Archives of Biochemistry and Biophysics 606 (2016) 143–150

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

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

Processing of A-form ssDNA by cryptic RNase H fold exonuclease PF2046

Junsoo Kim [a, b, 1], Gerelt-Od Sambalkhundev [a, b, 1], Sulhee Kim [a], Jonghyeon Son [a], Ah-reum Han [a], Sul-Min Ko [a], Kwang Yeon Hwang [a, *], Woo Cheol Lee [a, b, **]

[a] Division of Biotechnology, Korea University, Anam-Dong, Seongbuk-gu, Seoul 136-713, Republic of Korea
[b] Institute for Life Sciences and Natural Resources, Korea University, Seoul 136-713, Republic of Korea

ARTICLE INFO

Article history:
Received 2 May 2016
Received in revised form
29 July 2016
Accepted 1 August 2016
Available online 3 August 2016

Keywords:
Exonuclease
ssDNA
A-form
RNase H

ABSTRACT

RNase H fold protein PF2046 of *Pyrococcus furiosus* is a 3′-5′ ssDNA exonuclease that cleaves after the second nucleotide from the 3′ end of ssDNA and prefers poly-dT over poly-dA as a substrate. In our crystal structure of PF2046 complexed with an oligonucleotide of four thymidine nucleotides (dT_4), PF2046 accommodates dT_4 tightly in a groove and imposes steric hindrance on dT_4 mainly by Phe220 such that dT_4 assumes the A-form. As poly-dA prefer B-form due to the stereochemical restrictions, the A-form ssDNA binding by PF2046 should disfavor the processing of poly-dA. Phe220 variants display reduced activity toward poly-dA and the A-form appears to be a prerequisite for the processing by PF2046.

1. Introduction

Pyrococcus furiosus is a hyperthermophilic archaeon and can grow optimally at 98 °C and tolerate temperature of 103 °C [1]. Owing to resistance to heat and radiation, *P. furiosus* is an attractive experimental model for DNA repair research. One study showed that *P. furiosus* can tolerate ionizing radiation because of an unknown DNA repair system [2]. In an effort to elucidate the mechanism of this robust DNA repair system, ssDNA-specific nucleases have been screened for these unusual properties, and thus PF2046 was identified as an ssDNA-specific exonuclease [3]. The structure of PF2046 has been determined previously and found to contain an RNase H fold [4], which was not inferred from its amino acid sequence. Because of this homology to the RNase H family of proteins, PF2046 was predicted to function as a ribonuclease. Structural analysis, however, did not lead to functional annotation and the function of PF2046 was identified later in the aforementioned study involving the search for DNA repair enzymes in *P. furiosus* [3]. In that study, the authors were looking for ssDNA-specific nucleases in a *P. furiosus* genomic library expressed in *Escherichia coli* and found PF2046 to have a 3′-5′ ssDNA-specific exonuclease activity. PF2046 has unique substrate specificity: it prefers a series of dT over a series of dC or dA. Furthermore, PF2046 shows an unusual catalytic mechanism, namely, it removes dinucleotide units from the 3′ end. Recently, structures of PhoExo I (PDB ID: 4YOX), a PF2046 homolog from *Pyrococcus abyssi* have been reported to explain the structural basis of two-nucleotide removal [5].

The RNase H fold is defined as three layers—α/β/α structure with a five-strand β-sheet—the second strand being antiparallel to the rest, according to the SCOP database [6]. RNase H is classified into three classes depending on the amino acid sequence homology [7]. RNases HI and HIII are Mg^{2+} dependent and cleave only the RNA strand of a RNA/DNA hybrid duplex. RNase HII is catalytically less efficient and nicks the 5′ side of a ribonucleotide present in DNA. Compared to RNase HI, RNase HII contains additional helices at the C terminus, which is involved in binding to a substrate [8]. RNase H containing proteins have other domains relevant for functions [9] and reaction mechanism follows two-metal ion catalysis [10]. In addition to simple ribonucleases, an RNase H fold is also found in the functional domain of larger proteins driving various cellular processes such as DNA replication, DNA repair, and post-transcriptional modifications involved in RNA interference [RNAi]

Abbreviations: dT_4, an oligo of four thymidine deoxynucleotides.

* Corresponding author.
** Corresponding author. Division of Biotechnology, Korea University, Anam-dong, Seongbuk-gu, Seoul 136-713, Republic of Korea.
E-mail addresses: chahong@korea.ac.kr (K.Y. Hwang), dreadco@gmail.com (W.C. Lee).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.abb.2016.08.001
0003-9861/

[11]. Phylogenetic analyses suggest that during evolution, RNase H proteins diverged to be readily detected by amino acid sequence homology; unexpected proteins with an RNase H fold are still being reported to take part in novel cellular processes. The best example is the RNase H fold proteins participating in RNAi, such as the argonaute protein of *P. furiosus* [12]. Its cryptic RNase H fold, also known as the Piwi domain, can bind double-stranded RNA and cleave complementary viral mRNA, thus implementing RNAi.

To understand the structural basis for the preference of PF2046 of a stretche of dTs over dAs, our aim was to determine the crystal structure of PF2046 complexed with poly-dT. The structure revealed that all three protomers of PF2046 bind ssDNA dT_4. Oligos bind mainly via interaction with sugar backbone atoms, and the N3 atom is hydrogen-bonded to the main-chain carbonyl group of Thr53, thus explaining the preference for thymidines over cytosine residues. The binding pocket for the 3′ nucleotide is flanked by Pro19 and Phe220, which are applying van der Waals strain to bases and give rise to the A-form DNA, which may pose further restriction on substrate binding.

2. Materials and methods

2.1. Overexpression and purification of PF2046

PF2046 was amplified from genomic DNA of *P. furiosus* DSM 3638 and ligated into pET-21a (Novagen, USA) resulting in an ORF with one additional Met residue at the N terminus of the protein. The expression vector was introduced into *E. coli* BL21 (DE3) and grown in 1.0 L of the Luria-Bertani (LB) medium in a baffled Erlenmeyer flask at 37 °C until optical density at 600 nm reached 0.6. The gene expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.25 mM final concentration) and the cells were further grown at 18 °C for 20 h and harvested. The cell pellet was washed with phosphate buffered saline (PBS) and stored at −20 °C until use. The cell pellet was thawed in 30 mL of PBS and lysed by VCX-750 sonicator (Sonics & Materials, USA) at room temperature (a 1-s pulse after a 5-s pause for 30 min at 50% power output), and the lysate was cleared by centrifugation at 29,693*g* for 1 *h*. The absence of refrigeration during the sonication elevated the lysate temperature to ~60 °C, which caused aggregation of heat-labile *E. coli* proteins; the resultant supernatant contained mostly PF2046. Nucleic acids in this fraction were removed by polyethyleneimine precipitation (final concentration 0.1%, w/v) and by centrifugation at 29,693*g* for 1 *h*. PF2046 was purified by anion exchange chromatography (Q-Sepharose HP 5 mL, GE Healthcare, USA) in 20 mM Tris-HCl buffer pH 8 using a gradient from 0 to 1 M NaCl. At the final purification step, PF2046 was isolated by size exclusion chromatography (Superdex S200, GE Healthcare, USA); the column was equilibrated and subjected to elution with the buffer 20 mM Tris pH 8 containing 100 mM NaCl. PF2046 was concentrated to ~25 mg/mL and stored at −80 °C until use. The protein concentration was measured by the Bradford method (Bio-Rad, USA).

2.2. Crystallization and determination of the crystal structure

PF2046 was crystallized in the presence of an oligodeoxynucleotide consisting of nine thymidine deoxynucleotides (dT_9). dT_9 was purchased in the lyophilized form (Cosmogenetech, Korea) and used without further purification. Prior to crystallization, PF2046 (1.13 mM in 20 mM HEPES pH 7, 100 mM NaCl) and dT9 (2.5 mM in H_2O) were mixed in the volume ratio of 3:2, yielding the ~1:2.5 M ratio, and the mixture was incubated on ice for 30 min. Crystallization conditions for the PF2046-ssDNA complex were selected by means of the PEGRx Crystal Screen Kit (Hampton Research, USA). Sitting-drop crystal screening was performed in a 96-well Intelli-Plate (Art Robbins Instruments, USA), where 0.5 μl of the PF2046-ssDNA complex was mixed with an equal volume of crystallization buffer from the kit. The crystals in the plate shape appeared after two days in the buffer composed of 0.1 M sodium malonate pH 8.0, 0.1 M Tris-HCl pH 8.0, and 30% (w/v) polyethylene glycol 1000 (PEG1000).

An X-ray diffraction dataset was collected at beamline BL-1A at the Photon Factory (Tsukuba, Japan) at wavelength 1.0000 Å. The crystal was mounted on a nylon loop and chilled to 100 K by a stream of nitrogen gas to reduce radiation damage. Due to the high concentration of PEG1000, the use of cryogenic buffer to prevent ice formation was not necessary. Diffraction images were processed and scaled by means of Mosflm [13] and Scala [14] in the CCP4 software suite [15]. The crystal structure was determined by molecular replacement in the MolRep software [16] using PF2046 structure without the nucleotide [PDB ID: 4O8U] as a search model. Crystal structure was modeled and refined in software packages XtalView [17] and CNS [18] (See Table 1). DNA conformation was analyzed in 3DNA [19] version 2.1. The final model coordinates were deposited in Protein Data Bank (PDB ID: 5CHI).

2.3. The exonuclease assay

During the purification of PF2046 and its variants for exonuclease assays, buffers were supplemented with 5 mM $MgCl_2$. In the poly-dA, poly-dC or poly-dT oligonucleotide digestion assay, 30-mer oligos labeled at the 5′ end by fluorescein isothiocyanate were purchased from Cosmogenetech (Republic of Korea) and used for fluorescence-based detection on a blue-LED transilluminator (IO Rodeo, USA). For exonuclease assays, 10 μl of reaction mix was composed of 20 mM Tris pH 8.0, 100 mM NaCl, 2 μM oligonucleotide and 5 nM PF2046 or its variants. The reaction proceeded at 75 °C for 10 min and was stopped by adding 2 μl of 500 mM EDTA pH 7. All reaction products were loaded and analyzed by 10% polyacrylamide gel electrophoresis (PAGE) in TAE (Tris, acetate, EDTA, pH 8.0) buffer. For pH dependent exonuclease assays, Tris pH

Table 1
Data collection and refinement statistics.

PF2046-dT_4	
Data collection	
PDB ID	5CHI
Space group	$P2_1$
Cell dimensions	a = 69.30, b = 85.96, c = 69.02 [Å], β = 118.4°
Resolution	50–2.47
No. of reflections	162,455
Unique reflections	24,745
Redundancy	6.57
Completeness [%]	96.5 [80.8]
I/σI	12.54 [2.10]
Rmeasure	0.145 [0.776]
CC[1/2] [%]	99.6 [73.8]
Refinement	
Resolution	34.5–2.47
R-factor	0.1824
R_{free}	0.2278
RMSD	
Bonds	0.010
Angles	1.358
B-factors	
Overall	44.66
Protein	44.19
dT_4	58.44
Mg^{2+}	41.35
Water	37.51

RMSD: root mean square deviation, PDB: Protein Data Bank.

8.0 was replaced with 20 mM MES pH 6.0 or 20 mM HEPES pH 7.0 buffer.

3. Results

3.1. Overall structure of PF2046 complexed with ssDNA (PF2046-dT_4)

The crystal structure of PF2046-dT_4 contains three PF2046 molecules in an asymmetric unit, which form a trimeric quaternary structure as observed previously in the apo structure (PDB ID: 4O8U) [4] (Fig. 1A). The three protomers is arranged in three-fold symmetry and interact by 875.7 Å^2 solvent-accessible area, approximately 8% of total surface area, as calculated by PISA program [20]. Each protomer contains a dT_4 at its active site (Fig. 1A). The structures of all protomers of the trimer are almost identical when superposed (RMSD for CA is 0.27–0.33 Å). Each protomer, however, showed different overall *B*-factors due to crystal contacts different from one another (data not shown). Among the three chains of polypeptides, chain A has the lowest overall *B*-factor.

During model building and refinement, we could recognize electron density of metal ions coordinated by four acidic residues in the presumed active site of chain A. Mg^{2+} was added to the activity assay in another study [3], whereas the divalent metal was not added during purification or crystallization; therefore, it is possible that cognate divalent metals, presumably Mg^{2+}, from the expression host *E. coli* may be the source of the electron density. As Mg^{2+} is a common divalent metal found in two-metal nucleases [10] and fit well into the density (data not shown), we modeled the density with Mg^{2+} ions. Besides, the presence of Mg^{2+} in the active site is less apparent for chains B and C. In chain B, Mg^{2+}-A has weak density, and in chain C, almost no density was observed (data not shown).

We could detect electron density only for four deoxynucleotides in the active site in the final F_o–F_c map. One study showed that PF2046 requires at least 4–6 nucleotides for activity [3]. According to that study, the oligonucleotide in the PF2046 structure is 4–6 nucleotides long, and nucleotides extending from the 5′ end, if present, are presumed to be in a random orientation. Accordingly, we will assume that a part of dT_9 corresponds to dT_4, and each thymidine is hereafter designated dT-1 through dT-4 from the 5′ end. When we compare our complex structure with apo PF2046 structure by Su et al. [4], the crystal structure of PF2046-dT_4 shows a conformational change due to the binding of the ligand (Fig. 1B). Helices A and C move toward dT_4 to accommodate the substrate. Two loops (L1 and L2) also change their conformation to accommodate the substrate.

3.2. Comparison of crystal structure of PF2046 and PhoExo I

Consistent with the amino acid sequence alignment between PF2046 and PhoExo I the structure of PF2046 is highly similar to that of PhoExo I (R.M.S.D of 0.421 Å) and sequence identity is about 77.2% (Fig. 1C). Likewise, oligonucleotides of PhoExo I interact in the same manner as PF2046. Residues of active site and binding pocket of between PF2046 and PhoExo I show that key residues located in same position (Fig. 1C). However, the guanidine moiety of Arg55 of PhoExo I is interacting with the pyrimidine ring of dT, while Arg55 in the corresponding position of PF2046 is facing the opposite direction and free of interaction with the ligand (Fig. 1D). Instead, the main chain carbonyl group of Thr53 in PF2046 provides hydrogen bonds for binding. The helices αB & αC flanking Ser102 of the PF2046 also deviate from PhoExo I. PhoExo I αB and αC are rather similar to apo structure of PhoExo I or PF2046.

3.3. The active site

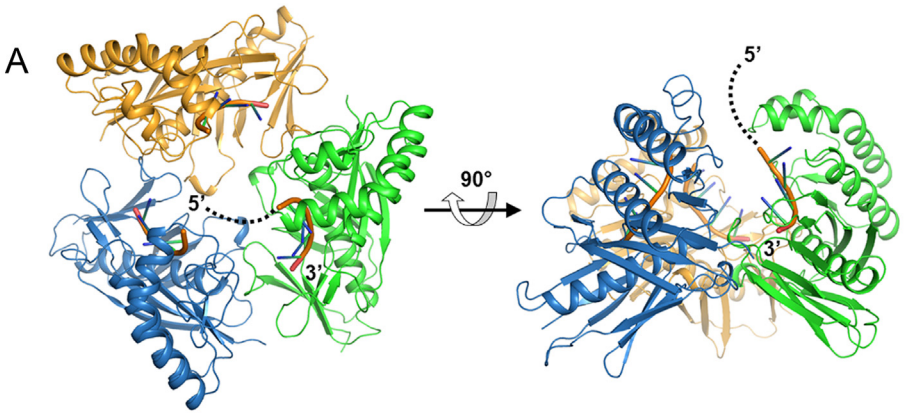
Among structures of RNase H-RNA/DNA duplex complexes, that from *B*acillus *halodurans* (PDB ID: 1ZBL) and LC11 (an RNase H1 from soil metagenomic library; PDB ID: 4H8K) were found to be the most similar to PF2046-dT_4 by using the DALI server. When compared to other RNase H, PF2046 has rearranged structures of helices C and D. Furthermore, PF2046 has an additional C-terminal greek-key motif for 3′-side end (Fig. 2A). Even though the structures shared overall similarity in terms of the RNase H fold, the substrate-binding mode showed a significant deviation. In particular, helices C and D in PF2046 show a notable departure from those of other RNase H fold proteins (Fig. 2A). In detail, helix C clashes with ssDNA in RNase H DNA-RNA heteroduplex and helix D correlates with the different ssDNA trajectory when compared to ssRNA in RNase H (Fig. 2A). These prominent features contribute to the binding of dT_4 instead of the binding of a DNA-RNA duplex by RNase H proteins. The distance between active-site Mg^{2+} ions and Phe220 (against which dT-4 is stacked) functions as a molecular ruler to determine the site of incision from the 3′ end (Fig. 1B). As a result, two nucleotides will be removed from the 3′ end to produce a dideoxynucleotide.

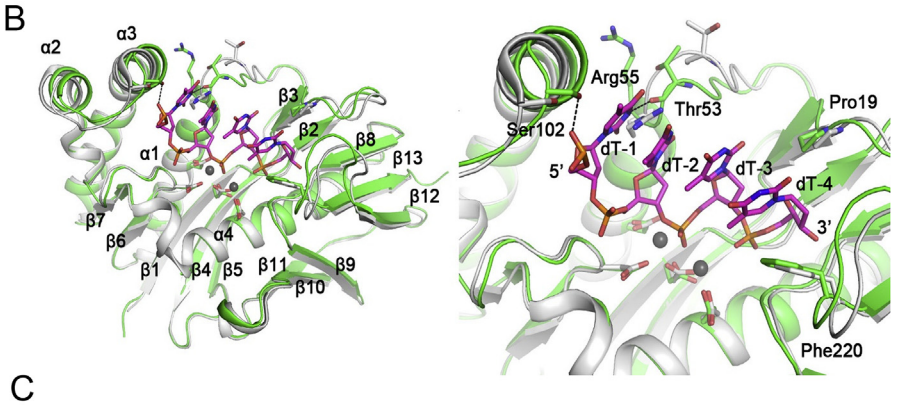
The active site contains four acidic amino acid residues coordinating two Mg^{2+} ions (Fig. 2B). The two metal ion centers are separated by 3.4 Å in chain A. Acidic residues in the active site are in the order Asp7-Glu61-Asp80-Glu145 (DEDE) as observed in RNase HIII (versus DEDD in RNase HI) (Fig. 2B). PhoExo I also has acidic residue and interact between acidic residue and ssDNA via Mg^{2+} ions [5]. To test whether these acidic amino acids are necessary for the enzymatic activity, we prepared two mutants—D7A and D80A—which coordinate Mg^{2+}-B that is interacting with the scissile bond and is thus supplying the water for hydrolysis. As expected, both variants completely lost the activity toward dT_{30} (Fig. 2C).

Compared to the metal binding by DNA-RNA duplex complexes of RNase Hs, PF2046 is notably different in terms of the Mg^{2+}-B binding. In contrast to direct coordination of Mg^{2+}-B in RNase HI, Glu61 of PF2046 is coordinating Mg^{2+}-B through two water molecules (Fig. 2D). We could identify a water molecule bound to Mg^{2+}-A, which may act as a nucleophile. Just as other RNase HI two-metal nucleases, PF2046 is lacking amino acids that would act as a hydrogen donor in the active site, and the water molecule bound to metal A is presumed to be the Lewis acid for activation (Fig. 2D) [21]. In the proposed reaction mechanism, a water molecule coordinated by Mg^{2+}-A may attack the phosphate to break the bond between the leaving group oxygen atom and the attacked phosphate atom (Fig. 2D).

3.4. PF2046-ssDNA interaction

In terms of sugar puckering, dT_4 is in the 3'-endo configuration (A-form). When conformational parameters are analyzed in the 3DNA software [19], a characteristic A-DNA relation between δ (70–90°) and χ torsion angles (−140° to −180°) is obtained [22]. Each end of dT_4 is responsible for most of the interaction with PF2046, whereas nucleotides in the middle (dT-2 and dT-3) interact mainly with the Mg^{2+} ions, and aside from that, do not interact with the protein. PF2046 and dT_4 interact mainly through hydrophobic interactions except for Thr53 and Ser102, which interact directly with dT-1, when assessed by the Ligplot program (Fig. 3A) [6]. In the DNA-binding pocket, the binding of purine rings is sterically inhibited due to the restricted binding pocket, especially by the strain on dT-4 base by Pro19 and Phe220 (Fig. 3A). PF2046





C

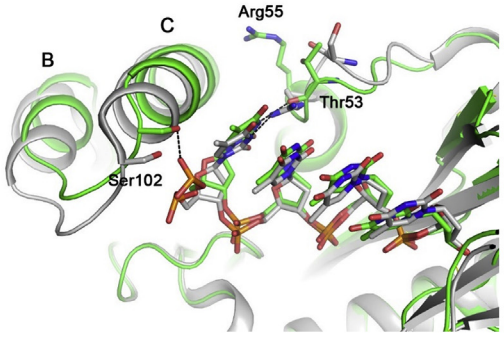
```
PF2046    1  MRIVAADTGGAVLDETFEPIGLIATVAVLVEKPYRSAKEVMVKYANPYDYDLTGRQAIRD  60
PhoExo1   1  MRIVAADTGGAVLDESFQPVGLIATVAVLVEKPYKTSKRFLVKYADPYNYDLSGRQAIRD  60

PF2046   61  EVLLAIELARKVKPDVIHLDSTLGGIELRKLDEPTIDALGISDKGKEVWKELSKDLQPLA 120
PhoExo1  61  EIELAIELAREVSPDVIHLDSTLGGIEVRKLDESTIDALQISDRGKEIWKELSKDLQPLA 120

PF2046  121  RKFWEETNIEIVAIGKSSVPVRIAEIYAGIYSAKWGIENVEKEGHLIIGLPRYMEVNIKD 180
PhoExo1 121  KKFWEETGIEIIAIGKSSVPVRIAEIYAGIFSVKWALDNVKEKGGLLVGLPRYMEVEIKK 180

PF2046  181  GKIIGRSLDPREGGLYGSAEVSVPSGVKWEIYPNPVARRFMIFEISSKR 229
PhoExo1 181  DKIIGKSLDPREGGLYGEVKTEVPQGIKWELYPNPLVRRFMVFEITSKS 229
```

D



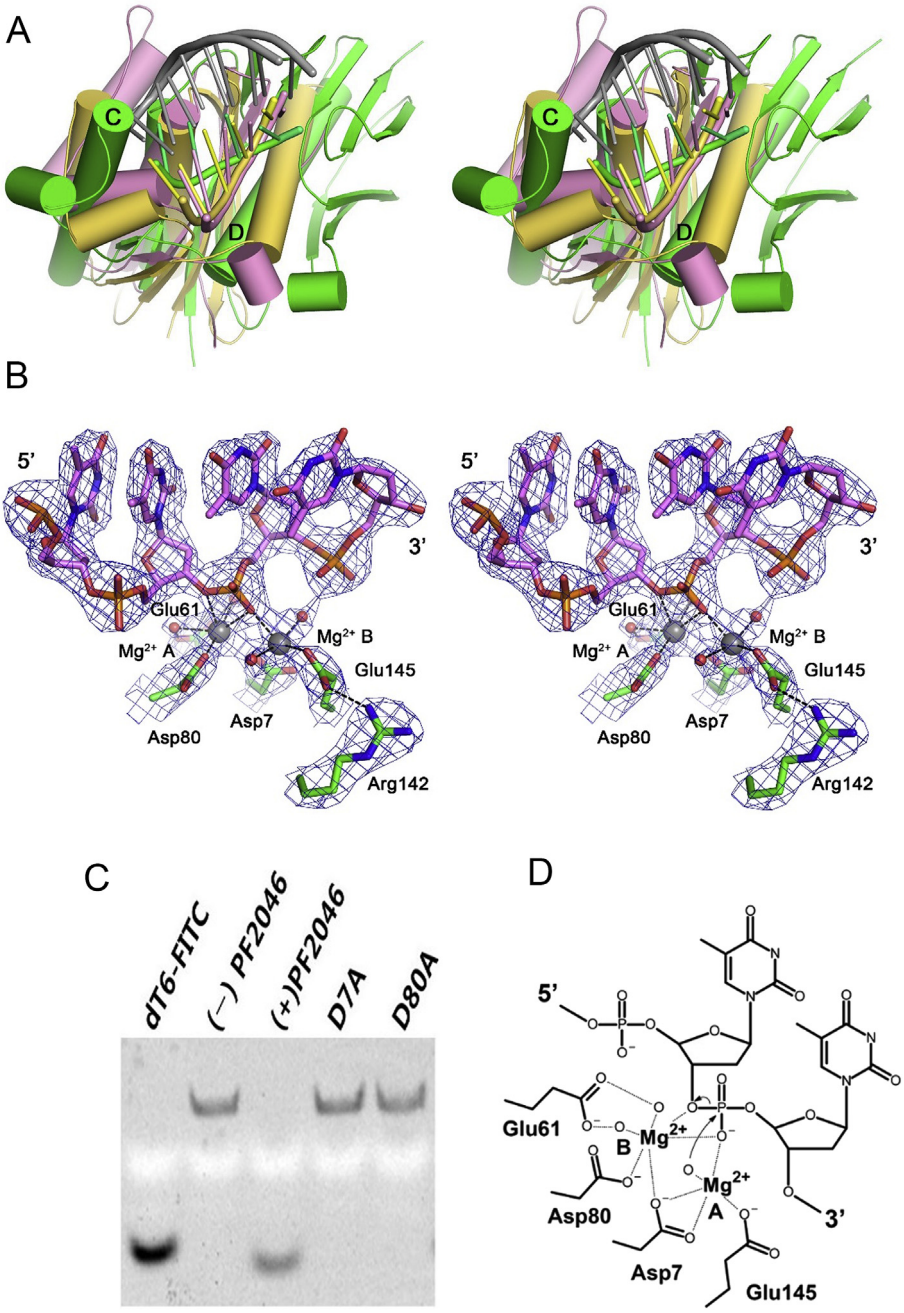


Fig. 2. A. Structural superposition with RNase H proteins. PF2046 (green), BhaRnase H. (pink) and LC11 (yellow) are depicted in ribbon models. DNA strand of DNA/RNA heteroduplexes bound to RNase Hs are colored in gray. **B**. Active-site configuration with a water molecule. The phosphate group of single-stranded DNA (ssDNA) interacts with amino acid residues via Mg^{2+} ion. **C**. The crucial role of the amino acid residues Asp7 and Asp80 involved in coordination of Mg^{2+}. Two variants D7A and D80A of PF2046 fail to process dT_{30}. **D**. The proposed reaction mechanism by PF2046. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shows strongly reduced reactivity toward cytosine nucleotides, and protonation states of the N3 atom may contribute to this specificity. In the structure, the N3 atom of dT-1 is hydrogen-bonded to the main-chain carbonyl group of Thr53 (Fig. 3B).

3.5. Activity assay of PF2046

The stacking interaction between dT-4 and Phe220 is an intriguing feature of substrate binding by PF2046. In the structure,

Fig. 1. A. Overall structure of PF2046-dT_4 and the structure rotated around the x-axis show that PF2046 binds the 3′ end of a single-stranded DNA (ssDNA). B. Structural superposition of apo structure of PF2046 (gray, PDB ID: 4O8U) and PF2046-dT4 (green). Conformational change enables main chain carbonyl group of Thr53 to form hydrogen bond with the N3 atom of thymidine nucleotide. **C**. Sequence alignment of PF2046 and PhoExo I. Identity of between the PF2046 and PhoExo I. **D**. Structural superposition of PF2046-dT_4 (green), and PhoExo I (gray). The residues Thr53, Arg55 and Ser102 of PF2046 and corresponding residues in PhoExo I are shown in sticks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

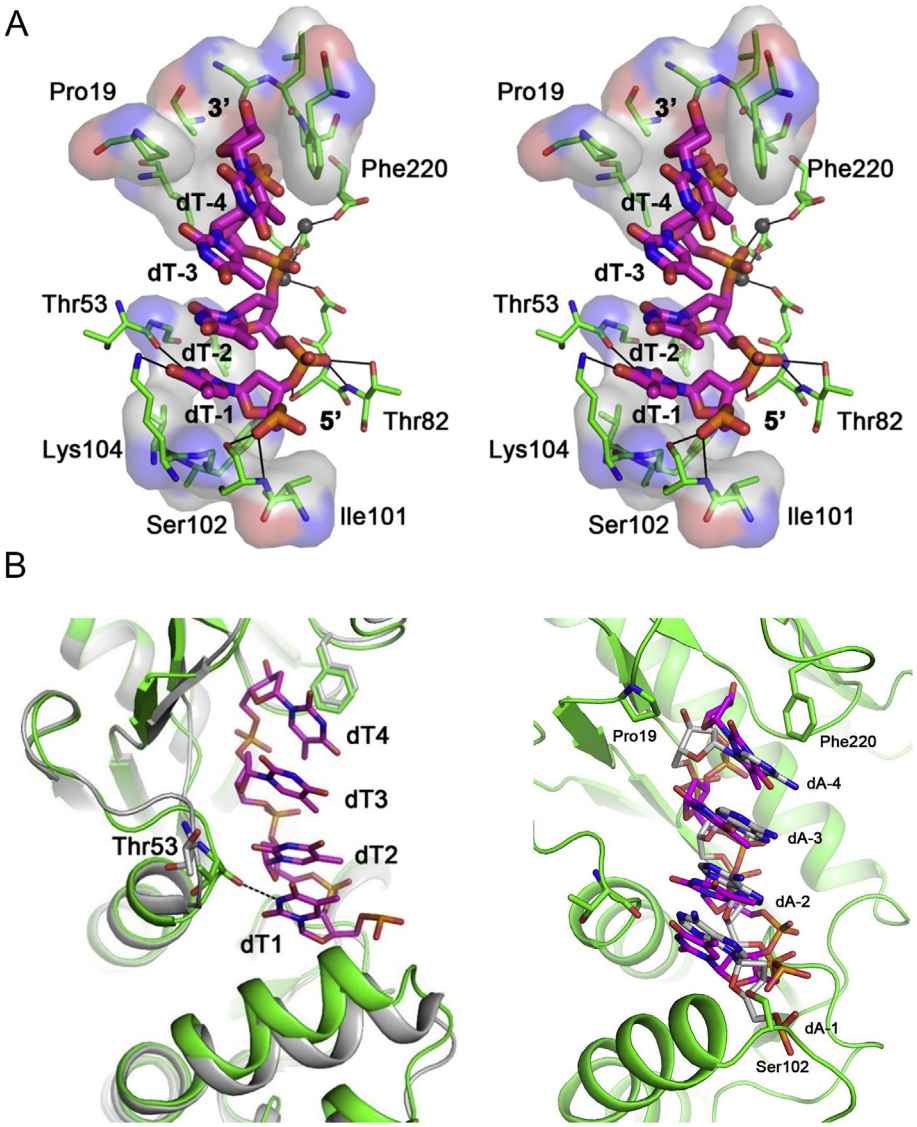


Fig. 3. A. Stereo surface representation of residues in direct contact with dT_4 binding pocket. **B**. Cytosine nucleotide is lacking in hydrogen in the corresponding N3 atom, which could affect its binding affinity. **C**. Representative B-form dA4 is superposed with dT4 in PF2046. Residues involved with ligand binding are represented in stick models.

Phe220 and Pro19 pose notable steric hindrance to dT-4 and contribute to the A-form of dT_4. When we model a typical B-form dA_4 into the binding porcket, Phe220 is found to be in sterical clash with dA-4 (Fig. 3C). Phe220 is invariably conserved among the homologs of PF2046. To understand this effect of steric hindrance by Phe220, we changed it to Leu or Ala. For dT_{30}, all variants were active at approximately the same level. Both variants, however, showed a much reduced activity toward dA_{30} (Fig. 4A).

To understand the effect of pH and cutting from 3′ to 5′ on PF2046 exonuclease activity, pH dependent activity was measured. For dT_{30}, activity at pH 6.0 was almost negligible while pH 7.0 or 8.0 displayed similar reactivity. Similarly, for dC_{30}, almost no reactivity was observed for pH 6.0, even though pH 7.0 or 8.0 showed similar reactivity (Fig. 4B). Due to the inactivity at lower pHs, whether the protonation state at N3 atom of cytosine is relevant for its lower affinity over thymidine could not be confirmed. Time dependent exonuclease activity assay shows that cut from 3′ to 5′ (Fig 4C). The dT_{30} which labeled on 5′ was cut by PF2046 from 3′ to 5′ and could show more and more broken dT_{30} in process of time. It means that PF2046 was able to cut from 3′ to 5′.

4. Discussion

DNA has a tendency to assume the B-conformation in an RNA/DNA duplex; this phenomenon is one of the reasons for selectivity of RNase H [23]. ssDNA is free of restraints of a Watson-Crick pairing and can form relatively diverse conformations. However, poly-dA tends to assume the B-form in solution [24]. The binding conformation of dT_4 in PF2046 is the A-form, which is mainly imposed by the base stacking with Phe220 and a similar binding conformation is expected with a poly-dA. This conformational restriction by PF2046 may reflect the lower activity toward poly-dA. A-form may also be a prerequisite for processing by PF2046 and the inability of Phe220 variants to impose the A-form on dA_{30} may have led to the inactivity (Fig. 4A).

The purpose of the +2 incision is not clear. On the product side, deoxynucleotides di-AMP and di-GMP are widely used as a second messenger by bacteria and archaea, and it is noteworthy that some of these second messengers are involved in the DNA damage response [25]. In humans, the thymidine dinucleotide activates a DNA repair response [26] although a similar activation mechanism

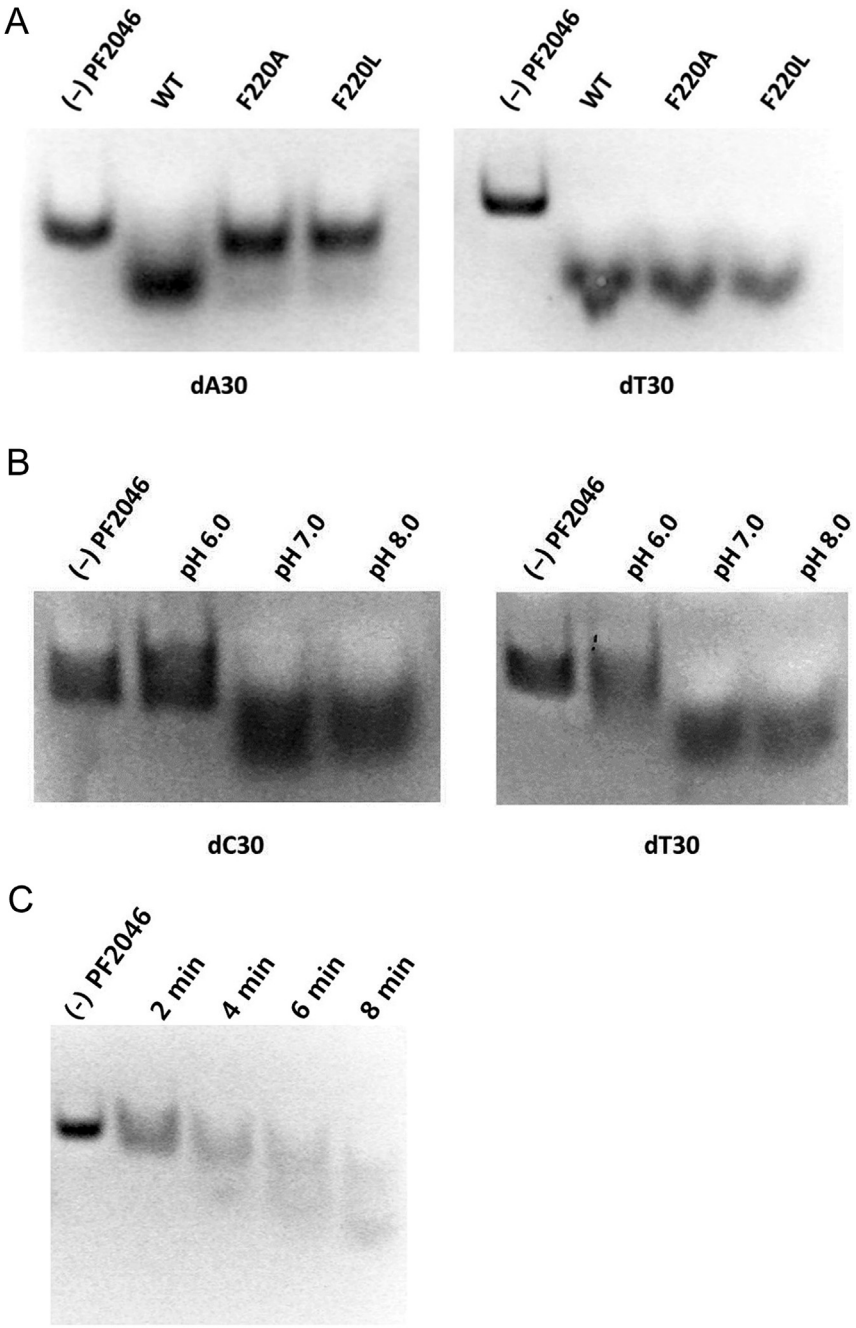


Fig. 4. A. A specificity assay with dT_{30} and dA_{30} and Phe220 variants (F220A and F220L). **B**. pH-dependent exonuclease activity on dC_{30} and dT_{30} by PF2046 was measured pH 6.0, 7.0 and 8.0. **C**. Time-dependent 3′ exonuclease activity assay on dT_{30}. PF2046 which labeled at 5′ shows cutting from 3′ to 5′.

has not been uncovered in archaea. On the basis of the propensity for production of thymidine dinucleotides during ssDNA degradation, it is tempting to speculate that PF2046 can generate a signal of DNA damage and induce a DNA repair response via a yet unknown pathway.

Another question arising from our structural analysis of PF2046 relates to its quaternary structure. Another exonuclease in the trimeric configuration has been reported: λ-exonuclease [27]. λ-exonuclease binds one copy of dsDNA per trimer, and the ssDNA product exits through a toroidal hole in the trimeric center. PF2046 collectively binds three oligonucleotides per trimer and there is no such a hole for the reaction products to pass through. In contrast to other RNase H proteins, the trimeric configuration of PF2046 restricts access of bulky substrates. The lack of dsDNA activity may be in part inherent in the ability to pass only ssDNA through the narrow entry cleft.

5. Conclusion

In conclusion, the crystal structure of the PF2046-dT_4 complex should advance the understanding of substrate binding and the reaction mechanism of PF2046. Our study illustrates a peculiar function of the RNase H fold in hyperthermophilic archaea. PF2046 is a full-fledged member of the RNase H family, and our analysis of the A-form of bound ssDNA suggests that PF2046 may also function as a ribonuclease of an ssRNA of unknown identity. Recently,

Miyazono et al. reported crystal structures of PhoExo I from *Pyrococcus horikoshii*, a homolog of PF2046 [5]. Structural comparison between PF2046 and PhoExo I indicated to differ about interaction with ssDNA by key residue such as Thr53 and Ser102. One possibility of the difference is the residual activity of PF2046 processing oligonucleotides into 4-base product to turn out as enzyme-product complex. Our structural and mutagenesis study suggests the conformational restrain on ssDNA can enable an RNaseH fold protein to function as a deoxyribonuclease. As thermophilic nucleases are attractive tools for particular molecular biology applications, this structure-functional study of PF2046 will be helpful in utilization of this interesting nuclease.

Acknowledgments

We thank supporting staff of beamline BL1A of the Photon Factory (Tsukuba, Japan) for the help with data collection. This research was supported by Basic Science Research Program through the gs2:National Research Foundation of Korea funded by the Ministry of Education (NRF-2013R1A1A2059835] to WCL. WCL and KYH were supported by the Korea University grants.

References

[1] G. Fiala, K.O. Stetter, Arch. Microbiol. 145 (1986) 56–61.
[2] J. DiRuggiero, N. Santangelo, Z. Nackerdien, J. Ravel, F.T. Robb, J. Bacteriol. 179 (1997) 4643–4645.
[3] K. Tori, S. Ishino, S. Kiyonari, S. Tahara, Y. Ishino, PLoS One 8 (2013) e58497.
[4] J. Su, Y. Li, N. Shaw, W. Zhou, M. Zhang, H. Xu, B.C. Wang, Z.J. Liu, Protein Cell 1 (2010) 453–458.
[5] K. Miyazono, S. Ishino, K. Tsutsumi, T. Ito, Y. Ishino, M. Tanokura, Nucleic Acids Res. 43 (2015) 7122–7136.
[6] R.A. Laskowski, M.B. Swindells, J. Chem. Inf. Model. 51 (2011) 2778–2786.
[7] N. Ohtani, M. Haruki, M. Morikawa, R.J. Crouch, M. Itaya, S. Kanaya, Biochemistry 38 (1999) 605–618.
[8] B.R. Chapados, Q. Chai, D.J. Hosfield, J. Qiu, B. Shen, J.A. Tainer, J. Mol. Biol. 307 (2001) 541–556.
[9] S.M. Cerritelli, R.J. Crouch, FEBS J. 276 (2009) 1494–1505.
[10] W. Yang, Nat. Struct. Mol. Biol. 15 (2008) 1228–1231.
[11] W. Yang, Q. Rev. Biophys. 44 (2011) 1–93.
[12] D.L. Theobald, R.M. Mitton-Fry, D.S. Wuttke, Annu. Rev. Biophys. Biomol. Struct. 32 (2003) 115–133.
[13] A.G.W. Leslie, H.R. Powell, Evolving Methods for Macromolecular Crystallography, Springer, 2007, pp. 41–51.
[14] P. Evans, Acta Crystallogr. D. Biol. Crystallogr. 62 (2006) 72–82.
[15] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Acta Crystallogr. D. Biol. Crystallogr. 67 (2011) 235–242.
[16] A. Vagin, A. Teplyakov, Acta Crystallogr. Sect. D. 66 (2010) 22–25.
[17] D.E. McRee, J. Struct. Biol. 125 (1999) 156–165.
[18] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Acta Crystallogr. Sect. D. 54 (1998) 905–921.
[19] X.J. Lu, W.K. Olson, Nat. Protoc. 3 (2008) 1213–1227.
[20] E. Krissinel, K. Henrick, J. Mol. Biol. 372 (2007) 774–797.
[21] M.E. Maguire, J.A. Cowan, Biometals 15 (2002) 203–210.
[22] R.E. Dickerson, H.R. Drew, B.N. Conner, R.M. Wing, A.V. Fratini, M.L. Kopka, Science 216 (1982) 475–485.
[23] M. Nowotny, S.A. Gaidamakov, R.J. Crouch, W. Yang, Cell 121 (2005) 1005–1016.
[24] J. Isaksson, S. Acharya, J. Barman, P. Cheruku, J. Chattopadhyaya, Biochemistry 43 (2004) 15996–16010.
[25] R.M. Corrigan, A. Grundling, Nat. Rev. Microbiol. 11 (2013) 513–524.
[26] M.S. Eller, T. Maeda, C. Magnoni, D. Atwal, B.A. Gilchrest, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 12627–12632.
[27] R. Kovall, B.W. Matthews, Science 277 (1997) 1824–1827.