Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The structure of *Bacillus subtilis* SPβ prophage dUTPase and its complexes with two nucleotides

Javier García-Nafría, Maria Harkiolaki,‡ Rebecca Persson,§ Mark J. Fogg and Keith S. Wilson*

Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, England

‡ Present address: Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Headington, Oxford OX3 7BN, England.
§ Present address: Semtech Metallurgy AB, Ideon Science Park, S-223 70 Lund, Sweden.

Correspondence e-mail: keith@ysbl.york.ac.uk

Received 15 December 2010
Accepted 25 January 2011

PDB References: YosS, apo form, 2xx6; dUDP complex, 2y1t; dUpNHpp–Mg^{2+} complex, 2xy3.

dUTPases are housekeeping enzymes which catalyse the hydrolysis of dUTP to dUMP in an ion-dependent manner. *Bacillus subtilis* has both a genomic and an SPβ prophage homotrimeric dUTPase. Here, structure determination of the prophage apoenzyme and of its complexes with dUDP and dUpNHpp–Mg^{2+} is described at 1.75, 1.9 and 2.55 Å resolution, respectively. The C-terminal extension, which carries the conserved motif V, is disordered in all three structures. Unlike all other trimeric dUTPases for which structures are available, with the exception of the *Bacillus* genomic enzyme, the aromatic residue covering the uridine and acting as the Phe-lid is close to motif III in the sequence rather than in motif V. This is in spite of the presence of an aromatic amino acid at the usual Phe-lid position in motif V. The alternative position of the Phe-lid requires a reconsideration of its role in the catalytic cycle of the enzyme. In the dUpNHpp–Mg^{2+} complex a water can be seen at the position expected for nucleophilic attack on the α-phosphate, in spite of motif V being disordered. Differences in the active site between the free enzyme and the dUDP and dUpNHpp–Mg^{2+} complexes shows that the triphosphate moiety needs to be in the *gauche* conformation to trigger the conformational changes that can be seen in both *B. subtilis* dUTPases.

1. Introduction

dUTPases are essential enzymes that catalyse the hydrolysis of dUTP into dUMP and PP_i in a metal-ion-dependent manner (Bertani *et al.*, 1963; Shlomai & Kornberg, 1978). Their function is to reduce the amount of dUTP in the cell nucleotide pool, as at high concentrations the promiscuity of DNA polymerase induces harmful uracil incorporation into DNA, resulting in thymine-less death (Ingraham *et al.*, 1986). In addition, since dUMP is a precursor of dTTP, its production promotes incorporation of dTTP rather than dUTP into DNA. dUTPases are classified into three families according to their oligomerization state: monomeric, homodimeric (also termed dUDPase/dUTPases) and homotrimeric. The homotrimeric family is the most extensively characterized, with structures of apoenzymes and ligand complexes from a wide range of species having been determined (reviewed in Vértessy & Tóth, 2009), and shares sequence and structural similarity to the monomeric family, while the dUDPase/dUTPases have quite different sequences and folds (Camacho *et al.*, 1997; Harkiolaki *et al.*, 2004).

Homotrimeric dUTPases form pyramidal trimers with the subunits related by a threefold axis around a central channel, whose characteristics vary with species (Takács *et al.*, 2004).

They possess three active sites, each of which is formed by five conserved motifs (McGeoch, 1990). All three subunits contribute to each active site: motif I, II and IV from one subunit, motif III from the adjacent subunit and motif V from the third subunit. Motifs I, II and III bind to the triphosphate and motif IV binds to the nucleoside. Motif V caps the active site upon ligand binding (Mol *et al.*, 1996) and is part of a C-terminal extension which is disordered in most structures determined to date, with the exception of a few nucleotide complexes: *Mycobacterium tuberculosis* dUTPase with Mg^{2+}–dUpNHpp, feline immunodeficiency virus (FIV) dUTPase with dUDP and the human enzyme with either dUMP, dUDP or dUTP (Prasad *et al.*, 2000; Chan *et al.*, 2004; Tóth *et al.*, 2007; Varga *et al.*, 2008; Mol *et al.*, 1996). Motif V is essential for activity (Vertessy, 1997; Freeman *et al.*, 2009), becoming ordered for catalysis (Varga *et al.*, 2007) and re-opening with the release of product. Motif V contains an aromatic residue which stacks over the uracil ring upon dUTP binding and was first termed the 'Phe-lid' in the human enzyme (Mol *et al.*, 1996). Mutants of this residue have been shown to be deficient in activity (Mol *et al.*, 1996; Freeman *et al.*, 2009). An aromatic residue (or in a few sequences a histidine) is present at an equivalent position in all trimeric dUTPase sequences reported to date, with the exception of the *Bacillus subtilis* genomic enzyme (García-Nafría *et al.*, 2010).

B. subtilis possesses two homotrimeric dUTPases, an SPβ prophage form (encoded by the *yosS* gene) and a genomic form (encoded by the *yncF* gene), the sequences of which are 93% identical (Fig. 1). For convenience, we henceforth refer to the two gene products as YosS and YncF, respectively. Previously, we have reported the structures of the YncF apoenzyme and of its complex with dUpNHpp and Ca^{2+} (an inhibitory metal), which revealed that the Phe-lid feature, Phe93, was located close to motif III in the sequence rather than in motif V and that its side chain flips its conformation to stack against the uridine in the dUpNHpp complex with a different conformation to that seen in the apoenzyme (García-Nafría *et al.*, 2010). Indeed, YncF does not have an aromatic amino acid at the expected position in motif V. In contrast, YosS has an aromatic residue both in the generally conserved motif V position and in the same position as that seen in YncF next to motif III. The main focus of the present study was to identify which of these aromatic residues functions as the Phe-lid in YosS and the implications of this for the mechanism of phosphorolysis.

A structure of apo YosS has been deposited (PDB code 2baz), but only the crystallization and preliminary X-ray analysis have been published (Li *et al.*, 2009). Here, we report the structures of apo YosS and its complexes with dUDP and dUpNHpp–Mg^{2+}. The structures pose a number of questions

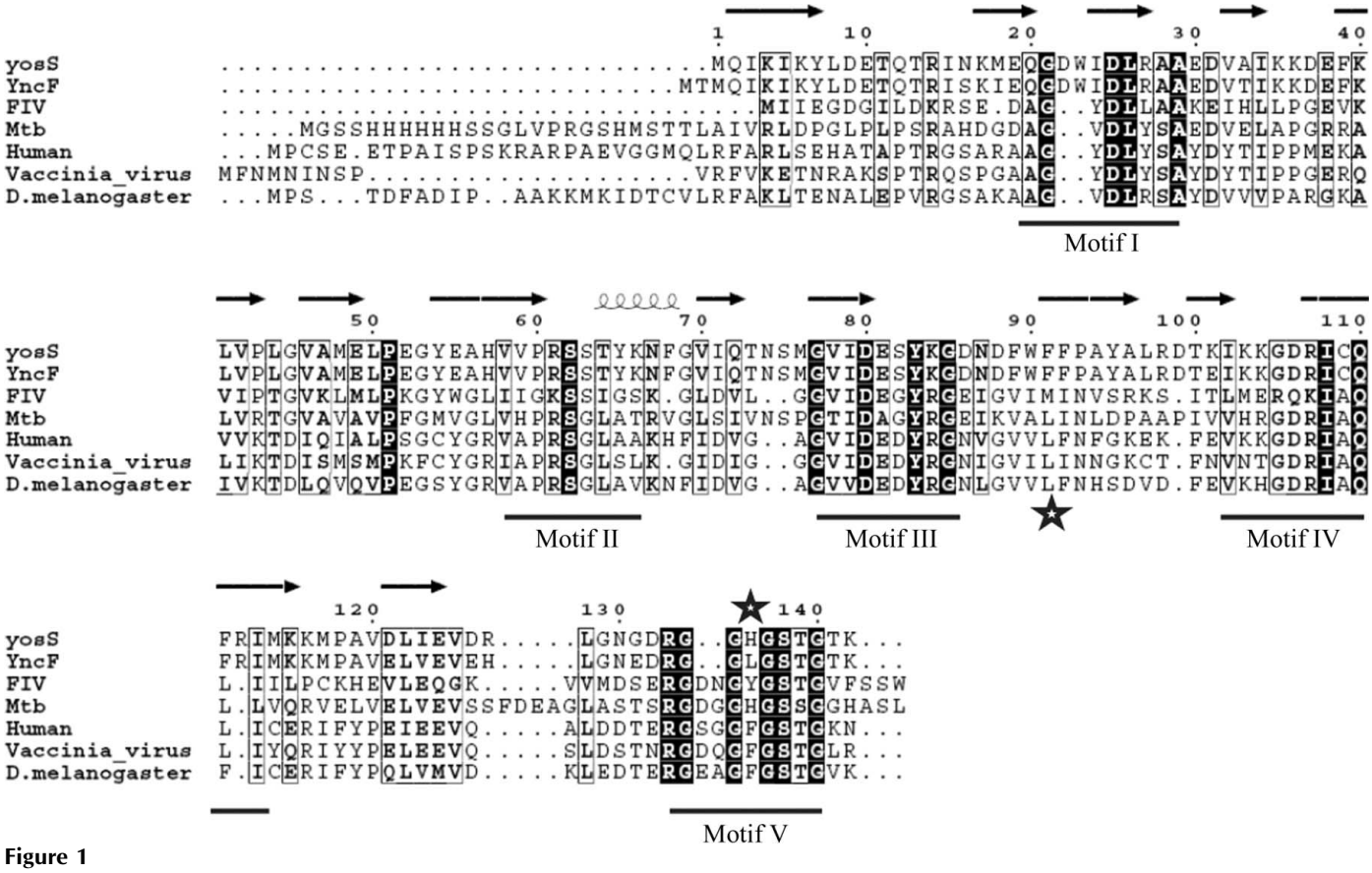


Figure 1
Sequence alignment of some representative dUTPases. Conserved residues are coloured in black, while residues in boxes indicate those with conserved physicochemical properties. The five conserved motifs are underlined and labelled. The positions of the Phe-lids are indicated by stars. YosS has two aromatic residues: one in the conserved position in motif V and one where the Phe-lid of the YncF is positioned.

about the ligand-specific rearrangements in the active site and confirm that the triphosphate moiety must be in the *gauche* conformation with bound metal in order to trigger the changes in the enzyme required for catalysis.

2. Methods

2.1. Protein overexpression and purification

The *B. subtilis yosS* gene coding for the SPβ prophage dUTPase was cloned, overexpressed in *Escherichia coli* and purified using phosphocellulose chromatography as described previously (Persson *et al.*, 2005). The resulting protein was dialysed against 10 m*M* MOPS pH 7 containing 1 m*M* β-mercaptoethanol. This construct was used for structure determination of the apoenzyme and its complex with dUDP.

For the determination of the structure of the dUpNHpp complex, a new construct was used. The *yosS* gene was amplified by polymerase chain reaction (PCR) from *B. subtilis* genomic DNA using KOD Hot Start DNA polymerase (Novagen) and complementary gene-specific primers, YosSF (5′-CAGGGACCA**CATATGC**AAATTAAAATCAAATAC-TTAGATGAAACAC-3′) and YosSR (5′-GAA**GGATCC**TT-ACTTCGTACCTGTTGAACCGAGTCCTCC-3′), designed to incorporate *Nde*I and *Bam*HI restriction-endonuclease recognition sequences (bold), respectively. The purified PCR fragment was digested with *Nde*I and *Bam*HI and ligated into similarly digested and purified plasmid pET-26b (Novagen). Recombinants were transformed into *E. coli* NovaBlue Giga-Singles cells (Novagen) and plated onto LB–agar medium to kanamycin (kan) resistance (30 µg ml^{-1}). Recombinant plasmids were purified from overnight cultures (10 ml LB medium, 30 µg ml^{-1} kan) of kan-resistant colonies. The presence of a gene insert was established by restriction-endonuclease digestion using the enzymes *Nde*I and *Bam*HI (New England Biolabs) and subsequently confirmed by DNA sequencing. The pET-26b*yosS* plasmid was transformed into *E. coli* B834 (DE3) for protein overexpression. Cultures of *E. coli* B834DE3/pET-26b*yosS* were grown overnight in 10 ml kan-supplemented (30 µg ml^{-1}) LB medium and used to inoculate 1 l kan-supplemented (100 µg ml^{-1}) ZYM-5052 auto-inducing medium (Studier, 2005) for protein overexpression. The resulting protein was purified using phosphocellulose chromatography as described previously (Persson *et al.*, 2005) and dialysed against 10 m*M* MOPS pH 7 containing 1 m*M* β-mercaptoethanol.

2.2. Crystallization

The apoenzyme was crystallized in a hexagonal form from 0.1 *M* imidazole–malate buffer pH 5.5, 30%(*v*/*v*) PEG 600 using protein concentrations ranging from 0.5 to 2.5 mg ml^{-1} (Persson *et al.*, 2001). The dUDP complex was crystallized from 0.1 *M* HEPES buffer pH 7.5, 10%(*v*/*v*) PEG 8000 and 8% ethylene glycol with 5 m*M* dUDP and 20 m*M* $SrCl_2$ using protein concentrations between 0.5 and 2.5 mg ml^{-1}.

The complex with dUpNHpp–Mg^{2+} was prepared by diluting the protein to 14 mg ml^{-1} in 50 m*M* Tris–HCl pH 8, 150 m*M* NaCl and adding dUpNHpp in a 3:1 ligand:protein molar ratio with 10 m*M* $MgCl_2$ followed by 30 min incubation at 277 K. The commercial Crystal Screen and Crystal Screen 2 (Hampton Research) and PACT (Molecular Dimensions) crystallization screens were used and diffraction-quality crystals of roughly 70 × 20 × 20 µm in size grew in 3.4 *M* 1,6-hexanediol, 0.1 *M* Tris pH 8.5, 0.2 *M* $MgCl_2$. X-ray data were collected to resolutions of 1.75 and 1.9 Å on beamline ID14-1 (wavelength 0.934 Å) at the ESRF for the apoenzyme and the dUDP complex, respectively, as described previously (Persson *et al.*, 2001) and to 2.55 Å resolution on beamline I02 (wavelength 0.979 Å) at the Diamond Light Source for the dUpNHpp–Mg^{2+} complex.

2.3. Structure solution and refinement

The structure of the apoenzyme was solved in space group $P6_3$ by molecular replacement with *AMoRe* (Trapani & Navaza, 2008) using FIV dUTPase as a search model. Phases were improved with *DM* (Cowtan & Main, 1998) and an initial model was built using *ARP*/*wARP* as described previously (Perrakis *et al.*, 2001). The structure of the dUDP complex was solved in space group $P2_12_12_1$ using the apoenzyme as a search model in *AMoRe* (Trapani & Navaza, 2008). Refinement of both structures was carried out with *REFMAC* (Murshudov *et al.*, 1997) using the TLS option, and manual correction, rebuilding and inspection were performed with *QUANTA* (Accelrys Inc., USA). The structures were re-refined and minor manual corrections were performed with more recent versions of *REFMAC* and *Coot* (Emsley & Cowtan, 2004), giving final R_{work}/R_{free} values of 17.14/21.86% and 14.8/17.6% for the apoenzyme and the dUDP complex, respectively. The structure of the dUpNHpp–Mg^{2+} complex was solved by molecular replacement using the apoenzyme as a search model in *Phaser* (McCoy *et al.*, 2007). Model completion and refinement were performed with *Coot* and *REFMAC*. The final residuals were R_{work} = 20.16% and R_{free} = 25.91%.

All residues in all three structures have good geometry and are in allowed regions of the Ramachandran plot, except for Ser75, the outlier geometry of which has biological importance as reported for the equivalent residues in other dUTPases (Mol *et al.*, 1996). The apo form has four protomers in the asymmetric unit with very similar folds, as shown by the low r.m.s.d. of the C^α positions (0.33 Å for 123 C^α atoms) calculated with *SSM* (Krissinel & Henrick, 2004). The dUDP and dUpNHpp–Mg^{2+} complexes both belong to space group $P2_12_12_1$ with two trimers in the asymmetric unit, again with similar folds (dUDP complex, 0.27 Å r.m.s.d. for 125 aligned C^α atoms of the six protomers; dUpNHpp–Mg^{2+} complex, 0.5 Å r.m.s.d. for 128 C^α atoms of the six protomers). The C-terminus is disordered in all protomers of all three structures from residues 129–132 onwards. In the structures of the dUpNHpp complex and the native enzyme there is a significant (7σ level) positive difference density peak in the three-fold channel where Asp37 and Lys3 from each chain converge. A buffer molecule or metal has been identified in the central channel of other dUTPases, but nothing could be modelled in

Table 1
Data collection and refinement.

	YosS	YosS–dUDP	YosS–dUpNHpp–Mg^{2+}
Source	ESRF	ESRF	Diamond Light Source
Beamline	ID14-1	ID14-1	I02
Wavelength (Å)	0.934	0.934	0.979
Data collection			
Space group	$P6_3$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters			
a (Å)	102.8	99.5	98.40
b (Å)	102.8	99.3	97.99
c (Å)	86.2	99.3	98.47
$\alpha = \beta$ (°)	90	90	90
γ (°)	120	90	90
Resolution (Å)	20–1.75 (1.81–1.75)	20–1.90 (1.97–1.90)	69.6–2.55 (2.59–2.55)
R_{merge} (%)	6.1 (69.9)	5.4 (45.6)	4.7 (58.6)
$\langle I/\sigma(I)\rangle$	26.4 (3.0)	25.4 (3.1)	37.06 (2.55)
Completeness (%)	99.9 (100)	99.8 (99)	99.8 (97.9)
Multiplicity	5.74 (2.8)	5.15	7.2 (6.4)
Refinement			
Resolution (Å)	1.75	1.90	2.55
No. of reflections	52846	78511	31651
R_{work}/R_{free}	17.14/21.86	14.8/17.6	20.16/25.91
No. of atoms			
Protein	4118	6269	6089
Ligand/ion	—	144	168
Water	477	512	91
B factors (Å^2)			
Protein	15.16	29.43	42.32
Ligand/ion	—	30.64	47
Water	25.10	35.44	38.38
R.m.s. deviations			
Bond lengths (Å)	0.024	0.026	0.017
Bond angles (°)	1.910	2.177	1.726

the present structure with a satisfactory chemistry. A summary of the data-collection and refinement statistics is shown in Table 1. The multiplicity in the high-resolution shell is missing for the dUDP complex: the images were integrated some years ago and this datum has been lost. However, the completeness, $\langle I/\sigma(I)\rangle$ and R_{merge} in the high-resolution shell confirm the data quality.

3. Results

3.1. Structure

There is no density for residues 16–21 in any of the chains in all three structures and they are presumed to be disordered, which is similar to the situation in YncF. The C-terminus, which contains the conserved motif V, is disordered from between residues 129 and 132 (with some variation between the protomers) to the end of the chain at residue 142.

3.2. Structure of the apoenzyme

Each chain contains 12 β-strands and one α-helix, forming a jelly-roll-like fold as reported for other dUTPases (Vértessy & Tóth, 2009; Fig. 2). β5 is split into two parts by an insertion including Ser75, the only Ramachandran outlier, thus creating a bulge that accommodates the uracil ring as seen in most other dUTPases, for example the human enzyme (Mol *et al.*, 1996). The three subunits form a pyramidal trimer around a threefold axis with extensive subunit interactions. The C-terminal β12 strand crosses over to the sheet in the adjacent subunit with extensive quaternary interactions around the central channel, which shows alternating layers of residues of hydrophobic and hydrophilic character.

3.3. The YosS–dUDP and YosS–dUpNHpp–Mg^{2+} complexes

The two structures are isomorphous, with six protomers in the asymmetric unit forming two biological trimers: *A*, *B* and *C*, and *D*, *E* and *F*. The fold shows no substantial changes from that of the apoenzyme as evidenced by the low 0.63 Å r.m.s.d. of the C^{α} atoms of the superposed trimers. In both complexes there are three ligands bound in each of the two trimers. The uracil-binding pocket is typical of that seen in other dUTPases. The uracil ring hydrogen bonds to the enzyme mimicking DNA base pairing, with the uracil O2 and N3 binding to the amino and carboxyl groups of the main-chain atoms of residue 91, respectively, while uracil O4 hydrogen bonds to the amino group of Asn74. The deoxyribose sugar is sandwiched between Tyr83 and Ile79. The aromatic ring of Tyr83 is responsible for the stereochemical discrimination for deoxyribose against ribose. Chemical derivatization of this residue has been shown to perturb ligand binding in the *E. coli* enzyme (Vertessy *et al.*, 1996).

3.4. dUDP complex

There is no metal bound to the ligand in this complex. This is different from the dUDP complex of equine infectious anaemia virus dUTPase (Dauter *et al.*, 1999), in which the strontium ions were clearly seen. We have no simple explanation for this difference, other than that the affinity for this metal may vary between the two proteins or the different crystallization buffers affected the relative binding constants.

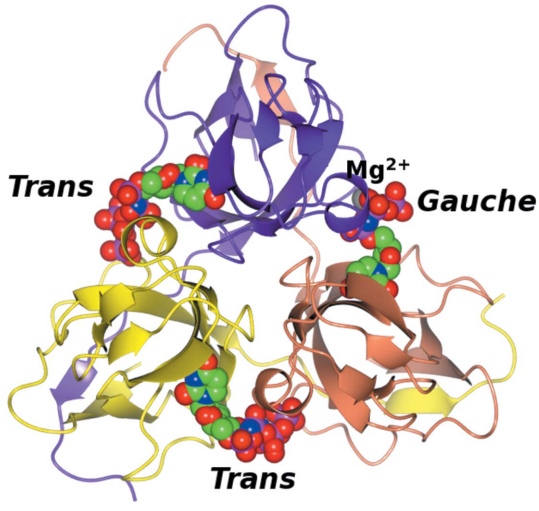


Figure 2
YosS trimer complexed with the dUpNHpp substrate analogue. The trimer is shown as ribbons coloured by chain. The overall fold of the trimer shows pairwise, central channel and arm-crossing interactions between the subunits. The dUpNHpp ligands, depicted as spheres, bind at the interface of the pairwise domain interactions. Two are in in the *trans* conformation and the third is *gauche*, with the bound Mg^{2+} ion shown as a grey sphere.

The α-phosphate is well ordered and in the same location in all six active sites. In contrast, the position of the β-phosphate varies. In three active sites (between chains *A*–*C*, *B*–*A* and *D*–*F*) it has a single well defined position, but this is the result of Lys101 and Lys103 from an adjacent trimer in the crystal packing against the α- and β-phosphates. We propose that the other three, open, active sites (between *C*–*B*, *F*–*E* and *E*–*D*) more closely represent the conditions in solution as they are not influenced by crystal contacts. In these sites, while it was only possible to model one position for the β-phosphate (with hydrogen bonds to surrounding waters), the surrounding residual difference density strongly indicates significant disorder with multiple conformations. The β-phosphate of the dUDP in *E. coli* dUTPase shows similar disorder, with alternate conformations and partial occupancy (Larsson *et al.*, 1996). The side chain of Ser62 also has alternative conformations.

3.5. dUpNHpp–Mg^{2+} complex

While dUpNHpp–Mg^{2+} is bound in all six active sites, in three of them (between chains *A*–*C*, *B*–*C* and *D*–*F*) the phosphate groups are in the *trans* conformation, with the other three being *gauche* (*A*–*B*, *D*–*E* and *F*–*E*). The *trans* conformation is assumed to be inactive as it does not promote nucleophilic attack on the α-phosphate by a water molecule (Chan *et al.*, 2004; Kovári *et al.*, 2008). In the present structure the *trans* conformation is stabilized by crystal contacts equivalent to those seen in the YncF–dUpNHpp–Ca^{2+} complex (García-Nafría *et al.*, 2010) and similar to those in the YosS–dUDP complex. We assume these *trans* ligands not to be biologically relevant and focus on the *gauche* form.

dUpNHpp is bound in a similar manner to dUDP, with two substantial differences. (i) The aromatic side chain of Phe91 flips roughly 180° in the *gauche* active sites (see below) in order to stack over the uracil at a distance of 3.6 Å, forming displaced face-to-face π–π interactions. This residue thus contributes two functions: the main-chain atoms are responsible for the accommodation of the uracil ring, while the side chain acts as the Phe-lid. In the *trans* active sites there is no density for the Phe91 chain, which is disordered. In reality it is probably split between two alternative conformations as in YncF, but the resolution of the present analysis is insufficient to confirm this. (ii) The triphosphate moiety is well ordered, with an associated magnesium ion. The β-phosphate is in a different position to that seen in the dUDP complex, with the P atoms in the two structures lying ∼4.5 Å from one another. This results in different interactions and induces different conformational changes in the enzyme (see below; Fig. 3).

The triphosphate-like moiety interacts extensively with the enzyme and with a number of waters in a similar manner to that seen in other dUTPases (Mol *et al.*, 1996; Chan *et al.*, 2004). The O2 of the α-phosphate is hydrogen bonded to the side-chain amino group of Gln110 and the main-chain amino group of Ser62. The OH of Ser62 contacts the α–β imino group of the ligand. The β-phosphate O2 hydrogen bonds to the hydroxyl group of Ser63 and the NH2 and N^{ε} atoms of Arg61. All three phosphates coordinate the magnesium ion as in other dUTPase–dUpNHpp complexes and also in DNA polymerase β (Sawaya *et al.*, 1997). There is only very low electron density at the sites expected for the three water molecules expected to complete the octahedral coordination of the magnesium, probably owing to the limited resolution. These waters are placed so as to contact the carboxyl group of Asp25, as seen in other dUTPases including YncF (García-Nafría *et al.*, 2010).

Density for the catalytic water can be seen at 0.8σ in an equivalent position to that in the YncF–dUpNHpp complex. It is positioned at a distance of 3.7 Å to perform in-line attack on the α-phosphate and also contacts the Val78 carbonyl O atom. Asp80, the residue that has been proposed to act as a general base and activate the catalytic water for dUTP hydrolysis (Barabás *et al.*, 2004), points away from the active site in the the same way as found in the YncF structure.

The coordinates for apo YosS and the dUDP and dUpNHpp–Mg^{2+} complexes have been deposited in the PDB with codes 2xx6, 2y1t and 2xy3, respectively.

4. Discussion

4.1. Overall structure

B. subtilis possesses two homotrimeric dUTPases: the genomic (YncF) and SPβ prophage (YosS) enzymes. In the present study, the structures of the YosS apoenzyme and its dUDP and dUpNHpp–Mg^{2+} complexes have been determined. The fold in all three structures is that of a typical homotrimeric dUTPase, with a minor difference being found in residues 16–21, which are disordered as in YncF. The structure of YosS closely resembles that of YncF, with an ∼0.5 Å r.m.s.d. for the C^{α} atoms of the trimer. The three protomers have a jelly-roll-like fold and form trimers around a threefold axis, with the central channel having an alternating hydrophobic and hydrophilic character similar to those of YncF (García-Nafría *et al.*, 2010) and FIV dUTPase (Prasad *et al.*, 1996). As the physicochemical character of the channel has been correlated with stability (Takács *et al.*, 2004), we would expect the latter dUTPases to have an intermediate stability between those with a hydrophobic and those with a hydrophilic channel. The C-terminal extension is disordered in all structures, even when the active site contains a triphosphate nucleotide analogue in the *gauche* conformation and a magnesium ion.

4.2. The ligand complexes and the active site

The binding mode of the ligands is similar to that in other dUTPases, with the uracil and deoxyribose making conserved contacts with the protein. The uracil-binding pocket is created by the hinge made by Ser75, as in other dUTPases such as the human enzyme (Mol *et al.*, 1996). The active sites in the YosS structures can be split into three structural categories.

(i) The active sites in the apoenzyme and the dUDP complex are identical; no significant changes in the conformations of the surrounding residues are induced by the dUDP inhi-

bitor. Even the Phe-lid is in the same open conformation, which indicates that binding of the uracil ring is not in itself sufficient to trigger flipping of the lid.

(ii) For the dUpNHpp–Mg^{2+} *trans* ligands there is little change in the active-site residues compared with the apo form or the dUDP complex. However, the Phe-lid now adopts two alternate conformations and is roughly equally split between the open and closed forms. The β-phosphate in the *trans* ligands occupies the same position as in the dUDP complex. The side chain of Ser62 has two alternative conformations in these first two types of active site.

(iii) The dUpNHpp–Mg^{2+} *gauche* ligands induce much more significant changes. A number of residues around the phosphate moiety (notably Arg61, Ser62, Ser63 and Arg107) adapt to accommodate the *gauche* triphosphate analogue; in parallel, the Phe-lid flips fully to stack over the uracil. In addition there is a small bending movement within the subunit carrying the Phe-lid (discussed below). Most of the conformational changes around the phosphates are in residues from the α-helix (residues ~62–68), which tilts somewhat relative to its position in the apoenzyme. Ser62, which has been proposed to play a dual role in dUTPases (Palmén *et al.*, 2008) and has alternate conformations in the apoenzyme, the dUDP complex and the *trans* dUpNHpp sites, becomes ordered and points towards the imino group of the ligand. It has been suggested that the β-OH of Ser62 destabilizes the ground-state complex while stabilizing the transition state. The resolution does not allow accurate positioning of the serine side chain in the dUpNHpp complex; however, similar arrangements were found in YncF. Ser63 moves to accommodate the ligand and interact with the β-phosphate. The side chain of Arg61 moves so as to contact the β-phosphate O2, which also hydrogen bonds

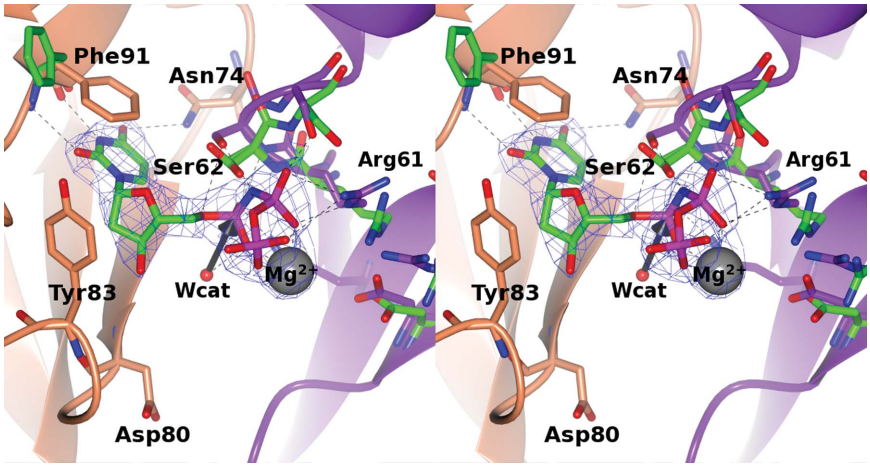
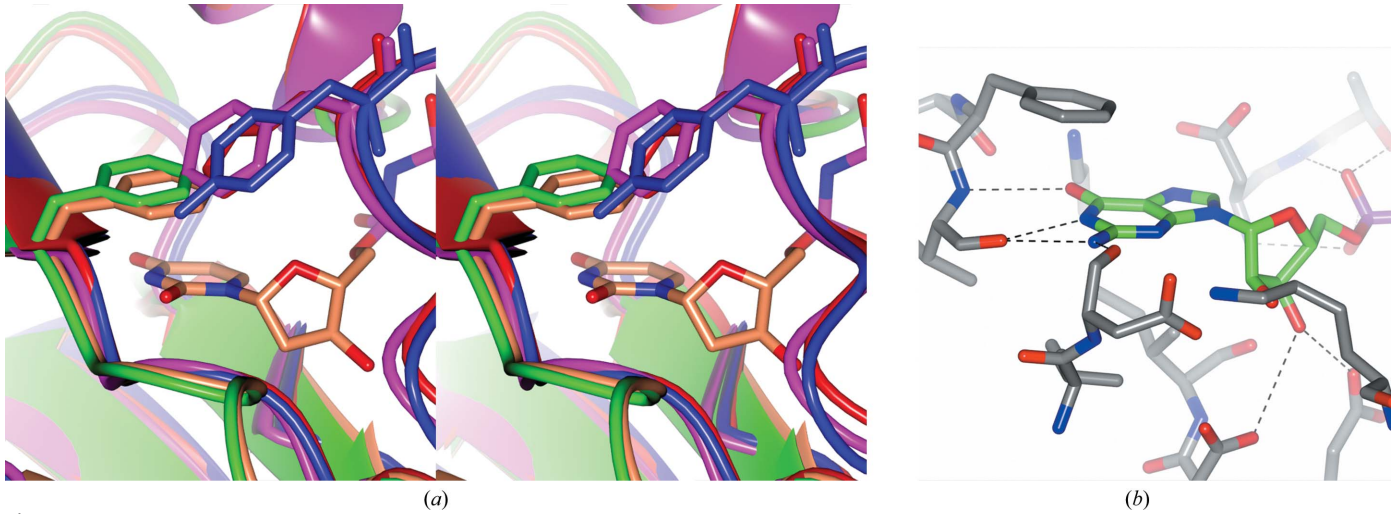


Figure 3
Stereoview of the YosS active site with *gauche* dUpNHpp–Mg^{2+}. Subunits *A* and *B* are coloured coral and purple, respectively, and are displayed as ribbons. dUpNHpp is displayed as cylinders and the magnesium ion is shown as a grey sphere, with the associated $2F_o - F_c$ electron density contoured at the 1σ level. The residues in the active site are depicted as cylinders, with the C atoms coloured according to subunit. Hydrogen bonds are shown as dashed lines. In the complex, Phe91 stacks against the uracil and Tyr83 against the deoxyribose. Residues shown as green cylinders correspond to the apo enzyme and reveal the rearrangements which occur upon ligand binding. The catalytic water is positioned to perform an in-line attack on the α-phosphate indicated by an arrow. The proposed general base, Asp80, points away from the active site.



(*a*) (*b*)

Figure 4
The Phe-lid in dUTPases and human HPGT. (*a*) Superposition of the human (magenta), *M. tuberculosis* (red), FIV (blue), YncF (green) and YosS (coral) dUTPases. The dUpNHpp is that from YosS, but its position is essentially the same in all four structures. The Phe-lid of the two *Bacillus* dUTPases is on a different secondary-structural element from the rest of the dUTPases. The *B. subtilis* enzymes have a face-to-face π–π interaction, while the rest have edge-to-face interactions. (*b*) The active site of human HPGT in complex with GMP. C atoms are shown as grey and green cylinders for the protein and the ligand, respectively. There is a face-to-face interaction between the guanine ring and the phenylalanine which is similar to the stacking seen in YosS and YncF.

to the NH2 and N^{ε} atoms. Arg107 moves closer to the ligand but does not make direct contacts, possibly because its role is to stabilize the negative charges from the phosphate chain. That these rearrangements only occur upon binding of the *gauche* form of the substrate is in keeping with this conformation being required for effective catalysis and is one factor in discrimination against dUDP as a substrate.

4.3. The Phe-lid

Motif V contains a conserved arginine and a glycine-rich sequence similar to the P-loop in ATPases and GTPases but with a very different three-dimensional structure and an aromatic residue. The latter, termed the Phe-lid, stacks over the uracil ring in those complexes in which the C-terminal extension and motif V take up an ordered conformation and can be a Phe, His or Tyr. Mutation of the Phe-lid to another aromatic residue does not influence the activity of the enzyme (Pecsi *et al.*, 2010). Recently, the *B. subtilis* YncF was shown to lack an aromatic residue at the expected position in motif V, with the role of the Phe-lid instead being provided by Phe93 positioned next to motif III (García-Nafría *et al.*, 2010). YosS shows 93% identity to YncF, but unlike the latter has an aromatic residue at the conserved position in motif V. However, the present YosS structures reveal that the Phe-lid residue, Phe91, is the same as in YncF and is close to motif III. In the dUpNHpp complexes of both enzymes the π–π interaction is displaced face-to-face rather than edge-to-face as seen in other dUTPases (Fig. 4*a*). The term Phe-lid was first used for human dUTPase (Mol *et al.*, 1996); the authors pointed out its similarity to human hypoxanthine-guanine phosphoribosyltransferase (HGPT) in complex with GMP (Eads *et al.*, 1994), which also has a phenylalanine side chain stacking over the base of the nucleotide. Indeed, the *B. subtilis* Phe-lids resemble this enzyme rather than other dUTPases, with the Phe-lid having a similar face-to-face stacking against the base (Fig. 4*b*).

It remains to be seen how the two bulky residues in YosS (Phe91 and His136) are arranged upon C-terminal closure, as there would be a steric clash between the two aromatic rings if they behaved as in most other dUTPases. One possibility is that the conserved histidine in motif V would displace the lid seen in the present YosS–dUpNHpp complex, but this change is less likely as the YncF dUTPase does not have a lid residue in motif V and has the same lid residue next to motif III. There is a two-residue deletion between the two conserved glycines in motif V in both *B. subtilis* dUTPases (Fig. 1). Based on the sequence alignment, this might be thought to distort the C-terminal structure and cause it to have a different conformation from that seen in the rest of the trimeric dUTPases. However, we propose a second explanation: in structures in which ordered C-terminal arms are visible these two deleted residues form part of a short loop, the deletion of which could readily be accommodated while maintaining the conformation of the rest of the arm. A structure of one of the *B. subtilis* enzymes with an ordered C-terminal arm is needed to answer this question, but we have not yet obtained such a complex.

4.4. Implications for catalysis

There are a number of structures of complexes of trimeric dUTPases with nucleotides and nucleotide analogues. Superposition of the ligands from complexes of the dUpNHpp substrate analogue with four dUTPases, two from *B. subtilis* (YosS and YncF; both in the *gauche* forms) and the human and *M. tuberculosis* enzymes are shown in Fig. 5. The uracil, deoxyribose and phosphate groups all have very similar orientations. In addition, the divalent metals (calcium in the YncF structure and magnesium in the other three) all lie within 0.5 Å of one another. The proposed nucleophilic waters show a larger variation in position of up to about 1 Å. In all four structures the Phe-lid is folded down over the uracil; in two of them it comes from motif V and in the two *Bacillus* enzymes it comes from close to motif III. An interesting question is raised by the observation that the C-terminal extension is ordered in the human and *Mycobacterium* enzymes but remains disordered in the *Bacillus* pair.

A close inspection of the differences between the native and complexed enzymes show that there are small but significant movements upon binding of the *gauche* ligand to both YosS and YncF (Fig. 6*a*). For both *Bacillus* enzymes superposing the native enzyme with the dUpNHpp complex reveals that the subunits around the *trans*-ligand sites show essentially identical conformations (Fig. 6*b*). In contrast, there are significant internal movements in going from the native to the *gauche*-liganded subunits. As described above, the helix (carrying the conserved residues Arg61, Ser62 and Ser63) tilts with respect to the free enzyme, removing steric conflicts and allowing both the adoption of the *gauche* conformation by the ligand and the simultaneous rearrangement of the side chains of Arg61, Ser62 and Ser63. The adjacent subunit carrying the Phe-lid also undergoes a small conformational change, with half of

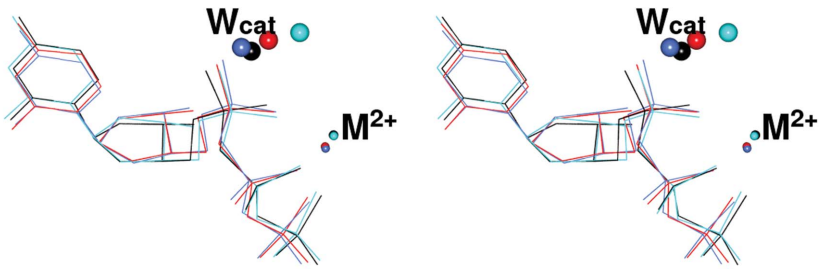


Figure 5
Stereoview of the superposition of dUpNHpp from its complexes with four trimeric dUTPases: YosS and YncF (PDB entry 2xce; García-Nafría *et al.*, 2010) from *B. subtilis* and the human (PDB entry 3ehw; E. Takács, O. Barabás & B. G. Vértessy, unpublished work) and *M. tuberculosis* (PDB entry 2py4; Varga *et al.*, 2008) enzymes. The ligand is coloured by atom type. The divalent-metal sites and the proposed nucleophilic waters are shown as spheres: YosS, cyan; YncF, black; human, light blue; *M. tuberculosis*, red.

the subunit bending towards the active site by ∼1 Å. This change is consistently seen across all three *gauche*-containing dUpNHpp ligand sites in both YncF and YosS. This is in marked contrast to the *trans* dUpNHpp sites and the dUDP sites, which all show no change compared with the free enzyme.

These coordinated changes in the YncF enzymes (which were selected since the structures have a higher resolution than the equivalent YosS structures) for the *gauche* ligands can clearly be seen in Fig. 6(*a*) and are in marked contrast to the native *trans*-ligand superposition in Fig. 6(*b*). It seems likely that these movements of the two subunits around the active sites binding *gauche* ligands are interlinked, but it is unclear which movement occurs first: the rearrangement around the phosphates on the right of Fig. 6(*a*) or that of the Phe-lid-carrying subunit on the left of the figure. In the latter, β9 which carries the Phe-lid does not move, but the antiparallel β4 moves closer to the ligand and rotates slightly with respect to β9 into a position which precludes the open native conformation of the Phe93 lid. The steric conflict is shown by the arrows in the figure, with a too close contact of 2.9 Å between Pro45 in the ligand complex and the position of Phe93 in the native structure. The equivalent distance in the *trans* complex, where there is no subunit movement, is 3.5 Å (Fig. 6*b*). This imposes the closed conformation of the Phe-lid stacked on the uracil in the *gauche* complex. This mechanism, in which the movement of the phosphates in one subunit is coordinated with the flipping of the Phe-lid in the adjacent one, may be unique to the *B. subtilis* dUTPases.

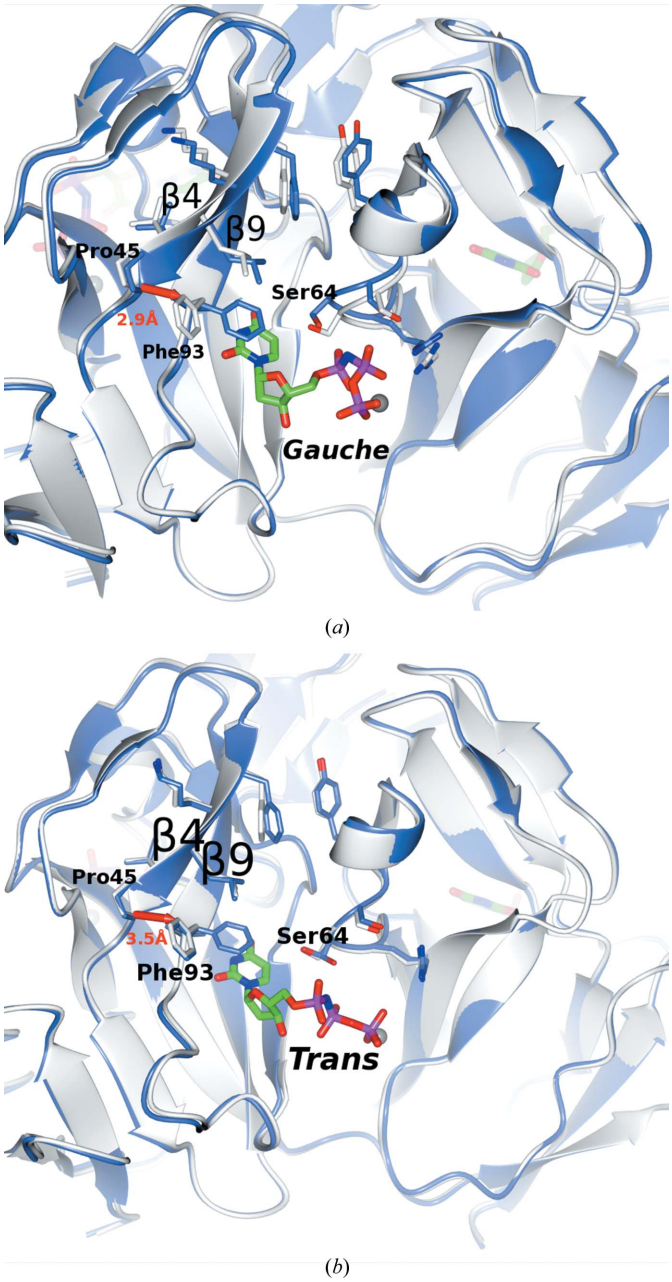


Figure 6
Conformational changes in *B. subtilis* dUTPase upon ligand binding. (*a*) An active site with a *gauche* dUpNHpp ligand (light blue) superposed on the free enzyme (white) in ribbon format. Key residues are displayed as cylinders in the same colour as the model to which they belong. The red arrow, with distance indicated, is from Pro45 C^{α} to the nearest C atom of the Phe93 side chain. There are small but clear changes in conformation induced by *gauche* ligand binding. (*b*) Equivalent view of an active site with a *trans* dUpNHpp ligand using the same colour scheme as in (*a*). There is no conformational change in the active site.

The new structures of native YosS and its complexes still leave unanswered some of the questions raised by YncF. The importance of motif V and the C-terminal extension in the *B. subtilis* enzymes seems to be different from other trimeric dUTPases. The Phe-lid residue had been proposed to play a role in expulsion of the dUMP product, which is a potential inhibitor. The positioning of the Phe-lid residue at the core of the protein and not in the flexible C-terminal extension effectively rules out such a role, at least in the *B. subtilis* dUTPases but probably in all the family. Biochemical and activity information are needed in order to characterize and provide a better understanding of the role of motif V in both *B. subtilis* dUTPases.

The catalytic water, which lies 3.7 Å from the α-phosphate, is positioned to perform an in-line nucleophilic attack. In contrast, the proposed general base Asp80 points away from the active site. A similar situation is seen in the monomeric Epstein–Barr virus dUTPase, which has its general base aspartate pointing away from the active site when the triphosphate nucleotide with associated metal is bound in the *gauche* conformation. In contrast, the aspartate points towards the active site in the dUMP complex (Tarbouriech *et al.*, 2005). It is likely that ordering of the C-terminus with motif V is required in order to correctly orient Asp80 for catalysis or there may be extra steps that are still to be identified.

As to the origin of both *B. subtilis* dUTPases, the fact that both have the Phe-lid residue in an unusual sequence position outside motif V, together with their high sequence similarity, suggests a horizontal transfer of the gene from the bacteria to the virus followed by its reinsertion as a prophage.

We should emphasize that a number of the conclusions reached in this paper are hypotheses based on the structural evidence and that future work using mutagenesis and kinetics is required for their verification. This is beyond the scope of the present study.

We thank the European Commission for funding through the SPINE2-COMPLEXES project LSHG-CT-2006-031220. We thank the staff at the Diamond Light Source (beamline I02) for provision of synchrotron facilities, Sam Hart for assistance in data collection at Diamond and John McGeehan for experimental assistance.

References

Barabás, O., Pongrácz, V., Kovári, J., Wilmanns, M. & Vértessy, B. G. (2004). *J. Biol. Chem.* **279**, 42907–42915.
Bertani, L. E., Haeggmark, A. & Reichard, P. (1963). *J. Biol. Chem.* **238**, 3407–3413.
Camacho, A., Arrebola, R., Peña-Diaz, J., Ruiz-Pérez, L. M. & González-Pacanowska, D. (1997). *Biochem. J.* **325**, 441–447.
Chan, S. *et al.* (2004). *J. Mol. Biol.* **341**, 503–517.
Cowtan, K. & Main, P. (1998). *Acta Cryst.* D**54**, 487–493.
Dauter, Z., Persson, R., Rosengren, A. M., Nyman, P. O., Wilson, K. S. & Cedergren-Zeppezauer, E. S. (1999). *J. Mol. Biol.* **285**, 655–673.
Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C. & Sacchettini, J. C. (1994). *Cell*, **78**, 325–334.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Freeman, L., Buisson, M., Tarbouriech, N., Van Der Heyden, A., Labbé, P. & Burmeister, W. P. (2009). *J. Biol. Chem.* **284**, 25280–25289.
García-Nafría, J., Burchell, L., Takezawa, M., Rzechorzek, N. J., Fogg, M. J. & Wilson, K. S. (2010). *Acta Cryst.* D**66**, 953–961.
Harkiolaki, M., Dodson, E. J., Bernier-Villamor, V., Turkenburg, J. P., González-Pacanowska, D. & Wilson, K. S. (2004). *Structure*, **12**, 41–53.
Ingraham, H. A., Dickey, L. & Goulian, M. (1986). *Biochemistry*, **25**, 3225–3230.
Kovári, J., Barabás, O., Varga, B., Békési, A., Tölgyesi, F., Fidy, J., Nagy, J. & Vértessy, B. G. (2008). *Proteins*, **71**, 308–319.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Larsson, G., Svensson, L. A. & Nyman, P. O. (1996). *Nature Struct. Biol.* **3**, 532–538.
Li, G.-L., Wang, J., Li, L.-F. & Su, X.-D. (2009). *Acta Cryst.* F**65**, 339–342.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McGeoch, D. J. (1990). *Nucleic Acids Res.* **18**, 4105–4110.
Mol, C. D., Harris, J. M., Mcintosh, E. M. & Tainer, J. A. (1996). *Structure*, **4**, 1077–1092.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Palmén, L. G., Becker, K., Bülow, L. & Kvassman, J. O. (2008). *Biochemistry*, **47**, 7863–7874.
Pecsi, I., Leveles, I., Harmat, V., Vertessy, B. G. & Toth, J. (2010). *Nucleic Acids Res.* **38**, 7179–7186.
Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). *Acta Cryst.* D**57**, 1445–1450.
Persson, R., Harkiolaki, M., McGeehan, J. & Wilson, K. S. (2001). *Acta Cryst.* D**57**, 876–878.
Persson, R., McGeehan, J. & Wilson, K. S. (2005). *Protein Expr. Purif.* **42**, 92–99.
Prasad, G. S., Stura, E. A., Elder, J. H. & Stout, C. D. (2000). *Acta Cryst.* D**56**, 1100–1109.
Prasad, G. S., Stura, E. A., McRee, D. E., Laco, G. S., Hasselkus-Light, C., Elder, J. H. & Stout, C. D. (1996). *Protein Sci.* **5**, 2429–2437.
Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J. & Pelletier, H. (1997). *Biochemistry*, **36**, 11205–11215.
Shlomai, J. & Kornberg, A. (1978). *J. Biol. Chem.* **253**, 3305–3312.
Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
Takács, E., Grolmusz, V. K. & Vértessy, B. G. (2004). *FEBS Lett.* **566**, 48–54.
Tarbouriech, N., Buisson, M., Seigneurin, J. M., Cusack, S. & Burmeister, W. P. (2005). *Structure*, **13**, 1299–1310.
Tóth, J., Varga, B., Kovács, M., Málnási-Csizmadia, A. & Vértessy, B. G. (2007). *J. Biol. Chem.* **282**, 33572–33582.
Trapani, S. & Navaza, J. (2008). *Acta Cryst.* D**64**, 11–16.
Varga, B., Barabás, O., Kovári, J., Tóth, J., Hunyadi-Gulyás, E., Klement, E., Medzihradszky, K. F., Tölgyesi, F., Fidy, J. & Vértessy, B. G. (2007). *FEBS Lett.* **581**, 4783–4788.
Varga, B., Barabás, O., Takács, E., Nagy, N., Nagy, P. & Vértessy, B. G. (2008). *Biochem. Biophys. Res. Commun.* **373**, 8–13.
Vertessy, B. G. (1997). *Proteins*, **28**, 568–579.
Vertessy, B. G., Persson, R., Rosengren, A. M., Zeppezauer, M. & Nyman, P. O. (1996). *Biochem. Biophys. Res. Commun.* **219**, 294–300.
Vértessy, B. G. & Tóth, J. (2009). *Acc. Chem. Res.* **42**, 97–106.