Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of deoxyuridine 5′-triphosphate nucleotidohydrolase from *Bacillus subtilis*

Rebecca Persson,* Maria Harkiolaki, John McGeehan† and Keith S. Wilson

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

† Present address: School of Biological Sciences, Biophysics Laboratories, University of Portsmouth, St Michael's Building, White Swan Road, Portsmouth, Hampshire PO1 2DT, England.

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

Received 4 March 2001
Accepted 9 May 2001

Single crystals of purified homotrimeric deoxyuridine 5′-triphosphate nucleotidohydrolase (dUTPase) from *Bacillus subtilis* have been grown under several different conditions using vapour diffusion. X-ray diffraction data have been collected using synchrotron radiation from three crystal forms of the unliganded enzyme and from enzyme cocrystallized with a substrate analogue and inhibitor, dUDP, and a metal ion, Sr^{2+}. The three crystal forms of unliganded enzyme belong to hexagonal ($P6_3$), orthorhombic ($P2_12_12$) and cubic ($P2_13$) space groups and data have been recorded to 1.75, 1.90 and 2.50 Å spacing, respectively. Crystals grown in the presence of dUDP and Sr^{2+} belong to the orthorhombic space group $P2_12_12_1$ and data were measured to 1.90 Å spacing. Solution of the hexagonal crystal form by molecular replacement using the dUTPase from feline immunodeficiency virus as a search model is in progress.

1. Introduction

Deoxyuridine 5′-triphosphate nucleotidohydrolase (dUTPase), an enzyme involved in nucleotide metabolism, catalyses the hydrolysis of dUTP to dUMP and inorganic pyrophosphate (Bertani *et al.*, 1961). dUTPase activity is essential for full viability in most biological systems, as maintenance of a low dUTP:dTTP ratio is required to safeguard the integrity of the genetic information stored in DNA (Kornberg & Baker, 1991). Perturbations in dUTPase activity result in erroneous incorporation of uracil by DNA polymerase into newly synthesized DNA, thus activating the base-excision repair pathway, which proceeds through single strand breaks (Tye & Lehman, 1977). If dUTP levels are abnormally high in a biological system, this repair system will expedite multiple and futile cycles of excision and reintroduction of uracil, resulting in increased recombination, DNA fragmentation and ultimately cell death (Richards *et al.*, 1984). Given the essentiality of the enzyme, its ubiquity and its involvement in a crucial step in nucleotide metabolism, dUTPase presents itself as an attractive potential target for chemotherapeutic, antiviral and antibacterial intervention.

Despite extensive structural and kinetic studies of various prokaryotic, eukaryotic and viral dUTPases, the detailed mechanism of how the enzyme facilitates the hydrolysis of dUTP is still poorly understood. However, it is known that all dUTPases studied so far are dependent on the presence of a divalent metal ion, preferably Mg^{2+}, for optimal activity (Greenberg & Somerville, 1962; Wohlrab & Francke, 1980; Camacho *et al.*, 2000). dUTPases are required to maintain exquisite specificity in choice of substrate, as imbalances in the nucleotide pool adversely affect accurate nucleotide metabolism. In particular, the enzyme is compelled to discriminate against the structurally most similar nucleotides, dTTP, dCTP and UTP, as they are all required for functional DNA or RNA synthesis. Kinetic studies do indeed demonstrate almost all known dUTPases to be highly specific for their substrate, dUTP (Larsson, Nyman *et al.*, 1996; Bergman *et al.*, 1998). Parasite-encoded dUTPases constitute the only known exception as they also accept dUDP, a potent inhibitor of other dUTPases, as a substrate (Camacho *et al.*, 2000).

Most of the amino-acid residues involved in substrate binding and catalysis have been identified through structural and mutational studies on various dUTPases. A large majority of dUTPases function as homotrimers. The structural layout of this group is illustrated by the four crystal structures determined so far, *i.e.* those of *Escherichia coli*, human, equine infectious anaemia virus and feline immunodeficiency virus (FIV) dUTPases in unliganded and liganded forms (Cedergren-Zeppezauer *et al.*, 1992; Larsson, Svensson *et al.*, 1996; Mol *et al.*, 1996; Prasad *et al.*, 1996, 2000; Dauter *et al.*, 1999). The homotrimers all contain three identical active sites formed at the subunit

interfaces, with each active site made up of residues contributed by all three subunits. One subunit provides residues vital for base and sugar recognition, while another is involved in phosphate interactions. The third subunit contributes a glycine-rich consensus motif positioned in the flexible C-terminal region, which is suggested to cover the active site and substrate during catalysis (Mol *et al.*, 1996). Within the homotrimer, each monomer contains five conserved motifs (McGeoch, 1990), which are all located at the active sites. Mammalian herpesvirus dUTPases function as monomers, still containing the conserved motifs, although in a rearranged order, but the polypeptide chain is longer. Owing to the similarities in sequence and metal-ion requirements between the monomeric and trimeric dUTPases it can be assumed that the major aspects of the reaction mechanism and perhaps also the layout of the active site are closely related. Parasite-encoded dUTPases, on the other hand, do not contain any traces of the conserved motifs and function as dimers (Camacho *et al.*, 2000).

In the current study, the *B. subtilis* dUTPase was chosen as a paradigm for studies of a dUTPase encoded by a Gram-positive bacterium. While this bacterial enzyme retains the general features of the five conserved motifs, some notable changes are observed in the amino-acid sequence in and adjacent to the motifs. The effects of these exchanges on the overall structure of the enzyme and on substrate recognition in particular require elucidation. Structural studies have therefore been initiated and here we report the crystallization and initial X-ray diffraction analysis of the enzyme. To facilitate crystallization and structural studies, the gene encoding *B. subtilis* dUTPase was cloned from chromosomal DNA, sequenced and expressed in *E. coli*. The recombinant product was purified in one step using cationic ion-exchange chromatography, as will be reported elsewhere. The enzyme is, as are most other bacterial dUTPases, active as a homotrimer and the amino-acid sequence of each monomer comprises 142 amino acids (16 kDa) as deduced from the nucleotide sequence of the dUTPase-encoding gene. Crystals of unliganded and potentially liganded enzyme were obtained, the latter grown in the presence of a substrate analogue and inhibitor, dUDP, and a metal ion, Sr^{2+}. The structure of such a complex should give insight into the nature of molecular interactions between the enzyme and its substrate and define residues participating in catalysis. The crystals proved to be suitable for X-ray diffraction analysis and molecular replacement is currently in progress using the dUTPase from FIV as a search model.

2. Materials and methods

2.1. Crystallization

The recombinant *B. subtilis* dUTPase was overexpressed in *E. coli* and purified to apparent homogeneity using phosphocellulose chromatography. Prior to crystallization, the pure enzyme was extensively dialysed against 10 m*M* MOPS pH 7 containing 1 m*M* β-mercaptoethanol and concentrated to 20 mg ml^{-1} using carboxymethylcellulose. Purity was checked by SDS–PAGE. The final protein concentration was estimated from the absorbance at 280 nm using a calculated extinction coefficient of $A_{280,0.1\%} = 1.101$.

Initial screening for nucleation and crystallization conditions was performed using sitting-drop vapour diffusion (Davies & Segal, 1971) at 290 K. Suitable conditions were identified using the commercially available sparse-matrix screening kits Crystal Screen and Crystal Screen II from Hampton Research (Jancarik & Kim, 1991). The conditions included in the 'footprint' screen (Stura *et al.*, 1992) were also investigated and then optimized. In all cases the droplets contained 0.25–1.25 µl protein, giving a protein concentration in the final 10 µl drop of 0.5–2.5 mg ml^{-1}, and 5 µl reservoir solution. Complex formation with dUDP and Sr^{2+} was enabled by inclusion of the substrate analogue and the metal salt ($SrCl_2$) to final concentrations of 5 and 20 m*M*, respectively, in the droplets. The droplets were equilibrated against a reservoir volume of 1 ml.

2.2. Data collection and processing

Prior to low-temperature data collection, crystals of unliganded enzyme were cryoprotected through sequential transfer through solutions of the mother liquor in which the polyethylene glycol (PEG) 600 concentration was increased in small steps to a final concentration of 30%. Crystals obtained from a drop containing 30% PEG 600 required no further addition of cryoprotectant and were harvested and vitrified directly from the drop. For vitrification, the crystals were mounted in loops and frozen in liquid nitrogen. A similar procedure was used for crystals of the putative complex, where a final concentration of 20% ethylene glycol was used.

X-ray diffraction data for *B. subtilis* dUTPase were measured using synchrotron radiation. Data for the orthorhombic and hexagonal crystal forms of unliganded enzyme and for the putatively liganded

Table 1
Crystallographic parameters and data-collection statistics.

	Native I (cubic)	Native II (orthorhombic)	Native III (hexagonal)	Potential complex with dUDP/Sr^{2+} (orthorhombic)
Beamline	14.1 (SRS)	14.1 (ESRF)	14.1 (ESRF)	14.1 (ESRF)
Image-plate system	ADSC Quantum 4R CCD	165 mm MAR CCD	165 mm MAR CCD	165 mm MAR CCD
Crystal-to-detector distance (mm)	170.1	140.0	130.0	139.2
Oscillation per image (°)	0.8	0.5	0.5	0.3
No. of images	146	180	180	620
Temperature (K)	100	100	100	100
Wavelength (Å)	1.244	0.934	0.934	0.934
Space group	$P2_13$	$P2_12_12$	$P6_3$	$P2_12_12_1$
Unit-cell parameters (Å,°)	$a = b = c = 116.9$, $\alpha = \beta = \gamma = 90$	$a = 116.0$, $b = 118.4$, $c = 60.8$, $\alpha = \beta = \gamma = 90$	$a = b = 102.8$, $c = 86.2$, $\alpha = \beta = 90$, $\gamma = 120$	$a = 99.5$, $b = c = 99.3$, $\alpha = \beta = \gamma = 90$
Asymmetric unit (monomers)	3	6	4	6
Packing density (V_M) (Å^3 Da^{-1})	2.8	2.2	2.0	2.5
Solvent content (%)	54.9	42.5	39.0	51.1
All data				
Resolution range (Å)	20–2.50	20–1.90	20–1.75	20–1.90
Total observations	487219	211897	303351	404962
Unique reflections	21068	67720	52846	78511
Completeness (%)	100	99.6	99.9	99.8
$I/\sigma(I)$	28.7	14.3	26.4	25.4
R_{merge}† (%)	9.0	8.3	6.1	5.4
Last shell				
Resolution range (Å)	2.59–2.50	1.97–1.90	1.81–1.75	1.97–1.90
Completeness (%)	100	99.6	100	99.0
$I/\sigma(I)$	7.34	2.1	3.0	3.1
R_{merge}† (%)	41.5	57.5	69.9	45.6

† $R_{merge} = \sum_{hkl} |I - \langle I \rangle| / \sum_{hkl} I$.

enzyme were collected at station 14.1 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France, while data for the cubic crystal form of unliganded enzyme were collected at station 14.1 at the Synchrotron Radiation Source (SRS), Daresbury, England. Table 1 specifies the details of the data-collection parameters used for each crystal form. The space groups and unit-cell parameters were determined using the autoindexing and parameter-refinement procedures of *DENZO* and all intensity data were integrated and reduced with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

Both uncomplexed and putatively complexed *B. subtilis* dUTPase crystallized rapidly (1–10 d) and reproducibly under several conditions found in the initial sparse-matrix and 'footprint' screens. Presence of PEG (of any molecular weight), ethylene glycol or 2-methyl-2,4-pentanediol (MPD) in the reservoir solution proved essential for nucleation and rapid crystal formation. The most promising crystals of unliganded enzyme were initially found using conditions included in a modified 'footprint' screen where crystals shaped as irregular plates were grown from a reservoir of 0.1 *M* imidazolemalate buffer pH 5.5, 33%(*v*/*v*) PEG 400. After optimization of pH and PEG concentration and size, good-quality single crystals of an orthorhombic crystal form used in data collection were grown from 0.1 *M* imidazolemalate buffer pH 4.0, 20%(*v*/*v*) PEG 600. Two other crystal forms, the cubic and hexagonal forms, on which data were also measured, were grown from 0.1 *M* imidazolemalate buffer pH 5.5 with 24 and 30%(*v*/*v*) PEG 600, respectively. Thus, subtle changes in PEG concentrations resulted in the appearance of crystals belonging to different forms. The size and shape of the crystals obtained varied notably even within each single drop for all three forms.

Crystals of putatively liganded enzyme suitable for X-ray diffraction studies were obtained using 0.1 *M* HEPES pH 7.5, 10%(*w*/*v*) PEG 8000 and 8%(*v*/*v*) ethylene glycol as a reservoir solution. These crystals were well shaped and had a typical edge length of 0.25–1 mm. A representative view of this crystal form is shown in Fig. 1.

Glycerol was initially tested as a cryoprotectant for all crystal forms but proved detrimental to the diffraction properties. Instead, stepwise increase of polyethylene glycol (for crystals of unliganded enzyme) or ethylene glycol (for crystals of putatively complexed enzyme) in the respective mother liquors was successfully used for cryoprotection.

The statistics for the X-ray diffraction data on all four crystal forms are summarized in Table 1. Initial efforts to construct a reliable model suitable for molecular-replacement experiments used as search models the *E. coli* and equine infectious anaemia virus dUTPase structures, which

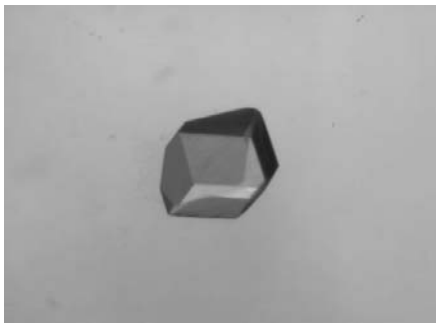

Figure 1
A crystal of *B. subtilis* dUTPase grown in the presence of dUDP and Sr^{2+}. The edge length of the crystal is approximately 0.5 mm.

share ~28 and ~21% sequence identity with the *B. subtilis* enzyme, respectively. However, this proved unsuccessful and the structure of dUTPase from FIV was tried instead (PDB code 1dut). This enzyme shares ~26% sequence identity with the *B. subtilis* dUTPase and a useful molecular-replacement solution was readily found; structure solution is now well under way for the hexagonal native crystal form.

In summary, the establishment of suitable cryoprotectant conditions together with synchrotron-radiation facilities enabled the collection of high-resolution diffraction data for crystals of *B. subtilis* dUTPase in unliganded and potentially liganded (with a substrate analogue and a metal ion) forms. In the past, crystallization of other trimeric dUTPases has not proved to be entirely straightforward (Persson *et al.*, 1997; Dauter *et al.*, 1998). Future structural studies of a dUTPase will therefore be aided by the ease with which high-diffraction quality crystals can be reproducibly generated with the *B. subtilis* enzyme under the conditions established in this study. It allows the determination of the structures of this dUTPase in complex with various substrate analogues and inhibitors, thereby providing further understanding of the features responsible for this essential enzyme's catalytic capability.

The authors are grateful to Dr Jim Brannigan and Dr Marek Brzozowski, YSBL, University of York, England for providing chromosomal *B. subtilis* DNA and advice on cryoprotectants, respectively. This work was supported by the European Commission by means of a Research Training Grant, contract number ERBBMH4CT985105 (to RP) and BIOMED contract number BMH4-CT97-2711. The YSBL is supported by a grant, number 87/SB09829, from the BBSRC as a Structural Biology Centre.

References

Bergman, A.-C., Nyman, P. O. & Larsson, G. (1998). *FEBS Lett.* **441**, 327–330.
Bertani, L. E., Häggmark, A. & Reichard, P. (1961). *J. Biol. Chem.* **236**, PC67–PC68.
Camacho, A., Hidalgo-Zarco, F., Bernier-Villamor, V., Ruiz-Perez, L. M. & Gonzalez-Pacanowska, D. (2000). *Biochem. J.* **346**, 163–168.
Cedergren-Zeppezauer, E. S., Larsson, G., Nyman, P. O., Dauter, Z. & Wilson, K. S. (1992). *Nature (London)*, **355**, 740–743.
Dauter, Z., Persson, R., Rosengren, A. M., Nyman, P. O., Wilson, K. S. & Cedergren-Zeppezauer, E. S. (1999). *J. Mol. Biol.* **285**, 655–673.
Dauter, Z., Wilson, K. S., Larsson, G., Nyman, P. O. & Cedergren-Zeppezauer, E. S. (1998). *Acta Cryst.* D**54**, 735–749.
Davies, D. R. & Segal, D. M. (1971). *Methods Enzymol.* **22**, 266–269.
Greenberg, G. R. & Somerville, R. L. (1962). *Proc. Natl Acad. Sci. USA*, **48**, 247–256.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kornberg, A. & Baker, T. A. (1991). *DNA Replication*, 2nd ed. New York: Freeman Press.
Larsson, G., Nyman, P. O. & Kvassman, J.-O. (1996). *J. Biol. Chem.* **271**, 24010–24016.
Larsson, G., Svensson, L. A. & Nyman, P. O. (1996). *Nature Struct. Biol.* **3**, 532–538.
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.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Persson, R., Rosengren, A. M., Nyman, P. O., Dauter, Z. & Cedergren-Zeppezauer, E. S. (1997). *Protein Pept. Lett.* **4**, 145–148.
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. & Stout, C. D. (1996). *Protein Sci.* **5**, 2429–2437.
Richards, R. G., Sowers, L. C., Laszlo, J. & Sedwick, W. D. (1984). *Adv. Enzyme Regul.* **22**, 157–185.
Stura, E. A., Nemerow, G. R. & Wilson, I. A. (1992). *J. Cryst. Growth*, **122**, 273–285.
Tye, B.-K. & Lehman, I. R. (1977). *J. Mol. Biol.* **117**, 293–306.
Wohlrab, F. & Francke, B. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 1872–1876.