Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of product-bound *Bacillus caldolyticus* uracil phosphoribosyltransferase confirms ordered sequential substrate binding

Anders Kadziola,[a]* Jan Neuhard[b] and Sine Larsen[a]

[a]Centre for Crystallographic Studies, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark, and [b]Department of Biological Chemistry, Institute of Molecular Biology, University of Copenhagen, Sølvgade 83H, DK-1307 Copenhagen, Denmark

Correspondence e-mail: anders@ccs.ki.ku.dk

Received 26 November 2001
Accepted 18 March 2002

PDB Reference: uracil phosphoribosyltransferase, 1i5e, r1i5esf.

Uracil phosphoribosyltransferase (UPRTase) is part of the salvage pathway that leads to the biosynthesis of UMP. It catalyzes the formation of UMP and pyrophosphate from uracil and α-D-5-phosphoribosyl-1-pyrophosphate. Unlike enzymes in the *de novo* synthesis of UMP, UPRTases have only been found in lower organisms and are therefore potential targets for the development of new antibiotics. UPRTase from *Bacillus caldolyticus* has been crystallized and the structure has been determined by isomorphous replacement and refined to 3.0 Å resolution. UPRTase from *B. caldolyticus* forms a dimer with the active sites pointing away from each other. A long arm from each subunit wraps around the other subunit, contributing half of the dimer interface. The monomer adopts the phosphoribosyltransferase type I fold, with a small C-terminal hood defining the uracil-binding site. The structure contains a well defined UMP molecule in the active site. The binding of UMP involves two sequence segments that are highly conserved among UPRTases. The first segment, Asp131–Ser139, contains the PRPP-binding consensus sequence motif known from other type I phosphoribosyltransferases and binds the ribose-5′-phosphate part of UMP. The second segment, Tyr193–Ala201, which is specific for uracil phosphoribosyltransferases, binds the uracil part of UMP through backbone contacts, partly mediated by a water molecule. Modelling of a PRPP–enzyme complex reveals that uracil can be activated to its tautomeric enol form by the complex. This is consistent with kinetic data, which display ordered sequential binding of substrates, with PRPP binding first. Based on this observation, a reaction mechanism is proposed.

1. Introduction

UMP is the common precursor of all pyrimidine nucleotides and may be synthesized either *de novo* from intermediates of central metabolism or *via* salvage of preformed pyrimidine nucleobases or nucleosides. The salvage pathways represent a significant energy saving for the cell and they define the sensitivity of the cell towards cytotoxic pyrimidine analogues (Neuhard, 1983). A key enzyme in microbial pyrimidine salvage is uracil phosphoribosyltransferase (UPRTase[1]). It catalyzes the formation of UMP and pyrophosphate from uracil and α-D-5-phosphoribosyl-1-pyrophosphate (PRPP) in

[1] Abbreviations: PRTase, phosphoribosyltransferase; UPRTase, uracil PRTase; OPRTase, orotate PRTase; H/G/XPRTase, hypoxanthine/guanine/xanthine PRTase; GPATase, glutamine phosphoribosylpyrophosphate amidotransferase; APRTase, adenine PRTase; PRPP, α-D-5-phosphoribosyl-1-pyrophosphate; R5P, ribose-5′-phosphate; PP_i, pyrophosphate; UMP, uridine-5′-phosphate; IMP, inosine-5′-phosphate; PEG, polyethylene glycol; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid); Orange-Pt, chloro(2,2′:6′,2″-terpyridine)platinum(II) chloride.

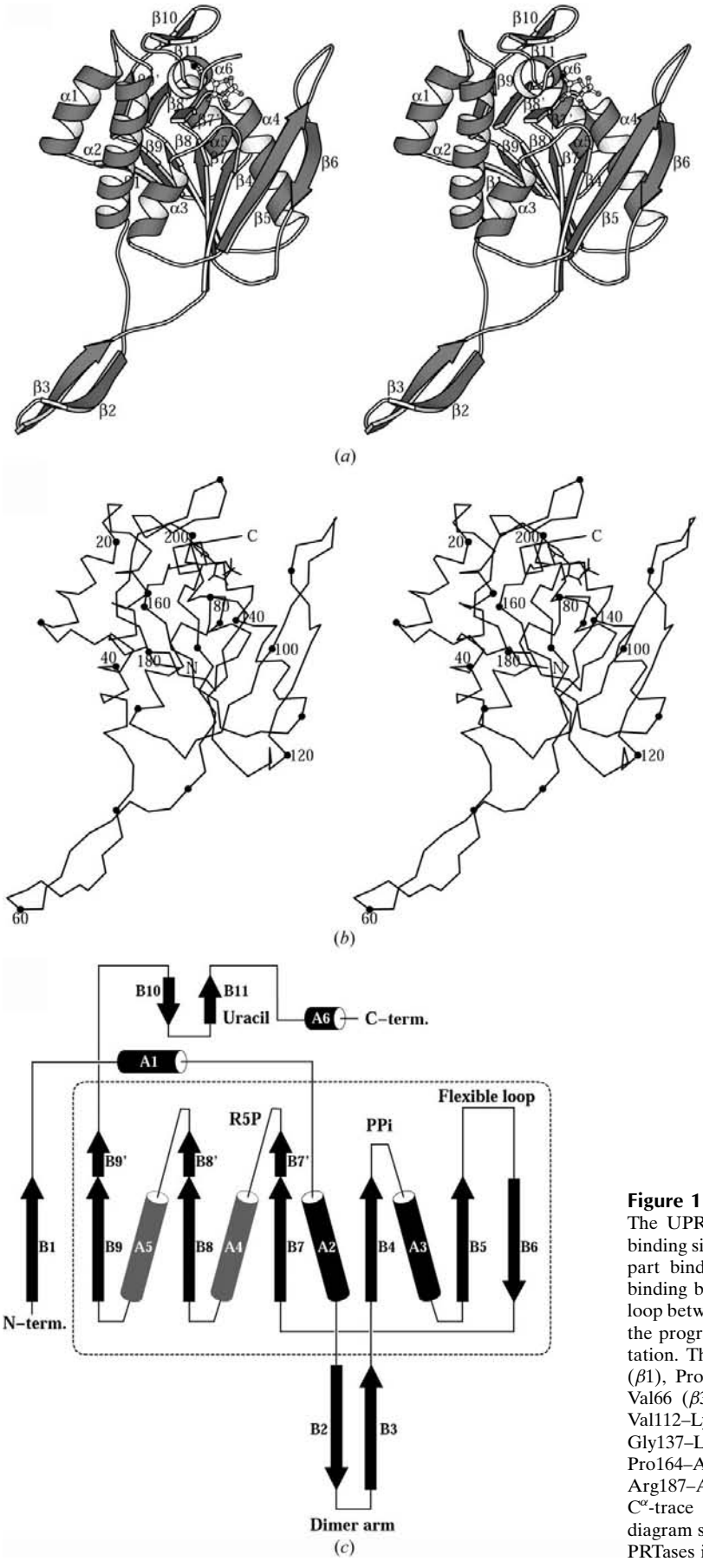

Figure 1
The UPRTase monomer fold with UMP. The core defines the PRPP-binding site common to all PRTasesm with the ribose-5′-phosphate (R5P) part binding between $\beta7'$ and $\alpha4$ and the pyrophosphate (PP_i) part binding between $\beta4$ and $\alpha3$. The uracil-binding site is defined by a long loop between $\beta9'$ and $\alpha6$. This figure and Figs. 6 and 7 were produced with the program *MOLSCRIPT* (Kraulis, 1991). (*a*) Ribbon stereo representation. The secondary-structure assignments are as follows: Ala3–Phe7 ($\beta1$), Pro10–Arg20 ($\alpha1$), Thr26–Ile44 ($\alpha2$), Glu51–Thr57 ($\beta2$), Ser60–Val66 ($\beta3$), Lys71–Ile77 ($\beta4$), Met84–Leu91 ($\alpha3$), Lys96–Tyr103 ($\beta5$), Val112–Lys117 ($\beta6$), Asp126–Asp131 ($\beta7$), Pro132–Leu134 ($\beta7'$), Gly137–Lys149 ($\alpha4$), Ser154–Cys159 ($\beta8$), Leu160–Ala163 ($\beta8'$), Pro164–Ala173 ($\alpha5$), Asp178–Ile181 ($\beta9$), Ala182–Asp185 ($\beta9'$), Arg187–Asn189 ($\beta10$), Tyr193–Val195 ($\beta11$), Ala201–Phe206 ($\alpha6$). (*b*) C^{α}-trace stereoview with dots for every ten residues. (*c*) Topology diagram showing the UPRTase fold. The PRTase core common to type I PRTases is framed with a dashed line.

the presence of Mg^{2+}. The phosphoribosyl transfer occurs with inversion at the anomeric C1 carbon of the ribose ring.

The enzyme has been partially characterized from eubacteria (Jensen & Mygind, 1996; Jensen *et al.*, 1997; Jyssum & Jyssum, 1979; McIvor *et al.*, 1983; Michell & Finch, 1979), archaea (Linde & Jensen, 1996) and lower eukaryotes (Alloush & Kerridge, 1994; Asai *et al.*, 1990; Carter *et al.*, 1997; Dai *et al.*, 1995; Natalini *et al.*, 1979; Plunkett & Moner, 1978). It appears to be absent from higher eukaryotes (Traut & Jones, 1996), although it has been reported to be present in pea seedlings (Bressan *et al.*, 1978).

The UPRTases from the Gram-negative bacterium *Escherichia coli* and from the Gram-positive thermophile *Bacillus caldolyticus* are the only prokaryotic UPRTases that have been purified to homogeneity (Jensen & Mygind, 1996; Jensen *et al.*, 1997). They show 45% amino-acid sequence identity and both were found to obey the kinetics of a sequential reaction mechanism, with PRPP binding before uracil and UMP leaving last (Lundegaard & Jensen, 1999; Jensen & Neuhard, unpublished experiments). The true K_M values for uracil and PRPP were found to be 0.5 and 58 μM, respectively, for the *E. coli* enzyme and 0.9 and 37 μM, respectively, for the *B. caldolyticus* enzyme. Despite these similarities, the two enzymes differ in quarternary structure in solution. Thus, the *E. coli* enzyme appears to be a pentamer or hexamer in its active form (Lundegaard & Jensen, 1999), whereas the enzymatically active form of the *B. caldolyticus* enzyme is a dimer (Jensen *et al.*, 1997).

UPRTases appear to belong to the type I phosphoribosyltransferases (PRTases), which contain a conserved 13-residue sequence segment termed the PRPP-binding motif (Hershey & Taylor, 1986; Hove-Jensen *et al.*, 1986). Type I PRTases are involved in the *de novo* synthesis of IMP and UMP, in the salvage of purines and pyrimidines and in the biosynthesis of the amino acids histidine and tryptophan (Jensen, 1983). Type II PRTases, which include the nicotinate and quinolinate phosphoribosyltransferases involved in pyridine nucleotide synthesis, do not show the conserved PRPP-binding motif in their amino-acid sequences (Eads *et al.*, 1997).

The three-dimensional structures of a number of type I PRTases have been determined, including OPRTases from *Salmonella typhimurium* (Scapin *et al.*, 1994, 1995) and *E. coli* (Henriksen *et al.*, 1996), GPATases from *B. subtilis* (Smith *et al.*, 1994; Chen *et al.*, 1997) and *E. coli* (Krahn *et al.*, 1997; Muchmore *et al.*, 1998), H/G/XPRTases from human (Eads *et al.*, 1994; Shi, Li, Tyler, Furneaux, Grubmeyer *et al.*, 1999; Balendiran *et al.*, 1999) and from the protozoan parasites *Tritrichomonas foetus* (Somoza *et al.*, 1996), *E. coli* (Vos *et al.*, 1997, 1998), *Trypanosoma cruzi* (Focia, Craig, Nieves-Alicia *et al.*, 1998; Focia, Craig & Eakin, 1998), *Toxoplasma gondii* (Schumacher *et al.*, 1999; Héroux *et al.*, 1999, 2000) and *Plasmodium falciparum* (Shi, Li, Tyler, Furneaux, Cahill *et al.*, 1999). The structures of UPRTase from *Tox. gondii* (Schumacher *et al.*, 1998) and APRTase from *Leishmania donovani* (Phillips *et al.*, 1999) have also recently been determined. The type I PRTases (Fig. 1) share a characteristic fold consisting of a core region with a flexible loop and a hood. The hood supplies the residues required for binding of the nucleobase and is either formed by N-terminal residues (OPRTase, APRTase) or by C-terminal residues (HGPRTase, UPRTase). The active-site cleft is situated between the hood and the core harbouring the PRPP-binding sequence. It has been proposed that enzyme-catalyzed phosphoribosyl-group transfer involves an oxocarbenium-ion intermediate/transition state (Goitein *et al.*, 1978; Tao *et al.*, 1996) which would require protection from solvent to prevent PRPP hydrolysis. The flexible loop is proposed to close down onto the active site during catalysis to protect the intermediate/transition state from hydrolysis (Henriksen *et al.*, 1996; Schumacher *et al.*, 1999). In the two-domain structure of PRPP synthetase, each domain adopts the type I PRTase fold even though the enzyme is not a PRTase (Eriksen *et al.*, 2000).

In the present paper, we describe the crystal structure of *B. caldolyticus* UPRTase bound to its product, UMP, at 3.0 Å resolution. Based on the structure, we have modelled an enzyme–substrate complex and propose a possible reaction mechanism.

2. Materials and methods

2.1. Crystallization

The recombinant UPRTase from *B. caldolyticus* used for crystallization was prepared as described by Jensen *et al.* (1997). Crystals were grown at room temperature by vapour diffusion using hanging or sitting drops. An initial crystal screen was set up as described by Jancarik & Kim (1991) and manufactured by Hampton Research. The protein solution contained UPRTase at a concentration of 20 mg ml^{-1} with 5 m*M* inorganic phosphate at pH 7.0 and hanging drops were composed of 3 µl protein solution and 3 µl reservoir. The best results from the crystal screen were hair-like structures and small needle-shaped crystals which appeared under conditions that had in common pH 7.5/8.5 and PEG 4000/6000 as precipitant. The crystallization conditions were optimized to give large three-dimensional crystals (Fig. 2). The crystals used for structure determination were grown at 293 K with the above-mentioned protein solution. Sitting drops were prepared from 7 µl protein solution and 7 µl reservoir solution from a 1 ml reservoir composed of 6–10% PEG 4000 with 0.1 *M* HEPES pH 7.5. Usually, crystals appeared after three weeks of equilibration. The maximal dimensions of the crystals, which appeared as truncated hexagonal bipyramids, were about 1.0 × 0.5 × 0.5 mm. Often, hair-like structures were observed in addition to crystals in the drop (Fig. 2).

With our in-house equipment (see below), the best crystals diffracted to about 3.6 Å resolution. The crystals belong to space group $P3_121$, with unit-cell parameters $a = b = 89.4$, $c = 163.7$ Å. Assuming a molecular weight of 22 817 Da and two molecules per asymmetric unit, V_M (Matthews, 1968) is calculated to be 4.1 Å^3 Da^{-1}, corresponding to 70% solvent in the crystal. This is consistent with a self-rotation function calculated with the program *AMoRe* (Collaborative Computational Project, Number 4, 1994; Navaza, 1994), which indicated the existence of a non-crystallographic twofold axis.

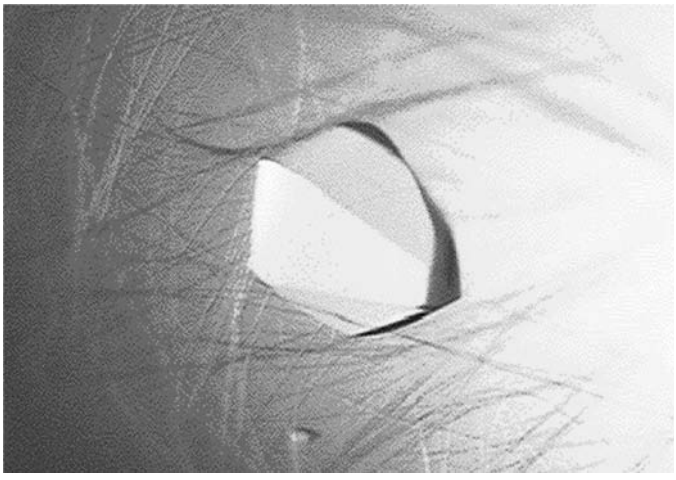

Figure 2
Typical morphological habit of the *B. caldolyticus* UPRTase crystals. At the macroscopic level, hair-like structures are often observed along with crystals. These fibre-like structures are also seen at the microscopic level in the crystal packing, which can be described as a stacking of chains formed by closely interacting dimers. Crystals were grown at 293 K by vapour diffusion from a 20 mg ml^{-1} protein solution containing 5 m*M* inorganic phosphate pH 7. Sitting drops were composed of 7 µl of the above-mentioned protein solution and 7 µl reservoir solution composed of 6–10% PEG 4000 and 0.1 *M* HEPES pH 7.5. Crystals appeared after about three weeks of equilibration against 1 ml reservoir solution.

Table 1
Data-collection statistics.

Values in parentheses are for the outermost resolution shell.

Data set	Native, in-house	$HgCl_2$, in-house	Orange-Pt, in-house	Native, synchrotron
Wavelength (Å)	1.5418	1.5418	1.5418	0.888
Resolution (Å)	40–3.6 (3.79–3.60)	40–4.5 (4.74–4.50)	40–4.0 (4.21–4.00)	40–3.0 (3.05–3.00)
Unique reflections	9126 (1305)	4569 (663)	6246 (919)	13156 (680)
Completeness (%)	99.4 (100.0)	95.0 (97.9)	93.9 (96.8)	85.5 (90.3)
Redundancy	5.5 (6.0)	3.7 (3.7)	3.5 (3.5)	4.8 (4.1)
$R_{\rm merge}$† (%)	10.7 (34.3)	19.6 (28.3)	12.0 (34.9)	7.1 (31.0)
$\langle I\rangle/\langle\sigma(I)\rangle$	6.6 (2.2)	3.6 (2.6)	5.7 (2.1)	21.4 (6.6)

† $R_{\rm merge} = \sum|I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections.

2.2. Derivatization and data collection

Derivatives were prepared by soaking with either 1 m*M* $HgCl_2$ or 0.5 m*M* Orange-Pt (Lawson, 1994) for one week by adding tiny amounts of the compound dissolved in water directly to the drops containing the crystals.

Data for one native crystal and the two soaked crystals mentioned above were collected at 288 K with our in-house equipment: an R-AXIS II imaging-plate system mounted on a Rigaku Rotaflex RU-200 copper rotating-anode generator operating at 50 kV and 180 mA with a graphite monochromator and a 0.5 mm collimator. The crystal-to-detector distance was 100 mm and images consisting of 2° oscillations were exposed for 30 min. Autoindexation and integration of intensities were performed with the program *DENZO* from the *HKL* package (Gewirth, 1994) and data reduction was performed with the programs *ROTAVATA* and *AGROVATA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Data-collection statistics are given in Table 1.

Owing to the relatively modest resolution obtained with the in-house equipment, an additional synchrotron data set was recorded at EMBL in Hamburg. Owing to difficulties in finding suitable cryoconditions, the crystal was cooled to 263 K without freezing; using a wavelength of 0.888 Å, diffraction extended to 3 Å resolution. Data were collected at a crystal-to-detector distance of 270 mm and each image was collected as a 1° oscillation with a dose of 3000 counts s^{-1}. Integration of intensities and data reduction was performed with the programs *DENZO* and *SCALEPACK*, respectively, from the *HKL* package (Gewirth, 1994). Data-collection statistics are again given in Table 1.

2.3. Structure determination

The structure was determined with the three in-house data sets mentioned above by multiple isomorphous replacement (MIR) in combination with direct-space molecular replacement using the OPRTase structure (Scapin *et al.*, 1994) as a model.

The isomorphous difference Patterson map for the Hg derivative was readily interpreted in terms of two major sites in the asymmetric unit. This method was not directly applicable to the Pt derivative; however, a difference Fourier map cross-phased with the Hg derivative revealed four sites in the asymmetric unit consistent with the difference Patterson map. Later in the procedure, two additional minor Hg sites were found. After having solved the structure, it turned out that the two major sites for Hg represented binding to Cys159*A* and Cys159*B* buried in the molecule and that the two minor sites were at His191*A* and His191*B* at the surface of the molecule. His191*A* and His191*B* were also the locations of all four Pt sites, indicating double conformations of His191*A* and His191*B* in this derivative.

Heavy-atom refinement and phase calculation were performed with the program *MLPHARE* (Collaborative Computational Project, Number 4, 1994; Otwinowski, 1991) using both isomorphous and anomalous data. Coordinates as well as real and anomalous occupancies were refined. *B* factors were fixed to 20 $Å^2$. The phasing ability of the Hg derivative was clearly much better than the Pt derivative. The handedness was determined with anomalous dispersion. Parallel refinement of coordinates x, y, z in space group $P3_121$ and coordinates $-x, -y, -z$ in space group $P3_221$ gave clear positive and negative anomalous occupancies, respectively, suggesting that the correct space group is $P3_121$.

The initial MIR phases, which extended to about 4.5 Å resolution, were improved with solvent flattening and histogram matching using the program *DM* (Collaborative Computational Project, Number 4, 1994). This produced an electron-density map which was difficult to interpret at first glance. The map skeleton calculated with *MAPMAN* (Kleywegt & Jones, 1996) did not show any helical features, which was surprising as the UPRTase was expected to be similar to the OPRTase structure, which does contain helices. The missing skeleton helices turned out to be an artifact arising from the modest resolution. The map did contain well defined helices, but the maximum density had a tendency to be concentrated at the helical axes, giving discontinuous lines in the skeleton instead of helical features. With some effort, two OPRTase monomers could manually be rotated and translated to give an approximate fit to the density by placing the C^α trace onto the skeleton. The positioning was consistent with both the positions of the heavy-atom sites and the results from the self-rotation function for the native data. From this starting structure, which contains important information about the overall folding and connectivity, a tentative polyalanine model of UPRTase was built into the electron-density map. This model was refined with *X-PLOR* (Brünger, 1992) with phases restrained to the *DM* values. The resulting calculated phases were combined with the already existing MIR phases using *SIGMAA* (Collaborative Computational Project, Number 4, 1994; Read, 1986) and again exposed to density modification with *DM*, now including twofold averaging of the

electron density. With this phase information for all reflections to a maximum resolution of 3.6 Å, an improved map was produced in which side-chain electron densities could be assigned to the sequence and the entire UPRTase model comprising 2 × 208 amino-acid residues was completed.

2.4. Refinement

A synchrotron data set extending to 3.0 Å resolution was collected at the EMBL in Hamburg from a capillary-mounted single crystal cooled to 263 K. Data-collection statistics are given in Table 1. With this data set, the structure was initially refined using the program *X-PLOR* (Brünger, 1992), while later refinements were performed with the program *CNS* (Brünger *et al.*, 1998) using twofold NCS restraints and bulk-solvent modelling. Between cycles of refinement the structure was manually corrected based on NCS-averaged $3mF_o - 2DF_c$ and $mF_o - DF_c$ maps (Collaborative Computational Project, Number 4, 1994; Read, 1986) using the graphics-display program *TURBO-FRODO* (Roussel & Cambillau, 1992). The final model consists of a dimer in the asymmetric unit, with each subunit consisting of amino acids Gly2–Lys209, one UMP molecule and a water molecule. This model with neighbour-restrained *B*-factor optimization refined to an *R* factor of 22.3% and an $R_{\rm free}$ of 27.0%. The model shows good stereochemistry, with r.m.s. deviations from standard values of 0.007 Å for bond distances and 1.43° for bond angles. A Ramachandran plot calculated with the program *PROCHECK* (Laskowski *et al.*, 1993) showed 83.3% of non-glycine and non-proline residues in the most favoured regions and 16.1% in the additionally allowed regions. Only one residue, Leu48, was found just outside in the generously allowed regions.

Figure 3
Stereoview rigid-body superpositioning of *T. gondii* UPRTase (^{TG}UPRTase) on *B. caldolyticus* UPRTase (^{BC}UPRTase) The two variants have a different twist, which means that only one subunit can be superposed at a time. The ^{TG}UPRTase (with bound UMP; PDB code 1upu) subunit *A* is superposed the ^{BC}UPRTase subunit *A*. The segments ^{TG}Asn35–^{TG}Met39, ^{TG}Thr42–^{TG}Leu78, ^{TG}Tyr102–^{TG}Arg126, ^{TG}Arg129–^{TG}Ile135, ^{TG}Lys145–^{TG}Lys186, ^{TG}Arg189–^{TG}Thr243 and ^{TG}U5P999 were fitted to ^{BC}Lys3–^{BC}Phe7, ^{BC}His9–^{BC}Thr45, ^{BC}Gly69–^{BC}Pro93, ^{BC}Lys96–^{BC}Leu102, ^{BC}Val112–^{BC}Lys153, ^{BC}Ser154–^{BC}The208 and ^{BC}U5P250, respectively, using all atoms for conserved residues and backbone atoms only for the rest. The gaps cover inserts and bulges which obviously did not fit well and include the dimer arm and the flexible loop. A total of 934 non-H atoms were paired, giving an r.m.s.d. of 1.24 Å. This figure and Figs. 4 and 5 were produced with the program *TURBO-FRODO* (Roussel & Cambillau, 1992).

3. Results and discussion

The crystal structure of uracil phosphoribosyltransferase from the thermophilic bacterium *B. caldolyticus* has been determined and refined to 3.0 Å resolution. The structure contains a dimer in the asymmetric unit, each monomer being composed of residues Gly2–Lys209. Since no UMP had been added to the crystallization mixture, an unexpected result of the structure determination was the presence of a UMP molecule in the active site, reflecting a strong product binding.

3.1. Monomer fold

The monomer possesses both β-strand and α-helix secondary-structural elements. We have assigned 11 β-strands and six α-helices (Fig. 1). The structural core contains a central seven-stranded β-sheet (β1, β9, β8, β7, β4, β5 and β6) surrounded by α-helices (α1–α5). Six of the β-strands form an extended parallel sheet with the flanking strand, β6, running antiparallel. This core is classified as a type I PRTase fold, which is the most abundant fold seen for PRTase structures.

The secondary-structural elements are, with three exceptions, connected by short loops, giving a rather compact fold. The antiparallel strands β5 and β6 are connected *via* a long loop known as the 'PRTase flexible loop'. Another structural feature, unique to the UPRTase structure, are two antiparallel β-strands (β2 and β3) that seem to be important for dimer stabilization. Finally, a C-terminal hood contains two short antiparallel β-strands (β10 and β11) and an α-helix (α6). As will be shown, this hood is also unique to the UPRTase structure and is involved in uracil recognition.

A feature common to PRTases is the use of PRPP as one of the substrates, suggesting the evolution of a common fold (the type I PRTase fold; Fig. 1) which can bind PRPP. In

UPRTase, this fold is represented by strands $\beta4$–$\beta9$ together with helices $\alpha2$–$\alpha5$. The PRPP recognition is shared by two subsites. One subsite (formed by the loop between strand $\beta7$ and helix $\alpha4$) binds the ribose-5′-phosphate (R5P) part of PRPP. The other subsite (formed by the loop between strand $\beta4$ and helix $\alpha3$) binds the pyrophosphate (PP_i) part of PRPP. A *cis*-peptide is often found in the PP_i-binding loop (residue *cis*-Arg79 in this structure). Another characteristic feature of the type I PRTase fold is an extension of three β-strands in the core. In UPRTase, strands $\beta7$, $\beta8$ and $\beta9$ have extensions $\beta7'$, $\beta8'$ and $\beta9'$, respectively. These extra β-strands form a small parallel sheet almost perpendicular to the core sheet. This special structure seems to define the architecture of the R5P-binding site, which extends from the C-terminal end of strand $\beta7$ to the N-terminal end of helix $\alpha4$.

3.2. The dimer

The crystal contains two molecules in the asymmetric unit organized as a tightly entwined dimer (Fig. 3). The dimer is symmetric and has an ellipsoid shape overall with dimensions of 29 × 48 × 88 Å. The accessible surface areas of the dimer and the monomer were calculated with the program *GRASP* (Nicholls *et al.*, 1991) to be 18 290 Å^2 and 11 097 Å^2, respectively, which showed that 1952 Å^2 (17.6%) of the monomer surface is involved in the dimer interaction. The dimer is held together by 12 hydrogen bonds, of which four are salt bridges: Lys22*A*/*B*–Glu56*B*/*A* (see also below) and Glu43*A*/*B*–Arg46*B*/*A*. The latter is in the middle of the dimer.

Strands $\beta2$ and $\beta3$ (Leu50–Ile67; hereafter referred to as the dimer arm) contribute about half of the dimer-interaction area. Without residues Leu50–Ile67, the interaction surface area is 941 Å^2 per monomer. The dimer arm has an uncharged face interacting with the corpus of the other monomer and a solvent-exposed face comprising exclusively charged residues. Interestingly, strand $\beta2$ is rich in negatively charged residues (Glu51, Glu52, Glu54 and Glu56), while strand $\beta3$ is rich in positively charged residues (Lys61, Arg63 and Lys65). The side chains are in favourable positions to form a zipper-like system of putative salt bridges (Lys65–Glu52–Arg63–Glu54–Lys61–Glu56) stabilizing the dimer arm. The last residue in this system, Glu56, is in turn involved in a salt bridge Glu56*A*/*B*–Lys22*B*/*A*, which holds the tip of the dimer arm and the corpus of the opposite monomer together.

3.3. Crystal packing

The modest resolution of the diffraction observed from the *B. caldolyticus* UPRTase crystals is not consistent with what would be expected from a thermophilic enzyme. However, analysis of the crystal packing gives an explanation for this observation. The most prominent packing contacts are long chains of dimers perpendicular to the crystallographic *c* axis.

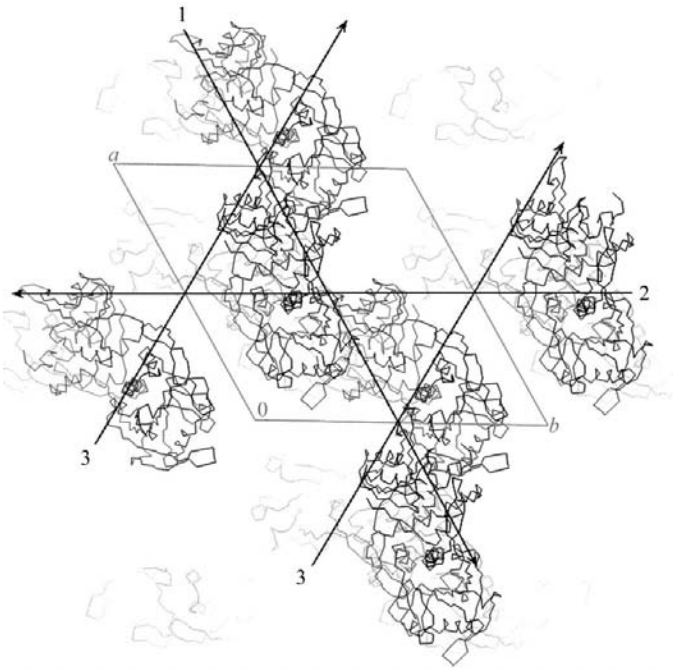

Figure 4
View of the crystal packing. Dimers interact strongly in long chains loosely stacked in layers perpendicular to the *c* axis. One layer is related to the next layer by a 60° rotation perpendicular to the *c*-axis direction followed by a translation of 1/3*c*.

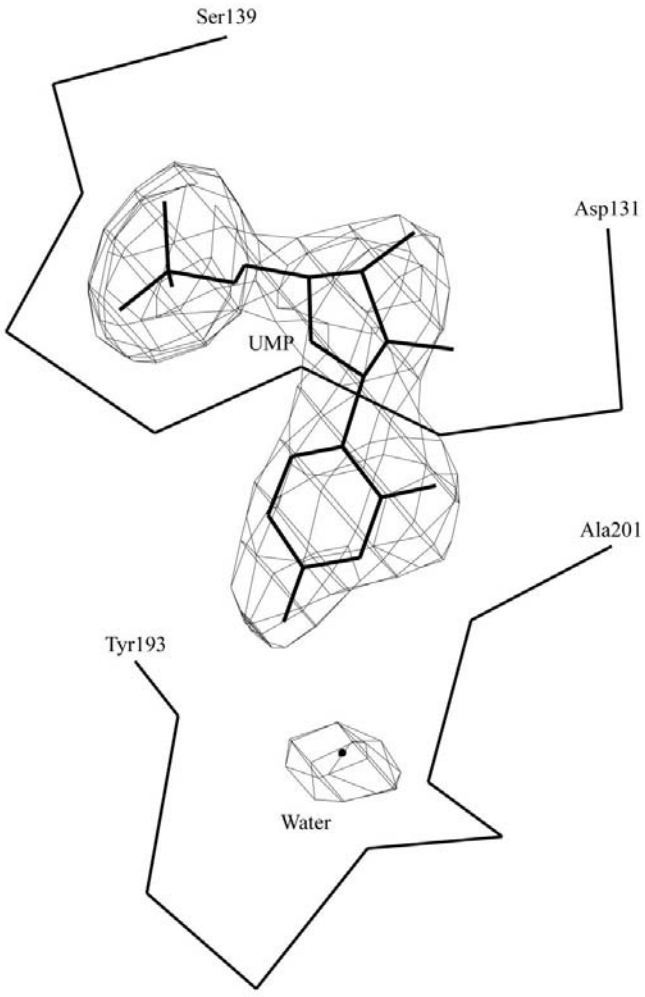

Figure 5
Residual density of UMP and water calculated before inclusion of these in the molecular model and superposed with the final model.

The chains loosely stack in the *c*-axis direction as alternating layers with a period of three, so that the chains from one layer form a 60° angle to the chains of the next layer (Fig. 4). There is no interaction between parallel chains in the same layer. Consistently, fibrous structures are often observed in the crystallization drop along with crystals (Fig. 2).

The interactions between dimers in a chain appear to be strong compared with the intermolecular crystal contacts normally observed in protein crystals. The interaction surface area, calculated as above, amounts to 1840 Å^2 or 10% of the dimer surface. This area is of the same magnitude as that which forms the dimers. The contacts comprise eight hydrogen bonds, of which four are between charged residues: Arg30*A*/*B*···Asp105*B*/*A* and Asp86*A*/*B*···Lys117*B*/*A*. These residues are not sequentially conserved among UPRTases.

3.4. The active site and UMP binding

After having constructed and refined the polypeptide chain, clear residual electron density was found at the active site (Fig. 5) despite there having been no addition of UMP to the crystallization mixture. When UMP was included in the refinement, a restrained temperature-factor optimization showed thermal *B* factors for UMP that were comparable with those of surrounding residues, implying that the UMP site is fully occupied. The unexpected presence of UMP in the crystals points to strong product binding. The conformation of UMP bound to UPRTase is the same as the conformation of UMP found in the small-molecule crystal structure of its sodium salt (Seshadri *et al.*, 1980) and which is classified as an *anti* conformation. This indicates that UMP binds to the enzyme in a relaxed conformation, which explains the strong binding and might be one of the driving forces towards the product.

The direct interaction between the enzyme and UMP involves the two segments Asp131–Ser139 and Tyr193–Ala201 (Figs. 5 and 6). The first segment is part of the type I PRTase core and contains the PRPP-binding consensus-sequence motif. Firstly, residue Asp131, which is situated at the end of β-strand β7, is hydrogen bonded to the ribose moiety of UMP. It binds *via* its carboxylate group, where the $O_{\delta1}$ and $O_{\delta2}$ atoms accept a proton from O2′ and O3′, respectively. There is then a stretch of structural support ($\beta7'$) with a few close van der Waals interactions to the uracil moiety, notably from the side chains of Met133 and Ala135. The last part of the segment defines the 5′-phosphate-binding loop, Ala135–Ser139, connecting β-strand $\beta7'$ and α-helix α4. The 5′-phosphate O atoms of UMP accept hydrogen bonds from the backbone atoms Ala135 N, Thr136 N, Gly137 N and Ser139 N, as well as from the side-chain atoms Thr136 O_γ and Ser139 O_γ.

The second segment involved in UMP binding is Tyr193–Ala201 at the C-terminus between β-strand β9 and α-helix α6. It interacts exclusively with the uracil part of UMP. This interaction involves backbone atoms and is partly mediated by a water molecule. The uracil O2 atom accepts a proton from Ala201 N, while N3 donates a proton to Gly199 O. O4 of uracil accepts protons from a water molecule and from Ile194 N. The water molecule in turn donates a proton to Ile194 O and accepts protons from Leu198 N and Gly199 N. This water molecule is, apart from lateral hydrogen bonds, confined to a partly hydrophobic environment. It is sandwiched between and has close van der Waals interactions with the side chain of Val59, situated at the tip of the dimer arm of the opposite monomer, and the side chain of Met133. The uracil-binding loop contains a *cis*-peptide bond between Val195 and Pro196. A very similar binding of uracil is seen in the structure of dUTPase from *E. coli* complexed with dUDP (Dauter *et al.*, 1998). Here, the binding of uracil also involves only backbone atoms and is mediated in the same way by a water molecule.

Comparison with orotate phosphoribosyltransferase complexed with orotate and PRPP (Scapin *et al.*, 1995) identifies the loop between β-strand β4 and α-helix α3 as a putative binding site for the pyrophosphate part of PRPP. This PP_i loop contains a *cis*-peptide bond between Leu78 and Arg79.

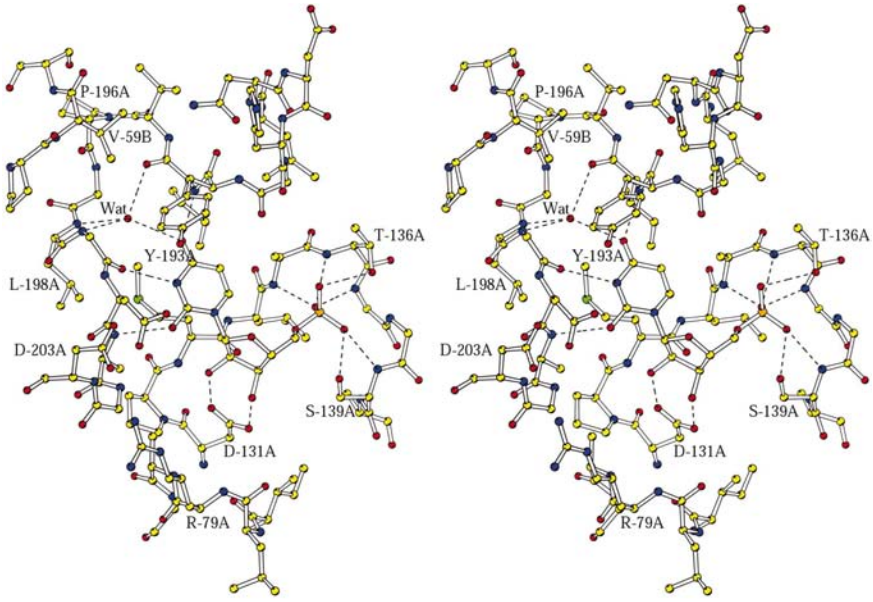

Figure 6
A close-up stereoview of the active site with the product UMP bound. Two parts of the enzyme contributes to the binding of UMP: the segment Asp131–Ser139, which includes the PRPP-binding motif characteristic of type I PRTases, binds the ribose-5-phosphate part of UMP, while the C-terminal segment Tyr193–Ala201 binds the uracil part. The binding of uracil is defined by backbone atoms only and is mediated by a characteristic water molecule.

3.5. Comparison with the *Tox. gondii* UPRTase structure

The sequence of *Tox. gondii* UPRTase (^{TG}UPRTase) has an N-terminal extension compared with that of *B. caldolyticus*

UPRTase (^{BC}UPRTase). In the area where the two sequences overlap and can be aligned, they share 75 out of 209 identical amino-acid residues.

Initially, the ^{TG}UPRTase (with bound UMP; PDB code 1upu) subunit *A* was superposed with the ^{BC}UPRTase subunit *A* (Fig. 3). Gaps, inserts and bulges which obviously did not fit well and which include the dimer arm and flexible loop were omitted from the fit. Backbone atoms together with side-chain atoms for conserved residues were included in the fit: a total of 934 non-H atom pairs with an r.m.s.d. of 1.24 Å.

Not surprisingly, the two structures superimpose best around the active site. UMP can be superposed with an r.m.s.d. of 0.21 Å for all 21 non-H atoms and in general its immediate surroundings fit within experimental accuracy. However, some differences are seen in connection with the hydrophobic interactions with the uracil ring. The substitutions ^{TG}Leu196/^{BC}Ile161 and ^{TG}Ile233/^{BC}Leu198 cause the ^{TG}Met166/^{BC}Met133 side chains to adopt different conformations and ^{TG}Tyr228/^{BC}Tyr193 to have different χ_2 angles. The ^{TG}Tyr228 and ^{BC}Tyr193 aromatic ring plane make angles of 95° and 55° with the uracil-ring plane for the ^{TG}UPRTase and ^{BC}UPRTase complexes, respectively.

Further away from the active site, the superpositional fit becomes poorer. The dimers have a different twist and cannot be superposed as a whole. When the upper subunits are superposed as described above (Fig. 3) the lower subunits do not fit and *vice versa*. In order to transform an upper fit into a lower fit a 16° rotation is required, followed by a 2.3 Å translation along the screw axis shown in Fig. 3. This means that the ^{TG}UPRTase dimer is in a tighter conformation than the ^{BC}UPRTase dimer and needs to be unwound slightly to fit the latter. It remains to be shown whether this difference arises from the species variation or reflects a kind of 'breathing' motion associated with the enzyme function. Here, it should be noted that the crystals of *Tox. gondii* and *B. caldolyticus* UPRTases were obtained under very different conditions using 1.0 *M* ammonium sulfate/phosphate and 6–10% PEG 4000 as precipitants, respectively. The authors have observed no 'breathing' motion in the four ^{TG}UPRTase structures deposited with the Protein Data Bank (PDB codes 1bd3, 1bd4, 1upf and 1upu).

Even though the dimer arm was not included in the superpositional fit, the upper part from the opposite subunit close to the active site matches well. The backbone adopts similar conformations until the conserved ^{TG}Pro91/^{BC}Pro58. An insertion in *Tox. gondii* (^{TG}Asp93) causes a bulge right after ^{TG}Leu92/^{BC}Val59 that compensates for the shorter ^{BC}Val59 side chain, so that the isopropyl groups of the two residues occupy approximately the same position within van der Waals interaction distance of ^{TG}Tyr228/^{BC}Tyr193. Another conserved feature is a salt-bridge anchor from the corpus to just below the tip: ^{TG}Lys55*A*–^{TG}Asp93*D*/^{BC}Lys22*A*–^{BC}Glu56*B*. The lysines are at equivalent positions in the corpus structures, while ^{BC}Glu56 and ^{TG}Asp93 are positioned before and after the conserved ^{BC}Pro58/^{TG}Pro91, respectively. Further down the dimer arm, the interactions with the corpus are much weaker in both structures. This produces the possibility that the dimer arm can slide on the corpus surface, leaving the dimer with a certain flexibility.

3.6. Insight into catalysis from modelling an enzyme–substrate complex

It has been demonstrated that UPRTases follow ordered sequential kinetics with the substrate PRPP binding first and the product UMP leaving last (Lundegaard & Jensen, 1999; Jensen & Neuhard, unpublished results). In order to understand this at a molecular level, we first modelled a PRPP–UPRTase complex. For simplicity, we have assumed that the R5P part of PRPP is close to that observed in the UMP–UPRTase complex, leaving only the PP_i part to be modelled. We therefore used the PRPP molecule (actually cPRPP, with O4 of the ribose ring substituted with CH_2) from the complex with *E. coli* XGPRTase (Vos *et al.*, 1998; PDB code 1a95) as a template (Fig. 7). Firstly, atoms C1, C2 and C5 of cPRPP were manually superposed with the equivalent atoms C1, C2 and O4 of UMP, letting cPRPP move as a rigid body. This was performed in order to obtain the correct geometry of the junction at C1. From there, the four torsion angles between C1 and the β P atom were allowed to move in order to dock the PP_i moiety to the enzyme. With these constraints, we could find only one sterically plausible conformation. This conformation had hydrogen-bond contacts between the β-phosphate group and the enzyme around *cis*-Arg79.

When PRPP is placed in the active site, it becomes clear that owing to steric reasons uracil can only enter parallel to the Tyr193

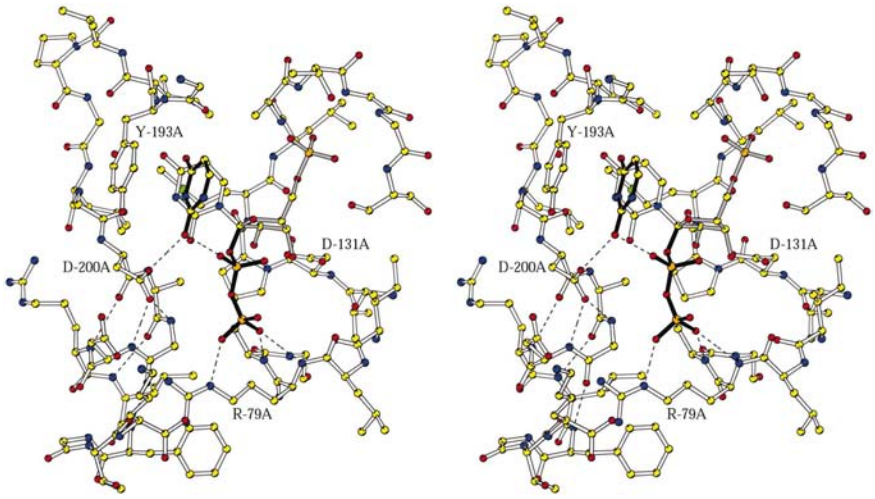

Figure 7
Stereoview of a hypothetical enzyme–substrate complex. The substrate molecules uracil and PRPP are shown (black bonds) superposed onto the actual enzyme–product complex (white bonds) for which the structure was determined. The ribose-5′-phosphate part common to substrate and product has grey bonds.

aromatic ring plane. Placing uracil to have a π-stacking with Tyr193 gives, in addition to the role of this residue in guiding uracil to its final position, a clue to how the substrate is activated. It is likely that uracil in this position will adopt its tautomeric enol form, with a double bond between N1 and C2 and the N1 proton repositioned to O2 (Fig. 8). The enol form is favoured by O2 being able to donate its proton in hydrogen bonding to the carboxylate group of Asp200 and the α-phosphate group of PRPP. This means that when the enol uracil enters further into the active site, the proton is in an ideal position to move towards OP1 of the α-phosphate group and thereby simultaneously activate uracil as a nucleophile and PP_i as a leaving group. As mentioned above, Val59 from the opposite subunit is in van der Waals contact with both Tyr193 and the water molecule in the uracil-binding site. Enzyme movement may deliver, *via* dimer-arm stimulation at the uracil-binding site, the activation energy required to move from an activated complex to the product. Flexibility in the site seems to be ensured by the conserved glycines 197 and 199, by the water molecule and in general by this part of the structure being a C-terminal protrusion. In entering the active site, it is likely that uracil orients itself by having its dipole moment aligned with that of α-helix 6.

In the present model, the conformation of the PP_i part of PRPP is consistent with observations made from HGPRTase in complex with the transition-state analogue ImmGP and Mg^{2+}–PP_i (Shi, Li, Tyler, Furneaux, Cahill *et al.*, 1999; Shi, Li, Tyler, Furneaux, Grubmeyer *et al.*, 1999). A shortcoming of our model is that we cannot directly infer the location of the magnesium ions. By comparison with the HGPRTase complex, candidates for two Mg^{2+}-binding sites would be between the PP_i moiety and Asp200 and Asp131, respectively (Fig. 7). In this context, one should recall that the PRPP-binding motif of UPRTases differs from that of all other PRTases by having the sequence Asp131-Pro132 instead of two acidic residues (Glu-Asp in OPRTase/HGPRTase and Asp-Asp in GPATase/APRTase). This may have an impact on the binding of Mg^{2+}. The reaction of type I PRTases has been proposed to proceed *via* a ribooxocarbenium-ion transition state (Grubmeyer & Wang, 1998; Shi, Li, Tyler, Furneaux, Cahill *et al.*, 1999; Shi, Li, Tyler, Furneaux, Grubmeyer *et al.*, 1999). The acidic residues of the PRPP-binding motif just mentioned (Asp-Pro, Glu-Asp or Asp-Asp close to the ribose ring) are obvious candidates for an overall stabilization of such a transition state. Because of the less acidic Asp131-Pro132 in UPRTases, one could speculate that other acidic groups such as Asp200 and Asp203 may also be involved in the overall stabilization of the transition state.

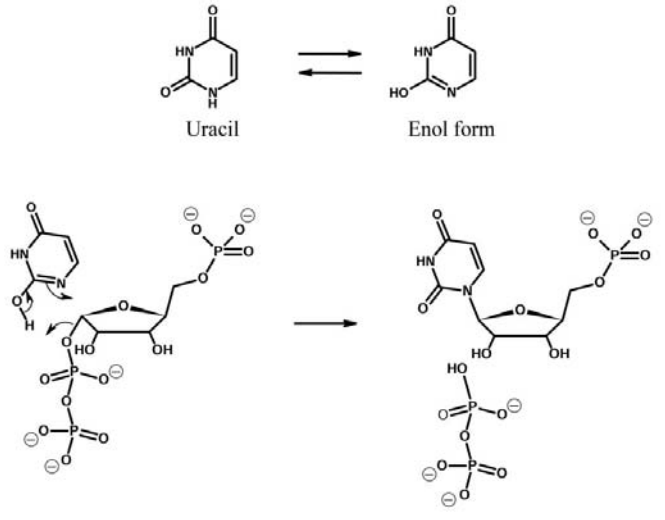

Figure 8
Proposed catalytic mechanism for uracil phosphoribosyltransferase: when PRPP is bound to the enzyme and uracil enters the active site (Fig. 7), uracil is stabilized as the enol tautomer. When in the enol form, uracil enters deeper into the active site and the shown electron translocation takes place, completing the reaction.

We are grateful to the EMBL Hamburg outstation c/o DESY for beam time and assistance during experiments. This research has been supported by the Danish National Research Foundation.

References

Alloush, H. M. & Kerridge, D. (1994). *Mycopathologia*, **125**, 129–141.
Asai, T., Lee, C. S., Chandler, A. & O'Sullivan, W. J. (1990). *Comput. Biochem. Physiol. B*, **95**, 159–163.
Balendiran, G. K., Molina, J. A., Xu, Y., Torres-Martinez, J., Stevens, R., Focia, P. J., Eakin, A. E., Sacchettini, J. C. & Craig, S. P. (1999). *Protein Sci.* **8**, 1023–1031.
Bressan, R. A., Murray, M. G., Gale, J. M. & Ross, C. W. (1978). *Plant Physiol.* **61**, 442–446.
Brünger, A. T. (1992). *X-PLOR Version* 3.1 *Manual.* Howard Hughes Medical Institute and Department of Biophysics and Biochemistry, Yale University, New Haven, CT, USA.
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.
Carter, D., Donald, R. G. K., Ross, D. & Ullman, B. (1997). *Mol. Biochem. Parasitol.* **87**, 137–144.
Chen, S., Tomchick, D. R., Wolle, D., Hu, P., Smith, J. L., Switzer, R. L. & Zalkin, H. (1997). *Biochemistry*, **36**, 10718–10726.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dai, Y.-P., Lee, C. S. & O'Sullivan, W. J. (1995). *Intern. J. Parasitol.* **25**, 207–214.
Dauter, Z., Wilson, K. S., Larsson, G., Nyman, P. O. & Cedergren-Zeppezauer, E. S. (1998). *Acta Cryst.* D**54**, 735–749.
Eads, J. C., Ozturk, D., Wexler, T. B., Grubmeyer, C. & Sacchettini, J. C. (1997). *Structure*, **5**, 47–58.
Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C. & Sacchettini, J. C. (1994). *Cell*, **78**, 325–331.
Eriksen, T. A., Kadziola, A., Bentsen A.-K., Harlow, K. W. & Larsen, S. (2000). *Nature Struct. Biol.* **7**, 303–308.
Focia, P. J., Craig, S. P. III & Eakin, A. E. (1998). *Biochemistry*, **37**, 17120–17127.
Focia, P. J., Craig, S. P. III, Nieves-Alicea, R., Fletterick, R. J. & Eakin, A. E. (1998). *Biochemistry*, **37**, 15066–15075.
Gewirth, D. (1994). *The HKL Manual: an Oscillation Data Processing Suite for Macromolecular Crystallography.* Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA.
Goitein, R. K., Chelsky, D. & Parsons, S. M. (1978). *J. Biol. Chem.* **253**, 2963–2971.
Grubmeyer, C. & Wang, G. (1998). *Paths Pyrimidines*, **6**, 1–11.

Henriksen, A., Aghajari, N., Jensen, K. F. & Gajhede, M. (1996). *Biochemistry*, **35**, 3803–3809.
Héroux, A., White, E. L., Ross, L. J., Davis, R. L. & Borhani, D. W. (1999). *Biochemistry*, **38**, 14495–14506.
Héroux, A., White, E. L., Ross, L. J., Kuzin, A. P. & Borhani, D. W. (2000). *Structure*, **8**, 1309–1318.
Hershey, H. V. & Taylor, M. (1986). *Gene*, **42**, 287–293.
Hove-Jensen, B., Harlow, K. W., King, C. J. & Switzer, R. L. (1986). *J. Biol. Chem.* **261**, 6765–6771.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jensen, H. C., Mikkelsen, N. & Neuhard, J. (1997). *Protein Expr. Purif.* **10**, 356–364.
Jensen, K. F. (1983). *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganisms*, edited by A. Munch-Petersen, pp. 1–25. London: Academic Press.
Jensen, K. F. & Mygind, B. (1996). *Eur. J. Biochem.* **204**, 637–645.
Jyssum, S. & Jyssum, K. (1979). *J. Bacteriol.* **138**, 320–323.
Kleywegt, G. J. & Jones, T. A. (1996). *Acta Cryst.* D**52**, 826–828.
Krahn, J. M., Kim, J. H., Burns, M. R., Parry, R. J., Zalkin, H. & Smith, J. (1997). *Biochemistry*, **36**, 11061–11068.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacArthur, M. V., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lawson, D. M. (1994). *Acta Cryst.* D**50**, 332–334.
Linde, L. & Jensen, K. F. (1996). *Biochim. Biophys. Acta*, **1296**, 16–22.
Lundegaard, C. & Jensen, K. F. (1999). *Biochemistry*, **38**, 3327–3334.
McIvor, R. S., Wohlhueter, R. M. & Plageman, P. G. W. (1983). *J. Bacteriol.* **156**, 192–197.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Michell, A. & Finch, L. L. (1979). *J. Bacteriol.* **137**, 1073–1080.
Muchmore, C. R. A., Krahn, J. M., Kim, J. H., Zalkin, H. & Smith, J. L. (1998). *Protein Sci.* **7**, 39–51.
Natalini, P., Ruggieri, S., Santarelli, I., Vita, A. & Magni, G. (1979). *J. Biol. Chem.* **254**, 1558–1563.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Neuhard, J. (1983). *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganisms*, edited by A. Munch-Petersen, pp. 95–148. London: Academic Press.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend: Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Phillips, C. L., Ullman, B., Brennan, R. G. & Hill, C. P. (1999). *EMBO J.* **18**, 3533–3545.
Plunkett, W. & Moner, J. G. (1978). *Arch. Biochem. Biophys.* **187**, 264–271.
Read, R. (1986). *Acta Cryst.* A**42**, 140–149.
Roussel, A. & Cambillau, C. (1992). *TURBO-FRODO*. Biographics and AFMB (Architecture et Fonction des Macromolécules Biologiques), Marseille, France.
Scapin, G., Grubmeyer, C. & Sacchettini, J. C. (1994). *Biochemistry*, **33**, 1287–1294.
Scapin, G., Ozturk, D. H., Grubmeyer, C. & Sacchettini, J. C. (1995). *Biochemistry*, **34**, 10744–10754.
Schumacher, M. A., Carter, D., Roos, D. S., Ullman, B. & Brennan, R. G. (1999). *Nature Struct. Biol.* **3**, 881–887.
Schumacher, M. A., Carter, D., Scott, D. M., Ross, D. S., Ullman, B. & Brennan, R. G. (1998). *EMBO J.* **17**, 3219–3232.
Seshadri, T. P., Viswamitra, M. A. & Kartha, G. (1980). *Acta Cryst.* B**36**, 925–927.
Shi, W., Li, C. M., Tyler, P. C., Furneaux, R. H., Cahill, S. M., Girvin, M. E., Grubmeyer, C., Schramm, V. L. & Almo, S. C. (1999). *Biochemistry*, **38**, 9872–9880.
Shi, W., Li, C. M., Tyler, P. C., Furneaux, R. H., Grubmeyer, C., Scramm, V. L. & Almo, S. C. (1999). *Nature Struct. Biol.* **6**, 588–593.
Smith, J. L., Zaluzec, E. J., Wery, J.-P., Niu, L., Switzer, R. L., Zalkin, H. & Satow, Y. (1994). *Science*, **264**, 1427–1433.
Somoza, J. R., Chin, M. S., Focia, P. J., Wang, C. C. & Fletterick, R. J. (1996). *Biochemistry*, **35**, 7032–7040.
Tao, W., Grubmeyer, C. & Blanchard, J. S. (1996). *Biochemistry*, **35**, 14–21.
Traut, T. W. & Jones, M. E. (1996). *Prog. Nucleic Acid Res.* **53**, 1–78.
Vos, S., De Jersey, J. & Martin, J. L. (1997). *Biochemistry*, **36**, 4125–4134.
Vos, S., Parry, R. J., Burns, M. R., De Jersey, J. & Martin, J. L. (1998). *J. Mol. Biol.* **282**, 875–889.