Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Substrate analogs induce an intermediate conformational change in *Toxoplasma gondii* adenosine kinase

Yan Zhang,[a] Mahmoud H. el Kouni[b] and Steven E. Ealick[c]*

[a]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA, [b]Department of Pharmacology and Toxicology, Center for AIDS Research, Comprehensive Cancer, University of Alabama at Birmingham, Birmingham, AL 35294, USA, and [c]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA

Correspondence e-mail: see3@cornell.edu

Received 29 September 2006
Accepted 19 October 2006

PDB Reference: adenosine kinase complex, 2a9y, r2a9ysf.

Adenosine kinase (AK) is a key enzyme in purine metabolism in the ubiquitous intracellular parasite *Toxoplasma gondii* and is a potential chemotherapeutic target for the treatment of *T. gondii* infections. To better understand the structure–activity relationship of 6-substituted purine ribosides, the structures of the *T. gondii* AK–N^6,N^6-dimethyladenosine (DMA) complex, the AK–DMA–AMP-PCP complex, the AK–6-methyl mercaptopurine riboside (MMPR) complex and the AK–MMPR–AMP-PCP complex were determined to 1.35, 1.35, 1.75 and 1.75 Å resolution, respectively. These structures reveal a conformation intermediate between open and closed, with a small lid-domain rotation of 12°. Residues Gly143-*X*-*X*-Gly146 undergo torsional changes upon substrate binding, which together with a Gly68-Gly69 switch induces a hinge bending of the lid domain. The intermediate conformation suggests that ATP binding is independent of adenosine binding. Orienting the γ-phosphate group of ATP into the optimal catalytic position may be the last step before the onset of chemical catalysis and may require the translocation of Arg136 following the complete closure of the lid domain. 6-Substituted purine-nucleoside analogs are accommodated in a hydrophobic cavity. Modification at the N^6 or C6 position of the nucleoside would affect the interactions with the surrounding residues and the binding affinity.

1. Introduction

Toxoplasma gondii is an intracellular protozoan parasite of warm-blooded animals, including humans. Close to one quarter of the US population over the age of 12 are infected by *T. gondii* (Jones *et al.*, 2001), which has an estimated impact of $7.7 billion on the economy (Buzby & Roberts, 1996). *T. gondii* infections are usually asymptomatic in adults with normal immune functions; however, congenital *T. gondii* infections may cause mental retardation, epilepsy and death (Daffos *et al.*, 1994; McAuley *et al.*, 1994). *T. gondii* can also cause severe and fatal encephalitis or disseminated toxoplasmosis among immune-suppressed persons (Leport *et al.*, 1996), including those with acquired immunodeficiency syndrome (AIDS; Luft & Remington, 1992; Mariuz *et al.*, 1994) and transplant recipients on immunosuppressive drugs (Couvreur *et al.*, 1992). Toxoplasmosis is the most frequent severe neurological infection among persons with AIDS (Jones *et al.*, 1999). Although the current widely used drugs for *T. gondii* infections that function by targeting folate metabolism are effective, they can induce severe side effects such as bone-marrow suppression and severe skin rashes (Dannemann *et al.*, 1992; Leport *et al.*, 1987). Therefore, the need to develop potent and less toxic chemotherapeutics against *T. gondii* is urgent.

Rational drug design is based on exploitation of fundamental biochemical and metabolic differences between pathogen and host. Like all the protozoan parasites, *T. gondii* lacks a *de novo* purine-biosynthesis pathway and must rely on purine salvage from the host cells to meet its purine requirements (el Kouni, 2003; Krug *et al.*, 1989). For nucleic acid biosynthesis, *T. gondii* must utilize salvaged adenine, adenosine, hypoxanthine and inosine because the parasites are capable of converting adenine to guanine nucleotides, but not the reverse (Krug *et al.*, 1989; Pfefferkorn & Pfefferkorn, 1978). In addition, adenosine is the preferred substrate among all the purine precursors and is incorporated into the nucleotide at a tenfold higher rate than the others (Krug *et al.*, 1989). The most important and the most active enzyme in adenosine metabolism in *T. gondii* is adenosine kinase (AK; EC 2.7.1.20). AK catalyzes the phosphorylation of adenosine to AMP, from which all other purine nucleotides can be synthesized. In mammalian cells, adenosine is predominantly subsequently catabolized to hypoxanthine by adenosine deaminase (EC 3.5.4.4) and purine nucleoside phosphorylase (EC 3.5.4.4; el Kouni, 2003). Neither of these two enzymes has significant activity in *T. gondii* (Krug *et al.*, 1989). The high activity and the importance of AK in adenosine metabolism in *T. gondii* make it an attractive chemotherapeutic target; however, deficiency of adenosine kinase was shown not to be lethal to the parasites (Pfefferkorn & Pfefferkorn, 1978; Sullivan *et al.*, 1999), indicating that inhibition of the enzyme will not lead to toxicity in toxoplasma. Therefore, our search for potential antitoxoplasmic drugs has focused on identifying subversive substrates that are specifically metabolized into a cytotoxic nucleotide by the AK of *T. gondii*, but not by that of humans. Intensive studies have been carried out to identify such potential subversive substrates, including 6-, 7- and N^6-substituted purine ribosides (Iltzsch *et al.*, 1995; Rais *et al.*, 2005; Yadav *et al.*, 2004).

T. gondii AK is a monomeric protein with 363 residues (Sullivan *et al.*, 1999). The structures of *T. gondii* AK in the apo form, in complex with adenosine and β,γ-methyleneadenosine 5′-triphosphate (AMP-PCP, a nonhydrolysable ATP analog) and in complex with 7-iodotubercidin and AMP-PCP have been determined (Schumacher *et al.*, 2000). Based on the sequence and structural homology, *T. gondii* AK is classified as a member of the ribokinase superfamily (Bork *et al.*, 1993; Schumacher *et al.*, 2000). The ribokinase superfamily was first identified based on sequence homology and was later expanded by the addition of some members sharing structural homology but lacking primary sequence homology. The structures of more than ten members of this superfamily have been solved, including those of *Escherichia coli* ribokinase (Sigrell *et al.*, 1998), human adenosine kinase (Mathews *et al.*, 1998), *Bacillus subtilis* 4-methyl-5-β-hydroxy-ethylthiazole kinase (Campobasso *et al.*, 2000), *Thermococcus litoralis* and *Pyrococcus furiosus* ADP-dependent glucokinase (Ito *et al.*, 2001, 2003), *Salmonella typhimurium* 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase (Cheng *et al.*, 2002), sheep pyridoxal kinase (Li *et al.*, 2002), *S. enterica* aminoimidazole riboside kinase (Zhang *et al.*, 2004), *Thermus thermophilus* 2-keto-3-deoxygluconate kinase (Ohshima *et al.*, 2004), *B. subtilis* YXKO protein (Zhang *et al.*, 2002) and human pyridoxal kinase (Cao *et al.*, 2006). The common structural feature of the enzymes in this superfamily is a central eight-stranded β-sheet that is flanked by eight structurally conserved α-helices, five on one side and three on the other. The active site is located in a shallow groove along one edge of the β-sheet, with the phosphate-acceptor hydroxyl group and the γ-phosphate of ATP close together in the middle of the groove and the substrate-binding and the ATP-binding sites at the ends. In addition to the common core structure, the structures of *T. gondii* AK reveal a novel catalytic mechanism. A conserved dipeptide Gly68-Gly69 undergoes a conformational change upon adenosine binding, resulting in a domain rotation of 30°. Additional local conformational changes induced by ATP binding complete the structural requirements for the catalysis (Schumacher *et al.*, 2000).

In order to better understand the structure–activity relationship of C6- and N^6-substituted purine-nucleoside analogs, we determined the structures of the *T. gondii* AK–N^6,N^6-dimethyladenosine (DMA) complex to 1.35 Å, the AK–DMA–AMP-PCP complex to 1.35 Å, the AK–6-methyl mercaptopurine riboside (MMPR) complex to 1.75 Å and the AK–MMPR–AMP-PCP complex to 1.75 Å. Surprisingly, all the structures reveal an intermediate conformational change with a domain rotation of 12°. The identification of key residues suggests the importance of the conformational changes to catalysis. This provides a structural basis for the design of effective and selective subversive substrates of *T. gondii* AK.

2. Experimental procedures

2.1. Protein expression, purification and crystallization

The AK gene was excised between the *Nde*I and *Sac*I sites of the previously described overexpression plasmid pET21a (Recacha *et al.*, 2000) and spliced into vector pET28a, which carries a sequence encoding an N-terminal polyhistidine tag. Both vectors were sequenced at the Biotechnology Resource Center at Cornell and showed the same errors. The mutation of Gly270 to a serine arises from one base change. The C-terminus is changed from FTSLPC to FTFTSG owing to several changes (single-base insertions and deletions) at the 3′ end of the gene, which alter the position of the stop codon. However, these changes in primary sequence do not alter the enzyme activity (data not shown). The plasmid was then transformed into *E. coli* strain BL21(DE3) (Novagen). The transformed cells were grown in 20 ml LB containing 50 µg ml^{-1} kanamycin at 310 K until the OD_{600} reached 0.8. This culture was inoculated into 1 l fresh LB containing the same antibiotic. The resulting culture was grown at 310 K until the OD_{600} reached 0.6. Expression was induced by addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 1 m*M* at 298 K. The cells were harvested after 20 h by centrifugation at 3000*g* for 20 min at 277 K and stored at 193 K.

Table 1
Summary of data-collection and processing statistics.

Values in parentheses are for the outer resolution shell.

	AK–MMPR	AK–MMPR–AMP-PCP	AK–DMA	AK–DMA–AMP-PCP
Wavelength (Å)	0.9790	1.5418	0.9795	0.9795
Resolution (Å)	1.75	1.75	1.35	1.35
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters				
a (Å)	60.35	60.49	60.29	60.07
b (Å)	61.70	61.49	61.52	61.59
c (Å)	91.96	91.60	91.83	91.65
No. of reflections	113729	135552	333095	391727
No. of unique reflections	32789	34865	72440	74024
Redundancy	3.5 (3.6)	3.9 (3.8)	4.6 (2.9)	5.3 (2.6)
Completeness (%)	92.6 (99.7)	99.1 (99.9)	95.5 (73.2)	98.3 (86.0)
R_{sym}† (%)	9.7 (41.3)	4.3 (28.6)	6.0 (37.8)	4.4 (30.3)
$I/\sigma(I)$	12.4 (4.2)	29.5 (4.7)	22.7 (2.1)	35.5 (2.8)

† $R_{sym} = \sum\sum_i |I_i - \langle I\rangle| / \sum\langle I\rangle$, where $\langle I\rangle$ is the mean intensity of N reflections with intensities I_i and common indices h, k and l.

All purification steps were carried out at 277 K. Cells were resuspended in 1/25th of the original culture volume of buffer *A* (50 m*M* sodium phosphate and 300 m*M* sodium chloride pH 8.0) and lysed by two passes through a French press at 172 MPa. The crude extract was centrifuged at 12 000*g* for 25 min and the supernatant was mixed for 1 h with TALON IMAC resin (BD) pre-equilibrated with buffer *A*. The protein-bound resin was poured into a column and then washed with buffer *B* (50 m*M* sodium phosphate, 300 m*M* sodium chloride, 10 m*M* imidazole pH 8.0). AK was eluted from the resin with buffer *C* (50 m*M* sodium phosphate, 300 m*M* sodium chloride, 300 m*M* imidazole pH 8.0) and buffer-exchanged into buffer *D* (20 m*M* Tris, 10 m*M* $MgCl_2$, 5 m*M* dithiothreitol pH 8.0) using an Econo-Pac 10DG column (Bio-Rad). The purity of AK was determined by Coomassie-stained SDS–PAGE and found to be greater than 95% (data not shown). The purified protein was then concentrated to 15 mg ml^{-1} using a 30 kDa cutoff concentrator (Amicon) and stored at 193 K. Protein concentrations were determined by the Bradford method (Bradford, 1976) using bovine serum albumin as the standard.

2.2. Crystallization

All crystallization experiments were carried out at 295 K using the hanging-drop vapor-diffusion method. 1 µl protein solution in buffer *D* containing 12 mg ml^{-1} AK and 4 m*M* substrate analog (MMPR or DMA) was mixed with an equal amount of reservoir solution and equilibrated against 0.4 ml reservoir solution. 2 m*M* or 1 m*M* AMP-PCP was also added to the protein solution for the AK–MMPR–AMP-PCP complex or the AK–DMA–AMP-PCP complex, respectively. The reservoir solution for the optimized conditions contained 19–21% PEG 4000, 0.2 *M* ammonium acetate, 0.1 *M* sodium citrate pH 6.0–6.5. Crystals appeared within 3 d and reached their maximum size (0.2 × 0.08 × 0.08 mm) in about 7 d.

The crystals belong to the orthorhombic space group $P2_12_12_1$. The crystals contain one molecule per asymmetric unit, with a Matthews number of 2.2 Å^3 Da^{-1} and a solvent content of 43.3%.

2.3. X-ray data collection and processing

All data were collected at cryogenic temperatures. The crystals were cryoprotected by a quick dip in a Paratone and paraffin mixture with a volume ratio of 1:2 and were then flash-frozen in liquid nitrogen.

The AK–MMPR data were collected to 1.75 Å at the NE-CAT 8-BM beamline at the Advanced Photon Source using a Quantum-315 CCD detector (Area Detector Systems Corp.) over a range of 100° using 30 s exposures for each 1.0° oscillation at a crystal-to-detector distance of 240 mm. The AK–MMPR–AMP-PCP data were collected to 1.75 Å at the home source using an R-AXIS IV image-plate system and a Rigaku RU300 rotating-anode X-ray generator operating at 50 kV and 100 mA. 240 frames with 0.5° oscillation each were collected using 7 min exposures at a crystal-to-detector distance of 110 mm. AK–DMA and AK–DMA–AMP-PCP data were collected to 1.35 Å at the F2 station of the Cornell High Energy Synchrotron Source using a Quantum-210 CCD detector (Area Detector Systems Corp.) and 30 s exposures for each 1.0° oscillation at a crystal-to-detector distance of 130 mm over a range of 120 and 150°, respectively.

The *HKL*-2000 suite (Otwinowski & Minor, 1997) of programs was used for integration and scaling of all data sets. Details of the data collection and processing are given in Table 1.

2.4. Structure determination

The structures of the AK–substrate analog complexes were determined using molecular replacement. Use of the previously published whole apo AK structure (PDB code 1lio) or the structure of the AK–adenosine–AMP-PCP complex (PDB code 1lii) as the search model was attempted, but failed. A solution was only obtained when the large main domain of the apo AK without the lid domain was used as the search model and the cross-rotation and translation functions in the *CNS* software package (Brünger *et al.*, 1998) were performed using the AK–MMPR data in the resolution range 10–4 Å. The correlation coefficient was 0.35. Subsequent refinement was performed using *CNS* followed by manual remodeling with the computer-graphics program *O* (Jones *et al.*, 1991). The lid region was manually built in as the phases were improved. Water molecules were included in the model in subsequent rounds of refinement. As the model improved, the density at the active site was still very strong and MMPR was modeled into the density. The structure of AK–MMPR was used as the starting model for the AK–MMPR–AMP-PCP, AK–DMA and AK–DMA–AMP-PCP complex structures. A similar refinement strategy was employed. For the structures of both AK–DMA and AK–DMA–AMP-PCP at 1.35 Å, *SHELXL* (Sheldrick & Schneider, 1997) was used for refinement at a later stage and *MI-fit* (Molecular Images Software) was used for manual rebuilding. The free *R* flags were carried over from *CNS* to *SHELXL*. Bond distance and angle restraints for

Table 2
Summary of data-refinement statistics.

	AK–MMPR	AK–MMPR–AMP-PCP	AK–DMA	AK–DMA–AMP-PCP
Resolution (Å)	50–1.75	50–1.75	10–1.35	10–1.35
Total No. of non-H atoms	2963	2925	3047	3047
No. of protein atoms	2667	2634	2744	2658
No. of ligand and ion atoms	43	54	49	76
No. of water O atoms	253	237	404	382
No. of reflections in refinement	30495	34001	62694	64552
No. of reflections in test set	2439	2568	5304	5332
R factor† (%)	14.0	19.4	13.3	13.7
R_{free}‡ (%)	20.4	23.0	19.6	19.6
R.m.s. deviations from ideal geometry				
Bonds (Å)	0.005	0.004	0.01	0.01
Angles (°)	1.2	1.2		
Angle distance (Å)			0.03	0.03
Ramachandran plot				
Most favored region (%)	92.7	93.4	94.4	94.1
Additionally allowed region (%)	6.6	5.9	5.0	5.0
Generously allowed region (%)	0	0	0	0.3
Disallowed region (%)	0.7	0.7	0.7	0.7

† R factor = $\sum_{hkl} \big||F_{obs}| - k|F_{calc}|\big| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. ‡ For R_{free}, the sum is extended over a subset of reflections (10%) excluded from all stages of refinement.

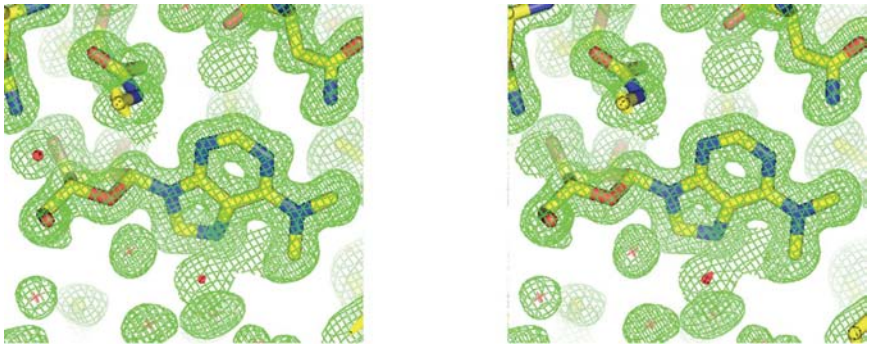

Figure 1
Stereoview of a $2F_{obs} - F_{calc}$ electron-density map at 1.35 Å for the DMA-binding site contoured at 1σ. The residues are shown in ball-and-stick representation, with O atoms in red, C atoms in yellow and N atoms in blue.

DMA were generated using the program *HIC-Up* (Kleywegt & Jones, 1998) and for MMPR and AMP-PCP using *PRODRG* (Schüttelkopf & van Aalten, 2004). In all the structures, one chloride ion and two sodium ions were built in based on the coordination geometry for each ion, bond distance and the refined *B* factors. An acetate ion was also included in the models of AK–MMPR and AK–DMA. The final refinement statistics are given in Table 2.

2.5. Figure preparation

Figures were prepared using *PyMOL* (DeLano, 2002), *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Bacon, 1997).

3. Results

3.1. Overall structure of the complexes

The crystal structures of *T. gondii* AK in complex with MMPR (AK–MMPR), with MMPR and AMP-PCP (AK–MMPR–AMP-PCP), with DMA (AK–DMA) and with DMA and AMP-PCP (AK–DMA–AMP-PCP) were determined by molecular replacement at 1.75, 1.75, 1.35 and 1.35 Å, respectively. All the crystals belong to the orthorhombic space group $P2_12_12_1$ and the asymmetric unit contains one molecule. The final models of all four structures contain residues 10–360, one chloride ion and two sodium ions. In addition, the structure of the AK–MMPR complex contains two MMPR molecules, the structure of the AK–MMPR–AMP-PCP complex contains one MMPR molecule and one AMP-PCP molecule, the structure of the AK–DMA complex contains two DMA molecules and the structure of the AK–DMA–AMP-PCP complex contains one DMA molecule at full occupancy, one DMA molecule at 40% occupancy and one AMP-PCP molecule at 60% occupancy. Pro12 adopts a *cis* conformation in all molecules. All the models have 99.3% of the residues in the allowed regions. Residues Arg83 and Ser198, which are located at non-standard turns and are represented by clear electron density, are in the disallowed region of the Ramachandran plot. A representative section of the electron density is shown in Fig. 1.

The complex structures consist of a small lid domain and a large main domain. The large main domain contains an α-β-α three-layer sandwich, which has been described in detail previously in other *T. gondii* AK structures (Schumacher *et al.*, 2000). The four structures of AK in complex with substrate analogs are essentially identical, with root-mean-square deviations (r.m.s.d.s) of 0.13–0.15 Å for pairwise comparison of the C^{α} atoms of all ordered residues.

3.2. Substrate analog-binding site

The active site of AK is located along one edge of the central β-sheet in the large main domain, with the adenosine-binding site near the lid domain and the ATP-binding site at the other end. In the AK–substrate analog complexes, DMA or MMPR is bound at the adenosine-binding site in the same way as adenosine. There are no significant differences in the substrate-analog binding regardless of the presence of AMP-PCP. Both DMA and MMPR assume a standard *anti* conformation, with a glycosidic torsion angle of −106 to −115°. The ribose takes an uncommon O4′-*endo* sugar pucker. One face of the ribose ring stacks against a highly conserved Gly-Gly dipeptide (Bork *et al.*, 1993; Schumacher *et al.*, 2000; Zhang *et al.*, 2004) and the other against the side chain of Leu138. The

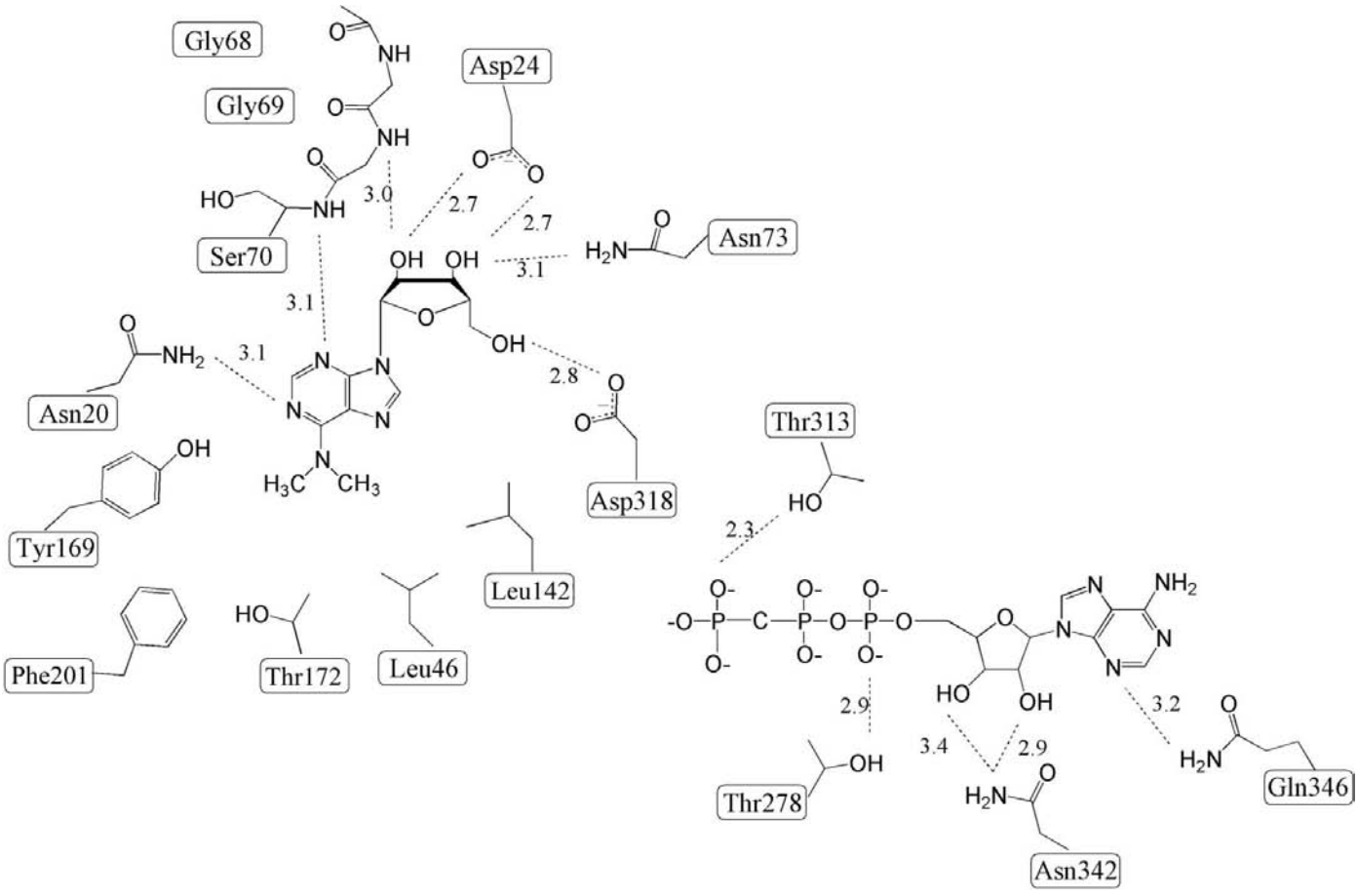

Figure 2
Schematic drawing of the active site of AK complexed with DMA and AMP-PCP. Key hydrogen bonds are indicated by dashed lines, with the corresponding donor–acceptor distance labeled.

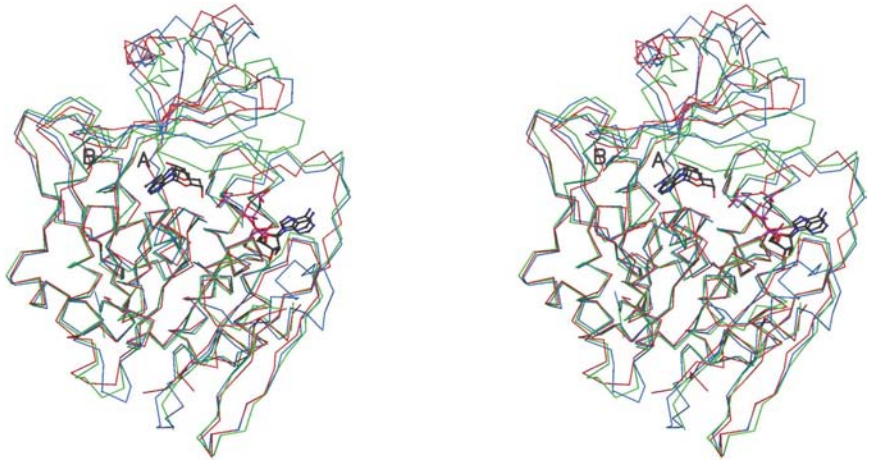

Figure 3
Stereoview of the superposition of the structures of *T. gondii* AK in apo form (red), in complex with DMA and AMP-PCP (blue) and in complex with adenosine and AMP-PCP (green) based on the C^{α} positions of all ordered residues in the large domain. The lid domain is at the top, while the main domain is at the bottom. DMA, adenosine and AMP-PCP are shown in ball-and-stick representation. The Gly68-Gly69 dipeptide is labelled A, while Gly143-*X*-*X*-Gly146 is labeled B.

ribose ring forms hydrogen bonds between the O2′, O3′ and O5′ O atoms and the side chains of residues Asp24, Asn73 and Asp318 and the main-chain N of residue Gly69 (Fig. 2).

Additional interactions between the substrate analogs and the active site come from the base moiety. The N3 atom of the purine ring forms a hydrogen bond to the amide N atom of Ser70 and N1 accepts a hydrogen bond from Asn20 N^{δ}. The C2 atom of the purine base makes van der Waals contacts to a chloride ion. The assignment of the chloride ion is determined by the refined *B* factor, bonding distances and the observation of a chloride ion at the same location in the previously solved *T. gondii* AK (Schumacher *et al.*, 2000) and human AK (Mathews *et al.*, 1998). The purine ring is also stabilized by stacking and van der Waals interactions with the side chains of Tyr169 and Ile22. The 6-dimethyl group and the 6-methyl group interact with the enzyme *via* van der Waals interactions with Leu46 and Leu142. In DMA-containing complexes, the N7 atom of the purine ring hydrogen bonds to a water molecule that in turn forms a hydrogen-bonding network with three nearby water molecules.

3.3. ATP analog-binding site

The ATP-binding site is located in a shallow groove within the large main domain, which is clarified by the binding of the nonhydrolysable ATP analog AMP-PCP in the AK–MMPR–AMP-PCP complex. DMA and MMPR appear to bind where the adenosine moiety of the AMP-PCP binds in the AK–DMA and AK–MMPR complexes, respectively. This phenomenon was also observed in the human AK–adenosine complex (Mathews *et al.*, 1998) and the *T. gondii* AK–adenosine complex (Schumacher *et al.*, 2000). In the *T. gondii* AK–DMA–AMP-PCP complex, both DMA and AMP-PCP are modeled at the ATP-binding site, with an occupancy of 40% for the former and 60% for the latter. DMA and MMPR at the ATP-binding site adopt different conformations compared with their partners at the adenosine-binding site. The sugar rings of DMA, MMPR and AMP-PCP are in the common C3′-*endo* conformation and the glycosidic torsion angles are in an *anti* conformation, with a value of −165 to −173°. The O2′ atom of the ribose moiety forms a hydrogen bond with Asn342 N^{δ} (Fig. 2), while the O3′ atom forms a hydrogen bond with a water molecule that in turn interacts with the carbonyl O atom of Asn282 and the amide N atom of Arg279. The N3 atom of the purine moiety forms a hydrogen bond with Gln346 N^{ε}. The N^6 atom of AMP-PCP in the AK–DMA–AMP-PCP and AK–MMPR–AMP-PCP complexes does not have any interactions with the enzyme or solvent molecules.

The tail of the phosphate group in AMP-PCP points toward the cleft between the large domain and the small lid domain. One α-phosphate O atom is hydrogen bonded to the hydroxyl group of Thr278 (Fig. 2). Two water molecules bridge the interactions between the two β-phosphate O atoms and one γ-phosphate O atom with the enzyme. One γ-phosphate O atom makes a hydrogen bond with the hydroxyl group of Thr313. No magnesium ion was found near the phosphate groups, although magnesium chloride was added during crystallization.

3.4. Sodium ion-binding sites

Two sodium ions were built into each model based on the coordination geometry for each site, the refined *B* factors and the presence of 350 m*M* sodium ion during the protein purification. The first sodium-binding site is adjacent to the anion hole (Schumacher *et al.*, 2000). Ribokinase (Andersson & Mowbray, 2002) and aminoimidazole ribokinase (Zhang *et al.*, 2004) have also been found to bind monovalent cations at the same position. The monovalent cations have been suggested to play a role in the activation of the enzyme (Andersson & Mowbray, 2002). The coordination geometry for the first sodium ion-binding site is best described as a distorted octahedron with one vertex missing for three of four structures, with the exception being the AK–MMPR complex, in which the sodium ion has only four coordination partners. The carbonyl O atoms of Val348, His351, Gly353 and one or two water molecules serve as vertices. The second sodium-binding site is located near the protein surface. The coordination geometry is very similar to the first binding site. The coordination number is four for the sodium ion in the AK–DMA and AK–MMPR complexes and five for the sodium ion in the AK–DMA–AMP-PCP and AK–MMPR–AMP-PCP complexes. The O^{δ} atom of Asn268, the carbonyl O atoms of Thr269 and Thr290 and one or two water molecules serve as vertices. The average observed distance between the sodium ions and O atoms is 2.38 ± 0.05 Å.

4. Discussion

4.1. G*XX*G switch, a supplemental mechanism to the GG switch

The phosphorylation of adenosine by *T. gondii* AK has been shown to require that the enzyme undergo significant conformational changes (Schumacher *et al.*, 2000). The binding of adenosine to the enzyme is proposed to induce a 30° rotation of the lid domain, resulting in a closed conformation and substrate sequestration. The GG conformational switch has been proposed as the trigger for the lid-domain rotation in AK (Schumacher *et al.*, 2000). Gly68 and Gly69, which form the base of the binding site for the ribose moiety, undergo large torsional changes upon the binding of adenosine (Figs. 3 and 4). These two glycine residues are located at one end of the hinge between the two domains. It is reasonable to suggest that the residues at the opposite end of the hinge may also have structural features that are related to the domain rotation. Leu142 is located near the other end of the hinge, with its side chain pointing into the substrate-binding pocket. If it were to maintain this position when adenosine is bound, the N^6 atom would be only 1.7 Å away from the C^{δ} atom of Leu142 (Fig. 4). Consequently, Leu142 moves away from the substrate to avoid unfavorable steric interactions. This movement is achieved by torsional changes of Gly143 and Gly146 adjacent to Leu142. The conformational changes of the G*XX*G motif trigger the formation of a 3_{10}-helix involving the two glycine residues (Figs. 3 and 4). In this scheme, both the GG switch and the coil to 3_{10}-helix transition associated with G*XX*G cause the small lid domain to rotate towards the large core domain while maintaining the tertiary structure of the small domain. The G*XX*G sequence is also conserved in human AK (Mathews *et al.*, 1998) and *E. coli* ribokinase (Sigrell *et al.*, 1999), which shows a similar substrate-induced lid-domain rotation.

4.2. Substrate analog-induced intermediate semi-closed conformational change

Compared with the previously solved structures of the apo form of AK (PDB code 1lio) and the AK–adenosine–AMP-PCP complex (PDB code 1lii), the AK–substrate analog complexes surprisingly reveal an intermediate conformation: semi-closed (or semi-open; Fig. 3). The small lid region undergoes a rigid-body rotation of about 12°, while the structures of the two domains remain essentially the same as for fully closed or fully open AK structures. The r.m.s.d. values are 1.0 and 1.5 Å for comparison of the C^{α} atoms of the large main domains of the AK–substrate analog complexes with the

apo enzyme and of the AK–substrate analog complexes with the AK–adenosine–AMP-PCP complex, respectively. The r.m.s.d. values are 0.6 and 1.2 Å for comparison of the small lid domains. Comparison of the substrate analog-bound complexes in the semi-closed conformation to the adenosine-bound complex in the fully closed conformation reveals that both the GG and G*XX*G regions have similar conformations. This implies that the torsional changes of GG and G*XX*G are sufficient to initiate the rigid-body rotation of the small lid domain, but not to cause complete closure.

The possibility that the semi-closed conformation is caused by crystal packing should be considered. Two helices, including residues 30–38 and 50–62, and the loop between residue 46 and residue 49 in the small lid region are involved in crystal packing, mainly by van der Waals interactions and several hydrogen bonds. The crystal-packing force applied on two helices would tend to favor a more closed lid, whereas the crystal contacts on the loop tend to favor a more open lid. While the effect of crystal packing is not entirely clear, the unsuccessful attempt to grow crystals of substrate-analog complexes under the same conditions used to obtain crystals of adenosine complexes suggests that the analog complexes may not have a fully closed lid in solution.

If the crystal packing does not significantly change the global conformation of the complexes, the 6- or N^6-substitutions are most likely to be responsible for the intermediate conformation. When adenosine is bound, the N^6 atom is hydrogen bonded to a water molecule that in turn hydrogen bonds to the hydroxyl group of Thr172 and the carbonyl O atom of Tyr169. When DMA or MMPR is bound, the side chains of Leu46 and Leu142 undergo significant movement (Fig. 4). Together with Tyr169, Thr172 and Phe201, these residues form a hydrophobic cavity to accommodate the hydrophobic 6- or N^6-substitutions (Fig. 5). The dimethyl group or methylthio group makes van der Waals interactions with the side chains of Leu46 and Leu142.

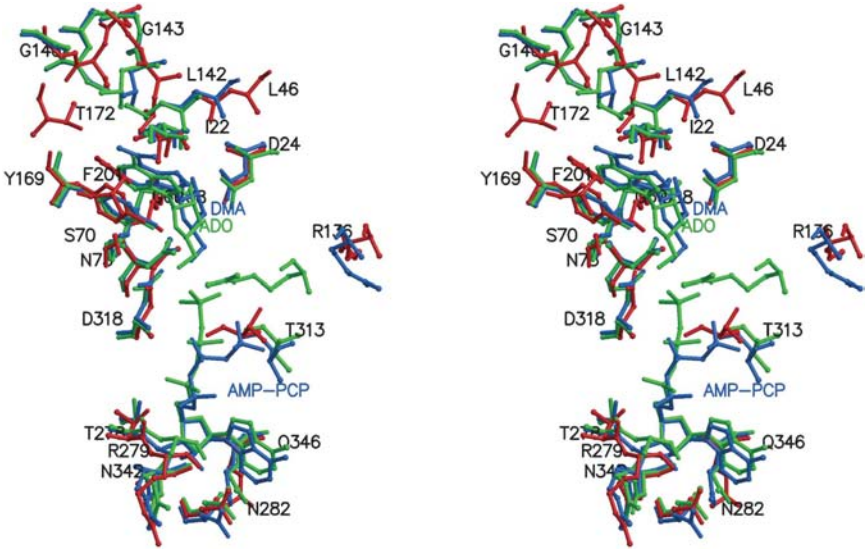

Figure 4
Stereoview of the active site of *T. gondii* AK. The key residues in the apo enzyme are shown as red sticks, those in the AK–DMA–AMP-PCP complex in blue and those in the AK–adenosine–AMP-PCP complex in green. Also shown are the substrate adenosine, DMA and AMP-PCP. The residues are labeled with their one-letter code and residue number. The superposition was based on the C^{α} positions of all ordered residues in the large domain.

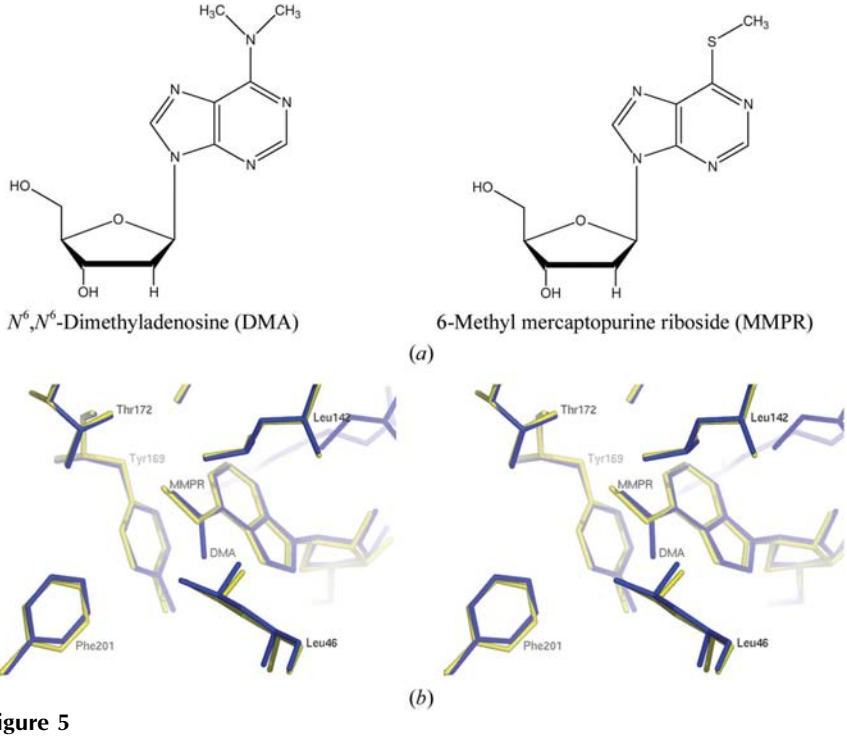

Figure 5
The structures of DMA and MMPR and their binding sites. (*a*) Schematic drawings of DMA and MMPR. (*b*) Stereoview of the binding site of N^6- and C6-substituted groups. The key residues in the DMA-bound complex are shown in blue; those in the MMPR-bound complex are yellow. The residues are labelled with their three-letter code and residue number.

4.3. Important role of Arg136 in orienting ATP for catalysis

Closure of the small lid domain brings Arg136, which is located on the small lid domain, into the active site (Schumacher *et al.*, 2000). The side chain of Arg136 anchors the tail of the ATP molecule (or AMP-PCP) by directly hydrogen bonding to the γ-phosphate O atoms and indirect interactions with the β-phosphate O atoms through water molecules. Therefore, the phosphate groups of ATP are kept in an extended conformation in which the γ-phosphate is located near the anion hole and oriented toward the 5′-OH group of the substrate

with a phosphorus–oxygen distance of 3.1 Å (Schumacher *et al.*, 2000; Zhang *et al.*, 2004). In the substrate analog-bound complexes, for which the active-site lid domain is only partially closed, Arg136 points away from the active site. Lacking the anchoring interactions with Arg136, the phosphate tail of AMP-PCP bends into a relaxed 'C'-shaped conformation near the purine base. In this conformation, the γ-phosphate points away from both the 5′-OH group of substrate analog and the anion hole. The distance between the γ-phosphorus and the 5′-O atom of the substrate analog is about 8 Å (Fig. 4).

Residues Gly315, Ala316, Gly317 and Asp318 form the anion hole, which is one of the most highly conserved regions in the entire ribokinase superfamily (Schumacher *et al.*, 2000; Zhang *et al.*, 2004). The anion hole helps to stabilize the accumulated negative charge during the phosphate-group transfer. The formation of anion hole in the substrate-analog complexes is not affected by the nonproductive conformations of Arg136 or the nonoptimal binding of the ATP analog. This suggests that the orientation of ATP into the catalytic position by Arg136 may be the last step before the initiation of catalysis.

4.4. Implications of the intermediate conformation for catalytic mechanism

The substrate analog-bound complexes contain most of the elements for chemical catalysis, including substrate (or analogs), ATP (or analog) and the anion hole. However, they lack translocation of the key residue Arg136 into the active site and the correct orientation of the γ-phosphate of ATP. The active site is not completely occluded from the solvent owing to the semi-closed lid domain. Thus, the structures of the analog complexes represent an intermediate conformation prior to the initiation of catalysis. When the adenosine is bound, Gly68-Gly69 and Gly143-*XX*-Gly146 undergo conformational changes that initiate the closing of the small lid domain. Simultaneously, binding of ATP induces the creation of the anion hole. In the substrate-analog complex the enzyme is trapped at an early stage with both substrate and ATP bound and with the lid in a semi-closed conformation. In contrast to the substrate-analog complexes, formation of the adenosine complex results in full lid closure and translocation of Arg136 into the active site, where it anchors ATP in its catalytically competent geometry. At this point, the substrate and γ-phosphate of ATP are also shielded from the solvent. In contrast to the previous proposal that adenosine binding is a prerequisite for ATP binding (Schumacher *et al.*, 2000), the intermediate conformation implies that ATP binding may be independent of substrate binding. This possibility has also been suggested for aminoimidazole ribokinase (Zhang *et al.*, 2004), another member of the ribokinase superfamily.

4.5. Structure–activity relationships

DMA and MMPR are N^6-substituted or C6-substituted adenosine analogs and are alternative substrates for *T. gondii* AK (unpublished results), with relatively high affinities of 55 and 22 μM, respectively (Iltzsch *et al.*, 1995). However, the structures of DMA- and MMPR-bound complexes reveal an unfavorable catalytic geometry resulting from the partially closed lid domain, the relaxed conformation of ATP and the large distance between the 5′ hydroxyl group and the γ-phosphate group. In contrast, the binding environment for the 6-substituents of the modified adenosine appears to be favorable. In the case of adenosine, the exocyclic 6-amino group is surrounded by a hydrophobic cavity at the active site of AK. In the case of modified adenosines, the N^6-substitution of DMA and C6-substitution of MMPR make hydrophobic as well as van der Waals interactions with Leu46 and Leu142. Further modification at the N^6 or C6 position of adenosine would affect the interactions with the surrounding residues. For example, 6-benzylthioinosine, with a bulky C6-substituent, has been shown to be an excellent subversive substrate of *T. gondii* AK and a prototype of antitoxoplasmic agents (el Kouni *et al.*, 1999; Rais *et al.*, 2005; Yadav *et al.*, 2004). The large hydrophobic benzylthio group is expected to extend out of an opening lined by Leu46, Leu142, Thr172 and Phe201. Based on the structures of MMPR-bound complexes, the thio group would be approximately at the opening which the nitrobenzyl group extends through. The dimensions of this opening are roughly 6 × 5, 5 × 5 and 4 × 5 Å for the apo enzyme, DMA- or MMPR-bound complexes and adenosine-bound complex, respectively. The second dimension lies between Leu46 and Leu142 or between Thr172 and Phe201. The closure of the lid domain changes only one dimension of this opening. The binding of 6-benzylthioinosine to *T. gondii* AK may require the residues in the opening to adopt a different comformation in order to accommodate a benzyl group, while still allowing full lid closure and translocation of Arg136. Efforts are under way to design new subversive substrates for *T. gondii* AK based on these structure–activity relationships.

This work is based upon research conducted at the Northeastern Collaborative Access Team beamline 8-BM of the Advanced Photon Source, supported by Award RR-15301 from the National Center for Research Resources at the National Institute of Health, and at the Macromolecular Diffraction at Cornell High Energy Synchrotron Source, supported by award RR-01646 from the National Center for Research Resources at the National Institute of Health. We thank Dr Cynthia Kinsland for the preparation of the His-tagged AK overexpression plasmid. We thank Drs Angela V. Toms, Yang Zhang and Christopher Lehmann for assistance with data collection and Leslie Kinsland for assistance with the preparation of this manuscript. This work was also supported by NIH grant AI052838 (MHK and SEE).

References

Andersson, C. E. & Mowbray, S. L. (2002). *J. Mol. Biol.* **315**, 409–419.
Bork, P., Sander, C. & Valencia, A. (1993). *Protein Sci.* **2**, 31–40.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M.,

Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Buzby, J. C. & Roberts, T. (1996). *Food Rev.* **19**, 20–25.
Campobasso, N., Mathews, I. I., Begley, T. P. & Ealick, S. E. (2000). *Biochemistry*, **39**, 7868–7877.
Cao, P., Gong, Y., Tang, L., Leung, Y. C. & Jiang, T. (2006). *J. Struct. Biol.* **154**, 327–332.
Cheng, G., Bennett, E. M., Begley, T. P. & Ealick, S. E. (2002). *Structure*, **10**, 225–235.
Couvreur, J., Tournier, G., Sardetfrismand, A. & Fauroux, B. (1992). *Presse Med.* **21**, 1569–1574.
Daffos, F., Mirlesse, V., Hohlfeld, P., Jacquemard, F., Thulliez, P. & Forestier, F. (1994). *Lancet*, **344**, 541.
Dannemann, B., McCutchan, J. A., Israelski, D., Antoniskis, D., Leport, C., Luft, B., Nussbaum, J., Clumeck, N., Morlat, P., Chiu, J., Vilde, J. L., Orellana, M., Feigal, D., Bartok, A., Heseltine, P., Leedom, J. & Remington, J. (1992). *Ann. Intern. Med.* **116**, 33–43.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.
el Kouni, M. H. (2003). *Pharmacol. Ther.* **99**, 283–309.
el Kouni, M. H., Guarcello, V., al Safarjalani, O. N. & Naguib, F. N. M. (1999). *Antimicrob. Agents Chemother.* **43**, 2437–2443.
Iltzsch, M. H., Uber, S. S., Tankersley, K. O. & el Kouni, M. H. (1995). *Biochem. Pharmacol.* **49**, 1501–1512.
Ito, S., Fushinobu, S., Jeong, J. J., Yoshioka, I., Koga, S., Shoun, H. & Wakagi, T. (2003). *J. Mol. Biol.* **331**, 871–883.
Ito, S., Fushinobu, S., Yoshioka, I., Koga, S., Matsuzawa, H. & Wakagi, T. (2001). *Structure*, **9**, 205–214.
Jones, J. L., Hanson, D. L., Dworkin, M. S., Alderton, D. L., Fleming, P. L., Kaplan, J. E. & Ward, J. (1999). *Arch. Dermatol.* **135**, 897–902.
Jones, J. L., Kruszon-Moran, D., Wilson, M., McQuillan, G., Navin, T. & McAuley, J. B. (2001). *Am. J. Epidemiol.* **154**, 357–365.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kleywegt, G. J. & Jones, T. A. (1998). *Acta Cryst.* D**54**, 1119–1131.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Krug, E. C., Marr, J. J. & Berens, R. L. (1989). *J. Biol. Chem.* **264**, 10601–10607.
Leport, C., Derouin, F., Morlat, P., Chene, G. & Vilde, J. L. (1996). *Med. Mal. Infect.* **26**, 437–440.
Leport, C., Vilde, J. L., Katlama, C., Regnier, B., Matheron, S. & Saimot, A. G. (1987). *Ann. Med. Interne (Paris)*, **138**, 30–33.
Li, M.-H., Kwok, F., Chang, W.-R., Lau, C.-K., Zhang, J.-P., Lo, S. C. L., Jiang, T. & Liang, D.-C. (2002). *J. Biol. Chem.* **277**, 46385–46390.
Luft, B. J. & Remington, J. S. (1992). *Clin. Infect. Dis.* **15**, 211–222.
McAuley, J. *et al.* (1994). *Clin. Infect. Dis.* **18**, 38–72.
Mariuz, P., Bosler, E. M. & Luft, B. J. (1994). *Infect. Dis. Clin. North Am.* **8**, 365–381.
Mathews, I. I., Erion, M. D. & Ealick, S. E. (1998). *Biochemistry*, **37**, 15607–15620.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Ohshima, N., Inagaki, E., Yasuike, K., Takio, K. & Tahirov, T. H. (2004). *J. Mol. Biol.* **340**, 477–489.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pfefferkorn, E. R. & Pfefferkorn, L. C. (1978). *J. Parasitol.* **64**, 486–492.
Rais, R. H., Al Safarjalani, O. N., Yadav, V., Guarcello, V., Kirk, M., Chu, C. K., Naguib, F. N. & el Kouni, M. H. (2005). *Biochem. Pharmacol.* **69**, 1409–1419.
Recacha, R., Talalaev, A., DeLucas, L. J. & Chattopadhyay, D. (2000). *Acta Cryst.* D**56**, 76–78.
Schumacher, M. A., Scott, D. M., Mathews, I. I., Ealick, S. E., Roos, D. S., Ullman, B. & Brennan, R. G. (2000). *J. Mol. Biol.* **296**, 549–567.
Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* D**60**, 1355–1363.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Sigrell, J. A., Cameron, A. D., Jones, T. A. & Mowbray, S. L. (1998). *Structure*, **6**, 183–193.
Sigrell, J. A., Cameron, A. D. & Mowbray, S. L. (1999). *J. Mol. Biol.* **290**, 1009–1018.
Sullivan, W. J. Jr, Chiang, C. W., Wilson, C. M., Naguib, F. N., el Kouni, M. H., Donald, R. G. & Roos, D. S. (1999). *Mol. Biochem. Parasitol.* **103**, 1–14.
Yadav, V., Chu, C. K., Rais, R. H., Al Safarjalani, O. N., Guarcello, V., Naguib, F. N. M. & el Kouni, M. H. (2004). *J. Med. Chem.* **47**, 1987–1996.
Zhang, R. G., Grembecka, J., Vinokour, E., Collart, F., Dementieva, I., Minor, W. & Joachimiak, A. (2002). *J. Struct. Biol.* **139**, 161–170.
Zhang, Y., Dougherty, M., Downs, D. M. & Ealick, S. E. (2004). *Structure*, **12**, 1809–1821.