Published in final edited form as:
Science. 2014 November 28; 346(6213): 1127–1130. doi:10.1126/science.1258973.

Mechanism of DNA Interstrand Crosslink Processing by Repair Nuclease FAN1

Renjing Wang[1], **Nicole Persky**[1], **Barney Yoo**[2], **Ouathek Ouerfelli**[2], **Agata Smogorzewska**[3], **Stephen J. Elledge**[4,5], and **Nikola P. Pavletich**[1,5,*]

[1]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

[2]Molecular Pharmacology and Chemistry Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

[3]Laboratory of Genome Maintenance, The Rockefeller University, New York, NY 10065, USA

[4]Department of Genetics and Howard Hughes Medical Institute, Harvard Medical School, Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA

[5]Howard Hughes Medical Institute

Abstract

DNA interstrand crosslinks (ICLs) are highly toxic lesions associated with cancer and degenerative diseases. ICLs can be repaired by the Fanconi Anemia (FA) pathway and through FA-independent processes involving the FAN1 nuclease. Here, FAN1-DNA crystal structures and biochemical data reveal that human FAN1 cleaves DNA successively at every third nucleotide. *In vitro*, this exonuclease mechanism allows FAN1 to excise an ICL from one strand through flanking incisions. DNA access requires a 5'-terminal phosphate anchor at a nick or 1–2 nucleotide flap, and is augmented by a 3' flap, suggesting FAN1 action is coupled to DNA synthesis or recombination. FAN1's mechanism of ICL excision is well suited for processing other localized DNA adducts as well.

DNA interstrand crosslinks (ICLs) covalently link together the two strands of the double helix (1). ICLs can be repaired by the Fanconi Anemia (FA) pathway of proteins mutated in the FA cancer predisposition syndrome (2), and also by FA-independent processes (1, 3, 4). The FA pathway is activated when a replication fork stalls upon encountering an ICL, leading to unhooking of the ICL through flanking incisions on one strand, translesion DNA synthesis across the unhooked ICL, removal of the ICL remnants by additional incisions, and homologous recombination (2, 5). The FAN1 nuclease (6–9) is a candidate for mediating FA-independent repair (1, 3, 4). While FAN1 mutations result in defective ICL repair, chromosomal instability, and hypersensitivity to a wide range of ICL-inducing agents (6–13), they do not cause FA (10–12). Instead, they cause a kidney degeneration disorder, karyomegalic interstitial nephritis (KIN) (11). FAN1 cleaves branched DNA structures with a preference for 5' flap substrates, and also exhibits 5' to 3' exonuclease activity on a variety

*To whom correspondance should be addressed. pavletin@mskcc.org.

of dsDNA and ssDNA substrates (6–9). The functional implications of these broadly-defined activities, whether FAN1 processes ICLs directly and in which manner, and the structural mechanism of ICL unhooking by nucleases are unknown.

To address these questions, we first investigated FAN1's DNA-binding specificity for various lengths of 5' flaps consisting of thymidine nucleotides (nts). We found that FAN1 exhibits specificity for a 5' phosphate group at flap of 1 to 2 nts or at a nick, with dissociation constants (K_d) of 10.3 nM for the 1-nt flap (5'pT_1), 121 nM for 2 nts (5'pT_2), and 182 nM for the nick (pNi) (fig. S1A). The corresponding substrates lacking the 5' phosphate group fail to bind appreciably until micromolar FAN1 concentrations (fig. S1A). Similarly, increasing the flap length beyond two nucleotides reduces FAN1 affinity substantially (fig. S1A). We also tested the effect of adding a 3' flap of 8 thymidine nts to the high-affinity 5'pT_1 (5'pT_1/3'T_8), 5'pT_2 (5'pT_2/3'T_8) and 5'pNi (5'pNi /3'T_8), and found that this increased their FAN1 affinity by a factor of ~25 to ~180 (K_d values of 0.4, 1.4, and 1.0 nM, respectively; fig. S1B). By contrast, adding a 3'T_8 flap to the low-affinity 5'pT_8 (5'pT_8/3'T_8) minimally enhanced its affinity, with the resulting 377 nM K_d a factor of ~1000 weaker than that of 5'pT_1/3'T_8 (fig. S1B).

Based on these findings, we determined the 2.9 Å structure of a 5'pG_1/3'T_4 DNA bound to human FAN1 (residues 364 to 1017), N-terminally truncated to remove the ubiquitin-binding UBZ domain and subsequent unstructured segment, as well as the structure of apo-FAN1 (fig. S2A, and table S1). FAN1 has a bi-lobal structure consisting of a 223-residue N-terminal domain (NTD) and a 415-residue C-terminal domain (CTD) that contains the PD-(D/E)XK nuclease motif (Fig. 1A and B). The DNA adopts a V-shaped structure, with a kink at the nick. The pre-nick and post-nick duplexes have an overall B-DNA conformation, with a 76° angle between them. The pre-nick duplex and 3' flap are bound by the NTD, while the 5' flap and post-nick duplex are bound by the CTD. There are no major conformational changes between the apo- and DNA-bound FAN1 structures (fig. S2B).

The NTD consists of a helical domain, a winged-helix (WH) DNA-binding domain, and the predicted SAP domain (Fig. 1C). These form a continuous surface that binds to a 9-bp duplex segment leading to the 3' flap, and to a 2-nt segment of the flap (Fig. 1A, fig. S3A to C). The duplex contacts involve only phosphodiester groups (fig. S3A to C), except for the flap-proximal base pair, which is contacted at both its phosphodiester and base groups (Fig. 2A). These base contacts block the DNA from extending as a regularly-stacked duplex, analogous to the helix-breaking wedges observed with other structure-specific nucleases (14, 15).

Only the first two nucleotides of the 3'T_4 flap are ordered. These extend away from the duplex and are bound by FAN1 (Fig. 2A). The base and ribose groups of the first nucleotide are contacted by Tyr374, Val577 and Arg581. The phosphate group of the second nucleotide binds to a pocket at the amino-terminus of an alpha helix and hydrogen bonds to a backbone amide group (Tyr374) in a partially buried environment. These 3' flap contacts are consistent with addition of a single-nucleotide 3' flap to 5'pT_1 resulting in roughly half of the affinity increase compared to an 8-nt flap (5'pT_1/3'T_1 K_d of 1.9 nM, fig. S1B).

The CTD starts with a 180-residue helical repeat domain that belongs to the TPR family, followed by a nuclease domain that has the ~120-residue αββββαβ fold of the PD-(D/E)XK superfamily of Mg^{2+}-dependent nucleases (16, 17) but which additionally contains a 125-residue helical insertion (Fig. 1C). The CTD has a wide and deep groove, with the TPR domain on one side and the nuclease and insertion domains on the other, into which the post-nick duplex binds (Fig. 1B). FAN1 contacts the 5' phosphate group of the flap and eight phosphodiester groups clustered at the flap-distal portion of the duplex (Fig. 2B). The backbone of the one-nucleotide flap extends orthogonally from the duplex. The 5' phosphate group abuts a highly basic pocket at the bottom of the CTD groove, with its three non-bridging oxygen atoms hydrogen bonding with the side chains of Arg706, Arg952, His742 and Lys986 (Fig. 2C and D, and fig. S4A). These contacts explain FAN1's strict requirement for a 5' phosphate group. Accordingly, the Arg706Ala/Arg952Ala double mutation reduces FAN1's affinity for 5'pT_1/3'T_8 by a factor of ~400 (210 nM K_d, fig. S4B). After the 5' phosphate group, the next two phosphodiester groups make relatively few contacts, while a non-bridging oxygen of the third phosphodiester group coordinates the Ca^{2+} ion at the nuclease active site (Fig. 2D).

The Ca^{2+} ion, which does not support nuclease activity (18), has a coordination shell similar to other nucleases (17, 18) (fig. S4C-E). This indicates that cleavage occurs at the third phosphodiester group from the 5' phosphate (n+3 register) and results in a 5' phosphorylated product, both of which we confirmed by liquid chromatography-coupled mass spectroscopy (LC-MS) of a nuclease reaction.

In additional co-crystal structures we determined (table S1 and fig. S5A), the 5' phosphate groups of 5'pT_1/3'T_8, 5'pT_1 and 5'pNi bind to the same FAN1 pocket and contact the same four side chains as in the 5'pG_1/3'T_4 structure (fig. S5A and C). By contrast, the pocket is devoid of DNA in the co-crystal structure of the 5'-hydroxyl 5'T_2, consistent with FAN1's specificity for a terminal phosphate. The 5'pT_1/3'T_8, 5'pT_1 and 5'pNi structures also suggest some inherent flexibility in the register of the scissile phosphate relative to the 5' phosphate group. In the 5'pT_1/3'T_8 and 5'pT_1 structures, an extra nucleotide is extruded out in a loop as the 5' flap backbone transitions to the duplex, resulting in an n+4 scissile phosphate register (fig. S5B and C). And in the 5'pNi structure, a shifted trajectory for the duplex results in an n+2 register (fig. S5C and D).

The importance of a 5' phosphate group on a short flap/nick is recapitulated in FAN1 nuclease assays. In the presence of 150 mM NaCl and with 100 nM substrate, 5'pT_1/3'T_8, pNi/3'T_8 and 5'pT_2/3'T_8 are cleaved starting with 0.1 nM FAN1, whereas the longer 5' flap 5'pT_8/3'T_8 substrate is not noticeably cleaved even with 100 nM FAN1 (Fig. 3A, fig. S6). 5'pT_8/3'T_8 cleavage required very low ionic strength, such as the 25 mM NaCl used in previous studies, but its extent was still substantially lower than that of 5'pT_1/3'T_8 (Fig. 3A, fig S6).

The n+3 scissile phosphate register suggests that FAN1's reported 5' to 3' exonuclease activity results from successive incisions releasing 3-nt fragments. In the time course of a 1 nM stoichiometric FAN1-5'pT_1/3'T_8 reaction, an 18-nt product corresponding to an n+3 incision appears first, its level peaking at 8 min then declining as it gets consumed (Fig. 3B).

A 15-nt product (two n+3 incisions) appears second and peaks at 16–32 min, while a 12-nt product (third incision) appears last. The 5'pNi/3'T_8 substrate produces a similar pattern of sequential cuts with a major n+3 spacing (Fig. 3B). Minor products corresponding to n+4 and n+2 incisions likely reflect the alternate scissile phosphate registers of the 5'pT_1/3'T_8 and 5'pNi co-crystal structures, respectively (fig. S5C).

Successive incisions do not occur in a detectably processive manner (fig. S7A), but meet the criteria of a distributive activity. FAN1 binds to a 3-nt gap substrate (pGap_3/3'T_8), which corresponds to the first incision product of 5'pNi/3'T_8, with a 3.6 nM K_d, compared to 1.0 nM for 5'pNi/3'T_8 (fig. S7B), and also cleaves it to generate products equivalent to the second and third products of pNi/3'T_8 (Fig. 3C). To determine the effect of increasing gap length, we assayed a 5'pT_1/3'T_8 substrate containing a long, 35-bp post-nick duplex. The sequential incisions stalled after four cuts even in a 100 nM stoichiometric reaction (fig. S7C), suggesting that a gap length beyond ~12 nts of ssDNA, which presumably loops out at the NTD-CTD interface, substantially reduces FAN1 binding efficiency. In vivo, the stalling could be inconsequential if FAN1 acts at a site with a complementary 3' flap, or in cooperation with a distributive polymerase, either of which can close the emergent gap and restore high affinity binding.

The n+3 exonuclease mechanism is well suited for unhooking an ICL through two successive incisions flanking it. To test this, we synthesized a stable ICL consisting of two cytosines linked at their N4 positions through a triazole moiety (19) (Fig. 4A, and fig. S8). We first assayed a triazole ICL with the crosslinked cytosine positioned 6 nts after the 5' phosphorylated nick of a 5'pNi/3'T_8 substrate (ICL-6). In a 1 nM stoichiometric reaction, FAN1 robustly cleaved past the ICL, giving rise to post-ICL products that match those of the unmodified control substrate (Crl-6) with similar rates (Fig. 4B and C). We confirmed that this is due to dual incisions 5' and 3' to the ICL by LC-MS analysis of the 64 minute reaction, which identified a major molecular species (37 % of the LC-UV quantitation; table S2) consisting of the continuous strand, the triazole linkage and a three nucleotide fragment of Cyt, Gua and Thy (11653.995 and 11653.387 Da observed and calculated mw, respectively; table S2). The ICL-6 products (Fig. 4D) also demonstrate that FAN1 can cleave immediately 5' to the ICL, with the 5' phosphodiester group of the crosslinked cytosine at the scissile position.

We next addressed whether FAN1 can unhook the triazole ICL irrespective of its register relative to the 5' phosphate group. For this, we tested ICLs with the crosslinked cytosine positioned 5, 4 or 3 nts after the 5' phosphorylated nick (ICL-5, ICL-4 and ICL-3). FAN1 unhooked ICL-4 and ICL-3 at faster rates compared to ICL-6, consistent with this requiring only two incisions compared to three, while the rate of ICL-5 unhooking was intermediate (Fig. 4B to D). LC-MS and denaturing PAGE of the ICL-5 reaction also indicated that FAN1 can cleave immediately 3' to the ICL (Fig. 4B to D, and table S2).

Diverse ICLs can cause a wide range of DNA-structure distortions, but these are mostly limited to one or two base pairs flanking the lesion (20). The structurally and biochemically observed +/− 1 nt shifts in the scissile phosphate register, coupled to the paucity of FAN1-DNA contacts in the immediate vicinity of the scissile phosphate group, suggest that the

FAN1 active site may have substantial flexibility for accommodating such distortions. Our DNA-binding and cleaving data strongly suggests that FAN1 acts at sites that have a 5' phosphate group on a short flap or nick, possibly generated by an upstream nuclease. Such sites may also arise from the regression of longer 5' phosphorylated flaps with the concomitant generation of a complementary 3' flap. If FAN1 accesses the DNA at such a site, it can keep excising with 3' flap annealing until it reaches the ICL and unhooks it.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Supported by HHMI (NP and SE) and R01HL120922 (AS). AS is a Rita Allen Foundation Scholar and recipient of a Burroughs Wellcome Career Award and Doris Duke Clinical Scientist Development Award. The PDB accession codes are 4RI8 (5'pG_1/3'T_4), 4RI9 (5'pT_1/3'T_8), 4RIA (5'pNi), 4RIB (5'pT_1), 4RIC (5'T_2) and 4RID (apo-FAN1).

References

1. Clauson C, Schärer OD, Niedernhofer L. Cold Spring Harb Perspect Biol. 2013; 5:a012732. [PubMed: 24086043]
2. Kim H, D'Andrea AD. Genes Dev. 2012; 26:1393–1408. [PubMed: 22751496]
3. Williams HL, Gottesman ME, Gautier J. Trends Biochem. Sci. 2013; 38:386–393. [PubMed: 23830640]
4. Deans AJ, West SC. Nat. Rev. Cancer. 2011; 11:467–480. [PubMed: 21701511]
5. Long DT, Räschle M, Joukov V, Walter JC. Science. 2011; 333:84–87. [PubMed: 21719678]
6. Kratz K, et al. Cell. 2010; 142:77–88. [PubMed: 20603016]
7. Liu T, Ghosal G, Yuan J, Chen J, Huang J. Science. 2010; 329:693–696. [PubMed: 20671156]
8. MacKay C, et al. Cell. 2010; 142:65–76. [PubMed: 20603015]
9. Smogorzewska A, et al. Mol. Cell. 2010; 39:36–47. [PubMed: 20603073]
10. Yoshikiyo K, et al. Proc. Natl. Acad. Sci. U. S. A. 2010; 107:21553–21557. [PubMed: 21115814]
11. Zhou W, et al. Nat. Genet. 2012; 44:910–915. [PubMed: 22772369]
12. Trujillo JP, et al. Blood. 2012; 120:86–89. [PubMed: 22611161]
13. Fontebasso Y, Etheridge TJ, Oliver AW, Murray JM, Carr AM. DNA Repair. 2013; 12:1011–1023. [PubMed: 24192486]
14. Tsutakawa SE, et al. Cell. 2011; 145:198–211. [PubMed: 21496641]
15. Orans J, et al. Cell. 2011; 145:212–223. [PubMed: 21496642]
16. Kinch LN, Ginalski K, Rychlewski L, Grishin NV. Nucleic Acids Res. 2005; 33:3598–3605. [PubMed: 15972856]
17. Pingoud A, Fuxreiter M, Pingoud V, Wende W. Cell. Mol. Life Sci. 2005; 62:685–707. [PubMed: 15770420]
18. Hadden JM, Déclais AC, Carr SB, Lilley DM, Phillips SE. Nature. 2007; 449:621–624. [PubMed: 17873858]
19. Presolski SI, Hong V, Cho SH, Finn MG. J. Am. Chem. Soc. 2010; 132:14570–14576. [PubMed: 20863116]
20. Muniandy PA, Liu J, Majumdar A, Liu ST, Seidman MM. Crit. Rev. Biochem. Mol. Biol. 2010; 45:23–49. [PubMed: 20039786]

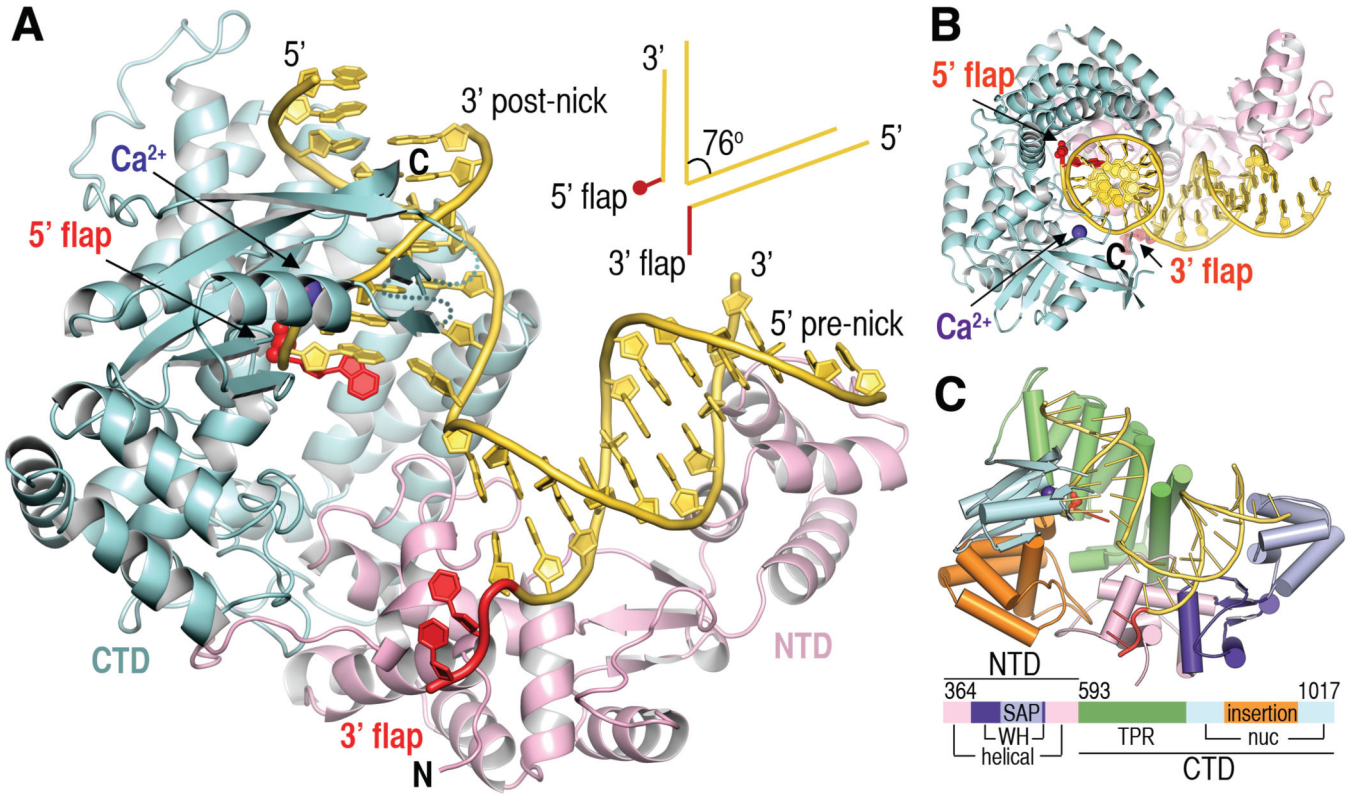

Fig. 1. Overall Structure of FAN1–5'pG$_1$/3'T$_4$

(**A**) Cartoon representation with the NTD colored pink, CTD cyan, dsDNA yellow, flaps and the 5' phosphate spheres red, and Ca^{2+} purple. The inset shows a line drawing of the DNA.

(**B**) View looking down the vertical axis of (A).

(**C**) The six different domains of FAN1 are colored according to the linear representation.

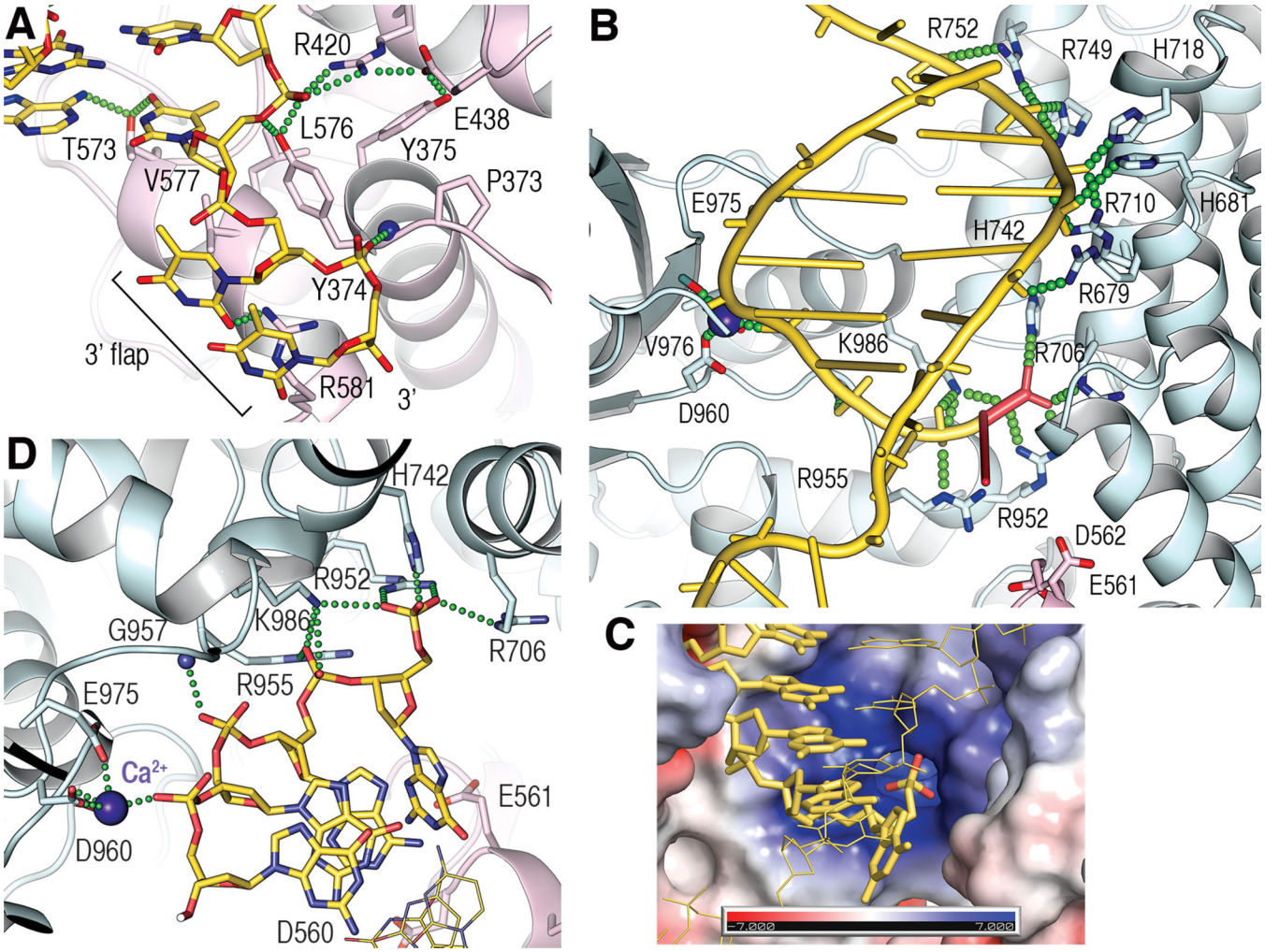

Fig. 2. FAN1-DNA contacts

(**A**) Close-up view of 3' flap contacts showing side chains involved in either hydrogen bond (green dotted lines) or van der Waals contacts.

(**B**) Overview of the CTD-DNA contacts, with the 5' flap and its phosphate group in red.

(**C**) Surface representation of the 5' phosphate-binding pocket, colored by the electrostatic potential (+/− 7kT), with the scissile strand shown as yellow sticks, the 5' phosphate oxygen atoms in red, and the continuous strand as thin yellow lines.

(**D**) Close-up view of contacts to the 5' phosphorylated flap and the three following nucleotides of the scissile strand. Continuous strand shown as thin lines.

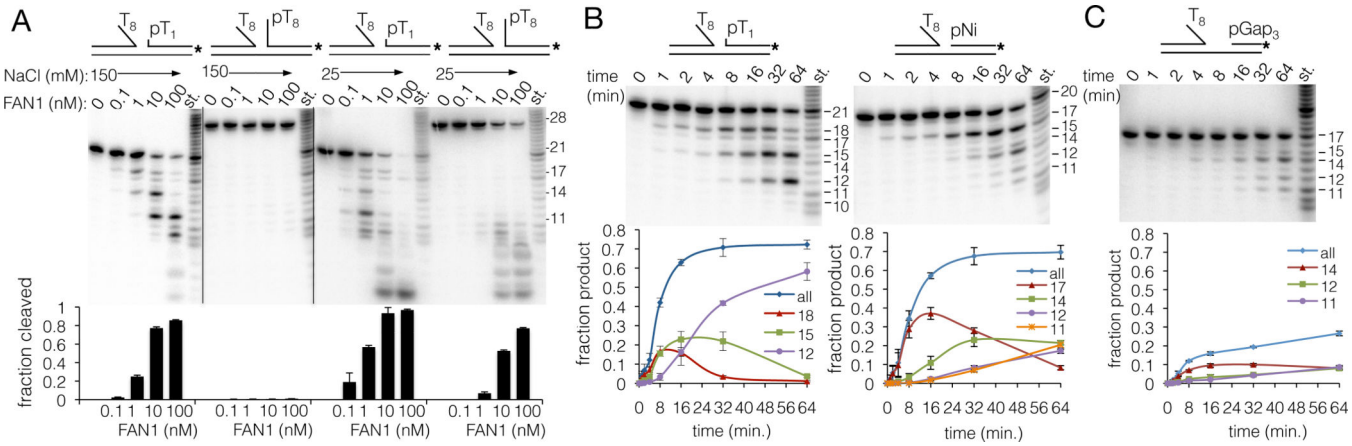

Fig. 3. Nuclease activity

(**A**) Denaturing PAGE of nuclease reactions with the indicated substrates (100 nM), FAN1 and NaCl concentrations. The substrate scissile strand and mw standards (st.) are 3’ α-^{32}P dATP labeled (asterisk). Column graph shows the sum of all cleaved bands divided by total DNA, as the mean +/− s.d. from three experiments.

(**B**) Nuclease time course of 5’pT_1/3’T_8 (left) and 5’pNi/3’T_8 (right) in 150 mM NaCl. Bands and the corresponding mw standards are marked. Fractional amounts of each major product and their total (“all”) are graphed as the mean +/− s.d. from three experiments.

(**C**) Nuclease time course of pGap_3/3’T_8, as in (B).

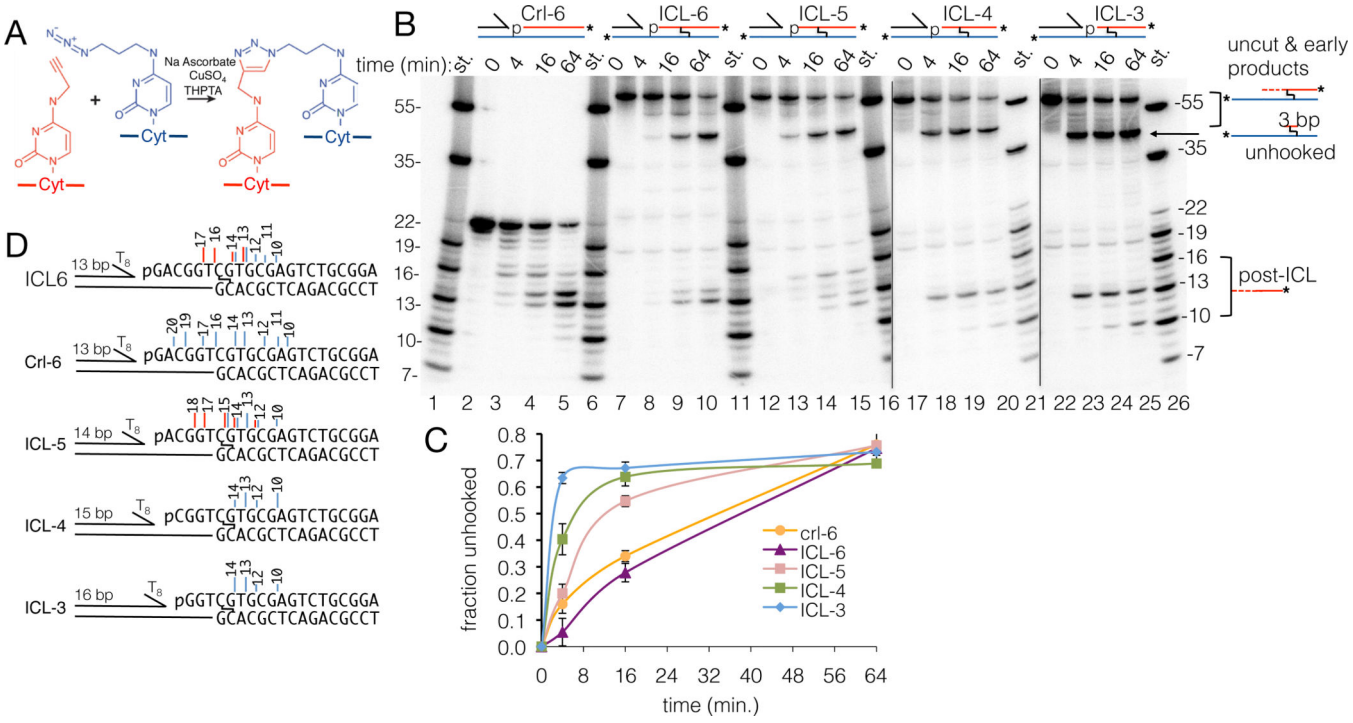

Fig. 4. ICL unhooking

(A) The Cu-catalyzed azide-alkyne cycloaddition and the expected 1,4-isomer of the triazole ICL.

(B) Nuclease time course of the indicated ICL substrates and the unmodified control substrate (Crl-6), all in 150 mM NaCl. The uncut substrates and the products of the pre-ICL cuts (“uncut & early products”) are not resolved due to the crosslinked nature of the long oligonucleotides (fig. S8). The identity of the unhooked ICL band (“unhooked”) was confirmed by LC-MS of the band eluted from a PAGE gel of the ICL-5 reaction (table S2).

(C) Quantitation of (B), graphing the ratio of “unhooked” bands to “unhooked” plus “uncut”, corrected for the presence of two ^{32}P-labeled strands in “uncut”. The Crl-6 curve shows total cuts, so it initially rises faster than the ICL-6 curve that does not reflect the unresolved cuts 1 and 2.

(D) Vertical lines mark incisions identified by PAGE (blue) and LC-MS (red), with line-height reflecting relative product abundance, and numbers product length (nts).