Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

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

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]*

[a]Programa de Pós-Graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, [b]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, [c]Center for Applied Toxinology, Institute Butantan, São Paulo, SP 05503-900, Brazil, [d]Laboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil, and [e]Centro 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@pucrs.br, walterfa@df.ibilce.unesp.br

Received 24 August 2004
Accepted 5 October 2004

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

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.

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 (Vonrhein *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 (β1) and first α-helix (α1), containing the conserved sequence G*XXXX*GKT/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 *ZZDXX*G 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 *ZZZ*TGGG 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 I/VD*XXX*(*X*)*X*P 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 m*M* Na HEPES buffer pH 7.5 containing 0.5 *M* NaCl and 5.0 m*M* $MgCl_2$. This protein solution was brought to 13.0 m*M* 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 α = 54.85, β = 85.32, γ = 91.90°. 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 $F_{obs} - F_{calc}$ 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 *CCP*4 (Collaborative Computational Project, Number 4, 1994).

The molecular-surface areas were calculated using the program *Swiss PDB Viewer* v.3.7 (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 m*M* $MgCl_2$, 13.0 m*M* ADP and 13.0 m*M* 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 (NQIIHMLESN) 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 Å^2, whereas that for side-chain atoms is 35.68 Å^2 (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 = $100 \times \sum |F_{obs} - F_{calc}| / \sum F_{obs}$, the sums being taken over all reflections with $F/\sigma(F) > 2\sigma(F)$. ‡ 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 Å^2 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

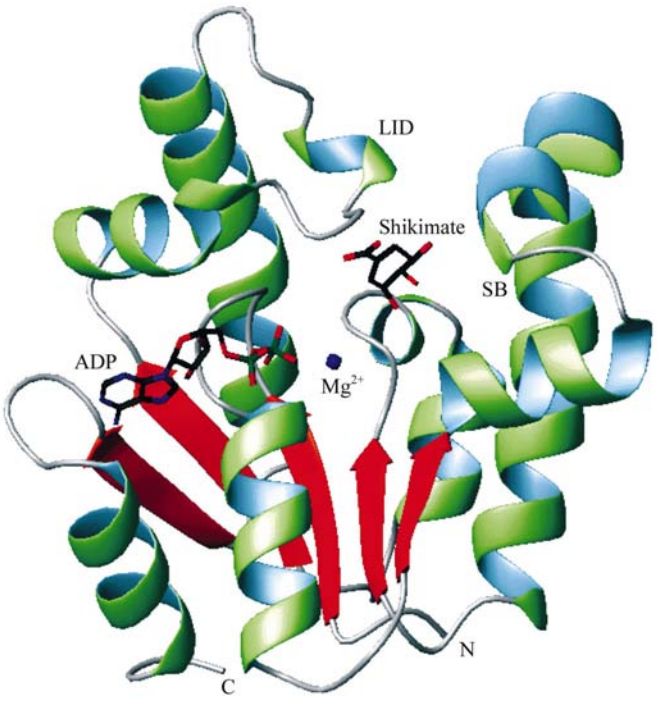

Figure 1
MtSK complexed with Mg^{2+}, ADP and shikimate. The LID (residues 112–124) and SB (residues 33–61) domains are responsible for large conformational changes during catalysis and are shown. The figure was prepared with *MOLMOL* (Koradi *et al.*, 1996).

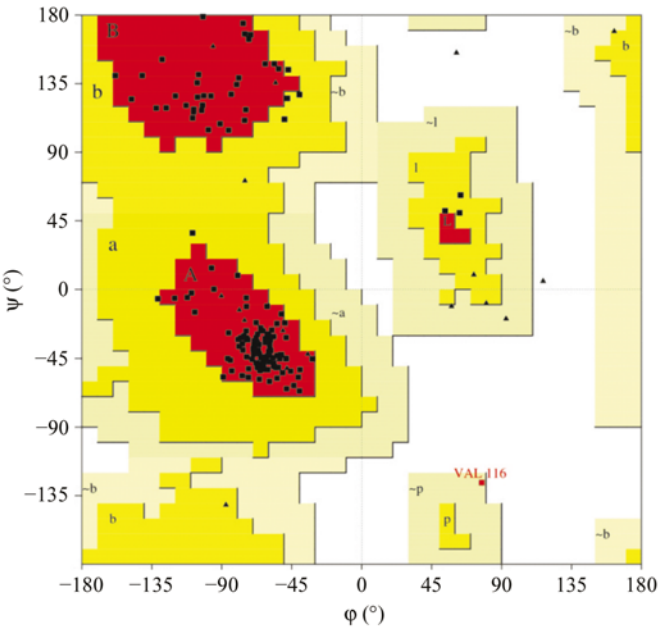

Figure 2
Ramachandran diagram for MtSK–MgADP–shikimate. 93.4% of non-glycine residues lie in the most favoured regions and no residues are in the disallowed region.

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/Mg^{2+}

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 R*XX*(*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/VD*XXX*(*X*)*X*P sequence stretch (Gu *et al.*, 2002). This motif forms a loop that wraps around the adenine moiety of ATP, connecting the β5-strand with the C-terminal α8-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

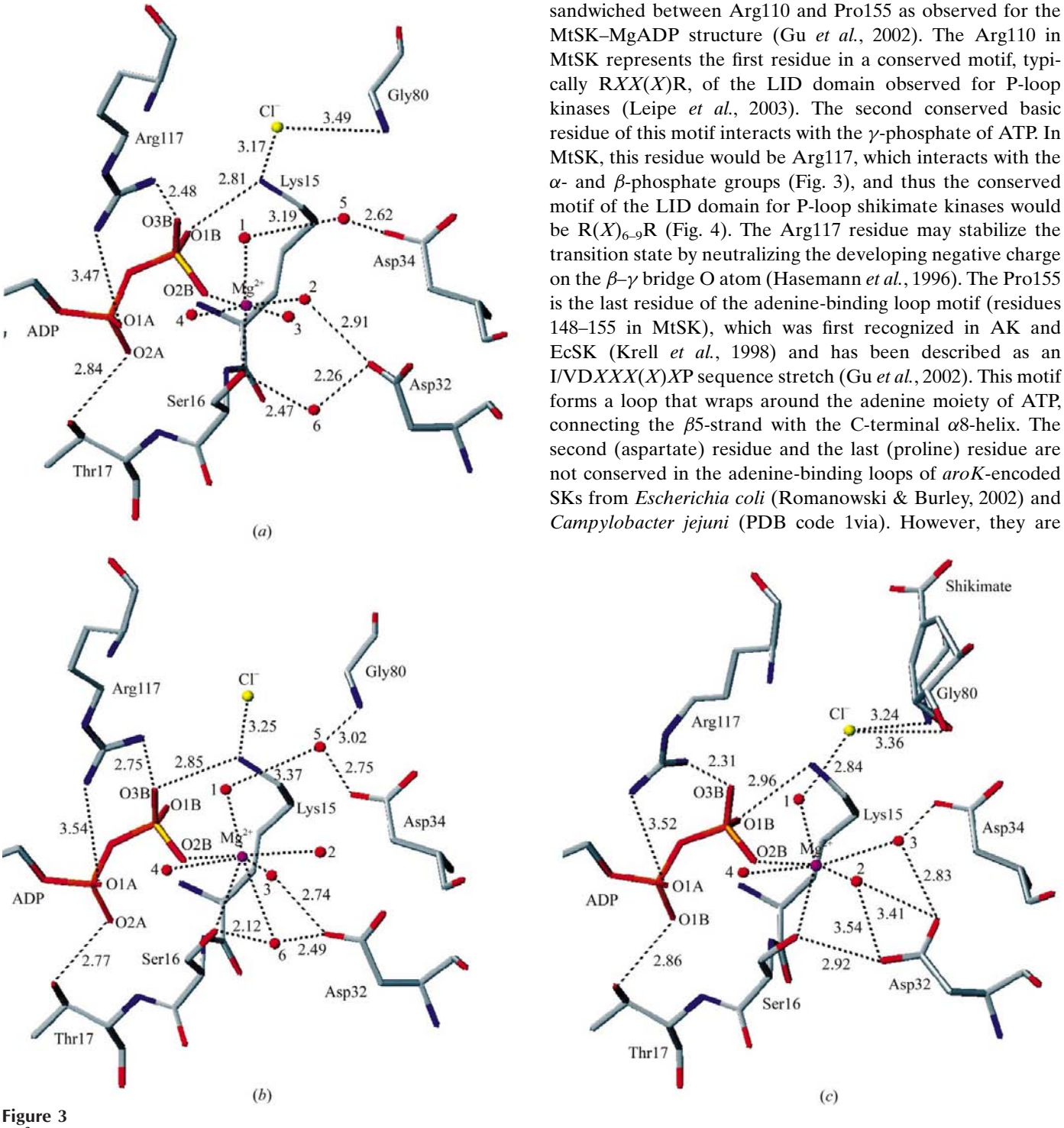

Figure 3
Mg^{2+} coordination. (*a*) 1l4y, (*b*) 1l4u (Gu *et al.*, 2002), (*c*) MtSK–MgADP–shikimate. For simplicity, only protein residues, atoms involved in binding of Mg^{2+} water and ADP are shown. Broken black lines represent hydrogen bonds or Mg^{2+} coordination. The distances are in Å. The 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*)*X*P 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 Mg^{2+}. In the MtSK–MgADP structure (PDB code 1l4y) a typical six-coordination is observed for Mg^{2+} (Fig. 3*a*), 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. 3*b*). The structure of MtSK–MgADP–shikimate shows a distorted six-coordinated Mg^{2+} (Fig. 3*c*). Binding distances to Mg^{2+} are shown in Table 3.

The Mg^{2+} 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 Mg^{2+} 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 Mg^{2+} 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 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 Mg^{2+}-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 Mg^{2+} 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 Mg^{2+}. Interestingly, the two functions of the aspartate of the *ZZDXX*G Walker B motif (hydrogen bonding to Mg^{2+} 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 β2 (Asp32 in MtSK) and the first position of helix α2 (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 *ZZZ*TGGG (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

Table 3
Binding of Mg^{2+} 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

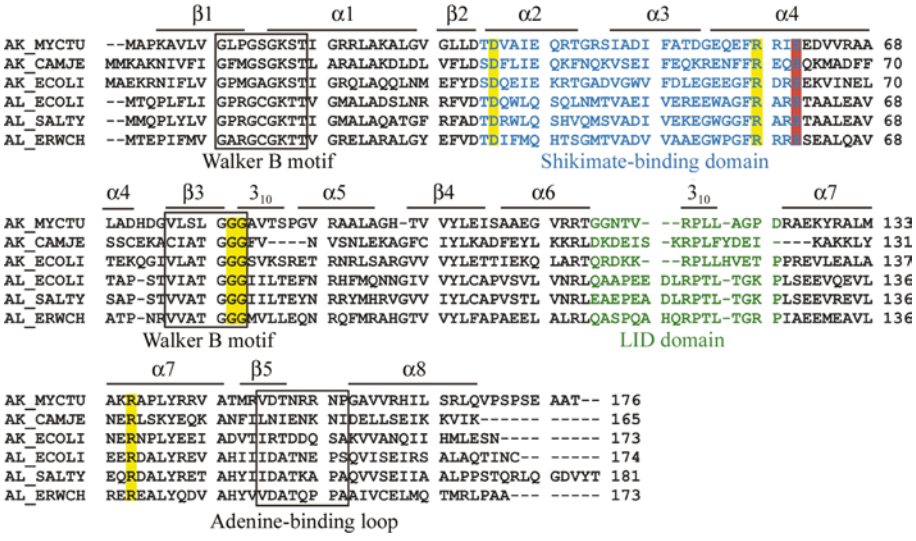

Figure 4
Alignment of shikimate kinases showing that the residues identified in binding of shikimate are conserved. The secondary-structure assignment for MtSK–shikimate is shown above the sequence and the locations of linear motifs are labelled. Shaded regions in yellow are the residues involved in 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*.

Table 4
Direct and water-mediated hydrogen bonding of shikimate in MtSK.

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		

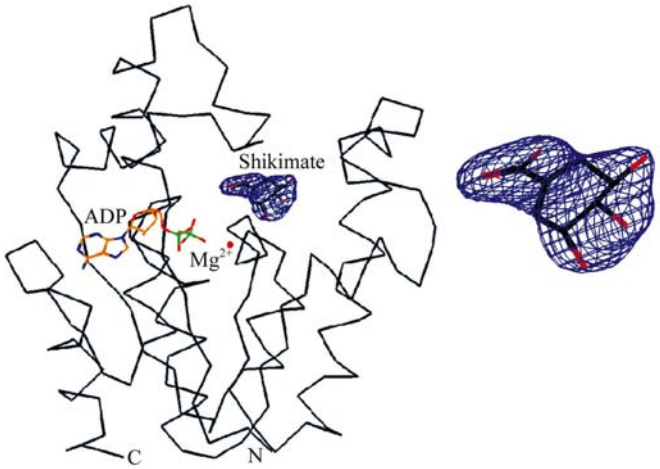

Figure 5
$F_{obs} - F_{calc}$ electron density contoured at 3.0σ showing the binding of shikimate. The figure was prepared with *XtalView* (McRee, 1999).

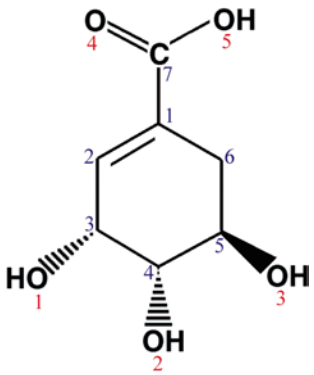

Figure 6
The molecular structure of shikimate shows the carboxyl group and three hydroxyl groups. C atoms are numbered in blue and O atoms in red.

(Mg^{2+}) to the β- and γ-phosphate groups, whereas departure of the leaving group will be most favoured in a structure with metal binding to the α- and β-phosphate groups (Jencks, 1975*a*). In enzyme reactions, the enzyme may facilitate the reaction and contribute to its specificity simply by favouring the binding of metal in the most favourable possible structure for the particular reaction that is being catalyzed. Although the mechanism of action of MtSK is still unknown, the interaction between Gly80 and the chloride ion may indicate that the latter occupies the γ-phosphate position as the chemical reaction proceeds, thereby suggesting that the interaction between Gly80 and γ-phosphate of ATP may play a role in the catalytic mechanism of MtSK, as discussed below.

3.2. Shikimate binding

The shikimate-binding domain, which follows strand β2, consists of helices α2 and α3 and the N-terminal region of helix α4 (residues 33–61). A peak of more than 3σ in the final $F_{obs} - F_{calc}$ difference Fourier map clearly indicates the position of bound shikimate in the electron density (Fig. 5).

The chemical structure of shikimic acid [3*R*-(3α,4α,5β)]-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.

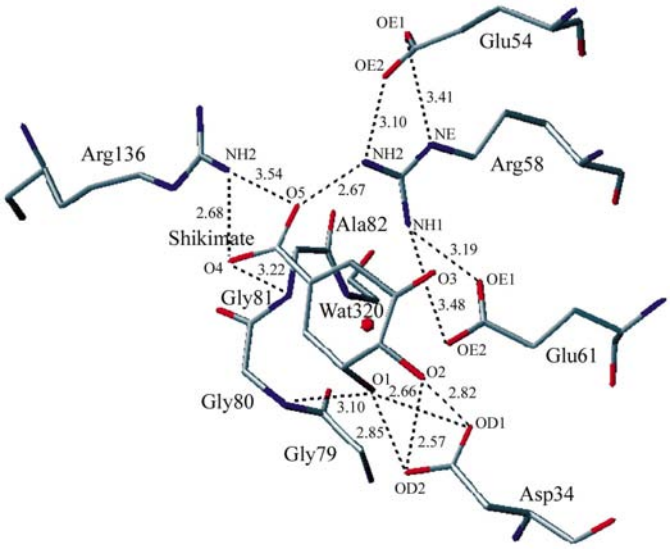

Figure 7
Polar interactions involved in the binding of shikimate to the MtSK active site. Hydrogen bonds between shikimate and protein groups and interactions between protein groups are represented as broken black lines. The interaction distances are in Å.

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 (1l4u and 1l4y; 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 (1l4y) 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)		
		1l4y	1l4u	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 (1l4y and 1l4u) 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

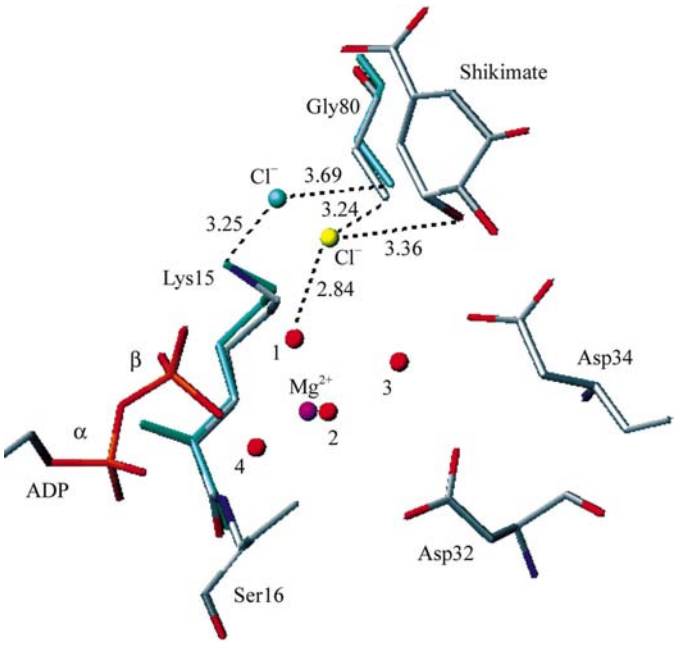

Figure 8
Chloride ions bound to the active sites of the MtSK–MgADP–shikimate and MtSK–MgADP (1l4u; Gu *et al.*, 2002) structures. For clarity, only residues Lys15, Ser16, Asp32, Asp34 and Gly80, shikimate, the α- and β-phosphates of ADP, Mg^{2+} and Cl^- ions and Mg^{2+}-coordinated waters are shown. 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 1l4u are coloured cyan.

1.52 Å in MtSK–MgADP–shikimate compared with the MtSK–MgADP 1l4u and 1l4y 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 (Vonrhein *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 1l4u and 1l4y, 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 α7 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, Mg^{2+}, shikimate, Cl^{-} ions and water molecules were removed prior to calculation. The calculated values are approximately 7246 $Å^2$ for MtSK complexed with MgADP (1l4u) and 6915 $Å^2$ for MtSK complexed with MgADP and shikimate. Thus, approximately 330 $Å^2$ 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.

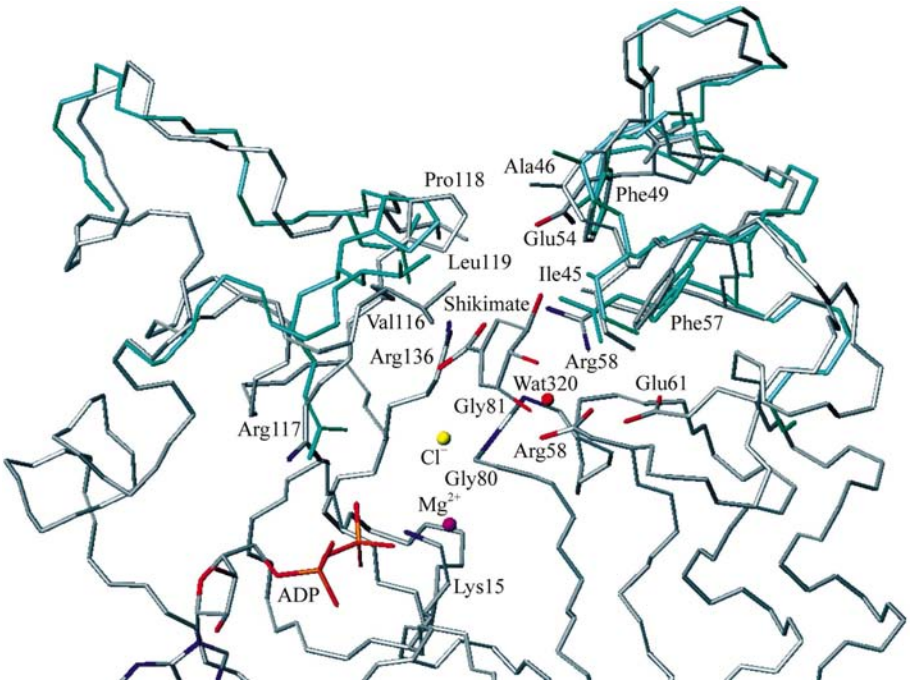

Figure 9
Shikimate binding in the MtSK–MgADP–shikimate structure. The C^{α} traces of MtSK–MgADP–shikimate (grey) and MtSK–MgADP (cyan) (1l4u; Gu *et al.*, 2002) were superimposed. The side chains with large shifts owing to shikimate binding are shown for both structures: Val116, Pro118 and Leu119 from LID domain and Ile45, Ala46, Glu54, Phe57 and Arg58 from the SB domain. For clarity, only residues from the LID and SB domains of 1l4u (cyan) are shown.

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 m*M*) decreased to 0.3 m*M* in the presence of 1.5 m*M* 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

and ATP to be 1.7 and 2.6 m*M*, 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, 1975*a*). 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.

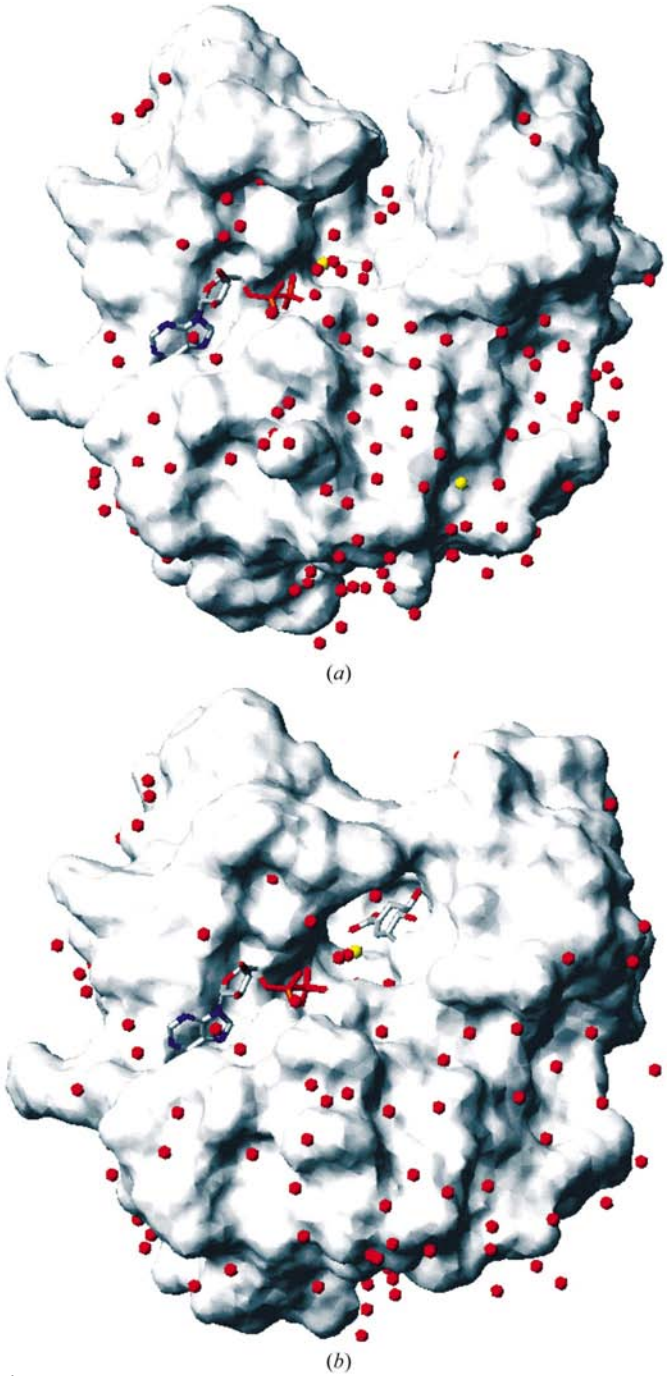

Figure 10
Molecular surface of (*a*) MtSK–MgADP (1l4u; Gu *et al.*, 2002) and (*b*) MtSK–MgADP–shikimate structures. Chloride ion is coloured yellow in MtSK–MgADP and MtSK–MgADP–shikimate. For water molecules only O atoms are shown.

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 Romanoswski 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.

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, Mg^{2+} 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.* D**50**, 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.* D**57**, 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. (1975*a*). In *Catalysis in Chemistry and Enzymology*. New York: Dover.

Jencks, W. P. (1975*b*). *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* v.3.0. *Program to Check the Stereochemistry Quality of Protein Structures. Operating Instructions.* http://www.ccp14.ac.uk/ccp/ccp14/ftp-mirror/llnlrupp/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.* A**50**, 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., Schlauderer, 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.