nature structural & molecular biology

Crystal structures of aprataxin ortholog Hnt3 reveal the mechanism for reversal of 5′-adenylated DNA

Yong Gong[1,3], Deyu Zhu[2,3], Jingjin Ding[1,3], Chuan-Na Dou[1], Xiaoming Ren[1], Lichuan Gu[2], Tao Jiang[1] & Da-Cheng Wang[1]

Aprataxin is a DNA deadenylase that resolves DNA 5′-AMP termini and reverses abortive DNA ligation. The crystal structures of *Schizosaccharomyces pombe* aprataxin Hnt3 in its apo form and in complex to dsDNA and dsDNA–AMP reveal how Hnt3 recognizes and processes 5′-adenylated DNA in a structure-specific manner. The bound DNA adopts a 5′-flap conformation that facilitates 5′-AMP access to the active site, where AMP cleavage occurs by a canonical catalytic mechanism.

Human aprataxin (APTX) is a neurodegenerative disease–related protein whose dysfunction causes ataxia with oculomotor apraxia 1 (AOA1)[1–3]. It is now known that APTX has a direct and unique role in DNA repair, catalyzing the nucleophilic release of adenylate groups from 5′-AMP termini at DNA nicks or breaks[4]. This finding provides a molecular rationale for AOA1. Recent studies have revealed some critical molecular and biochemical properties of APTX in DNA repair[5–10]. APTX is a unique member of the histidine triad (HIT) superfamily, in which the HIT motif HΦHΦHΦΦ (where Φ is a hydrophobic residue) works as the active site[11]. The specificity of APTX for DNA adenylates is achieved by its HIT and zinc-finger (ZF) domains, which, in both yeast and humans, is a defining feature of aprataxin orthologs that are functionally conserved[4,9]. It is noteworthy that the first position of the HIT motif is not strictly conserved in aprataxins[9]. The underlying mechanism for reversal of 5′-adenylated DNA by aprataxins has remained unresolved because of the lack of solved structures of the enzymes bound to DNA. The *S. pombe* aprataxin ortholog Hnt3, which we have studied here, is an ortholog of human APTX (**Fig. 1a**) and has a similar role in DNA repair pathways[4,12]. Therefore, we determined the crystal structures of apo Hnt3 and Hnt3–DNA and Hnt3–DNA–AMP complexes at 1.8, 1.9 and 2.1 Å resolution (**Supplementary Table 1**), respectively, as described in **Supplementary Methods**.

The apo Hnt3 structure shows that the HIT domain and ZF domain are packed together to form a compact structure (**Supplementary Fig. 1a**). The HIT domain of Hnt3 contains a typical HIT domain and a conserved region immediately downstream[7] that we here together refer to as the eHIT domain ('e' for extended), which is characterized by a six-stranded β-sheet flanked by seven helices. The zinc-finger motif (Cys200, Cys203, His217 and Glu221) coordinates a zinc ion. The superposition of the structures of Hnt3 and human HINT1 (PDB 1KPF[13]) suggests that the HIT-domain structure is highly conserved across the HIT superfamily. By contrast, there is an obvious difference in the C-terminal segment, which may be related to the distinct substrate specificities of Hnt3 and other HIT members (**Supplementary Fig. 1b**).

The asymmetric units of both Hnt3-DNA and Hnt3-DNA-AMP co-crystals contain two copies of the Hnt3–DNA complex, each consisting of two protein molecules bound to one double-stranded DNA (dsDNA) (**Fig. 1b–d**). The structures of Hnt3 in the apo and complex states are virtually identical (r.m.s. deviation of ~0.6 Å), indicating that it uses a rigid, preformed surface to interact with the dsDNA. The observed dsDNA in our structures contains one nucleotide that flips out of the DNA duplex at the 5′ end of each strand and a thirteen–base pair (bp) duplex segment that broadly adopts a B-form conformation (**Supplementary Fig. 1c**). Notably, the two Hnt3 molecules (monomers A and B) within the complex do not interact with each other, and they bind the DNA duplex independently as monomers. Moreover, a noncrystallographic pseudo two-fold symmetry axis is located in the center of the DNA duplex, indicating that the interactions of both monomers with the DNA are almost identical. For clarity, we will mainly describe the protein-DNA interactions of monomer A with the dsDNA in the Hnt3–DNA–AMP complex.

Hnt3 constitutes a distinct molecular platform for DNA recognition and processing through three essential parts. The first is a unique DNA end–binding cleft formed by the eHIT domain, which interacts with the 5′ end of the short DNA strand (**Fig. 2a–c**). Four conserved residues (Phe65, Lys67, Lys161 and His165) act as a gate of the cleft to fix the DNA backbone through interactions with nucleotides $T3_S$ and $T4_S$ (the subscript 'S' means short strand). The N-terminal loop, β4–β5 loop, α6–β6 loop and α7 helix cooperatively interact with the first and second base pairs, allowing for structure-specific recognition of 5′ nucleotides and causing the first base pair to flip out of the double helix. Residue Phe34 forms a stacking interaction with nucleotide $A2_S$, end-capping the exposed $A2_S$–$T2_L$ base pair. The F34A mutant shows markedly decreased DNA-binding affinity (**Supplementary Fig. 2**), suggesting that Phe34 has an essential role in DNA recognition. It is noteworthy that the matched $C15_S$-$G15_L$ base pair at the blunt end is also disrupted, and the nucleotides $T1_S$ and $G15_L$ at both ends form an identical 5′-flap conformation adjacent to the HIT active site (**Figs. 1c,d** and **2c**). Moreover, the 5′ flap contributes ~28% (~215 Å^2) of the ~760 Å^2 of the dsDNA solvent-accessible surface area that is buried each time

[1]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. [2]State Key Laboratory of Microbial Technology, School of Life Science, Shandong University, Jinan, China. [3]These authors contributed equally to this work. Correspondence should be addressed to T.J. (tjiang@sun5.ibp.ac.cn) and D.-C.W. (dcwang@ibp.ac.cn).

Received 16 May; accepted 25 July; published online 9 October 2011; doi:10.1038/nsmb.2145

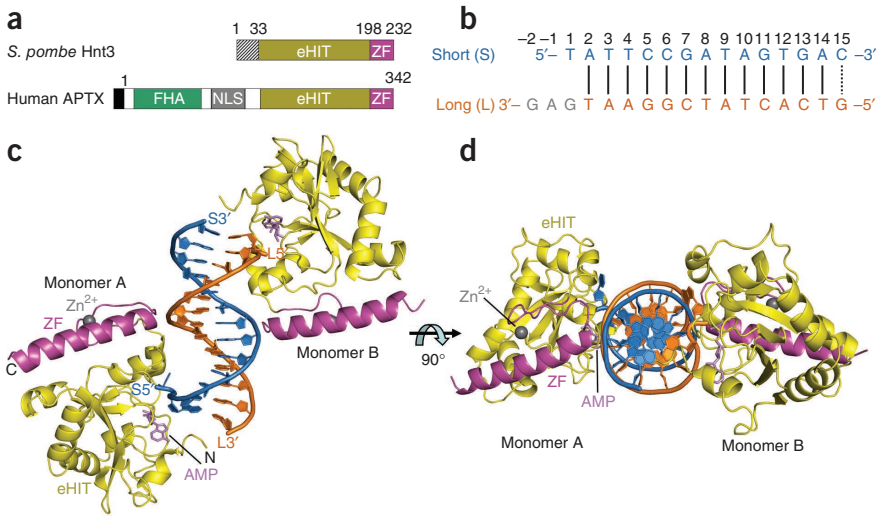

Figure 1 Overall structure of the Hnt3–dsDNA–AMP ternary complex. (**a**) Schematic representations of the Hnt3 and human APTX domains. Sequences not included in the crystallized Hnt3 are marked with hatched lines. The black box of human APTX denotes the additional 14 amino acids (aas) that are present at the N-terminal in the longest (356 aa) alternative splice variant. (**b**) Sequences of the dsDNA used in the co-crystallization. The bases at position 1 are mismatched, and the paired bases at 15 are disrupted in the co-crystals. (**c**) Ribbon diagram of Hnt3 in complex with dsDNA and AMP (in the ratio of 2:1:2, respectively). eHIT, yellow; ZF, magenta; short strand of the dsDNA, blue; long strand, orange. Zn^{2+} and AMP are shown as a gray sphere and violet sticks, respectively. (**d**) Ribbon diagram of the complex rotated by 90° about the horizontal axis relative to **c**.

the Hnt3 monomer binds, indicating that the 5′ flap contributes substantially to Hnt3 DNA substrate affinity.

The second part of the molecular platform is the ZF domain, which contacts the DNA backbone of the long DNA strand in a sequence-independent manner (**Fig. 2b**,**d**), instead of binding to specific bases as observed in most other ZF–DNA complexes[14], and supports a general proofreading role for Hnt3 in DNA repair. All of the ZF-dsDNA interactions are made outside the grooves of DNA and on their surface, consistent with the rapid on-off kinetics observed in surface plasmon resonance (SPR) experiments (**Supplementary Fig. 2a**,**b**). Notably, these phosphate-binding residues (Arg209, Phe211 and Thr212) are not well conserved (for example, in our study, Arg209 of Hnt3 was missing in other aprataxin orthologs) (**Supplementary Fig. 3**), which might affect the DNA-binding activity and in turn influence the properties of the different aprataxin proteins.

The third part of the molecular platform is the conserved HIT active site (Asn145, His147, His149 and the additional fourth histidine His71)[9], which binds an AMP molecule (**Fig. 3a**,**b** and **Supplementary Fig. 4a**). The Hnt3-AMP interactions are similar to those observed in the human HINT1 structures (**Supplementary Fig. 4b**). His71 (equivalent to His201 of APTX) interacts with His149 and Lys67 through hydrogen bonds. These bonds could be disrupted by the mutation of His71 and, in turn, impair the DNA binding and deadenylation activity of proteins, which explains why more than half of the mutations linked to AOA1 map to His201 or to its immediate vicinity (**Supplementary Fig. 3**). By analogy with the HINT1 complex (PDB 1KPE[13]), the conserved His138 may contact the α- and β-phosphates of the 5′,5′-adenylyl pyrophosphoryl DNA (AppDNA) substrate, indicating an important role for His138 in the catalytic reaction but not in DNA binding (**Supplementary Fig. 2c** and **Supplementary Fig. 4b**). Notably, the ternary complex structure shows the detailed characteristics of the catalytic products (**Fig. 3b**). Two water molecules (1W and 2W) located near the His147 residue are conserved in our three Hnt3 and HINT1 structures. Therefore, our structures provide evidence for the previously proposed two-step catalytic mechanism[9], in which aprataxin first breaks the phosphodiester bond of the AppDNA substrate by transferring the AMP onto the second histidine residue of the HIT motif (His147 for Hnt3 and His260 for APTX) and next

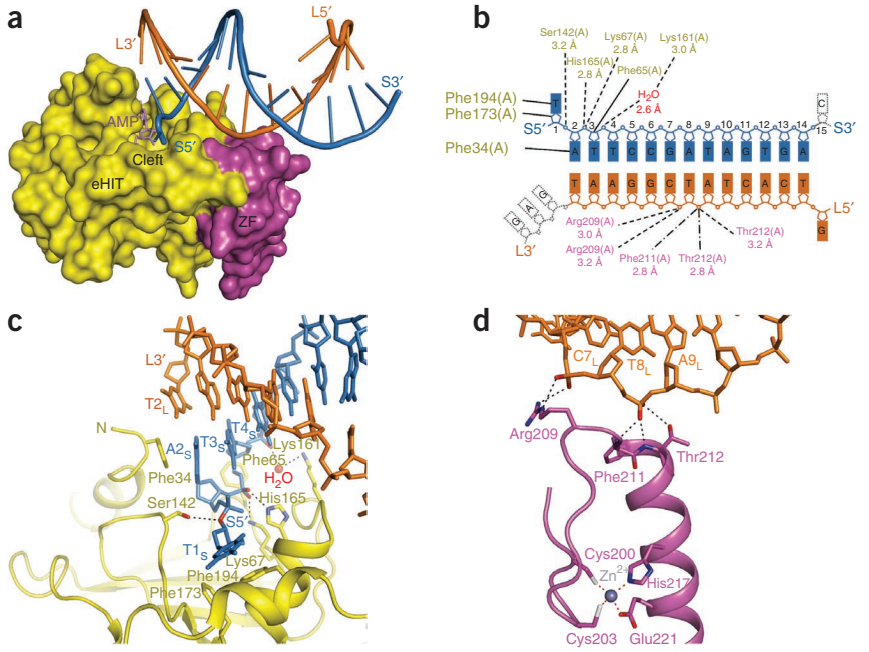

Figure 2 Interactions between Hnt3 and dsDNA. (**a**) Surface representation of Hnt3. The eHIT domain forms a DNA-binding cleft adjacent to the active site. Two DNA strands are bound through the cleft and ZF domain. AMP is bound in the active site. (**b**) Schematic diagram summarizing the Hnt3-DNA interactions. (A), monomer A. Hydrogen bonds mediated by side chains and main chains, dashed and dashed-and-dotted lines, respectively; stack interactions, black lines. The four nucleotides disordered in the structure are shown inside dotted gray boxes. (**c**) Closeup view of the interactions between the cleft and 5′ end of the short DNA strand. DNA-contacting residues are shown as stick representations. Hydrogen bonds are shown as black dashed lines. (**d**) Closeup view of the ZF-DNA interface in the same representations as those in **c**. The conserved zinc-finger motif coordinating a zinc ion is also shown.

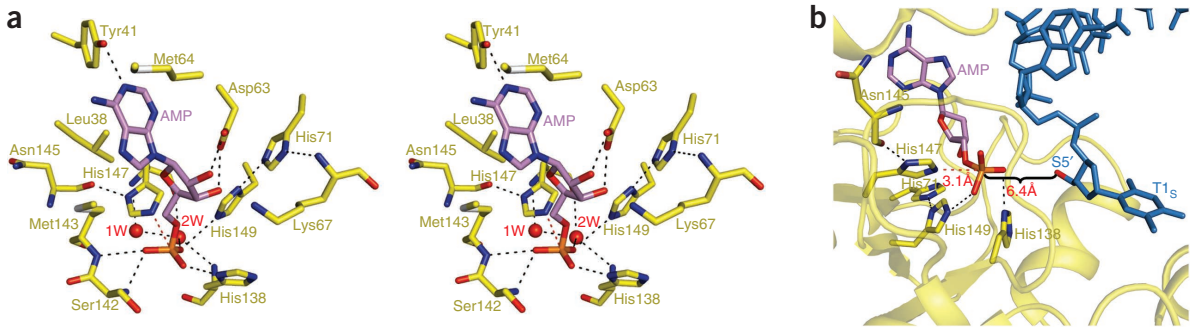

Figure 3 AMP recognition by Hnt3. (**a**) Stereo view of the Hnt3-AMP interactions. AMP-contacting residues are shown as stick representations. Hydrogen bonds are shown as black dashed lines. Two structurally conserved water molecules are represented as sphere models. (**b**) The structure shows the characteristics of the cleavage products. The critical catalytic residue His147 is ~3.1 Å from the α-phosphate of AMP. The 5′-flap nucleotide points its 5′-hydroxyl group to the α-phosphate of AMP, with a distance of ~6.4 Å between them.

activates a water molecule to hydrolyze the phosphohistidine bond to release AMP and regenerate the enzyme.

Our structure-based deductions described above are generally consistent with previous biochemical studies[4–6,9]. We therefore suggest that our complex structures reflect the physiological assembly mode for DNA deadenylation. They reveal that Hnt3 binds and processes 5′-adenylated DNA in a structure-specific manner (**Supplementary Fig. 5**). Hnt3 binds asymmetrically to the DNA almost at the 3′ side of the adenylate through its two 'molecular hands', the eHIT domain and the ZF domain, covering the major groove of the DNA and spanning a ~9-bp DNA duplex region. These interactions result in a predictable DNA bend proximal to the 5′ side of the adenylate and the disruption of the 5′-adenylated base pair at nicks. This frayed 5′ nucleotide adopts a flap conformation to facilitate 5′-AMP docking at the active site, which uses a canonical HIT superfamily catalytic mechanism to cleave the AMP. In addition, these structures illuminate the substrate specificity of Hnt3 as a DNA nick sensor, which explains why aprataxins prefer to bind 5′-adenylated nicks and double-strand breaks. These results support the idea that the HIT–ZF domain is sufficient for efficient DNA processing and that the FHA domain of APTX may coordinate interactions with other proteins during DNA repair[9,15,16]. Future studies will determine how these protein interactions regulate the enzymatic activity of APTX. Nevertheless, the high-resolution structures reported here not only provide a framework for understanding the mechanism of DNA deadenylation but also shed new light on the molecular basis of AOA1 (see **Supplementary Fig. 6** and **Supplementary Discussion**).

Accession codes. Protein Data Bank: Coordinates and structure factor files have been deposited with the accession codes 3SP4 (apo Hnt3), 3SPD (Hnt3–DNA) and 3SPL (Hnt3–DNA–AMP).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We are grateful to Y. Chen for her assistance in Biacore analysis. We thank the staff at the Shanghai Synchrotron Radiation Facility beamline 17U for assistance with data collection, and we thank our colleagues for critical comments on the manuscript. This work was funded by Chinese Ministry of Science and Technology '973' grants 2011CB910302 and 2011CB910304 to T.J. and D.-C.W.; National Natural Science Foundation of China grants 31021062 and 31025009 to T.J. and 31000330 to D.Z.; and Chinese Academy of Sciences grant KSCX2-EW-J-3 to J.D. and D.-C.W.

AUTHOR CONTRIBUTIONS

Y.G., D.Z., J.D., T.J. and D.-C.W. designed the research; Y.G., D.Z., J.D., X.R. and C.-N.D. carried out the experiments; Y.G., D.Z., J.D., T.J. and D.-C.W. analyzed the data and wrote the paper; and all authors contributed to editing the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Aicardi, J. *et al. Ann. Neurol.* **24**, 497–502 (1988).
2. Date, H. *et al. Nat. Genet.* **29**, 184–188 (2001).
3. Moreira, M.C. *et al. Nat. Genet.* **29**, 189–193 (2001).
4. Ahel, I. *et al. Nature* **443**, 713–716 (2006).
5. Kijas, A.W., Harris, J.L., Harris, J.M. & Lavin, M.F. *J. Biol. Chem.* **281**, 13939–13948 (2006).
6. Rass, U., Ahel, I. & West, S.C. *J. Biol. Chem.* **282**, 9469–9474 (2007).
7. Rass, U., Ahel, I. & West, S.C. *Cell* **130**, 991–1004 (2007).
8. Caldecott, K.W. *Nat. Rev. Genet.* **9**, 619–631 (2008).
9. Rass, U., Ahel, I. & West, S.C. *J. Biol. Chem.* **283**, 33994–34001 (2008).
10. Sykora, P., Croteau, D.L., Bohr, V.A. & Wilson, D.M. III. *Proc. Natl. Acad. Sci. USA* **108**, 7437–7442 (2011).
11. Brenner, C. *Biochemistry* **41**, 9003–9014 (2002).
12. Deshpande, G.P. *et al. DNA Repair (Amst.)* **8**, 672–679 (2009).
13. Lima, C.D., Klein, M.G. & Hendrickson, W.A. *Science* **278**, 286–290 (1997).
14. Wolfe, S.A., Nekludova, L. & Pabo, C.O. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 183–212 (2000).
15. Clements, P.M. *et al. DNA Repair (Amst.)* **3**, 1493–1502 (2004).
16. Becherel, O.J. *et al. Nucleic Acids Res.* **38**, 1489–1503 (2010).