Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

High-*syn* conformation of uridine and asymmetry of the hexameric molecule revealed in the high-resolution structures of *Shewanella oneidensis* MR-1 uridine phosphorylase in the free form and in complex with uridine

Tatyana N. Safonova,[a]* Sergey N. Mikhailov,[b] Vladimir P. Veiko,[a] Nadezhda N. Mordkovich,[a] Valentin A. Manuvera,[c] Cyril S. Alekseev,[b] Mikhail V. Kovalchuk,[d] Vladimir O. Popov[a,d] and Konstantin M. Polyakov[a,b]

[a]Bach Institute of Biochemistry, Russian Academy of Sciences, 33 Leninskii Ave., Moscow 119071, Russian Federation, [b]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 32 Vavilova St., Moscow 119991, Russian Federation, [c]Research Institute for Physico-Chemical Medicine, Federal Medical–Biological Agency, 1a Malaya Pirogovskaya St., Moscow 119435, Russian Federation, and [d]National Research Centre 'Kurchatov Institute', 1 Pl. Akademika Kurchatova, Moscow 123182, Russian Federation

Correspondence e-mail: tn_safonova@inbi.ras.ru

Received 5 September 2014
Accepted 31 October 2014

PDB references: uridine phosphorylase, 4r2x; uridine phosphorylase in complex with uridine, 4r2w

Uridine phosphorylase (UP; EC 2.4.2.3), a key enzyme in the pyrimidine-salvage pathway, catalyzes the reversible phosphorolysis of uridine to uracil and ribose 1-phosphate. Expression of UP from *Shewanella oneidensis* MR-1 (SoUP) was performed in *Escherichia coli*. The high-resolution X-ray structure of SoUP was solved in the free form and in complex with uridine. A crystal of SoUP in the free form was grown under microgravity and diffracted to ultrahigh resolution. Both forms of SoUP contained sulfate instead of phosphate in the active site owing to the presence of ammonium sulfate in the crystallization solution. The latter can be considered as a good mimic of phosphate. In the complex, uridine adopts a high-*syn* conformation with a nearly planar ribose ring and is present only in one subunit of the hexamer. A comparison of the structures of SoUP in the free form and in complex with the natural substrate uridine showed that the subunits of the hexamer are not identical, with the active sites having either an open or a closed conformation. In the monomers with the closed conformation, the active sites in which uridine is absent contain a glycerol molecule mimicking the ribose moiety of uridine.

1. Introduction

Uridine phosphorylase (UP; EC 2.4.2.3) belongs to the nucleoside phosphorylases, which are involved in nucleic acid metabolism and play a key role in the salvage pathway of nucleoside biosynthesis. These enzymes are found in virtually all organisms and vary only slightly in primary structure. They are used in industry for the synthesis of drugs and commercially important nucleosides (Utagawa, 1999; Li *et al.*, 2010; Mikhailopulo & Miroshnikov, 2010, 2011).

In the presence of orthophosphate, UP catalyzes the reversible phosphorolysis of uridine to the free heterocyclic base (uracil) and α-D-ribose-1-phosphate, and normally the equilibrium shifts to the nucleoside (Pugmire & Ealick, 2002; Lewkowicz & Iribarren, 2006; Lashkov *et al.*, 2011). Uridine phosphorylase is ubiquitously present in various organisms including prokaryotes, yeasts and higher organisms. The level of this enzyme is often elevated in tumours (Krenitsky *et al.*, 1965; Kanzaki *et al.*, 2002), owing to which it can be used as a marker for cancer diagnosis (Watanabe & Uchida, 1995). The amino-acid sequence of UP is highly conserved amongst bacterial and vertebrate UPs (Pugmire & Ealick, 2002).

Uridine phosphorylase from *Shewanella oneidensis* MR-1 (SoUP) consists of 252 amino-acid residues and has a molecular weight of 27 kDa. *S. oneidensis* MR-1 is a free-living Gram-negative γ-proteobacterium (Alteromonadales order)

and is a facultative anaerobe capable of surviving and proliferating in both aerobic and anaerobic conditions. The activity of SoUP is 2.5 times higher (Mordkovich *et al.*, 2013) than that of *Escherichia coli* UP and is similar to that of UP from *Salmonella typhimurium*.

Of the protein structures that have been deposited in the Protein Data Bank, SoUP has the highest identity to UP from *Vibrio cholerae* O1 Biovar El Tor (76%; PDB entry 3o6v, 1.70 Å resolution; Center for Structural Genomics of Infectious Diseases, unpublished work), UP from *S. typhimurium* (76%; PDB entry 2hsw, 1.50 Å resolution; V. I. Timofeev, M. V. Dontsova, A. G. Gabdoulkhakov, A. M. Mikhailov, A. A. Lashkov, V. Voelter & G. S. Kachalova, unpublished work) and UP from *E. coli* (75%; PDB entry 1rxy, 1.70 Å resolution; Caradoc-Davies *et al.*, 2004), and much lower identity to UP from *Streptococcus pyogenes* (37%; PDB entry 3qpb, 1.82 Å resolution; Tran *et al.*, 2011), bovine UP (31%; PDB entry 3ku4, 2.10 Å resolution; Paul *et al.*, 2010) and human UP (26%; PDB entry 3euf, 1.90 Å resolution; Roosild *et al.*, 2009). The three-dimensional structure of SoUP has not yet been studied. The only structure of UP in complex with uridine available in the PDB was solved at 2.91 Å resolution (PDB entry 2hwu; V. I. Timofeev, A. G. Gabdulkhakov, M. V. Dontsova, W. Voelter, G. S. Kachalova & A. M. Mikhailov, unpublished work).

Here, we report the X-ray crystal structures of uridine phosphorylase from *S. oneidensis* MR-1 in the free form and in complex with the natural substrate uridine solved and refined at 0.95 and 1.6 Å resolution, respectively. These data made it possible to determine the geometric parameters of the uridine molecule with high accuracy and to compare the structures of the active sites in the free form and in the complex. The uridine was shown to be bound in the active site of the enzyme in the high-energy high-*syn* conformation with a nearly planar ribose ring.

2. Experimental

2.1. Expression and purification of uridine phosphorylase from *S. oneidensis* MR-1

The cloning of the *udp* gene from *S. oneidensis* MR-1, the construction of the recombinant producer strain and the expression, isolation and purification of recombinant SoUP have been described previously (Mordkovich *et al.*, 2012). The enzyme solution was lyophilized and stored at −20°C. The primary structure of recombinant SoUP was confirmed by mass-spectrometric analysis carried out on an UltrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Ettlingen, Germany) in the Joint Use Center of the A. N. Bach Institute of Biochemistry of the Russian Academy of Sciences.

The hexameric structure of recombinant SoUP in solution was confirmed by gel filtration (Mordkovich *et al.*, 2013). SoUP was dissolved (to a final concentration of 0.3 mg ml^{-1}) in phosphate-buffered saline (PBS) and analyzed on a Superdex 200 Tricorn 10/300 column (column volume 23.5 ml; GE Healthcare) using recombinant UP from *E. coli*, which was produced earlier (Veiko *et al.*, 1994), as the reference protein. The specific activity of the enzyme was determined as described by Leer *et al.* (1977).

2.2. Crystallization

The search for crystallization conditions for SoUP in the free form and in complex with uridine has been described in detail previously (Safonova *et al.*, 2012). Crystals were grown by the hanging-drop vapour-diffusion method at room temperature using kits from Hampton Research. The drops were composed of equal volumes (1 µl) of reservoir solution and protein solution (the lyophilized protein was dissolved in Milli-Q water at a concentration of 20 mg ml^{-1}). The first crystals appeared within one week. It should be noted that although we tested about 250 initial crystallization conditions, we only obtained crystals in the presence of high concentrations of sulfate.

Crystals of SoUP in complex with the substrate (uridine) were grown by cocrystallization using the hanging-drop vapour-diffusion method at room temperature. The screening of crystallization conditions for the complex of SoUP with uridine was carried out taking into account the best hits for SoUP in the free form. Uridine was used in a 100-fold molar excess with respect to SoUP. The largest crystals were obtained using 0.75 *M* ammonium sulfate, 0.075 *M* bis-tris pH 5.5, 0.75%(*w*/*v*) PEG 3350, 25%(*v*/*v*) glycerol as the reservoir solution. The crystal dimensions were 0.30 × 0.25 × 0.30 mm.

We also performed a crystal-growth experiment under microgravity conditions in an attempt to improve the quality of the crystals of SoUP in the free form. The conditions for crystallization under microgravity were optimized and converted to the free-interface diffusion method by performing experiments in glass capillaries. The preliminary conditions thus chosen were used for crystallization under microgravity on board the Russian Segment of the International Space Station (ISS). We succeeded in growing a high-quality crystal using 0.75 *M* ammonium sulfate, 0.1 *M* bis-tris pH 5.5, 0.75%(*w*/*v*) PEG 3350, 25%(*v*/*v*) glycerol. The crystal dimensions were 0.60 × 0.50 × 0.50 mm. As will be shown below, the crystal grown under microgravity gave ultrahigh-resolution diffraction.

2.3. X-ray data collection and processing

An X-ray data set for SoUP in the free form was collected at 100 K from a high-quality crystal grown under microgravity to a resolution of 0.95 Å using an ADSC Quantum 315 CCD detector at a wavelength of 0.800 Å on the BL41XU beamline at the SPring-8 synchrotron-radiation facility, Japan. Since the crystal was grown in the presence of 25%(*v*/*v*) glycerol, no cryoprotection was required.

An X-ray diffraction data set for SoUP complexed with uridine was collected to a resolution of 1.6 Å using a MAR165 detector at a wavelength of 0.812 Å on EMBL beamline BW7A at the DESY synchrotron, Hamburg. Since the crystal was grown in the presence of 25%(*v*/*v*) glycerol, it was simply

Table 1
Data-collection and refinement statistics for SoUP in the free state and in complex with uridine.

	Free form	Complex with uridine
Data collection		
Space group	$P2_1$	$P2_1$
No. of hexamers per unit cell, Z	2	2
Unit-cell parameters (Å, °)	a = 91.54, b = 95.93, c = 91.61, β = 120.0	a = 91.92, b = 96.48, c = 91.93, β = 120.01
Resolution range (Å)	30.0–0.95 (1.00–0.95)	19.5–1.60 (1.70–1.60)
$R_{\rm meas}$† (%)	8.7 (70.8)	6.5 (25.6)
$R_{\rm p.i.m.}$‡ (%)	4.2 (43.2)	3.9 (14.6)
$I/\sigma(I)$	12.2 (2.6)	10.8 (3.5)
Mn[$I/\sigma(I)$]	11.9 (2.2)	10.1 (3.3)
Mn(I) half-set correlation $CC_{1/2}$ (%)	99.8 (66.5)	99.8 (77.1)
Data completeness (%)	96.1 (93.4)	99.6 (84.6)
Measured reflections	4165930 (472372)	433165 (58145)
Unique reflections	834827 (99302)	191905 (25676)
Multiplicity	5.3 (4.8)	2.2 (1.9)
Solvent content (%)	38.4	43.2
Matthews coefficient (Å^3 Da^{-1})	2.01	2.18
Refinement		
No. of twin domains	3	3
Twin fractions	0.80, 0.10, 0.10	0.52, 0.32, 0.16
$R_{\rm cryst}$ (%)	14.9	17.7
$R_{\rm free}$ (%)	16.4	19.2
Cruickshank's DPI for coordinate error (Å)	0.004	0.019
Maximal estimated error (Å)	0.019	0.061
R.m.s.d. from ideal values		
Bond lengths (Å)	0.014	0.018
Bond angles (°)	1.69	1.93
No. of protein atoms	21683	10900
No. of water molecules	1092	554
No. of uridine molecules	0	1
No. of sulfate molecules	8	8
No. of glycerol molecules	19	6
No. of chloride ions	4	0
Ramachandran plot statistics (%)		
Most favoured	90.5	89.7
Additionally allowed	9.1	9.9
Generously allowed	0.4	0.5
Disallowed	0	0
Average B factors (Å^2)		
Protein	8.6	17.3
Water	17.4	21.1

† $R_{\rm meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{\rm p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $N(hkl)$ is the total number of times that a given reflection was measured.

flash-cooled in liquid nitrogen immediately prior to X-ray data collection.

All diffraction data sets were processed using the *XDS* software package (Kabsch, 2010). The data-collection statistics are summarized in Table 1.

2.4. Structure solution and refinement

Firstly, we solved the structure of SoUP in complex with uridine by the molecular-replacement method using *MOLREP* (Vagin & Teplyakov, 2010) from the *CCP*4 suite (Winn *et al.*, 2011) with the structure of *E. coli* UP (PDB entry 1rxy; Caradoc-Davies *et al.*, 2004) as the starting model. Although the unit-cell parameters are consistent with the hexagonal system (Table 1), both hexagonal and monoclinic space groups ($P6_3$ and $P2_1$) can be chosen based on systematic absences. Despite the fact that $R_{\rm meas}$ for this data set at 2.0 Å resolution was 18.2 and 10.7% in the hexagonal and monoclinic space groups, respectively, we made an attempt to solve the structure in space group $P6_3$. Structure solution with *MOLREP* in space group $P6_3$ gave two subunits per asymmetric unit. However, structure refinement with *REFMAC*5 (Murshudov *et al.*, 2011; Winn *et al.*, 2011) in space group $P6_3$ converged to R = 33.4% ($R_{\rm free}$ = 37.1%). Hence, we chose space group $P2_1$, and all subsequent calculations were carried out in this space group. Structure solution in space group $P2_1$ gave six subunits per asymmetric unit. Refinement in the monoclinic space group without twinning converged to R = 23.8% ($R_{\rm free}$ = 27.8%). Structure solution for SoUP in the free form by the molecular-replacement method was not performed since the crystals of the free form are isomorphous with the crystals of the complex.

The statistical properties of the diffraction intensities were estimated using the *CTRUNCATE* program from the *CCP*4 suite, which calculates a number of statistics from the intensity data, such as moments and cumulative intensity distributions, to check for possible twinning based on the H-test (Yeates, 1988) and the L-test (Padilla & Yeates, 2003). The calculations showed possible twinning for both X-ray data sets. The L-tests for the X-ray diffraction data obtained for SoUP in the free form and in the complex are displayed in Fig. 1. The L-test statistics are 0.417 and 0.401, respectively (the expected value for detwinned data is 0.5, while that for a perfect twin is 0.375), thus attesting to the partial twinning. Based on the H-test, three domains with the same symmetry operations (for the first domain, h, k, l and $-h, k, -l$ are equivalent; for the second domain, $-h-l, k, h$ and $h+l, k, -h$ are equivalent; for the third domain, $l, k, -h-l$ and $-l, k, h+l$ are equivalent) and different degrees of twinning can be distinguished. The domain fractions are 0.8:0.1:0.1 and 0.52:0.32:0.16 for the free enzyme and for the complex, respectively. As can be seen from these data, the crystal of the free form is more free from twinning. It should be noted that the plot of the experimental second moments ($\langle I^2\rangle/\langle I\rangle^2$) *versus* the resolution shows no twinning for the SoUP crystals under consideration. This is apparently associated with the fact that the twin axis is parallel to the noncrystallographic screw axis (the threefold axis, which is parallel to the crystallographic twofold axis in space group $P2_1$) and, as was demonstrated by Lunin, this situation can corrupt the intensity statistics (Lunin, 2012).

The final refinement was performed with *REFMAC*5 taking the twinning into account. The contribution of H atoms of the protein molecule and the ligands to the X-ray scattering was taken into account in the refinement of the free form of SoUP. For the structure of SoUP in the free form, the refinement at atomic resolution was performed with anisotropic displacement parameters for non-H atoms. The twin refinement converged to R = 14.9% ($R_{\rm free}$ = 16.4%) and R = 17.7% ($R_{\rm free}$ = 19.2%) for the free enzyme and the complex, respectively. Therefore, refinement taking twinning into account made it possible to substantially improve the atomic models. In the course of the refinement, we performed manual corrections of the model using *Coot* (Emsley & Cowtan, 2004). In the final refinement cycles, the occupancies of some side chains adopting double conformations, as well as of some sulfate and glycerol molecules, were manually specified. The *B* factors of these moieties were taken to be approximately equal to those of the surrounding atoms, and the occupancies were then varied in order to eliminate residual electron-density peaks in the vicinity of these moieties. The structure-refinement statistics are given in Table 1. The electron density in the active-site region for the free form and the complex is shown in Fig. 2. The coordinates and structure factors of the free form and the complex have been deposited in the Protein Data Bank (PDB entries 4r2x and 4r2w, respectively). The figures were prepared using the *CCP*4*mg* molecular-graphics software (McNicholas *et al.*, 2011) and the *loggraph* utility of the *CCP*4 program suite (Winn *et al.*, 2011).

3. Results and discussion

3.1. Overall structure

The quaternary structure of SoUP, like the structures of all prokaryotic UPs determined to date, is a hexamer comprised of six subunits both in solution and in the crystal, as shown by protein electrophoresis and X-ray crystallography. The toroidal hexamer is approximately 100 Å in diameter and 50 Å in height, with a central channel of 18 Å in diameter which narrows to 5 Å at both ends. The hexamer consists of three homodimers, which are the major structural and functional units since amino-acid residues of both subunits of the homodimer are involved in the active site of the enzyme (Gly23, Arg27, Arg88, Thr91, Gln163, Arg165 and Met194 of one subunit and His5 and Arg45 of the other subunit). The active sites in the two subunits of the dimer are separated by a distance of about 22 Å. The distance between the active sites of two adjacent subunits from adjacent dimers is ∼42 Å.

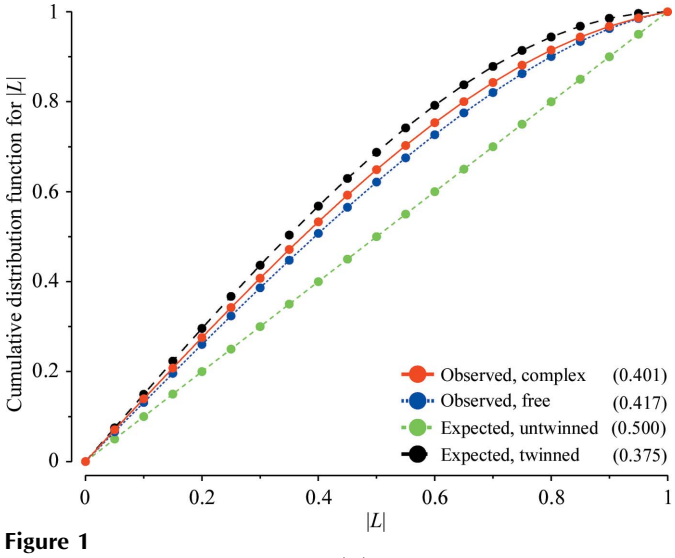

Figure 1
Cumulative distribution function for $|L|$, where $L = (I_1 - I_2)/(I_1 + I_2)$. The expected values for detwinned data are in green, the expected values for the perfect twin are in black and the observed data for the free form and the complex are in blue and red, respectively.

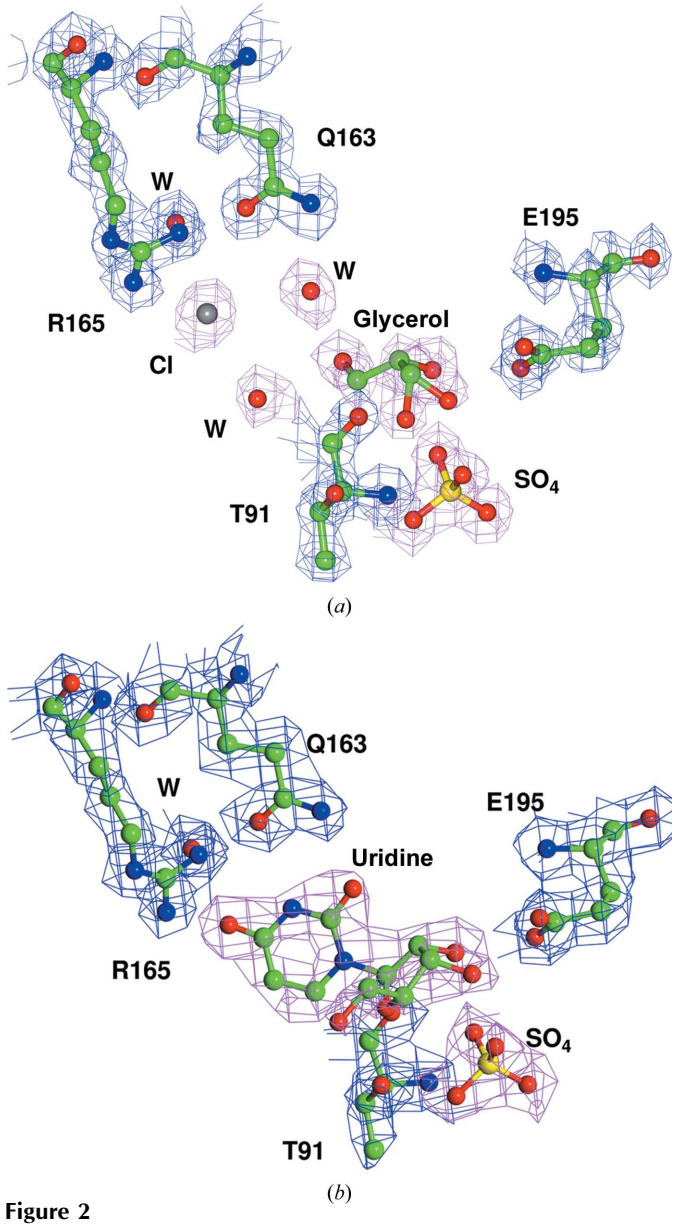

Figure 2
Active site of SoUP for (*a*) the free form and (*b*) the complex with uridine. The $2F_o - F_c$ electron density is contoured at the 1σ level. The electron density for the ligands and amino-acid residues is shown in magenta and blue, respectively.

In the crystal structures of both the complex and the free form (space group $P2_1$) there is one hexamer per asymmetric unit. Hence, there are six crystallographically independent subunits. The overall structure of the hexamer is shown in Fig. 3. The hexamer is stabilized by hydrophobic interactions and hydrogen bonds between the amino-acid residues of adjacent homodimers. In all three functional dimers of the hexamer, the contacts between the subunits as analyzed with *PISA* (Krissinel & Henrick, 2007) are very similar. The same is true for the contacts between the subunits of adjacent dimers. The tertiary structure of the subunits is very similar to those observed for other bacterial uridine phosphorylases.

In the crystal structure, the hexamers form a two-layer packing. Within each layer, subunits *A*, *C* and *E* belonging to different hexamers (see Fig. 3 for the numbering scheme of the subunits) form contacts with each other *via* the regions around residues 36–38 in the vicinity of a noncrystallographic threefold axis. Subunits *B*, *D* and *F* form contacts in the vicinity of another noncrystallographic threefold axis, but these contacts differ from those described above. Thus, the region around residues 36–38 and the C-terminus of subunit *B* are involved in contacts with the C-terminus of subunit *F* and with the region around residues 36–38 of subunit *D*, respectively. The layers in the packing are shifted with respect to one another. The contacts between the subunits from different layers will be considered below.

The subunits in SoUP are not identical. A comparison of the subunits of the enzyme in the free form showed that the subunits can be divided into two groups with closed and open conformations, which is manifested in the movement of the main chain around Thr91 (residues 88–93). This main-chain region is disordered over two positions corresponding to the closed and open conformations in subunits *B* and *E* of the free form (Fig. 4). These are adjacent subunits, but they belong to different functional dimers (see Fig. 3), whereas this region in the other four subunits is ordered and corresponds to the closed conformation. The movement of this region is accompanied by the movement of the main-chain regions 100–105 and 212–219. In the complex with uridine, the open conformation is also observed for subunits *B* and *E* (the region around Thr91 in subunit *B* almost coincides with that in subunits *A*, *C*, *D* and *F*, but the main-chain regions 100–105 and 212–219 correspond to the open conformation). It should be noted that in the complex the main chain around Thr91 is ordered in all subunits. Therefore, the binding of uridine probably facilitates the ordering of the main chain in the vicinity of the active site.

The four subunits with the closed conformation can be further divided into two groups: those in which the loop comprising residues 223–236 is disordered and is not observed in the electron-density maps (subunits *A* and *F*) and those in which this loop is ordered (subunits *C* and *D*). This loop closes the active-site entrance. This situation is observed in crystals of both the free form and the complex. The subunits with the ordered loop belong to the same functional dimer. In the above-mentioned subunits *B* and *E* with an open conformation, the loop 223–236 is also not observed in the electron-density maps. Therefore, there are three types of subunits: *B* and *E* (open conformation, disordered loop 223–232), *C* and *D* (closed conformation, ordered loop), and *A* and *F* (closed conformation, disordered loop 227–236). In each such pair, the subunits form similar contacts with hexamers from adjacent layers, whereas the interhexamer contacts between the layers

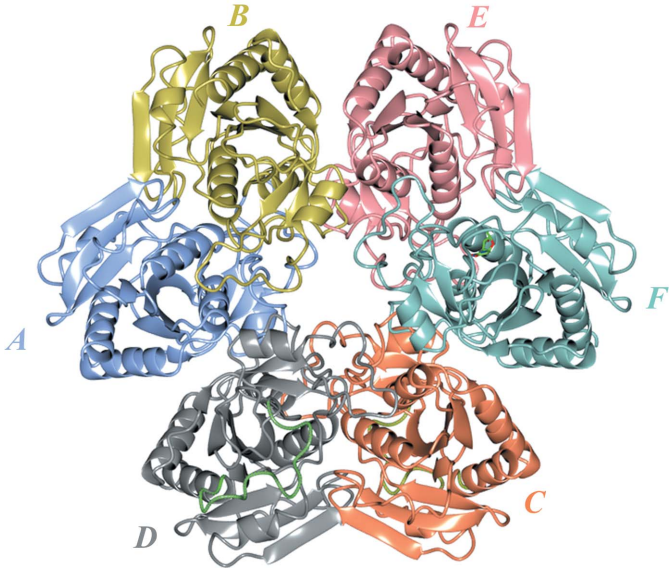

Figure 3
Overall structure of the hexamer of uridine phosphorylase. The loop which is ordered in subunits *C* and *D* and is partially disordered in the other subunits (see the text) is shown in green; the uridine molecule in the active site of subunit *F* is also shown in green.

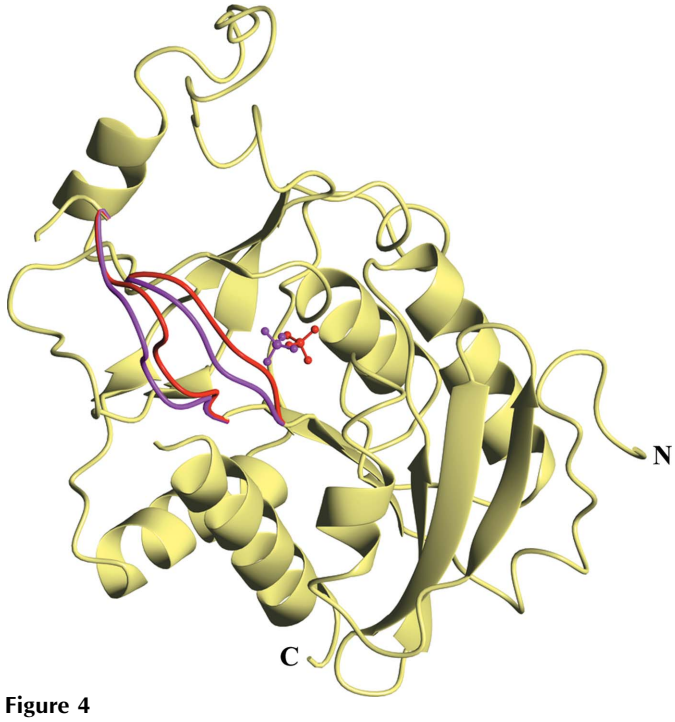

Figure 4
Subunit *B* of free-form SoUP. Two positions of the main chain, regions 88–93 and 212–219, and two positions of the disordered sulfate ion are shown; the closed conformation is in red and the open conformation is in magenta.

differ for different pairs. Thus, the region around residue 222 in subunits *C* and *D* with the ordered loop forms contacts with the region around residues 171–172 in subunits *B* and *E*, respectively, whereas the corresponding region in subunits *B* and *E* forms contacts with the region around residues 228–229 of subunits *C* and *D*, respectively; subunits *A* and *F* do not form such contacts.

The functional dimers can be classified into two types: the dimer consisting of subunits *C* and *D*, which have a closed conformation and an ordered loop 223–236, and the dimers composed of subunits *A* and *B* and of subunits *E* and *F*, in which one of the subunits has a closed conformation whereas the other has an open conformation, and in which the loop 223–236 is disordered. The dimer composed of subunits *C* and *D* is somewhat more stable compared with the other two dimers. For this dimer, according to *PISA* (Krissinel & Henrick, 2007), the $\Delta^{i}G$, which indicates the solvation free-energy gain upon formation of the interface, is −19.8 and −19.7 kcal mol^{-1} in the complex and the free form, respectively, whereas the corresponding energies for the other dimers vary in the range from −17.5 to −17.9 kcal mol^{-1}. The difference in the $\Delta^{i}G$ values may be attributed to the fact that residues 225 and 226 are involved in the formation of the interface between subunits *C* and *D*, whereas this region is disordered in the other four subunits. This is additional evidence that the dimers are not identical.

In the structure of uridine phosphorylase from *V. cholerae* (PDB entry 4k6o; Prokofev *et al.*, 2014; 1.17 Å resolution), the region around the corresponding threonine (Thr93) in two subunits is also disordered and corresponds to both the closed and open conformations, whereas the other four subunits have the closed conformation. In the structure of wild-type purine nucleoside phosphorylase containing phosphate ions in the active site (Mikleušević *et al.*, 2011), the active site has a closed conformation in two subunits and an open conformation in four subunits. As in SoUP, the two subunits which differ in conformation from the other four are next to each other but belong to different functional dimers. Additionally, the authors of the cited study found two types of dimers in wild-type purine nucleoside phosphorylase containing sulfate ions in the active site, with the subunits of one of the dimers being related by a crystallographic twofold axis.

3.2. The active site

The active site is located at the interface between two subunits of the homodimer and includes highly conserved amino-acid residues mainly from one subunit (subunit *F* in the complex of SoUP with uridine) and also from another subunit (*E*) of the functional dimer (Fig. 5). It should be noted that the uridine molecule occupies the active site in only one subunit (*F*) of the hexamer (Figs. 2*b* and 6), whereas sulfate ions are located in all six subunits. Additionally, there are two sulfate ions on a crystallographic threefold axis coordinated by Arg175 of three subunits.

This site can be divided into three subsites: the pyrimidine-binding, ribose-binding and phosphate-binding sites. The pyrimidine-binding site is formed by the residues Phe159, Gln163, Arg165, Tyr192, Glu193 and Met194 of subunit *F* (Fig. 6). The residues Gln163 and Arg165, which form hydrogen bonds to the pyrimidine moiety, play a key role in the recognition of the pyrimidine moiety of the substrate. The O2 and N3 atoms of the uracil base form hydrogen bonds to Gln163. The O4 atom of uracil forms a hydrogen bond to Arg165 and a water molecule (2.94 Å), which in turn is hydrogen-bonded to Arg220 (OW···NH1, 2.79 Å) and the above-mentioned Gln163 (OW···OE1, 2.86 Å). The conformation of Arg165 is fixed not only by the hydrogen bond to the uracil base, but also by hydrogen bonds to Gln163, Glu224, and Phe159. The orientations of all residues involved in the

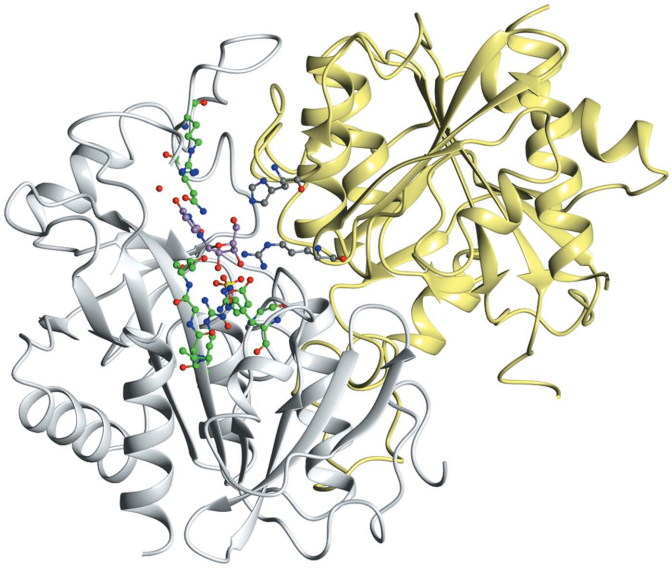

Figure 5
Functional dimer of SoUP in the structure of the complex formed by subunits *E* and *F*. The active-site residues, uridine and sulfate are shown. The uridine molecule is in lilac, the active-site residues of subunit *F* are in green and the active-site residues of subunit *E* are in grey.

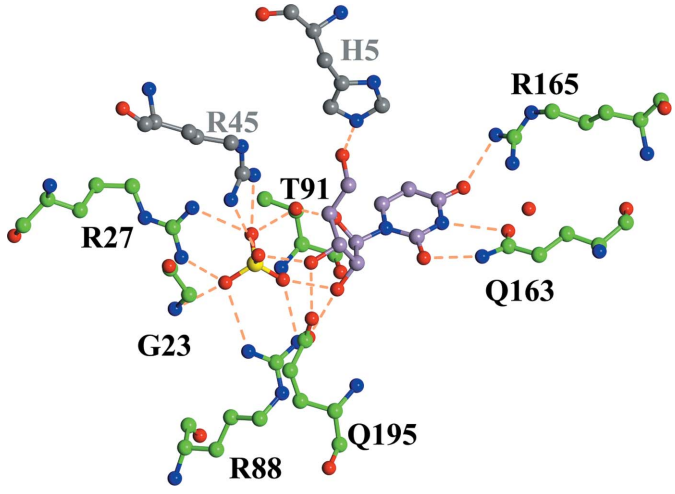

Figure 6
The active site containing uridine and a sulfate ion. Residues belonging to subunits *E* and *F* are shown in grey and green, respectively. Hydrogen bonds are shown as orange broken lines.

Table 2
Hydrogen bonds between uridine and amino-acid residues.

Atom of uridine	Atom of amino-acid residue	Hydrogen-bond distance (Å)
O2	Gln163 NE2	3.06
N3	Gln163 OE1	2.90
O4	Arg165 NH2	2.78
O2′	Arg88 NH2	3.30
O2′	Glu195 OE1	2.66
O2′	Met194 N	3.23
O3′	Glu195 OE2	2.86
O4′	Thr91 OG1	2.98
O5′	His5 NE2†	2.53

† The His5 residue belongs to subunit *E*.

binding of the uracil base remain unchanged in all six subunits. However, it should be noted that Glu224, which is involved in hydrogen bonding to Arg165 in subunits *A*, *C*, *D* and *F*, was not located in the electron-density maps for subunits *B* and *E*.

The pyrimidine-binding active site also contains a structural water molecule. In the complex with uridine, this molecule is present in all subunits except for subunit *E* and is hydrogen-bonded to the side chains of Gln163 and Arg220 (Figs. 2 and 6). In the active site of subunit *F* this water molecule is additionally hydrogen-bonded to O4 of uridine. Therefore, it can be hypothesized that the absence of this water molecule in subunit *E* is associated with the binding of uridine in subunit *F*, which causes structural changes in another subunit (subunit *E*) of the homodimer. For instance, the movement of the amino-acid region 212–219 is accompanied by a change in the conformation of the side chain of Arg220, which is more distant from the active site and cannot be involved in hydrogen bonding to the structural water molecule. Additionally, Arg165, which forms a hydrogen bond to uridine in subunit *F*, also moves somewhat away from the active site in subunit *E*. This difference in the position of Arg165 results in this residue in subunit *F* being additionally stabilized *via* a strong hydrogen bond to Glu224 (2.84 Å), owing to which the flexible loop is more ordered in subunit *F* compared with subunit *E*. In the free form, the structural water molecule is present in all six subunits, where it is also hydrogen-bonded to Gln163 and Arg220. This is additional evidence that the binding of the uridine molecule may be responsible for the absence of the structural water in subunit *E*.

The ribose-binding site is formed by the Arg88, Thr91, Met194 and Glu195 residues of one subunit of the homodimer and the His5 residue of the adjacent subunit. The ribose moiety is held in place by hydrogen bonds formed to the above-mentioned residues. It should be noted that the conformation of the side chain of Arg88 is identical in all subunits because it is fixed by hydrogen bonds to the Glu193 residue and the sulfate ion. The hydrogen-bond lengths formed by the uridine molecule are listed in Table 2.

Arginine residues and a threonine residue play a key role in the formation of the phosphate-binding site. The latter is located closer to the surface. The binding of the sulfate ion involves interactions with the side chains of Arg27, Arg88 and Thr91 and the main-chain N atoms of Gly23 and Thr91 of one subunit and Arg45 of another subunit (Table 2, Fig. 6). The

Table 3
Hydrogen bonds formed by sulfates in three types of subunits in complex with uridine.

Atom		Hydrogen-bond distance (Å)		
Sulfate	Amino-acid residue/ uridine/glycerol	Subunit *F*	Subunit *E*	Subunit *C*
O1	Arg88 NH2	2.75	3.03	2.91
	Thr91 N	3.07	—	3.05
	Uridine O2′/glycerol O1	2.75	—	2.86
O2	Arg88 NH1	3.12	—	3.01
	Arg27 NH2	3.13	—	2.99
	Gly23 N	2.77	2.92	2.83
O3	Arg45 NH1†	2.95	2.89	2.85
	Uridine O3′	2.54	—	—
O4	Thr91 OG1	2.57	—	2.66
	Arg27 NH1	2.88	—	2.92
	Arg45 NH2†	3.06	3.13	3.02

† The Arg45 residue belongs to the adjacent subunit of the homodimer.

sulfate ion also forms strong hydrogen bonds to the ribose O2′ and O3′ atoms.

The sulfates are in somewhat different environments in different subunits of the hexamer (Table 3). In the complex with uridine, the most substantial differences are observed between subunits *F* and *E*, which form a homodimer. Recall that subunit *F* is that in which the uridine molecule is located. As opposed to subunit *F*, the sulfate ion in subunit *E* is only hydrogen-bonded to the side-chain atoms of Arg88 and the main-chain O atom of Gly23, as well as to Arg45 of subunit *F*, while there are no hydrogen bonds to Arg27 and Thr91. Additionally, the hydrogen bonds to the sulfate in subunit *F* containing uridine in the active site are slightly shorter compared with those observed in subunit *E*. This sulfate ion has half occupancy. In agreement with the classification of subunits considered above, the sulfate-binding site in subunit *B* is also half occupied. It should be noted that differences between the sulfate-binding sites are also observed in the crystal structure of the free form of SoUP, in which the sulfate ion is disordered in the same two subunits (*B* and *E*), which are characterized by both the closed and open conformations (see above; Fig. 4). In these subunits, the active sites contain glycerol molecules with an occupancy of 0.8, and Arg88 is disordered over two positions. The sulfate ion with higher occupancy (0.7) in subunit *B* and the sulfate ion in one half-occupied site in subunit *E* form hydrogen bonds similar to those in the other four subunits, whereas the sulfate ion with lower occupancy (0.3) in subunit *B* and the sulfate ion at another half-occupied site in subunit *E* are hydrogen-bonded only to Thr91, Arg45 of the adjacent subunit and water molecules. In addition to the amino-acid residues, the sulfate ions in subunits *A*, *C*, *D* and *F* are also hydrogen-bonded to glycerol molecules.

Additionally, the active sites of four subunits (*A*, *C*, *D* and *F*), *i.e.* those having a closed conformation in the free form, contain a chloride ion (Fig. 2*a*). The latter ion is hydrogen-bonded to the structural water. It should be noted that subunits *B* and *E*, in which Cl^- is absent, are those which have a disordered region around Thr91 and have both the closed and open conformations (see above).

Glycerol, which was present in the reservoir solution used for crystallization, was found to be involved in the active sites in both the free form of SoUP and its complex with uridine. This molecule mimics the ribose ring of the substrate (Fig. 7). In the complex of SoUP, a glycerol molecule is present in three subunits (*A*, *C* and *D*), *i.e.* in the subunits with the closed conformation. The glycerol molecule forms a hydrogen bond to the main-chain O atom of Thr91, two strong hydrogen bonds to O1E and OE2 of Glu195 and two hydrogen bonds to the sulfate ion. In the free form of SoUP, each active site contains a glycerol molecule (Fig. 2*a*). In subunits *B* and *E*, in which the active site has both closed and open conformations, the glycerol molecule has an occupancy of 0.8 and forms hydrogen bonds to Glu195 and one of the two positions of the disordered residue Arg88. In the other four subunits, glycerol is present with full occupancy and forms a short hydrogen bond to Thr91 (2.6 Å), as well as hydrogen bonds to the side chain of Glu195 and the sulfate ion. As mentioned above, these residues are also involved in the hydrogen bonding to the ribose moiety of uridine (see Table 2). It should be noted that glycerol-mimicking substrates have also been found, for example, in β-glycosidase from *Acidolobus saccharovorans* (Trofimov *et al.*, 2013).

In a recent study (Paul *et al.*, 2010), the authors reported the observation of an unexpected species in the active site of uridine phosphorylase from *E. coli*. The crystals were grown by the cocrystallization of UP with 5-fluorouridine and sulfate in the presence of glycerol. The electron-density map clearly showed the presence of three species, which were interpreted as 5-fluorouracil, sulfate and glycal. However, the electron density for the C1′ atom of the putative glycal was not observed in any of the six subunits. Meanwhile, a glycerol molecule could be fitted well into this electron density. Hence, it cannot be ruled out that this species is actually glycerol, which was present in the solution and which mimics part of the ribose moiety, as was observed for SoUP in the free state.

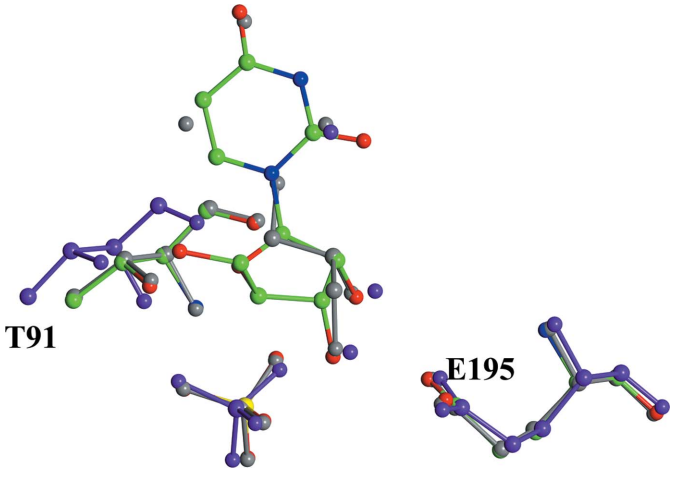

Figure 7
Superposition of the active sites of subunit *F* containing uridine (green), subunit *A* containing glycerol (grey) and subunit *E* with the open conformation (purple) in the structure of the complex of SoUP.

As mentioned above, the chain region including the active-site residue Thr91, which is involved in hydrogen bonding to uridine and sulfate, along with its adjacent amino-acid sequence residues, changes its position in subunit *E* compared with all of the other five subunits of the complex. Thus, it is located at a greater distance from the sulfate (3.47 Å in subunit *E* *versus* 2.57 Å in subunit *F*), resulting in a larger active-site pocket. When the *E* and *F* subunits are superimposed, the distance between the C^{α} atoms of Thr91 of these subunits is 1.37 Å. It should be noted that different positions of the chain region including the active-site threonine have also been observed, for example, in the structure of uridine phosphorylase complexed with anhydrouridine (Lashkov *et al.*, 2010). However, the authors did not discuss this fact.

Let us also note that in the structure of UP from *E. coli* the substrate molecules were also found only in part of the active sites (Caradoc-Davies *et al.*, 2004).

3.3. Conformation of uridine in complex with uridine phosphorylase

In the uridine molecule, the base can rotate about the glycosidic bond that links the C1′ atom of the sugar moiety to the N1 atom of the pyrimidine. The orientation of the base relative to the sugar is defined by the glycosidic bond torsion angle O4′—C1′—N1—C2 (according to the IUPAC nomenclature) denoted χ. There are two energetically favourable orientations of the sugar with respect to the base, *syn* or *anti*, characterized by χ angles in the ranges 0–90° and 270–360° for *syn* and 90–270° for *anti* (Saenger, 1984). The rotation about the glycosidic linkage is hindered by steric encumbrance between the base H6 proton and the ribose H2′ proton (the conformation corresponding to this barrier is referred to as high-*anti*) and between the C2 carbonyl group and the H2′ proton (the relevant conformation is referred to as high-*syn*) (Mikhailov *et al.*, 1999). The former barrier is relatively low and can be overcome at room temperature, whereas the latter barrier is much higher (43–46 kJ mol^{-1}), as estimated by NMR spectroscopy (Wegner & Jochims, 1979). For pyrimidine nucleosides, the *anti* conformation is much more favourable than the *syn* conformation because the latter gives rise to a short contact between the O2 atom of the base and the C2′ atom of the ribose.

In SoUP, uridine adopts a high-energy high-*syn* conformation (χ = 134°; see Fig. 8). The distance between the carbonyl O2 atom and the C2′ atom of the ribose is 2.92 A (the distance between H2′ and O2 can be estimated as 1.8 Å). There is also a very short contact between the O2′ and O3′ atoms (2.25 Å).

It should be noted that the structure of the complex with uridine has previously only been determined for UP from *S. typhimurium* (PDB entry 2hwu; V. I. Timofeev, A. G. Gabdulkhakov, M. V. Dontsova, W. Voelter, G. S. Kachalova & A. M. Mikhailov, unpublished work). In this structure, the uridine conformation is somewhat similar to that observed in SoUP. However, since the structure of *S. typhimurium* UP was solved at low resolution (2.91 Å), a detailed comparison of the

major determinants of uridine binding in the two structures is not reasonable.

The ribose ring in SoUP adopts the least favorable O4′-*exo* conformation in terms of steric interactions, in which the substituents at C2′ and C3′ are eclipsed and both bulky substituents, the –CH_2OH group and the heterocyclic base, are in axial positions. This geometry orients the nucleoside so that the C1′ atom is near one of the O atoms of the phosphate ion. Additionally, the ribose ring is highly flattened, as can be seen from the torsion angles (C1′—C2′—C3′—C4′ is −2° and the largest endocyclic torsion angle is 20°). The O4′ atom deviates from the plane passing through the other four atoms of the five-membered ring by 0.3 Å. This is highly unusual for nucleosides, for which the nonplanar conformation of the furanose ring is energetically favourable.

Therefore, the substrate is bound in the active site of SoUP in the energetically unfavourable strained high-*syn* conformation. This is possible owing to the presence of interactions between the substrate and the protein environment (active-site residues) mainly through hydrogen bonds. It should be emphasized that all of the functional groups of the uridine molecule are involved in an extensive hydrogen-bond network (see Fig. 6 and Table 2). Thus, the O4 atom of uridine forms a hydrogen bond to Arg165 NH2, the O2 atom forms a hydrogen bond to the NE2 atom of Gln163, and the latter residue also forms a hydrogen bond to the N3 atom of uridine *via* its OE1 atom. Additionally, there is a hydrogen bond between the O4 atom of uridine and a water molecule, which in turn is involved in hydrogen bonding to Arg220 NH1. It should be noted that this water molecule is also rigidly fixed by a hydrogen bond to Arg165. In addition, there is a hydrogen bond between Arg165 NH1 and Gln163 OE1, which are involved in hydrogen bonding to uridine. The ribose ring is held in place by hydrogen bonds between O4′ and Thr91 OG1, O2′ and Arg88 NH2, O2′ and Met194 N, O2′ and Glu195 OE1, O3′ and Glu195 OE2, and O5′ and His5 NE2. It has been shown (Gao *et al.*, 2006) that the high-energy conformation of uridine is also stabilized through electrostatic interactions. In SoUP, electrostatic interactions can occur between the substrate molecule and the Arg165 and Glu163 residues.

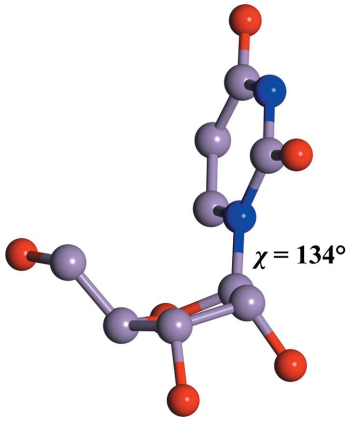

Figure 8
High-*syn* conformation of uridine; the ribose ring adopts a highly flattened envelope conformation with the O4′ atom as the flap.

Based on an analysis of the relevant literature on the conformations of nucleosides in solution and the energy barriers to rotation, the difference between the ground state of uridine and uridine complexed with UP may be estimated to be 15 kcal mol^{-1}. This estimate is supported by a recent publication (Miyahara *et al.*, 2014). As a result of the binding in a high-energy conformation, the glycosidic bond is weakened.

Our observations are consistent with the results of single point energy calculations for the binding of purine nucleosides to purine nucleoside phosphorylases (PNPs; Erion *et al.*, 1997), which predicted the binding conformation to enhance phosphorolysis through ligand strain. A high-energy conformation was also found in the PNP structures of the NP-I family (Pugmire & Ealick, 2002). It should be noted that PNP and UP share general mechanistic features. Thus, nearly all of the functional groups of the substrate form hydrogen bonds to active-site amino acids, rather large regions of the enzymes undergo conformational changes upon substrate binding, and the substrates are bound in an energetically unfavourable conformation. Based on a high degree of similarity of both the sequence and structure of the ribose and phosphate-binding residues between these enzymes, it was suggested that uridine phosphorylase has a similar catalytic mechanism to that of PNP (Caradoc-Davies *et al.*, 2004).

4. Conclusions

A comparison of the high-resolution X-ray structures of uridine phosphorylase from *S. oneidensis* MR-1 in the free form and in complex with the natural substrate uridine showed that the subunits of the hexamer are not identical, with the active sites having either an open or a closed conformation. The subunits can be divided into three types according to their structural features.

The uridine molecule is present in only one subunit of the hexamer (subunit *F*). A structural water molecule is present in the pyrimidine-binding active sites of all subunits both in the free form and in the complex, except for subunit *E* in the complex. Since the latter subunit is involved in the functional dimer with the uridine-containing subunit *F*, it can be hypothesized that the absence of this water molecule in subunit *E* is associated with the binding of uridine in subunit *F*, which causes structural changes in subunit *E*. In the monomers with the closed conformation, a glycerol molecule mimicking the ribose moiety of uridine is present in all active sites except for that containing uridine.

The disorder of the main chain around the active-site residue Thr91 observed in two subunits in the free form of SoUP is absent in the complex. Therefore, the binding of uridine probably contributes to the ordering of the main chain in the vicinity of the active site.

The uridine molecule has the energetically unfavourable high-*syn* conformation with a nearly planar ribose ring.

This work was supported by the Russian Science Foundation (RSCF, grant No. 14-24-00172). SNM and KMP were

supported by the 'Molecular and Cellular Biology' Program of the Russian Academy of Sciences.

References

Caradoc-Davies, T. T., Cutfield, S. M., Lamont, I. L. & Cutfield, J. F. (2004). *J. Mol. Biol.* **337**, 337–354.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Erion, M. D., Stoeckler, J. D., Guida, W. C., Walter, R. L. & Ealick, S. E. (1997). *Biochemistry*, **36**, 11735–11748.

Gao, X.-F., Huang, X.-R. & Sun, C.-C. (2006). *J. Struct. Biol.* **154**, 20–26.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kanzaki, A., Takebayashi, Y., Bando, H., Eliason, J. F., Watanabe, S., Miyashita, H., Fukumoto, M., Toi, M. & Uchida, T. (2002). *Int. J. Cancer*, **97**, 631–635.

Krenitsky, T. A., Mellors, J. W. & Barclay, R. K. (1965). *J. Biol. Chem.* **240**, 1281–1286.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Lashkov, A. A., Zhukhlistova, N. E., Gabdoulkhakov, A. H., Shtil, A. A., Efremov, R. G., Betzel, C. & Mikhailov, A. M. (2010). *Acta Cryst.* D**66**, 51–60.

Lashkov, A. A., Zhukhlistova, N. E., Seregina, T. A., Gabdoulkhakov, A. G. & Mikhailov, A. M. (2011). *Crystallogr. Rep.* **56**, 560–589.

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

Lewkowicz, E. S. & Iribarren, A. M. (2006). *Curr. Org. Chem.* **10**, 1197–1215.

Li, N., Smith, T. J. & Zong, M.-H. (2010). *Biotechnol. Adv.* **28**, 348–366.

Lunin, V. Y. (2012). *Acta Cryst.* D**68**, 732–734.

McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). *Acta Cryst.* D**67**, 386–394.

Mikhailopulo, I. A. & Miroshnikov, A. I. (2010). *Acta Naturae*, **2**, 36–59.

Mikhailopulo, I. A. & Miroshnikov, A. I. (2011). *Mendeleev Commun.* **21**, 57–68.

Mikhailov, S. N., Lysov, Yu. P. & Yakovlev, G. I. (1999). *Mol. Biol. (Russia)* **33**, 340–354.

Mikleušević, G., Štefanić, Z., Narczyk, M., Wielgus-Kutrowska, B., Bzowska, A. & Luić, M. (2011). *Biochimie*, **93**, 1610–1622.

Miyahara, T., Nakatsuji, H. & Wada, T. (2014). *J. Phys. Chem. A*, **118**, 2931–2941.

Mordkovich, N. N., Manuvera, V. A., Veiko, V. P. & Debabov, V. G. (2012). *Biotekhnologiia (Mosc.)*, **1**, 21–30.

Mordkovich, N. N., Safonova, T. N., Manuvera, V. A., Veiko, V. P., Polyakov, K. M., Alekseev, K. S., Mikhailov, S. N. & Popov, V. O. (2013). *Dokl. Biochem. Biophys.* **451**, 187–189.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.

Paul, D., O'Leary, S. E., Rajashankar, K., Bu, W., Toms, A., Settembre, E. C., Sanders, J. M., Begley, T. P. & Ealick, S. E. (2010). *Biochemistry*, **49**, 3499–3509.

Prokofev, I. I., Lashkov, A. A., Gabdoulkhakov, A. G., Dontsova, M. V., Seregina, T. A., Mironov, A. S., Betzel, C. & Mikhailov, A. M. (2014). *Acta Cryst.* F**70**, 60–63.

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

Roosild, T. P., Castronovo, S., Fabbiani, M. & Pizzorno, G. (2009). *BMC Struct. Biol.* **9**, 14.

Saenger, W. (1984). *Principles of Nucleic Acid Structure.* New York: Springer.

Safonova, T. N., Mordkovich, N. N., Polyakov, K. M., Manuvera, V. A., Veiko, V. P. & Popov, V. O. (2012). *Acta Cryst.* F**68**, 1387–1389.

Tran, T. H., Christoffersen, S., Allan, P. W., Parker, W. B., Piškur, J., Serra, I., Terreni, M. & Ealick, S. E. (2011). *Biochemistry*, **50**, 6549–6558.

Trofimov, A. A., Polyakov, K. M., Tikhonov, A. V., Bezsudnova, E. Y., Dorovatovskii, P. V., Gumerov, V. M., Ravin, N. V., Skryabin, K. G. & Popov, V. O. (2013). *Dokl. Biochem. Biophys.* **449**, 99–101.

Utagawa, T. (1999). *J. Mol. Catal. B Enzym.* **6**, 215–222.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Veiko, V. P., Siprashvili, Z. Z., Ratmanova, K. I., Gul'ko, L. B., Mironov, A. A., Andriukhina, R. V. & Debabov, V. G. (1994). *Dokl. Akad. Nauk*, **339**, 819–821.

Watanabe, S. & Uchida, T. (1995). *Biochem. Biophys. Res. Commun.* **216**, 265–272.

Wegner, G. & Jochims, J. C. (1979). *Chem. Ber.* **112**, 1941–1947.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yeates, T. O. (1988). *Acta Cryst.* A**44**, 142–144.