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

Overexpression, crystallization and preliminary X-ray crystallographic analysis of the RNA polymerase domain of primase from *Streptococcus mutans* strain UA159

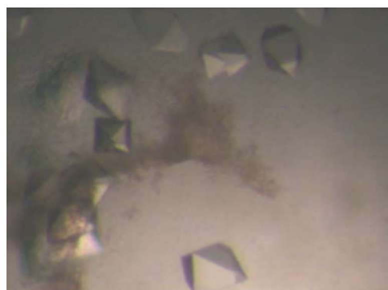
Dong-Won Im, Tae-O Kim, Ha Yun Jung, Ji Eun Oh, Se Jin Lee and Yong-Seok Heo*

Department of Chemistry, Konkuk University, Hwayang-dong, Gwangjin-gu, Seoul 143-701, Republic of Korea

Correspondence e-mail: ysheo@konkuk.ac.kr

Received 30 September 2011
Accepted 23 November 2011

Primase is the enzyme that synthesizes RNA primers on single-stranded DNA during normal DNA replication. In this study, the catalytic core domain of primase from *Streptococcus mutans* UA159 was overexpressed in *Escherichia coli*, purified and crystallized. Diffraction data were collected to 1.60 Å resolution using a synchrotron-radiation source. The crystal belonged to space group $P4_1$ or $P4_3$, with unit-cell parameters $a = b = 52.63$, $c = 110.31$ Å. The asymmetric unit is likely to contain one molecule, with a corresponding V_M of 1.77 $Å^3$ Da^{-1} and a solvent content of 30.7%.



© 2012 International Union of Crystallography
All rights reserved

1. Introduction

DNA replication is a complicated process that requires a large molecular machine known as the replisome, comprising one helicase and three copies of both the primase and the polymerase. Because the unwound strands of DNA are antiparallel, the replisome must accommodate two distinct modes of operation. The leading strand is synthesized in the 5′→3′ direction continuously in the same direction as the replication fork. The other new strand, the lagging strand, is also synthesized in the 5′→3′ direction. However, it can only be made discontinuously, as Okazaki fragments, as single-stranded parental DNA becomes newly exposed as the replication fork. During bacterial DNA replication, DNA polymerase catalyzes the addition of nucleotides to the 3′-OH group of the sugar moiety. Because DNA polymerases are incapable of *de novo* synthesis, a short oligonucleotide called a primer is required to provide an exposed 3′-OH group for the initiation of DNA synthesis. Primase is a single-stranded DNA-dependent RNA polymerase required for the initiation of DNA replication by synthesizing RNA primers, which can be subsequently elongated by DNA polymerases (Frick & Richardson, 2001). During lagging-strand synthesis, primase binds transiently with the replicative helicase and thereby limits the size of the nascent primers to ~9–14 nt, and this association of helicase and primase is an important driving force to initiate a new cycle of Okazaki fragment synthesis and to terminate the preceding fragment synthesis (Zechner *et al.*, 1992).

Bacterial primases are composed of three domains: an N-terminal zinc-binding domain, an RNA polymerase catalytic domain and a C-terminal helicase-binding domain (Tougu *et al.*, 1994). To date, crystal structures have been determined of the bacterial RNA polymerase catalytic domains of primase from *Escherichia coli* and *Aquifex aeolicus*, which show sequence identities of 35 and 32% to the domain of *Streptococcus mutans*, respectively (Keck *et al.*, 2000; Podobnik *et al.*, 2000; Corn *et al.*, 2005, 2008). These structures showed an active-site architecture that is unrelated to other DNA or RNA polymerase folds, which approximate the shape of a right hand with domains that are referred to as the fingers, the thumb and the palm domain, holding DNA with high processivity and fidelity. This explains the functional distinction of primase from 'classic' polymerases, which include reduced processivity and lower fidelity of primase. However, no structure of the primase catalytic domain from a Gram-positive bacterium has yet been determined.

Table 1
Data-collection statistics.

Values in parentheses are for the last shell.

Synchrotron	PF, beamline BL-5A
Wavelength (Å)	1.00
Resolution range (Å)	50.0–1.60 (1.63–1.60)
Space group	$P4_1$ or $P4_3$
Unit-cell parameters (Å)	$a = b = 52.63$, $c = 110.31$
No. of unique reflections	39133
Multiplicity	6.9 (6.5)
Completeness (%)	99.1 (99.5)
Molecules per asymmetric unit	1
$V_{\rm M}$ (Å^3 Da^{-1})	1.77
Solvent content (%)	30.7
Average $I/\sigma(I)$	49.2 (5.4)
$R_{\rm merge}$† (%)	3.7 (29.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection *hkl* and $\langle I(hkl)\rangle$ is the average intensity of reflection *hkl*.

To provide further structural data with regard to the catalytic mechanism of primase and the structural characteristics of the protein from Gram-positive bacteria, we have crystallized and performed X-ray crystallographic experiments on the catalytic core domain of primase from the Gram-positive bacterium *S. mutans*, which is most strongly associated with dental caries.

2. Materials and methods

2.1. Cloning, protein expression and purification

The gene fragment encoding the catalytic core domain (amino acids 105–380) of primase was amplified from genomic DNA of *S. mutans* strain UA159 by the polymerase chain reaction (PCR) using specific primers. The forward primer contained an *Nde*I restriction site (bold) and had the sequence 5′-GGG CCC **CAT ATG** AAA AAG GAA CAT CCA CAT CAA-3′, while the reverse primer contained a *Xho*I site (bold) and had the sequence 5′-GGG CCC **CTC GAG** ATC AAT ATT TTC AGG TTT TAG ATG TT-3′. The PCR product was then subcloned between *Nde*I and *Xho*I sites of pET-22b vector (Novagen, USA). This construct contains an additional hexahistidine tag (LEHHHHHH) at the C-terminus for purification purposes. The recombinant plasmid was transformed into *E. coli* strain BL21 (DE3) (Novagen) and the cells were grown in a shaking incubator at 310 K in LB broth medium supplemented with 50 µg ml^{-1} ampicillin. Protein expression was induced by adding 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) when the cells reached an optical density at 600 nm of about 0.6 and the cells were cultured at 293 K for ~16 h. Cultured cells were harvested by centrifugation at 3000*g* for 30 min at 277 K. The cell pellet was resuspended in binding buffer (20 m*M* Tris pH 8.0, 300 m*M* NaCl) and disrupted by sonication at 277 K. The crude lysate was centrifuged at 25 000*g* for 1 h at 277 K. The supernatant was then loaded onto an Ni^{2+}-chelated HisTrap HP column (GE Healthcare, USA) which had been pre-equilibrated with binding buffer. After washing with wash buffer (20 m*M* Tris pH 8.0, 300 m*M* NaCl, 50 m*M* imidazole), the bound protein was eluted with elution buffer (20 m*M* Tris pH 8.0, 300 m*M* NaCl, 400 m*M* imidazole). The eluted protein was dialyzed for 6 h at 277 K in buffer *A* (20 m*M* Tris pH 7.0, 30 m*M* NaCl) and loaded onto a HiTrap Heparin HP column (GE Healthcare, USA) which had been pre-equilibrated with buffer *A*. The bound protein was eluted with a buffer consisting of 20 m*M* Tris pH 7.0 and 1.0 *M* NaCl and was then dialyzed for 6 h at 277 K in buffer consisting of 20 m*M* Tris pH 7.0, 0.3 *M* NaCl, 5 m*M* DTT and 7% glycerol. The purified protein was concentrated to 25 mg ml^{-1} by Bradford assay and the purity of the protein was examined by 12% SDS–PAGE and determined to be >95% pure.

2.2. Crystallization and data collection

Crystallization of the protein was initiated by crystal screening at 293 K by the hanging-drop vapour-diffusion method using 24-well VDX plates (Hampton Research, USA) with a drop consisting of 1 µl protein solution concentrated in the dialysis buffer and 1 µl well solution equilibrated against 500 µl well solution. Commercial screening kits from Hampton Research and Emerald BioSystems (Crystal Screen, Crystal Screen 2, Index, SaltRx and Wizard I and II) were used for the preliminary screening. Suitable-sized crystals were obtained within a week under the following condition: 0.05 *M* Na HEPES pH 7.0, 0.005 *M* magnesium chloride, 25% PEG MME 550 (Fig. 1). The crystals were cryoprotected by soaking them for 3 s in a

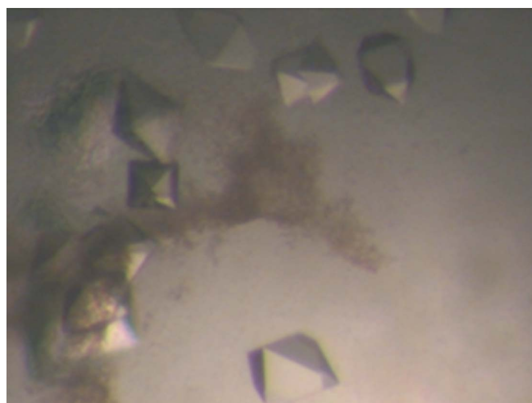


Figure 1
A crystal of the catalytic core domain of primase from *S. mutans* strain UA159 grown using 0.05 *M* Na HEPES pH 7.0, 0.005 *M* magnesium chloride, 25% PEG MME 550. The crystal dimensions are approximately 0.2 × 0.1 × 0.1 mm.

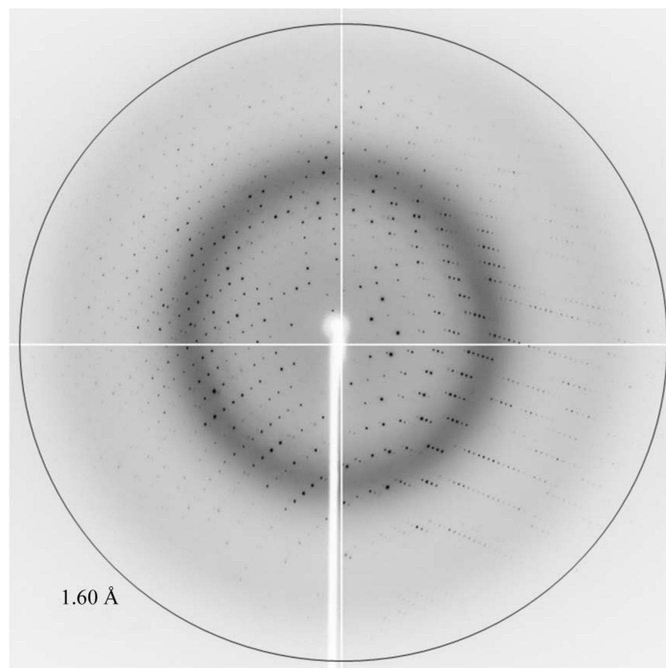


Figure 2
A diffraction image of the crystal of the catalytic core domain of primase from *S. mutans* strain UA159. The resolution limit is indicated as a circle with its value (1.60 Å).

cryoprotectant solution containing an additional 20%(*v*/*v*) ethylene glycol and were cryocooled in liquid nitrogen. Frozen crystals were mounted on the goniometer in a stream of cold nitrogen at 100 K. X-ray diffraction data were collected from a frozen crystal using an ADSC Quantum CCD 210r detector on beamline BL-5A at the Photon Factory, Japan. A total rotation range of 180° was covered with 1.0° oscillations and 5 s exposure per frame. The wavelength of the synchrotron X-ray beam was 1.00 Å and the crystal-to-detector distance was set to 140 mm. X-ray diffraction data were collected to 1.60 Å resolution (Fig. 2). Data were indexed, integrated, scaled and merged using the *HKL*-2000 software package (Otwinowski & Minor, 1997).

3. Results and discussion

The RNA polymerase domain (amino acids 105–380) of primase from *S. mutans* strain UA159 was cloned, overexpressed, purified and crystallized for structural studies. X-ray diffraction data from the crystal indicated that it belonged to space group $P4_1$ or $P4_3$ on the basis of systematic absences, with unit-cell parameters $a = b = 52.63$, $c = 110.31$ Å. Data-collection statistics are provided in Table 1. The Matthews coefficient suggested the presence of one molecule in the crystallographic asymmetric unit, with a V_M of 1.77 Å^3 Da^{-1} and a solvent content of 30.7% (Matthews, 1968). The molecular-replacement method was used in an attempt to solve the structure using the crystal structures of the catalytic core domain of *E. coli* primase (PDB entries 1dd9, 1eqn and 3b39; Keck *et al.*, 2000; Podobnik *et al.*, 2000; Corn *et al.*, 2008) and *A. aeolicus* primase (PDB entry 2au3; Corn *et al.*, 2005), which share 35% and 32% sequence identity, respectively. However, none of our attempts provided a clear solution. Therefore, SeMet-labelled crystals were obtained under the same condition as used for the native crystals and the phase problem will be solved by the MAD method (Hendrickson *et al.*, 1990) using selenium as an anomalous scatterer.

We thank Dr Y. Yamada at beamline BL-5A at Photon Factory for his assistance. This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST; No. 2011-0003605) and Pohang Accelerator Laboratory (PAL).

References

Corn, J. E., Pease, P. J., Hura, G. L. & Berger, J. M. (2005). *Mol. Cell*, **20**, 391–401.
Corn, J. E., Pelton, J. G. & Berger, J. M. (2008). *Nature Struct. Mol. Biol.* **15**, 163–169.
Frick, D. N. & Richardson, C. C. (2001). *Annu. Rev. Biochem.* **70**, 39–80.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Keck, J. L., Roche, D. D., Lynch, A. S. & Berger, J. M. (2000). *Science*, **287**, 2482–2486.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **277**, 307–326.
Podobnik, M., McInerney, P., O'Donnell, M. & Kuriyan, J. (2000). *J. Mol. Biol.* **300**, 353–362.
Tougu, K., Peng, H. & Marians, K. J. (1994). *J. Biol. Chem.* **269**, 4675–4682.
Zechner, E. L., Wu, C. A. & Marians, K. J. (1992). *J. Biol. Chem.* **267**, 4045–4053.