Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of *Escherichia coli* tryptophanase

Shao-Yang Ku,[a,b] Patrick Yip[b] and P. Lynne Howell[a,b]*

[a]Structural Biology and Biochemistry, Research Institute, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada, and [b]Department of Biochemistry, Faculty of Medicine, University of Toronto, Medical Science Building, Toronto, Ontario M5S 1A8, Canada

Correspondence e-mail: howell@sickkids.ca

Received 16 March 2006
Accepted 26 May 2006

PDB Reference: *E. coli* tryptophanase, 2c44, r2c44sf.

Pyridoxal 5′-phosphate (PLP) dependent tryptophanase has been isolated from *Escherichia coli* and its crystal structure has been determined. The structure shares the same fold with and has similar quaternary structure to *Proteus vulgaris* tryptophanase and tyrosine-phenol lyase, but is found in a closed conformation when compared with these two enzymes. The tryptophanase structure, solved in its apo form, does not have covalent PLP bound in the active site, but two sulfate ions. The sulfate ions occupy the phosphoryl-binding site of PLP and the binding site of the α-carboxyl of the natural substrate tryptophan. One of the sulfate ions makes extensive interactions with both the transferase and PLP-binding domains of the protein and appears to be responsible for holding the enzyme in its closed conformation. Based on the sulfate density and the structure of the *P. vulgaris* enzyme, PLP and the substrate tryptophan were modeled into the active site. The resulting model is consistent with the roles of Arg419 in orienting the substrate to PLP and acidifying the α-proton of the substrate for β-elimination, Lys269 in the formation and decomposition of the PLP quinonoid intermediate, Arg230 in orienting the substrate–PLP intermediates in the optimal conformation for catalysis, and His463 and Tyr74 in determining substrate specificity and suggests that the closed conformation observed in the structure could be induced by substrate binding and that significant conformational changes occur during catalysis. A catalytic mechanism for tryptophanase is proposed. Since *E. coli* tryptophanase has resisted forming diffraction-quality crystals for many years, the molecular surface of tryptophanase has been analyzed in various crystal forms and it was rationalized that strong crystal contacts occur on the flat surface of the protein and that the size of crystal contact surface seems to correlate with the diffraction quality of the crystal.

1. Introduction

Tryptophan indole-lyase (EC 4.1.99.1), also known as tryptophanase, catalyzes the *in vivo* degradation of L-tryptophan to indole, pyruvate and ammonia *via* a β-elimination mechanism (Snell, 1975). This reaction is pyridoxal 5′-phosphate (PLP, vitamin B_6) dependent and tryptophanase is only active when a PLP molecule is covalently attached to a lysine residue in the active site. Tryptophanase appears to have promiscuous substrate specificity because it also degrades many other β-substituted L-amino acids and their analogs (Watanabe &

Snell, 1977), such as the colorimetric substrate analog *S*-(*o*-nitrophenyl)-L-cysteine (Suelter *et al.*, 1976). Tryptophanase can also degrade D-tryptophan in the presence of high concentrations of ammonium phosphate (Shimada *et al.*, 1996, 1997) and can degrade diastereomeric mixtures of cysteine *S*-conjugates (Wakabayashi *et al.*, 2004). This apparent lack of stereochemical specificity has made the enzyme a model system for the development of an engineered enzyme useful for biodegradation of pollutants. In addition, tryptophanase is important industrially as it can be used to economically synthesize L-tryptophan and related amino acids *in vitro* in the presence of excess products (Watanabe & Snell, 1972).

Tryptophanase is overexpressed and becomes one of the most abundant proteins in *Escherichia coli* when the cells experience alkaline stress (Blankenhorn *et al.*, 1999). The up-regulation of the tryptophanase gene *tnaA* by stress factor σ_s (Lacour & Landini, 2004) promotes the production of indole, which serves as an extracellular signaling molecule in *E. coli* (Wang *et al.*, 2001) and many other bacteria such as *Proteus vulgaris* and various species of *Morganella* (Sonnenwirth, 1980). The use of chemical signals for bacterial communication (quorum sensing) is a widespread phenomenon. In many pathogens, such as *Pseudomonas aeruginosa*, quorum sensing turns on the expression of virulence genes (Passador *et al.*, 1993; Latifi *et al.*, 1995). Likewise, the chemical signal indole facilitates quorum sensing in Enterobacteriaceae during environmental stress and has been shown to regulate biofilm formation in *E. coli* (Di Martino *et al.*, 2002, 2003).

E. coli tryptophanase functions as a tetramer of four 52.8 kDa subunits and while the dimeric form of the enzyme can be isolated (Raibaud & Goldberg, 1976) or induced by cold (Erez *et al.*, 1998), it is inactive. In the active form of tryptophanase PLP is covalently attached to Lys270, but it can be removed by dialysis to produce the apo form of the enzyme (Metzler *et al.*, 1991). To further understand the molecular properties of *E. coli* tryptophanase, its crystal structure has been actively pursued, but these efforts have been hampered by difficulties in forming diffraction-quality crystals (Tani *et al.*, 1990; Kogan *et al.*, 2004; Kawata *et al.*, 1991; Dementieva *et al.*, 1994; Newton *et al.*, 1965). Although until now the structure of *E. coli* tryptophanase has remained elusive, the structure of its homolog from *P. vulgaris* has been solved (Isupov *et al.*, 1998). Since *P. vulgaris* and *E. coli* tryptophanase share 51% sequence identity (Altschul *et al.*, 1997), their three-dimensional structures, active sites and other important residues should be highly similar. However, experimental evidence suggests that the two enzymes may have subtle structural differences as some of their kinetic and spectral properties differ (Zakomirdina *et al.*, 2002). For example, based on spectroscopic studies *E. coli* tryptophanase has been suggested to have a more hydrophobic active site than the *P. vulgaris* enzyme (June *et al.*, 1981).

We present here the first structure of *E. coli* tryptophanase in the public domain. The enzyme has been isolated from its native host and crystallized in its apo form. The structure shows distinct conformational differences compared with the *P. vulgaris* enzyme and explains previous mutagenesis data.

2. Material and methods

2.1. Expression and purification

During the production of the recombinant form of *Leishmania donovani S*-adenosylhomocysteine (SAH) hydrolase (Yang & Borchardt, 2000), native tryptophanase from *E. coli* strain JM109 {genotype *e*14-(*Mcr*A−) *rec*A1 *end*A1*gyr*A96 *thi*-1 *hsd*R17(rK− mK+) *sup*E44 *rel*A1 (*lac-pro*AB) [F′ *tra*D36 *pro*AB *lac*I^{q}Z M15]} was also overexpressed and copurified with our intended target. The pellets of *E. coli* JM109 cells were lysed by sonication in 10 m*M* potassium phosphate pH 7.0, 1 m*M* EDTA and 5 m*M* β-mercaptoethanol (buffer *A*). After the cell debris had been removed by centrifugation at 39 200*g* for 30 min at 277 K, the lysate was loaded onto a 50 ml diethylaminoethyl (DEAE) cellulose (Whatman) anion-exchange column that had been pre-equilibrated with buffer *A*. Partially purified protein was eluted by applying one column volume of 25 m*M* succinate-malic acid buffer pH 5.0. Ammonium sulfate was then added to the elutant to a final concentration of 1.3 *M* and the mixture was incubated on ice for 15 min. After the precipitated protein had been removed by centrifugation at 39 200*g* for 15 min, additional ammonium sulfate was added to the supernatant to a final concentration of 3.6 *M*. After incubation on ice for 30 min, the precipitated tryptophanase was collected by centrifugation at 39 200*g* for 15 min, redissolved in buffer *A* and dialyzed in 4 l of the same buffer overnight. The protein was further purified by size-exclusion chromatography (Superdex 200 HR10/30; Pharmacia) pre-equilibrated with buffer *A*. The last step of purification was on a Mono Q anion-exchange column (MonoQ HR5/5; Pharmacia) pre-equilibrated with buffer *A*; the protein was eluted with a linear gradient to buffer *A* and 0.4 *M* NaCl in five column volumes. Salt was removed from the purified protein by dialysis against 4 l buffer *A* overnight and the protein was concentrated to 13 mg ml^{-1} prior to crystallization. A single but diffuse protein band was observed using SDS–PAGE (data not shown). Sample purity was subsequently assessed by ES–TOF mass spectrometry (see below).

2.2. Crystallization and protein identification

Initial protein crystals were obtained from Hampton Research Crystal Screen 1. With some optimization, *E. coli* tryptophanase was crystallized from a 4 µl drop of equal volumes of 13 mg ml^{-1} protein solution in buffer *A* and crystallization precipitant [1.6 *M* $(NH_4)_2SO_4$, 0.5%(*v*/*v*) PEG 400 and 0.1 *M* sodium HEPES pH 7.9]. In a hanging-drop experiment, the solution drop was suspended over a 0.5 ml reservoir of the precipitant at room temperature and long rod-shaped crystals grew in about a month. The crystals were subsequently soaked in 75%(*v*/*v*) precipitant solution mixed with 25%(*v*/*v*) of glycerol for 10 s and flash-frozen in a stream of nitrogen gas at 110 K prior to data collection. The crystals diffracted weakly to ∼3 Å using Cu $K\alpha$ radiation and were therefore taken to Station X8C, National Synchrotron Light Source (NSLS). A set of 360° of data was collected at a wavelength of 1.0 Å with 90 s exposure per 0.5° oscillation.

The data were processed using *HKL*2000 (Otwinowski & Minor, 1997). The crystal diffracted to 2.8 Å and belonged to space group $P4_12_12$, with unit-cell parameters $a = b = 215.5$, $c = 107.6$ Å. A tetramer was estimated to be present in the asymmetric unit ($V_M = 2.96$ Å^3 Da^{-1}; Matthews, 1968) with a solvent content of 58.4%. The data statistics are summarized in Table 1.

Since we expected to crystallize *L. donovani* SAH hydrolase, we attempted to solve the structure by molecular replacement (MR) using the human enzyme as the search model. The failure to find an MR solution prompted us to analyze the identity of the protein that had been crystallized. The protein solution was therefore analyzed by ES–TOF MS; the MS spectrum exhibited a major peak of 52.8 kDa and a minor peak of 47.6 kDa. Protein crystals were subsequently dissolved in water and blotted on a PVDF membrane and analyzed by N-terminal sequencing. The first 11 amino acids were found to be Met-Glu-(Asp/Asn)-Phe-Lys-His-Leu-Pro-Glu-Pro-Phe. Based on the molecular weight and the N-terminal sequence, *E. coli* tryptophanase was unambiguously identified by *BLAST* (Altschul *et al.*, 1997).

2.3. Structure determination and refinement

The structure of *E. coli* tryptophanase was solved by MR using the *P. vulgaris* enzyme (PDB code 1ax4) as the search model. All rotation and translation matrices were calculated using the MR module in *CNS* (Brünger *et al.*, 1998). The initial σ_A-weighted $2F_o - F_c$ map was difficult to interpret but was dramatically improved by fourfold NCS averaging and density modification. Cycles of structure refinement were iterated between *CNS* (Brünger *et al.*, 1998) and manual modeling using *Xfit* (McRee, 1999). The refinement procedure consisted of iterations of rigid-body refinement, torsion-angle simulated annealing using a maximum-likelihood (ML) amplitude (MLF) target and grouped *B*-factor refinement. During the later stages of refinement, maps were calculated without NCS averaging, but the NCS restraints were still imposed between the four monomers. All refinement steps were monitored using R_{free} with 5% cross-validated data. The structure was refined to R and R_{free} values of 24.1 and 26.5%, respectively, before any ligand or water molecules were added. Water density was then modeled using the iterative procedure implemented in *ARP*/*wARP* (Morris *et al.*, 2003) and *REFMAC* (Murshudov, 1997). Water molecules with unrealistic *B* factors and unusually bulky density were examined and later modeled as sulfate or potassium ions depending on their size and chemical environment. A total of six potassium ions were modeled in the tryptophanase tetramer. We did not find convincing electron density for covalently attached PLP on Lys270 in any of the four subunits in the asymmetric unit; the structure of the tryptophanase presented here is thus the apo form of the enzyme. The first four residues at the N-terminus (residues 1–4) and the last residue of the C-terminus (residues 471) are disordered; residues that lacked side-chain density were truncated to alanine and are listed in Table 1.

Table 1
Diffraction data and crystallographic refinement statistics.

Values in parentheses are for the highest resolution shell (2.91–2.80 Å).

Space group	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 215.5$, $c = 107.5$
Wavelength (Å)	1.0
Resolution (Å)	152–2.8
Average redundancy	6.6
R_{merge}† (%)	9.3 (24.3)
Completeness (%)	99.8 (98.8)
$\langle I\rangle/\langle\sigma(I)\rangle$	26.9 (8.7)
Data with $\langle I\rangle/\langle\sigma(I)\rangle > 3\sigma$ (%)	92.3 (77.1)
No. of reflections (working/test)	59116/3143
No. of protein/ligand atoms	14282/46
No. of water molecules	518
R_{cryst}/R_{free}‡ (%)	19.5/21.8
R.m.s. deviation from ideal values	
Bond length (Å)	0.006
Bond angle (°)	1.37
Dihedral (°)	22.34
Improper (°)	0.79
Average *B* factor (Å^2)	13.8
Ramachandran plot§	
Total favored (%)	96.1
Total allowed (%)	99.6
σ_A-Estimated coordinate error based on R_{free} (Å)	0.34
Residues truncated to Ala	
Subunit *A*	5, 29, 85, 88–89, 94, 115, 119, 127, 156, 164, 174, 178, 182, 218, 246, 374, 406–407, 409, 444, 447, 459–460, 469
Subunit *B*	5, 29, 33, 85, 88–89, 115, 119–120, 127, 156, 164, 174, 182, 218, 246, 285, 374, 406–407, 409, 447, 459–460
Subunit *C*	20, 29, 33, 85, 88, 94, 118–120, 127, 136, 151, 156, 174, 181, 218, 244, 246, 285, 374, 406–407, 409, 447, 459–460
Subunit *D*	5, 20, 28–29, 32–33, 88, 94, 115, 119, 124, 127, 137, 156, 164, 178, 181–182, 218, 246, 285, 371, 374, 387, 396, 406–407, 409, 444, 447, 450, 452, 459–460

† $R_{merge} = \sum\sum|I(k) - \langle I\rangle|/\sum I(k)$, where $I(k)$ and $\langle I\rangle$ represent the diffraction intensity values of the individual measurements and the corresponding mean values. The summation is over all unique measurements. ‡ $R_{cryst} = \sum|F_{obs} - F_{calc}|/\sum|F_{obs}|$; R_{free} is R_{cryst} for a 5% cross-validated test data set. § Ramachandran plot in *MolProbity* (Lovell *et al.*, 2003).

The final *E. coli* tryptophanase model contains 1867 residues, eight sulfate ions and six potassium ions in four subunits. A total of 518 water molecules were modeled in the asymmetric unit. The final model has R and R_{free} values of 19.5 and 21.8%, respectively. Structure validation was performed using *MolProbity* (Lovell *et al.*, 2003). The refinement and structure-validation statistics are summarized in Table 1.

2.4. Structural analysis

The protein sequences were aligned using the program *T-Coffee* (Notredame *et al.*, 2000) and displayed using *ESPript* (Gouet *et al.*, 1999). The secondary-structural elements were defined from the coordinates using *DSSP* (Kabsch & Sander, 1983). Structural alignments were performed using *DaliLite* (Holm & Park, 2000). The C^α-displacement plot was calculated using a Python (http://www.python.org) script that parses files from the *T-Coffee* sequence alignment and *DaliLite* structural alignment and outputs the displacement for plotting

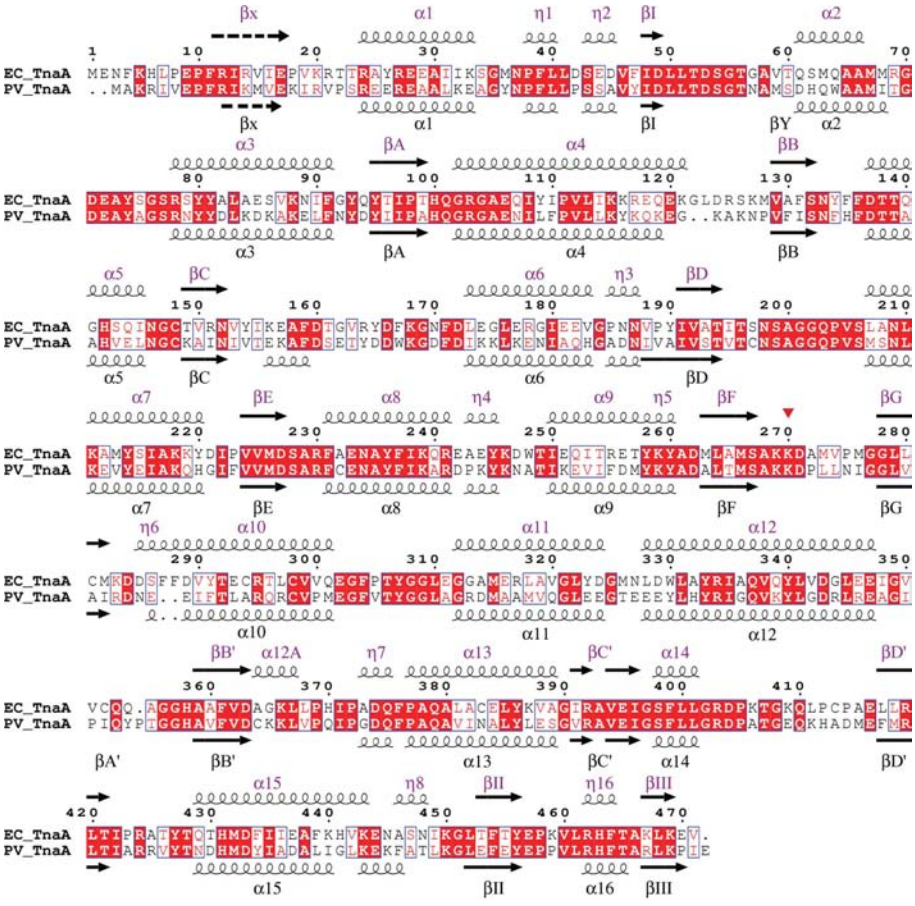

Figure 1
Sequence alignment of *E. coli* and *P. vulgaris* tryptophanase. The secondary-structural elements of *E. coli* tryptophanase have been labeled to be consistent with those of the *P. vulgaris* enzyme. Any unlabeled or seemingly mislabeled secondary-structural elements arise from the inconsistency of secondary-structural assignment between the *DSSP* algorithm and the *GCG* algorithm that was used to define the secondary structure of *P. vulgaris* tryptophanase (Isupov *et al.*, 1998). In the alignment, helices are presented by curly lines and β strands by arrows; η represents a 3_{10}-helix. Strand βx is only classified as strand by *DSSP* when the structure is a tetramer and hence is shown as a dashed arrow. Conserved residues are framed in a red box, while chemically similar residues are framed in a white box. The residue labeled with a red triangular arrow in *E. coli* tryptophanase is Lys270, which has been shown to covalently attach to the PLP; this corresponds to Lys266 of the *P. vulgaris* enzyme.

in the program *Grace* (http://plasma-gate.weizmann.ac.il/Grace/). Modeling and energy minimization were performed in *Swiss-PDB Viewer* (Guex & Peitsch, 1997) using the *GROMOS* force field (van Gunsteren & Mark, 1992). The surface potential was calculated using the *Adaptive Poisson-Boltzmann Solver* (*APBS*; Baker *et al.*, 2001). Figs. 2, 3(*a*), 4 and 5 were prepared using *PyMOL* (http://www.pymol.org).

3. Results and discussion

3.1. Production of apo *E. coli* tryptophanase

During the overexpression of recombinant *L. donovani* SAH hydrolase, native *E. coli* tryptophanase was also expressed in abundance by the *E. coli* strain JM109. This was probably because the bacteria experienced environmental stress during growth and found the need to produce indole as a stress response (Blankenhorn *et al.*, 1999; Di Martino *et al.*, 2003; Lacour & Landini, 2004). The enzyme was copurified with the intended *L. donovani* SAH hydrolase owing to their similar physical properties. The hydrolase is a homotetramer of 47.8 kDa subunits with an estimated pI of 6.0, while the tryptophanase is a homotetramer of 52.8 kDa with an estimated pI of 6.2. Although the ionization efficiency of proteins is hard to predict, the tryptophanase was the highest peak in the ES–TOF mass spectrum and appears to be the major species in the protein mixture. N-terminal sequencing of the dissolved crystals unambiguously confirmed that *E. coli* tryptophanase had been preferentially crystallized. Since PLP was never supplemented during the purification steps, the cofactor was lost during dialysis, resulting in the inactive apo form of the enzyme.

3.2. Overall structure

E. coli tryptophanase and the previously solved *P. vulgaris* enzyme share 51% sequence identity. Not surprisingly, the two proteins have the same fold and highly similar secondary-structural profiles. For ease of comparison, we have adopted the same secondary-structure nomenclature established for the *P. vulgaris* enzyme (Isupov *et al.*, 1998). Structural elements that are unique to *E. coli* tryptophanase are labeled with 'A' after the helix or strand number (*e.g.* helix α12A; Fig. 1). Like its *P. vulgaris* homolog, each monomer of *E. coli* tryptophanase has a PLP-dependent transferase fold, which consists of a small capping domain comprised of residues 1–50 and 329–471 and a large PLP-binding domain comprised of residues 60–328 (Fig. 2*a*). The two domains are connected by two hinges: a loop (residues 51–59) between strand βI and helix α2 and the N-terminal end of helix α12. The cleft region between the two domains forms one part of the active site where the Lys270–PLP covalent adjunct in the large PLP-binding domain would be buried in the active form of the enzyme. In the apo form of the tryptophanase structure presented here, however, no PLP was found near Lys270; instead, the active site was found to be occupied by two sulfate ions (Fig. 2*a*). When the PLP-binding site of one monomer engulfs the potassium ion-binding site of the other monomer, the complete active site is formed (Fig. 2*b*). This dimer is

referred to as the 'catalytic dimer' (Isupov *et al.*, 1998), although tryptophanase is only active as a homotetramer (Snell, 1975). The 'catalytic dimer' dimerizes to form a *D*2 tetramer *via* interactions between their N-terminal βx strands and their α2 helices (Fig. 2*c*).

3.3. Monovalent cation-binding site

The potassium ion plays a role in stabilizing the oligomeric structure of tryptophanase, as cold dissociation of the enzyme has been shown to occur slowly in the presence of K^+ ions and much faster in its absence (Erez *et al.*, 1998). Other monovalent cations such as NH_4^+ or Rb^+ also activate tryptophanase activity *in vitro* (Hogberg-Raibaud *et al.*, 1975; Suelter & Snell, 1977), while Na^+ or Li^+ ions have been shown to inhibit the enzyme activity (Snell, 1975). Four K^+ ion-binding sites are observed in the *E. coli* tryptophanase tetramer as well as in the *P. vulgaris* enzyme (Fig. 2; red spheres). Each of these four K^+ ions interacts with the O^{ε} atom of Glu72 in one subunit and the backbone carbonyl O atoms of Gly55 and Pro275 in the second subunit of the catalytic dimer. These four seemingly evolutionarily conserved K^+ ions are also called the 'activating K^+ ion' as they are close to the active site and are essential for tryptophanase activity (Isupov *et al.*, 1998). Owing to limited resolution of the *E. coli* tryptophanase structure, the coordinating water molecules around the K^+ ions lack clear electron density. In the *P. vulgaris* enzyme, the activating K^+ ion is coordinated by the equivalent protein residues and three water molecules with an irregular coordination geometry. In the *E. coli* tryptophanase tetramer, two additional potential K^+-binding sites are observed in the interface between two subunits of the 'catalytic dimers' (Fig. 2, orange spheres). In each binding site, the cation is coordinated by Gln107 O^{ε}, the backbone O atom of Gln301 and Glu302 O^{ε} of one subunit, along with the same three atoms from the second subunit of the 'catalytic dimer'. In the *P. vulgaris* enzyme, two water molecules are observed in this binding site. These K^+ ions are distant from the active site and appear to play a structural role. While the ions are absent in *P. vulgaris* tryptophanase, it is tempting to speculate that the additional K^+ ions may be involved in enhancing the stability of the 'catalytic dimer'.

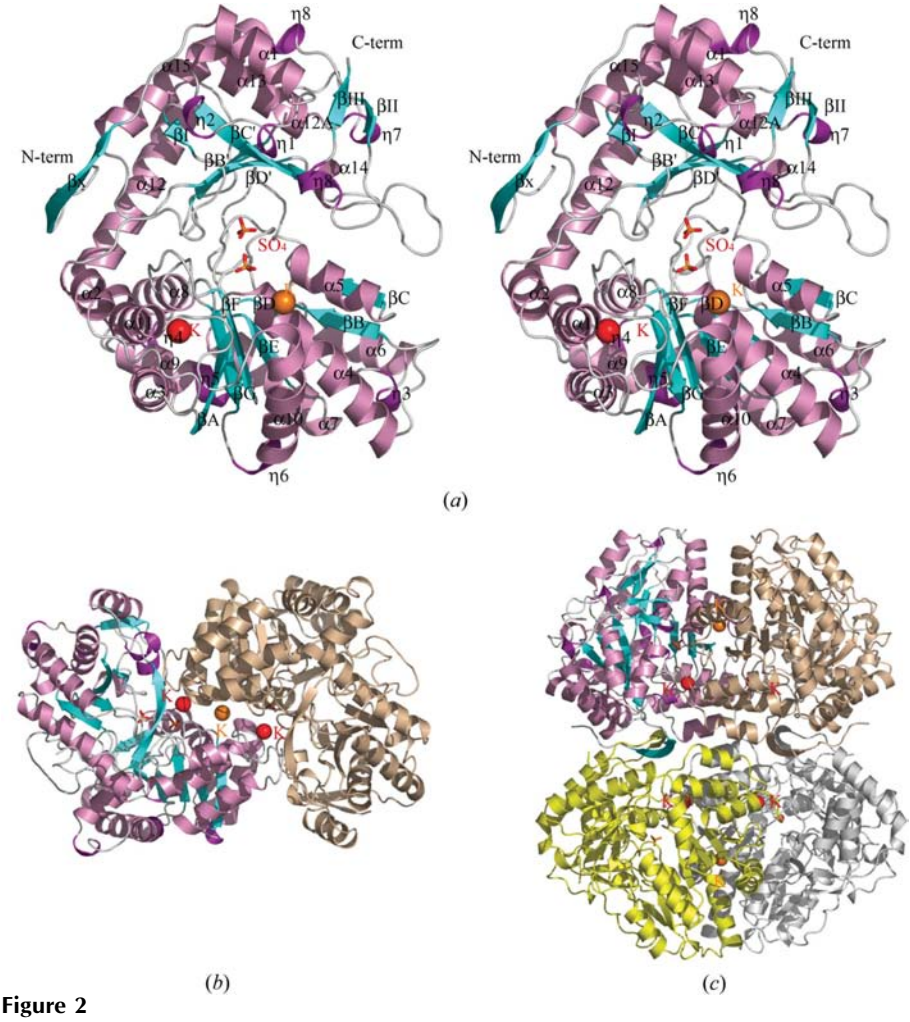

Figure 2
The structure of *E. coli* tryptophanase. (*a*) Stereo ribbon representation of the *E. coli* tryptophanase monomer with α-helices colored in violet, 3_{10}-helices in purple, β-strands in cyan and loops in gray. Two bound sulfate ions in the active site are shown in stick representation with O atoms colored in red and S atoms in orange. The 'activating potassium ions' (see text) that are present in both *E. coli* and *P. vulgaris* tryptophanase are shown as red spheres. The potassium ions that are unique to the *E. coli* enzyme are shown as orange spheres and are located in the interface between two monomers of the 'catalytic dimer'. The secondary-structural elements are labeled as in Fig. 1. (*b*) The structure of the 'catalytic dimer'. Subunit *A* is shown using the same color scheme as in (*a*) and subunit *B* is shown in wheat. (*c*) The structure of the *D*2 tetramer. Subunits *A* and *B* are shown using the same color scheme as in (*b*). Subunits *C* and *D* are shown in yellow and gray, respectively.

3.4. Conformational variations

E. coli and *P. vulgaris* tryptophanase have the same tertiary and quaternary structures and are structurally similar to another PLP-dependent transferase, tyrosine phenol-lyase (TPL; Antson *et al.*, 1993). To date, none of these enzymes have been solved with substrate (tryptophan or tyrosine) bound in the active site. The structure of TPL has been solved both in its apo form (Antson *et al.*, 1993) and in complex with PLP and the inhibitor hydroxylphenyl propionic acid (PDB code 2tpl; B. Sundararaju, A. A. Antson, R. S. Phillips, T. V. Demidkina, M. V. Barbolina, P. Gollnick, G. G. Dodson & K. S. Wilson, unpublished results), while the structure of *P. vulgaris* tryptophanase has been determined with a PLP molecule covalently bound to Lys266. When the

structure of *P. vulgaris* tryptophanase is compared with the structures of TPL both with or without ligands bound, the enzyme was described as being in a 'closed' conformation (Isupov *et al.*, 1998; Antson *et al.*, 1993). Although the four subunits of the *P. vulgaris* enzyme have slightly different domain orientations, resulting in an r.m.s. C^{α} deviation of about 3 Å between each subunit, our structure of apo *E. coli* tryptophanase shows a more 'closed' conformation relative to any one of the subunits of the *P. vulgaris* enzyme (Fig. 3*a*). When the large PLP-binding domains of the two enzymes are structurally aligned, the C^{α} trace of the small capping domain is displaced by 4–7 Å, with the exception of two regions consisting of residues 358–361 and 420–440 (Fig. 3*b*), both of which are located near the hinge region between the two domains. Overall, such large concerted C^{α} displacements strongly suggest rigid-body domain movement.

In addition to the apparent domain closure, residues 122–128 and 283–288 of the PLP-binding domain also deviate considerably in the structural alignment of the two enzymes (Fig. 3). These two regions of local conformational variation correspond to poorly conserved regions between the two enzymes (Fig. 1) and are distant from the active site. Interestingly, in the *E. coli* structure residues in the 122–128 loop make crystal contacts in the lattice, suggesting that crystal packing may be responsible for the observed loop movements in this region.

3.5. Sulfate binding in the active site

Two sulfate ions lie in the active site between the two domains of *E. coli* tryptophanase: one sulfate ion interacts with the large PLP-binding domain, while the other interacts with both the large and the small capping domains. The sulfate ion in the large domain, $SO_4(1)$, overlaps with the sole sulfate ion found in the active site of apo TPL (PDB code 1tpl; Antson *et al.*, 1993) and the phosphate moiety of PLP in the active site of holo TPL (PDB code 2tpl); this sulfate also aligns well with the phosphate moiety of PLP in the active site of *P. vulgaris* tryptophanase (PDB code 1ax4; Fig. 3*a*). PLP is not found in the active site of *E. coli* tryptophanase; instead, the sulfate ion $SO_4(1)$ is found. Given the resolution of the data and the spherical density of the sulfate ion, it is difficult to determine unambiguously the directionality of the sulfate O atoms; however, it appears likely that $SO_4(1)$ interacts with the hydroxyl of Ser267, the backbone amides of Gly102 and Arg103 and the guanidium group of Arg103 (Fig. 4*a*). Residues Ser267, Gly102 and Arg103 of the *E. coli* enzyme correspond to Ser263, Gly100 and Arg101 of *P. vulgaris* tryptophanase and these residues directly interact with the phosphoryl O atoms of PLP in the *P. vulgaris* enzyme. Thus, we propose that $SO_4(1)$ in the *E. coli* tryptophanase structure occupies the phosphate-binding site of PLP. Based on the sulfate density and the structural alignment of the PLP-binding domains, we used the geometry of the Lys266–PLP adjunct found in the *P. vulgaris* enzyme to model the Lys270–PLP adjunct in the active site of *E. coli* enzyme (Fig. 4*b*).

The second sulfate ion, $SO_4(2)$, lies just below the small capping domain and forms potential hydrogen bonds of less than 3.0 Å with the hydroxyl of Thr52 and the guanidium group of Arg419. The sulfate ion $SO_4(2)$ also interacts with residues from the large PLP-binding domain, forming short hydrogen bonds with the side-chain amide of Asn198, the guanidium group of Arg203 and the terminal amine of Lys270 (Fig. 4*a*). Owing to its extensive interactions with both domains of the protein, this sulfate ion may be responsible for holding the two domains in the closed conformation.

Sulfate ions can bind to basic cavities in proteins and in many cases mimic the carboxyl groups of the substrate. For example, in the $\delta1$ or $\delta2$ crystallins, argininosuccinate lyase homologs, a sulfate ion is observed in the active site; this sulfate superimposes well with the carboxyl moiety of its natural substrate, argininosuccinate (Sampaleanu *et al.*, 2001).

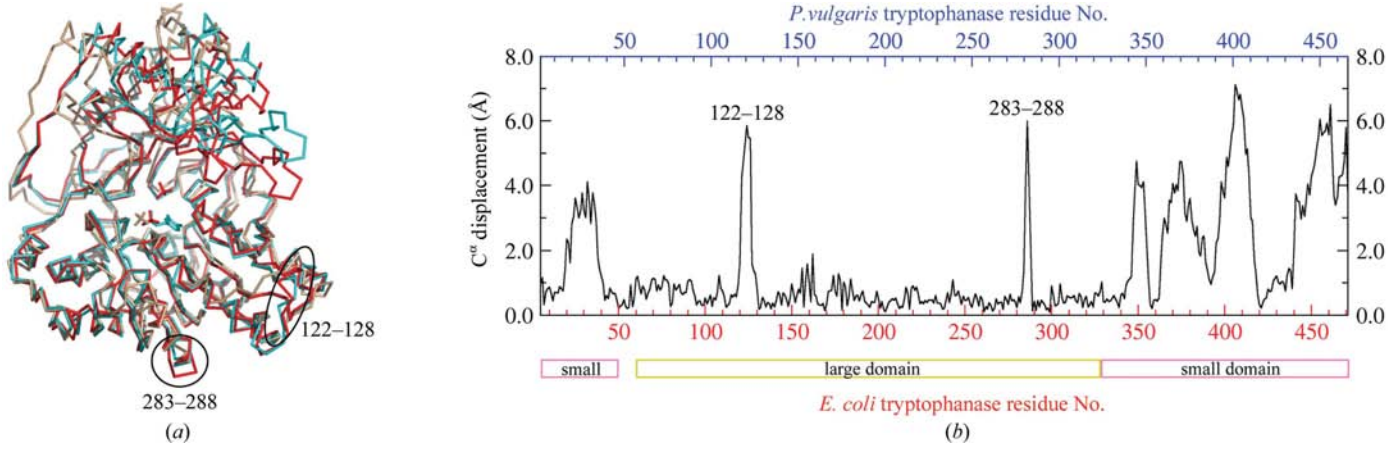

Figure 3
Structural alignment of tryptophanase structural neighbors. (*a*) The C^{α} structural alignment of the large PLP-binding domains of *Citrobacter intermedus* tyrosine phenol-lyase in wheat (PDB code 1tpl), *P. vulgaris* tryptophanase in blue (PDB code 1ax4) and *E. coli* tryptophanase in red. The single sulfate ion found in tyrosine phenol-lyase, the lysyl-PLP found in *P. vulgaris* tryptophanase and the two sulfate ions found in *E. coli* tryptophanase are shown as wheat, blue and red sticks, respectively. (*b*) The C^{α} displacements between *E. coli* and *P. vulgaris* tryptophanase as a function of residue number based on the structural alignment in (*a*). The domain assignment refers to the *E. coli* tryptophanase enzyme.

In *E. coli* tryptophanase, the sulfate $SO_4(2)$ interacts strongly with Arg419, which has been predicted to bind the α-carboxyl of the diverse tryptophanase substrates (Phillips & Doshi, 1998). Arg416 of *E. coli* tryptophanase corresponds to Arg414 of *P. vulgaris* tryptophanase and Arg404 of *C. freundii* TPL. In the holo structure of TPL, Arg404 interacts with the carboxyl moiety of the TPL inhibitor hydroxylphenyl propionic acid, which mimics the α-carboxyl of the TPL substrate tyrosine. Thus, is it likely that the $SO_4(2)$ sulfate in the *E. coli* tryptophanase structure occupies the binding site of α-carboxyl of the natural substrate tryptophan. The catalytic mechanism of tryptophanase requires that the α-amino of tryptophan must be near the imine group of Lys270–PLP in order to form the external tryptophan–PLP imine intermediate (Snell, 1975; Phillips, 1991). Based on this geometric requirement and the $SO_4(2)$ sulfate density near Lys419, we modeled the substrate tryptophan into the active site and performed an energy minimization (Fig. 4*b*).

3.6. Active-site model

The structure of holo tryptophanase has not been reported. In addition to this study, past crystallization attempts also found that *E. coli* tryptophanase preferentially crystallizes from ammonium sulfate (Tani *et al.*, 1990; Kogan *et al.*, 2004; Kawata *et al.*, 1991; Dementieva *et al.*, 1994; Newton *et al.*, 1965). As sulfate ions appear to be competitive inhibitors of the enzyme, we expect that there will be challenges to producing crystals of the tryptophanase in complex with tryptophan. Thus, based on the sulfate density observed and our structural superposition with the PLP-bound form of *P. vulgaris* tryptophanase, we modeled PLP and tryptophan into the active site of the *E. coli* enzyme (Fig. 4*b*). This model is consistent with the results of previous biochemical studies. Firstly, the phosphoryl moiety of PLP interacts extensively with Arg103 and the nearby Lys269. Based on kinetics studies of K269R mutants, the electrostatic interaction between the PLP and Lys269 has been proposed to be important for the formation and decomposition of the PLP quinonoid intermediate, although Lys269 itself is not an essential catalytic residue (Phillips *et al.*, 1991). Secondly, in the active-site model, Arg230 interacts with α-carboxyl of tryptophan and is not near the PLP. Studies on the equivalent residue Arg226 in the *P. vulagris* tryptophanase demonstrate that this residue is not important for tautomerization of the Lys266–PLP internal imine but is important for the optimal conformation of the substrate–PLP intermediates (Kulikova *et al.*, 2003). Thirdly, His463 and Tyr74 of *E. coli* tryptophanase were shown to be important for the substrate specificity of L-tryptophan as Y74F and H463F mutants showed vanishing activity for L-tryptophan but elevated activity for other amino-acid analogs (Phillips *et al.*, 2002). Furthermore, both H463F and Y74F mutants are unable to bind L-tryptophan, but partial denaturing and refolding of equimolar mixtures of the two mutants rescued the tryptophanase activity for L-tryptophan. This observation suggests that in the active site His463 and Tyr74 come from different subunits. Upon dissociation and reassociation of tryptophanase subunits, the hybrid dimers of H463F and Y74F would regenerate one native His463/Tyr74 binding site (Phillips *et al.*, 2002). This is consistent with our model, where His463 from

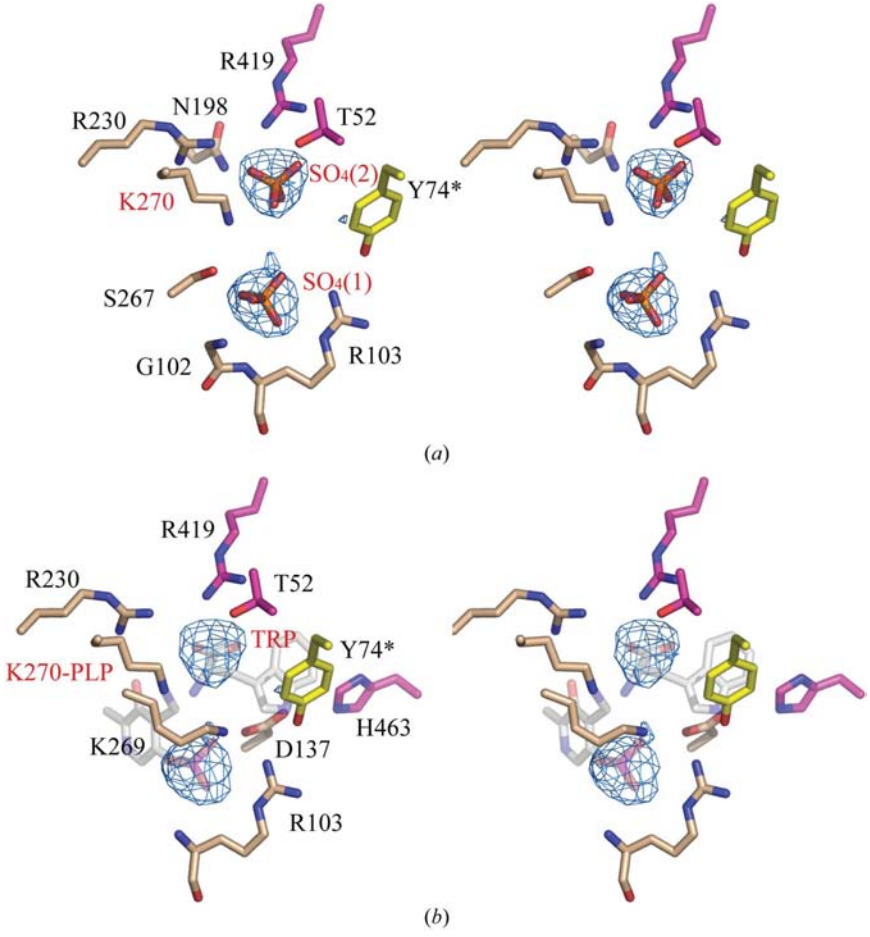

Figure 4
Active site of *E. coli* tryptophanase. (*a*) Stereo representation of the sulfate-binding site. The carbon elements of residues from the large PLP-binding domain of the enzyme are colored in wheat, while those from the small capping domain are colored in magenta. Residue Tyr74 is donated from the second subunit of the 'catalytic dimer' and is labeled with an asterisk and colored yellow. The blue σ_A-weighted $F_o - F_c$ sulfate OMIT map is contoured at 3σ. All residues, except for Tyr74, are within 3.5 Å or less of the sulfate ion. Since the sulfate density is spherical at 2.8 Å, it is difficult to predict its exact orientation, which has been modeled by geometric restraints only. Any of the residues depicted, except Tyr74, could potentially interact with the sulfate ion. (*b*) Stereo representation of the active-site model with the bound PLP and tryptophan displayed as transparent sticks. The residues in (*b*) are important active-site residues that have been implicated previously as being important for the activity of *E. coli* tryptophanase.

one subunit and Tyr74 from the other subunit both interact with the indole ring of L-tryptophan (Fig. 4*b*). Lastly, based on the kinetics studies of the D133A and H458A mutants of *P. vulgaris* tryptophanase, Asp133 is likely to form a hydrogen bond with the indole N atom of the substrate and His458 forms a hydrogen bond to Asp133 (Demidkina *et al.*, 2003). Residues Asp133 and His483 of *P. vulgaris* tryptophanase correspond to Asp137 and His463 of the *E. coli* enzyme, respectively. Our active-site model shows that the proposed hydrogen-bonding scheme is indeed plausible, although the exact hydrogen-bonding geometry and distances cannot be determined at this point.

3.7. Implications for enzymatic mechanism

We described above the domain closure of *E. coli* tryptophanase relative to the *P. vulgaris* enzyme and TPL. Although the two domains are held by a sulfate ion, we argue that the closed conformation observed is relevant for catalysis. Since this sulfate is likely to occupy the α-carboxyl binding site of the tryptophanase substrate and makes interactions with residues of both domains of the enzyme (Fig. 4*b*), we propose that this closed conformation would be induced by substrate binding. The domain closure seems to be important for catalysis because it brings the α-amino group of the substrate closer to the reactive center of the lysyl-PLP to form an external imine that would otherwise be slow to form in the open conformation. An alternative interpretation could be that this closed conformation is unique to the *E. coli* enzyme and may explain some of the differences observed in the kinetic and spectral properties of the *E. coli* and *P. vulgaris* enzymes. Spectroscopic studies have suggested that the active site of *E. coli* tryptophanase has a more hydrophobic microenvironment than that of the *P. vulgaris* enzyme (June *et al.*, 1981) and indeed the active site in our closed conformation is certainly less accessible to solvent than the active site of the *P. vulgaris* enzyme. While our results do suggest that domain movement is required for catalysis, whether this is a general phenomenon required in related enzymes is unclear, as TPL undergoes little domain movement when the inhibitor hydroxylphenyl propionic acid is bound (Antson *et al.*, 1993).

E. coli tryptophanase is an acidic protein and its surface is largely negatively charged; the active site, however, is a deep basic cavity (Fig. 5*a*). As the substrate tryptophan (TRP) enters the active site, the α-amino group of the TRP needs to be deprotonated not only to reduce the electrostatic repulsion in the basic cavity but also to act as a nucleophile to attack the CN imine carbon of Lys270–PLP. The nucleophilic attack of Lys270–PLP results in the formation of a PLP–TRP external aldimine intermediate. Arg419 would be an essential residue for this process as it not only orients the substrate to PLP but also neutralizes the α-carboxyl of TRP to make the α-proton acidic enough to be extracted by a strong base, converting the enzyme from the external aldimine intermediate form to the quinonoid intermediate form. Phillips *et al.* (2002) proposed that Tyr74 of the adjacent subunit donates a proton to the indole moiety of PLP–TRP, making the indole a better leaving

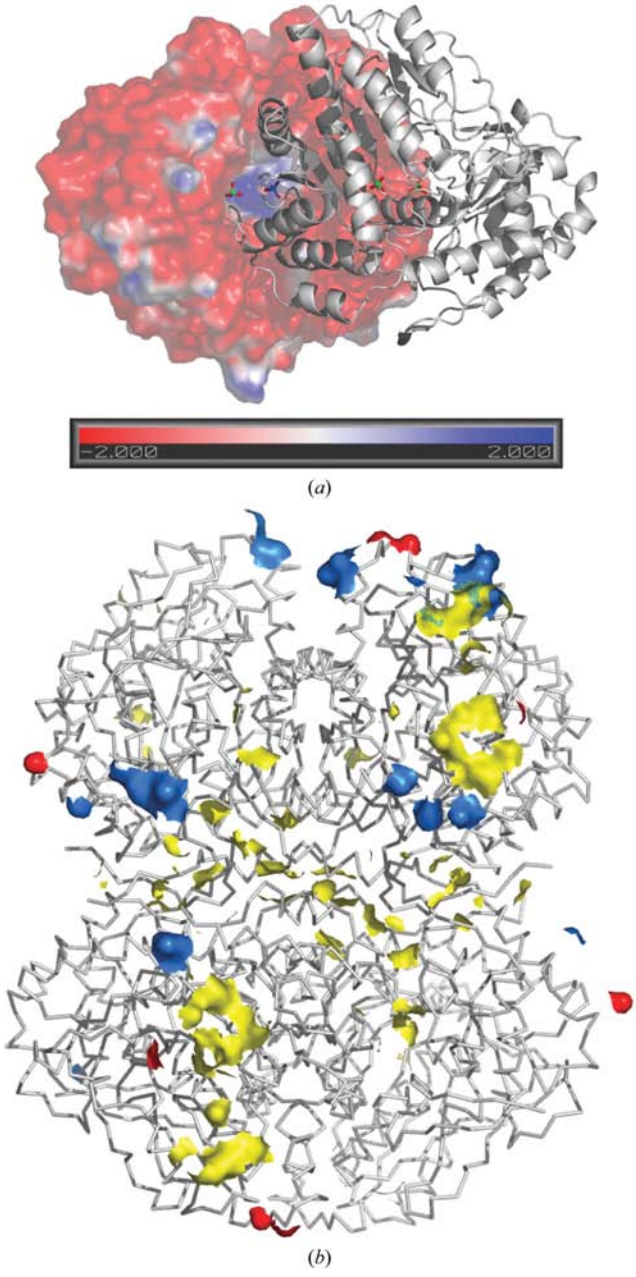

Figure 5
Surface properties of tryptophanase. (*a*) Surface potential of tryptophanase monomer with positive potential shown in blue and negative in red. The scale of the surface potential is dimensionless and ranges from $-2kT/e$ to $+2kT/e$. The second subunit of the 'catalytic dimer' is also shown as a white cartoon as a reference. Since the surface potential was calculated using only one subunit of the protein, some systematic errors in the estimation of the electrostatic potential could occur in the regions buried by the other subunit. The sulfate ions can be seen in the basic cavity of the overall acidic protein. (*b*) The crystal contact surface of tryptophanase. The overall structure of the tryptophanase tetramer is shown as white lines. The red surface represents the atomic crystallographic contacts of less than 3 Å in the structure of *E. coli* tryptophanase in space group $P4_12_12$ with four subunits per asymmetric unit. The yellow surface represents the estimated atomic crystallographic contacts of less than 3 Å if the enzyme were in in a *F*222 cell with one subunit per asymmetric unit. The blue surface represents the atomic crystallographic contacts of less than 3 Å in the structure of *P. vulgaris* tryptophanase in space group $P2_12_12_1$ with four subunits per asymmetric unit.

group; β-elimination of PLP–TRP then occurs to form an α-aminoacrylate intermediate *via* an imino quinonoid transition state (Phillips *et al.*, 2002). Subsequent nucleophilic attack of Lys270 on the CN imine carbon of the PLP–α-aminoacrylate intermediate results in Lys270–PLP and the release of the unstable α-aminoacrylate into solution, where it is converted to pyruvate and ammonium in an aqueous environment.

3.8. Crystal packing

Crystallization is the most poorly understood bottleneck in the structure determination of macromolecules; this is particularly true for the crystallization of *E. coli* tryptophanase. The first crystals of *E. coli* tryptophanase, although not grown for diffraction studies, were described in 1965 (Newton *et al.*, 1965). The first crystals grown for diffraction purposes were obtained in 1990 and somewhat optimized in 1991. These crystals were reported to belong to a hexagonal space group (unit-cell parameters $a = b = 153.3$, $c = 141.8$ Å, $\gamma = 120^\circ$) and diffracted to 7 Å resolution (Tani *et al.*, 1990; Kawata *et al.*, 1991). The diffraction quality of the crystals of *E. coli* tryptophanase was subsequently improved to 3.4 Å resolution in 1994; these crystals belonged to space group $P4_32_12$ (unit-cell parameters $a = b = 113.8$, $c = 231.7$ Å), with two subunits per asymmetric unit (Dementieva *et al.*, 1994). The recombinant form of *E. coli* tryptophanase was crystallized again in the orthorhombic space group $F222$ (unit-cell parameters $a = 118.4$, $b = 120.1$, $c = 171.2$ Å), with one subunit per asymmetric unit (Kogan *et al.*, 2004). We unintentionally crystallized non-recombinant tryptophanase in yet another space group $P4_12_12$ (unit-cell parameters $a = b = 215.5$, $c = 107.6$ Å), with four subunits per asymmetric unit. Interestingly, most of the *E. coli* tryptophanase crystals were grown using ammonium sulfate as the major precipitant. Although many crystal forms have been observed, it is only recently that diffraction-quality crystals of *E. coli* tryptophanase have been obtained. Since the most important variable in protein crystallization is the protein itself, we have attempted to rationalize the difficulty of crystallizing the enzyme by analyzing the molecular properties of tryptophanase rather than the macroscopic properties of the crystallization conditions.

E. coli tryptophanase has been crystallized in hexagonal, tetragonal and orthorhombic space groups. Our crystals in space group $P4_12_12$ diffract weakly, probably because of their small crystal contact area. The entire tetrameric asymmetric unit only has eight atomic crystal contacts that are less than 3 Å and only four residues, Gly122, Asn186, Ser206 and Glu444, are involved in the lattice packing. From the coarse grid surface-area calculation in *POPS* (Fraternali & Cavallo, 2002), the eight crystal contacts constitute only 0.6% of the total surface area of the asymmetric unit (Fig. 5*b*; red surface). Although the structure of *E. coli* tryptophanase in space group $F222$ is not yet available, we can deduce the $F222$ lattice packing by shifting the centroid of the 222 ($D2$) tryptophanase tetramer to the origin of an orthorhombic cell ($a = 118.4$, $b = 120.1$, $c = 171.2$ Å) and by rotating the tetramer so that the NCS axes correspond to the crystallographic axes. Without further refinement in real space or reciprocal space, we have not attempted to predict the exact side-chain atoms that are involved in crystal contacts. Nevertheless, the protein monomer in the $F222$ cell would form 99 atomic contacts that are less than 3 Å to the neighboring molecules; 38 of these 99 contacts are not involved in native tetramer formation. We calculate that the surface area involved in crystal packing, but not in tetramer formation, is about 4.9% of total surface area (Fig. 5*b*; yellow surface). *P. vulgaris* tryptophanase has been crystallized in space group $P2_12_12_1$ with four subunits per asymmetric unit. An entirely different set of 14 residues are involved in forming crystal contacts in this space group and constitute 1.8% of the total surface area of the tetrameric protein in the asymmetric unit (Fig. 5*b*; blue surface). Lattice formation, like protein oligomerization, is a special case of protein–protein interaction. We expect that the general rules observed for protein–complex formation should hold for lattice formation. As only the size of the interacting area correlates well with the affinity of the protein–protein interaction (Ponstingl *et al.*, 2000), so the size of the surface that is involved in the crystal contacts, at least for tryptophanase, seems to correlate with the diffraction quality of the crystal.

4. Conclusions

The structure of *E. coli* tryptophanase from its natural source has been determined and described here for the first time. The structure shows a closed conformation that we propose resembles the holo form of the enzyme and therefore may be relevant in catalysis. Since the substrate has never been observed in tryptophanase or related lyases, we have modeled the tryptophan and PLP into the active site. The model adequately explains much of the existing biochemical data. Further structural data will be required, however, to explain the enzyme's apparent lack of stereospecificty. This property of tryptophanase could be used to design a more robust enzyme for broad-spectrum biodegradation and other industrial applications. The enzyme has been resistant to the formation of diffraction-quality crystals until recently (Kogan *et al.*, 2004) and we hypothesize that diffraction quality is directly related to surface area of crystal contact.

We thank Jinsong Zhang and Ronald T. Borchardt for providing the JM109 *E. coli* strain and the plasmid of *L. donovani* SAH hydrolase. We also thank G. David Smith for insightful discussion. This work is supported in part by a research grant from the Canada Institute of Health Research (No. 13337). PLH and S-YK are the recipients of a Canadian Institutes of Health Research Investigator Award and Natural Sciences and Engineering Research Council of Canada postgraduate scholarship, respectively. Station X8C at National Synchrotron Light Source, Brookhaven National Laboratory is supported by the United States Department of Energy and Multi-User maintenance grants from the Canada Institutes of

Health Research and the Natural Sciences and Engineering Research Council of Canada.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.

Antson, A. A., Demidkina, T. V., Gollnick, P., Dauter, Z., von Tersch, R. L., Long, J., Berezhnoy, S. N., Phillips, R. S., Harutyunyan, E. H. & Wilson, K. S. (1993). *Biochemistry*, **32**, 4195–4206.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Blankenhorn, D., Phillips, J. & Slonczewski, J. L. (1999). *J. Bacteriol.* **181**, 2209–2216.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Dementieva, I. S., Zakomirdina, L. N., Sinitzina, N. I., Antson, A. A., Wilson, K. S., Isupov, M. N., Lebedev, A. A. & Harutyunyan, E. H. (1994). *J. Mol. Biol.* **235**, 783–786.

Demidkina, T. V., Zakomirdina, L. N., Kulikova, V. V., Dementieva, I. S., Faleev, N. G., Ronda, L., Mozzarelli, A., Gollnick, P. D. & Phillips, R. S. (2003). *Biochemistry*, **42**, 11161–11169.

Di Martino, P., Fursy, R., Bret, L., Sundararaju, B. & Phillips, R. S. (2003). *Can. J. Microbiol.* **49**, 443–449.

Di Martino, P., Merieau, A., Phillips, R., Orange, N. & Hulen, C. (2002). *Can. J. Microbiol.* **48**, 132–137.

Erez, T., Gdalevsky, G., Torchinsky, Y. M., Phillips, R. S. & Parola, A. H. (1998). *Biochim. Biophys. Acta*, **1384**, 365–372.

Fraternali, F. & Cavallo, L. (2002). *Nucleic Acids Res.* **30**, 2950–2960.

Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.

Gunsteren, W. F. van & Mark, A. E. (1992). *Eur. J. Biochem.* **204**, 947–961.

Hogberg-Raibaud, A., Raibaud, O. & Goldberg, M. E. (1975). *J. Biol. Chem.* **250**, 3352–3358.

Holm, L. & Park, J. (2000). *Bioinformatics*, **16**, 566–567.

Isupov, M. N., Antson, A. A., Dodson, E. J., Dodson, G. G., Dementieva, I. S., Zakomirdina, L. N., Wilson, K. S., Dauter, Z., Lebedev, A. A. & Harutyunyan, E. H. (1998). *J. Mol. Biol.* **276**, 603–623.

June, D. S., Suelter, C. H. & Dye, J. L. (1981). *Biochemistry*, **20**, 2714–2719.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Kawata, Y., Tani, S., Sato, M., Katsube, Y. & Tokushige, M. (1991). *FEBS Lett.* **284**, 270–272.

Kogan, A., Gdalevsky, G. Y., Cohen-Luria, R., Parola, A. H. & Goldgur, Y. (2004). *Acta Cryst.* D**60**, 2073–2075.

Kulikova, V. V., Zakomirdina, L. N., Bazhulina, N. P., Dementieva, I. S., Faleev, N. G., Gollnick, P. D. & Demidkina, T. V. (2003). *Biochemistry (Mosc.)*, **68**, 1181–1188.

Lacour, S. & Landini, P. (2004). *J. Bacteriol.* **186**, 7186–7195.

Latifi, A., Winson, M. K., Foglino, M., Bycroft, B. W., Stewart, G. S., Lazdunski, A. & Williams, P. (1995). *Mol. Microbiol.* **17**, 333–343.

Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.

McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Metzler, C. M., Viswanath, R. & Metzler, D. E. (1991). *J. Biol. Chem.* **266**, 9374–9381.

Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.

Murshudov, G. N. (1997). *Acta Cryst.* D**53**, 240–255.

Newton, W. A., Morino, Y. & Snell, E. E. (1965). *J. Biol. Chem.* **240**, 1211–1218.

Notredame, C., Higgins, D. G. & Heringa, J. (2000). *J. Mol. Biol.* **302**, 205–217.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Passador, L., Cook, J. M., Gambello, M. J., Rust, L. & Iglewski, B. H. (1993). *Science*, **260**, 1127–1130.

Phillips, R. S. (1991). *Biochemistry*, **30**, 5927–5934.

Phillips, R. S. & Doshi, K. J. (1998). *Eur. J. Biochem.* **255**, 508–515.

Phillips, R. S., Johnson, N. & Kamath, A. V. (2002). *Biochemistry*, **41**, 4012–4019.

Phillips, R. S., Richter, I., Gollnick, P., Brzovic, P. & Dunn, M. F. (1991). *J. Biol. Chem.* **266**, 18642–18648.

Ponstingl, H., Henrick, K. & Thornton, J. M. (2000). *Proteins*, **41**, 47–57.

Raibaud, O. & Goldberg, M. E. (1976). *J. Biol. Chem.* **251**, 2820–2824.

Sampaleanu, L. M., Vallee, F., Slingsby, C. & Howell, P. L. (2001). *Biochemistry*, **40**, 2732–2742.

Shimada, A., Kogure, H., Shishido, H. & Nakamura, I. (1997). *Amino Acids*, **12**, 379–383.

Shimada, A., Shishido, H. & Nakamura, I. (1996). *Amino Acids*, **11**, 83–89.

Snell, E. E. (1975). *Adv. Enzymol. Relat. Areas Mol. Biol.* **42**, 287–333.

Sonnenwirth, A. C. (1980). *Microbiology*, 3rd ed., edited by B. D. Davis, R. Dulbecco, H. N. Eisen & H. S. Ginsberg, pp. 645–672. New York: Harper & Row.

Suelter, C. H. & Snell, E. E. (1977). *J. Biol. Chem.* **252**, 1852–1857.

Suelter, C. H., Wang, J. & Snell, E. E. (1976). *Anal. Biochem.* **76**, 221–232.

Tani, S., Tsujimoto, N., Kawata, Y. & Tokushige, M. (1990). *Biotechnol. Appl. Biochem.* **12**, 28–33.

Wakabayashi, H., Wakabayashi, M., Eisenreich, W. & Engel, K. H. (2004). *J. Agric. Food Chem.* **52**, 110–116.

Wang, D., Ding, X. & Rather, P. N. (2001). *J. Bacteriol.* **183**, 4210–4216.

Watanabe, T. & Snell, E. E. (1972). *Proc. Natl Acad. Sci. USA*, **69**, 1086–1090.

Watanabe, T. & Snell, E. E. (1977). *J. Biochem. (Tokyo)*, **82**, 733–745.

Yang, X. & Borchardt, R. T. (2000). *Arch. Biochem. Biophys.* **383**, 272–280.

Zakomirdina, L. N., Kulikova, V. V., Gogoleva, O. I., Dementieva, I. S., Faleev, N. G. & Demidkina, T. V. (2002). *Biochemistry (Mosc.)*, **67**, 1189–1196.