Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of orotate phosphoribosyltransferase from the caries pathogen *Streptococcus mutans*

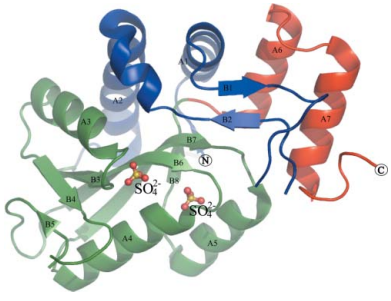
Chao-Pei Liu,[a,b] Rui Xu,[a] Zeng-Qiang Gao,[a] Jian-Hua Xu,[a] Hai-Feng Hou,[a] Li-Qin Li,[a,b] Zhun She,[a,b] Lan-Fen Li,[c] Xiao-Dong Su,[c] Peng Liu[a] and Yu-Hui Dong[a]*

[a]Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, People's Republic of China, [b]Graduate School of the Chinese Academy of Sciences, Beijing 100049, People's Republic of China, and [c]National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Science, Peking University, Beijing 100871, People's Republic of China

Correspondence e-mail: dongyh@ihep.ac.cn

Received 25 December 2009
Accepted 11 March 2010

PDB Reference: orotate phosphoribosyltransferase, 3dez.



Orotate phosphoribosyltransferase (OPRTase) catalyzes the OMP-forming step in *de novo* pyrimidine-nucleotide biosynthesis. Here, the crystal structure of OPRTase from the caries pathogen *Streptococcus mutans* is reported at 2.4 Å resolution. *S. mutans* OPRTase forms a symmetric dimer and each monomer binds two sulfates at the active sites. The structural symmetry of the sulfate-binding sites and the missing loops in this structure are consistent with a symmetric catalysis mechanism.

1. Introduction

The Gram-positive bacterium *Streptococcus mutans*, a member of the viridans streptococci, is the most probable aetiological agent of dental caries (Loesche, 1986) and is also a possible causative agent of subacute bacterial endocarditis (Ullman *et al.*, 1988). The SMU.1221 gene from *S. mutans* encodes an orotate phosphoribosyltransferase consisting of 209 residues with a molecular weight of about 23 kDa. Orotate phosphoribosyltransferase (OPRTase; EC 2.4.2.10) plays an essential role in the *de novo* and recycling pathways of purine, pyrimidine and pyridine metabolism as well as in the biosynthesis of histidine and tryptophan (Musick, 1981). It catalyzes the Mg^{2+}-dependent reaction between α-D-5-phosphoribosyl-1-pyrophosphate (PRPP) and orotic acid (OA) to yield pyrophosphate (PP_i) and β-*N*-riboside monophosphate (OMP). Despite in-depth enzymological studies (Victor *et al.*, 1979; Bhatia *et al.*, 1990; Wang *et al.*, 1999; McClard *et al.*, 2006) and the publication of many crystal structures of OPRTase both in the apo form and bound to substrates and inhibitors (Scapin *et al.*, 1994, 1995; Henriksen *et al.*, 1996; Gonzalez-Segura *et al.*, 2007), its precise enzymatic mechanism remains elusive. A recent investigation using isothermal titration calorimetry (ITC) and ^{31}P NMR spectroscopy revealed that OMP or PRPP bind to *Saccharomyces cerevisiae* OPRTase in a 1:1 molar ratio. Therefore, the authors concluded that the enzymatic mechanism was an alternating-sites model of catalysis (McClard *et al.*, 2006). However, crystal structures of *S. cerevisiae* OPRTase showed the substrates to be bound in the active sites symmetrically, with the two highly conserved loops remaining in either the open or the closed conformation (Gonzalez-Segura *et al.*, 2007). In this paper, the 2.4 Å resolution structure of *S. mutans* OPRTase is presented; the structure is consistent with a symmetric catalysis mechanism.

2. Materials and methods

2.1. Cloning, expression and purification

The SMU.1221 gene was amplified from *S. mutans* genomic DNA by polymerase chain reaction (PCR; Saiki *et al.*, 1988). The sense primer was 5′-CGC**GGATCC**ATGACATTAGCAAAAGACATCG-CTC-3′ and the antisense primer was 5′-GTG**CTCGAG**TTAGTCT-TGCCAAGTCTCTTGGT-3′; the primers introduced *Bam*HI and *Xho*I restriction sites, respectively (shown in bold). The PCR-amplified fragment was then cloned into expression vector pET28a (Novagen) by conventional cloning methods with an N-terminal His_6

Table 1
Summary of data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Space group	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 73.80$, $c = 184.8$
Resolution (Å)	50–2.40 (2.49–2.40)
Completeness (%)	99.9 (100)
R_{merge}† (%)	7.1 (48)
Mean $I/\sigma(I)$	33.9 (3.8)
Redundancy	12.0 (10.9)
Refinement statistics	
Resolution (Å)	50–2.40
No. of unique reflections	19803
R_{work}/R_{free}‡ (%)	0.218/0.249
No. of atoms	
Protein	3089
Ligand/ion	20
Water	89
Average B values (Å^2)	
Protein	49.67
Ligand/ion	63.66
Water	47.06
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.40
Ramachadran plot statistics (%)	
Most favoured	91.0
Additional allowed	9.0

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ R_{work} and $R_{free} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$ for the working set and test set (5%) of reflections, respectively, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes for reflection hkl.

tag (MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGS). The constructed plasmid was transformed into *Escherichia coli* strain BL21 (DE3) for expression. The transformed *E. coli* strain BL21 (DE3) cells were cultured in Luria–Bertani (LB) medium containing 50 µg ml^{-1} kanamycin at 310 K until the OD_{600} reached 0.6. Gene expression was then induced with 0.5 m*M* isopropyl β-D-1-thio-galactopyranoside (IPTG) and the culture was incubated at 298 K for a further 12 h. The cells were harvested by centrifugation (8000 rev min^{-1}, 10 min, 277 K) and suspended in buffer *A* (20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl) containing 1 m*M* phenylmethylsulfonyl fluoride (PMSF). The resuspended cells were disrupted by sonication on ice and the cell debris was removed by centrifugation (16 000 rev min^{-1}, 60 min, 277 K). The supernatant was loaded onto a 5 ml Ni^{2+}-chelating affinity column (HiTrap, GE Healthcare, USA) pre-equilibrated with buffer *A*. Impurities were washed out with 10% buffer *B* (20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 500 m*M* imidazole) and the target protein was eluted with a linear gradient of 10–100% buffer *B*. The eluted protein was concentrated by ultrafiltration in an Amicon cell (Millipore, California, USA) and was purified to homogeneity by size-exclusion chromatography (HiLoad Superdex 75 XK16/60, GE Healthcare, USA) using buffer *C* (20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* DTT). From gel-filtration chromatography, it was deduced that the protein exists as a dimer in solution. The molecular mass of the purified protein was about 27 kDa, which was in agreement with the predicted mass (with an additional 4 kDa from the fusion part). The purity of *S. mutans* OPRTase was examined by SDS–PAGE at each step.

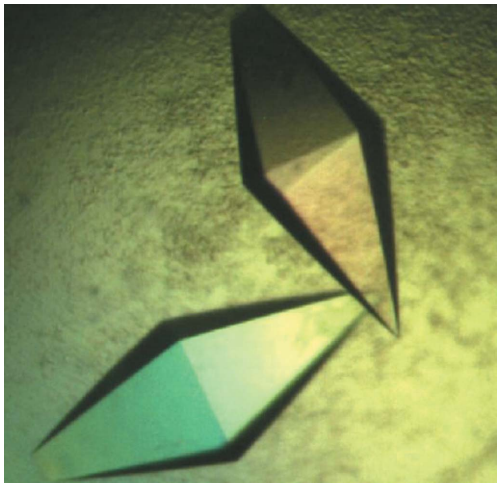


Figure 1
Crystals of *S. mutans* OPRTase.

2.2. Crystallization, data collection and processing

The purified protein was concentrated to 40 mg ml^{-1} by ultrafiltration (Millipore, California, USA). Crystallization conditions were screened by the hanging-drop vapour-diffusion method using the Index, Crystal Screen I and Crystal Screen II kits (Hampton Research, California, USA) at 293 K. 1 µl reservoir solution was mixed with 1 µl freshly purified protein solution and equilibrated against 500 µl reservoir solution. Crystals suitable for X-ray diffraction were obtained from a condition containing 15 mg ml^{-1} protein, 0.1 *M* Tris–HCl pH 8.6 and 2.3 *M* ammonium sulfate (Fig. 1). The crystals were cryoprotected by soaking them in a solution containing 20%(*v*/*v*) PEG 400, 0.1 *M* Tris–HCl pH 8.6 and 2.3 *M* ammonium sulfate for a few seconds; they were then flash-cooled in a 100 K nitrogen stream. A data set was collected at a wavelength of 1.0 Å on beamline 3W1A at BSRF (Beijing Synchrotron Radiation Facility, Beijing, People's Republic of China) using a MAR165 CCD detector. A complete data set was collected to a resolution of 2.4 Å. The diffraction data were indexed, integrated and scaled using *DENZO* and *SCALEPACK* from the *HKL*-2000 program package (Otwinowski & Minor, 1997). The *S. mutans* OPRTase crystal belonged to space group $P4_12_12$, with unit-cell parameters $a = 73.80$, $c = 184.8$ Å (Table 1), and contained two molecules per asymmetric unit, with a solvent content of about 48% (Matthews, 1968).

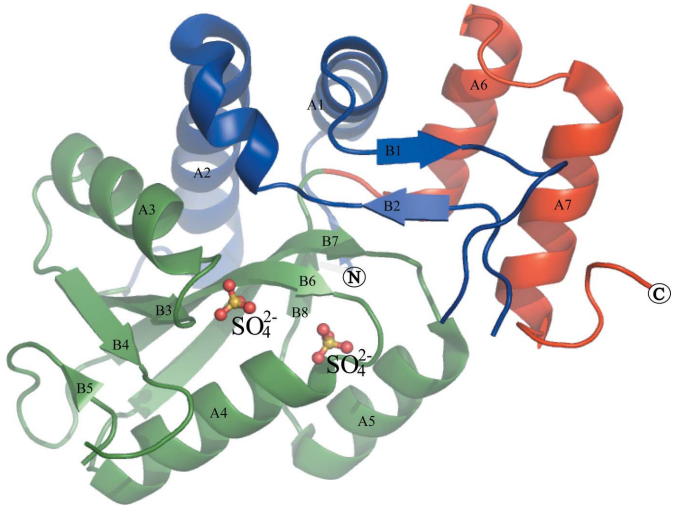


Figure 2
Structure of the *S. mutans* OPRTase monomer. The N-terminal hood domain (blue), including helices A1 and A2, extends across the upper portion of the figure. The C-terminal helices A6 and A7 (red) can be seen behind the hood structure. The α/β nucleotide-binding fold (green) occupies the lower portion of the figure. The two bound SO_4^{2-} ions are shown as ball-and-stick models.

2.3. Structure determination and refinement

The *S. mutans* OPRTase structure was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2005; Storoni *et al.*, 2004) as implemented in *CCP*4 (Collaborative Computational Project, Number 4, 1994). The initial search model was created from the monomer of *S. pyogenes* OPRTase (PDB code 2aee; C. Chang, H. Li, F. Collart & A. Joachimiak, unpublished work) using *CHAINSAW* (Stein, 2008). The sequences of the *S. mutans* and *S. pyogenes* OPRTases have 85% identity. Two significant molecular-replacement solutions were found with *Z* scores of 10.2 and 10.1. The structure was refined using *REFMAC*5 (Murshudov *et al.*, 1997, 1999) and was manually rebuilt with *Coot* (Emsley & Cowtan, 2004). Simulated annealing in *CNS* (Brünger *et al.*, 1998; Brunger, 2007) was used to help with fitting difficult loops. In the final stages of the refinement, electron density consistent with four SO_4^{2-} ions and 89 water molecules was fitted into difference electron density and the model converged with an *R* factor of 0.218 and an R_{free} of 0.249. The data-collection and refinement statistics are summarized in Table 1. The model quality was assessed using *PROCHECK* (Laskowski *et al.*, 1993): 91% of the residues were in the most favoured region of the Ramachandran plot, with the remainder being in the additional allowed area. All figures portraying protein models were prepared using *PyMOL* (DeLano Scientific, San Carlos, California, USA). The molecular coordinates and structure factors have been deposited in the Protein Data Bank with accession code 3dez.

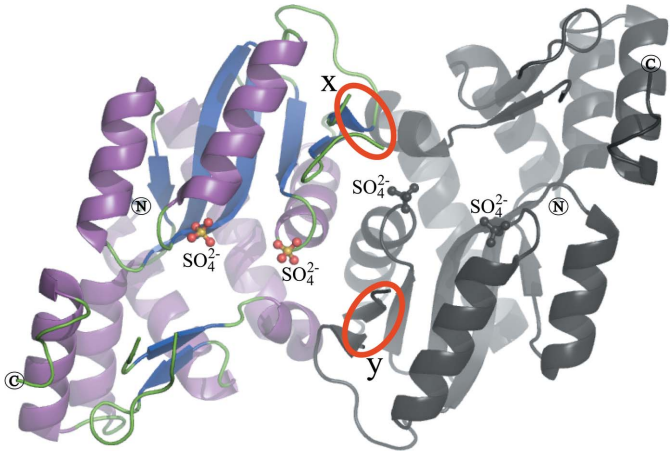


Figure 3
Architecture of the *S. mutans* OPRTase dimer. One monomer is coloured grey and the other is coloured according to the secondary structure. Four sulfates are shown as ball-and-stick models; the loops closing onto the active sites are disordered in the electron density of the *S. mutans* OPRTase structure. The flexible loops are highlighted by red circles and labelled 'x' and 'y'.

3. Results and discussion

3.1. Structure description

The monomer of *S. mutans* OPRTase contains 209 residues and is composed of eight β-strands (B1–B8) and seven α-helices (A1–A7) (Fig. 2). Its overall structure is similar to those of OPRTases from other species such as *S. typhimurium* (Scapin *et al.*, 1994), *E. coli* (Henriksen *et al.*, 1996), *S. cerevisiae* (Gonzalez-Segura *et al.*, 2007) and *S. pyogenes* (PDB code 2aee; C. Chang, H. Li, F. Collart & A. Joachimiak, unpublished work). The structure can be divided into the same three structural regions as defined in *S. typhimurium* OPRTase: the N-terminal, core and C-terminal regions. The N-terminal region (blue; residues 1–66) contains the same α-helices (A1 and A2) and two short antiparallel β-strands (B1 and B2). These two β-strands are known as the 'hood' and partially envelop the OA ring in the ternary complex, as shown in the *S. typhimurium* and *S. cerevisiae* structures. The core region (green; residues 67–73) contains a six-stranded twisted β-sheet (B3–B8), instead of the four strands found in the *S. typhimurium* and *E. coli* OPRTases and the five found in the *S. cerevisiae* OPRTase, surrounded by three α-helices (A3–A5). The C-terminal region (red; residues 174–209) is composed of two antiparallel α-helices (A6–A7) connected by a loop structure, which serves to stabilize the conformation of the hood. The *S. mutans* OPRTase structure is found to adopt a dimeric assembly in the protein crystal with approximate dimensions of 53 × 50 × 62 Å, as shown in Fig. 3. The two molecules (grey/coloured in Fig. 3) of the asymmetric unit related by a noncrystallographic twofold axis form a homodimer. The dimer interface consists of the regions Ile14–Tyr18 (loop), Arg40–Arg49 (helix A2) and Thr75–Arg111 (helix A3–loop–strand B4–loop–strand B5); 26 residues from each monomer directly interact with the other monomer. The two monomers in the asymmetric unit exhibit an average root-mean-square deviation (r.m.s.d.) of only 0.238 Å for all aligned C^{α} atoms (except for the flexible loop) using *SUPERPOSE* (secondary-structure matching) in *CCP*4 (Collaborative Computational Project, Number 4, 1994). Comparison of the monomers revealed that the two monomers were almost identical. The monomers in *E. coli*, *S. cerevisiae* and *S. pyogenes*

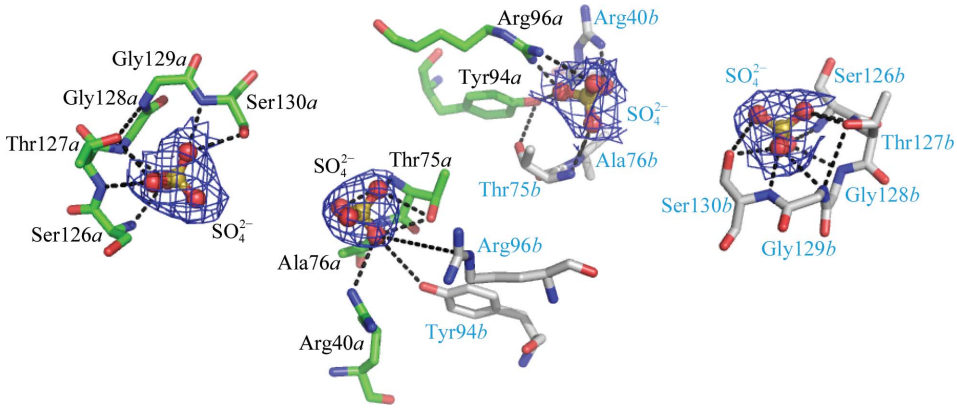


Figure 4
Hydrogen-bonding network in the SO_4^{2-}-binding sites. The σ_A-weighted $2F_o - F_c$ electron-density map of SO_4^{2-} is contoured at 1σ. Residues from the *a* chain (black) and *b* chain (cyan) around the SO_4^{2-}-binding sites are labelled. Tyr94*a* and Arg96*a* (or Tyr94*b* and Arg96*b*) are contributed to the SO_4^{2-}-binding site from the other subunit. The hydrogen-bond network is shown as black dashed lines.

OPRTases are essentially identical, with r.m.s.d.s of 0.482, 0.684 and 0.833 Å, respectively. In *S. typhimurium* OPRTase an identical dimer was formed along a crystallographic twofold axis in the apoenzyme and in the enzyme complexed with PRPP and OA. In the structure of *S. mutants* OPRTase two sulfates that inhibit enzyme activity have been found in both monomers close to the disordered loop regions.

3.2. The SO_4^{2-}-binding sites

A search model containing only one peptide chain without SO_4^{2-} ions was used in molecular replacement and four additional peaks of strong electron density were identified in the binding sites during the refinement procedure. These peaks were fitted as SO_4^{2-} ions from the ammonium sulfate in the crystallization reservoir. These SO_4^{2-} ions could be identified from large well resolved peaks in the σ_A-weighted $2F_o - F_c$ electron-density map at the 1σ level (Fig. 4). Residues from the *a* chain (black) and *b* chain (cyan) around the SO_4^{2-}-binding sites make similar hydrogen-bond interactions with sulfates in each monomer. Tyr94*a* and Arg96*a* (or Tyr94*b* and Arg96*b*) are contributed to the SO_4^{2-}-binding site from the other monomer. The hydrogen-bond networks (shown as black dashed lines in Fig. 4) in the SO_4^{2-}-binding sites are symmetric in both monomers. The binding of the sulfates mimics the binding of the pyrophosphate moiety of PRPP to OPRTase. It is similar to the *E. coli* OPRTase structure, in which sulfate has been found in both monomers in the same orientation and makes the same hydrogen-bond interactions with the monomer (Henriksen *et al.*, 1996). In the structures of the *S. typhimurium* and *S. cerevisiae* OPRTases bound to substrates the arrangements of the subunits of the dimer were identical and contained the same symmetric hydrogen-bond interactions with substrates (Scapin *et al.*, 1995; Gonzalez-Segura *et al.*, 2007).

3.3. Flexibility of the loops

A highly conserved loop (residues ^{96}RSKPKDHGAGN^{106}; indicated by the red circles labelled 'x' and 'y' in the two monomers in Fig. 3) was not determined in the model. These flexible loops with poor electron density extended into the solvent and were missing from the structure. In structures of OPRTases from various species, including apoenzymes as well as those containing substrates, products or inhibitors under different crystallization conditions (in the pH range 4.6–8.6), the loops consistently show considerable flexibility. The 'x' loop in one subunit and the 'y' loop in the other were missing in the *S. mutans* OPRTase structure (3dez; this work), with four sulfates binding in the active sites symmetrically (Fig. 5*a*). In the structure of *E. coli* OPRTase (1oro; Henriksen *et al.*, 1996) the 'y'

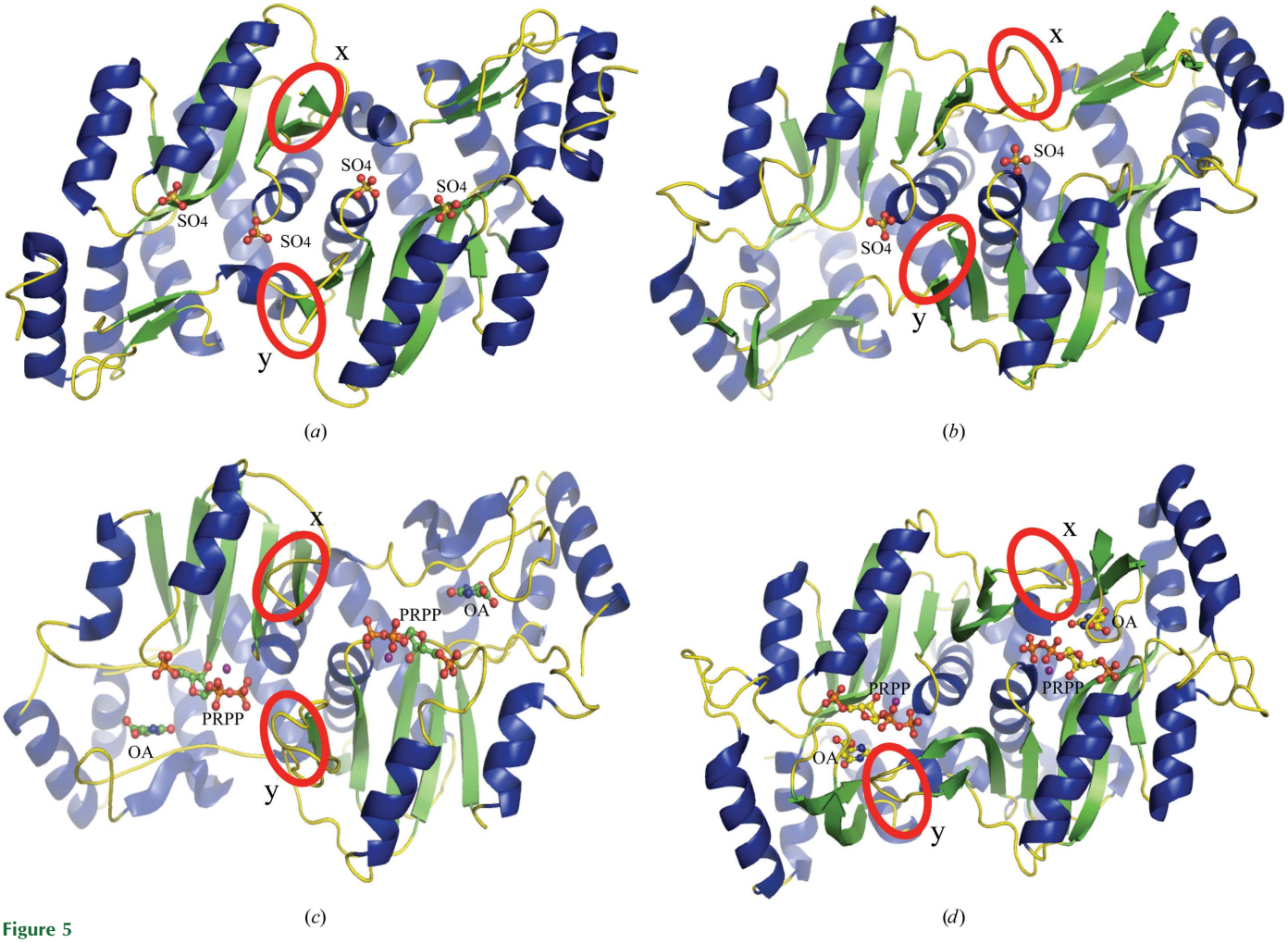


Figure 5
The flexibility of loops in published OPRTases from various species. (*a*) *S. mutans* OPRTase (PDB code 3dez). The 'x' and 'y' loops were both missing, with four sulfates binding in the active sites symmetrically. (*b*) *E. coli* OPRTase (PDB code 1oro). The 'y' loop was missing and the 'x' loop was closed, with two sulfates binding in the active sites symmetrically. (*c*) *S. cerevisiae* OPRTase (PDB code 2ps1). The 'x' and 'y' loops were both closed, with PRPP and OA binding in the active sites symmetrically. (*d*) *S. typhimurium* OPRTase (PDB code 1opr). The 'x' and 'y' loops were both open, with PRPP and OA binding in the active sites symmetrically.

loop was missing and the 'x' loop was closed, with two sulfates binding in the active sites symmetrically (Fig. 5*b*). In the structure of *S. cerevisiae* OPRTase (PDB code 2ps1; Gonzalez-Segura *et al.*, 2007) loops 'x' and 'y' were both closed, with PRPP and OA binding in the active sites symmetrically (Fig. 5*c*). In the structure of *S. typhimurium* OPRTase (1opr; Scapin *et al.*, 1995) both loops were open, with PRPP and OA binding in the active sites symmetrically (Fig. 5*d*). The other OPRTase structures [PDB codes 2pry and 2prz (Gonzalez-Segura *et al.*, 2007), 1sto (Scapin *et al.*, 1994), 2yzk (OPRTase from *A. pernix*; M. Kanagawa, S. Baba, S. Kuramitsu, S. Yokoyama, G. Kawai & G. Sampei, unpublished work), 2wns (the OPRTase domain of human uridine 5′-monophosphate synthase in complex with its substrate OMP; M. Moche *et al.*, unpublished work), 2p1z (phosphoribosyltransferase from *Corynebacterium diphtheriae*; C. Chang, H. Li, S. Clancy & A. Joachimiak, unpublished work) and 2aee (OPRTase from *S. pyogenes*; C. Chang, H. Li, F. Collart & A. Joachimiak, unpublished work)] in the Protein Data Bank are also symmetrical. Most crystal structures of OPRTase complexes show symmetrical full site occupancy and the two highly conserved loops remain in either the open or the closed conformation. These structures are consistent with a symmetric catalysis mechanism. The only exception found was PDB entry 1lh0 (*S. typhimurium* OPRTase; A. A. Federvo, K. Panneerselvam, W. Shi, C. Grubmeyer & S. C. Almo, unpublished work). This structure showed half site occupancy, with PRPP and OA bound in one site and OA only bound in the other. However, 1opr and 1lh0 describe the same protein bound with the same substrates but with different configurations of the dimers. The two structures observed in the crystals are representatives of two different conformations. One (1opr) has symmetric full site occupancy consistent with a symmetric catalysis mechanism. The other (1lh0) has asymmetric half site occupancy consistent with an asymmetric catalysis mechanism (Gonzalez-Segura *et al.*, 2007).

4. Summary and conclusions

In summary, the three-dimensional crystal structure of *S. mutans* OPRTase has been determined to 2.4 Å resolution by X-ray crystallography. The structure shows high overall similarity to those of other OPRTases. The sulfate-binding sites are structurally symmetric and two highly conserved loops are missing. The *S. mutans* OPRTase structure is consistent with a symmetric catalysis mechanism. In fact, the structures observed in the crystals may possibly not be representative of the conformations present in solution since the crystal packing may induce distortion in structures, especially in the flexible loop regions. Also, the pH values and ionic strength during crystallization are sufficient to induce subtle changes in the protein structure. Thus, further studies such as enzyme kinetics studies and solution structures are required to confirm the catalytic mechanism of *S. mutans* OPRTase.

This work was supported by grants from the National Natural Science Foundation of China (10979005) and the National Basic Research Program of China (2009CB918600).

References

Bhatia, M. B., Vinitsky, A. & Grubmeyer, C. (1990). *Biochemistry*, **29**, 10480–10487.
Brunger, A. T. (2007). *Nature Protoc.* **2**, 2728–2733.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gonzalez-Segura, L., Witte, J. F., McClard, R. W. & Hurley, T. D. (2007). *Biochemistry*, **46**, 14075–14086.
Henriksen, A., Aghajari, N., Jensen, K. F. & Gajhede, M. (1996). *Biochemistry*, **35**, 3803–3809.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Loesche, W. J. (1986). *Microbiol. Rev.* **50**, 353–380.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McClard, R. W., Holets, E. A., MacKinnon, A. L. & Witte, J. (2006). *Biochemistry*, **45**, 5330–5342.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Musick, W. D. (1981). *CRC Crit. Rev. Biochem.* **11**, 1–34.
Otwinowski, Z. & Minor, Z. (1997). *Methods Enzymol.* **276**, 307–326.
Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). *Science*, **239,** 487–491.
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.
Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Ullman, R. F., Miller, S. J., Strampfer, M. J. & Cunha, B. A. (1988). *Heart Lung*, **17**, 209–212.
Victor, J., Greenberg, L. B. & Sloan, D. L. (1979). *J. Biol. Chem.* **254**, 2647–2655.
Wang, G. P., Lundegaard, C., Jensen, K. F. & Grubmeyer, C. (1999). *Biochemistry*, **38**, 275–283.