced FMN, in which the isoalloxazine ring system may be bent into a “butterfly”. Furthermore, the lack of

specific contacts between CS and the FMN ribityl chain would also allow some flexibility of the FMN molecule [10]. The figure 3a shows the hydrogen bonds between the molecule de FMN and MtCS.

The FMN molecule seems not cause structural change in the MtCS as well as of the planar oxidized and highly puckered reduced flavins can be accommodated in the same active site without the need of changes in protein conformation [18].

The role of FMN in structure of chorismate synthase seems to be the stabilization of any electron-deficient intermediate of EPSP, indeed, the N5 of FMN is positioned directly below the C6-*pro*-R hydrogen of EPSP and is ideally positioned to remove this proton from the transient intermediate. Therefore, the flavin could be able to donate an electron and subsequently accept a hydrogen atom.

EPSP Binding

There are several polar interactions between EPSP and the MtCS. The EPSP site is an environment extremely basic, with the six arginine (R40, R41, R46, R49, R112, R139) and three histidine (H11, H115, H339) residues concentrated in a small, tightly enclosed binding site in the three corners of the site. The contact area for a molecule of EPSP is $119.0 \text{ \AA}^2$ and $163.0 \text{ \AA}^2$, for molecular model of MtCS and structure of SpMS, respectively. Recently, it was demonstrated by site-directed mutagenesis the role of the H11 and H115 [19]. The H115 serves to protonate the reduced FMN while the H11 protonates the leaving phosphate group of the substrate.

The analysis of the MtCS indicates a total of 11 interactions between the enzyme and EPSP (Figure 3b). The EPSP molecule can be divided in three moieties: carboxyl, enol-pyruvyl, phosphate. The enol-pyruvyl moiety interacts with three arginines residues (R40,

R46 and R139). The phosphate group forms interactions between H11, R49, E81 and R341. The carboxyl group forms hydrogen bonds between H115 and A138.

The molecule of EPSP can be responsible to induce changes in the structure of MtCS, unlike the FMN can be accommodated in the active site without change in protein conformation. The physical studies and spectroscopic in *Escherichia coli* showed that CS undergoes a major structural change when both FMN and EPSP, are bound [20, 21]. The binding of EPSP must cause protein conformational changes involving movements in some of the flexible loops regions resulting in a more apolar environment for the bound co-factor.

The CD spectrum of MtCS in the presence and absence of the FMN is shown in figure 5. The percentage of secondary structure for the two independent measurements and for model can be observed of table 2. The percentage of the secondary structure elements determined by CD is in good agreement with the molecular model of MtCS. These results show that the prediction of secondary structure by homology molecular modeling is reliable. The presence of oxidized FMN seems to cause minor effects on the CD spectrum. These results are like those obtained for chorismate synthase from *Escherichia coli*, where it was used oxidized FMN and EPSP. In the presence of substrate and co-factor the chorismate synthase from *E. coli* was not observed significant variation in the CD spectrum [5].

The EPSP molecule and FMN provide interesting targets for structure-based drug design, not alone against *M. tuberculosis*, but also, for other pathogen, once that specificity and affinity between enzyme and its inhibitors depend on directional hydrogen bonds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners [22-25]. Inhibitors could compete with the substrate and co-factor in the binding to

the enzyme, leaving the enzyme in unproductive conformations. Furthermore, the 5-deazaflavin [26] an flavin analog in which the nitrogen at the 5-position is replaced by carbon, and (6*R*)-6-fluoro-EPSP [27] in which the (6*R*)-hydrogen of EPSP is replaced by fluorine, results a lack of activity of the chorismate synthase [6, 28]. These are examples of molecules that can be rationally modified to optimize the inhibition of the enzyme.

Acknowledgements

This work was supported by grants from FAPESP (SMOLBNet 01/07532-0, 02/04383-7, 04/00217-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT). WFA (CNPq, 300851/98-7), and MSP (CNPq, 300337/2003-5) are researchers for the Brazilian Council for Scientific and Technological Development.

References

- [1] R. Bentley, The shikimate pathway—A metabolic tree with many branches, *Crit. Rev. Biochem. Mol. Biol.* **25** (1990), 307–384.
- [2] F. Roberts, C.W. Roberts, J.J. Johnson, D.E. Kyle, T. Krell, J.R. Coggins, G.H. Coombs, W.K. Milhous, S. Tzipori, D.J. Ferguson, D. Chakrabarti and R. McLeod, Evidence for the shikimate pathway in apicomplexan parasites, *Nature (London)* **395** 6699 (1998), 801–805.
- [3] K.M. Herrmann, L. M. Weaver, The shikimate pathway, *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* **50** (1999), 473–503.
- [4] S. Bornemann, Flavoenzymes that catalyse reactions with no net redox change. *Nat. Prod. Rep.* **19**, (2002), 761–772.

- [5] P. Macheroux, E. Schönbrunn, D.I. Svergun, V.V. Volkov, M.H.J. Koch, S. Bornemann, R.N.F. Thorneley,. Evidence for a major structural change in *Escherichia coli* chorismate synthase induced by flavin and substrate binding. *Biochem. J.* **335** (1998) 319–327.
- [6] P. Macheroux, J. Schmid, N. Amrhein, A. Schaller, A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway. *Planta* **207**, (1999), 325–334.
- [7] C. Ratledge, Nutrition, growth and metabolism, in: C. Ratledge, J. L. Stanford, (Eds.), *The biology of the mycobacteria*, Academic Press, London, 1982, Vol 1, 185-271.
- [8] M. V. B. Dias, F. Ely, F. Canduri, J. H. Pereira, J. Frazzon, L. A. Basso, M. S. Palma, W. F. De Azevedo Jr., D. S. Santos, Crystallization and preliminary X-ray crystallographic analysis of chorismate synthase from *Mycobacterium tuberculosis*. *Acta Crystal. Section D* **60**, (2004), 2003-2005.
- [9] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* **234** (1993), 779–815
- [10] J. Maclean, S. Ali, The structure of chorismate synthase reveals a novel flavin binding site fundamental to a unique chemical reaction, *Structure* **11** (2003), 1499-1511.
- [11] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchinson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thornton, PROCHECK V.3.0 program to check the stereo-chemistry quality of protein structures operating instructions. 1994.
- [12] Collaborative Computational Project No. 4, The CCP4 suite: program for protein crystallography, *Acta Crystallogr. D* **50** (1994) 760–763.

- [13] A.T. Brünger, X-PLOR Version 3.1: A System for Crystallography and NMR, Yale University Press, New Haven (1992).
- [14] J.U. Bowie, R. Luthy, D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure, *Science* **253** (1991) 164–170.
- [15] R. Luthy, J. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature* **356** (1992), 83–85.
- [16] G. Deléage, C. Geourjon. An interactive graphic program for calculating the secondary structures content of proteins from circular dichroism spectrum. *Comp. Appl. Biosc.* **9** (1993), 197-199.
- [17] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins* **11** (1991) (4) 281–296.
- [18] B.W. Lennon, C.H. Williams Jr., M.L. Ludwig, Crystal structure of reduced thioredoxin reductase from *Escherichia coli*: structural flexibility in the isoalloxazine ring of the flavin adenine dinucleotide cofactor. *Protein Sci.* **8** (1999) 2366–2379.
- [19] K. Kitzing, S. Auweter, N. Amrhein, P. Macheroux, Mechanism of chorismate synthase. role of the two invariant histidine residues in the active site. *J. Biol. Chem.* **279** (10), 9451-9461.
- [20] P. Macheroux, J. Petersen, S. Bornemann, D. Lowe, R. N. F. Thorneley, Binding of the oxidized, reduced, and radical flavin species to chorismate synthase. An investigation by spectrophotometry, fluorimetry, and electron paramagnetic resonance and electron nuclear double resonance spectroscopy. *Biochemistry* **35**(5) (1996), 1643 – 1652.

- [21] P. Macheroux, S. Bornemann, S. Ghisla, R. N. F. Thorneley, 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** (1996), 25850-25858.
- [22] W.F. De Azevedo Jr., S. Leclerc, L. Meijer, L. Havlicek, M. Strnad, S.-H. Kim. Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human CDK2 complexed with roscovitine, *Eur. J. Biochem.* **243** (1997) 518–526.
- [23] J.H. Pereira, F. Canduri, J.S. de Oliveira, N.J. da Silveira, L.A. Basso, M.S. Palma, W.F. deAzevedo, Jr., D.S. Santos, Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* **312** 3 (2003), pp. 608–614.
- [24] W.F. De Azevedo, F. Canduri, D.M. dos Santos, J.H. Pereira, M.V.B. Dias, R.G. Silva, M.A. Mendes, L.A. Basso, M.S. Palma, D.S. Santos, Structural basis for inhibition of human PNP by immucillin-H. *Biochem. Biophys. Res. Commun.* **309** (2003), pp. 922–927.
- [25] W.F. De Azevedo, Jr., H.J. Mueller-Dieckmann, U. Schulze-Gahmen, P.J. Worland, E. Sausville and S.-H. Kim , Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. *Proc. Natl. Acad. Sci. USA* **93** 7 (1996), 2735–2740.
- [26] C. T. Lauhon, P. A. Bartlett, Substrate analogs as mechanistic probes for the bifunctional chorismate synthase from *Neurospora crassa*. *Biochemistry* **33**(47) (1994) 14100-14108.
- [27] J. K. Sutherland, W. J. Watkins, J. P. Bailey, A. K. Chapman, G. M. Davies, The synthesis of 6a- and 6b-fluoroshikimic acids. *J. Chem. Soc.Chem. Commun.* (1989) 1386–1387.



[28] S. Bornemann, D.M. Lawson, R.N. Thorneley, A branch point in chorismate synthase research. *Structure* **11(12)** (2003) 1463-1465.

[29] R. Koradi, M. Billeter and K. Wüthrich, MOLMOL: A program for display and analysis of macromolecular structures. *J. Molecular Graphics*, **14(1)** (1996), 51-55.

[30] J.D. Thompson, D.G. Higgins, T.J. Gibson, 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** (1994), 4673-4680.

Table 1. Summary of structural results for MtCS and SpCS.

Analysis of the model of MtCS and structure of SpCS								
Proteins	3D Profile ^a			G-factor ^b			rmsd from ideal geometry	
	Total Score	Ideal Score	S _{ideal} Score	Torsion Angles	Covalent Geometry	Global	Bond Lengths (Å)	Bond Angles (°)
MtCS	157.49	179.29	0.88 S	-0.07	-0.34	-0.16	0.022	3.172
SpCS	163.85	171.92	0.95 S	0.00	-0.16	-0.03	0.027	2.558

^a Total Score: is the sum of the 3D-1D scores (statistical preferences) of each residue present in protein. Ideal Score: $S_{ideal} = \exp(-0.83+1.008 \times \ln(L))$; where L is number of amino acids. S_{ideal} Score: is compatibility of the sequence with their 3D structure. It is obtained Total Score / Ideal Score. S_{ideal} Score above $0.45S_{ideal}$.

^b Ideally, scores should be above - 0.5. Values below -1.0 may need investigation.

Table 2 – Percentage of the secondary structure elements for experimental data of CD and for MtCS molecular model.

	Molecular Model of MtCS ^a	MtCS with FMN ^b	MtCS without FMN ^b
β-sheet	18	18	16
α-helix	40	44	45
Coil	42	38	39

^a Secondary structure elements were calculated using the program MOLMOL [29]

^b Secondary structure elements were calculated using the program of CD analysis DICROPROT [16]

Figure 2. Ribbon diagram of the MtCS model completed with PDB and DSSP generated secondary structure.

Figure 1. The sequence alignment of MtCS and SpCS indicating the secondary structural elements. The sequence MtCS shows 44.1% of identity with the sequence of SpCS. The alignment was performed with the program CLUSTAL W [30].

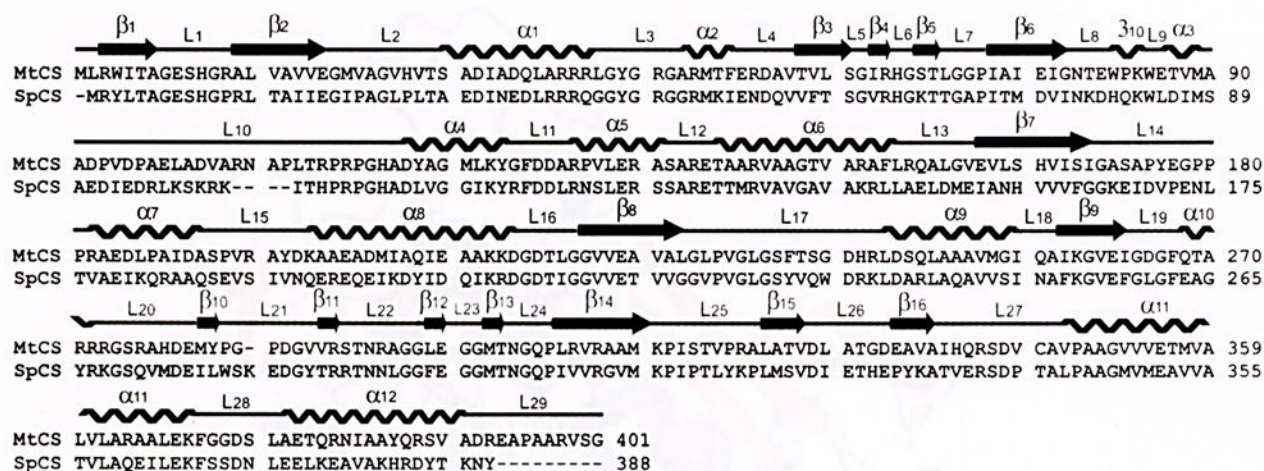


Figure 2. Ribbon diagram of the MtCS model complexed with FMN and EPSP generated by MolMol [29].

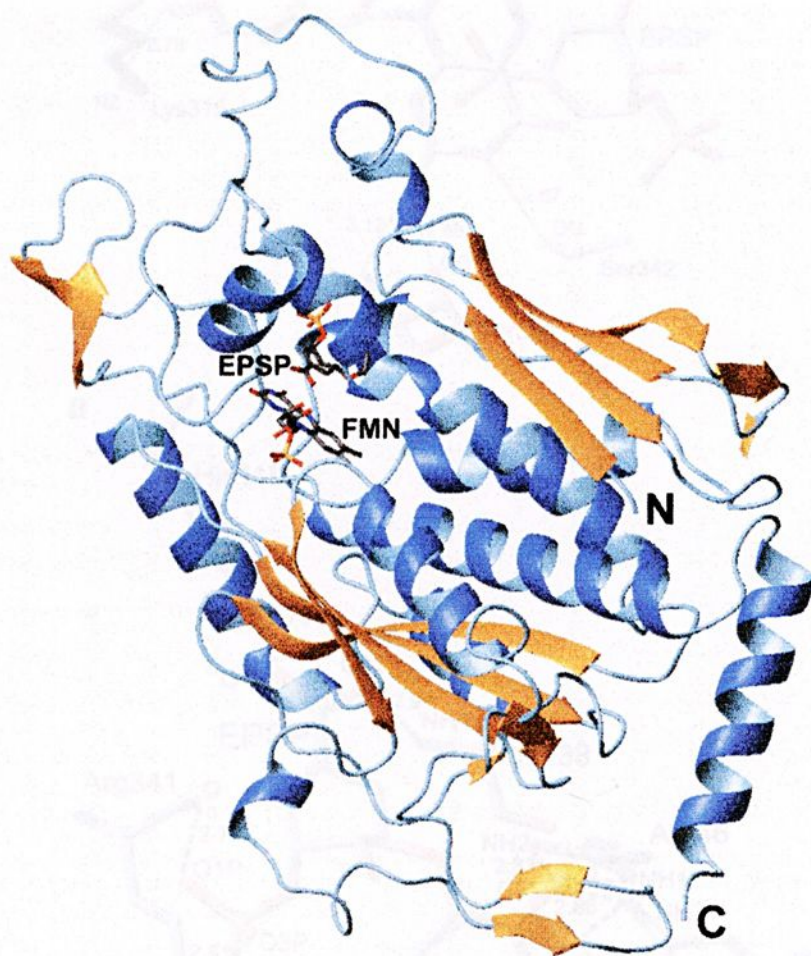


Figure 3. Diagrams showing the active sites of the MtCS; (A) active site of the FMN and (B) active site of the EPSP.

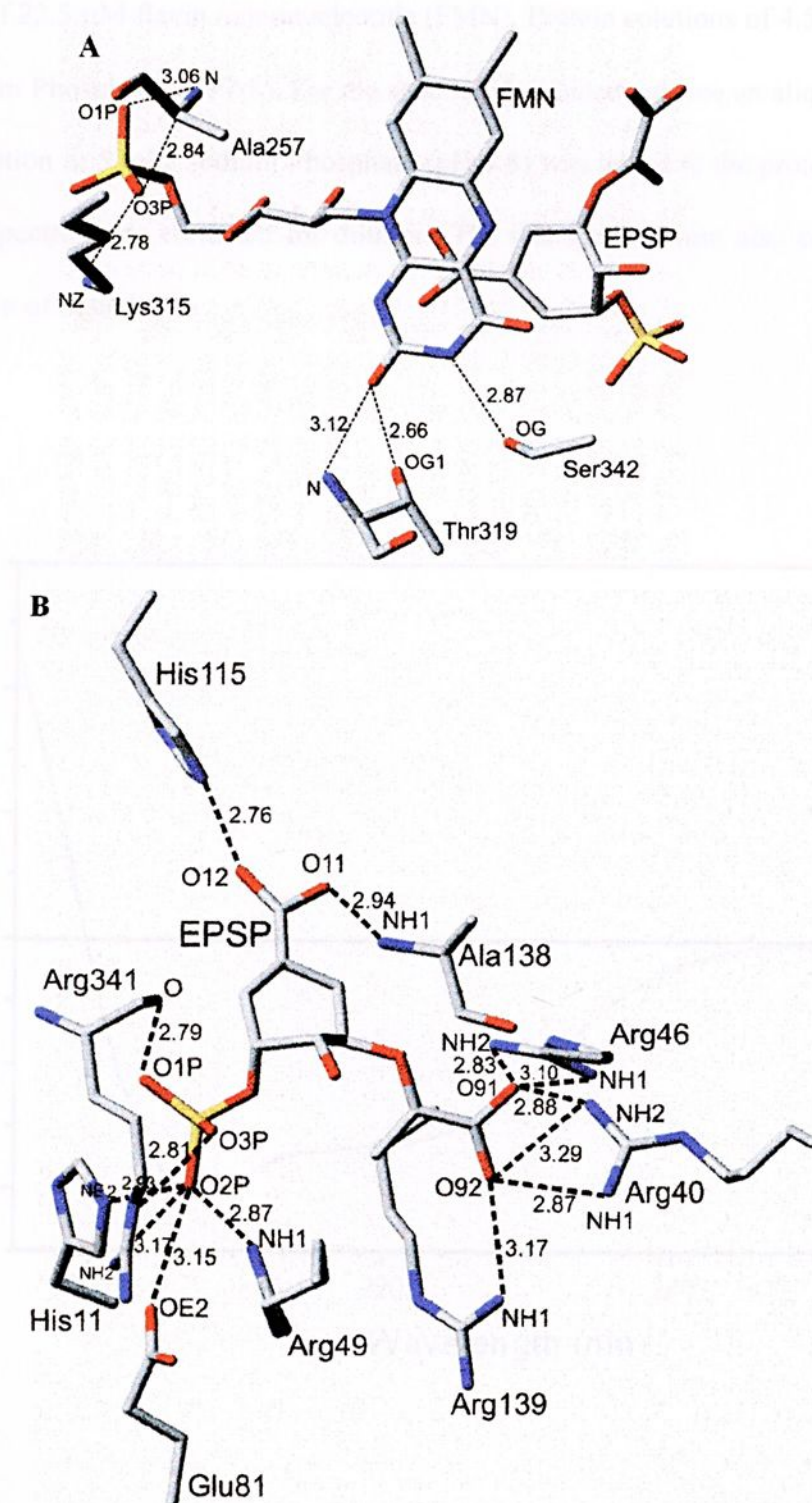


Figure 4 – Superimposed circular dichroism spectra of unliganded and liganded chorismate synthase from *Mycobacterium tuberculosis*. (—) Unliganded enzyme; (----) enzyme in the presence of 22.5 μM flavin mononucleotide (FMN). Protein solutions of 4.5 μM were in 5 mM Sodium Phosphate (pH 7.8). For the spectra of liganded enzyme an aliquot of a 10 mM ligand solution in 5 mM Sodium Phosphate (pH 7.8) was added to the protein solution and resulting spectra were corrected for dilution. The CD spectra were also corrected for the contribution of ligand.

