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

Crystallization and preliminary X-ray analysis of three dUTPases from Gram-positive bacteria

Gui-Lan Li,[a] Juan Wang,[a] Lan-Fen Li[a] and Xiao-Dong Su[a,b]*

[a]National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Science, Peking University, Beijing 100871, People's Republic of China, and [b]Peking University Shenzhen Graduate School, Shenzhen 518055, People's Republic of China

Correspondence e-mail: xdsu@pku.edu.cn

Received 16 December 2008
Accepted 20 February 2009

All organisms examined to date possess a dUTPase that performs the important function of efficiently hydrolyzing dUTP to dUMP in order to prevent the incorporation of dUTP into DNA. Three putative dUTPases from Gram-positive bacteria have been studied in this work. Two dUTPase-encoding genes, *yncF* and *yosS*, have been identified in *Bacillus subtilis*. The gene *dut*, encoding dUTPase from the dental pathogen *Streptococcus mutans*, was amplified from *S. mutans* genomic DNA. The three genes were cloned into expression vectors and overexpressed at high levels in *Escherichia coli*. Each protein was purified in two steps using chromatographic methods. Crystals of the YosS and YncF proteins and of *S. mutans* dUTPase were obtained using the vapour-diffusion method. X-ray diffraction data sets were collected from crystals of selenomethionine-labelled YosS and *S. mutans* dUTPase to resolutions of 2.3 and 1.7 Å, respectively. The crystal of native YncF diffracted to 2.7 Å resolution.

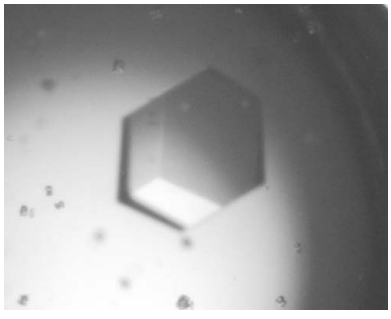
1. Introduction

Deoxyuridine triphosphate nucleotidohydrolase (dUTPase), which is responsible for the Mg^{2+}-dependent hydrolysis of dUTP to dUMP and inorganic pyrophosphate, is a ubiquitous enzyme in free-living organisms (Shlomai & Kornberg, 1978). This reaction controls the cellular dUTP:dTTP ratio, preventing the incorporation of dUTP into DNA, and provides dUMP as a precursor of dTTP synthesis (McIntosh & Haynes, 1997; Ladner, 2001; Vértessy & Tóth, 2009). The inactivation of dUTPase results in chromosome fragmentation and thymine-less cell death. Although three classes of dUTPase are known based on structural characteristics, the subgroup of homotrimeric dUTPases is quite predominant in nature and is found in eukaryotic, bacterial and retroviral organisms (Mol *et al.*, 1996; Barabás *et al.*, 2004, 2006; Samal *et al.*, 2007; Németh-Pongrácz *et al.*, 2007; Varga *et al.*, 2007). The homotrimeric fold of dUTPases shares five conserved sequence motifs that constitute the active site (Fig. 1).

Two putative dUTPase-encoding genes have been identified in *Bacillus subtilis* (Kunst *et al.*, 1997; Persson *et al.*, 2001, 2005). One is named *yosS* (location 2159 kb), corresponding to the integrated bacteriophage SPβ. The other, denoted *yncF* (location 1899 kb), is the host paralogue. The gene products of *yosS* and *yncF* constitute 142 and 144 residues, respectively, and have 93% sequence identity. To date, studies have been focused on systematic gene disruption in order to identify the essential bacterial functions encoded by the gene pair (Thomaides *et al.*, 2007).

In *Streptococcus mutans*, the principal pathogen in human dental caries (Mosci *et al.*, 1990), the dUTPase (SmdUTPase) is encoded by the gene *dut* and has a molecular weight of 16.1 kDa. Multiple sequence alignment shows that the dUTPases from *B. subtilis* and *S. mutans* exhibit the homotrimeric fold, although they share low sequence identities (Fig. 1).

In this work, we report the preparation, crystallization and preliminary X-ray analysis of the three dUTPases from the two above-mentioned Gram-positive bacteria. YosS and SmdUTPase have about 30% sequence identity to other known trimeric dUTPases and



therefore it has proven difficult to solve their structures using the molecular-replacement method. Selenomethionine-substituted proteins were therefore produced for two of the proteins. The structural and functional studies of these dUTPases will provide more detailed information on these important enzymes and may help us to design specific inhibitors for dUTPases (Nguyen *et al.*, 2005).

2. Materials and methods

2.1. Cloning and expression

The *B. subtilis yosS* and *yncF* genes and the *S. mutans dut* gene were amplified by PCR using *B. subtilis* and *S. mutans* genomic DNA as templates. The *yosS* gene was cloned into the pET21-DEST expression vector using Gateway cloning technology similar to that described by Ren *et al.* (2004), whereas the *yncF* and *S. mutans dut* genes were cloned into the vector pET28a (Novagen) at the *Bam*HI and *Xho*I restriction sites using the conventional cloning method. The correctness of these target vectors was verified by DNA sequencing. The recombinant plasmids containing N-terminal fusion His_6 tags were then transformed into *Escherichia coli* strain BL21 (DE3) cells for further expression experiments.

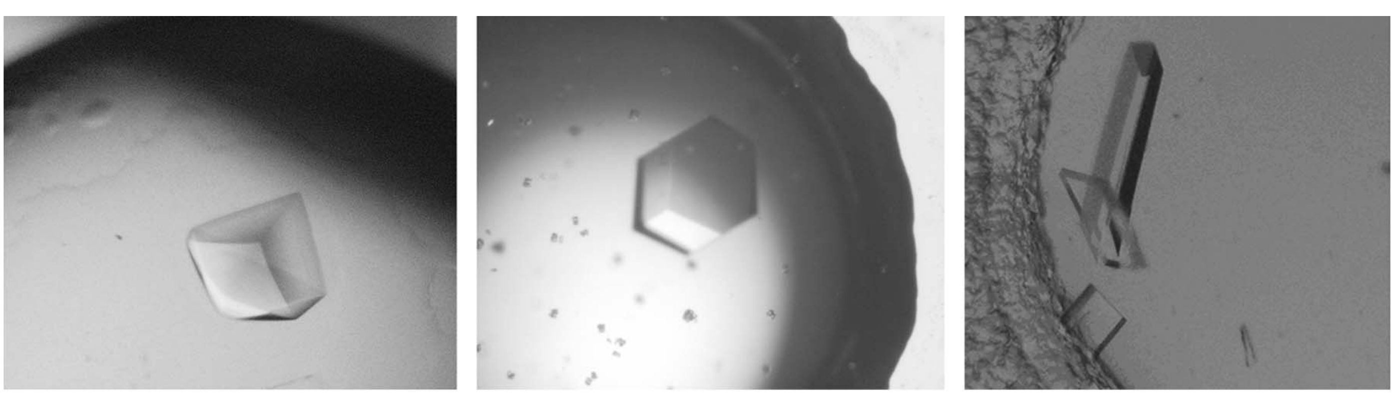
Selenomethionine-substituted protein cultivation was carried out for the YosS protein and SmdUTPase. Transformed cells grown overnight in 20 ml Luria–Bertani (LB) medium at 310 K were centrifuged at 2831*g* for 10 min at room temperature and then resuspended in 1 ml M9 minimal medium. The resuspended cells were grown in 1.0 l fresh M9 medium with 50 mg ml^{-1} ampicillin (for YosS expression) or 50 µg ml^{-1} kanamycin (for SmdUTPase expression) supplemented with 2 m*M* $MgSO_4$, 0.1 m*M* $CaCl_2$ and 0.4% glucose at 310 K. When the OD_{600} reached 0.6, 50 mg selenomethionine, leucine, isoleucine and valine and 100 mg lysine, threonine and phenylalanine were added to each of the cell cultures. After incubation for 30 min, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 m*M* and cultivation was continued for about 20 h at 291 K for expression of SeMet-YosS or SeMet-SmdUTPase. Cells were harvested by centrifugation at 4671*g* for 10 min and resuspended in buffer *A* (20 m*M* Tris–HCl, 500 m*M* NaCl pH 7.5).

E. coli cells containing pET28a-*yncF* were grown in LB medium with 50 µg ml^{-1} kanamycin until an OD_{600} of 0.6 was reached. The expression of target protein was then induced with 0.5 m*M* IPTG for 4 h at 310 K. The cells were then harvested and resuspended as described above.

```
                                                          **                                     **
YosS_bs   ------------------------MQIKIKYLDETQTRINKMEQGDWIDLRAAE--DVAIKKDEFKLVPLGVAMELPEGYEA-HVVPRSS 63
YncF_bs   ----------------------MTMQIKIKYLDETQTRISKIEQGDWIDLRAAE--DVTIKKDEFKLVPLGVAMELPEGYEA-HVVPRSS 65
 DUT_sm   -----------------MMKTRGFELITDYTDENLLPKRETAHAAGYDLKVAER--TEISAGAIVLVPTGVKAYMQVGEVL-YLFDRSS 69
 DUT_ec   -----------------MKKIDVKILDPRVGKEFPLPTYATSGSAGLDLRACLNDAVELAPGDTTLVPTGLAIHIADPSLAAMMLPRSG 72
 DUT_hs   MPCSEETPAISPSKRARPAEVGGMQLRFARLSEHATAPTRGSARAAGYDLYSAY--DYTIPPMEKAVVKTDIQIALPSGCYG-RVAPRSG 87
 DUT_vv   --------------------------MFVKETNRAKSPTRQSPGAAGYDLYSAY--DYTIPPGERQLIKTDISMSMPKFCYG-RIAPRSG 61

                       * **  *                          * *  *                          * *  *
YosS_bs   TYKNFGVIQTNSMGVIDESYKGDND-----FWFFPAYALRDTKIKKGDRICQFRIMKKMPAVDLIEVDRLGNGDRG--GHGSTGTK 142
YncF_bs   TYKNFGVIQTNSMGVIDESYKGDND-----FWFFPAYALRDTEIKKGDRICQFRIMKKMPAVELVEVEHLGNEDRG--GLGSTGTK 144
 DUT_sm   NPRKKGLVLINSVGVIDGDYYNNPNNEGHIFAQMKNMTDQTVVLEAGERVVQG------VFMPFLLIDGDKATGTRTGGFGSTGG- 148
 DUT_ec   LGHKHGIVLGNLVGLIDSDYQGQLM-----ISVWN-RGQDSFTIQPGERIAQM-IFVPVVQAEFNLVEDFDATDRGEGGFGHSGRQ 151
 DUT_hs   LAAKHFIDVG--AGVIDEDYRGNVG-----VVLFN-FGKEKFEVKKGDRIAQL-ICERIFYPEIEEVQALDDTERGSGGFGSTGKN 164
 DUT_vv   LSLKG-IDIG--GGVIDEDYRGNIG-----VILIN-NGKCTFNVNTGDRIAQL-IYQRIYYPELEEVQSLDSTNRGDQGFGSTGLR 137
```

Figure 1
Multiple sequence alignment of trimeric dUTPase homologues using the program *ClustalX* (Thompson *et al.*, 1997). The positions of the five conserved motifs are marked by black bars. The conserved residues in the motifs are marked with asterisks. YosS_bs, *B. subtilis* YosS; YncF_bs, *B. subtilis* YncF; DUT_sm, *S. mutans* dUTPase; DUT_ec, *E. coli* dUTPase (Barabás *et al.*, 2004); DUT_hs, *Homo sapiens* dUTPase (Mol *et al.*, 1996; Varga *et al.*, 2007); DUT_vv, vaccinia virus dUTPase (Samal *et al.*, 2007). The sequence identities between YosS and YncF, *S. mutans* dUTPase, *E. coli* dUTPase, human dUTPase and vaccinia virus dUTPase are 93, 28, 31, 32 and 30%, respectively.



(*a*) (*b*) (*c*)

Figure 2
(*a*) The best crystal of SeMet-YosS obtained using the hanging-drop vapour-diffusion method, grown in 2.0 *M* ammonium sulfate, 0.2 *M* potassium/sodium tartrate, 0.1 *M* trisodium citrate pH 5.6; the dimensions of the crystal are approximately 0.1 × 0.1 × 0.2 mm. (*b*) Crystal of native YncF obtained by the hanging-drop method with 0.2 *M* trimethylamine *N*-oxide dehydrate, 15%(*w*/*v*) PEG MME 2000, 0.1 *M* Tris–HCl pH 9.0; the dimensions of the crystal are approximately 0.2 × 0.1 × 0.1 mm. (*c*) Crystal of SeMet *S. mutans* dUTPase obtained using the sitting-drop method, grown in the condition 0.2 *M* $MgCl_2$, 25%(*w*/*v*) PEG 3350, 0.1 *M* Tris–HCl pH 8.5. The dimensions of the crystal are approximately 0.1 × 0.1 × 0.7 mm.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	SeMet-YosS			SeMet-SmdUTPase	
	Peak	Edge	Remote	Peak	Native YncF
Wavelength (Å)	0.9803	0.9809	1.100	0.9787	1.5418
Space group	$P4_3$			$C2$	$C2$
Unit-cell parameters (Å, °)	$a = b = 116.80$, $c = 57.0$			$a = 88.79$, $b = 53.54$, $c = 93.29$, $\beta = 113.87$	$a = 185.89$, $b = 53.02$, $c = 93.30$, $\beta = 94.17$
Resolution range (Å)	50–2.3 (2.4–2.3)			30–1.7 (1.8–1.7)	51–2.7 (2.8–2.7)
Completeness (%)	96.6 (97.7)	96.1 (96.7)	92.5 (94.8)	98.6 (96.1)	92.5 (94.8)
No. of observed reflections	176806	166810	149071	157686	68403
No. of unique reflections	64117	63833	59530	43830	23426
Average $I/\sigma(I)$	11.5 (4.8)	11.7 (5.4)	7.9 (2.5)	15.28 (8.97)	5.36 (1.4)
$R_{\rm merge}$† (%)	8.0 (25.3)	7.7 (21.9)	10.0 (49.4)	4.7 (12.3)	9.42 (32.5)
$V_{\rm M}$ (Å^3 Da^{-1})	2.0	2.0	2.0	2.1	2.4

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

2.2. Protein purification and crystallization

All cells were lysed using a sonicator (JY-92, Ningbo, China) on ice and centrifuged twice (15 570*g*, 277 K, 35 min) to remove debris. The supernatant was collected and loaded onto a 5 ml HiTrap Ni^{2+}-chelating affinity column (GE Healthcare, USA) pre-equilibrated with buffer *A*. The target protein (SeMet-YosS, SeMet-SmdUTPase or native YncF) was eluted with a linear gradient of 50–500 m*M* imidazole. The eluted protein was further purified by running a 16/60 HiLoad Superdex 75 column (GE Healthcare, USA) equilibrated with 20 m*M* Tris–HCl buffer containing 150 m*M* NaCl pH 7.5. The samples of SeMet-YosS, SeMet-SmdUTPase and YncF were concentrated by the ultrafiltration method (Amicon Ultra-15, 10 kDa molecular-weight cutoff, Millipore, USA) to about 30, 15 and 20 mg ml^{-1}, respectively.

All crystallization experiments were performed at 289 K using the vapour-diffusion technique. Hampton Research Crystal Screen, Crystal Screen 2 and Index Screen kits (Hampton Research, USA) were used for initial screening conditions. 1 µl protein solution was mixed with 1 µl reservoir solution and equilibrated against 500 µl reservoir solution for SeMet-YosS or YncF using the hanging-drop method or 100 µl reservoir solution for SeMet-SmdUTPase using the sitting-drop method.

2.3. X-ray diffraction data collection

Selenium multiple-wavelength anomalous dispersion (MAD) data were collected for Se-YosS and selenium single-wavelength anomalous dispersion (SAD) data were collected for SeMet-SmdUTPase using MAR 165 CCD detectors on beamline ID14-1 of Beijing Synchrotron Radiation Facility (BSRF), China and on beamline I911-3 of MAX-II-lab, Lund, Sweden, respectively. Both crystals were flash-frozen to 100 K in a cold nitrogen-gas stream during data collection. Diffraction data were processed using the *XDS* program (Kabsch, 1993).

The diffraction data for native YncF were collected on a Bruker SMART 6000 CCD detector using Cu $K\alpha$ radiation from a Bruker–Nonius FR591 rotating-anode generator. The data were then processed using the Bruker *PROTEUM* online software.

3. Results

The three target proteins were expressed in *E. coli* BL21(DE3) in soluble form and were purified to homogeneity after gel filtration. All three proteins are homotrimers as judged from the gel-filtration results. The molecular weights of the three monomeric proteins (each about 16.0 kDa plus 4.0 kDa fusion peptide) shown on SDS–PAGE are consistent with their predicted molecular weights.

The best SeMet-YosS crystals were obtained within 3 d using the condition 2.0 *M* ammonium sulfate, 0.1 *M* trisodium citrate pH 5.6 and 0.2 *M* potassium/sodium tartrate; phase separation appeared in the crystal drop (Fig. 2*a*). The data set for SeMet YosS was 99.3% complete to 2.3 Å resolution, with an overall $R_{\rm merge}$ of 8.4%. The space group was determined to be $P4_3$, with unit-cell parameters $a = b = 116.80$, $c = 57.0$ Å. Crystals of SeMet-SmdUTPase that were sufficiently large for diffraction were obtained in the condition 0.2 *M* $MgCl_2$, 25%(*w*/*v*) polyethylene glycol (PEG) 3350, 0.1 *M* Tris–HCl pH 8.5 (Fig. 2*b*). The crystal of SeMet-SmdUTPase diffracted to 1.7 Å resolution and belonged to space group *C*2. The unit-cell parameters were $a = 88.79$, $b = 53.54$, $c = 93.29$ Å, $\beta = 113.87°$. Both the SeMet-YosS and the SeMet-SmdUTPase crystals are expected to contain three monomeric molecules per asymmetric unit; the $V_{\rm M}$ values were 2.0 and 2.1 Å^3 Da^{-1}, corresponding to solvent contents of 38.3% and 41.2%, respectively (Matthews, 1968). Data-collection statistics are summarized in Table 1.

The crystal of YncF appeared after about two weeks in 0.2 *M* trimethylamine *N*-oxide dehydrate, 15%(*w*/*v*) polyethylene glycol monomethyl ether 2000 (PEG MME 2000), 0.1 *M* Tris–HCl pH 9.0 (Fig. 2*c*). The crystal diffracted to 2.7 Å resolution and belonged to space group *C*2 (unit-cell parameters $a = 185.89$, $b = 53.02$, $c = 93.30$ Å, $\beta = 94.17°$). Assuming the presence of six molecules per asymmetric unit gave a $V_{\rm M}$ value of 2.4 Å^3 Da^{-1} and a corresponding solvent content of 48.1% (Table 1).

This work was supported by a grant (30530190) from the National Natural Science Foundation of China. Peking University's 985 and 211 grants are gratefully acknowledged. We thank the beamline staff at BSRF and at MAX-II-lab, Lund, Sweden for help with data collection.

References

Barabás, O., Németh, V. & Vértessy, B. G. (2006). *Acta Cryst.* F**62**, 399–401.
Barabás, O., Pongrácz, V., Kovári, J., Wilmanns, M. & Vértessy, B. G. (2004). *J. Biol. Chem.* **279**, 42907–42915.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kunst, F. *et al.* (1997). *Nature (London)*, **390**, 249–256.
Ladner, R. D. (2001). *Curr. Protein Pept. Sci.* **2**, 361–370.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McIntosh, E. M. & Haynes, R. H. (1997). *Acta Biochim. Pol.* **44**, 159–171.
Mol, C. D., Harris, J. M., McIntosh, E. M. & Tainer, J. A. (1996). *Structure*, **4**, 1077–1092.

Mosci, F., Perito, S., Bassa, S., Capuano, A. & Marconi, P. F. (1990). *Minerva Stomatol.* **39**, 413–429.
Neméth-Pongrácz, V., Barabás, O., Fuxreiter, M., Simon, I., Pichová, I., Rumlová, M., Zábranská, H., Svergun, D., Petoukhov, M., Harmat, V., Klement, E., Hunyadi-Gulyás, E., Medzihradszky, K. F., Kónya, E. & Vértessy, B. G. (2007). *Nucleic Acids Res.* **35**, 495–505.
Nguyen, C., Kasinathan, G., Leal-Cortijo, I., Musso-Buendia, A., Kaiser, M., Brun, R., Ruiz-Pérez, L. M., Johansson, N. G., González-Pacanowska, D. & Gilbert, I. H. (2005). *J. Med. Chem.* **48**, 594–5954.
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.
Ren, H., Liang, Y., Li, R., Ding, H., Qiu, S., Lu, S., An, J., Li, L., Luo, M., Zheng, X. & Su, X.-D. (2004). *Acta Cryst.* D**60**, 1292–1294.
Samal, A., Schormann, N., Cook, W. J., DeLucas, L. J. & Chattopadhyay, D. (2007). *Acta Cryst.* D**63**, 571–580.
Shlomai, J. & Kornberg, A. (1978). *J. Biol. Chem.* **253**, 3305–3312.
Thomaides, H. B., Davison, E. J., Burston, L., Johnson, H., Brown, D. R., Hunt, A. C., Errington, J. & Czaplewski, L. (2007). *J. Bacteriol.* **189**, 591–602.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). *Nucleic Acids Res.* **25**, 4876–4882.
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.
Vértessy, B. G. & Tóth, J. (2009). *Acc. Chem. Res.* **42**, 97–106.