Structural Insights of the ssDNA Binding Site in the Multifunctional Endonuclease AtBFN2 from Arabidopsis Thaliana

Tsung-Fu Yu (尹宗富)a,b, Manuel Maestre-Reyna b, Chia-Yun Ko (柯嘉昀)c,d, Tzu-Ping Ko (柯子平)b, Yuh-Ju Sun (孫玉珠)c, Tsai-Yun Lin (林彩雲)c, Jei-Fu Shaw (蕭介夫)c,d,e, and Andrew H.-J. Wang (王惠鈞)a,b,c,e

a Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu 30013, Taiwan
b Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan
c Agricultural Biotechnology Research Center, Academia Sinica, Taipei 11529, Taiwan
d Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan
e Department of Biological Science and Technology, I-Shou University, Kaohsiung 84001, Taiwan
albert.yul1128@msa.hinet.net

Abstract

The multi S1/P1 nuclease AtBFN2 (EC 3.1.30.1) encoded by the Arabidopsis thaliana At1g68290 gene is a glycoprotein that digests RNA, ssDNA, and dsDNA. AtBFN2 depends on three zinc ions for cleaving DNA and RNA at 3’-OH to yield 5’-nucleotides. In addition, AtBFN2’s enzymatic activity is strongly glycan dependent. Plant Zn2+-dependent endonucleases present a unique fold, and belong to the Phospholipase C (PLC)/P1 nuclease superfamily. In this work, we present the first complete, ligand-free, AtBFN2 crystal structure, along with sulfate, phosphate and ssDNA co-crystal structures. With these, we were able to provide better insight into the glycan structure and possible enzymatic mechanism. In comparison with other nuclease, the AtBFN2/ligand-free and AtBFN2/PO4 models suggest a similar, previously proposed, catalytic mechanism. Our data also confirm that the phosphate and vanadate can inhibit the enzyme activity by occupying the active site. More importantly, the AtBFN2/A5T structure reveals a novel and conserved secondary binding site, which seems to be important for plant Zn2+-dependent endonucleases. Based on these findings, we propose a rational ssDNA binding model, in which the ssDNA wraps itself around the protein and the attached surface glycan, in turn, reinforces the binding complex.

Keywords - X-ray diffraction; nuclease; phosphodiesterase; tri-nuclear metal enzyme; N-glycosylation.

Introduction

Nucleases catalyze the cleavage of phosphodiester bonds in nucleic acids, and can be classified into DNases and RNases [1]. The multi S1/P1 nucleases have a broad specificity. They can cleave both single-stranded DNA (ssDNA) and RNA, but show a lower affinity for dsDNA. The S1/P1 nuclease class takes its name from P1 nuclease from Penicillium citrinum [2] and Nuclease S1 from Aspergillus oryzae [3,4], which are both glycosylated endonucleases, and share 51% sequence identity (68% homology). S1/P1 endonucleases are also present in plants, and are involved in programmed cell death (PCD) via hydrolysis of genomic DNA [5]. Plant S1/P1 endonucleases are divided into two broad types according to their metal ion cofactor: Ca2+-dependent and Zn2+-dependent endonucleases (EC 3.1.30.1) [6]. Ca2+-dependent endonucleases prefer ssDNA over RNA at neutral pH, and may be involved in plant immunity [7]. Acidic pH leads to an increase in activity of Zn2+-dependent endonucleases including Arabidopsis thaliana multifunctional Nuclease 1 (AtBFN1 or ENDO1) [8,9], AtBFN2 [10], BEN1 [11], CEL1 [12], ZEN1 [13], ABN1 and HBN1 [14]. N-Glycosylation patterns are highly conserved within the plant S1/P1 endonucleases [15], although these sites are not consistent with those in P1 nuclease [16]. AtBFN2 (also called ENDO2) is a 34.5 kDa glycoprotein with three glycosylation sites (Asn91, Asn110, and Asn184). While the P1 nuclease has a higher specificity for RNA over ssDNA, in a previous study we showed that the opposite is true for AtBFN2 [10]. Structurally, Zn2+-dependent endonucleases present a unique fold, and are classified into the P1/S1 nuclease family (pfam ID PF02265) [17], itself part of the Phospholipase C (PLC)/P1 nuclease superfamily (pfam ID CL0368) [18]. The active site within the P1/S1 nuclease family is very well conserved [19,20]. It contains nine zinc cluster interacting amino acids (1 tryptophan; 3 aspartic acids and 5 histidines) [15]; the conservation of this metal cluster extends into the PC-PLC/P1 endonuclease superfamily despite low sequence identity [21]. Despite extensive efforts to understand how Zn2+-dependent endonuclease bind and digest ssDNA [16,22], it is still not clear how, or even whether, these endonucleases unwind and cleave dsDNA. Further, the mechanism by which endonucleases specifically interact with ssDNA remains obscure. All co-crystal structures show partially digested DNA in the active site [16], or crystalline artifacts outside of it [17]. Additionally, though it has been previously suggested that glycosylation is relevant for protein structural integrity [10], a role for glycans in substrate binding or modulation of active site shape may also be important.

To address these open questions, we extended our previous studies based on AtBFN2 [16]. We improved the resolution of our earlier sulfate co-crystal [16] from 1.76 Å to 1.22 Å, thus yielding more insight into the glycan structure. We obtained a ligand-free AtBFN2 structure, as well as a phosphate co-crystal structure in a new unit cell with space group P1. We also found that phosphate and its transition state analog vanadate were both capable of inhibiting enzyme activity. Finally, we obtained a detailed ssDNA•AtBFN2 co-crystal structure by soaking the crystals with a thiophosphorylated ssDNA analog. In this structure a large secondary ssDNA
binding site was observed. Based on these findings, we propose a rational model for ssDNA•AtBFN2 binding, in which the ssDNA wraps itself around the protein and the attached surface glycan, in turn, reinforces the complex, acting not unlike an elastic band.

Experiments
Cloning, overexpression, mutagenesis and purification of AtBFN2

The ENDO2 cDNA was amplified by PCR with primer pairs B2-11F and B2-12R to include a C-terminal 6 x His-tag (Table S1). The ENDO2-His-tag was cloned into pBI121 binary vector to form CaMV35SP::ENDO2-hisOE, and introduced into Agrobacterium tumefaciens strain GV3101::pMP90 by electroporation. AtBFN2 was purified following the method of Ko et al. with minor modifications [10]. Purified AtBFN2 was concentrated with 10 kDa Vivaspain turbo for crystallization.

Crystallization, data collection, and structure determination

Ligand-free AtBFN2 crystals grew within 5 to 7 days in drops prepared by mixing 1 μl of the reservoir solution (0.1 M Tris pH 8.5, 0.2M NaOAc, 30% (w/v) PEG 4000) with 1 μl of the protein (6.0 mg/ml). All AtBFN2/PO4; AtBFN2/PO4 and AtBFN2/A5T crystals grew within 3 to 4 weeks in drops prepared by mixing 1 μl of the reservoir solution (0.1 M Hepes, pH 7.5; 0.2 M LiSO4 and 30% (w/v) PEG 3350 for AtBFN2/PO4; 0.1 M Hepes, pH 7.5; 0.2 M Na2HPO4 and 30% (w/v) PEG 3350 for AtBFN2/A5T) with 1 μl of the protein (8.5 mg/ml). The AtBFN2/A5T co-crystals were produced by soaking AtBFN2/PO4 crystals with thiophosphorylated nucleotides: d[A(s)A(s)A(s)A(s)A(s)T], denoted as A5T. AtBFN2/PO4 crystals were washed in 0.1 M Hepes, pH 7.5 and then soaked in 0.1 M Hepes, pH 7.5 and 30% (w/v) PEG 3350 for 30 min. Subsequently, they were flash-cooled to 100 K in a stream of cold nitrogen.

The X-ray diffraction images were collected using the SPXF beamline BL13B1, BL13C1, and BL15A1 of the National Synchrotron Radiation Research Center (NSRRC) in Taiwan, and were then processed using HKL2000 [23]. Prior to use in structure determination, 5% randomly selected reflections were set aside as our Rfree reference [24]. The statistics of data collection are summarized in Table 1.

ssDNA activity assay

ssDNA activity was measured in triplicate according to the method of Ko et al. [10]. All assays were performed in 100 mM Tris pH 7 buffer, to which different putative inhibitors were added (100 mM Na2SO4; 100 mM LiCl; 100 mM Li2SO4; 1 mM and 10 mM Na3VO4; 1 mM and 10 mM and 100 mM NaH2PO4).

Results and Discussion
Extended N-glycan structures

The protein portion of the sulfate co-crystal structure (4CXO, Table 1) presented in this work is in very good agreement with our previously published AtBFN2•PO4 co-crystal structure (3W52), which showed the overall fold of this endonuclease (Fig. 1A) [16]. The improved diffraction quality of 4XCP, however, allowed us to refine the structure to include almost all putative glycans, as described by Ko et al. [10], with only six moieties still missing (Fig. 1). As with the 3W52, 4XCP shows three N-glycosilation sites, at Asn91, Asn110, and Asn184.

The glycan bound to Asn91 is especially well defined and we can discern both N-acetylglucosamines (GlcNAc), and a further six mannose moieties (Man) (Fig. 1B). Much like in 3W52, GlcNAc1 interacts extensively with the polypeptide, forming hydrogen bonds with Arg83, Tyr92, and Gln95. Further, the methyl group of GlcNAc1 is located within a small aliphatic pocket formed by Leu57, and Ala88. GlcNAc2, on the other hand, interacts via π-ring stacking with Trp53. Man1-5 do not interact with the protein. Man6, on the other hand, folds backwards into the protein, and interacts with the side chain of Gln108 and with the carbonyl group of GlcNAc2 (Fig. 1B).

While we were able to refine only one additional mannose moiety on the glycan bound to Asn184 (Fig. 1D), we obtained a much more detailed structure for the one attached to Asn110. It was thus possible to observe two more Man residues, which extend towards the solvent, and do not interact directly with the protein (Fig. 1C). A fucose moiety (Fuc) is attached to GlcNAc1 via an α-1,3 glycoside bond, but it also makes no direct contacts with the protein.

The ssDNA•AtBFN2 complex structure reveals a secondary, glycan dependent ssDNA binding site

By soaking our crystals in a 0.2 mM solution of the thio-ssDNA analog ArT (sequence: d[A(s)A(s)A(s)A(s)A(s)T]), we managed to obtain the first ssDNA•AtBFN2 co-crystal structure (4CXO, Table 1). Here we observed two distinct binding pockets. A single deoxythymidine monophosphate residue was bound to the active site, in much the same way as has been described for P1 nuclease [17]. We were unable to refine any further nucleotides at this site, indicating that the rest of the hexanucleotide was probably highly flexible. The thymine moiety was stacked against Tyr59 and lodged within a pocket comprising the peptide backbones of His130, Lys137, and Gly138, and the side-chains of Tyr59, Asn61, and Leu129 (Fig. 2A). Asn61 bound the thymine carbonyl oxygen 4 via a hydrogen bond. The deoxyribose was bound to Asn140 at the ring oxygen, while the 3'-end was tightly coordinated to Zn2, displacing the Wat3, which was present in the other three structures. Under these conditions, we also found a single sulfate anion within the binding pocket, which had displaced the other waters, including catalytically relevant Wat1.

The secondary binding site is completely novel, and much larger than the active site pocket. In it, elements from both the protein and the N-glycan at Asn91 participate in the binding of a dinucleotide composed of one adenosine, and one thymidine (Fig. 2B). As with the active site, a large part of the ssDNA molecule could not be refined. However, since thymidine was present in both binding sites, it follows that the two binding sites are
occupied by distinct molecules. At the 5'-end of the single stranded dinucleotide, the adenosine is bound by an extensive network of hydrogen bonds and hydrophobic interactions, with Tyr70 stacking with the purine moiety. Further, the amide belonging to Asn71 coordinates the aromatic nitrogen 7 and the primary amine group on the adenine ring. The deoxyribose portion of the adenosine interacts via a π-ring interaction with Tyr90, while the phosphate backbone attracts the positive side-chain of Lys201.

A wrapped ssDNA binding model

The secondary binding site possibly reveals a new and crucial role for glycosylation at Asn91, since GlcNAc1 is an important partner in binding the dinucleotide. A reasonable hypothesis would be that a longer ssDNA strand would be capable of interaction with further glycans. Thus, we propose that N-glycosylation at Asn91 has not only a role in maintaining structural integrity, but also in substrate binding. While it was only possible to observe two nucleotides from our short DNA sequence within the secondary binding pocket, the relative position of these, in comparison with the adenosine in the active site, is suggestive of a long, snaking DNA molecule wrapped around the surface of AtBFN2 (Fig. 3). Based on DelPhi electrostatic calculations [34], we propose that there are two further ssDNA binding pockets (pockets two and four in Fig. 3). Pocket two is occupied in the sulfate co-crystal structure by sulfate (Fig. 1A), suggesting that it should be possible for the ssDNA backbone to be bound in a similar fashion. On the other hand, pocket four, an extension of pocket three, is positively charged, and contains several aromatic residues (Figure 3).

The physiological role of the Tyr site on the P1 nuclease (Tyr144 and Tyr155) has not yet been confirmed, and there is still doubt as to its relevance outside the crystal environment, since the key amino-acids are poorly conserved [17]. Therefore we sought to evaluate the degree of conservation, in comparison to the Tyr site, of our experimental and proposed secondary binding sites by performing an alignment based on several plant Zn²⁺-dependent nucleases, and the more distantly related P1 nuclease, and BcPLC (Fig. 2C). The majority of the amino-acids involved in the experimental secondary binding site (pocket three) are well conserved within the plant Zn²⁺-dependent nucleases, and the ones with a lesser degree of conservation are replaced by residues capable of similar interactions. For example, Asn71, was conservatively replaced by serine in the consensus sequence (Fig. 2C). On the other hand, Lys201 is most often replaced by aspartate, which is incapable of interacting with the phosphate backbone. However, aspartate side-chains can rotate freely, resulting in low repulsion against the nucleic acid binding. Additionally, there may be alternative pathways by which the binding pocket may attain a similar geometry. For example, the tomato multifunctional endonuclease TBN1 presents a Gln instead of Tyr at position 90 (Fig. 2C). The TBN1 structure (3SNG) [22] also comprises an N-terminal extension (lacking in AtBFN2) of α helix 11 with an exposed phenylalanine. Thus, it is possible that the side-chains of Gln and Phe are able to structurally compensate for the absence of Tyr90. Furthermore, the computed binding sites (pockets two and four) exhibit high sequence similarity, with conserved amino acid substitutions, e.g. Arg to Lys. Considering the high levels of sequence and structure conservation in the proposed ssDNA secondary binding sites, it is highly likely that they have some biological significance in plant Zn²⁺-dependent endonucleases. Since, however, their levels of sequence and structure conservation in AtBFN2, and P1 nuclease are low, and since both exhibit different specificities (weaker ssDNA binding to P1 nuclease) the DNA binding mechanism for distantly related endonucleases might differ considerably.

Conclusion

In this work, we presented structural and functional data confirming that phosphate and vanadate act as inhibitors of AtBFN2 by occupying the active site, and possibly displacing catalytic water. We also obtained high resolution structures with well defined N-glycans at positions 91, 110, and 184 of the protein. Our most important finding, however, was the strong evidence for a secondary ssDNA binding site, for which the N-glycan at position 91 might be crucial. This binding site seems to be important within plant Zn²⁺-dependent and/or ssDNA digesting endonucleases, yet might not have similar relevance for other organisms.
Acknowledgments

We would like to thank the Transgenic Plant Core Facility, Academia Sinica for assistance in plant transformation and National Synchrotron Radiation Research Center for X-ray data collection. This work was supported by Academia Sinica, and Ministry of Science and Technology grants: NSC-101-2313-B-214-002-MY3 (to J.-F. Shaw), (NSC-103-3113-P-008-001, and NSC-101-2319-B-001-003 (to A. H.-J. Wang).

References

1. Yang W (2011) Nucleases: diversity of structure, function and mechanism. Q Rev Biophys 44: 1–93.
2. Fujimoto M, Kuninaka A, Yoshino H (1974) Identity of phosphodiesterase and phosphomonoesterase activities with nuclease P1 (a nuclease from Penicillium citrinum). Agric Biol Chem. 38: 785–790.
3. Vogt VM (1980) Purification and properties of S1 nuclease from Aspergillus. Methods Enzym 65: 248–255.
4. Oleson AE, Sasakuma M (1980) S1 nuclease of Aspergillus oryzae: a glycoprotein with an associated nucleotidase activity. Arch Biochem Biophys 204: 361–370.
5. McCabe PF, Leaver CJ (2000) Programmed cell death in cell cultures. Plant Mol Biol 44: 359–368.
6. Sugiyama M, Ito J, Fukuda H (2002) ZEN1 is a key enzyme in the degradation of nuclear DNA during programmed cell death of tracheary elements. Plant Cell 14: 3201–3211.
7. Sugiyama M, Ito J, Aoyagi S, Fukuda H (2000)
8. McCabe PF, Leaver CJ (2000) Programmed cell death
9. Farage-Barhom S, Perl-Treves R, et al. (2013) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 69: 213–226.
10. Otwinowski Z, Minor W (1997) Processing of X-raydiffraction data collected in oscillation mode. Methods Enzym 276: 307–326.
11. Ko CY, Lai YL, Liu WY, Lin CH, Chen YT, et al. (2013) Plant multifunctional nuclease TBN1 with unexpected phospholipase activity: structural study and reaction-mechanism analysis. Acta Crystallogr D Biol Crystallogr 69: 213–226.
12. Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Structure 18: 1137–1140.
13. Ito J, Fukuda H (2002) ZEN1 is a key enzyme in the degradation of nuclear DNA during programmed cell death of tracheary elements. Plant Cell 14: 3201–3211.
14. Podzimek T, Matousek J, Lipovova P, Prouckova P, Spiwok V, et al. (2011) Biochemical properties of three plant nuclease with anticancer potential. Plant Sci 180: 343–351.
15. Till BJ, Burnet C, Comai L, Henikoff S (2004) Mismatch cleavage by single-strand specific nucleases. Nucleic Acids Res 32: 2632–2641.
16. Chou TL, Ko TP, Ko CY, Lin TY, Guo RT, et al. (2013) Mechanistic insights to catalysis by a zinc-dependent bi-functional nuclease from Arabidopsis thaliana. Biocatal Agric Biotechnol 2: 191–195.
17. Romier C, Dominguez R, Lahm A, Dahl O, Suck D (1998) Recognition of single-stranded DNA by nuclease P1: high resolution crystal structures of complexes with substrate analogs. Proteins 32: 414–424.
18. Finn RD, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, et al. (2006) Pfam: clans, web tools and services. Nucleic Acids Res 34: D247–51.
19. Mackawa K, Tsuchioka S, Digo B, Sakiyama F (1991) Primary structure of nuclease P1 from Penicillium citrinum. Eur J Biochem 200: 651–661.
20. Volbeda A, Lahm A, Sakiyama F, Suck D (1991) Crystal structure of Penicillium citrinum P1 nuclease at 2.8 Å resolution. EMBO J 10: 1607–1618.
21. Hough E, Hansen LK, Birknes B, Jynge K, Hansen S, et al. (1989) High-resolution (1.5 Å) crystal structure of phospholipase C from Bacillus cereus. Nature 338: 357–360.
22. Koval T, Lipovova P, Podzimek T, Matousek J, Dusko J, et al. (2013) Plant multifunctional nuclease TBN1 with unexpected phospholipase activity: structural study and reaction-mechanism analysis. Acta Crystallogr D Biol Crystallogr 69: 213–226.
23. Otwinowski Z, Minor W (1997) Processing of X-raydiffraction data collected in oscillation mode. Methods Enzym 276: 307–326.
24. Brünger AT (1993) Assessment of phase accuracy by cross validation: the free R value. Acta Crystallogr D Biol Crystallogr 49: 24–36.
25. Brünger AT, Adams PD, Clore GM, Delano WL, Gros P, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54: 905–921.
26. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67: 235–242.
27. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126–2132.
28. Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ (1999) Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr D Biol Crystallogr 55: 247–255.
29. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67: 355–367.
30. Chen VB, Arendall 3rd WB, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66: 12–21.