STRUCTURAL, BIOCHEMICAL AND *IN VIVO* INVESTIGATIONS OF THE THREONINE SYNTHASE FROM *MYCOBACTERIUM TUBERCULOSIS**

Adrian Suarez Covarrubias[1], Martin Högbom[1Ω], Terese Bergfors[1], Paul Carroll[2], Karin Mannerstedt[3], Stefan Oscarson[3], Tanya Parish[2], T. Alwyn Jones[1] and Sherry L. Mowbray[4]*

[1]Department of Cell & Molecular Biology, Uppsala University, SE-751 24 Uppsala, Sweden

[2]Institute for Cell and Molecular Science, Barts and the London School of Medicine and Dentistry, London, UK,

[3]Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91, Stockholm, Sweden,

[4]Department of Molecular Biology, Swedish University of Agricultural Sciences, SE-751 24 Uppsala, Sweden.

[Ω] Current address: Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, SE-10691 Stockholm, Sweden

Running title: Threonine synthase from *Mycobacterium tuberculosis*

*Corresponding author: Sherry Mowbray, Department of Molecular Biology, Swedish University of Agricultural Sciences, Biomedical Center, Box 590, S-751 24 Uppsala, Sweden, Fax. +46-18-536971; E-mail: mowbray@xray.bmc.uu.se

Data deposition: The atomic coordinates and structure factors have been deposited at the Protein Data Bank, www.pdb.org (PDB ID code 2D1F).

Summary

Threonine biosynthesis is a general feature of prokaryotes, eukaryotic microorganisms and higher plants. Since mammals lack the appropriate synthetic machinery, instead obtaining the amino acid through their diet, the pathway is a potential focus for the development of novel antibiotics, anti-fungal agents and herbicides. Threonine synthase, a pyridoxal 5-phosphate-dependent enzyme, catalyses the final step in the pathway, in which L-homoserine phosphate and water are converted into threonine and inorganic phosphate. In the present publication we report structural and functional studies of *Mycobacterium tuberculosis* threonine synthase, the product of the *rv1295* (*thrC*) gene. The structure gives new insights into the catalytic mechanism of threonine synthases in general, specifically by suggesting the direct involvement of the phosphate moiety of the cofactor, rather than the inorganic phosphate product, in transferring a proton from C4' to Cγ in the formation of the α,β-unsaturated aldimine. It further provides a basis for understanding why this enzyme has a higher pH optimum than has been reported elsewhere for threonine synthases, and gives rise to the prediction that the equivalent enzyme from *Thermus thermophilus* will exhibit similar behavior. A deletion of the relevant gene generated a strain of *M. tuberculosis* that requires threonine for growth; such auxotrophic strains are frequently attenuated *in vivo*, indicating that threonine synthase is a potential drug target in this organism.

Keywords

Threonine biosynthesis, tuberculosis, enzyme mechanism, X-ray structure, drug target

Introduction

Tuberculosis was viewed as largely under control a century ago, but it has reemerged as a public health burden of enormous proportions. The World Health Organization (http://www.who.org) estimates that 1.8 billion people are currently infected with *Mycobacterium tuberculosis,* the microorganism causing the disease, and that there are 9 million new active cases and 2 million deaths resulting each year. The current "short-term" treatment for tuberculosis lasts a minimum of 4 months. Given the expense and difficulties in successfully completing such cures (partly due to their unpleasant side effects) and the increasing incidence of drug-resistant strains, additional targets for drug design are urgently required. Resistance is a particularly frightening problem, given the fact that no new anti-tubercular drugs have been brought to market since the early 1960's.

One critical step in the process of devising new treatments for tuberculosis is the identification of essential proteins or pathways that are valid targets for drug design. The determination of the genome sequence of *M. tuberculosis*[1] provided an enormous boost to the effort, since about 60% of the encoded proteins' functions could thereby be assigned. Subsequent studies have attempted to identify genes that are likely to be required for bacterial survival and infection.[2,3] We have used such investigations, together with other biological data, to select *M. tuberculosis* gene products that are interesting subjects for detailed structural and functional studies.

The *M. tuberculosis rv1295* gene (also designated as *thrC* in that organism) is annotated at the Pasteur Institute's TubercuList server (http://genolist.pasteur.fr/TubercuList) as being essential by Himar1-based transposon mutagenesis in the H37Rv strain[3] and as coding for a probable threonine synthase (TS; EC 4.2.3.1). TSs catalyze the last step in the threonine biosynthetic pathway, in which (2*S*)-*O*-phosphohomoserine (OPH) and water are converted into threonine and inorganic phosphate *via* a pyridoxal 5-phosphate (PLP) dependent mechanism. The reaction has been described as one of the most sophisticated of those involving PLP, as it is believed to proceed through all of the known intermediates.[4,5] Since threonine biosynthesis is absent in mammals (which can instead obtain the amino acid from their diet), but is a general feature of prokaryotes, eukaryotic microorganisms and higher plants,

TS is an interesting target for the development of novel antibiotics, anti-fungal agents and herbicides. Some substrate analogues have previously been tested as putative mechanism-based inhibitors.[5-7]

TSs are categorized as PLP-dependent enzymes of fold-type II;[8-10] they have further been divided into class I and class II subfamilies based on multiple sequence alignments.[6] Class I enzymes include TSs from higher plants, cyanobacteria, archaebacteria, and eubacterial *Mycobacteria, Aquificaceae* and *Bacillus* species. In addition to the catalytic regions, plant class I TSs have an additional N-terminal domain that mediates their allosteric activation by S-adenosylmethionine. Class II comprises enzymes from fungi and the eubacterial groups of *Proteobacteria* and coryneform bacteria, which are characterized by a C-terminal extension of about 50 residues.

Four three-dimensional structures of TS are currently available. Two represent class I enzymes, from *Arabidopsis thaliana* (AtTS[11]) and *Thermus thermophilus* HB8 (TtTS[12]). Two others, the enzymes from *Saccharomyces cerevisiae* (ScTS[13]) and *Escherichia coli* (EcTS, an as yet unpublished structure found in the PDB with entry code 1VB3), belong to class II. In the present publication, we report structural and functional studies of the class I enzyme encoded by the *M. tuberculosis rv1295* (*thrC*) gene, MtTS.

Results and Discussion

Biochemical studies

His-tagged protein corresponding to the full-length *rv1295* (*thrC*) gene (which encodes for 367 residues) was over-expressed in *E. coli* and purified. The yellow sample was subjected to UV-visible spectroscopy, which indicated that it contained bound PLP, as deduced from the presence of absorption peaks near 335 and 411 nm (Figure 1(a)). The latter peak was much larger, which was presumed to reflect a predominance of the protonated (ketoenamine) over unprotonated (enolimine) form of the Schiff base that links PLP to an active-site lysine in the apoenzyme. The spectra were not significantly affected by changes in pH, consistent with earlier reports for both ScTS[13] and TtTS.[12] Although spectra were not shown for ScTS, those for TtTS exhibited an even larger proportion of the 411 nm peak than was observed here. By contrast, AtTS[14] and EcTS[15] spectra were reported to be sensitive to pH.

Since the *rv1295* gene was annotated as encoding a putative TS in Tuberculist, but no previous investigations had been directed at the actual protein, a brief examination of the kinetic properties was conducted. This was considered to be particularly important in light of close relationships at the sequence level to a number of non-TS enzymes (see below). OPH was therefore synthesized starting from benzyl (*2S*)-2-(benzyloxycarbonyl)amino butanoate using a modification of previous methods.[16] Phosphorylation of the free hydroxyl group using phosphorus trichloride and imidazole followed by addition of benzyl alcohol and *in situ* oxidation with *meta*-chloroperbenzoic acid produced benzyl-(*2S*)-*N*-(benzyloxycarbonyl)-phosphohomoserine dibenzyl ester (72% yield), which was deprotected by catalytic hydrogenolysis and purified by C18 reversed phase silica gel chromatography to give the final product (99% yield).

The production of inorganic phosphate from OPH showed that MtTS is an active TS. At pH 8.4 and 25 °C, the K_m was determined to be 1.2 ± 0.1 mM, and the specific activity was 1.05 ± 0.15 μmol.min^{-1}.mg^{-1} (corresponding to a k_{cat} of 0.13 s^{-1}), well in line with published data for other TS enzymes. For example, TtTS is reported to have a specific activity of 4.25 μmol.min^{-1}.mg^{-1} at 70 °C, pH 7.0,[12] while EcTS[15] at pH 8.0 exhibits a K_m of 0.5 mM, and a specific activity of 9.3 μmol.min^{-1}.mg^{-1}, which corresponds to a k_{cat} of 7.3 s^{-1}. However, comparison of catalytic constants is complicated by the fact that MtTS activity is strongly influenced by pH (Figure 1(b)). pH optima have been reported to be 7.5, 8.0 and 8.2 for EcTS,[7] AtTS[14] and ScTS,[13] respectively. MtTS, by contrast, is inactive at pHs of 7 or less, and has greatest activity at high pH; controls indicated that the effect was due to an enzymatic rather than non-enzymatic mechanism. Similar results were obtained with different substrate concentrations, including the highest attainable (5 mM, near-saturating) OPH concentration, suggesting that the effect is on k_{cat}, rather than K_m. We conclude that the optimum pH for a given TS is not correlated with the class I versus II distinctions, and instead is determined by features of its particular active site. Thus MtTS activity seems to be dependent on the presence of an active-site group with a very high pK_a; this group seems to be distinct from the protonated form of the Schiff base, since that is unchanged in the absorbance spectra at various pHs.

As would be predicted from the lack of an N-terminal extension, MtTS activity is not influenced by the addition of S-adenosylmethionine, the allosteric activator of AtTS.

MtTS loses less than 10% of the initial activity during a 10-minute incubation at 65 °C. At 80 °C, 90% of the activity was lost over the same time period. We conclude that the enzyme shows good temperature stability, given that *M. tuberculosis* is a mesophilic microorganism. However, we were unable to find stability data for other TSs for comparison.

Structure of MtTS

MtTS crystals possessed the symmetry of space group $P6_1$ with two molecules expected in the asymmetric unit. The structure was solved by molecular replacement, using a homology model based on TtTS (PDB code 1UIN[12], 57 % amino acid sequence identity) in the rotation and translation searches. The final model, refined to 2.5 Å resolution, comprises 696 amino acids, 54 water molecules and two PLP molecules. The final R_{factor} and R_{free} were 18.7 and 25.6%, respectively. In both protein molecules in the asymmetric unit, density was interpretable for residues 10 through 358. Other details concerning data collection, refinement and stereochemistry of the final model are presented in Table 1.

Overall structure

The structure of MtTS has a homodimeric organization in which the two subunits are related by a non-crystallographic 2-fold axis (Figure 2). Each subunit is composed of three domains. Following the conventions established by previous authors, the largest domain is an α/β sandwich built by two segments of the polypeptide chain, i.e. residues Gln10-Gly84 and Val165-Pro341. The small domain is an insertion in the larger one, and also possesses an α/β architecture; it consists of residues Gln85 to Ser164. At the C-terminal end of the molecule, the so-called swap domain is made up of amino acids Ser342 to Gly358. This unit provides an additional strand and helix, which pack onto the small domain of the other subunit. Each subunit buries approximately 3200 $Å^2$ in the dimer interaction (20% of the protein surface), which is at the high end of the range observed for dimers of similar size and complexity.[17] The subunit surfaces are highly complementary in shape; the interface is calculated to have a gap volume index of 1.5,

significantly better than the 2.2 average observed for homodimers.[17] The surface is quite hydrophobic, including 72% nonpolar residues. There were relatively few hydrogen bonds, and no salt links. The combined observations suggest that MtTS is found only as a dimer in solution, consistent with the approximate size of 70 kDa estimated from gel filtration. The two subunits of the asymmetric unit can be superimposed with a root-mean-square difference of 0.6 Å when all Cα atoms are compared, indicating that their structures are fundamentally the same. The non-crystallographic symmetry is well preserved, except for the helix centered on residue 150, which shows some small differences that are probably due to crystal lattice contacts.

PLP binding site

Although PLP was not added during the purification or crystallization, electron density corresponding to the cofactor was observed between the large and small domains in each subunit (Figure 2). PLP is covalently connected by a Schiff-base linkage to Lys69 (Figure 3). Further, its pyridine nitrogen (N1) is hydrogen bonded to the hydroxyl group of Thr326, while the C3-hydroxyl group interacts with the amide group of Asn95. That it is the amide nitrogen rather than the oxygen that interacts with PLP cannot be determined from the present structure. However, the ambiguity is resolved by two hydrogen bonds observed in a TsTS structure described below. On this basis, the oxygen of Asn95 is expected to interact with Lys331, and the nitrogen with Glu296 in MtTS. The phosphate group is coordinated by hydrogen bonds to the backbone nitrogen atoms of Gly196, Asn197, Ala198, and Asn200; the placement of these groups at the N-terminal end of an α-helix is well-suited to the phosphate's negative charge. Water-mediated hydrogen bonds between the phosphate and Ile201-N and Pro194-O are also observed. Van der Waals interactions are contributed by the side chains of Phe68 and Ala249, and by Gly327. All of these interacting residues are drawn from the large domain except Asn95, which is part of the small domain.

Although not restrained as such in refinement, electron density in both active sites suggests that the C3-C4-C4'-NZ unit is nearly planar, i.e. the internal aldimine bond of the Schiff base is part of the conjugated π system of the PLP ring. Thus, consistent with the spectroscopic results, the nitrogen atom of the Schiff base is expected to be protonated and involved in a hydrogen bond with O3.

Comparison with other structures

In comparisons between MtTS and structures in the PDB using DALI,[18] the class I threonine synthase, TsTS, gave the highest Z-scores. Three different crystal structures of TtTS have been reported, all with bound PLP.[12] One corresponds to the enzyme in a covalent complex with the substrate analogue L-(+)-2-amino-5-phosphonopentanoic acid (AP5), seen in PDB entry 1V7C. The present MtTS structure is most similar to the other two TtTS structures, 1IUN and 1IUM, which lack the inhibitor; the former (shown in Table 2 and Figure 3) was used in the molecular replacement here. Large movements are observed in TtTS when it binds to, and reacts with, the substrate analogue. The small domain, moving together with the associated swap domain from the other subunit, rotates by ~25°, closing the active site cleft. The orientation of the PLP changes slightly, as the lysine side chain is released, and a covalent bond is formed between the cofactor and AP5. The ester oxygen of the OPH substrate is replaced by a methylene group in the inhibitor, which prevents further reaction. Most interactions between the PLP and the large domain are preserved after the domain rotation; additional smaller changes, primarily in side chain conformations, help shape the substrate-binding site.

Using the TtTS-AP5 complex, we can predict the interactions between MtTS and substrate. As is typical for fold-type II PLP enzymes, all active site residues are derived from a single subunit. One of the α-carboxylate oxygens of OPH is expected to interact with backbone amide nitrogens of residues 95 and 96. The other carboxylate oxygen will hydrogen bond to the side-chain hydroxyl group of Ser92 as well as the backbone amide nitrogen of Thr93. The main-chain conformation of the latter must change significantly for the interaction to occur, as was seen for TtTS. Thr96 will interact with the ester oxygen as well as the phosphate group of substrate. Additional interactions with the phosphate group will be provided by Lys69, Asn163, backbone-nitrogen and side-chain hydroxyl groups of Ser164, Arg169 and Asn197.

Another class I threonine synthase, AtTS, gave the next highest Z-score (Table 2). While PLP was observed in the AtTS structure,[20] the binding of the substrate was modeled, rather than experimentally determined. Most features of AtTS are, however, expected to resemble those of TtTS.

Interestingly, MtTS is more similar in structure to several PLP-dependent enzymes with non-TS function than it is to the class II enzymes EcTS and ScTS (Table 2). Indeed, a BLAST search[21] with the sequence of MtTS did not even detect the presence of ScTS in the database. EcTS and ScTS share 31% sequence identity, and both are monomeric.[6,22] Monomers are rare among PLP-dependent enzymes; however, given that all residues of an active site are drawn from a single molecule, a dimer does not appear to be a prerequisite for function. The structural elements corresponding to the swap domain of the class I enzymes are present in the class II enzymes. However, they are derived from a different region of the sequence, and from the same molecule rather than the second subunit of the dimer. These and other fundamental differences in the structures make sequence alignments (and associated calculations of sequence identity) between class I and class II enzymes problematic; the alignments are by definition flawed.

The EcTS structure is a complex with the substrate-like inhibitor 2-oxo-5-phosphonopentanoic acid and a sulfate ion; no PLP is bound. Like the AP5-bound form of TtTS, the EcTS structure exhibits a closed form of the enzyme. The inhibitor binds at the same position as the very similar moiety in the TtTS-AP5 complex, while the sulfate ion binds at the position of the phosphate group of the PLP. The ScTS structure has only PLP bound. Comparison of these structures with the class I enzymes shows that the positions of bound PLP and substrate are equivalent in the two classes and that their active site residues are well conserved. Indeed, the PLP binding site and orientation appear to be conserved for all fold-type II PLP enzymes.

Mechanism

The expected reaction mechanism of MtTS is outlined in Figure 4. First, OPH binds to the open PLP-bound form seen in the current structure. A conformational change ensues that gives rise to a closed form similar to that seen for the inhibitor-bound structures of TtTS and EcTS. OPH approaches the *re*-face of PLP, with its α-amino group directed towards the Schiff base that links the cofactor and Lys69. The nucleophilic attack of the α-amino group of the substrate on C4' of the cofactor results in the formation of an external aldimine, and releases the neutral side chain of Lys69. Lys69 then abstracts the proton from

the Cα of OPH, to produce the carbanion intermediate, as proposed by Omi *et al.*[12] It is thought that the most important function of the cofactor is to act as an electron sink to delocalize, and thereby stabilize, the negative charge of this intermediate. Lys69 now donates its proton to C4' of the cofactor, to give the ketimine. The ε-amino group of Lys69 is again neutral, and so can abstract the pro-R hydrogen from the Cβ atom of the substrate to produce the enamine. Lys69 then transfers this proton to the phosphate, allowing its γ-elimination, and yielding a β,γ-unsaturated ketimine.

The proton from C4' must now be transferred to Cγ to form the α,β-unsaturated aldimine. If the mode of binding of the inhibitors reflects that of the substrate, the side chain of the lysine is too distant (~ 5Å) from Cγ in the TtTS-AP5 and EcTS structures to perform this transfer, and the geometry is poor. Therefore, Omi *et al.* proposed that the inorganic phosphate acts as a shuttle, in a form of product-assisted catalysis. However, the closest distance from the substrate's phosphate group to C4' is also ~5 Å, and substantial rearrangements in the protein structures seem to be required for the inorganic phosphate to attain the required proximity and geometry for proton transfer between C4' and Cγ. The observed interactions between the phosphate and protein suggest that the phosphate itself is unlikely to move freely. A more satisfying mechanism is suggested to us by the fact that the 5'-phosphate of PLP is also near both C4' and Cγ in the TtTS-AP5 structure, with distances of 3.5 and 4.6 Å (as shown in Figure 5), and moreover, the necessary distances and geometry can be attained via movements in the intermediate. Note that at this step, the β,γ-unsaturated ketimine lacks the phosphate group, and so some repositioning is expected in this part of the molecule, which seem likely to push it away from the released phosphate, and toward the PLP phosphate. Another possible candidate for the shuttle is the side chain hydroxyl of Thr93, but this group is poorly positioned for the desired function and is not well conserved in TS sequences. Therefore, it seems more probable that its role is in stabilizing the closed form of the enzyme via interactions to 249-O and a nearby water. Participation of the phosphate group of PLP in acid-base catalysis is not unprecedented. A similar role has been suggested in the action of serine dehydratase,[23] one of the enzymes included in Table 2. Like TS, this enzyme is believed to use the carbanion rather than the quinonoid intermediate, and the arrangement of groups in the active site is remarkably similar to that of

MtTS. In a different structural/functional context, the role of the 5'-phosphate of PLP in glycogen phosphorylase catalysis is well established.[24]

In the next step of the reaction, Lys69 abstracts a proton from a water molecule at the *si*-face side of the cofactor, so forming a hydroxyl ion.[12] In the MtTS structure, this water corresponds to W47 and W12 in molecules A and B, respectively. The hydroxyl ion then attacks the Cβ atom of the substrate, while Lys69 simultaneously donates its proton to Cα. The result is the addition of a water molecule to the Cα–Cβ unsaturated bond, producing the second external aldimine, in which a Schiff base now links PLP and the product threonine. In the final step of the reaction, the internal Schiff base is reformed between PLP and Lys69. Many interactions with the phosphates of OPH and PLP, as well as the carboxylate of OPH, are expected to keep these groups tethered to the enzyme as the situation in the rest of the active site changes during catalysis. Thus, the phosphate product is probably only released when the enzyme's active site reopens, and threonine leaves.

Most PLP-dependent enzymes are believed to use the cofactor to stabilize the negative charge at Cα of a quinonoid intermediate (see inset, Figure 4) in the transition state.[25] However, as is typical of TS structures, the observed hydrogen-bonding pattern in MtTS suggests that N1 is not protonated. Ser299 accepts a hydrogen bond from Gly327-N, and is therefore likely to donate a hydrogen bond to Thr326, which in turn donates a hydrogen bond to N1. On this basis, Omi *et al.* proposed that the reaction proceeds through the carbanion, rather than the quinonoid. The threonine is replaced by an acidic group in other PLP enzymes, which is expected to improve the ability of the pyridine ring of the cofactor to act as an electron sink. The neutral linkage in TSs, however, has the advantage that this aspect of the reaction will not be as sensitive to pH as when the charged group is present.[26]

We are not aware of any data concerning the pH optimum of TtTS, but values for AtTS, ScTS and EcTS have been reported to be near pH 8, in contrast to the very high pH optimum we observe for MtTS (Figure 1(b)). Such an involved reaction mechanism offers many possible reasons for the differences in behavior, although some factors can be ruled out. For example, the nature of the residue interacting with N1 cannot explain the effect, as it is a neutral residue in all the known TSs. The α-amino group of the substrate must be unprotonated to carry out a nucleophilic attack on C4' of the cofactor. Normal pK_as for

such groups are in the range of 7-8, apparently too low to explain the pH/activity data. Additionally, this would be expected to affect all of the TSs to a similar degree.

The spectroscopic data (Figure 1(a)) suggest that the PLP molecule bound to MtTS is found almost exclusively as the protonated (ketoenamine) form of the Schiff base that links it to Lys69 (Figure 4). TtTS spectra exhibit even less of the 335 (unprotonated) peak than is seen for MtTS. However, the protonation of this group does not appear to be changing for either protein at pHs for which increasing activity is observed for MtTS. ScTS spectra are also reported to be invariant in the range of pH 6-9, and yet the optimum pH for that enzyme is near 8. Thus the protonation state of the Schiff-base nitrogen does not appear to explain the pH/activity data, and differences in the pH optima of the TSs must be due to other aspects of their active sites. Inspection of the structures suggests that substitutions in several active-site residues should be considered (Figure 5). First, Asn95 of MtTS, which interacts with O3, is equivalent to Asn87 of TtTS, but is replaced by aspartic acid residues in AtTS, ScTS and EcTS. Second, Glu296 of MtTS is also a glutamate in TtTS, but varies in both position and sequence (glutamate, cysteine, aspartate) in the other structures. Third, Ala298 of MtTS is equivalent to Ala289 of TtTS, but is replaced by histidine residues in the other three enzymes. A fourth change in the active site is not expected to change the catalytic properties of either cofactor or enzyme. Asp332 of MtTS (Asp323 in TtTS) points away from the active site, interacting with Thr335 (Thr326 in TtTS). This residue is replaced by a phenylalanine, the side chain of which does lie within the active site in AtTS, ScTS and EcTS.

Of the other three changes, the asparagine/aspartate switch at Asn95 of MtTS is by far the most interesting, because of the proximity to O3 of PLP. It was suggested earlier that the aspartate of AtTS and ScTS (and by implication EcTS) must carry a proton, to allow the interaction with a deprotonated O3. The pK_a of this aspartyl residue will also be influenced by its interaction with a nearby conserved lysine side chain (Lys331 of MtTS). At pHs that are above the pK_a of the aspartate, its charge will be negative and the formation of the negatively charged carbanion intermediate will be disfavored. The asparagine residue observed at this position in MtTS, however, will continue to stabilize the intermediate at higher pHs. We predict that TtTS also has a relatively high activity at high pH. These conclusions are consistent with the suggestion of the kinetic experiments that the pH behavior of MtTS is based on an effect on k_{cat},

rather than K_m.

The role of Glu296 in MtTS (and TtTS) is to support the conformation of Asn95 in the closed form of the enzyme. Although it lies near C2' of the PLP, the lack of conservation among the other proteins suggests that it is not a primary determinant of the pH optimum. The replacement of Ala298 of MtTS with a histidine in the enzymes with lower pH optima is of possible interest, since the histidine lies near C2 and C2'. As this residue takes on different conformations and interactions in the structures available, its involvement in local hydrogen bonding networks is difficult to evaluate; in any case, at higher pHs, its charge would be zero, which would not have a negative impact on catalysis.

Inspection of a large number of related sequences indicates that the class I sequences can be divided into two sub-families. AtTS represents one group, while MtTS and TtTS belong to another that is much more common. The patterns discussed above are characteristic of the sub-families: an aspartate to asparagine substitution near O3 is tightly coupled with the histidine to alanine (or another small amino acid such as glycine or serine), and phenylalanine to aspartate, replacements. If an asparagine interacts with the O3 of PLP, the glutamate that supports it is also present. The replacement of the histidine with a smaller residue makes room for this glutamate. These appear to be the only changes in the active sites. Thus we predict that most known TS sequences correspond to enzymes with high pH optima, provided that they are correctly folded. The fact that the AtTS-like pattern of active site conservation applies to ScTS and EcTS indicates that it has broader relevance over the class I/II boundaries.

Several factors influence whether a given enzyme will be a TS or instead catalyze a different type of reaction. Obviously, the shape of the particular active site dictates which substrates can be bound, a restraint that is tightened in the closed form of the enzyme. Dunathan[27] suggested that for a bond to be broken it must be aligned perpendicularly to the pyridine ring of the cofactor in the transition state of the reaction. The carbanion that results is then stabilized by conjugation in the extended π system. The reaction of a TS represents a combination of γ-elimination and β-replacement. In the TtTS-AP5 structure, the phosphate group is aligned with the PLP, in precisely the manner required for γ-elimination. The deprotonations that result in formation of the carbanion intermediate and eneamine also fulfill this criterion. Thus the conformation(s) allowed for the bound substrate to a large degree dictates which

reactions are preferred.

Considerations for drug design

Birds and mammals are unable to synthesize threonine, but are able to acquire it through their diet. The enzymes that synthesize this amino acid in plants, bacteria and fungi are therefore potentially interesting targets for new antibacterial agents and herbicides, provided that they are essential. Data from transposon libraries predict that MtTS is necessary for *M. tuberculosis* growth in normal culture medium.[28] If the most important function of the enzyme is in threonine biosynthesis, mutants should be viable if the growth medium is supplemented with L-threonine. We used our robust two-step homologous recombination method[29] to test this hypothesis. A non-replicating vector carrying an in-frame, unmarked deletion of the *rv1295* (*thrC*) gene was constructed and electroporated into wild-type *M. tuberculosis*. One single cross-over transformant was used to generate double cross-over strains in the presence of L-threonine. We screened the double cross-over strains for the desired deletion by PCR (Figure 6); strains carrying the deletion allele were obtained and confirmed by Southern analysis (not shown). One strain generated in this way was tested for auxotrophy, and as predicted, was unable to grow on plates in the absence of L-threonine. These results confirm both that the gene was deleted, and that MtTS plays a specific role in the biosynthesis of threonine. Thus MtTS has potential as a novel drug target if adequate levels of threonine are not available to the host during infection. We note further that the deletion strain grew poorly, even in the presence of 50 μg/ml L-threonine. Previous work with auxotrophs using model infection systems has suggested that amino acids are not widely available *in vivo*, since numerous *M.* tuberculosis auxotrophs are attenuated[30-36]. It seems likely that a threonine auxotroph would also be attenuated in vivo. Tests of survival in macrophages are the next step in the validation of this hypothesis. The identification of specific inhibitors of MtTS will also be a key step in determining whether this enzyme is an effective target for antitubercular drugs. Clearly, the better understanding of the structure and mechanism of MtTS resulting from the present study will aid that process. The issue of specificity is particularly important, since humans have many PLP-containing enzymes that are essential. A related question concerns the effects on MtTS of compounds directed toward other mycobacterial PLP-dependent

enzymes. Branched-chain amino acid aminotransferases are one example for which available inhibitors have been shown to have an effect on M. tuberculosis.[37] However, given the simple nature of these compounds, it seemed likely that they could be non-specific. We tested four such compounds here, and found no inhibition of MtTS, even at concentrations as high as 1 mM. This suggests that specificity may not be a serious issue with these particular compounds, and that generally, the needed specificity for MtTS or other PLP enzymes can be attained. One known inhibitor of the branched-chain amino acid aminotransferase, gabapentin, is already widely used to relieve neuropathic pain with relatively mild side effects.

Materials and Methods

Cloning, protein expression and purification

The sequence corresponding to the ORF *rv1295* was amplified by PCR from *M. tuberculosis* DNA strain H37Rv [1] using the primer pair 5'- ATG ACC GTC CCG CCG ACG GCC-3' (forward) and 5'- CCG CGC TTG CGA TCG CC -3' (reverse) with Pfu polymerase. A second PCR reaction to introduce an N-terminal 6-histidine tag was performed with the same reverse primer and the forward primer 5'-ATG GCT CAT CAT CAT CAT CAT CAT GGT ACC GTC CCG CCG ACG GCC-3', using the product from the first PCR as template. The PCR product was incubated with dNTPs and Taq polymerase to create an A-overhang, then purified and ligated into the vector pCR®T7 using the TOPO® Cloning pCR® kit (Invitrogen). The plasmid was transformed into Top10 cells (Invitrogen) by heat shock and grown on Luria-Bertani agar plates containing 50 μg/mL ampicillin. Positive clones were transformed into *E. coli* BL21/AI, where the protein was over-expressed after induction with 0.02% arabinose at 37º for 2 h. Cells were harvested by centrifugation. The cell pellet was washed with buffer (10 mM NaH_2PO_4 (pH 7.5), 150 mM NaCl, 1 mM EDTA) and stored at –70 ^{o}C. The clone was verified by DNA sequencing.

For purification, thawed cells were re-suspended in native lysis buffer (50 mM NaH_2PO_4 pH 8, 300 mM NaCl, 20 mM imidazole, 10% glycerol) with 0.01 mg/mL RNAse, 0.02 mg/mL DNAse, 1 mM PMSF and lysed using a One Shot cell disruptor (Constant Systems, Ltd). The cell-free extract was incubated with pre-equilibrated Ni-NTA agarose (Qiagen) slurry for 45 min at 4º C; the resin was washed

and Rv1295 protein eluted with 250 mM imidazole. The eluted protein pool was applied to a size-exclusion chromatography column (Hiload™ 16/60 Superdex™ 75, Pharmacia Biotech) using a buffer containing 150 mM NaCl, 20 mM Tris, HCl, pH 8.0. Purity of the material was evaluated by SDS-PAGE (PhastSystem™, GE Healthcare). Protein samples were concentrated using Vivaspin concentrators (Vivascience) with a molecular weight cut-off of 10 kDa. Final storage was at -20°, in a buffer containing 20 mM Tris-HCl, pH 7.5, 150 mM NaCl. Protein concentration was estimated by the BCA method (BioRad).

Synthesis of OPH

PCl_3 (0.170 mL, 1.94 mmol) and Et_3N (0.88 mL, 6.32 mmol) in CH_2Cl_2 were added to a solution of imidazole (397 mg, 5.84 mmol) in CH_2Cl_2 at 0 °C. The solution was stirred at 0 °C for 10 min, after which benzyl (2S)-2-(benzyloxycarbonyl)amino butanoate (167 mg, 0.486 mmol; synthesized according to previous methods [38]) in CH_2Cl_2 was added. After 30 min at room temperature, benzyl alcohol was added and the mixture was stirred for another 40 min. The mixture was cooled to 0 °C and meta-chloroperbenzoic acid (727 mg, 2.92 mmol) was added and the reaction was stirred for 1 h. The reaction mixture was diluted with CH_2Cl_2 and washed with $Na_2S_2O_3$ (10%, aq.), 1 M HCl, $NaHCO_3$ (sat aq.), dried (Na_2SO_4) filtered and concentrated (reduced pressure at <40 °C). Purification by silica gel chromatography (35-70 μm, Millipore, toluene/EtOAc 3:1) gave benzyl-(2S)-N-(benzyloxycarbonyl)-phosphohomoserine dibenzyl ester (212 mg, 0.351 mmol, 72%); ^{1}H NMR (recorded in $CDCl_3$ at 25 °C on a Varian 300 MHz instrument): 7.32 (m, 20H, Ar), 5.55 (d, 1H,NH), 4.99-5.12 (m, 8H, 4xArCH$_2$O), 4.48 (d, 1H, CHNH), 4.07 (m, 2H, OCH$_2$CH$_2$), 2.10 (m, 2H, CH$_2$CH$_2$CH); MALDI-TOF MS (recorded on a Bruker Biflex III instrument using 2',4',6'-trihydraoyacetophenone trihydrate as matrix): calculated for $C_{33}H_{34}NO_8P$ ($[M+Na]^+$) 626.19, found 626.70.

Palladium on activated carbon was added to the product of the previous reaction (212 mg, 0.351 mmol) in ethanol and the mixture was hydrogenolyzed at atmospheric pressure overnight. The reaction mixture was filtered and concentrated, and the residue purified on a C18 reversed phase column (silica gel

C18 60A, 40-60 μm), $H_2O \rightarrow H_2O$/MeOH 5:1) to afford OPH (69 mg, 0.347 mmol, 99%), the ^1H NMR spectrum (recorded as above) of which was identical to that reported earlier.[16]

Assays

Threonine synthase activity was measured by analyzing the phosphate released from OPH. Measurements were routinely performed in 0.3 mL samples containing 100 mM TAPS buffer, pH 8.4, 0.1 mM pyridoxal phosphate, 2 mM OPH, 0.01 mg/mL human serum albumin and 0.5-15 μg MtTS at room temperature. The reactions were initiated by adding purified enzyme. Ten μL aliquots were removed at 0, 10 and 20 min and assayed for inorganic phosphate by a modification of the method of Lanzetta et al.[39] Malachite green-ammonium molybdate color reagent (160 μL) was added to each 10-μL aliquot, and a minute later 20 μL sodium citrate (34%) were added; the detergent was omitted from the original procedure. Inorganic phosphate concentration was estimated by measuring the absorbance at 660 nm.

The pH dependence of MtTS catalysis was assayed by pre-incubating samples in different buffers, then measuring activity at the same pH. Temperature stability was investigated by pre-incubating samples of MtTS in 10 mM TAPS buffer, pH 8.4, for 10 minutes at 42 ^oC, 65 ^oC and 80 ^oC, then assaying the residual activity.

Inhibitory effects were assessed for the following branched-chain amino acid transferase inhibitors (obtained from Sigma-Aldrich, Inc.), at final concentrations of 0.1 and 1.0 mM: O-allylhydroxylamine hydrochloride, O-benzylhydroxylamine hydrochloride, O-tert-butylhydroxylamine hydrochloride, and gabapentin.

Crystallographic studies

MtTS was crystallized using the sitting-drop vapor-diffusion method at 20 °C. Final conditions were as follows: 1 μL of protein solution (22.5 mg/mL in storage buffer) was mixed with 1 μL of reservoir solution containing 200 mM sodium acetate, 10 mM NaCl and 20-30% PEG 3350 at pH 7.9. Hexagonal crystals with maximum dimensions of ~0.03 mm grew within 2-3 months. Crystals were cryoprotected using the mother liquor supplemented with 20% glycerol and flash-cooled in liquid nitrogen. Diffraction

data were collected at 100 K at the ID14-4 beam line at the European Synchrotron Radiation Facility in Grenoble, France. The data, which showed some anisotropy, were processed and scaled using Denzo and SCALEPACK.[40]

The structure was solved by molecular replacement, using MOLREP[41] with a search model based on the TtTS structure (PDB code 1UIN[12]); in the modified model, all conserved residues were kept, and the rest were mutated to fit the Rv1295 sequence. Both enantiomorphic space groups, $P6_1$ and $P6_5$, were evaluated. The choice of $P6_1$ was confirmed by the presence of electron density for the PLP cofactor, which was not included in the search model. Rigid body refinement was followed by alternating cycles of non-crystallographic symmetry-restrained TLS refinement in REFMAC5[42,43] and manual rebuilding in O,[44] which included the addition of the covalently-bound PLP cofactor. Data collection and refinement statistics are presented in Table 1. Atomic coordinates and structure factor data have been deposited at the PDB with accession code 2D1F.

Essentiality studies

The delivery vector pTACK31G was constructed as follows. First, the upstream flanking region was PCR-amplified using primers ThrF1C 5' AAG CTT TGA CCG ACG ACA TTG AAG AG and ThrR1C 5' CCG CAT GCA GAT TAG TTG CCG CGA TGA G. Then the downstream flanking region was PCR-amplified using primers ThrF2B 5' GCA TGC CGG CAA CGG TCT TAA GGA T and ThrR2 5' GGT ACC CTC TTA CTG TCG GGC TCC AG. Resulting fragments were cloned as *Hin*dIII-*Sph*I and *Sph*I-*Kpn*I fragments into p2NIL *Hin*dIII-*Kpn*I [29]. Finally, the gene cassette (*lacZ, hyg, sacB*) from pGOAL19 [29] was added as a *Pac*I fragment. The resulting vector was electroporated into *M. tuberculosis* [45] and single cross-overs isolated on 7H10 medium plus 10% OADC supplement (oleic acid, bovine serine albumin, dextrose, catalase; Becton Dickinson) agar plates containing 100 µg/ml hygromycin, 20 µg/ml kanamycin, 50 µg/ml X-gal. One single cross-over strain was streaked onto plates without antibiotics. A small sample of cells was recovered and serial dilutions were plated onto medium containing 50 µg/ml L-threonine, 2% w/v sucrose and 50 µg/ml X-gal. White, sucrose-resistant colonies were streaked out to test for auxotrophy and kanamycin/hygromycin sensitivity. PCR primers ThrD1 5' ACT GGA AGG AAC

CGG CTT AT and ThrD2 5' CAG AAG GCA ACA ATG CTT GA were used to amplify the *rv1295* (*thrC*) region to screen for the presence of the wild type (1.3kb) or deletion (0.5 kb) allele. L-threonine was added at 50 μg/ml where required.

Other methods

DALI[18] was used to find the most similar structures in the PDB. Coordinates of proteins with Z scores of 20 or more were obtained and compared with MtTS using DaliLite. Large-scale comparisons of proteins related to MtTS were carried out with BLAST[21] (with default settings) and SAM-02.[46] Structural figures were prepared with O[44] and rendered with POV-Ray (http://www.povray.org) using the Molray interface.[47] Other figures were prepared using Excel (Microsoft Corp) and ChemDraw (CambridgeSoft Corp).

Acknowledgements

The authors are grateful to Anna Suarez Larsson for valuable input on the manuscript, Anna Hartridge and Karen Adie for technical assistance, and Micael Jacobsson (iNovacia, Sweden) for help with the stereochemical restraints for refinement of the Schiff base. Mikael Widersten provided useful insights into the reaction mechanisms of PLP-dependent enzymes. Our tuberculosis-related work is supported by funding from the Foundation for Strategic Research (SSF), the Swedish Research Council (VR), the EU Sixth Framework Programme NM4TB CT:018923, and Uppsala University.

References

1. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Barrell, B. G. & et al. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-44.
2. Sassetti, C. M. & Rubin, E. J. (2003). Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* **100**, 12989-94.
3. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77-84.
4. Flavin, M. & Slaughter, C. (1960). Threonine synthetase mechanism: studies with isotopic hydrogen. *J Biol Chem* **235**, 1112-8.
5. Laber, B., Lindell, S. D. & Pohlenz, H. D. (1994). Inactivation of Escherichia coli threonine synthase by DL-Z-2-amino-5-phosphono-3-pentenoic acid. *Arch Microbiol* **161**, 400-3.
6. Laber, B., Maurer, W., Hanke, C., Grafe, S., Ehlert, S., Messerschmidt, A. & Clausen, T. (1999). Characterization of recombinant Arabidopsis thaliana threonine synthase. *Eur J Biochem* **263**, 212-21.
7. Farrington, G. K., Kumar, A., Shames, S. L., Ewaskiewicz, J. I., Ash, D. E. & Wedler, F. C. (1993). Threonine synthase of Escherichia coli: inhibition by classical and slow-binding analogues of homoserine phosphate. *Arch Biochem Biophys* **307**, 165-74.
8. Mehta, P. K. & Christen, P. (2000). The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Adv Enzymol Relat Areas Mol Biol* **74**, 129-84.
9. Schneider, G., Käck, H. & Lindqvist, Y. (2000). The manifold of vitamin B6 dependent enzymes. *Structure Fold Des* **8**, R1-6.
10. Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. (1995). Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci* **4**, 1291-304.

11. Thomazeau, K., Curien, G., Dumas, R. & Biou, V. (2001). Crystal structure of threonine synthase from Arabidopsis thaliana. *Protein Sci* **10**, 638-48.

12. Omi, R., Goto, M., Miyahara, I., Mizuguchi, H., Hayashi, H., Kagamiyama, H. & Hirotsu, K. (2003). Crystal structures of threonine synthase from Thermus thermophilus HB8: conformational change, substrate recognition, and mechanism. *J Biol Chem* **278**, 46035-45.

13. Garrido-Franco, M., Ehlert, S., Messerschmidt, A., Marinkovic, S., Huber, R., Laber, B., Bourenkov, G. P. & Clausen, T. (2002). Structure and function of threonine synthase from yeast. *J Biol Chem* **277**, 12396-405.

14. Curien, G., Job, D., Douce, R. & Dumas, R. (1998). Allosteric activation of Arabidopsis threonine synthase by S-adenosylmethionine. *Biochemistry* **37**, 13212-21.

15. Laber, B., Gerbling, K. P., Harde, C., Neff, K. H., Nordhoff, E. & Pohlenz, H. D. (1994). Mechanisms of interaction of Escherichia coli threonine synthase with substrates and inhibitors. *Biochemistry* **33**, 3413-23.

16. Barclay, F., Chrystal, E. & Gani, D. (1996). Synthesis of (2S)-O-phosphohomoserine and its C-2 deuteriated and C-3 chirally deuteriated isotopomers: probes for the pyridoxal phosphate dependent threonine synthase reaction. *J. Chem. Soc. Perkin Trans* **1**, 683-689.

17. Jones, S. & Thornton, J. M. (1996). Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* **93**, 13-20.

18. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138.

19. Curien, G., Dumas, R., Ravanel, S. & Douce, R. (1996). Characterization of an Arabidopsis thaliana cDNA encoding an S-adenosylmethionine-sensitive threonine synthase. Threonine synthase from higher plants. *FEBS Lett* **390**, 85-90.

20. Mas-Droux, C., Biou, V. & Dumas, R. (2006). Allosteric threonine synthase. Reorganization of the pyridoxal phosphate site upon asymmetric activation through S-adenosylmethionine binding to a novel site. *J Biol Chem* **281**, 5188-96.

21. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

22. Szczesiul, M. & Wampler, D. E. (1976). Regulation of a metabolic system in vitro: synthesis of threonine from aspartic acid. *Biochemistry* **15**, 2236-44.

23. Yamada, T., Komoto, J., Takata, Y., Ogawa, H., Pitot, H. C. & Takusagawa, F. (2003). Crystal structure of serine dehydratase from rat liver. *Biochemistry* **42**, 12854-65.

24. Palm, D., Klein, H. W., Schinzel, R., Buehner, M. & Helmreich, E. J. (1990). The role of pyridoxal 5'-phosphate in glycogen phosphorylase catalysis. *Biochemistry* **29**, 1099-107.

25. Eliot, A. C. & Kirsch, J. F. (2004). Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu Rev Biochem* **73**, 383-415.

26. Kaiser, J. T., Clausen, T., Bourenkow, G. P., Bartunik, H. D., Steinbacher, S. & Huber, R. (2000). Crystal structure of a NifS-like protein from Thermotoga maritima: implications for iron sulphur cluster assembly. *J Mol Biol* **297**, 451-64.

27. Dunathan, H. C. (1966). Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc Natl Acad Sci U S A* **55**, 712-6.

28. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2001). Comprehensive identification of conditionally essential genes in mycobacteria. *Proc Natl Acad Sci U S A* **98**, 12712-7.

29. Parish, T. & Stoker, N. G. (2000). Use of a flexible cassette method to generate a double unmarked Mycobacterium tuberculosis tlyA plcABC mutant by gene replacement. *Microbiology* **146 (Pt 8)**, 1969-75.

30. Gordhan, B. G., Smith, D. A., Alderton, H., McAdam, R. A., Bancroft, G. J. & Mizrahi, V. (2002). Construction and phenotypic characterization of an auxotrophic mutant of Mycobacterium tuberculosis defective in L-arginine biosynthesis. *Infect Immun* **70**, 3080-4.

31. Hondalus, M. K., Bardarov, S., Russell, R., Chan, J., Jacobs, W. R., Jr. & Bloom, B. R. (2000). Attenuation of and protection induced by a leucine auxotroph of Mycobacterium tuberculosis. *Infect Immun* **68**, 2888-98.

32. Pavelka, M. S., Chen, B., Kelley, C. L., Collins, F. M. & Jacobs, W. R. (2003). Vaccine efficacy of a lysine auxotroph of Mycobacterium tuberculosis. *Infection and Immunity* **71**, 4190-4192.

33. Lee, S., Jeon, B. Y., Bardarov, S., Chen, M., Morris, S. L. & Jacobs, W. R., Jr. (2006). Protection elicited by two glutamine auxotrophs of Mycobacterium tuberculosis and in vivo growth phenotypes of the four unique glutamine synthetase mutants in a murine model. *Infect Immun* **74**, 6491-5.

34. Sambandamurthy, V. K., Wang, X., Chen, B., Russell, R. G., Derrick, S., Collins, F. M., Morris, S. L. & Jacobs, W. R., Jr. (2002). A pantothenate auxotroph of Mycobacterium tuberculosis is highly attenuated and protects mice against tuberculosis. *Nat Med* **8**, 1171-4.

35. Smith, D. A., Parish, T., Stoker, N. G. & Bancroft, G. J. (2001). Characterization of auxotrophic mutants of Mycobacterium tuberculosis and their potential as vaccine candidates. *Infect Immun* **69**, 1142-50.

36. Tullius, M. V., Harth, G. & Horwitz, M. A. (2003). Glutamine synthetase GlnA1 is essential for growth of *Mycobacterium tuberculosis* in human THP-1 macrophages and guinea pigs. *Infect Immun* **71**, 3927-36.

37. Venos, E. S., Knodel, M. H., Radford, C. L. & Berger, B. J. (2004). Branched-chain amino acid aminotransferase and methionine formation in Mycobacterium tuberculosis. *BMC Microbiol* **4**, 39.

38. Wang, Q. & Linhardt, R. J. (2003). Synthesis of a serine-based neuraminic acid C-glycoside. *J Org Chem* **68**, 2668-72.

39. Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. (1979). An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem* **100**, 95-7.

40. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326.

41. Vagin, A. & Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. *Acta Crystallogr.* **D56**, 1622-4.

42. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.* **D53**, 240-255.

43. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr.* **D57**, 122-33.

44. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**, 110-119.

45. Hinds, J., Mahenthiralingam, E., Kempsell, K. E., Duncan, K., Stokes, R. W., Parish, T. & Stoker, N. G. (1999). Enhanced gene replacement in mycobacteria. *Microbiology* **145 (Pt 3)**, 519-27.

46. Karplus, K., Barrett, C. & Hughey, R. (1998). Hidden Markov models for detecting remote protein homologies. *Bioinformatics* **14**, 846-56.

47. Harris, M. & Jones, T. A. (2001). Molray - a web interface between O and the POV-Ray ray tracer. *Acta Crystallogr.* **D57**, 1201-3.

48. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-42.

49. Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr.* **A47**, 392-400.

50. Kleywegt, G. J. & Jones, T. A. (1996). Phi/Psi-cology: Ramachandran revisited. *Structure* **4**, 1395-1400.

51. Gallagher, D. T., Gilliland, G. L., Xiao, G., Zondlo, J., Fisher, K. E., Chinchilla, D. & Eisenstein, E. (1998). Structure and control of pyridoxal phosphate dependent allosteric threonine deaminase. *Structure* **6**, 465-75.

52. Burkhard, P., Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F. & Jansonius, J. N. (1998). Three-dimensional structure of O-acetylserine sulfhydrylase from Salmonella typhimurium. *J Mol Biol* **283**, 121-33.

53. Rhee, S., Parris, K. D., Hyde, C. C., Ahmed, S. A., Miles, E. W. & Davies, D. R. (1997). Crystal structures of a mutant (betaK87T) tryptophan synthase alpha2beta2 complex with ligands bound to the active sites of the alpha- and beta-subunits reveal ligand-induced conformational changes. *Biochemistry* **36**, 7664-80.

54. Oda, Y., Mino, K., Ishikawa, K. & Ataka, M. (2005). Three-dimensional structure of a new enzyme, O-phosphoserine sulfhydrylase, involved in l-cysteine biosynthesis by a hyperthermophilic archaeon, Aeropyrum pernix K1, at 2.0A resolution. *J Mol Biol* **351**, 334-44.

55. Yao, M., Ose, T., Sugimoto, H., Horiuchi, A., Nakagawa, A., Wakatsuki, S., Yokoi, D., Murakami, T., Honma, M. & Tanaka, I. (2000). Crystal structure of 1-aminocyclopropane-1-carboxylate deaminase from Hansenula saturnus. *J Biol Chem* **275**, 34557-65.

Figure legends

Figure 1. Spectroscopic and kinetic properties of MtTS. (a) Absorbance spectra were recorded at 25 ^{o}C with a Beckmann spectrophotometer using an optical path of 1.0 cm. MtTS (0.7 mg/mL, 18.7 μM) was observed in the following buffers, all at 20 mM concentration and including 150 mM NaCl: sodium cacodylate (pH 6.4), Tris-HCl (pH 7.5), Tris-HCl (pH 8.5) and CAPSO (pH 9.6). The insert shows the spectroscopic properties of PLP (0.1 mM) alone in solution under the same conditions. (b) MtTS activity (0.7 μM) with 1 or 5 mM OPH was assayed after incubation for 10 minutes in 10 mM of the following buffers, all except Tris with a sodium counter-ion: cacodylate (pH 6.0), HEPES (pH 7.0, 7.6), Tris-HCl (pH 8.0), CAPSO (pH 8.6, 9.0, 9.6) and CAPS (pH 10.0, 11.0). In each case, the protein was diluted 1/15 into the stated buffer.

Figure 2. Overall structure of dimeric MtTS. The N and C termini of one subunit are labeled. In the same subunit, the domains are colored as follows: large domain (green), small domain (red), swap domain (cyan); the active site PLP of this subunit is colored magenta.

Figure 3. Comparison of MtTS and TtTS active sites. Active-site residues of MtTS and the bound PLP are shown in gold, and equivalent residues of TtTS are gray. Two important hydrogen bonds mentioned in the text are indicated by black dotted lines. Electron density for the PLP in the last $2|F_o|-|F_c|$ map prior to its inclusion in the model, contoured at 1 σ = 0.2 e/Å^3.

Figure 4. Proposed reaction mechanism of MtTS. Residues of particular importance in the reaction described in the main text are indicated. Atoms of the PLP cofactor and the substrate are labeled in the first panel.

Figure 5. Features of particular interest in the TS catalytic mechanism. A stereo view illustrates how the involvement of the PLP 5'-phosphate group in the conversion of β,γ-unsaturated ketimine to α,β-unsaturated aldimine is suggested by the proximity of the relevant groups in the modeled reaction

intermediate, based on superposition of the Tt-AP5 structure (shown by dotted lines). Residues that could influence the pH optimum, as discussed in the text, are shown for MtTS (gold carbons) and EcTS (cyan carbons).

Figure 6. Construction of an *rv1295* (*thrC*) deletion strain of *M. tuberculosis*. The chromosomal location of the gene is shown. The flanking regions indicated by bars were amplified and used to construct a deletion delivery vector carrying an in-frame, unmarked deletion of *thrC*. Deletion strains of *M. tuberculosis* were generated as described as described in the experimental section. PCR primers (indicated by arrows) were used to amplify the *thrC* region; wild type (1.3 kb) and deletion (0.5 kb) products are indicated. Lane 1 – markers; Lane 2 – *thrC* deletion double cross-over strain; Lane 3-6 – wild-type *thrC* double cross-over strains.

Footnotes

Abbreviations used: TS, threonine synthase; MtTS, *Mycobacterium tuberculosis* threonine synthase; AtTS, threonine synthase from *Arabidopsis thaliana*; TtTS, threonine synthase from *Thermus thermophilus HB8*; EcTS, threonine synthase from *Escherichia coli*; ScTS, threonine synthase from yeast (*Saccharomyces cerevisiae*); PLP, pyridoxal 5-phosphate; OPH, (*2S*)-*O*-phospho-L-homoserine; AP5, L-(+)-2-amino-5-phosphonopentanoic acid; PDB, Protein Data Bank[48]).

Table 1. Data collection and refinement statistics.

Data collection statistics	
Beamline	ESRF ID14:4
Wavelength (Å)	0.939
Space group	$P6_1$
Cell parameters: (Å)	56.0, 56.0, 368.4
(degrees)	90, 90, 120
Resolution (Å)	20.0-2.50 (2.59-2.50)[1]
No. of unique reflections	19,700
Redundancy	4.2 (3.0)[1]
Completeness (%)	92.6 (56.4)[1]
R_{sym} (%)	6.9 (34.6)[1]
$\langle I \rangle / \langle \sigma I \rangle$	13.9 (2.2)[1]
Refinement statistics	
Resolution (Å)	20.0-2.50
No. of unique reflections	19,700
No. of reflection in test set	1047
R_{work} (%)	18.7
R_{free} (%)	25.6
No. of atoms (mean B-factor, $Å^2$)	
Protein	5084 (40.8)
Solvent	54 (50.1)
Ligand	29 (42.4)
RMSD from ideal values[2]	

bond length (Å)	0.013
bond angle (°)	1.33
Ramachandran outliers, %[3]	3.1

[1]Values in parentheses refer to statistics in the highest resolution shell.

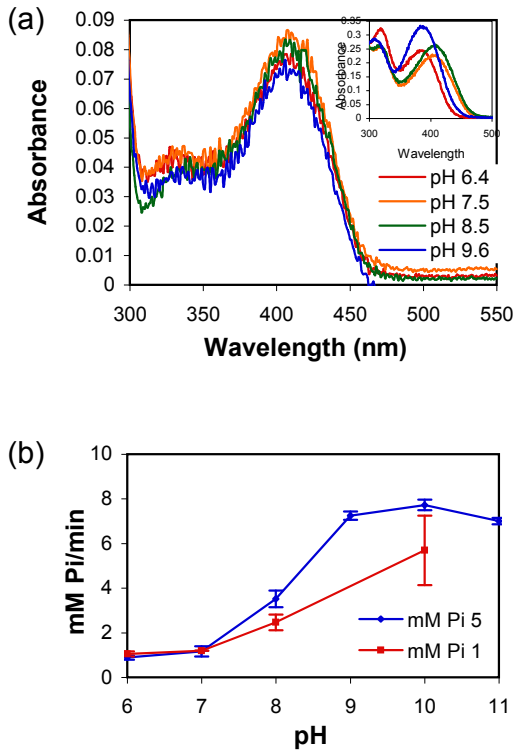
[2] Ideal values from Engh and Huber.[49]

[3] Calculated using a strict-boundary Ramachandran plot.[50]

Table 2. Comparison of MtTS with related structures. Clusters of related structures were identified using DALI.[18] The TtTS structure was added manually; it was not in the original list, as it had been assigned to the same cluster by the program. Similarly, the 2C2B structure replaces 1E5X from the DALI list, as it is more similar in conformation. All structures were subsequently analyzed with DaliLite. Where more than one chain was present, the letter following the PDB code indicates which was used. Nr, number of residues in the sequence compared; Nc, number of Cα atoms considered to be equivalent; Ic, sequence identity within these regions; *Z*, Dali *Z*-score.

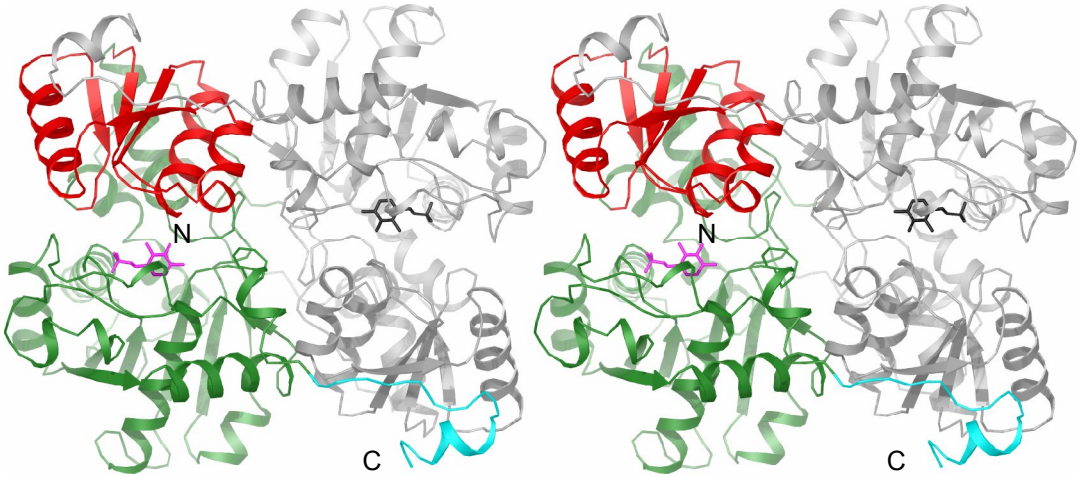
Protein (source)	Nr	Nc	Ic (%)	r.m.s. distance	Z	PDB code, reference
TtTS	350	344	58	1.3	53.1	1UIN-A[12]
AtTS	443	343	37	1.8	45.5	2C2B-A[20]
Threonine deaminase	494	305	19	3.8	34.1	1TDJ[51]
Serine racemase	318	290	20	2.9	33.7	1V71, Goto, unpublished
O-acetylserine sulfhydrylase	315	297	19	2.5	32.9	1OAS-A[52]
Tryptophan synthase	395	305	18	2.8	32.2	2TYS-B[53]
L-serine dehydratase	315	283	21	2.8	30.8	1PWE-A[23]
EcTS	428	303	20	3.3	29.0	1VB3, Omi *et al.*, unpublished
O-phosphoserine sulfhydrylase	382	283	21	3.0	28.7	1WKV-A[54]
ScTS	503	301	18	3.1	25.6	1KL7-A[13]
1-aminocyclopropane-1-carboxylate deaminase	341	282	16	3.0	24.8	1F2D-A[55]

Figure

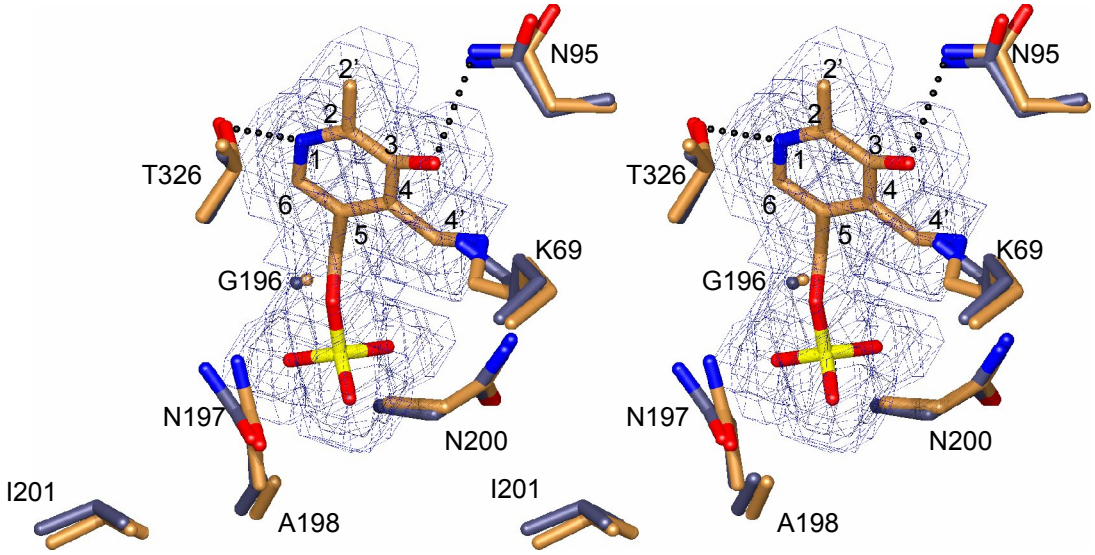

(a)
Absorbance
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
300
350
400
450
500
550
Wavelength (nm)
pH 6.4
pH 7.5
pH 8.5
pH 9.6
Absorbance
Wavelength

(b)
mM Pi/min
10
8
6
4
2
0
6
7
8
9
10
11
pH
mM Pi 5
mM Pi 1

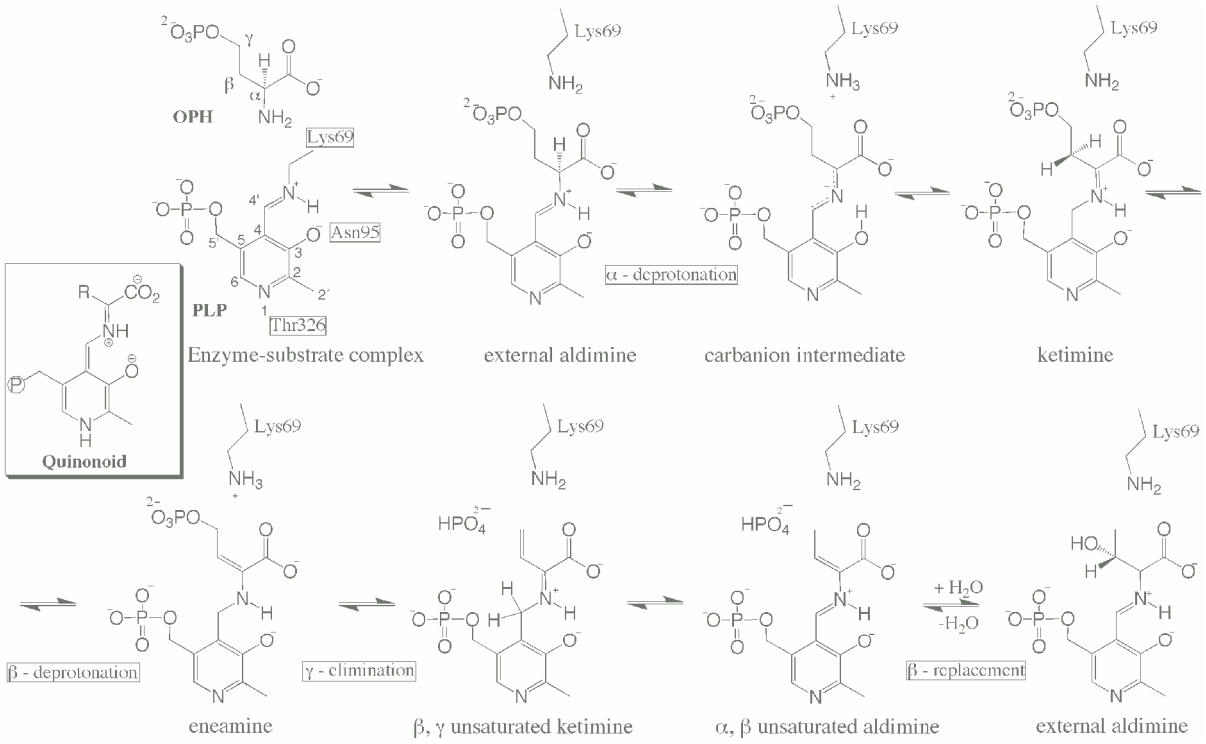
Figure



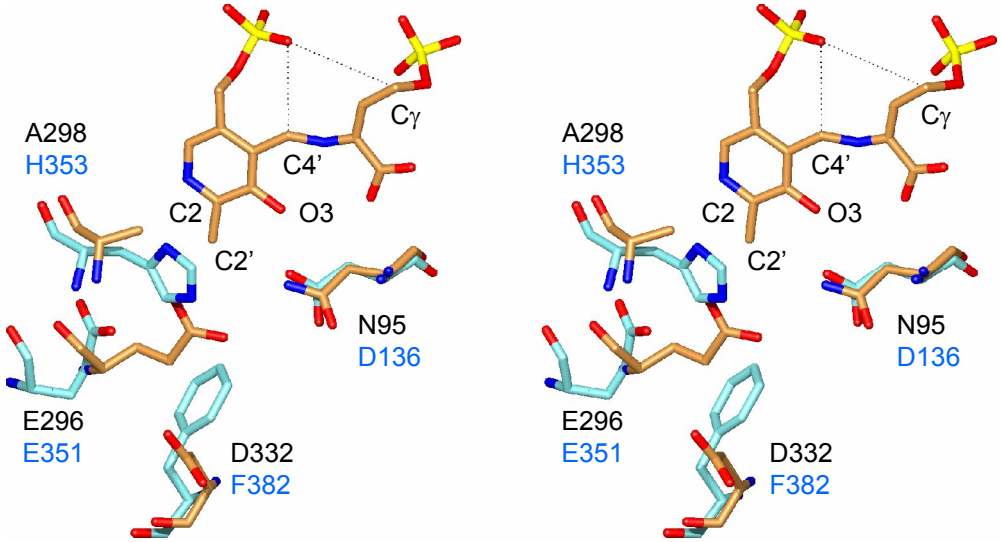
Figure


N95
2'
2
3
1
T326
6
4
4'
5
K69
G196
N197
N200
I201
A198
N95
2'
2
3
1
T326
6
4
4'
5
K69
G196
N197
N200
I201
A198

Figure


OPH
PLP
Lys69
Asn95
Thr326
Quinonoid
Enzyme-substrate complex
external aldimine
α - deprotonation
carbanion intermediate
ketimine
β - deprotonation
eneamine
γ - elimination
β, γ unsaturated ketimine
α, β unsaturated aldimine
+ H2O
-H2O
β - replacement
external aldimine

Figure


A298
H353
Cγ
C4'
C2
O3
C2'
N95
D136
E296
E351
D332
F382
A298
H353
Cγ
C4'
C2
O3
C2'
N95
D136
E296
E351
D332
F382

Figure