

# Structure and enzymatic mechanism of a moonlighting dUTPase

Ibolya Leveles,<sup>a,b</sup> Veronika Németh,<sup>a</sup> Judit E. Szabó,<sup>a</sup> Veronika Harmat,<sup>c</sup> Kinga Nyíri,<sup>a,b</sup> Ábris Ádám Bendes,<sup>a,b</sup> Veronika Papp-Kádár,<sup>a,b</sup> Imre Zagyva,<sup>a</sup> Gergely Róna,<sup>a</sup> Olivér Ozohanics,<sup>d</sup> Károly Vékey,<sup>d</sup> Judit Tóth<sup>a</sup> and Beáta G. Vértessy<sup>a,b,\*</sup>

<sup>a</sup>Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, 29 Karolina Street, 1113 Budapest, Hungary, <sup>b</sup>Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics, 4 Szent Gellért tér, 1111 Budapest, Hungary, <sup>c</sup>Hungarian Academy of Sciences–Eötvös Loránd University Protein Modeling Research Group, Institute of Chemistry, Eötvös Loránd University, Budapest, Hungary, and <sup>d</sup>Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Pusztaszeri út 59-67, 1025 Budapest, Hungary

Correspondence e-mail:  
vertessy.beata@ttk.mta.hu

Genome integrity requires well controlled cellular pools of nucleotides. dUTPases are responsible for regulating cellular dUTP levels and providing dUMP for dTTP biosynthesis. In *Staphylococcus*, phage dUTPases are also suggested to be involved in a moonlighting function regulating the expression of pathogenicity-island genes. Staphylococcal phage trimeric dUTPase sequences include a specific insertion that is not found in other organisms. Here, a 2.1 Å resolution three-dimensional structure of a  $\phi$ 11 phage dUTPase trimer with complete localization of the phage-specific insert, which folds into a small  $\beta$ -pleated mini-domain reaching out from the dUTPase core surface, is presented. The insert mini-domains jointly coordinate a single  $Mg^{2+}$  ion per trimer at the entrance to the threefold inner channel. Structural results provide an explanation for the role of Asp95, which is suggested to have functional significance in the moonlighting activity, as the metal-ion-coordinating moiety potentially involved in correct positioning of the insert. Enzyme-kinetics studies of wild-type and mutant constructs show that the insert has no major role in dUTP binding or cleavage and provide a description of the elementary steps (fast binding of substrate and release of products). In conclusion, the structural and kinetic data allow insights into both the phage-specific characteristics and the generally conserved traits of  $\phi$ 11 phage dUTPase.

Received 14 March 2013

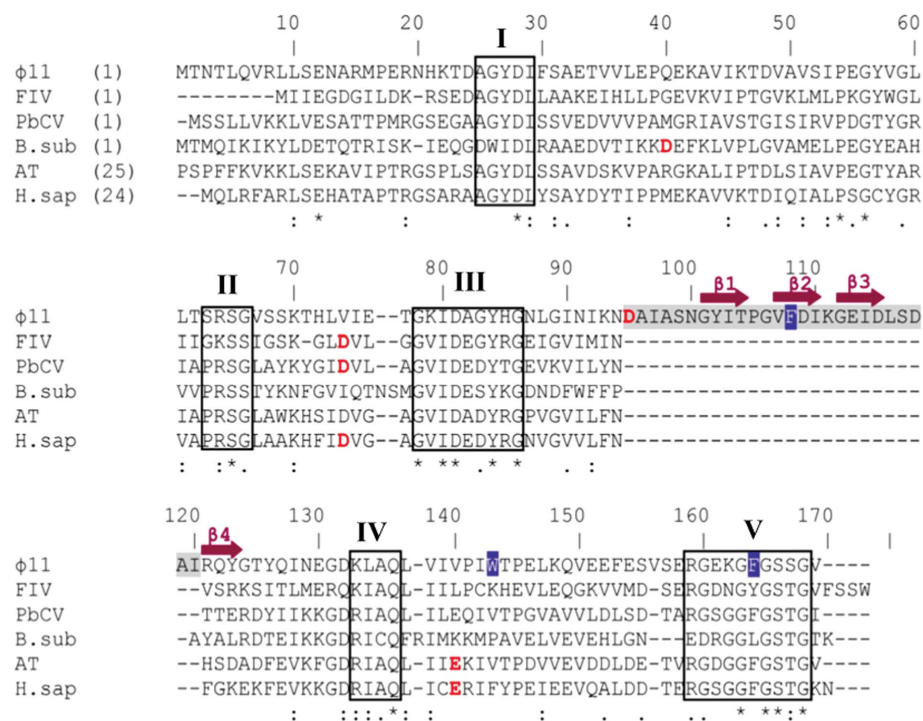
Accepted 29 July 2013

**PDB Reference:**  $\phi$ 11 phage dUTPase, 4gv8

## 1. Introduction

Genome integrity ultimately depends on dedicated mechanisms of DNA damage recognition and repair, as well as the regulation of cellular nucleotide pools to allow the synthesis of DNA with high fidelity (Langerak & Russell, 2012; Rouse & Jackson, 2002; Niida *et al.*, 2010). Although the core pathways are well conserved from bacteria to lower and higher eukaryotes, the set of protein macromolecules involved in these processes varies among different evolutionary branches, reflecting the adaptation of the organisms to cellular and environmental conditions. Viruses represent a specific group from this point of view: while importantly relying on the set of genes and proteins endogenous to the host organism, viruses also encode several key factors in their own genome that interact with the DNA-repair apparatus of the host cell (Chaurushiya & Weitzman, 2009).

A frequently occurring mistake in DNA is the presence of uracil, either from cytosine deamination or thymine-replacing uracil incorporation (Vértessy & Tóth, 2009; Castillo-Acosta *et al.*, 2012; Wilson *et al.*, 2012; Galperin *et al.*, 2006). Two enzyme families, uracil-DNA glycosylases and dUTPases, are involved in eliminating or preventing the incorporation of erroneous uracil, respectively (Doseth *et al.*, 2012; Visnes *et al.*, 2009; Pecsli *et al.*, 2012; Muha *et al.*, 2012). Several viruses encode



**Figure 1**  
Sequence alignment of trimeric dUTPases coordinating divalent metal ions within the threefold central channel:  $\Phi$ 11, *S. aureus*  $\phi$ 11 phage; FIV, *Feline immunodeficiency virus*; PbCV, *Paramecium bursaria Chlorella virus*; B. sub, *Bacillus subtilis*; AT, *Arabidopsis thaliana*; H. sap, *Homo sapiens*. Sequences were initially aligned using *ClustalW* and the alignment was manually fine-tuned based on the three-dimensional structures. The metal-ion-coordinating residues are shown in red. Conserved motifs I–V are boxed. The residues used in fluorescence experiments are shown in blue.  $\beta$ -Strands of the phage-specific insert are indicated as purple arrows.

representatives of both of these families, and it has been shown that viral genes for these enzymes have significant roles in the efficiency of the virus in invading host cells, especially when the expression of the respective host genes is down-regulated (*e.g.* in resting or differentiated cells; Payne & Elder, 2001). The dUTPase enzyme family constitutes trimeric, dimeric and monomeric enzymes, where the trimeric and monomeric representatives both show a well conserved  $\beta$ -pleated jelly-roll fold and five conserved motifs involved in enzyme action (Vértessy & Tóth, 2009; Persson *et al.*, 2001; Freeman *et al.*, 2009; Tarbouriech *et al.*, 2005). Dimeric dUTPases belong to another protein family with no structural similarities (Moroz *et al.*, 2004; Harkiolaki *et al.*, 2004). Interestingly, annotated genome databases of staphylococcal phages report that phages may encode dUTPase genes either from the trimeric or the dimeric family, depending on the specific phage. It was recently shown that staphylococcal phage dUTPases may be responsible for inducing horizontal gene transfer of the mobile genetic elements termed staphylococcal pathogenicity islands (SAPIs; Tormo-Más *et al.*, 2010). SAPIs are encoded within the bacterial genome and are under the regulation of repressor proteins (StIs) specific for the different SAPIs (Novick & Subedi, 2007; Ruzin *et al.*, 2001, Tormo *et al.*, 2008). As such, the StI repressor of SAPIbov1 down-regulates the expression of the proteins responsible for SAPIbov1 mobilization. Helper phages are

already known to relieve this repression and to induce the excision and replication of SAPIs (Maiques *et al.*, 2007). The dUTPases encoded within the helper phages have been implicated in this task (Tormo-Más *et al.*, 2010). An in-depth structural and functional characterization of phage dUTPases is indispensable to understand this biological phenomenon of de-repression, which is also important from a biomedical point of view since SAPIs encode virulence factors and toxins and their spreading among different strains is a major biomedical challenge (Novick & Subedi, 2007). A recent study reported three-dimensional structures of the  $80\alpha$  phage dUTPase and its complexes at 2.8–3.1 Å resolution. These structures provided reliable information on the trimeric protein fold but have not yet allowed full insight into protein–ligand interactions (Tormo-Más *et al.*, 2013). It has also been reported that the  $\phi$ 11 phage dUTPase (see Fig. 1 and Supplementary Fig. S2<sup>1</sup> for sequence alignments) is much more efficient in de-repression compared with the  $80\alpha$  phage dUTPase, making it an interesting target for detailed investigations (Tormo-Más *et al.*, 2010, 2013).

In the present work, we describe the structure of the  $\phi$ 11 phage trimeric dUTPase at 2.1 Å resolution using X-ray crystallography and also characterize its enzyme-kinetic cycle. We analyze the phage-specific structural characteristics, compare it with the  $80\alpha$  phage dUTPase structure and show that the phage-specific insert forms four short  $\beta$ -strands organised in an antiparallel fashion. The orientation of this mini-domain is aided by coordination of a  $Mg^{2+}$  ion, accommodated by one aspartate residue from the insert in each of the three subunits. Using fluorescence spectroscopy with different site-specific tryptophan probes, as well as steady-state and transient kinetics, we provide a description of the catalytic cycle of the phage dUTPase.

## 2. Materials and methods

### 2.1. Cloning, mutagenesis, expression and purification of recombinant proteins

The cDNA of wild-type  $\phi$ 11 phage dUTPase was amplified from an expression vector previously described in Leveles *et al.* (2011) and was cloned into the *EcoRI/XhoI* restriction sites of the pGEX-4T-1 vector in frame with the N-terminal

<sup>1</sup> Supplementary material has been deposited in the IUCr electronic archive (Reference: DW5050). Services for accessing this material are described at the back of the journal.

glutathione-S-transferase tag, as well as with a thrombin cleavage site and a 6×His tag between the glutathione-S-transferase tag and the dUTPase sequence. This fusion construct was expressed using the *Escherichia coli* BL21 (DE3) Rosetta expression system. 0.5 l LB medium was inoculated with Rosetta cells from a 5 ml overnight culture and grown at 310 K until the OD<sub>600</sub> reached 0.4. After 30 min cooling at 303 K, protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside for 4 h at 303 K. The cells were harvested by centrifugation at 4000g for 20 min and stored at 193 K. Thawed cells were resuspended in 15 ml 1× PBS buffer pH 7.3 containing 200 mM NaCl, 5 mM MgCl<sub>2</sub>, 1% Triton X-100, 3 μg ml<sup>-1</sup> RNase, 3 μg ml<sup>-1</sup> DNase and EDTA-free cOMplete ULTRA protease-inhibitor preparation (Roche, Switzerland). The solution was sonicated and centrifuged and the supernatant was applied onto a pre-equilibrated benchtop glutathione-agarose affinity-chromatography column. After washing the column with ten volumes of buffer, the dUTPase was eluted using overnight thrombin cleavage to remove the glutathione-S-transferase tag, resulting in protein preparations of >95% homogeneity as judged by SDS-PAGE. This construct contained a GSPEFHSHHHHGS N-terminal extension preceding the original dUTPase sequence. We checked that this extension did not change the dUTPase enzymatic activity compared with the previously investigated tag-free dUTPase (Leveles *et al.*, 2011) and will use the abbreviation φ11DUT<sup>WT</sup> to refer to it in this study. The molecular mass of the φ11DUT<sup>WT</sup> protein is 19 853 Da as calculated from the amino-acid sequence. Throughout this study, the numbering of the dUTPase sequence is according to the original cDNA (UniProt ID Q8SDV3), starting with the physiologically encoded initial methionine.

The following mutants were created by QuikChange site-directed mutagenesis (Stratagene) using suitable mutagenic primers (listed in Supplementary Table S1): φ11DUT<sup>F108W</sup>, φ11DUT<sup>F164W</sup>, φ11DUT<sup>E158STOP</sup> and φ11DUT<sup>Δ101–122</sup>. All constructs were verified by DNA sequencing of both strands. Mutants were expressed and purified as for the wild-type proteins. All protein preparations were >95% pure as judged by SDS-PAGE.

Protein concentrations were determined using Bradford's assay or by UV spectrometry using A<sub>280</sub><sup>0.1%</sup> values of 0.79, 1.08, 1.08, 0.84 and 0.81 ml mg<sup>-1</sup> cm<sup>-1</sup> for the wild type and the φ11DUT<sup>F108W</sup>, φ11DUT<sup>F164W</sup>, φ11DUT<sup>E158STOP</sup> and φ11DUT<sup>Δ101–122</sup> mutants, respectively, as calculated from the amino-acid sequence. Concentrations are given in subunits throughout the text (molecular-mass values were calculated based on amino-acid sequence).

## 2.2. Crystal structure determination

Crystals of φ11 phage dUTPase in the presence of 1 mM α,β-imido-dUTP and 5 mM Mg<sup>2+</sup> were grown by the hanging-drop vapour-diffusion method at 293 K as described previously (Leveles *et al.*, 2011). A diffraction data set was collected from a single crystal in a nitrogen cryostream at 100 K on the PXIII beamline of the Swiss Light Source,

**Table 1**

Crystallographic data-collection and refinement statistics.

Values in parentheses are for the outer shell.

Data collection	
Space group	P4 <sub>3</sub> 22
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 109.66, <i>c</i> = 169.27
Resolution range (Å)	47.10–2.10 (2.15–2.10)
Total No. of reflections	813098 (56753)
No. of unique reflections	60917 (4440)
Multiplicity	13.3 (12.8)
Completeness (%)	100.0 (100.0)
<i>I</i> (σ( <i>I</i> ))	34.37 (3.07)
<i>R</i> <sub>meas</sub> <sup>†</sup>	0.058 (0.978)
Refinement	
No. of dUTPase subunits in asymmetric unit	6
No. of protein atoms	6800
No. of ligand atoms	168
No. of waters	376
No. of Mg <sup>2+</sup> ions	8
<i>R</i> <sub>cryst</sub> / <i>R</i> <sub>free</sub> <sup>‡</sup>	0.186/0.232
Luzzati plot coordinate error (Å)	0.286
Average <i>B</i> factors (Å <sup>2</sup> )	
Wilson <i>B</i> factor	47.78
Protein atoms	62.64
Ligand atoms	56.42
Water	54.26
Mg <sup>2+</sup> ions	51.40
R.m.s. deviations from ideal values	
Bond lengths (Å)	0.017
Bond angles (°)	1.89
Ramachandran plot analysis§, residues in (%)	
Favoured region	98.01
Allowed region	1.88
Disallowed region	0.11

<sup>†</sup>  $R_{meas} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - (I(hkl))| / \sum_{hkl} \sum_i I_i(hkl)$ . <sup>‡</sup>  $R_{cryst} = \sum_{hkl} (|F_{obs}| - |F_{calc}|) / \sum_{hkl} |F_{obs}|$ , where  $F_{obs}$  and  $F_{calc}$  are the observed and calculated structure-factor amplitudes, respectively.  $R_{free}$  is equivalent to  $R_{work}$  but calculated using 5% of the total reflections which were chosen randomly in thin resolution shells and omitted from refinement for cross-validation purposes. § Ramachandran plot analysis was carried out using *MolProbity* (Chen *et al.*, 2010).

Villigen, Switzerland (PILATUS 2M detector, wavelength 1.00 Å, φ range of 180° with an increment of 0.1°; crystal-to-detector distance 210 mm). Data were processed to 2.1 Å resolution using the *XDS* and *XSCALE* programs (Kabsch, 2010). The data-collection statistics are summarized in Table 1.

The structure was solved by molecular replacement using *MOLREP* v.10.2.35 (Vagin & Teplyakov, 2010) from the *CCP4* package (Winn *et al.*, 2011). The structure of mycobacterial dUTPase (PDB entry 3hza; Pecsí *et al.*, 2010) was used as a model for molecular replacement for a different crystal form of φ11 dUTPase (Leveles *et al.*, 2011; space group *I*23, resolution 2.98 Å, one dUTPase subunit in the asymmetric unit); however, the insertion loop could not be unambiguously built into electron density. The crystallographic dUTPase trimer generated from this partially built low-resolution structure was used as a model for molecular replacement of the present structure. The asymmetric unit contains six subunits (two dUTPase trimers); the Matthews coefficient was 2.31 Å<sup>3</sup> Da<sup>-1</sup>.

Model building was carried out with *Coot* v.0.6.2 (Emsley *et al.*, 2010) using graphical comparison of noncrystallographically related subunits (NCS ghost control option). The model was refined by likelihood-based refinement using *REFMAC* v.5.6.0117 (Murshudov *et al.*, 2011), in which TLS refinement (Winn *et al.*, 2001) was carried out for quaternary-

structure regions (six TLS groups each consisting of core residues of one subunit, the C-terminal arm region of the next subunit interacting with these and  $\text{Mg}^{2+}$ - $\alpha,\beta$ -imido-dUTP bound at the active site); noncrystallographic restraints were introduced for regions with similar conformations of each subunit within both trimers (residues 4–19, 24–82, 85–97, 101–154 and  $\text{Mg}^{2+}$ - $\alpha,\beta$ -imido-dUTP). Water molecules were included in the model by the water-picking mode of *Coot* (within hydrogen-bonding distance of hydrogen-bond donor atoms,  $mF_o - DF_c$  density greater than  $3.0\sigma$ ; checked for refined *B* factors of less than  $80 \text{ \AA}^2$ ).

In this model, the phage dUTPase-specific insert could be nicely located. However, the N-terminal His tag and the C-terminal 15 residues could not be located in the electron-density map, with the sole exception of residue Phe164, which makes a close interaction with the uracil ring of the nucleoside triphosphate ligand. Flexibility of the His tag is a commonly observed phenomenon, while flexible behaviour of the dUTPase C-terminal arm containing the conserved motif V is also frequently observed even in the presence of bound dUTP substrate or its slowly hydrolyzing analogue  $\alpha,\beta$ -imido-dUTP. In addition to protein atoms, the substrate analogue  $\alpha,\beta$ -imido-dUTP was readily observed in the maps, together with a  $\text{Mg}^{2+}$  ion coordinating to the phosphate chain of the nucleoside triphosphate; however, the phosphate chain and coordinated  $\text{Mg}^{2+}$  ion have higher *B* factors than the deoxyuridine moiety. An additional  $\text{Mg}^{2+}$  ion was also located within the central channel of each trimer.

The final model comprises two dUTPase trimers with  $\alpha,\beta$ -imido-dUTP bound at each active site. A total of 311 water molecules were refined in the final structure. Data-collection statistics and model quality are compiled in Table 1. Structural model figures were prepared using *PyMOL* (<http://www.pymol.org>). The coordinates and structure factors have been deposited in the Protein Data Bank under accession code 4gv8.

### 2.3. Mass spectrometry

Samples of  $\phi$ 11 phage dUTPase were analyzed on a Waters Q-ToF Premier-type mass spectrometer in positive electrospray ionization mode. Mass spectra were obtained under native conditions using protein solution containing  $\phi$ 11 phage dUTPase at a concentration of  $60 \mu\text{M}$  in  $10 \text{ mM}$  ammonium hydrogen carbonate buffer pH 7.8. This allows transfer of the oligomeric species present in the solution into the gas phase (Benesch *et al.*, 2007; Grandori *et al.*, 2009). The capillary voltage was 2800 V, the sampling-cone voltage was 128 V and the temperature of the source was kept at 363 K. Mass spectra were recorded in the mass range 1500–8000 *m/z*.

### 2.4. Steady-state kinetics experiments

Proton release during the transformation of dUTP into dUMP and  $\text{PP}_i$  was followed at 559 nm at 293 K using a Jasco V550 spectrophotometer (Takács *et al.*, 2010; Varga *et al.*, 2007). Reaction mixtures contained  $80 \text{ nM}$  enzyme in  $1 \text{ mM}$  HEPES–HCl pH 7.5 containing  $40 \mu\text{M}$  dUTP,  $5 \text{ mM}$   $\text{MgCl}_2$ ,

$150 \text{ mM}$  KCl,  $40 \mu\text{M}$  phenol red pH indicator. The initial velocity was determined from the slope of the first 10 s of the progress curve.

### 2.5. Fluorescence measurements

Fluorescence spectra and intensity titrations were recorded on a Jobin Yvon Spex FluoroMax-3 spectrofluorometer in  $20 \text{ mM}$  HEPES pH 7.5 buffer containing  $1 \text{ mM}$   $\text{MgCl}_2$ ,  $300 \text{ mM}$  NaCl and  $10 \text{ mM}$   $\beta$ -mercaptoethanol at 293 K. Tryptophan residues were excited at 295 nm. Emission spectra were recorded between 320 and 400 nm. The excitation and emission slits were 1 and 5 nm, respectively. Additional fluorescence or inner filter effects imposed on the measured intensities during titration experiments were corrected by subtracting the intensity of the assay buffer. Titration data were recorded by using a  $5 \mu\text{M}$  constant protein concentration and varying the  $\alpha,\beta$ -imido-dUTP concentration. Results were fitted to the quadratic binding equation describing 1:1 stoichiometry for the dissociation equilibrium with no cooperativity,

$$y = s + \frac{A\{(c + x + K) - [(c + x + K)^2 - 4cx]^{1/2}\}}{2c}, \quad (1)$$

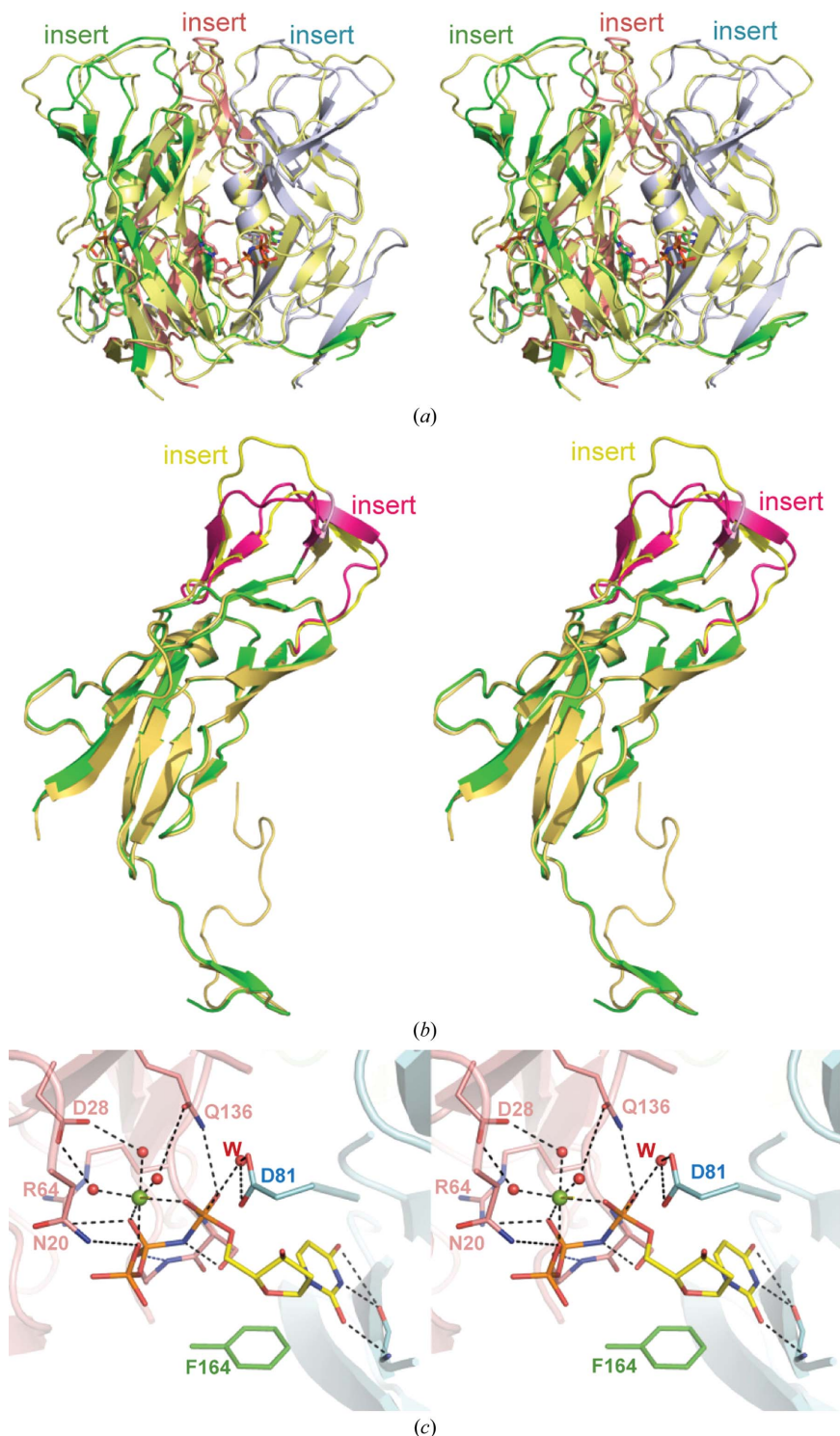
where *x* is the ligand concentration and *y* is the fluorescence intensity,  $s = y$  at  $x = 0$ , *A* is the amplitude of the fluorescence intensity change, *c* is the enzyme concentration and *K* is the dissociation constant ( $K_d$ ; as before; Takács *et al.*, 2010; Varga *et al.*, 2008; Tóth *et al.*, 2007).

### 2.6. Transient kinetics experiments

Stopped-flow measurements were carried out using an SX-20 (Applied Photophysics, UK) stopped-flow apparatus as described previously (Tóth *et al.*, 2007). Trp fluorescence was excited at 295 nm and emission was selected with a 320 nm longpass filter. Equal volumes ( $50 \mu\text{l}$ ) of dUTPase enzyme and dUTP solutions in  $20 \text{ mM}$  HEPES pH 7.5 buffer containing  $1 \text{ mM}$   $\text{MgCl}_2$ ,  $300 \text{ mM}$  NaCl,  $10 \text{ mM}$   $\beta$ -mercaptoethanol and  $0.1 \text{ mM}$  PMSF were mixed. For active-site titration the curves were measured upon mixing various concentrations of dUTP with  $20 \mu\text{M}$  enzyme (post-mixing concentration). 4–5 traces were recorded for each data point. The detector sensitivity was set to high voltage = 357 during measurement. For dUTP binding the curves were measured upon mixing various concentrations of dUTP with  $0.5 \mu\text{M}$  enzyme (post-mixing concentration) and 7–8 traces were recorded. The detector sensitivity was set to high voltage = 460 during measurement. All measurements were performed at 293 K. Time courses were analyzed by *Origin 7.5* (OriginLab, Northampton, Massachusetts, USA).

### 2.7. Phylogenetic studies

A total of 61 sequences of *Staphylococcus aureus* phage dUTPases were retrieved from the GenBank database with the homology *BLAST* search and using GenBank accession numbers extracted from the literature (Kahánková *et al.*, 2010). Global alignments were carried out using *ClustalW* with



**Figure 2**  
 Three-dimensional structural characteristics of  $\phi 11$  phage dUTPase (stereoview). (a) Overall three-dimensional structure of the protein as a ribbon model of colour-coded subunits (salmon, blue and green) with ligand molecules shown as ball-and-stick models (yellow C atoms and atomic colouring otherwise). The  $80\alpha$  phage dUTPase structure (PDB entry 3zez; Tormo-Más *et al.*, 2013; ribbon model in yellow) is superimposed in order to highlight the position of the phage-specific insert. (b) Ribbon model of a subunit with the phage-specific insert highlighted in purple, superimposed on the respective subunit of  $80\alpha$  phage dUTPase (PDB entry 3zez). The four  $\beta$ -strands of the insert mini-domain are annotated in Fig. 1. (c) Active-site close-up with the catalytically important residues and the  $Mg^{2+}$ - $\alpha,\beta$ -imido-dUTP ligand shown as ball-and-stick models, while the rest of the protein is presented as a ribbon model. Mg is shown in green, ligand C atoms in yellow, protein C atoms according to subunit colours and other atoms with atomic colouring. Hydrogen-bonding interactions are shown as dashed black lines.

automatic penalty preferences. Two distinct supergroups were identified as corresponding to trimeric and dimeric dUTPases, respectively (Galperin *et al.*, 2006; Moroz *et al.*, 2004, 2005; Vértessy & Tóth, 2009). The groups were confirmed with *BLAST* searches against known dUTPase group members. 35 homotrimeric dUTPases were selected and further aligned with *ClustalX* 2.1 using modified penalty preferences (gap opening, 50; gap extension, 1) and using the BLOSUM weighting matrix. The phylogenetic tree was generated using *ClustalX* 2.1 with the neighbour-joining algorithm (Brinkman & Leipe, 2001). Bootstrap calculations, iterated 1000 times, were also made. Tree building was carried out with *FigTree* v.1.3.1. The dUTPases were grouped into six subgroups based on the tree topology and the bootstrap values (Hillis & Bull, 1993). The groups were colour-coded for better recognition.

### 3. Results and discussion

#### 3.1. Oligomerization and three-dimensional structure of $\phi 11$ phage dUTPase

Fig. 1 shows an alignment of selected dUTPases representing major evolutionary groups of viruses as well as prokaryotes and eukaryotes. The specific insert segment of over 20 residues present in  $\phi 11$  phage dUTPase as well as in other staphylococcal phage dUTPases (Leveles *et al.*, 2011; Tormo-Más *et al.*, 2010, 2013) is situated between the dUTPase conserved motifs III and IV. Disregarding this insert, the  $\phi 11$  phage dUTPase sequence corresponds to the usual well conserved sequence characteristics of dUTPases with regard to both the presence of the conserved motifs I–V and the spacing between these motifs (Vértessy & Tóth, 2009).

To investigate the three-dimensional structure of this protein, we used X-ray crystallography on protein crystals grown in the presence of  $\alpha,\beta$ -imido-dUTP, a well characterized isosteric and slowly hydrolysable substrate analogue, as well as  $Mg^{2+}$  ion, a cofactor that accommodates the nucleoside triphosphate substrate bound in the dUTPase active site (Kovári *et al.*, 2008; Barabás *et al.*, 2004; García-Nafaría *et al.*, 2011; Chan *et al.*, 2004). The structure was determined to 2.1 Å resolution from single

crystals using synchrotron radiation, also allowing insights into details of side-chain and ligand conformations.

The structural model built based on the crystallographic data is shown in Fig. 2. The structure shows the characteristic trimeric dUTPase antiparallel  $\beta$ -pleated fold of the subunits (Fig. 2*a*; see also Supplementary Fig. S1). The three subunits attach to each other with the same characteristics as in other trimeric dUTPases, forming three active sites at the subunit interfaces. The fold of the trimeric  $\phi$ 11 phage dUTPase is very similar to that of the 80 $\alpha$  phage dUTPase, as shown by

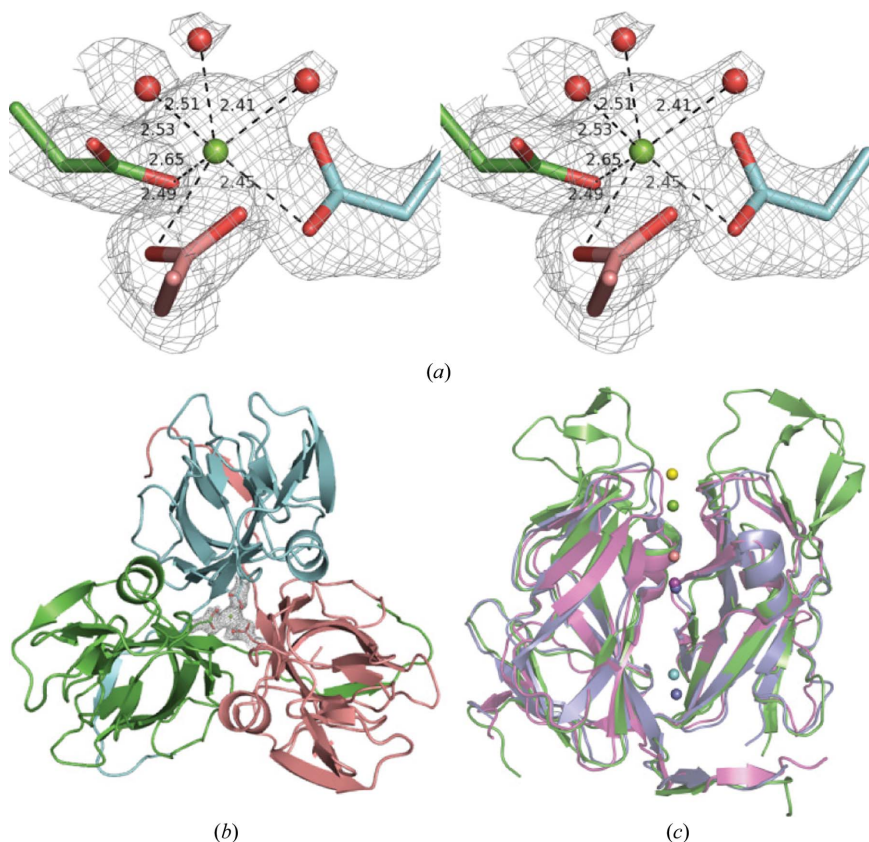
superimposition (Figs. 2*a* and 2*b*); the only minor difference is observed in the orientation of one of the loops of the phage-specific insert.

The phage-specific insert is clearly observed in the density maps as a segment that forms a characteristic  $\beta$ -pleated arrangement looping out from the surface of the trimer (see Fig. 2). The insert consists of four short  $\beta$ -strands ( $\beta$ 1– $\beta$ 4; Fig. 1), arranged in pairs of two, constituting an antiparallel organization of a distorted Greek key. The  $\beta$ 2 strand of this  $\beta$ -pleated ‘mini-domain’ also contacts a short  $\beta$ -strand of the conserved dUTPase core fold (Fig. 2*b*). This connection presumably helps to orient the phage-specific mini-domain.

The  $\text{Mg}^{2+}$ - $\alpha,\beta$ -imido-dUTP ligand is accommodated within the active site (Fig. 2*c*) by making a set of contacts to the conserved side chains of the different motifs in a manner highly similar to those found in human and mycobacterial dUTPases (Chan *et al.*, 2004; Varga *et al.*, 2007, 2008; Mol *et al.*, 1996). The active-site Asp81 coordinates a water molecule posed for in-line attack at the  $\alpha$ -P atom of  $\alpha,\beta$ -imido-dUTP. The residues of the C-terminal arm of the enzyme, which contains the conserved motif V, are mostly missing in the electron-density map; however, one side chain (Phe164) could be built in the map. A flexible character of motif V is a common observation in numerous trimeric dUTPases, preventing its localization even in the presence of the nucleoside triphosphate ligand in many cases (Barabás *et al.*, 2004, 2006; García-Nafria *et al.*, 2010, 2011; Samal *et al.*, 2007; Németh-Pongrácz *et al.*, 2007). The fact that the Phe164 side chain is still observable is most possibly owing to the fact that it creates an aromatic stacking interaction with the uracil ring of the ligand (Pecsi *et al.*, 2010; Mol *et al.*, 1996) and hence its conformation is stabilized.

### 3.2. The phage-specific insert coordinates a $\text{Mg}^{2+}$ ion and does not affect dUTPase enzymatic activity

With the exception of the  $\beta$ -stranded connection between strand  $\beta$ 2 of the mini-domain and a short  $\beta$ -strand of the dUTPase fold, there is no other contact between the insert and the other parts of the protein in the presently observed conformation. However, the tip of the mini-domain approaches the dUTPase surface and, depending on the specific side-chain conformations of some surface-oriented flexible residues, some contacts can be



**Figure 3**

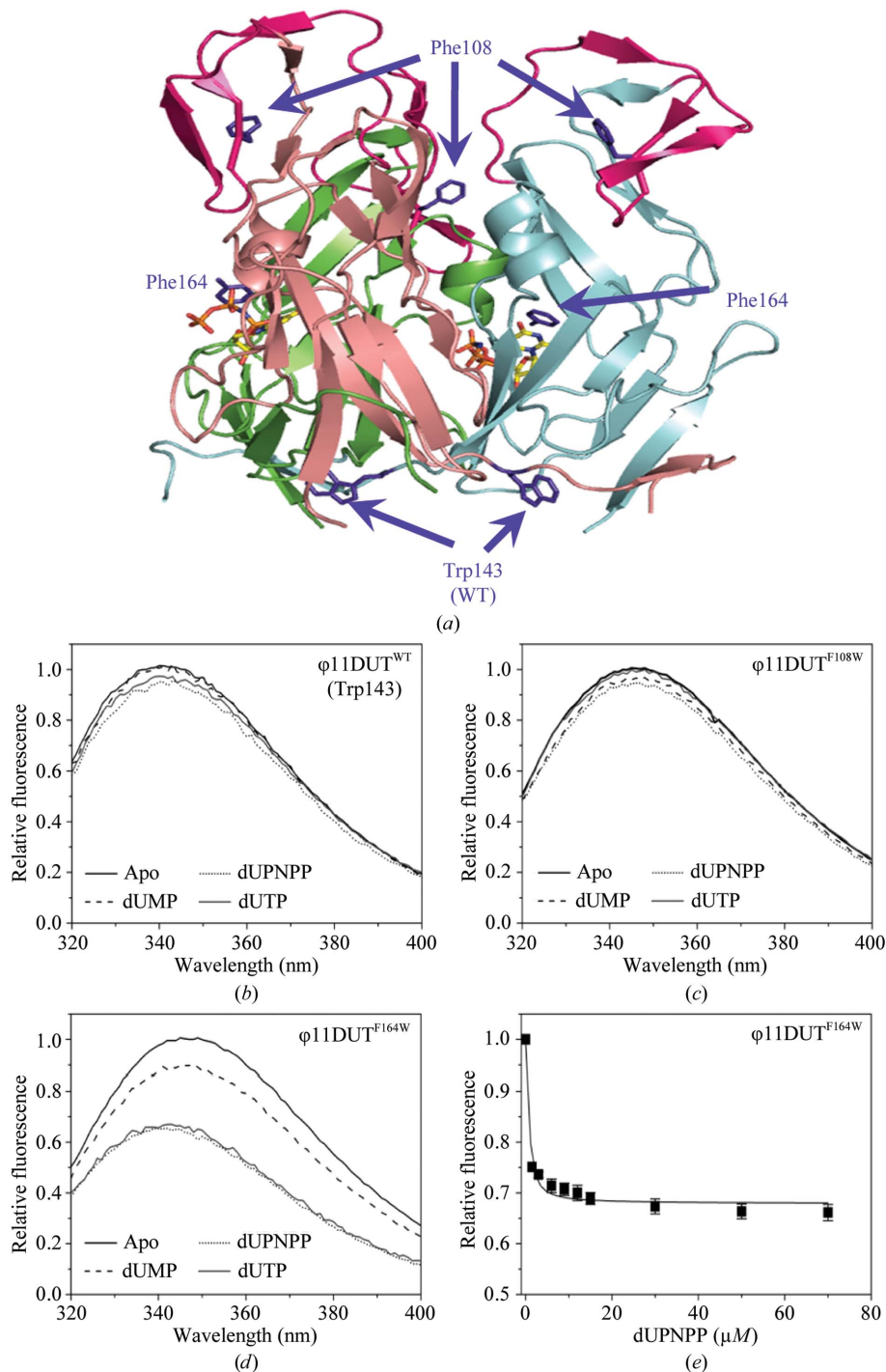
Coordination of  $\text{Mg}^{2+}$  within the central channel of the dUTPase trimer. (a) OMIT map around the  $\text{Mg}^{2+}$  ion in the threefold channel of  $\phi$ 11 phage dUTPase (stereoview). The  $\text{Mg}^{2+}$  ion, the coordinating side chains and waters are shown as ball-and-stick models. The  $\text{Mg}^{2+}$  ion is shown in green, the C atoms of the three aspartic acids are coloured according to subunit colours and other atoms and waters are shown in atomic colouring. Hydrogen-bonding interactions are shown as dashed black lines. The  $2F_o - F_c$  composite map is restricted to show the  $\text{Mg}^{2+}$  ion, the three Asp95 side chains and the three water molecules coordinating to the metal ion. (b) Bottom-up image of the trimeric  $\phi$ 11 phage dUTPase, highlighting the position of the  $\text{Mg}^{2+}$  ion in the inner channel coordinated by the first amino acid (Asp95) of the phage-specific insert. The protein is shown as a ribbon model and the coordinating side chains and the  $\text{Mg}^{2+}$  ion are shown as ball-and-stick models. Protein chain and C atoms are coloured according to subunit colours, aspartic acid O atoms are shown in red and the  $\text{Mg}^{2+}$  ion is shown in green. The  $2F_o - F_c$  simulated-annealed composite OMIT map around the  $\text{Mg}^{2+}$  ion is also shown. (c) Superimposed structural view of the inner-channel location of the  $\text{Mg}^{2+}$  ion ligand in dUTPases from different organisms, namely (from top to bottom) *Bacillus subtilis* (yellow; PDB entry 2xcd; García-Nafria *et al.*, 2010); *S. aureus*  $\phi$ 11 phage (green; PDB entry 4gv8; this work); *Paramecium bursaria Chlorella virus* (salmon; PDB entry 3c2t; Homma & Moriyama, 2009); *Feline immunodeficiency virus* (magenta; PDB entry 1dut; Prasad *et al.*, 1996); *Arabidopsis thaliana* (cyan; PDB entry 2p9o; M. Bajaj & H. Moriyama, unpublished work); and human (dark purple; PDB entries 3ehw and 3arn; two  $\text{Mg}^{2+}$  sites; E. Takács, O. Barabás & B. G. Vértessy, unpublished work; Miyahara *et al.*, 2012). Proteins are shown as ribbon models, while  $\text{Mg}^{2+}$  ions are shown as balls. Only three different subunit chains of the six superimposed trimeric protein structures are shown for better visualization (the metal-ion-coordinating side chains are coloured in red in the sequence alignment in Fig. 1).

made within this structural context [e.g. between residues Asp111 or Lys113 of the insert and Lys70, a conserved basic residue in most dUTPases, or between Ile112 of the insert and the C-terminal conserved dUTPase motif V (residue Val169)].

At the N-terminal segment of the mini-domain insert, the three Asp95 residues (shown in red in the alignment in Fig. 1), one from each subunit, participate in coordinating an additional  $Mg^{2+}$  ion per trimer. The electron density for the metal ion and its coordinating atoms are clearly localized within the

threefold central channel of the dUTPase trimer (Figs. 3a and 3b). In addition to the three Asp95 carboxyl groups, three water molecules contribute to the hexagonal coordination sphere around the metal ion, in agreement with the geometry usually reported for  $Mg^{2+}$  sites. It is of interest to mention that in the recently published 80 $\alpha$  phage dUTPase structure a  $Ni^{2+}$  ion was tentatively localized at this site (with 0.33 occupancy and a high *B* factor; Tormo-Más *et al.*, 2013).

Localization of a  $Mg^{2+}$  ion within the threefold central channel of trimeric dUTPases has already been reported in enzymes from several sources (Prasad *et al.*, 2000; Takács *et al.*, 2009; Mol *et al.*, 1996; García-Nafría *et al.*, 2010). It is also known that, for example, in human dUTPases the  $Mg^{2+}$  ion within the central channel may contribute to structural stability (Kovári *et al.*, 2004; Varga *et al.*, 2007). In Fig. 1, we aligned all of the dUTPase sequences in which such divalent metal-ion sites have been reported and indicated those glutamate or aspartate residues that were implicated in coordinating the  $Mg^{2+}$  ion in the respective three-dimensional structures. We constructed a structural alignment and present a superimposed structural view of the respective location of these  $Mg^{2+}$  sites in the different enzymes (Fig. 3c). In the presently investigated  $\phi$ 11 phage dUTPase, the Asp95 residue that coordinates the metal ion is the first residue of the phage-specific insert, hence this site is specific to the phage dUTPase. We also found that this residue is present in all staphylococcal phage dUTPases that possess the characteristic dUTPase motifs I–V (Supplementary Figure S1). Interestingly, this metal-ion site is located at the beginning of the outreaching phage-specific mini-domain insertion. We propose that metal binding may stabilize the conformation of the insertions and may therefore play a role in insert-specific function. Accordingly, it has been reported that a



**Figure 4** Active-site tryptophan labelling allows the detection of substrate and product binding to  $\phi$ 11 phage dUTPase. (a) The sites of the three aromatic residues serving as fluorophores in the present study. (b), (c) and (d) show the fluorescence spectra of the  $\phi$ 11DUT<sup>WT</sup>,  $\phi$ 11DUT<sup>F108W</sup> and  $\phi$ 11DUT<sup>F164W</sup> enzymes, respectively, in the absence of ligands or in the presence of 1 mM dUMP, 100  $\mu$ M  $\alpha,\beta$ -imido-dUTP or 500  $\mu$ M dUTP (saturating concentrations). (e) Fluorescence intensity titration upon  $\alpha,\beta$ -imido-dUTP binding to the  $\phi$ 11DUT<sup>F164W</sup> construct. Data were fitted using a quadratic equation (see equation 1) to yield  $K_d = 0.32 \mu$ M. Error bars show standard deviations ( $n = 2$ ).

staphylococcal phage dUTPase carrying a glutamate mutation at this position failed to induce the pathogenicity island (Tormo-Más *et al.*, 2010). Our present three-dimensional structure suggests a plausible explanation for this finding: a glutamate at this position cannot produce the same geometry since it possesses an additional  $-\text{CH}_2-$  moiety compared with the aspartate and would therefore be sterically incompetent at this site.

Elimination of the phage-specific insert from the phage  $\phi 11$  dUTPase was stated to result in a catalytically inactive mutant (Tormo-Más *et al.*, 2010). Our presently determined three-dimensional structure in the crystal phase visualizes this insert region as a  $\beta$ -pleated mini-domain within the vicinity of the active site but without any detectable interaction with the active site and with no detectable major interference with the folding of the dUTPase trimer. Since static crystal structures inherently do not provide information concerning flexible movements of protein segments, it was important to design solution-phase experiments.

Tryptophan fluorescence is a frequently used sensitive indicator of protein conformation and conformational changes that occur upon ligand binding. Previously, it has been shown that exchange of the conserved aromatic residue (Phe or His) within motif V to Trp can adequately report on the different ligand-bound states during the dUTPase reaction cycle (Pécsi *et al.*, 2011; Varga *et al.*, 2007; Tóth *et al.*, 2007). It has been widely reported that motif V undergoes a shift from a flexible to a more ordered conformation during the dUTPase catalytic cycle as it flips over the substrate to contribute to formation of the catalytically competent complex (Vertessy *et al.*, 1998; Vertessy, 1997; Mol *et al.*, 1996; Chan *et al.*, 2004). The characteristic changes in the motif V Trp sensor fluorescence can be quantitatively followed during binding of the substrate analogue  $\alpha,\beta$ -imido-dUTP or the product dUMP. It was also shown that a Trp residue that does not interact directly with the substrate but is in the vicinity of motif V similarly shows characteristic changes in its fluorescence parameters during dUTP hydrolysis catalyzed by dUTPase (Lopata *et al.*, submitted), most probably by sensing the conformational change of motif V. Based on these observations, we created the F108W mutation ( $\phi 11\text{DUT}^{\text{F108W}}$ ) to introduce a Trp residue into the phage-specific insert (Figs. 1 and 4a).

The steady-state activity of this mutant did not change compared with the wild type ( $\phi 11\text{DUT}^{\text{WT}}$ ), as shown in Table 2. We performed fluorescence measurements to determine whether binding of the substrate analogue  $\alpha,\beta$ -imido-dUTP or the product dUMP induces any fluorescence spectral changes in the different Trp mutants. Wild-type  $\phi 11$  phage dUTPase also contains a Trp between motifs IV and V (Trp143; Figs. 1 and 4a). We first tested whether the fluorescence of this Trp143 changes upon substrate or product binding. This tryptophan residue is not part of the active site and does not contact motif V; therefore it may not sense ligand binding and can be used as a negative control. As expected, none of the ligands induced a significant intensity change or shift in the  $\lambda_{\text{max}}$  upon ligand binding (Fig. 4b, Supplementary Table S2). Next we tested the  $\phi 11\text{DUT}^{\text{F108W}}$  mutant that

**Table 2**

Steady-state catalytic properties ( $k_{\text{cat}}$  and  $K_{\text{m}}$ ) determined for wild-type and mutant  $\phi 11$  dUTPase proteins.

Data were measured by the steady-state assay proton-release assay and represent averages and errors of 5–8 parallel measurements. ND, not determined.

Enzyme construct	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ ( $\mu\text{M}$ )
$\phi 11\text{DUT}^{\text{WT}}$	$5.5 \pm 1.6$	$1.2 \pm 0.5$
$\phi 11\text{DUT}^{\Delta 101-122}$	$4.5 \pm 1.3$	$2.5 \pm 0.4$
$\phi 11\text{DUT}^{\text{F164W}}$	$5.0 \pm 1.2$	$1.6 \pm 0.3$
$\phi 11\text{DUT}^{\text{F108W}}$	$5.3 \pm 1.8$	ND
$\phi 11\text{DUT}^{\text{E158STOP}}$	$<10^{-3}$	ND

contains a tryptophan fluorophore also within the phage-specific insert, and found that this protein also did not show any fluorescence spectral changes upon binding of either nucleotide ligands (Fig. 4c, Supplementary Table S2), indicating that the insert may not contact motif V during when this segment flips over the active site or their interaction may not change upon ligand binding. Although tryptophan fluorescence is a very useful and sensitive measure of protein conformation, it is still an indirect technique in which a lack of signal is indicative of a lack of conformational change but does not constitute direct solid proof.

To complement the fluorescence measurements with a more direct approach, we wondered whether removal of the insert would have any influence on the catalytic properties of  $\phi 11$  phage dUTPase. Earlier studies concluded that removal of the insert is deleterious to enzymatic activity (Tormo-Más *et al.*, 2010); however, in this case the truncated construct might have been also compromised in folding. Structural knowledge of this protein enabled us to design a truncated construct in which folding may not be perturbed (construct  $\phi 11\text{DUT}^{\Delta 101-122}$ ); in this construct, the first six residues of the insert are still present to potentially serve as a linker segment. We tested the enzymatic activity of this protein and found that both the catalytic rate constant ( $k_{\text{cat}}$ ) and the Michaelis–Menten constant ( $K_{\text{m}}$ ) are only slightly changed compared with the wild-type enzyme (Table 2). We therefore conclude, in contrast to the previous report (Tormo-Más *et al.*, 2010), that the phage-specific insert does not have any significant effect on the catalytic rate constant of dUTPase and its removal does not interfere with the folding of the protein.

### 3.3. Catalytic mechanism of $\phi 11$ phage dUTPase

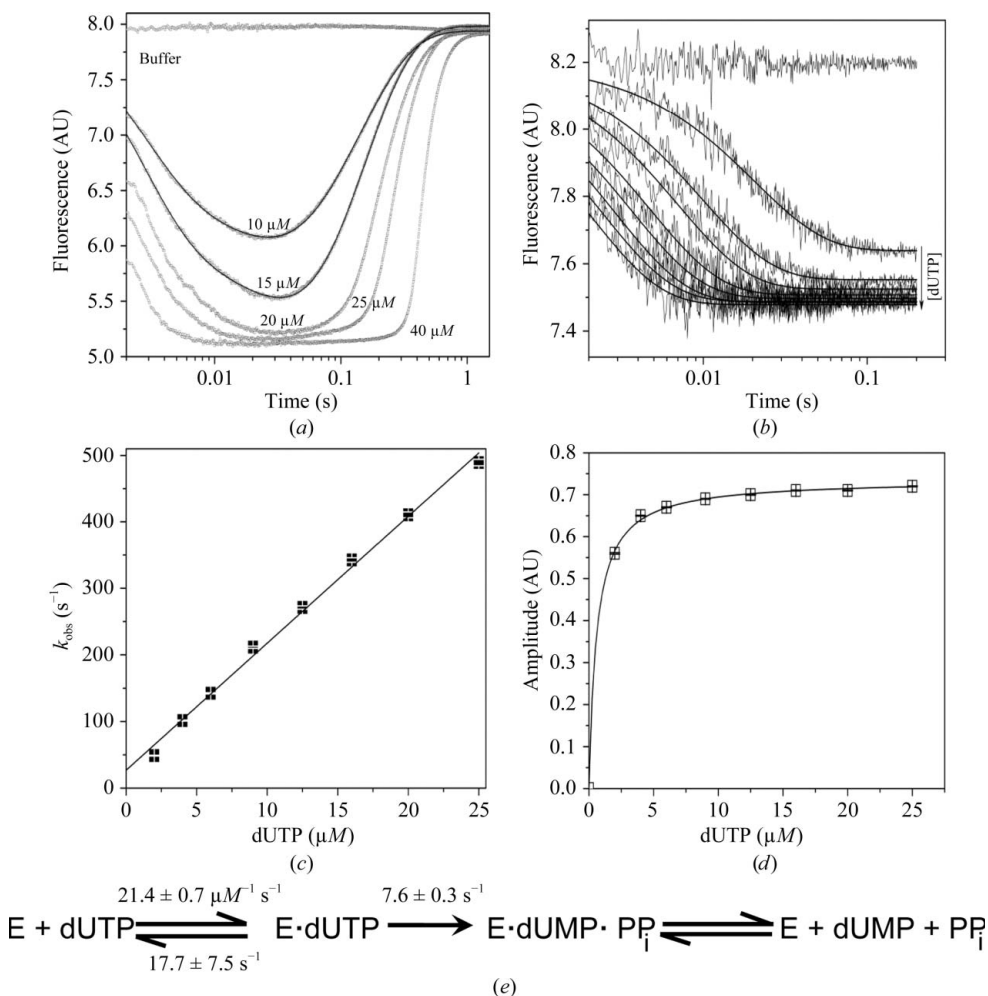
In order to analyze the kinetics of the elementary steps of the dUTPase enzymatic reaction for the  $\phi 11$  phage enzyme and to determine whether this specific protein is similar or different from other dUTPases (human or bacterial), we exchanged the phenylalanine residue within motif V to tryptophan (yielding the  $\phi 11\text{DUT}^{\text{F164W}}$  protein), similarly to the previous cases in which this fluorophore was successfully used to analyze the catalytic cycle (Pécsi *et al.*, 2011; Merényi *et al.*, 2011; Tóth *et al.*, 2007).

This active-site mutant  $\phi 11\text{DUT}^{\text{F164W}}$  had the same  $k_{\text{cat}}$  and  $K_{\text{m}}$  values as the wild type (Table 2) and showed very similar



signal changes upon ligand binding compared with other dUTPases (Fig. 4*d*, Supplementary Table S2). The fluorescence signal quenching upon ligand binding is somewhat less extensive compared with the cases of human or *Mycobacterium tuberculosis* dUTPase (Supplementary Table S2), presumably owing to the presence of the second tryptophan residue (Trp143) in  $\phi$ 11 phage dUTPase that does not change

its spectral parameters (Fig. 4*b*) and therefore ‘dilutes’ the signal. To confirm that the fluorescence change is specific to the ligand binding, we also performed titration with the substrate analogue  $\alpha,\beta$ -imido-dUTP (Fig. 4*e*). The titration yielded  $K_d = 0.32 \mu\text{M}$  for the  $\phi$ 11 phage dUTPase- $\alpha,\beta$ -imido-dUTP complex, which is in good agreement with previous data ranging from 0.5 to 2  $\mu\text{M}$  for other dUTPases (Pécsi *et al.*, 2011; Takács *et al.*, 2010; Varga *et al.*, 2007; Tóth *et al.*, 2007).



**Figure 5**

Transient kinetics studies describe the elementary steps of the enzymatic reaction. (a) Active-site titration of  $\phi$ 11DUT<sup>F164W</sup> enzyme by stopped-flow kinetics.  $\phi$ 11DUT<sup>F164W</sup> enzyme was mixed with equal volumes of varying concentrations of dUTP substrate in the stopped-flow apparatus. The post-mixing concentration of dUTP is indicated on the graph. The smooth line represents triple exponential fits to the single-turnover reaction curves. The reaction curves have two phases with decreasing fluorescence and one with increasing fluorescence. These fluorescence phases were previously determined to be (i) substrate binding, (ii) isomerization and (iii) product release, with the rate determined by the spectrally invisible hydrolysis step (Tóth *et al.*, 2007). A triple exponential fit to the reaction curve yielded  $k_{\text{observed}}$  ( $k_{\text{obs1}}$ ) = 375  $\text{s}^{-1}$ ,  $k_{\text{obs2}}$  = 52  $\text{s}^{-1}$  and  $k_{\text{H}}$  = 7.9  $\text{s}^{-1}$  for 10  $\mu\text{M}$  dUTP and  $k_{\text{obs1}}$  = 396  $\text{s}^{-1}$ ,  $k_{\text{obs2}}$  = 34  $\text{s}^{-1}$  and  $k_{\text{H}}$  = 7.6  $\text{s}^{-1}$  for 15  $\mu\text{M}$  dUTP. (b) shows a representative measurement of dUTP binding. Fluorescence time courses were recorded during the initial binding phase of the reaction at various dUTP concentrations, with a constant dUTPase ( $\phi$ 11DUT<sup>F164W</sup>) concentration of 0.5  $\mu\text{M}$  (post-mixing concentration). Smooth lines are single exponential fits to the experimental curves. (c) and (d) show the analysis of three independent measurements presented in (b). In (c) black scatter data points show the observed rate constants from the exponential fits. Error bars represent the SD for  $n = 3$ . The smooth line is a linear fit to the data, yielding rate constants  $k_{\text{on}} = 21.4 \pm 0.7 \mu\text{M}^{-1} \text{ s}^{-1}$ ,  $k_{\text{off}} = 17.7 \pm 7.5 \text{ s}^{-1}$  and  $K_d = 0.8 \pm 0.5 \mu\text{M}$ . These rate constants are of the same order of magnitude as for other dUTPases, although somewhat lower, probably owing to the higher salt concentration used in present experiments. In (d) open black scatter data points show the total amplitude change of the fluorescence signal upon dUTP substrate binding. Fitting of a hyperbolic function to the data yielded  $K_d = 0.3 \pm 0.03 \mu\text{M}$ . (e) shows a simplified kinetic model of  $\phi$ 11DUT<sup>F164W</sup> dUTP hydrolysis.

As we obtained a characteristic fluorescence change upon the binding of substrate and product to  $\phi$ 11DUT<sup>F164W</sup>, we also performed fast kinetic active-site titration by stopped flow to observe the intrinsic catalytic constant and real active-site concentration (Fig. 5*a*). Active-site titration of  $\phi$ 11DUT<sup>F164W</sup> shows similarly progressive fluorescent curves to those reported for human and *M. tuberculosis* dUTPases (Pecsi *et al.*, 2010; Tóth *et al.*, 2007). The close similarity of the reaction curves that can be fitted with three exponential phases implies that the kinetic mechanism of  $\phi$ 11DUT<sup>F164W</sup> is also similar to those of the above-mentioned dUTPases (Pecsi *et al.*, 2010; Tóth *et al.*, 2007). Accordingly, the main reaction steps are (i) fast substrate binding, (ii) relatively slow substrate-induced isomerization to the rate-limiting hydrolysis step (iii) and (iv) rapid release of the products. The results show that  $k_{\text{hydrolysis}}$  ( $k_{\text{H}}$ ) is  $7.6 \pm 0.3 \text{ s}^{-1}$  for  $\phi$ 11DUT<sup>F164W</sup>. Based on the titration, the major part ( $\geq 75\%$ ) of the protein preparation is active, indicating that our constructs are fully competent for detailed analysis of structure and function. The measured catalytic rate constant and the measured active proportion of the enzyme preparation are in good agreement with the steady-state activity (Table 2).

Transient kinetic measurement of dUTP binding under pseudo-first-order conditions (Figs. 5*c* and 5*d*) showed fast substrate binding and a submicromolar  $K_d$ . These results further support that the kinetic mechanism of the

reaction is similar to that reported for other dUTPases (for a simplified version of the previously proposed kinetic mechanism, see Fig. 5e). In summary, our ligand-binding, steady-state and transient kinetic studies showed that irrespective of the presence of the species-specific insert, the  $\phi$ 11 phage dUTPase is also a highly efficient enzyme. Its catalytic cycle involves fast substrate binding and product release, with the rate-determining step being the hydrolysis of dUTP. The enzyme is characterized by a  $k_{\text{cat}}$  of  $7.6 \text{ s}^{-1}$  that allows it to quickly remove dUTP from the nucleotide pool.

#### 4. Conclusions

Our data showed that the phage-specific insert segment of  $\phi$ 11 phage dUTPase folds into a  $\beta$ -pleated mini-domain resembling a distorted Greek-key motif. This small structural motif is very common in protein-folding cores and has also been used for the design of artificial mini-proteins (Dowd *et al.*, 2002; Zheng *et al.*, 2011; Hatfield *et al.*, 2011). Comparison of the presently available sequences of staphylococcal phages encoding trimeric dUTPases (Supplementary Fig. S2) reveals that the presence of such inserts seems to be general and that subgroups of phages can be distinguished based on the segment characteristics (Supplementary Fig. S3). Although the insert sequences are varied, the fold adopted by the insert is very similar in  $\phi$ 11 and  $80\alpha$  phage dUTPases (Fig. 2; Tormo-Más *et al.*, 2013). Within the insert, conserved residues can also be identified; the aspartate responsible for coordinating a  $\text{Mg}^{2+}$  ion (Asp95 in  $\phi$ 11 phage dUTPase) is strictly conserved. The structure of the staphylococcal phage  $80\alpha$  dUTPase, belonging to a distinct group (Supplementary Figs. S2 and S3), was also shown to contain the  $\beta$ -pleated mini-domain, although the low resolution prevented the reliable identification of an intrachannel  $\text{Mg}^{2+}$  ion (Tormo-Más *et al.*, 2013).

To complement the crystal structure analysis, we have shown using kinetics and ligand-binding studies in solution that the phage-specific insert has no effect on the enzymatic activity and that its conformation may not be sensitive to the changes during the catalytic cycle. Using different site-specific fluorophores, we determined that the phage dUTPase is also a highly efficient enzyme characterized by fast substrate binding and efficient catalysis. These characteristics are even more important in the context of several staphylococcal strains in which a genomic copy of dUTPase could not be identified (Chua *et al.*, 2010; Holden *et al.*, 2010; Baba *et al.*, 2008). Hence, these strains rely on helper phage dUTPases to provide a nucleotide pool with the correct balance between dUTP and dTTP to help genomic integrity.

We thank Dr Attila Reményi for data collection at SLS. This work was supported by the Hungarian Scientific Research Fund OTKA NK 84008, NK83857, NK101072 and K109486, the Baross Program of the New Hungary Development Plan (3DSTRUCT, OMFB-00266/2010 REG-KM-09-1-2009-0050), the Hungarian Academy of Sciences (TTK IF-28/2012) and the European Commission FP7 Biostruct-X project (contract No. 283570).

#### References

- Baba, T., Bae, T., Schneewind, O., Takeuchi, F. & Hiramoto, K. (2008). *J. Bacteriol.* **190**, 300–310.
- Barabás, O., Németh, V. & Vértessy, B. G. (2006). *Acta Cryst.* **F62**, 399–401.
- Barabás, O., Pongrácz, V., Kovári, J., Wilmanns, M. & Vértessy, B. G. (2004). *J. Biol. Chem.* **279**, 42907–42915.
- Benesch, J. L., Ruotolo, B. T., Simmons, D. A. & Robinson, C. V. (2007). *Chem. Rev.* **107**, 3544–3567.
- Brinkman, F. S. L. & Leipe, D. D. (2001). *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, edited by A. D. Baxevanis & B. F. F. Oulette, pp. 323–358. Hoboken: John Wiley & Sons.
- Castillo-Acosta, V. M., Aguilar-Pereyra, F., Bart, J. M., Navarro, M., Ruiz-Pérez, L. M., Vidal, A. E. & González-Pacanowska, D. (2012). *DNA Repair*, **11**, 986–995.
- Chan, S. *et al.* (2004). *J. Mol. Biol.* **341**, 503–517.
- Chaurushiya, M. S. & Weitzman, M. D. (2009). *DNA Repair*, **8**, 1166–1176.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Chua, K., Seemann, T., Harrison, P. F., Davies, J. K., Coutts, S. J., Chen, H., Haring, V., Moore, R., Howden, B. P. & Stinear, T. P. (2010). *J. Bacteriol.* **192**, 5556–5557.
- Doseth, B., Ekre, C., Slupphaug, G., Krokan, H. E. & Kavli, B. (2012). *DNA Repair*, **11**, 587–593.
- Dowd, C. S., Leavitt, S., Babcock, G., Godillot, A. P., Van Ryk, D., Canziani, G. A., Sodroski, J., Freire, E. & Chaiken, I. M. (2002). *Biochemistry*, **41**, 7038–7046.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Freeman, L., Buisson, M., Tarbouriech, N., Van der Heyden, A., Labbé, P. & Burmeister, W. P. (2009). *J. Biol. Chem.* **284**, 25280–25289.
- Galperin, M. Y., Moroz, O. V., Wilson, K. S. & Murzin, A. G. (2006). *Mol. Microbiol.* **59**, 5–19.
- García-Nafria, J., Burchell, L., Takezawa, M., Rzechorzek, N. J., Fogg, M. J. & Wilson, K. S. (2010). *Acta Cryst.* **D66**, 953–961.
- García-Nafria, J., Harkiolaki, M., Persson, R., Fogg, M. J. & Wilson, K. S. (2011). *Acta Cryst.* **D67**, 167–175.
- Grandori, R., Santambrogio, C., Brocca, S., Invernizzi, G. & Lotti, M. (2009). *Biotechnol. J.* **4**, 73–87.
- Harkiolaki, M., Dodson, E. J., Bernier-Villamor, V., Turkenburg, J. P., González-Pacanowska, D. & Wilson, K. S. (2004). *Structure*, **12**, 41–53.
- Hatfield, M. P., Murphy, R. F. & Lovas, S. (2011). *J. Phys. Chem. B*, **115**, 4971–4981.
- Heck, A. J. R. (2008). *Nature Methods*, **5**, 927–933.
- Hillis, D. M. & Bull, J. J. (1993). *Syst. Biol.* **42**, 182–192.
- Holden, M. T., Lindsay, J. A., Corton, C., Quail, M. A., Cockfield, J. D., Pathak, S., Batra, R., Parkhill, J., Bentley, S. D. & Edgeworth, J. D. (2010). *J. Bacteriol.* **192**, 888–892.
- Homma, K. & Moriyama, H. (2009). *Acta Cryst.* **F65**, 1030–1034.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kahánková, J., Pantůček, R., Goerke, C., Růžičková, V., Holochová, P. & Doškař, J. (2010). *Environ. Microbiol.* **12**, 2527–2538.
- Kovári, J., Barabás, O., Takács, E., Békési, A., Dubrovay, Z., Pongrácz, V., Zagyva, I., Imre, T., Szabó, P. & Vértessy, B. G. (2004). *J. Biol. Chem.* **279**, 17932–17944.
- Kovári, J., Barabás, O., Varga, B., Békési, A., Tölgyesi, F., Fidy, J., Nagy, J. & Vértessy, B. G. (2008). *Proteins*, **71**, 308–319.
- Langerak, P. & Russell, P. (2012). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 3562–3571.
- Leveles, I., Róna, G., Zagyva, I., Bendes, Á., Harmat, V. & Vértessy, B. G. (2011). *Acta Cryst.* **F67**, 1411–1413.

- Maiques, E., Úbeda, C., Tormo, M. Á., Ferrer, M. D., Lasa, Í., Novick, R. P. & Penadés, J. R. (2007). *J. Bacteriol.* **189**, 5608–5616.
- Merényi, G., Kovári, J., Tóth, J., Takács, E., Zagya, I., Erdei, A. & Vértessy, B. G. (2011). *Nucleosides Nucleotides Nucleic Acids*, **30**, 369–390.
- Miyahara, S. *et al.* (2012). *J. Med. Chem.* **55**, 2970–2980.
- Mol, C. D., Harris, J. M., McIntosh, E. M. & Tainer, J. A. (1996). *Structure*, **4**, 1077–1092.
- Moroz, O. V., Harkiolaki, M., Galperin, M. Y., Vagin, A. A., González-Pacanoska, D. & Wilson, K. S. (2004). *J. Mol. Biol.* **342**, 1583–1597.
- Moroz, O. V., Murzin, A. G., Makarova, K. S., Koonin, E. V., Wilson, K. S. & Galperin, M. Y. (2005). *J. Mol. Biol.* **347**, 243–255.
- Muha, V., Horváth, A., Békési, A., Pukáncsik, M., Hodoscsek, B., Merényi, G., Róna, G., Batki, J., Kiss, I., Jankovics, F., Vilmos, P., Erdélyi, M. & Vértessy, B. G. (2012). *PLoS Genet.* **8**, e1002738.
- 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. D67*, 355–367.
- Németh-Pongrácz, V., Barabás, O., Fuxreiter, M., Simon, I., Pichová, I., Rumlová, M., Záborská, H., Svergun, D., Petoukhov, M., Harmat, V., Klement, E., Hunyadi-Gulyás, E., Medzihradzsky, K. F., Kónya, E. & Vértessy, B. G. (2007). *Nucleic Acids Res.* **35**, 495–505.
- Ngounou Wetie, A. G., Sokolowska, I., Woods, A. G., Roy, U., Loo, J. A. & Darie, C. C. (2013). *Proteomics*, **13**, 538–557.
- Niida, H., Shimada, M., Murakami, H. & Nakanishi, M. (2010). *Cancer Sci.* **101**, 2505–2509.
- Novick, R. P. & Subedi, A. (2007). *Chem. Immunol. Allergy*, **93**, 42–57.
- Payne, S. L. & Elder, J. H. (2001). *Curr. Protein Pept. Sci.* **2**, 381–388.
- Pecsi, I., Hirmondo, R., Brown, A. C., Lopata, A., Parish, T., Vértessy, B. G. & Tóth, J. (2012). *PLoS One*, **7**, e37461.
- Pecsi, I., Leveles, I., Harmat, V., Vértessy, B. G. & Toth, J. (2010). *Nucleic Acids Res.* **38**, 7179–7186.
- Pécsi, I., Szabó, J. E., Adams, S. D., Simon, I., Sellers, J. R., Vértessy, B. G. & Tóth, J. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 14437–14442.
- Persson, R., Cedergren-Zeppezauer, E. S. & Wilson, K. S. (2001). *Curr. Protein Pept. Sci.* **2**, 287–300.
- 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.
- Prasad, G. S., Stura, E. A., Elder, J. H. & Stout, C. D. (2000). *Acta Cryst. D56*, 1100–1109.
- Rouse, J. & Jackson, S. P. (2002). *Science*, **297**, 547–551.
- Ruzin, A., Lindsay, J. & Novick, R. P. (2001). *Mol. Microbiol.* **41**, 365–377.
- Samal, A., Schormann, N., Cook, W. J., DeLucas, L. J. & Chattopadhyay, D. (2007). *Acta Cryst. D63*, 571–580.
- Takács, E., Barabás, O., Petoukhov, M. V., Svergun, D. I. & Vértessy, B. G. (2009). *FEBS Lett.* **583**, 865–871.
- Takács, E., Nagy, G., Leveles, I., Harmat, V., Lopata, A., Tóth, J. & Vértessy, B. G. (2010). *FEBS Lett.* **584**, 3047–3054.
- Tarbouriech, N., Buisson, M., Seigneurin, J.-M., Cusack, S. & Burmeister, W. P. (2005). *Structure*, **13**, 1299–1310.
- Tormo, M. Á., Ferrer, M. D., Maiques, E., Úbeda, C., Selva, L., Lasa, Í., Calvete, J. J., Novick, R. P. & Penadés, J. R. (2008). *J. Bacteriol.* **190**, 2434–2440.
- Tormo-Más, M. Á., Donderis, J., García-Caballer, M., Alt, A., Mir-Sanchis, I., Marina, A. & Penadés, J. R. (2013). *Mol. Cell*, **49**, 947–958.
- Tormo-Más, M. Á., Mir, I., Shrestha, A., Tallent, S. M., Campoy, S., Lasa, Í., Barbé, J., Novick, R. P., Christie, G. E. & Penadés, J. R. (2010). *Nature (London)*, **465**, 779–782.
- Tóth, J., Varga, B., Kovács, M., Málnási-Csizmadia, A. & Vértessy, B. G. (2007). *J. Biol. Chem.* **282**, 33572–33582.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D66*, 22–25.
- Varga, B., Barabás, O., Kovári, J., Tóth, J., Hunyadi-Gulyás, E., Klement, E., Medzihradzsky, 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.
- Vértessy, B. G. (1997). *Proteins*, **28**, 568–579.
- Vértessy, B. G., Larsson, G., Persson, T., Bergman, A. C., Persson, R. & Nyman, P. O. (1998). *FEBS Lett.* **421**, 83–88.
- Vértessy, B. G. & Tóth, J. (2009). *Acc. Chem. Res.* **42**, 97–106.
- Visnes, T., Doseth, B., Pettersen, H. S., Hagen, L., Sousa, M. M., Akbari, M., Otterlei, M., Kavli, B., Slupphaug, G. & Krokan, H. E. (2009). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 563–568.
- Wilson, P. M., LaBonte, M. J., Lenz, H.-J., Mack, P. C. & Ladner, R. D. (2012). *Mol. Cancer Ther.* **11**, 616–628.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D67*, 235–242.
- Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst. D57*, 122–133.
- Zheng, W., Qi, B., Rohrdanz, M. A., Caflisch, A., Dinner, A. R. & Clementi, C. (2011). *J. Phys. Chem. B*, **115**, 13065–13074.