Biochemical and Biophysical Research Communications 483 (2017) 52–57

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Crystal structure of Deep Vent DNA polymerase

Yasushi Hikida [a, b, c], Michiko Kimoto [a, d, e], Ichiro Hirao [a, d, e, **], Shigeyuki Yokoyama [a, b, c, *]

[a] RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan
[b] Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[c] RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan
[d] Division of Structural and Synthetic Biology, RIKEN Center for Life Science Technologies, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan
[e] Institute of Bioengineering and Nanotechnology, 31 Biopolis Way, The Nanos, #04-01, 138669, Singapore

ARTICLE INFO

Article history:
Received 22 December 2016
Accepted 3 January 2017
Available online 4 January 2017

Keywords:
DNA polymerase
Unnatural base pair
X-ray crystallography
Polymerase chain reaction

ABSTRACT

DNA polymerases are useful tools in various biochemical experiments. We have focused on the DNA polymerases involved in DNA replication including the unnatural base pair between 7-(2-thienyl)imidazo[4,5-*b*]pyridine (**Ds**) and 2-nitro-4-propynylpyrrole (**Px**). Many reports have described the different combinations between unnatural base pairs and DNA polymerases. As an example, for the replication of the **Ds**–**Px** pair, Deep Vent DNA polymerase exhibits high efficiency and fidelity, but Taq DNA polymerase shows much lower efficiency and fidelity. In the present study, we determined the crystal structure of Deep Vent DNA polymerase in the apo form at 2.5 Å resolution. Using this structure, we constructed structural models of Deep Vent DNA polymerase complexes with DNA containing an unnatural or natural base in the replication position. The models revealed that the unnatural **Ds** base in the template-strand DNA clashes with the side-chain oxygen of Thr664 in Taq DNA polymerase, but not in Deep Vent DNA polymerase.

1. Introduction

DNA polymerases are widely used in molecular biology techniques, such as polymerase chain reaction (PCR) and sequencing. DNA polymerases are divided into 6 subfamilies (A, B, C, D, X, and Y) [1–3]. Family B, a large and well-studied family, contains single-subunit DNA polymerases from Archaea, Bacteria, and Eukarya. In studies of unnatural base pairs, DNA polymerases are used in PCR to amplify DNAs containing the unnatural bases [4–7]. One of the highly specific, unnatural base pairs that functions as a third base pair for PCR amplification is the 7-(2-thienyl)imidazo[4,5-*b*]pyridine (**Ds**) and 2-nitro-4-propynylpyrrole (**Px**) pair (Fig. 1a) [8]. For the PCR amplification of DNA fragments involving the **Ds**–**Px** pair, the amplification fold and the retention rate of **Ds**–**Px** were examined by using the family-B DNA polymerases: Deep Vent, AccuPrime Pfx, Pfx 50, Phusion HF, Pfu, Pwo SY, Vent, and N9°, and the family-A DNA polymerases: TITANIUM Taq and Taq [9]. Among them, Deep Vent DNA polymerase, a thermostable archaeal family-B DNA polymerase from *Pyrococcus* strain GB-D [10], was found to be highly specific with a low misincorporation rate of 0.005%/bp/replication, and highly efficient with an amplification rate of ~10^{10}-fold by 40 cycles of PCR [9]. In contrast, Taq DNA polymerase exhibited much lower selectivity in PCR [9].

In the present study, we determined the crystal structure of Deep Vent DNA polymerase at 2.5 Å resolution, and examined why Deep Vent DNA polymerase is superior to Taq DNA polymerase with respect to the PCR amplification of the **Ds**–**Px** pair. Model structures of its complexes with the template DNA and the incoming nucleoside triphosphate containing the **Ds**–**Px** pair clarified that neither the template **Ds** base nor the incoming **Px** base causes a steric clash with Deep Vent DNA polymerase. In contrast, the template **Ds** base sterically clashed with the side chain of Thr664 in Taq DNA polymerase, which should be one of the reasons for its much lower amplification and retention rates. We then compared our **Ds**–**Px** model with another unnatural base pair, **NaM**–**5SICS** (Fig. 1b) [11], in the crystal structure of its complex with Taq DNA polymerase [4].

* Corresponding author. RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
** Corresponding author. Institute of Bioengineering and Nanotechnology, 31 Biopolis Way, The Nanos, #04-01, 138669, Singapore.
E-mail addresses: ichiro@ibn.a-star.edu.sg (I. Hirao), yokoyama@riken.jp (S. Yokoyama).

http://dx.doi.org/10.1016/j.bbrc.2017.01.007

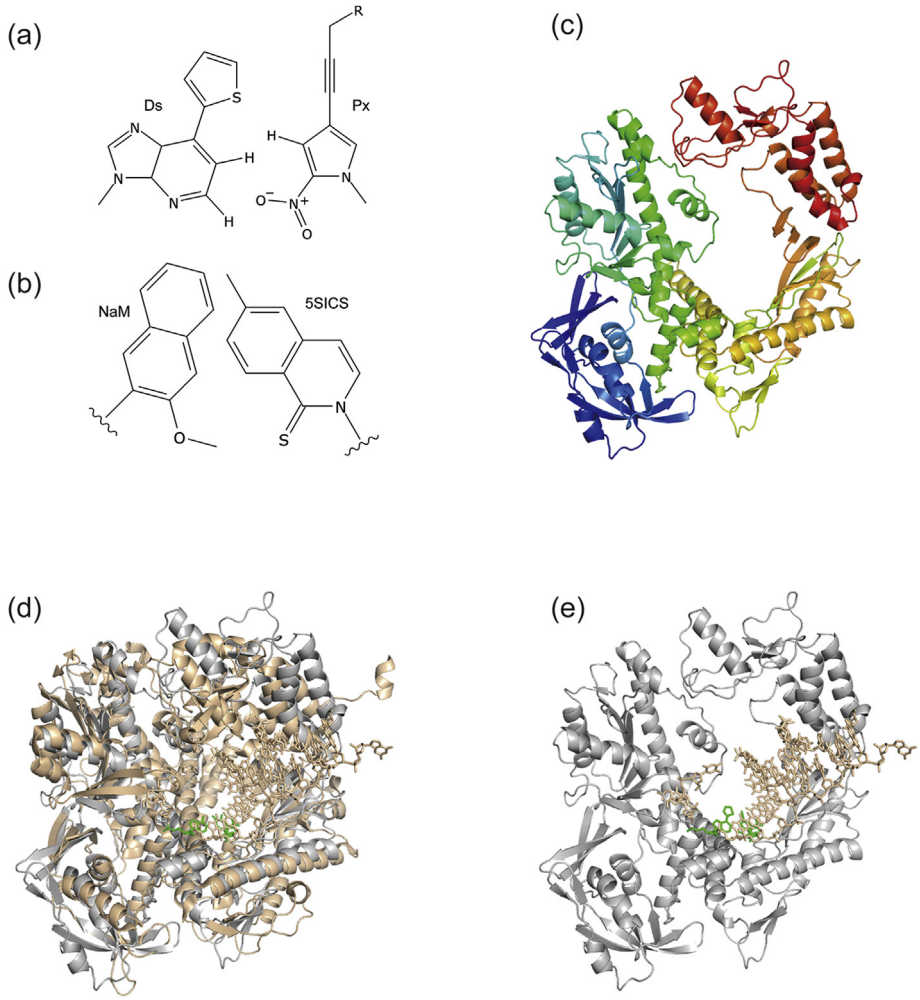

Fig. 1. Overall structure of Deep Vent DNA polymerase. (a) The **Ds**–**Px** pair. In this article, R = H in **Px**. (b) The **NaM**–**5SICS** pair. (c) Overall structure of Deep Vent DNA polymerase in a cartoon model, colored blue to red from the N-terminus to the C-terminus. (d) Superimposition of Deep Vent DNA polymerase (gray) and RB69 DNA polymerase in complex with the non-template DNA, the template DNA with A at the templating position, the non-template DNA, and the incoming deoxythymidine triphosphate (base pair: green, overall: wheat). (e) A model of Deep Vent DNA polymerase in complex with the non-template DNA, the template DNA with **Ds** (green) at the templating position, and the incoming **Px** triphosphate (d**Px**TP, green). These figures were rendered with PyMol, Version 1.4.1 [23].

2. Materials and methods

2.1. Deep Vent DNA polymerase

Deep Vent DNA polymerase from *Pyrococcus* Strain GB-Dt [10] was prepared in the same manner as the commercial product (Deep Vent DNA Polymerase, Catalog #M0258L, New England BioLabs, Inc., https://www.neb.com/products/m0258-deep-ventr-dna-polymerase). The sample was dialyzed against the protein buffer [10 mM Tris–HCl buffer (pH 7.6) containing 100 mM KCl, 1 mM DTT, and 0.1 mM EDTA], and was concentrated to 9.6 mg/ml.

2.2. Crystallization, data collection, and structure determination

The crystals of Deep Vent DNA polymerase were obtained by the sitting-drop vapor diffusion method in 96-well plates at 293 K, by mixing 500 nL of the protein solution with 500 nL of the crystallization buffer [number 18 of the MembFac crystal screening kit (Hampton Research) or 0.1 M sodium citrate tribasic dehydrate (pH 5.6), 0.1 M lithium sulfate monohydrate, and 12% (w/v) polyethylene glycol 6000].

X-ray diffraction data were collected at 100 K from crystals flash-cooled in a cryo-protectant solution, prepared by mixing equal volumes of the 2 × protein buffer [20 mM Tris–HCl buffer (pH 7.6) containing 200 mM KCl, 2 mM DTT, and 0.2 mM EDTA] and the 2 × crystallization buffer [0.2 M sodium citrate tribasic dehydrate (pH 5.6), 0.2 M lithium sulfate monohydrate, 24% (w/v) polyethylene glycol 6,000, and 64% (w/v) glycerol], at beamline BL5A of the Photon Factory (KEK, Tsukuba, Japan), and were processed using the HKL2000 programs.

The crystals belonged to the space group $P2_12_12_1$, with unit-cell parameters $a = 77.0$, $b = 112.0$, $c = 111.8$ Å. The phase was determined by molecular replacement, using the program MOLREP from the CCP4 suite [12] and the coordinates of DNA polymerase from *Pyrococcus furiosus* (PDB ID: 2JGU) [13] as the search model. The atomic model was built and adjusted using the program Coot [14], with reference to the coordinates of the DNA polymerase from *Pyrococcus furiosus* (PDB ID: 4AHC) [15], and was refined using CNS [16,17]. According to the wwPDB X-ray Structure Validation [18], the final model has 96% of the residues in the favored regions of the Ramachandran plot, 3% in the allowed regions, and 1% in the outlier regions. The asymmetric unit contains one protein molecule, with a calculated solvent content of 54%. The statistics for the X-ray diffraction data and the refined structure are provided in Table 1. The coordinates and the structure factors have been deposited in the Protein Data Bank, with the PDB ID: 5H12.