Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of *Escherichia coli* uridine phosphorylase at 2.0 Å

F. Temple Burling,† Ryan Kniewel, John A. Buglino, Tanya Chadha, Andrew Beckwith and Christopher D. Lima*

Biochemistry Department and Structural Biology Program, Weill Medical College of Cornell University, New York, NY 10021, USA

† Current address: Department of Physics, Carthage College, Kenosha, WI 53140, USA.

Correspondence e-mail: lima@pinky.med.cornell.edu

Received 30 July 2002
Accepted 15 October 2002

PDB Reference: uridine phosphorylase, 1lx7, r1lx7sf.

The 2.0 Å crystal structure has been determined for *Escherichia coli* uridine phosphorylase (UP), an essential enzyme in nucleotide biosynthesis that catalyzes the phosphorolytic cleavage of the C—N glycosidic bond of uridine to ribose-1-phosphate and uracil. The structure determination of two independent monomers in the asymmetric unit revealed the residue composition and atomic details of the apo configurations of each active site. The native hexameric UP enzyme was revealed by applying threefold crystallographic symmetry to the contents of the asymmetric unit. The 2.0 Å model reveals a closer structural relationship to other nucleotide phosphorylase enzymes than was previously appreciated.

1. Introduction

Uridine phosphorylase catalyzes the phosphorolytic cleavage of the C—N glycosidic bond of uridine to ribose-1-phosphate and uracil (Leer *et al.*, 1977; Vita *et al.*, 1986). Uridine phosphorylase is a member of the pyrimidine nucleoside phosphorylase (PyNP) class of enzymes that catalyze the general reaction

$$\text{pyrimidine nucleoside} + \text{phosphate} \longleftrightarrow \text{ribose-1-phosphate} + \text{pyrimidine base}.$$

Functionally related to the PyNP proteins are the purine nucleoside phosphorylase (PNP) enzymes that catalyze the analogous reaction

$$\text{purine nucleoside} + \text{phosphate} \longleftrightarrow \text{ribose-1-phosphate} + \text{purine base}.$$

Together, these two enzyme types comprise the nucleoside phosphorylase (NP) class of proteins. Nucleoside phosphorylases are involved in essential biochemical salvage pathways in the cell that provide free purine and pyrimidine bases for subsequent nucleotide biosynthesis, enabling a less costly alternative to *de novo* nucleotide biosynthesis.

Nucleoside phosphorylases are also able to inactivate certain purine and pyrimidine nucleoside analogs that posses anti-tumor activity (Morgunova *et al.*, 1995; Pugmire & Ealick, 2002). Thus, the discovery of selective inhibitors for both PyNP and PNP enzymes could lead to enhanced therapeutic activity for these nucleoside analogs. More detailed views provided by high-resolution structures of the active sites for both PyNP and PNP enzymes could aid in the development of such compounds. To this end, we report the determination of a 2.0 Å apo structure of *Escherichia coli* uridine phosphorylase (UP; EC 2.4.2.3). The New York Structural Genomics Research Consortium (NYSGRC; http://www.nysgrc.org) has

targeted highly conserved enzyme families for structure determination as part of the national effort in structural genomics. At the time of target selection, UP represented a highly conserved protein family not represented in the PDB, so it was selected as a target for crystallographic structure determination.

2. Materials and methods

2.1. Isolation, expression, crystallization and structure determination

UP was selected as a unique target for our structural genomics effort by virtue of its conservation (>25% sequence identity) between several bacterial, human and mouse genomes. The coding region for the *E. coli* UP enzyme was amplified from *E. coli* genomic DNA by PCR, ligated into a modified version of pET28b, expressed in *E. coli* BL21 DE3 Codon Plus RIL (Stratagene) and purified using Ni–NTA–agarose resin (Qiagen). UP was further purified by gel filtration (Superdex200, Pharmacia), eluting with an apparent molecular weight consistent with a hexamer. Fractions were analyzed by SDS–PAGE, pooled and concentrated to 10.0 mg ml^{-1} (10 m*M* Tris–HCl pH 8, 50 m*M* NaCl, 1 m*M* DTT). Selenomethionine-substituted (SeMet) UP was generated by expressing UP in B834(DE3) cells (Hendrickson *et al.*, 1990).

96-well crystallization trials were conducted that produced diffraction-quality crystals. SeMet UP protein crystals were refined and grown by hanging-drop vapor diffusion against a well solution containing 10% PEG 4K, 0.1 *M* MES pH 6.5 and 5% glycerol to final dimensions of 0.3 × 0.3 × 0.3 mm. The data presented here were obtained from UP crystallized in space group *R*3 (unit-cell parameters $a = b = 151.4$, $c = 48.2$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$). Diffraction data collection was accomplished using cryopreserved crystals (30% glycerol in mother liquor).

Diffraction experiments took place at beamline X9A at the National Synchrotron Light Source and the data were processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) and input to *SOLVE*, *RESOLVE* (Terwilliger & Berendzen, 1999) and the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) to calculate a 2.0 Å SAD phase set. *RESOLVE* automatically determined the twofold NCS operators relating the *A* and *B* monomers and this solvent-modified NCS-averaged electron density was manually traced using *O* (Jones *et al.*, 1991; see Table 1). The protomer model was initially fitted into the two respective positions within the asymmetric unit using *BRUTPTF* (see http://www.nysgrc.org/molrep for details) and the subsequent model containing two monomers was refined without NCS restraints in *CNS* (Brünger *et al.*, 1998).

The final model contained 474 amino-acid residues comprised of residues 4–253 from monomer *A* and residues 3–162, 183–221 and 233–253 from monomer *B*. The structure revealed two independent determinations of the active site since the model contained two monomers in the asymmetric unit. The native oligomeric state for *E. coli* UP is a hexamer and this oligomeric state can be reconstructed by transforming the *A* and *B* monomers within the asymmetric unit by the *R*3 crystallographic threefold axis (for details, see PDB code 1lx7). A similar oligomeric state for *E. coli* UP has been described in previous reports (Zhao, 1991; Morgunova *et al.*, 1995). Additional information on the expression and crystallization of *E. coli* UP can be found at http://www.nysgrc.org under target-identification code T24.

Table 1
Crystallographic data and refinement statistics.

Crystal characteristics and data-collection statistics	
Unit-cell parameters (Å, °)	$a = b = 151.4$, $c = 48.2$, $\alpha = \beta = \gamma = 90$
Space group	*R*3
Molecules per asymmetric unit	2
X-ray source	NSLS X9A beamline
λ_1 (SeMet peak)	
Wavelength (Å)	0.9790
Resolution (Å)	20.0–2.0
No. of observations	747216
No. of reflections†	49742
Completeness‡ (%)	89.2 (78.6)
Mean $I/\sigma(I)$‡	27.8 (15.6)
$R_{\rm merge}$ on I‡§	3.8 (8.0)
Cutoff criteria	$I < 0\sigma(I)$
SOLVE figure of merit¶	0.29 (20.0–2.0 Å resolution) for 24381 reflections
RESOLVE figure of merit¶	0.53 (20.0–2.0 Å resolution) for 24381 reflections
Model and refinement statistics	
Data set used in structure refinement	Structure factors derived from *SOLVE*
Resolution range (Å)	20.0–2.0
No. of reflections	26441 (25126 in working set; 1315 in test set)
Completeness (%)	95.1 (90.3 in working set; 4.7 in test set)
Cutoff criterion	$\|F\| > 0.0$
No. of amino-acid residues/atoms	474/3512
No. of water atoms	361
$R_{\rm cryst}$††‡‡	0.182 (0.204)
$R_{\rm free}$‡‡	0.215 (0.234)
Root-mean-square deviations	
Bond lengths (Å)	0.006
Bond angles (Å)	1.20
B factor main chain/side chain (Å^2)	1.26/2.34
Ramachandran plot statistics§§	
Residues in most favored regions	366 (90.1%)
Residues in additional allowed regions	38 (9.4%)
Residues in generously allowed regions	2 (0.5%)
Residues in disallowed regions	0 (0.0%)

† MAD data completeness treats Bijvoët mates independently. ‡ Values in parentheses are for the highest resolution shell (2.07–2.0 Å). § $R_{\rm merge} = \sum_{hkl}\sum_i |I(hkl)_i - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i \langle I(hkl)_i\rangle$. ¶ Figure of merit calculated using *SOLVE*/*RESOLVE*. †† $R_{\rm cryst} = \sum_{hkl} |F_o(hkl) - F_c(hkl)| / \sum_{hkl} |F_o(hkl)|$, where F_o and F_c are the observed and calculated structure factors, respectively. ‡‡ Values in parentheses are for the highest resolution shell (2.13–2.0 Å). §§ Computed with *PROCHECK* (*CCP*4 suite; Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

3.1. Structure of *E. coli* UP

E. coli UP was expressed, purified, crystallized and characterized by X-ray crystallography (see §2). A previous 2.5 Å structure of *E. coli* UP was determined in a monoclinic space group and published (Morgunova *et al.*, 1995), although the

data were not deposited in the PDB until this year (PDB code 1k3f). Additionally, a 3.0 Å structure of *E. coli* UP determined in a trigonal space group was described in a PhD thesis, but the coordinates were never deposited in the PDB (Zhao, 1991). The 2.5 Å structure of *E. coli* UP was utilized in a comparison with several purine nucleoside phosphorylase structures in a recent review (Pugmire & Ealick, 2002) including *E. coli* PNP (Mao *et al.*, 1998; PDB code 1ecp), bovine PNP (bPNP) and human PNP (hPNP) (Pugmire & Ealick, 2002). *E. coli* UP and *E. coli* PNP share only 26.7% sequence identity and *E. coli* UP and *E. coli* PNP both align with less than 20% sequence identity to either bPNP or hPNP (Pugmire & Ealick, 2002). Nonetheless, these four proteins share a common fold and are part of a family of proteins that Pugmire and Ealick term the nucleoside phosphorylase-I (NP-I) family. The striking similarity in fold between *E. coli* UP (a PyNP) and *E. coli* PNP led Pugmire and Ealick to propose that these two proteins evolved from a common ancestor. It is interesting to note that a second family of nucleoside phosphorylases, termed the NP-II family by Pugmire and Ealick, are structurally distinct from the NP-I enzymes. The NP-II family consists of enzymes that are specific for thymine in higher organisms, but will catalyze nucleoside phosphorylation reactions of both thymine and uridine in lower organisms (Pugmire & Ealick, 2002).

Neither the 2.0 Å *E. coli* UP structure reported here nor the previously reported 2.5 Å *E. coli* UP structure contains substrate at the active site. The *E. coli* PNP structure is also unbound. However, there are several examples of substrate-bound and substrate-analog-bound structures of bPNP (Morgunova *et al.*, 1995; Pugmire & Ealick, 2002). The exceptional conservation of protein folds between the bacterial NPs and the mammalian PNPs permits identification of the probable active-site residues of *E. coli* UP and *E. coli* PNP based on three-dimensional alignment with substrate-bound bPNP structures.

We aligned our 2.0 Å UP structure to the 2.5 Å UP structure and the structure of *E. coli* PNP using *DALI* (Holm & Sander, 1993). The root-mean-square deviations (r.m.s.d.) between monomers of the 2.5 Å UP structure and monomer *A* from our structure ranged from 1.1 to 1.5 Å^2 over 236 amino acids (excluding disordered loops). The alignment of monomer *A* from our *E. coli* UP structure and monomer *A* from the *E. coli* PNP structure revealed an overall r.m.s.d. of 2.0 Å^2 over 223 amino acids. Based on these alignments, we find several striking differences between the 2.0 Å structure of *E. coli* UP reported here and the previous 2.5 Å *E. coli* UP structure in the vicinity of the putative active site (Fig. 1). Specifically, four critical active-site residues in the 2.0 Å structure are observed in similar conformations to those in the *E. coli* PNP structure. These observations are in contrast to the structure of the active-site residues observed in the 2.5 Å *E. coli* UP structure. Greater structural homology observed here between *E. coli* UP and *E. coli* PNP lends further support to the hypothesis of Pugmire and Ealick that these enzyme classes evolved from a common ancestor.

3.2. *E. coli* UP active site

The additional structural homologies observed between the 2.0 Å structure of *E. coli* UP and *E. coli* PNP are as follows. In the previous 2.5 Å *E. coli* UP structure, the authors note that Arg48 is disordered. The corresponding Arg43 in *E. coli* PNP is ordered and reaches into the phosphate-binding site of an adjacent subunit. In the 2.0 Å structure reported here, Arg48 appears well ordered and in a conformation similar to that of Arg43 in the *E. coli* PNP structure. In addition, His8 from monomer *B* was observed in a conformation essentially identical to that of His4 in *E. coli* PNP. The 2.0 Å structure also suggests that Glu196 has a similar conformation to Glu179, as observed in *E. coli* PNP. The previous 2.5 Å *E. coli* UP structure revealed Glu196 to be in a different conformation from that observed in the 2.0 Å structure.

The most striking differences observed between the two *E. coli* UP structures are the divergent conformations between the loops comprised of amino-acid residues 163–177. The loop conformation observed in the 2.0 Å structure positions two residues, Gln166 and Arg168, into the putative active-site cleft, amino acids that point out of and away from the binding site in the previous structure. Tyr163 is also located within this loop. In the 2.0 Å structure, Tyr163 superimposes well with *E. coli* PNP Tyr160, whereas Tyr163 in the 2.5 Å structure

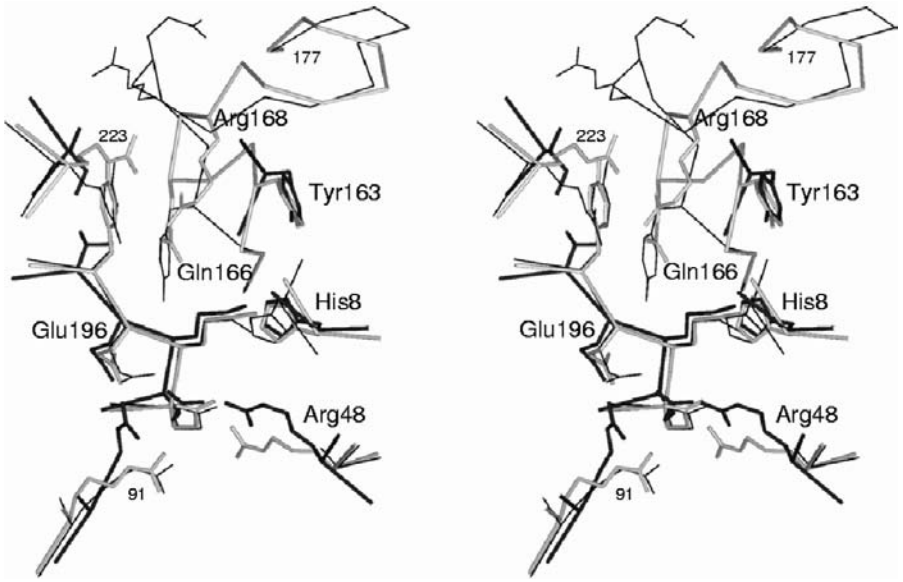

Figure 1
Stereo diagram of a representative UP active site. Three NP-I family members are represented: the 2.0 Å structure of *E. coli* UP reported here (in grey thick lines), the previously reported 2.5 Å structure of *E. coli* UP (in thin black lines) and the 2.0 Å structure of *E. coli* PNP (in thick black lines). Specific amino-acid residues mentioned in the text are numbered for the 2.0 Å structure of *E. coli* UP. This figure was prepared using *SETOR* (Evans, 1993).

superimposes onto Gln166 of the 2.0 Å structure and would be predicted to make very close contacts to a modeled pyrimidine ring within the putative ligand-binding site.

4. Conclusions

Homology-modeling studies of *E. coli* and mammalian purine phosphorylase along with *E. coli* uridine phosphorylase structures reveal a common fold and permit identification of active-site residues in the apo uridine phosphorylase (Pugmire & Ealick, 2002). The 2.0 Å resolution structure of the *E. coli* UP suggests greater structural similarity to *E. coli* purine phosphorylase than was previously surmised based on an earlier 2.5 Å structure of *E. coli* uridine phosphorylase (Morgunova *et al.*, 1995). Combined with previous observations, the high-resolution structure of *E. coli* UP lends further support to the hypothesis that members of the related but functionally distinct NP-I enzyme class evolved from a common ancestor.

We thank the staff of beamline X9A at the National Synchrotron Light Source for their support. This work was supported in part by a Young Investigator award made to CDL from the Arnold and Mabel Beckman Foundation and NIH structural genomics pilot center grant 1P50 GM62529.

References

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.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Evans, S. V. (1993). *J. Mol. Graph.* **11**, 134–138.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.

Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–118.

Leer, J. C., Hammer-Jespersen, K. & Schwartz, M. (1977). *Eur. J. Biochem.* **75**, 217–224.

Mao, C., Cook, W. J., Zhou, M., Federov, A. A., Almo, S. C. & Ealick, S. E. (1998). *Biochemistry*, **37**, 7135–7146.

Morgunova, E. Y., Mkhailov, A. M., Popov, A. N., Blagova, E. V., Smirnova, E. A., Vainshtein, B. K., Mao, C., Armstrong, S. R., Ealick, S. E., Komissarov, A. A., Linkova, E. V., Burlakova, A. A., Mironov, A. S. & Debabov, V. G. (1995). *FEBS Lett.* **367**, 183–187.

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

Pugmire, M. J. & Ealick, S.A. (2002). *Biochem. J.* **361**, 1–25.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Vita, A., Amici, A., Cacciamani, T., Lanciotti, M. & Magni, G. (1986). *Int. J. Biochem.* **18**, 431–436.

Zhao, B. (1991). PhD dissertation, University of Alabama at Birmingham, USA.