THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 32, pp. 23473–23481, August 10, 2007

Structural Insights into Catalysis and Inhibition of *O*-Acetylserine Sulfhydrylase from *Mycobacterium tuberculosis*

*CRYSTAL STRUCTURES OF THE ENZYME α-AMINOACRYLATE INTERMEDIATE AND AN ENZYME-INHIBITOR COMPLEX**[S]

Received for publication, April 27, 2007, and in revised form, May 30, 2007 Published, JBC Papers in Press, June 13, 2007, DOI 10.1074/jbc.M703518200

Robert Schnell[‡1], **Wulf Oehlmann**[§], **Mahavir Singh**[§¶], **and Gunter Schneider**[‡2]

From the [‡]*Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden,* [§]*LIONEX GmbH, Inhoffenstrasse 7, 38124 Braunschweig, Germany and the* [¶]*Helmholtz Center for Infection Research, Inhoffenstrasse 7, 38124 Braunschweig, Germany*

Cysteine biosynthetic genes are up-regulated in the persistent phase of *Mycobacterium tuberculosis*, and the corresponding enzymes are therefore of interest as potential targets for novel antibacterial agents. *cysK1* is one of these genes and has been annotated as coding for an *O*-acetylserine sulfhydrylase. Recombinant CysK1 is a pyridoxal phosphate (PLP)-dependent enzyme that catalyzes the conversion of *O*-acetylserine to cysteine. The crystal structure of the enzyme was determined to 1.8 Å resolution. CysK1 belongs to the family of fold type II PLP enzymes and is similar in structure to other *O*-acetylserine sulfhydrylases. We were able to trap the α-aminoacrylate reaction intermediate and determine its structure by cryocrystallography. Formation of the aminoacrylate complex is accompanied by a domain rotation resulting in active site closure. The aminoacrylate moiety is bound in the active site via the covalent linkage to the PLP cofactor and by hydrogen bonds of its carboxyl group to several enzyme residues. The catalytic lysine residue is positioned such that it can protonate the Cα-carbon atom of the aminoacrylate only from the *si*-face, resulting in the formation of L-cysteine. CysK1 is competitively inhibited by a four-residue peptide derived from the C-terminal of serine acetyl transferase. The crystallographic analysis reveals that the peptide binds to the enzyme active site, suggesting that CysK1 forms an bi-enzyme complex with serine acetyl transferase, in a similar manner to other bacterial and plant *O*-acetylserine sulfhydrylases. The structure of the enzyme-peptide complex provides a framework for the design of strong binding inhibitors.

Gene expression and proteome analysis of various models of dormancy in *Mycobacterium tuberculosis* have shown that the genes involved in sulfate transport (1), sulfate reduction (PAPS pathway), and cysteine biosynthesis (2–5) are up-regulated. In the persistent phase, the pathogen is surviving within macrophages (6, 7), where it is exposed to oxidative stress and reactive nitrogen intermediates as a cellular response to pathogen invasion, leading to oxidation and *S*-nitrosylation of cysteine residues (8). Bacteria in the group designated as *Actinomycetales*, including mycobacteria, produce mycothiol (1-d-*myo*-inosityl-2-(*N*-acetyl-cysteinyl) amino-2-deoxy-α-D-glucopyranoside) as their principal low molecular-mass thiol, which plays the role of glutathione and contains a cysteine moiety (9). Thereby the first defense line in oxidative stress is directly linked to the availability of cysteine. Inhibition of cysteine biosynthesis therefore appears to be an attractive route to novel antibacterials that may also be active against persistent *M. tuberculosis*.

In plants, some archaea, and most eubacteria, the *de novo* biosynthesis of L-cysteine starts from serine and involves two enzymes. Serine acetyl transferase (SAT,[3] EC 2.3.1.30), also denoted CysE, catalyzes the formation of *O*-acetylserine (OAS) from acetyl-CoA and serine. This metabolite is then converted to cysteine through the elimination of acetate and addition of hydrogen sulfide by the enzyme *O*-acetylserine sulfhydrylase (OASS, EC 2.5.1.47) (for recent reviews, see Refs. 10 and 11). The sulfur source is supplied by the sulfate reduction (APS/PAPS) pathway. In most bacteria, pyridoxal phosphate (PLP)-dependent OASS is found as two isoenzymes, OASS-A (also denoted CysK) and OASS-B (denoted CysM). These isoenzymes show 25–45% identity in amino acid sequence but exhibit characteristic differences in their substrate specificity with respect to the sulfur donor. CysK uses hydrogen sulfide as sulfur source, whereas the characterized CysM proteins tend to accept thiosulphate and larger substrates as sulfur source (12, 13). The genome of *M. tuberculosis* H37Rv encodes three genes that were annotated as OASS (the TubercuList World-Wide Web Server). CysK1 (Rv2334) is homologous to OASS-A, and CysM (Rv1336) is homologous to OASS-B. The third isoenzyme, encoded by Rv0848, shows similarities to the other two

* This work was supported by the European Commission Contract LSHP-CT-2005-018729. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

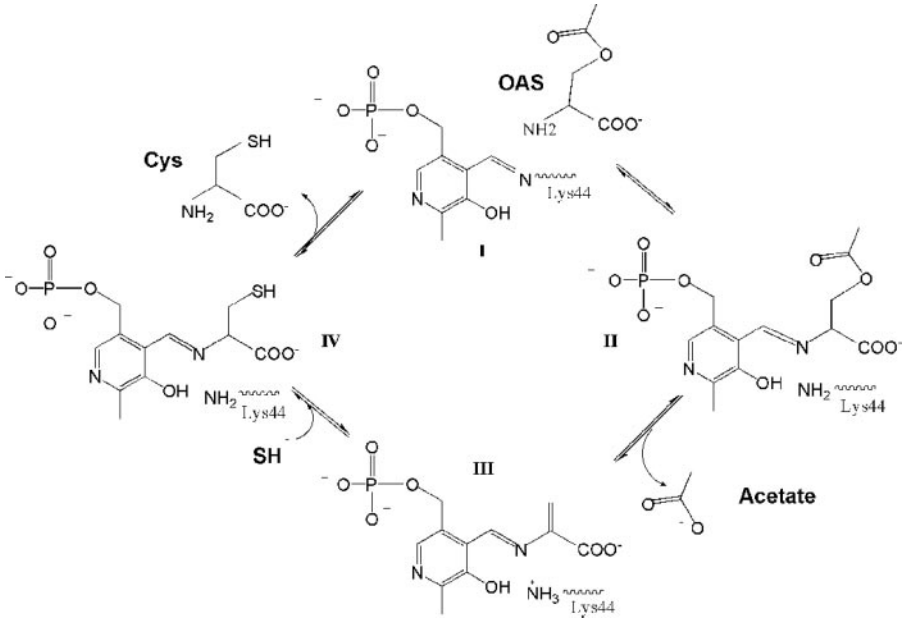
The atomic coordinates and structure factors (code 2Q3B, 2Q3D, and 2Q3C) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains two supplemental figures.

[1] Supported by the David and Astrid Hagelén Foundation.

[2] To whom correspondence should be addressed: Dept. of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden. E-mail: Gunter.Schneider@ki.se.

[3] The abbreviations used are: SAT, serine acetyl transferase; OAS, *O*-acetyl-L-serine; OASS, *O*-acetylserine sulfhydrylase; MPD, 2-methyl-2,4-pentane-diol; PLP, pyridoxal phosphate; MOPS, 4-morpholinepropanesulfonic acid; Bis-Tris, 2-(bis(2-hydroxyethyl)amino)-2-(hydroxymethyl)propane-1,3-diol.



SCHEME 1. **Schematic representation of the reaction catalyzed by CysK1.**

enzymes to the same extent and is therefore designated as either CysK2 or CysM2.

The crystal structure analysis of OASS from *Salmonella typhimurium* (14) *Haemophilus influenzae* (15), *Escherichia coli* (16), and *Arabidopsis thaliana* (17) revealed that it belongs to the fold type II class of vitamin B_6-dependent enzymes (18, 19). The enzymatic mechanism of OASS follows a ping-pong mechanism typical for PLP-dependent enzymes (10, 11, 20). The reaction can be divided into two halves (Scheme 1). In the resting state, the cofactor forms an internal Schiff base with the invariant catalytic lysine residue (compound **I** in Scheme 1). The incoming substrate then forms an external Schiff base with PLP (**II**) followed by a β-elimination in which acetate is released and a proton is abstracted from the α position, most likely by the lysine side chain (21). A similar role of the invariant lysine residue has been proven in the homologous tryptophan synthase β-subunit (22). The product of the first half-reaction is the α-aminoacrylate intermediate (23), covalently linked to PLP (**III**). The second half-reaction starts with a nucleophilic attack of the sulfide (or HS^-) on the β-carbon of the aminoacrylate intermediate, and the α-carbon is reprotonated (20, 24), resulting in cysteine bound as external Schiff base (**IV**). Finally, the product is released via regeneration of the internal aldimine. Formation of the external aldimine appears to be coupled to a large conformational change that leads to the closure of the active site, as suggested by the structure of the *S. typhimurium* CysK K41A mutant in complex with methionine, a substrate analogue (25). However, in the corresponding structure of OASS from *A. thaliana*, the enzyme also remained in the open conformation after formation of the covalent PLP-methionine adduct (17). The latter is assumed to mimic the external aldimine (**II**).

In this work, we describe the cloning, purification, characterization, and 1.8 Å x-ray structure of CysK1 from *M. tuberculosis*. We were also able to trap the α-aminoacrylate in the crystal and report the structure of this reaction intermediate to 2.2 Å resolution, thus providing novel structural insights in the reaction cycle of CysK1. We also determined the crystal structure of CysK1 from *M. tuberculosis* bound to an inhibitory four-residue peptide derived from the C terminus of *M. tuberculosis* CysE (SAT, Rv2335). The structure of this inhibited form of CysK1 may provide the basis for the design of strong binding inhibitors of this enzyme.

EXPERIMENTAL PROCEDURES

Gene Cloning and Protein Production—The DNA sequence coding for CysK1 from *M. tuberculosis* (*Rv2334*) was amplified by PCR and cloned using upstream NdeI and downstream HindIII restriction sites. Subsequently, the sequence was verified, and the fragment was cloned into the expression vector pET28a (Novagen), resulting in a cleavable six-histidine tag at the N terminus of the polypeptide chain. In this construct, thrombin cleavage results in a recombinant protein with three additional amino acids (Gly-Ser-His) at the N terminus.

E. coli BL21(DE3) carrying the expression construct pET-His6CysK1 was cultivated in 1.5 liters of LB medium supplemented with kanamycin (30 μg/ml) at 21 °C. At an A_{600} value of 0.5–0.6, gene expression was induced by the addition of 0.1 mM isopropyl-1-thio-β-D-galactopyranoside. After ~24 h, the cells were harvested and then resuspended in a buffer consisting of 10 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 10 mM imidazole. Cells were disrupted by freeze-thaw cycles, lysozyme, and DNase I treatment and sonication. The clarified lysates were loaded on a nickel-nitrilotriacetic acid column (Qiagen) with a column volume of 1.8 ml. After extensive washing with 10 mM imidazole, remaining proteins were eluted with an imidazole step gradient. Pure His_6-CysK1 eluted in fractions containing 100–200 mM imidazole. His_6-CysK1-containing fractions were pooled, loaded on a desalting PD10 column (GE-Healthcare), and eluted in a buffer of 25 mM Tris-HCl, pH 8.0, and 150 mM NaCl. The N-terminal His_6 tag was removed by thrombin cleavage. Residual non-cleaved His-CysK1 was removed by passage through a nickel-nitrilotriacetic acid column. The flow-through containing the cleaved CysK1 protein was further purified using an ion-exchange step on a Q-Sepharose (Amersham Biosciences) column and NaCl gradient elution. The CysK1-containing fractions (200–250 mM NaCl) were pooled and loaded on a Superdex 200 column equilibrated with 25 mM Tris-HCl, pH 8.0, and 150 mM NaCl. This procedure resulted in homogeneous CysK1 as verified by SDS chromatography. CysK1 protein was concentrated to 10 mg/ml using an Amicon centrifugation device with a 3-kDa molecular mass cut-off. Aliquots of the protein preparations were flash-frozen in liquid nitrogen and stored at −80 °C until further use.

Spectrophotometry—UV-visible absorbance spectra of CysK1 were recorded at a concentration of 3.3 mg/ml using a Jasco-V65 spectrophotometer. Spectra of the covalent aminoacrylate complex were obtained after the addition of 300 μM *O*-acetylserine to the enzyme solution.

Enzyme Activity—Recombinant CysK1 was assayed for *O*-acetylserine sulfhydrylase activity spectrophotometrically at 560 nm by monitoring the formation of cysteine using the acid-ninhydrin method (26). The reactions were carried out in 1500 μl of 100 mM MOPS buffer at pH 7.0. *O*-Acetyl-L-serine and sodium sulfide (dissolved in 1 mM NaOH) were added to final concentrations of 10 and 0.25 mM, respectively, and the reactions were started by the addition of the enzyme (0.025 μg) to the assay mixtures. Aliquots of 250 μl were taken at various time points, and the reaction was stopped by the addition of trichloroacetic acid (final concentration 16.6%). Following centrifugation, the amount of product was determined spectrophotometrically using a standard curve for cysteine in the range 0.04–0.8 mM. Specific activities were derived from the linear part of the time curves, and each measurement was carried out in triplicate at 30 and 37 °C, respectively.

Peptide Binding Assay and Inhibition Studies—The chemically synthesized oligopeptide DFSI was obtained from the GenScript Corp., Scotch Plains, NJ. The binding interaction was probed in a ThermoFluor experiment (27) in the buffer consisting of 25 mM Tris-HCl, pH 8.0, and 150 mM NaCl in the presence of the dye Sypro Orange (Sigma). CysK1 at a concentration of 0.4 μg/μl (12 pmol/μl) in a total volume of 25 μl was mixed with the DFSI peptide at concentrations of 0.0016, 0.016, 0.16, and 1.6 mM. Fluorescence at the wavelength 575 nm was recorded in the temperature interval 20–90 °C.

The initial rate of cysteine formation by CysK1 was determined at 22 °C in the absence and presence of various DFSI peptide concentrations, and varying the OAS concentration in the range 1.0, 2.0, 5.0, 8.0, and 10 mM at a constant sodium sulfide concentration of 1.0 mM in 100 mM MOPS buffer, pH 7.0. At various time intervals, aliquots of the reaction mixture were removed and assayed for L-cysteine formation as described above. The mode of inhibition was determined from double reciprocal plots, and the K_I value for the peptide was derived using the formalism described in (28).

Crystallization and Data Collection—Crystallization screening was carried out using the vapor diffusion method and a Phoenix crystallization robot. After extensive screening and optimization, two conditions were established that resulted in single crystals of CysK1. Both contained unusually high 2-methyl-2,4-pentane-diol (MPD) concentrations. Crystal form I was obtained by mixing 3 μl of protein solution (10 mg/ml) with 2 μl of reservoir solution (Bis-Tris-propane, pH 7.5, and 60–65% MPD) in sitting drops. The rod-like crystals diffracted, however poorly (~6 Å), and were not characterized further. Crystal form II was obtained by mixing 2 μl of protein solution (10 mg/ml) with 2 μl of reservoir solution (sodium HEPES, pH 7.2–8.2, and 80% MPD) in either sitting or hanging drops. The diamond-shaped single crystals reproducibly diffracted to 1.8–2.0 Å resolution.

The CysK1 α-aminoacrylate complex was prepared by a combination of soaking and freeze-trapping. Sitting drops containing CysK1 crystals obtained at pH 8.2 were layered cautiously with a droplet of mother liquor containing the substrate *O*-acetylserine at concentrations of 1 and 2 mM, respectively. At various time intervals (1 mM OAS concentration after 10, 15, 20, 30, 35, and 30 min; at 2 mM OAS concentration after 10 and 20 min), crystals were removed, flash-frozen, and stored in liquid nitrogen until x-ray analysis. Due to the high MPD concentration in the mother liquor, no cryoprotectant was necessary.

Co-crystallization of CysK1 with the DFSI-peptide was achieved using the crystallization conditions established for the holoenzyme. CysK1 at a concentration of 10 mg/ml was mixed with the peptide at various peptide concentrations (0.1, 1.0, 4.0, 5.0, 10.0 mM) and preincubated at 22° for 30–60 min before setting up the crystallization drops using the conditions described above.

Diffraction data for crystals of the holoenzyme were collected at beamline ID14:3 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Diffraction data for the peptide complex were collected at the beamline ID14:1 of the ESRF. Diffraction data sets for the CysK1-aminoacrylate complexes were collected at beamline 9/11 at MAX-lab (Lund, Sweden). All data sets were collected in a nitrogen gas stream at 110 K.

The x-ray data were processed and scaled with the programs MOSFLM and SCALA from the CCP4 suite (29). The native crystals belong to the tetragonal space group $P4_12_12$ with cell dimensions $a = b = 71.0$ Å and $c = 179.6$ Å. Soaking of the crystals with OAS resulted in a minor change of cell parameters to $a = b = 71.7$ Å and $c = 181.0$ Å. The co-crystallization of CysK1 with the DFSI peptide resulted in crystals with cell dimensions $a = b = 72.45$ and $c = 178.93$. The statistics of the data sets are given in Table 1.

Molecular Replacement and Crystallographic Refinement—The structure was determined by molecular replacement using the program MOLREP (30). The coordinates for OASS from *A. thaliana* (17) (Protein Data Bank accession code 1Z7W) with the cofactor PLP and the bound sulfate ion omitted were used as the search model. The best solution had a score of 0.435 and an *R*-factor of 48.3%, with a single monomer in the asymmetric unit. To monitor the behavior of the refinement process, 5% of the x-ray data were removed for the calculation of R-free. Initial cycles of restrained refinement using Refmac5 (31) resulted in a drop of the *R*-factor by 12.8%. The correctness of the molecular replacement solution was confirmed by electron density for the cofactor, including the covalent linkage to the active site lysine residue appearing at the expected positions.

Manual rebuilding of the model was carried out with the program Coot (32) based on σ-A weighted $2F_o - F_o$ and $F_o - F_o$ electron density maps (33). Manual adjustment of the model was interspersed with rounds of refinement by Refmac5 (31). Water molecules were added based on peak heights, shape of the electron density, temperature factor, and capability to form hydrogen bonds to surrounding protein residues and/or other water molecules. The final model contains amino acid residues 1–300 from CysK1, one PLP molecule, four MPD molecules, one chloride ion, and 205 water molecules. The last 10 residues at the C terminus did not have defined electron density.

TABLE 1
Data collection and refinement statistics

Parameter	CysK1-native	CysK1-aminoacrylate complex	CysK-DFSI complex
Data collection			
Beam line	ID14-3 (ESRF)	9/11 (MAX-lab)	ID14-1 (ESRF)
Wavelength (Å)	0.931	1.000	0.934
Resolution (Highest resolution shell)	37.9-1.80 (1.90-1.80)	71.80-2.20 (2.32-2.20)	59.66-2.10 (2.21-2.10)
No. of observed reflections	310962 (45355)	131874 (18830)	134681 (19632)
No. of unique reflections	43621 (6244)	24607 (3550)	26571 (3941)
$I/\sigma(I)$	16.4 (3.1)	14.2 (2.4)	11.5 (3.3)
Completeness	100.0 (100.0)	98.9 (99.8)	93.2 (96.2)
R_{sym}	0.094 (0.542)	0.091 (0.499)	0.089 (0.352)
B-factor from Wilson plot (Å^2)	21.2	43.0	31.0
Refinement			
Resolution interval			
R_{factor} (%)	17.4 (21.7)	19.4 (30.2)	18.7 (26.7)
R_{free} (%)	19.5 (24.0)	24.9 (31.8)	20.7 (29.6)
No. of protein atoms	2277	2297	2240
No. of cofactor/ligand atoms	15	21	49
No. of solvent atoms (MPD, water, chloride l)	236	51	171
Overall *B*-factor (Å^2)	24.6	56.1	36.5
Average *B*-factor protein atoms (Å^2)	23.1	56.3	36.0
Average *B*-factor ligand atoms (Å^2)	16.8	42.5	33.0
Average *B*-factor solvent atoms (Å^2)	38.6	52.9	42.9
r.m.s.[a] deviation bonds (Å)	0.01	0.012	0.009
r.m.s.[a] deviation angles (degrees)	1.28	1.42	1.20
Ramachandran plot			
Percentage of residues in most favorable regions (excluding Gly and Pro)	92.4	91.4	
Percentage of residues in additional allowed regions (excluding Gly and Pro)	7.6	8.6	

[a] r.m.s., root mean square.

Due to the change in cell dimensions upon soaking with OAS, molecular replacement was used to solve the structure of the CysK1 α-aminoacrylate complex. An initial CysK1 model was used as search model. The electron density maps calculated from the data set collected from a crystal soaked in 1 mM OAS for 30 min clearly indicated formation of the α-aminoacrylate intermediate. Refinement of the complex was carried out using a similar protocol as outlined above. At the final stage, TLS refinement (34) was used with 20 TLS groups as defined by the TLSMD server (35), which resulted in a 2% drop in the free *R*-factor. The final model contains residues 1–306 from CysK1, the amino-acrylate PLP covalent complex, one MPD, and 47 crystallographic water molecules.

The structure of the CysK1-peptide complex was solved by molecular replacement using the initially refined CysK1 structure as a search model. The electron density calculated after the initial refinement rounds indicated the presence of the peptide bound at the enzyme active site. Refinement of the complex was carried out using a similar protocol as outlined above. The final model comprises CysK1 residues 1–300, the four-residue-long peptide, one MPD molecule, and 163 water molecules.

The protein models were analyzed with PROCHECK (36) to monitor the stereochemistry. Details of the refinement and protein models are given in Table 1. Figures were made using the program PyMOL. The sequence alignment was made by ClustalW (37).

RESULTS AND DISCUSSION

Characterization of CysK1 from M. tuberculosis—CysK1 was produced with a hexahistidine tag at the N terminus. The tag was removed by thrombin leaving three extra residues Gly-Ser-His on the N terminus of the purified protein. Recombinant CysK1 has a yellow color and shows the typical absorption spectrum of PLP-dependent enzymes (38). The absorbance peak at 410 nm is characteristic of the internal Schiff base formed between the cofactor and the active site lysine residue (Fig. 1).

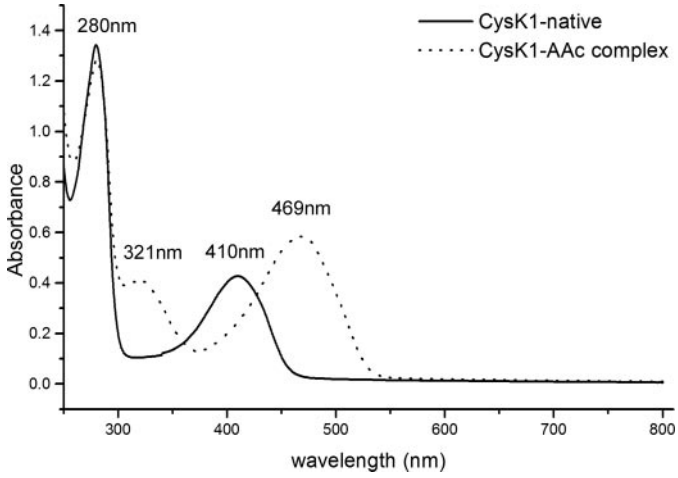


FIGURE 1. **UV-visible spectra of recombinant holo-CysK1 and the complex with the α-aminoacrylate reaction intermediate bound at the active site as external Schiff base.** The spectra were recorded in 25 mM Tris-HCl buffer, pH 8.0, with an enzyme concentration of 3.3 mg/ml. The spectrum of the covalent aminoacrylate complex was obtained after the addition of 300 μM *O*-acetylserine to the enzyme solution.

Gel filtration chromatography suggested that CysK1 is a dimer in solution and that the enzyme migrated in a native gel as a single well defined band (data not shown). Purified recombinant CysK1 is catalytically active as an *O*-acetylserine sulfhydrylase. The specific activity was determined with *O*-acetyl-L-serine and sodium sulfide as substrates. The turnover number was 211 s^{-1} at 30 °C and 352 s^{-1} at 37 °C. These values compare well with those determined for the related enzyme from *S. typhimurium*, which showed 280 s^{-1} at 25 °C as turnover rate for the reaction (39).

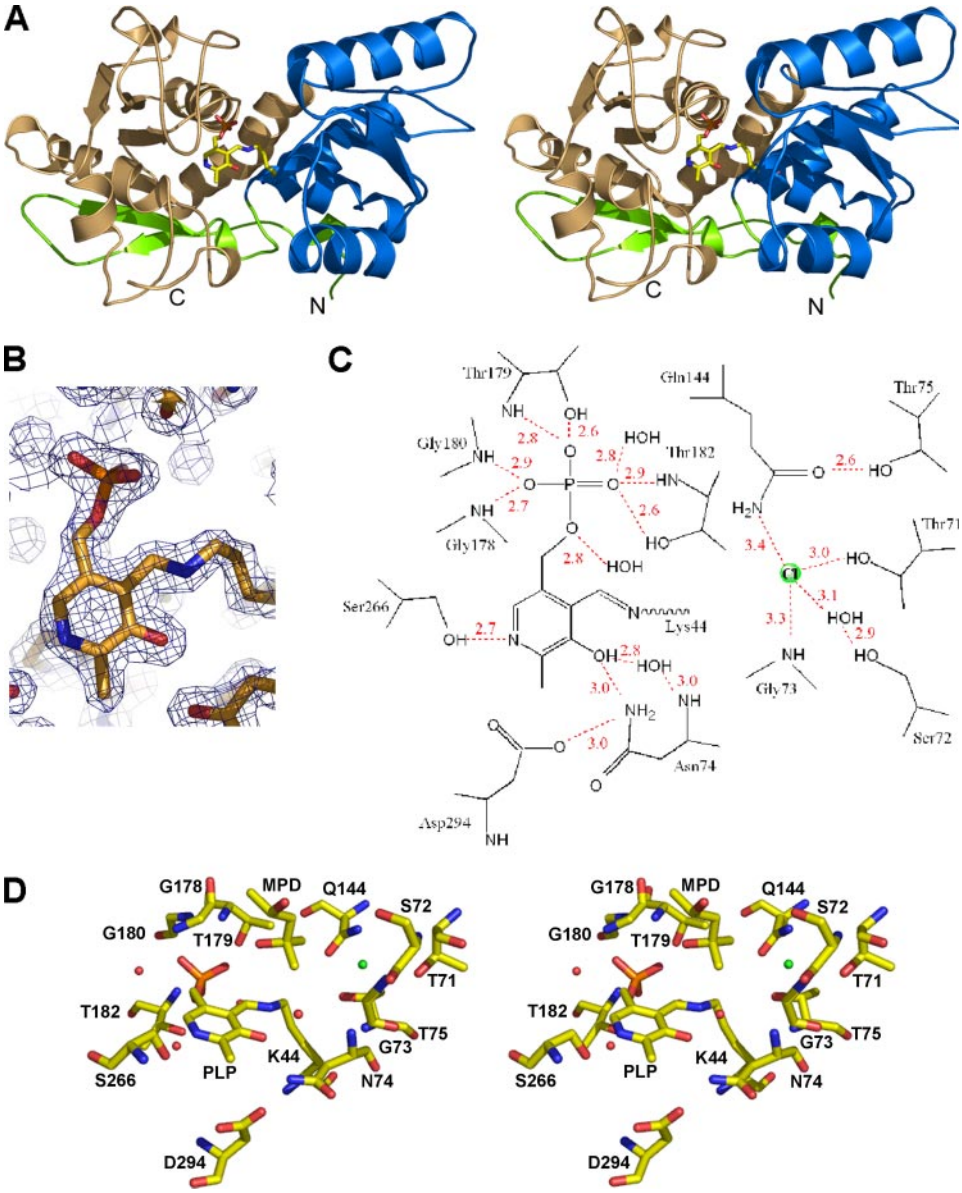


FIGURE 2. **Structure of holo-CysK1.** *A*, stereo view of the overall structure of CysK1. The N-terminal fragment that forms part of the C-terminal domain is shown in *green*, the N-terminal domain is shown in *blue*, and the C-terminal domain is shown in *light brown colors*. The cofactor PLP forming the internal Schiff base with lysine 44 is shown as a *stick model*. *B*, the internal aldimine at the active site of CysK1. The PLP cofactor forming the internal Schiff base with lysine 44 is shown as a *stick model*. The $2F_o - F_o$ electron density map is contoured at 1.8σ. *C*, schematic representation of the coordination of the cofactor and the anion binding site in holo-CysK1. Hydrogen bonds are indicated by *dashed lines*. *D*, stereo view of the active site in CysK1. The amino acid side chains, the bound MPD molecule, and PLP are shown as *stick models*, and water molecules are shown as *red spheres*. The *green sphere* represents the bound chloride ion.

Overall Structure of M. tuberculosis CysK1—The asymmetric unit contains a single copy of the CysK1 subunit, corresponding to a solvent content of 65%. The subunit consists of two domains, each having a α/β-fold (Fig. 2*A* and supplemental Fig. S1), typical of the fold type II PLP enzymes (18, 19). The N-terminal domain (residues 43–145) is built up by a central four-stranded parallel β-sheet, which is flanked by three α-helices on one side and one α-helix on the other side. The C-terminal domain consists of residues 11–42 and 151–304 that fold into a six-stranded mixed β-sheet sandwiched by four α-helices. In the crystal, the enzyme is a dimer, with the molecular axis coinciding with the 2-fold crystallographic axis. The buried surface at the dimer interface is 2130 Å^2, which corresponds to 15.7% of the total surface of one monomer.

CysK1 is related in fold to other PLP enzymes of the fold type II class, *i.e.* the β-subunit of tryptophan synthase (40), threonine deaminase (41), and OASS (14–17). The closest structural relatives of *M. tuberculosis* CysK1 are OASS from *S. typhimurium* (14) and *A. thaliana* (17) with root mean square deviation values after structural superposition of 0.81 Å (294 equivalent Cα-atoms) and 0.7 Å (298 equivalent Cα-atoms), respectively.

Pyridoxal Phosphate Binding Site—The PLP binding site is located in a cleft formed between the C-terminal ends of the β-sheets of the two domains within one subunit (Fig. 2*A*). The cofactor is bound via a covalent linkage to the side chain of Lys^{44} (Fig. 2*B*). The phosphate group forms hydrogen bonds to the highly conserved glycine- and threonine-rich loop (^{178}GTGGT^{182}) at the N terminus of helix 6 and three water molecules (Fig. 2, *C* and *D*). The aromatic ring of PLP is positioned at a crossover connection in the β-sheet of the C-terminal domain. The N-1 atom of the pyridine ring is hydrogen-bonded to the side chain of Ser^{266}, located in another highly conserved region (^{264}GISSGAA^{270}) of the polypeptide chain. The 3′-hydroxyl group of PLP forms a hydrogen bond to the side chain of Asn^{74} from the highly conserved sequence stretch ^{69}EPTSGNTG^{76} and to one water molecule. Related OASS enzymes (14–17) show basically identical interaction networks with the PLP cofactor.

The Active Site in the CysK1 Holoenzyme—The active site of the CysK1 holoenzyme contains several ligands from the crystallization solution. An MPD molecule is bound in the active site pocket and, with the exception of a hydrogen bond to carbonyl oxygen of Gly^{222}, mostly forms hydrophobic interactions with residues lining the pocket. The electron density map further contained a strong spherical electron density close to the internal aldimine linkage and the conserved loop (^{71}TSGNT^{75}) at the N terminus of helix α2, which was modeled as a chloride ion, mostly likely originat-

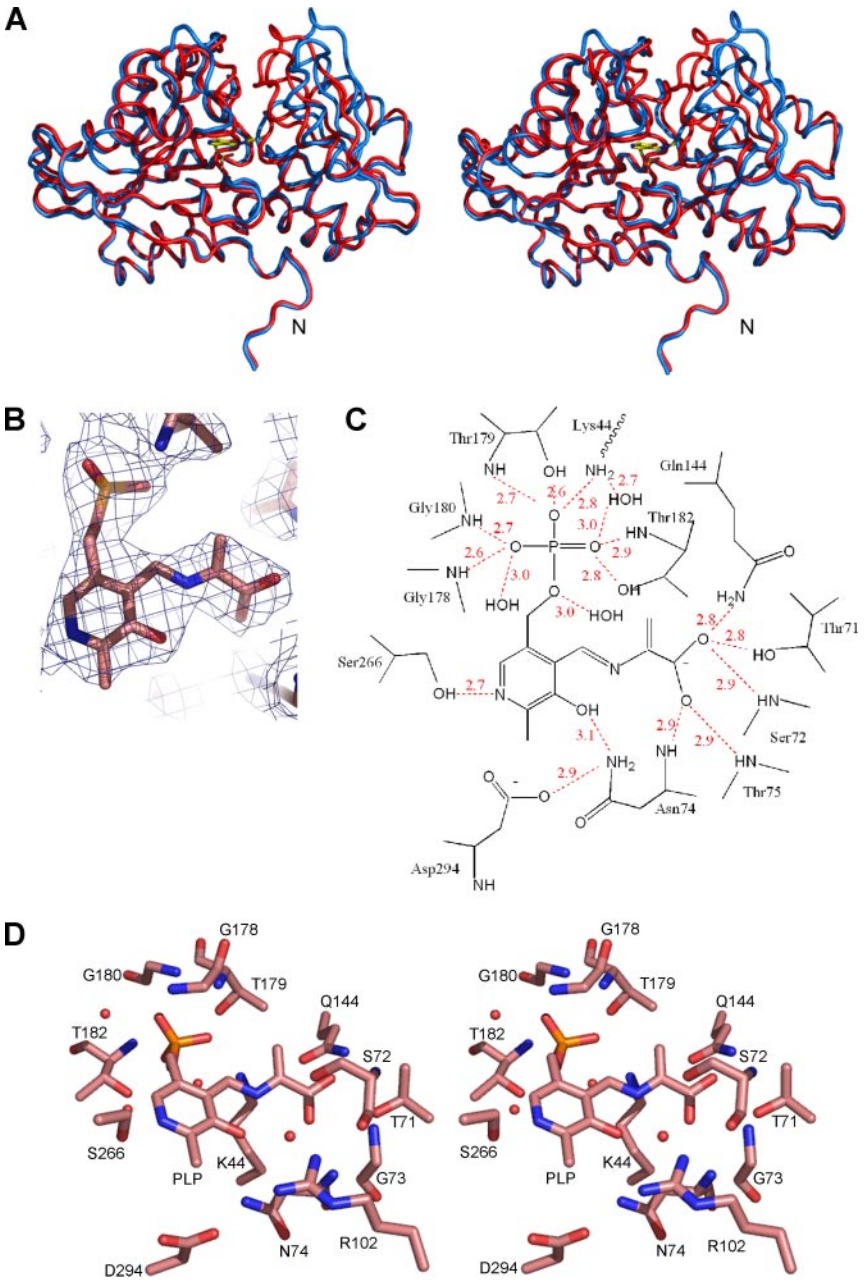


FIGURE 3. **Structure of the complex of CysK1 with α-aminoacrylate.** *A*, superposition of holo-CysK1 (*blue*) and the CysK1-aminoacrylate complex (*red*) illustrating the conformational changes of the N-terminal domain leading to the closure of the active site. The covalent intermediate forming the external Schiff base with PLP is shown as a *stick model* (*yellow*). *B*, part of the $2F_o - F_o$ electron density map, contoured at 1.2σ, showing formation of the α-aminoacrylate intermediate. *C*, schematic showing the interactions of the reaction intermediate with enzyme residues in the active site. Hydrogen bonds are indicated by *dashed lines*. *D*, stereo view of the active site of the CysK1-α-aminoacrylate intermediate complex. The *red spheres* indicate the positions of bound water molecules.

ing from the crystallization buffer (Fig. 2, *C* and *D*). A similar anion binding site has also been found in other *O*-acetylserine sulfhydrylases (14, 17).

Structure of the CysK1-α-Aminoacrylate Complex—The formation of the aminoacrylate Schiff base intermediate, the product of the first half-reaction, was observed in solution by spectrophotometry. The UV-visible spectrum of CysK1 in the presence of 300 μM OAS (Fig. 1) showed the absorbance maxima at 321 and 469 nm, which are characteristic for the formation of the α-aminoacrylate intermediate (23, 38). The subsequent addition of 0.25 mM sodium sulfide resulted in the fast disappearance of the band at 469 nm, indicating decomposition of the intermediate.

The ~60 nm shift of the absorbance maximum in the spectrum corresponds to a visible color change, which was also observed in the crystals during soaking experiments with OAS. Diffraction data collected from such a crystal after arresting the reaction by flash-freezing were used to solve the structure of CysK1 with the trapped α-aminoacrylate intermediate. Most notably, parts of the N-terminal domain had moved significantly upon formation of the reaction intermediate and had to be rebuilt completely. The largest conformational changes (maximal atomic displacements of 6 Å) involve the loop region consisting of the conserved sequence stretch ^{71}TSGNT^{75}, the loop between strand β5 and helix α3 (residues 94–101) and between strand β6 and helix α4 (residues 117–124) that fold over the active site (Fig. 3*A*). The changes in cell dimensions in the crystals containing the reaction intermediate most likely reflect the conformational changes of this part of the polypeptide chain. The domain rotation (Fig. 3*A*) results in the closure of the active site, which becomes less accessible and thus protects the reaction intermediate from solvent.

The electron density map clearly showed that the internal Schiff base linkage between PLP and the catalytic lysine residue was broken and that the external Schiff base with the α-aminoacrylate had formed in the crystal (Fig. 3*B*). We did not observe any electron density beyond the Cβ-carbon atom of the external Schiff base, indicating that the acetate moiety had been expelled from the substrate, OAS. The carbon atoms of the aminoacrylate moiety and the nitrogen involved in the Schiff base linkage are planar, as expected for the aminoacrylate intermediate. One of the oxygen atoms of the carboxylate group is within hydrogen-bonding distance to the side chains of Gln^{144} and Thr^{71} and the main chain nitrogen atom of Ser^{72}. The second oxygen atom is tethered to the backbone amide groups of Asn^{74} and Thr^{75} (Fig. 3, *C* and *D*). Most of these residues are

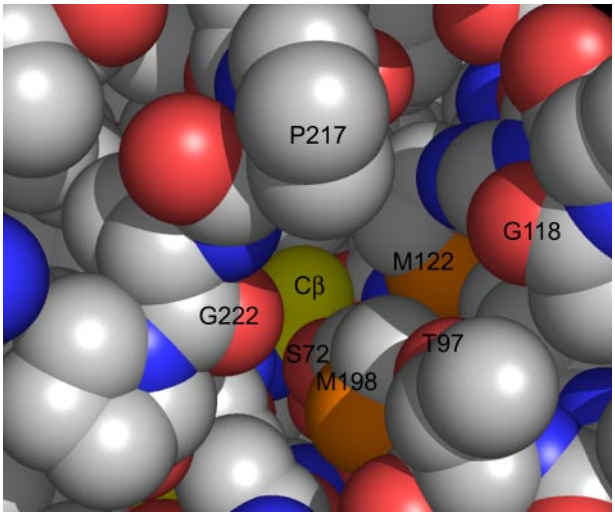


FIGURE 4. **Space-filling model of CysK1, highlighting the narrow tunnel that leads to the active site.** The only solvent-accessible atom is the Cβ-carbon atom of the α-aminoacylate-intermediate, shown in *yellow*.

part of the anion binding site in the holoenzyme, but as a result of the loop movement, they are closer to the active site and interact with the carboxyl group of the aminoacrylate.

The formation of the reaction intermediate results in a number of subtle changes in the active site. The pyridine ring of PLP is tilted by ~19° along an axis passing through the 2′-5′-carbon atoms. A small shift in the position of the phosphate group was found (0.5 Å), which leads to slightly different interactions when compared with the structure of the holoenzyme. One of the phosphate oxygen atoms forms a hydrogen bond with the ϵ-amino group of the catalytic residue Lys^{44}, and it is likely that the protonated state of the Lys^{44}, crucial in the second half-reaction, is ensured by this interaction. Another phosphate oxygen atom acquires a water molecule as a new hydrogen-bonding partner. In total, the number of hydrogen bonds of the phosphate group increases from seven to nine and is thus firmly anchored in the active site.

In the structure of the CysK1-aminoacrylate complex, the Cβ-carbon atom points into the active site pocket and is the only atom of the intermediate accessible for a nucleophilic attack by hydrogen sulfide through a narrow tunnel extending from the enzyme surface (Fig. 4). The ϵ-amino group of Lys^{44} is within 4.3 Å distance of the Cα-atom of the aminoacrylate but can easily adopt a position suitable for the proposed role in protonation of the Cα-carbon atom (20, 24) by a simple rotation along the Cδ-Cϵ carbon bond. Lys^{44} is positioned such that protonation of the Cα-carbon can only occur from the *si*-face of the intermediate. The channel from the surface leading to the Cβ-carbon atom (Fig. 4) directs the nucleophilic attack of the sulfide toward the *re*-face of the intermediate, leading to an *anti*-addition, resulting in the formation of L-cysteine.

In previous attempts to obtain structures of *O*-acetylserine sulfhydrylases with bound substrate and intermediates, active site mutants where the catalytic lysine residue had been replaced by alanine were employed. The structure analysis of such mutants from *S. typhimirium* (25) and *A. thaliana* (17) revealed that an external aldimine had formed via a reaction with a methionine residue during protein production (21). This structure represented a mimic of the first intermediate, the species preceding the formation of the α-aminoacrylate in the reaction cycle of OASS (compound **II** in Scheme 1). Although the methionine complex of OASS from *A. thaliana* (crystallized in 4.5 M sodium formate) did not show any major structural rearrangements upon ligand binding, the *S. typhimirium* (crystallized at pH 7.0, Tris buffer) enzyme showed a similar conformational change as seen in CysK1 from *M. tuberculosis*. The external methionine Schiff base adduct in *S. typhimurium* OASS and the aminoacrylate Schiff base in *M. tuberculosis* CysK1 overlap well upon superposition, in particular the Cα, Cβ-carbon atoms and the carboxyl group (supplemental Fig. S2*A*). Consistent with the key role of Thr^{71}, Asn^{74}, Ser^{72}, and Gln^{144} in recognizing and binding the carboxyl group of the aminoacrylate intermediate as suggested by the structure analysis, replacement of these amino acids in OASS from *A. thaliana* leads to mutants with a 10^3-10^5-fold lower catalytic efficiency than wild-type enzyme (17).

It is of interest to note that in tryptophan synthase from *S. typhimurium,* the only other member of the fold type II PLP enzymes with a known structure of an aminoacrylate external Schiff base (42), the interactions of the intermediate with the enzyme are rather different from that in CysK1 and are not conserved (supplemental Fig. S2*B*). Interactions are made to the side chain of Thr^{110} and backbone nitrogens from Ala^{112}, Gly^{113}, and Gln^{114}. These residues are structurally equivalent to the substrate binding loop (residues 71–75) of OASS enzymes but are not conserved. A notable difference in the recognition and binding of the aminoacrylate moiety in CysK1 and tryptophan synthase is the absence of Gln^{144} in the latter. The opening of the hydrophobic channel that connects the two active sites in the heterodimeric tryptophan synthase is located at approximately this position and allows access of indole, formed in the α-subunit, to the PLP cofactor.

M. tuberculosis CysK1 Interacts with a SAT-derived C-terminal Peptide—OASS associates with SAT to form a bi-enzyme complex, in which the OASS activity is down-regulated (43, 44). Crystallographic studies of OASS from *H. influenzae* and *A. thaliana* (15, 45) have shown that an essential part of the interactions between the two enzymes most likely involves binding of the C-terminal four residues of SAT to the active site of OASS. We therefore addressed the question whether CysK1 from *M. tuberculosis* is able to form similar interactions with SAT and whether binding of a peptide derived from the C terminus of *M. tuberculosis* SAT (CysE) would lead to inhibition of CysK1. The addition of a peptide comprising the last four residues (DFSI) of CysE to CysK1 gave a shift in the thermal denaturation curve of the enzyme by 5 °C. Inhibition studies showed the competitive nature of the inhibition with a K_I value of 5.0 μM (Fig. 5). These observations indicated formation of an enzyme-peptide complex, and the DFSI peptide was therefore used for co-crystallization.

The electron density maps calculated from x-ray data collected on crystals of CysK1 obtained in the presence of 1.0 and 4.0 mM DFSI peptide clearly showed that the peptide was indeed bound to CysK1 (Fig. 6, *A* and *B*). The peptide is bound in the active site cleft between the two domains and extends from the enzyme surface into the interior of the active site. It

blocks the active site cleft completely and also acts like a wedge, preventing complete domain rotation seen in the enzyme-intermediate complex. CysK1 is thus trapped in an open conformation similar to the holoenzyme.

The peptide interacts with CysK1 through direct hydrogen bonds as well as hydrophobic interactions (Fig. 6*A*). Two main chain atoms of the peptide are engaged in hydrogen bonds. The carbonyl oxygen of Ser^{3*} (the asterisk denotes residues from the DFSI peptide) interacts with the main chain nitrogen atom from Gly^{222}, and the amide of Ser^{3*} forms an indirect hydrogen bond to the carbonyl oxygen atom of Ile^{223} via a water molecule. Residue Asp^{4*} is firmly held in place through interactions with the side chain of Lys^{215} and the main chain amide of Met^{122}. The side chain of Phe^{2*} binds to a hydrophobic pocket lined by Met^{122}, Phe^{145}, Ala^{225}, and Phe^{227}. Finally, Ile^{4*} is buried in the interior of the active site and packs against Phe^{145}, Gly^{178}, Gly^{222}, and Ala^{225}. The carboxylate group of Ile^{4*} occupies the position of the carboxylate of the aminoacrylate moiety in the CysK1-intermediate complex. Hydrogen bonds are formed to main chain nitrogen atoms from residues of the conserved sequence stretch ^{71}TSGNT^{75}, the serine loop. Additional hydrogen bonds are made to the side chain of Gln^{144}.

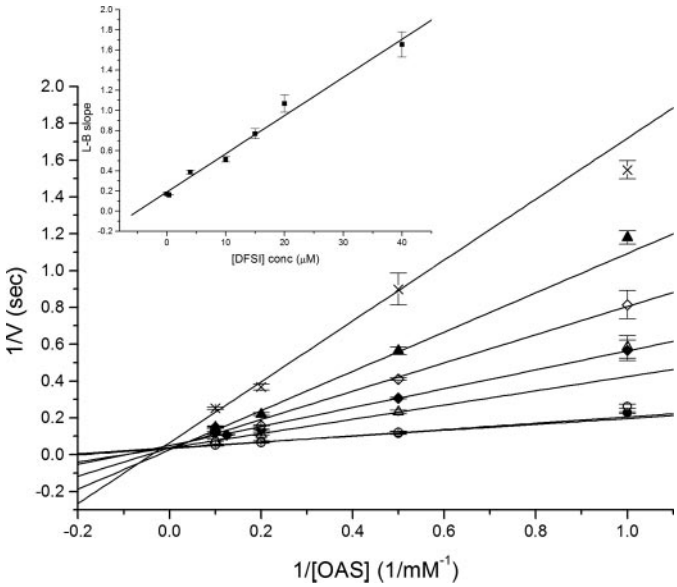


FIGURE 5. **Inhibition of CysK1 by the DFSI peptide.** Double reciprocal plots of the initial rate of the CysK1 activity in the presence of different fixed concentrations of the DFSI peptide illustrating the competitive nature of inhibition. The linear curves were fitted to the data points (average of measurements from three independent reactions) at DFSI concentrations 0.4 μM (●), 4 μM (△), 10 μM (♦), 15 μM (◇), 20 μM (▲), 40 μM (×), and without inhibitor (○). In the *inset*, the values for the slopes of the double reciprocal plots are plotted against the peptide concentrations.

A comparison of the CysK1-DFSI complex with similar peptide complexes of OASS from *H. influenzae* and *A. thaliana* (15, 45) reveals that peptide recognition is mediated via side chain interactions in a sequence-specific manner. The only enzyme-peptide interaction common to the three enzymes is the binding mode of the conserved C-terminal isoleucine residue via specific hydrogen bonds to its carboxyl group and hydrophobic interactions with the side chain. The residues of the remaining peptide residues are either not conserved, or in cases where they are, their interactions with the enzyme are different. For instance, Ser^{3*} corresponds to asparagine in *H. influenzae* or valine in *A. thaliana*, and the hydrogen bond formed between the side chains of Ser^{3*} and Ser^{72} is specific for the *M. tuberculosis* Cysk1-peptide complex. The N-terminal peptide residue Asp^{1*}, although conserved in *A. thaliana*, has a different conformation and interacts with different enzyme residues.

In conclusion, the strong inhibition and the unique interactions of the peptide with CysK1 suggest this complex to be an attractive scaffold for the rational design of CysK1 inhibitors

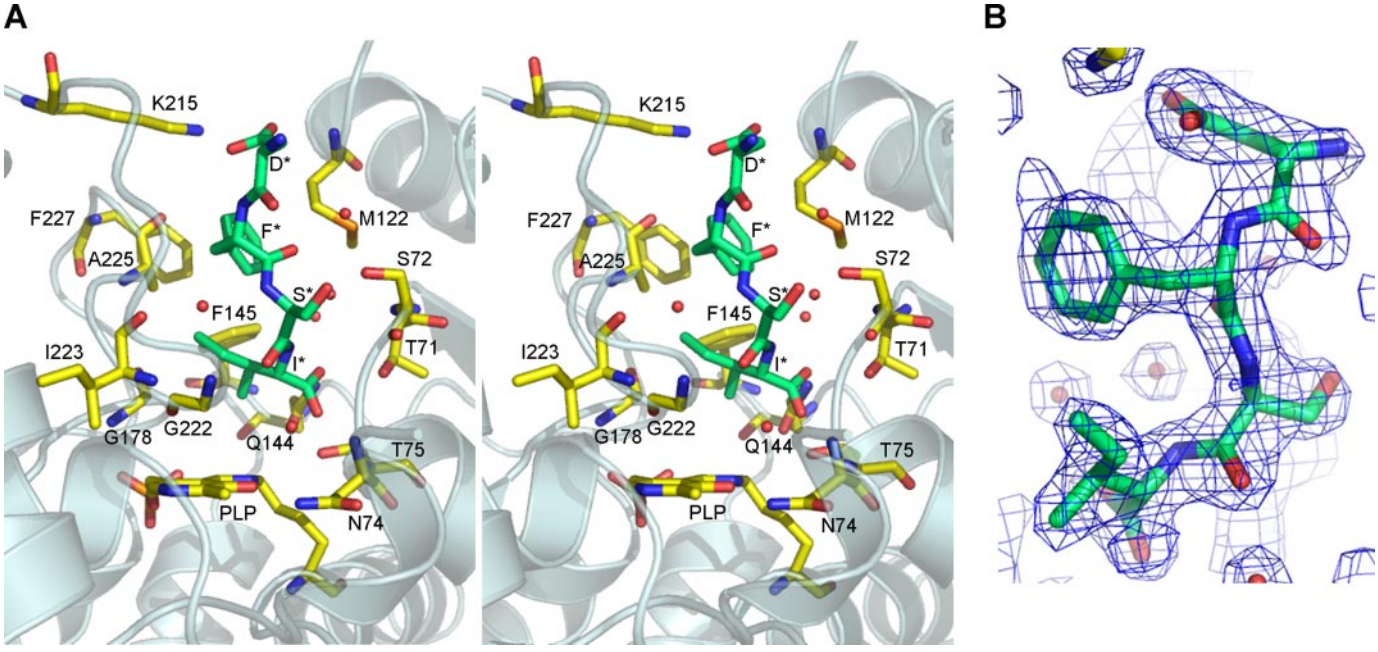


FIGURE 6. **Structure of the CysK1-peptide complex.** *A*, stereo view of the active site of CysK1 with the bound DFSI peptide, shown in *green*. CysK1 amino acid residues are shown in *yellow*, and water molecules are shown as *red spheres*. *B*, part of the $2F_o - F_o$ electron density map at the peptide binding site in CysK1, contoured at 1.4σ.

specific for *M. tuberculosis*. Furthermore, we have shown that CysK1 is able to interact with the C-terminal peptide from SAT in a similar manner as seen in OASSs, where formation of a bi-enzyme complex has been shown. This suggests that the *M. tuberculosis* enzyme may form such a complex *in vivo* and be subject to a similar mechanism of regulation.

Acknowledgments—We gratefully acknowledge access to synchrotron radiation at the ESRF, Grenoble, France and MAX-lab, Lund University, Sweden.

REFERENCES

1. Wooff, E., Michell, S. L., Gordon, S. V., Chambers, M. A., Bardarov, S., Jacobs, W. R., Jr., Hewinson, R. G., and Wheeler, P. R. (2002) *Mol. Microbiol.* **43,** 653–663
2. Manganelli, R., Voskuil, M. I., Schoolnik, G. K., Dubnau, E., Gomez, M., and Smith, I. (2002) *Mol. Microbiol.* **45,** 365–374
3. Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butche, P. D., Nathan, C., and Schoolnik, G. K. (2003), *J. Exp. Med.* **198,** 693–704
4. Hampshire, T., Soneji, S., Bacon, J., James, B.W., Hinds, J., Laing, K., Stabler, R.A., Marsh, P.D., and Butcher, P.D. (2004) *Tuberculosis* (*Edinb.*) **84,** 228–238
5. Pinto, R., Tang, Q. X., Britton, W. J., Leyh, T. S., and Triccas, J. A. (2004) *Microbiology* (*Read.*) **150,** 1681–1686
6. Wayne, L. G., and Sohaskey, C. D. (2001) *Annu. Rev. Microbiol.* **55,** 139–163
7. McKinney, J. D., Bentrup, K., Munoz-Elias, E. J., Miczak, A., Chen, B., Chan, W. T., Swenson, D., Sacchettini, J. C., Jacobs, W. R., Jr., and Russell, D. G. (2000) *Nature* **406,** 735–738
8. Rhee, K. Y., Erdjument-Bromage, H., Tempst, P., and Nathan, C. F. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 467–472
9. Bornemann, C., Jardine, M. A., Spies, H. S., and Steenkamp, D. J. (1997) *Biochem. J.* **325,** 623–629
10. Tai, C-H., and Cook, P. F. (2001) *Acc. Chem. Res.* **34,** 49–59
11. Rabeh, W. M., and Cook, P. F. (2004) *J. Biol. Chem.* **279,** 26803–26806
12. Nakamura, T., Iwahashi, H., and Eguchi, Y. (1984) *J. Bacteriol.* **158,** 1122–1127
13. Maier, T. H. (2003) *Nat. Biotechnol.* **21,** 422–427
14. Burkhard, P., Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F., and Jansonius, J. N. (1998) *J. Mol. Biol.* **283,** 121–133
15. Huang, B., Vetting, M. W., and Roderick, S. L. (2005) *J Bacteriol.* **187,** 3201–3205
16. Claus, M. T., Zocher, G. E., Maier, T. H., and Schulz, G. E. (2005) *Biochemistry* **44,** 8620–8626
17. Bonner, E. R., Cahoon, R. E., Knapke, S. M., and Jez, J. M. (2005) *J. Biol. Chem.* **280,** 38803–38813
18. Grishin, N. V., Phillips, M. A., and Goldsmith, E. J. (1995) *Protein Sci.* **4,** 1291–1304
19. Schneider, G., Käck, H., and Lindqvist, Y. (2000) *Structure* (*Lond.*) **8,** R1–R6
20. Rabeh, W. M., Alguindigue, S. S., and Cook, P. F. (2005) *Biochemistry* **44,** 5541–5550
21. Rege, V. D., Kredich, N. M., Tai, C. H., Karsten, W. E., Schnackerz, K. D., and Cook, P. F. (1996) *Biochemistry* **35,** 13485–13493
22. Lu, Z., Nagata, S., McPhie, P., and Miles, E. W. (1993) *J. Biol. Chem.* **268,** 8727–8743
23. Schnackerz, K. D., Ehrlich, J. H., Giesemann, W., and Reed, T. A. (1979) *Biochemistry* **18,** 3557–3563
24. Tai, C. H., Nalabolu, S. R., Simmons, J. W., III, Jacobson, T. M., and Cook, P. F. (1995) *Biochemistry* **34,** 12311–12322
25. Burkhard, P., Tai, C. H., Ristroph, C. M., Cook, P. F., and Jansonius, J. N. (1999) *J. Mol. Biol.* **291,** 941–953
26. Gaitonde, M. K. (1967) *Biochem. J.* **104,** 627–633
27. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N., and Nordlund, P. (2006) *Anal. Biochem.* **357,** 289–298
28. Cornish-Bowden, A. (2004) *Fundamentals of Enzyme Kinetics*, Portland Press Ltd., London, UK
29. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
30. Vagin, A., and Teplyakov, A., (1997) *J. Appl. Crystallogr.* **30,** 1022–1025
31. Murshudov, G., Vagin, A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–253
32. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132
33. Read, R. J. (1986) *Acta Crystallogr. Sect. A* **42,** 140–149
34. Painter, J., and Merritt, E. A. (2006) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62,** 439–450
35. Painter, J., and Merritt, E. A. (2006) *J. Appl. Crystallogr.* **39,** 109–111
36. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. (1993) *J. Appl. Crystallogr.* **26,** 282–291
37. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
38. Karsten, W. E., and Cook, P. F. (2002), *Methods Enzymol.* **354,** 223–237
39. Hwang, C. C, Woehl, E. U., Minter, D. E., Dunn, M. F., and Cook, P. F. (1996) *Biochemistry* **35,** 6358–6365
40. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) *J. Biol. Chem.* **263,** 17857–17871
41. Gallagher, D. T., Gilliland, G. L., Xiao, G., Zondlo, J., Fisher, K. E., Chinchilla, D., and Eisenstein, E. (1998) *Structure* (*Lond.*) **6,** 465–475
42. Schneider, T. R., Gerhardt, E., Lee, M., Liang, P. H., Anderson, K. S., and Schlichting, I. (1998) *Biochemistry* **37,** 5394–5406
43. Droux, M., Ruffet, M. L., Dounce, R., and Job, D. (1998) *Eur. J. Biochem.* **255,** 235–245
44. Mino, K., Hiraoka, K., Imamura, K., Sakiyama, T., Eisaki, N., Matsuyama, A., and Nakanishi, K. (2000) *Biosci. Biotechnol. Biochem.* **64,** 1874–1880
45. Francois, J. A., Kumaran, S., and Jez, J. M. (2006) *Plant Cell* **18,** 3647–3655