ACADEMIC PRESS

Biochemical and Biophysical Research Communications 295 (2002) 142–148

BBRC

www.academicpress.com

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.

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.).

0006-291X/02/$ - see front matter
PII: S0006-291X(02)00632-0

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* shikiamte 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/Mg^{2+} we superposed the atomic coordinates of the ADP/Mg^{2+} 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

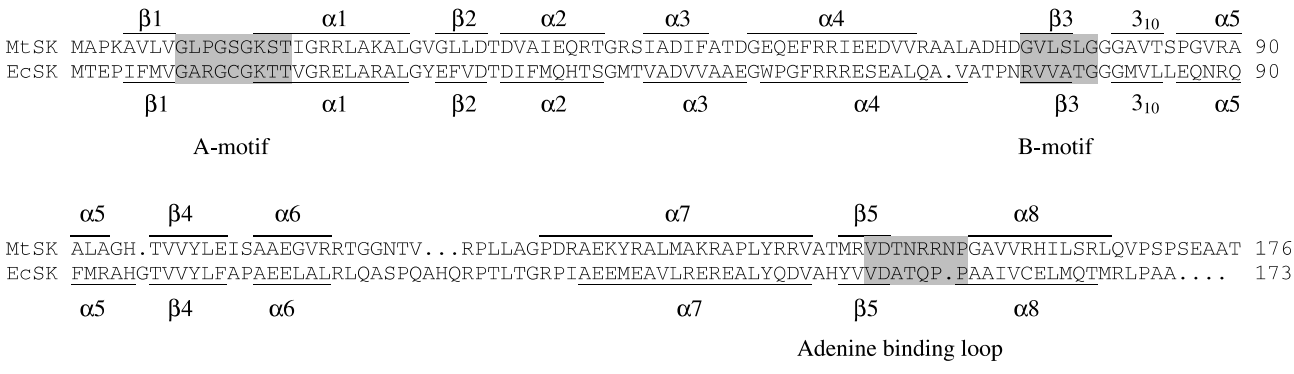

Fig. 1. The sequence alignment of ecSK and mtSK indicating the secondary structural elements. The sequence mtSK shows 34% of identity with the sequence of ecSK. The alignment was performed with the program CLUSTAL V [31].

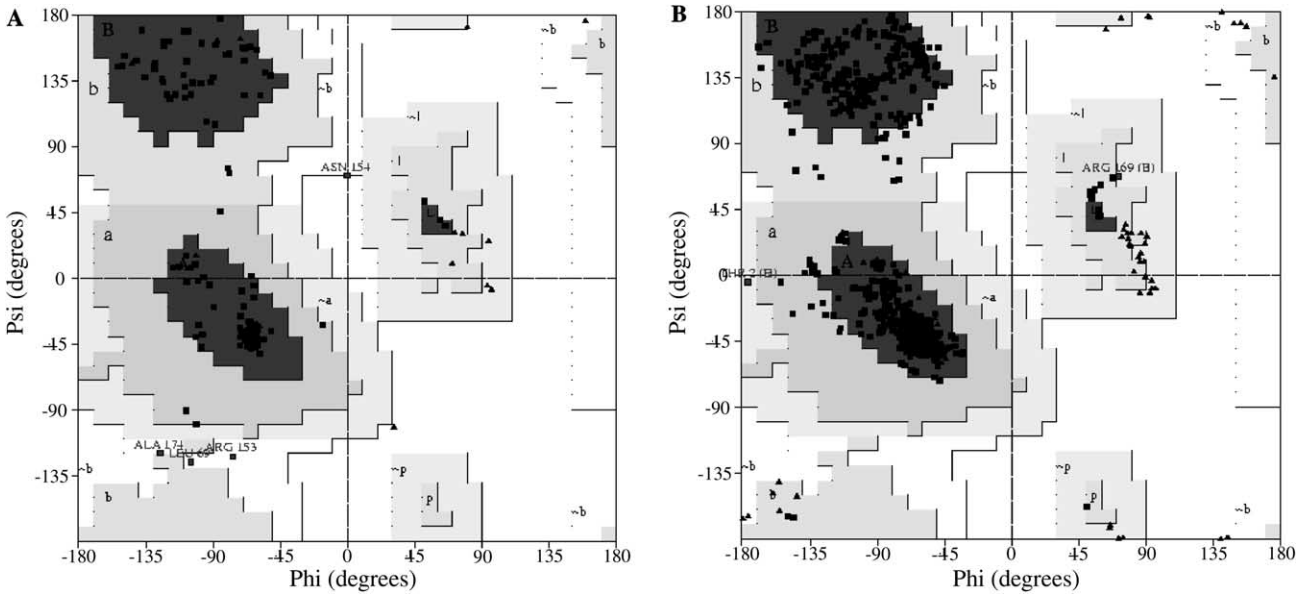

Fig. 2. (A) Ramachandran diagram ϕ–ψ plots for the mtSK structure and (B) for three crystallographic SK structures solved to resolution better than 2.6 Å.

remaining 8.9% in the additional allowed regions. The same analysis for three crystallographic ecSK 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 (β1–β5) presents the strand order 23145. The β-strands are flanked on either side by α helices (α1 and α8 on one side, α4, α5, and α7 on the other). Fig. 3 shows a schematic diagram of the mtSK structure, with shikimate and ADP/Mg^{2+} 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 β1 and α1 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 confomation 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 ecSK structure [8].

ADP/Mg^{2+}-binding site

The molecular model for ternary complex mtSK–shikimate–ADP/Mg^{2+} indicates that ADP/Mg^{2+} is

Fig. 3. Ribbon diagram of the mtSK structure with shikimate and ADP/Mg^{2+} 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/Mg^{2+}. 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 β2, consists of helices α2 and α3 and the N-terminal region of helix α4. 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

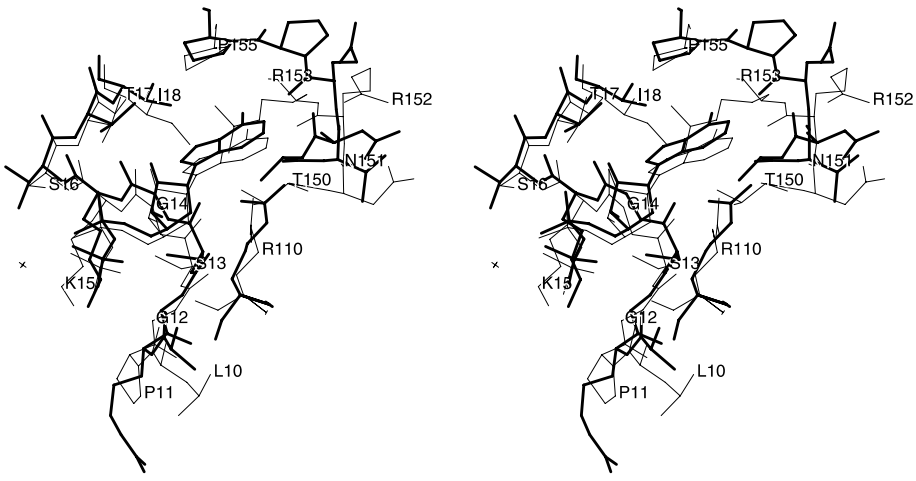

Fig. 4. Superimposed binding pockets of the ATP-binding site of mtSK (thick line) and ecSK (thin line).

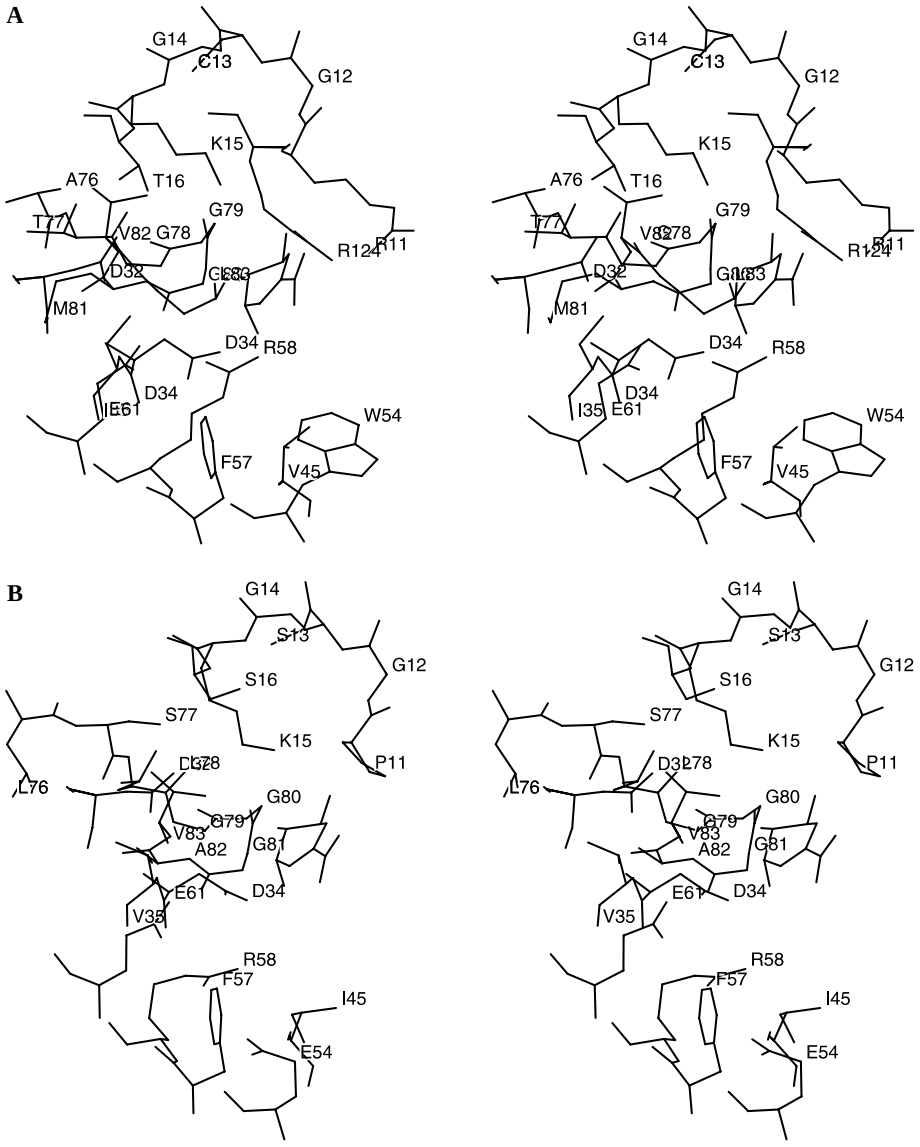

Fig. 5. Main residues involved in contact with shikimate in the complexes (A) mtSK:shikimate and (B) ecSK:shikimate.

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 is 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

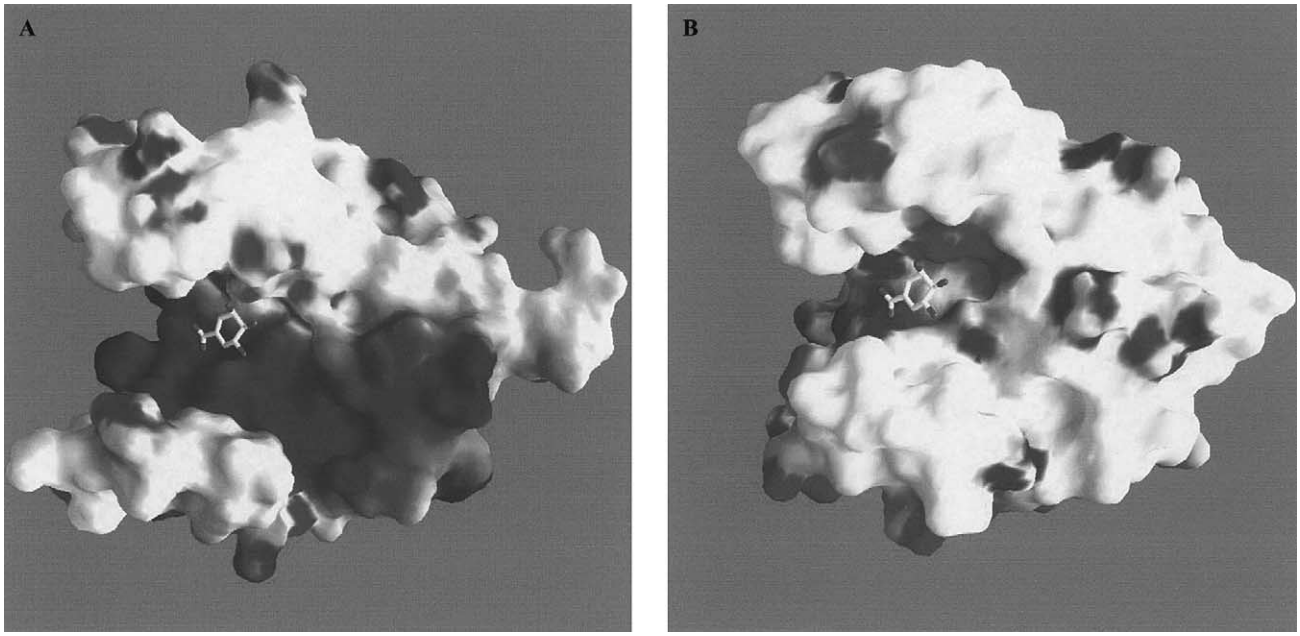

Fig. 6. (A) Molecular surfaces for mtSK and (B) ecSK complexed with shikimate generated with GRASP [24].

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. Barrel, 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. Thorton, 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. Schlauderer, 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 Silveira, 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.