PNAS
www.pnas.org

Supplementary Information for

Structures and single-molecule kinetic analysis of the mycobacterial motor-nuclease AdnAB illuminate the mechanism of DNA double-strand break resection

Ning Jia, Mihaela C. Unciuleac, Chaoyou Xue, Eric C. Greene,
Dinshaw J. Patel, and Stewart Shuman

Corresponding authors: Stewart Shuman, Dinshaw Patel
Emails: s-shuman@ski.mskcc.org; pateld@MSKCC.ORG

This PDF file includes:

METHODS

Recombinant AdnAB proteins. Wild-type AdnAB heterodimer and mutant heterodimers AdnA(D934A)-AdnB(D1014A) (dual nuclease-inactive mutant), AdnA(D255A)-AdnB (A motor domain motif II mutant), and AdnA-AdnB(W325A-R326A) (B motor domain motif III mutant) were produced in *E. coli* by co-expression of His_{10}-Smt3-tagged AdnA and un-tagged AdnB subunits as described previously (1). All purification steps were performed at 4˚C. Soluble extracts of 2-liter bacterial cultures were prepared by suspending the frozen and thawed cell pellet in 30 ml lysis buffer (50 mM Tris-HCl, pH 8.0, 10% sucrose, 250 mM NaCl), sonicating the suspension on ice for 10 min, and then removing insoluble material by centrifugation for 1 h at 16000 rpm. The resulting supernatants were applied to 10 ml columns of nickel-NTA resin (Qiagen) that had been equilibrated with lysis buffer. The columns were washed with lysis buffer and then eluted stepwise with 20, 100, 300, and 1000 mM imidazole in buffer A (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 10% glycerol). The polypeptide compositions of the eluate fractions were monitored by SDS-PAGE. Recombinant AdnAB was recovered predominantly in 300 mM imidazole eluate. The His_{10}Smt3 tag was removed from AdnAB by treatment with the Smt3-specific protease Ulp1 (at a AdnAB:Ulp1 ratio of ~500:1) during overnight dialysis against buffer A containing 20 mM imidazole. The tag-free AdnAB was separated from the tag by passage over a second nickel-NTA agarose column equilibrated in 20 mM imidazole in buffer A. AdnAB was recovered in the flow-through fraction. The protein concentrations were determined by using the Biorad dye reagent with BSA as the standard.

Preparation of AdnAB•DNA complexes. The DNA ligand used was a 70-mer synthetic oligonucleotide designed to form the forked duplex structure depicted below, comprising a 19-bp duplex with a 5-nucleotide hairpin on one end and a 6-nucleotide 5' tail and a 21-nucleotide 3' tail on the other end. To promote hairpin formation, the DNA stock solution was heated for 5 min at 95°C and quenched on ice prior to incubation with AdnAB.

```
    TT                    TTTTTT 5'
C   ATCGTCACGAGCGTAATCT
CC  TAGCAGTGCTCGCATTAGA
                          TTTTGTTTTTTTAGCGGTTTT 3'
```

In one set of experiments, a mixture of 23 μM AdnA(D934A)-AdnB(D1014A), 48 μM forked DNA, 1 mM AMPPNP, and 3 mM $MgCl_2$ in buffer A was incubated for 30 min on ice. In a separate set of experiments, a mixture of 10 μM wild-type AdnAB, 32 μM forked DNA, 1.6 mM AMPPNP, and 8 mM $MgCl_2$ in buffer A was incubated for 30 min on ice. The samples were gel filtered through a Superdex-200 column equilibrated in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl. The coinciding peak

A_{260} and A_{280} fractions were pooled, concentrated by centrifugal ultrafiltration, and used immediately to prepare the cryo-EM grids.

Cryo-EM sample preparation and data acquisition. Aliquots (3 µl) of ~1.5 mg/ml gel-filtered AdnAB complexes were applied onto glow-discharged UltrAuFoil 300 mesh R1.2/1.3 grids (Quantifoil). Grids were blotted for 1.5 s at ~100% humidity and flash frozen in liquid ethane using an FEI Vitrobot Mark IV. Images were collected on a FEI Titan Krios electron microscope operated at an acceleration voltage of 300 kV with a Gatan K2 Summit detector with a 0.8613 Å pixel size and 8.0 electrons per pixel per second. The defocus range was set from -1.0 µm to -2.5 µm. Dose-fractionated images were recorded with a per-frame exposure time of 200 ms and a dose of ~2.16 electrons per $Å^2$ per frame. Total accumulated dose was ~86 electrons per $Å^2$.

Cryo-EM image processing. For the AdnA(D934A)-AdnB(D1014A)•DNA•AMPPNP complex dataset, motion correction was performed with MotionCor2 (2) Contrast transfer function parameters were estimated by Ctffind4 (3). All other steps of image processing were performed by RELION 2.1 (4). Templates for automated particle selection were generated from 2D-averages of ~2,000 manually picked particles. Automated particle selection resulted in 1,149,116 particles from 2,008 images. After three rounds of 2D classification, a total of 291,139 particles were selected for 3D classification using the initial model generated by RELION as reference. Two separated 3D classes with good secondary structural features were selected for further refinement. One class and the corresponding 80,396 particles were polished using RELION 3D auto-refine and particle polishing, yielding an electron microscopy map of AdnA(D934A)-AdnB(D1014A)•AMPPNP with a resolution of 3.5 Å. And the other class and the corresponding 61,579 were polished using RELION 3D auto-refine and particle polishing, yielding an electron microscopy map of AdnA(D934A)-AdnB(D1014A)•DNA•AMPPNP with a resolution of 3.5 Å.

The dataset for the wild-type AdnAB•DNA complex, was processed by the same procedure as above. Briefly, 393,393 particles were auto-picked from 2,407 images, 60,108 particles were selected for the final 3D reconstruction after two rounds of 2D and 3D classification, resulting in a AdnAB•DNA complex map with at an overall resolution of 3.5 Å.

All resolutions were estimated using RELION ‘post-processing’ by applying a soft mask around the protein density and the Fourier shell correlation (FSC) = 0.143 criterion. Local resolution estimates were calculated from two half data maps using ResMap (5). Further details related to data processing and refinement are summarized in Table S1 and shown schematically in Figure S2.

Model building and refinement. For AdnA(D934A)-AdnB(D1014A)•AMPPNP complex, models for AdnA and AdnB were first generated by SWISS-MODEL online server. Then all docked models were manually rebuilt in COOT (6) to fit the density map. Other parts of the complex were built based on the bulky side chains to register the sequence. Residues in parts of complex were modeled as poly-alanines, or the side chains of parts of the complex were deleted due to the poor side-chain densities.

For the AdnA(D934A)-AdnB(D1014A)•DNA•AMPPNP-DNA and AdnAB•DNA complexes, the model of the AdnA(D934A)-AdnB(D1014A)•AMPPNP complex was docked into the cryo-EM density map using UCSF Chimera (7) and then manually rebuilt in COOT according to densities of each map. All models were refined against summed maps using phenix.real_space_refine (8) by applying geometric and secondary structure restraints. All figures were prepared with PyMol (http://www.pymol.org) or Chimera (7). The statistics for model refinement are shown in Table S1.

Single molecule DNA curtain assays. All experiments were performed with a custom-built prism-type total internal reflection fluorescence (TIRF) microscope (Nikon) equipped with a 488-nm laser (Coherent Sapphire, 200 mM) and two Andor iXon EMCCD cameras. Flowcells and single-tethered dsDNA curtains were prepared as described previously (9) using bacteriophage λ DNA (48.5 kb; NEB Cat No. N3011S) that was biotinylated at one end. All single molecule experiments were performed at 37˚C in resection buffer (20 mM Tris-HCl, pH 8.0, 1 mM DTT, 2 mM $MgCl_2$, 1 mM ATP, 0.2 mg/ml BSA). DNA was aligned along chromium barriers under a constant buffer flow at 0.15 ml/min. dsDNA molecules were stained with 0.1 mM YOYO-1 for 1 min. Free YOYO-1 was flushed out of the flow cell at 0.5 ml/min for 2 min. End resection was then immediately initiated by the injection of 5 nM AdnAB through a 150 µl loop at 0.15 ml/min and unbound AdnAB was flushed from the sample chamber. End resection was then visualized by the gradual loss of YOYO-1 signal. Image acquisition was started immediately prior to the protein injections and continued for the duration of the resection assays (typically 5-7 minutes). All images were acquired at 1 frame per 0.5 s or 2 s with 0.1 s integration time using an EMCCD camera, and the illumination lasers were shuttered between each image to minimize photo-bleaching.

Optical microscopy Image processing and data analysis. Raw TIFF images were imported as image stacks to ImageJ. Images were corrected for drift using the StackReg function in ImageI. Kymographs were then generated from the corrected image stacks by defining a 1-pixel wide region of interest (ROI) encompassing individual dsDNA molecules and these kymographs were used for analysis of AdnAB processivity and velocity as described (9). AdnAB end resection events from the kymographs were fitted with a linear function to calculate AdnAB translocation

velocity for each individual molecule. AdnAB translocation velocities were then obtained from Gaussian fits of the distributions of observed velocities.

Table S1. Cryo-EM Statistics for Data Collection and Model Refinement.

	AdnAB$^{(nuc\text{-}dead)}$ AMPPNP	AdnAB$^{(nuc\text{-}dead)}$ AMPPNP-DNA	AdnAB$^{(WT)}$ AMPPNP-DNA
Data collection			
Magnification	22,500	22,500	22,500
Voltage (kV)	300 kV	300 kV	300 kV
Electron exposure (e^-/ Å^2)	86	86	86
Defocus rang (μm)	–1.0 to –2.5	–1.0 to –2.5	–1.0 to –2.5
Pixel size (Å)	0.8613	0.8613	0.8613
Initial particles (no.)	1,149,116	1,149,116	393,393
Final particles (no.)	80,396	61,579	60,108
Map resolution (Å)	3.5	3.5	3.5
FSC threshold	0.143	0.143	0.143
Map sharpening B factor (Å^2)	–116.50	–114.77	–119.50
Refinement			
Model resolution (Å)	3.52	3.59	3.56
FSC threshold	0.5	0.5	0.5
Model composition			
Protein residues	1,749	1,688	1,353
Nonhydrogen atoms	12,004	11,699	10,104
R.m.s. deviations			
Bond lengths (Å)	0.005	0.012	0.014
Bond angles (°)	0.950	1.255	1.324
Validation			
MolProbity score	1.60	1.90	1.78
Clashscore	3.45	3.01	2.68
Poor rotamers (%)	0.4	1.9	1.7
Ramachandran plot			
Favored (%)	92.7	92.7	92.1
Allowed (%)	7.3	7.3	7.9
Disallowed (%)	0	0	0
PBD ID	6PPJ	6PPR	6PPU

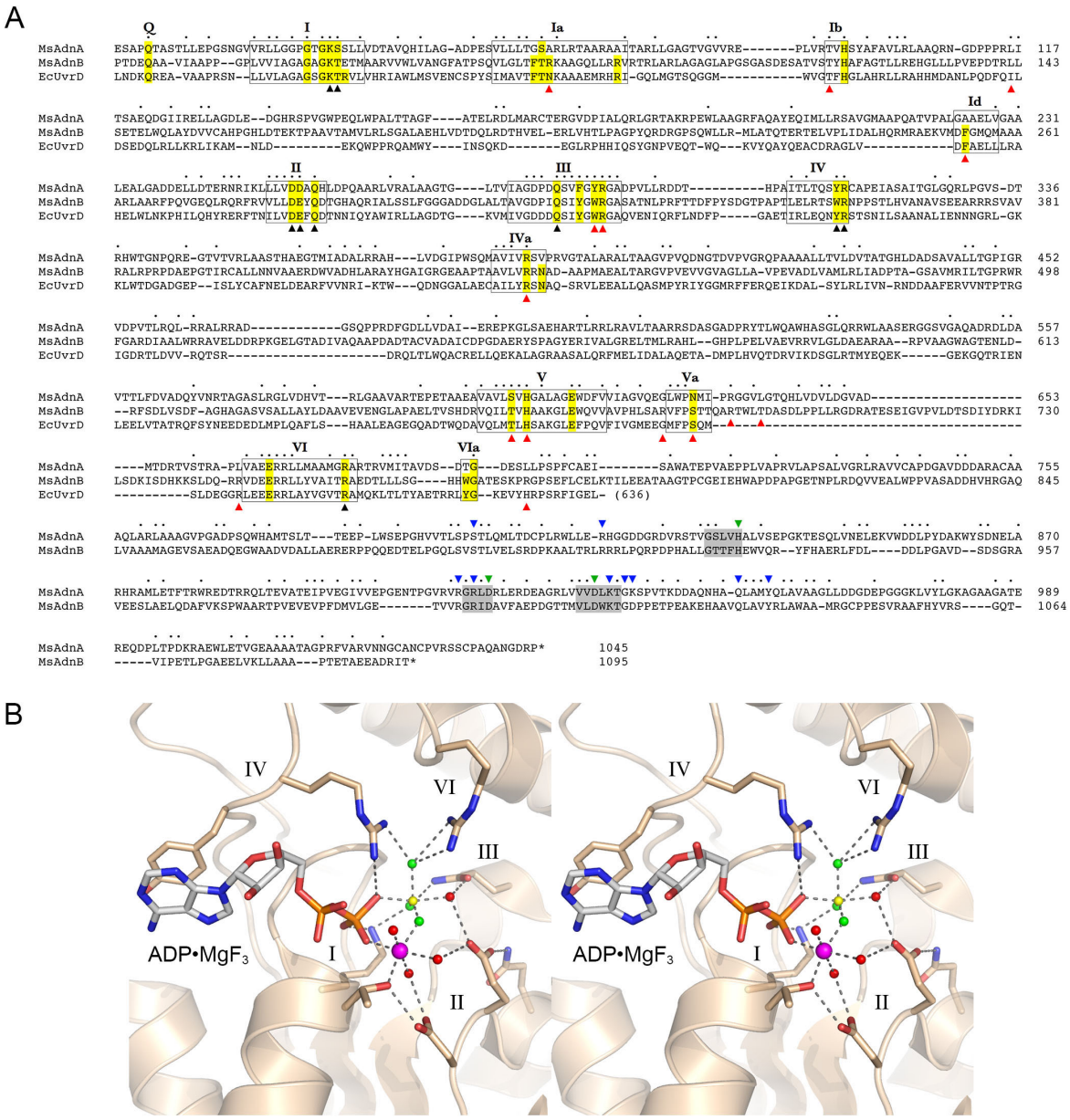
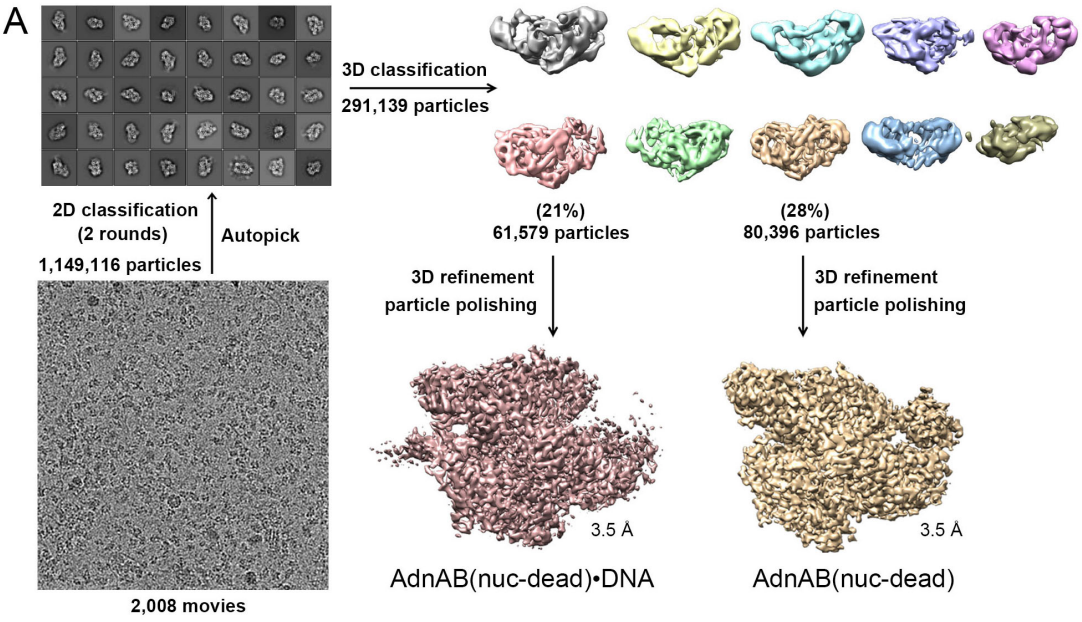
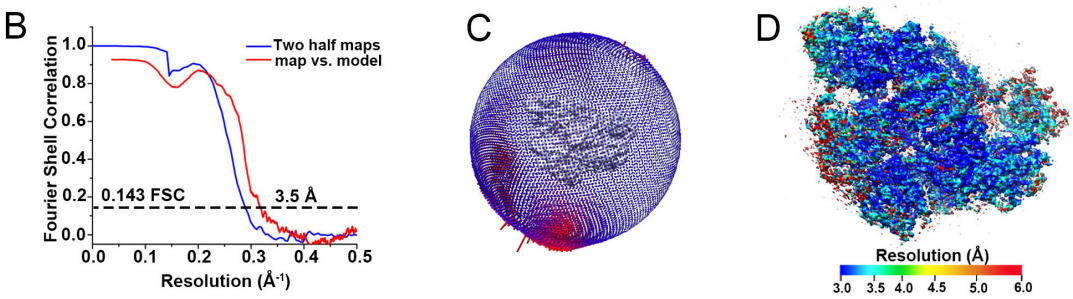


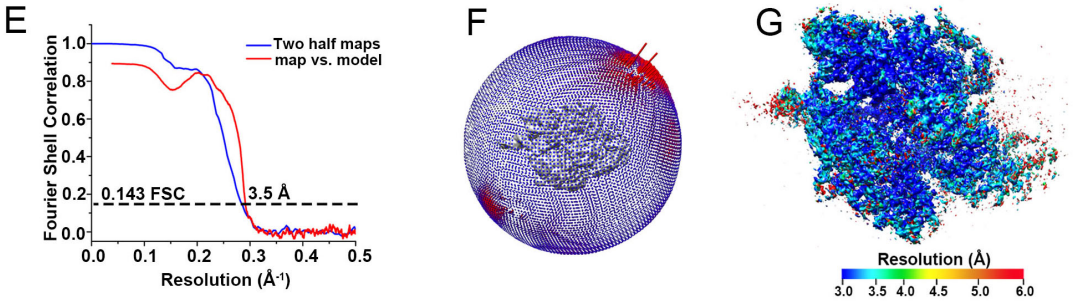
Figure S1. AdnA and AdnB subunits and conservation of their ATPase domain with UvrD. (A) The amino acid sequences of *M. smegmatis* (Ms) AdnA and AdnB are aligned to each other. The N-terminal ATPase domains are aligned to *E. coli* (Ec) UvrD helicase. The ATPase/helicase motifs (10) are denoted in boxes. Amino acids at the ATPase active site and DNA-binding site of UvrD that are conserved in AdnA or AdnB are highlighted in yellow. The C-terminal nuclease domains of AdnA and AdnB contain counterparts of three motifs of RecB nuclease (gray shading). Constituents of the AMPPNP-binding site in AdnA revealed by the AdnAB cryo-EM structure are denoted by black arrowheads below the alignment. AdnB residues that contact the 3' tracking strand or duplex junction of the forked DNA ligand are denoted by red arrowheads. AdnA nuclease domain residues that contact the displaced 5' strand or duplex junction are indicated by blue arrowheads above the alignment. AdnA nuclease residues that bind the Mg cofactor are indicated by green arrowheads. Positions of side chain identity/similarity in all of the aligned proteins and indicated by • above the sequences. (B) Stereo view of the ATPase active site of UvrD (pdb 2IS6) as a transition state mimetic in complex with Mg^{2+}, ADP, and a trigonal planar MgF_3 complex with an apical water nucleophile. Atomic contacts of conserved motif side chains are denoted by dashed lines.



AdnAB(nuc-dead)-AMPPNP apoenzyme



AdnAB(nuc-dead)-AMPPNP–DNA complex



AdnAB(WT)-AMPPNP–DNA complex

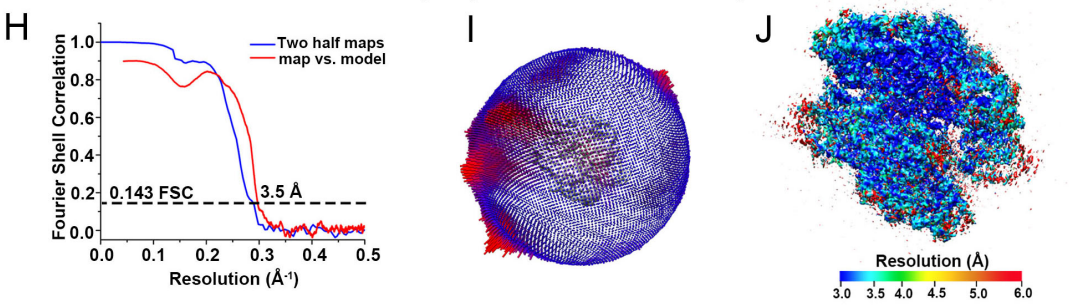


Figure S2. Cryo-EM analysis of AdnAB in complexes with AMPPNP and forked duplex DNA.

(A) Flow chart of cryo-EM image processing for nuclease-inactive AdnA(D934A)-AdnB(D1014A) heterodimers that had been preincubated with AMPPNP and forked DNA duplex. Two of the indicated 3D classes were refined to yield models in which AdnAB(nuclease-inactive) was either bound to DNA (shown at left) or was DNA-free (shown at right). Cryo-EM data for wild-type AdnAB that had been preincubated with AMPPNP and forked DNA duplex were processed and refined via the same procedures.

(B, E, H) Fourier Shell Correlation (FSC) curves of the AdnAB(nuclease-inactive)-AMPPNP DNA-free enzyme (B), the AdnAB(nuclease-inactive)-AMPPNP-DNA complex (E), and the AdnAB(WT)-AMPPNP-DNA complex (H) between two half maps that were calculated from two half datasets (blue curves), and between the cryo-EM map and corresponding model (red curves).

(C, F, I) Euler angle distribution of cryo-EM particles for calculating the final EM map of the AdnAB(nuclease-inactive)-AMPPNP DNA-free enzyme (C), the AdnAB(nuclease-inactive)-AMPPNP-DNA complex (F), and the AdnAB(WT)-AMPPNP-DNA complex (I). The position of each sphere relative to the density map (gray in the center) corresponds to its angular assignment, with the radius of sphere proportional to the number of particles in that orientation.

(D, G, J) Final 3D reconstructed map of the AdnAB(nuclease-inactive)-AMPPNP DNA-free enzyme (D), the AdnAB(nuclease-inactive)-AMPPNP-DNA complex (G) and the AdnAB(WT)-AMPPNP-DNA complex (J) colored according to local resolution.

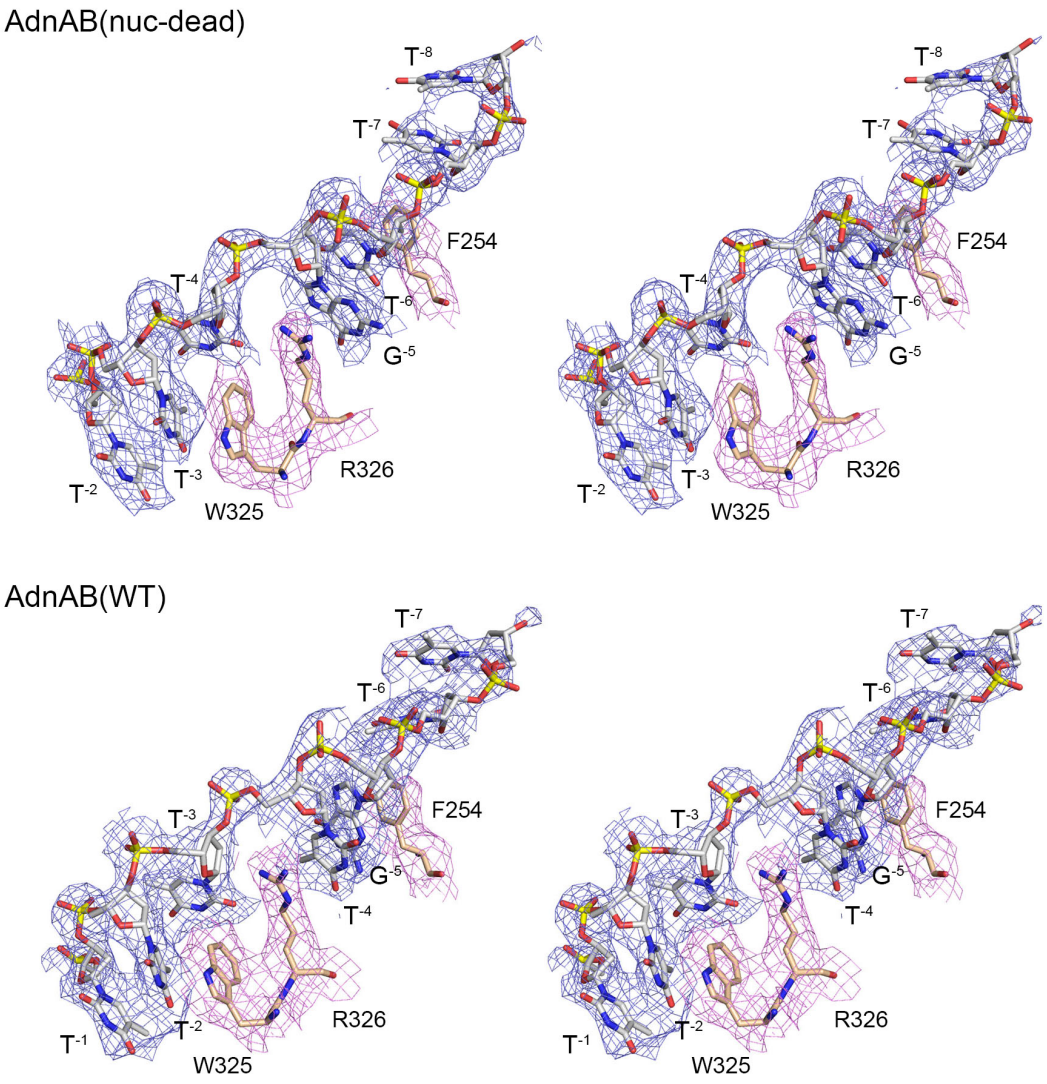


Figure S3. Stereo views of the 3' ssDNA tails transiting through the AdnB motor domain. The AdnAB(nuclease-inactive)–DNA complex is shown in the top panel; the AdnAB(WT)-DNA complex is shown in the bottom panel. Electron density for the ssDNA is rendered as blue mesh. Electron density for AdnB amino acids Trp325, Arg326, and Phe254 that stack on the nucleobases is depicted as pink mesh. The nucleobases are numbered as in Figs. 3B and 4B.

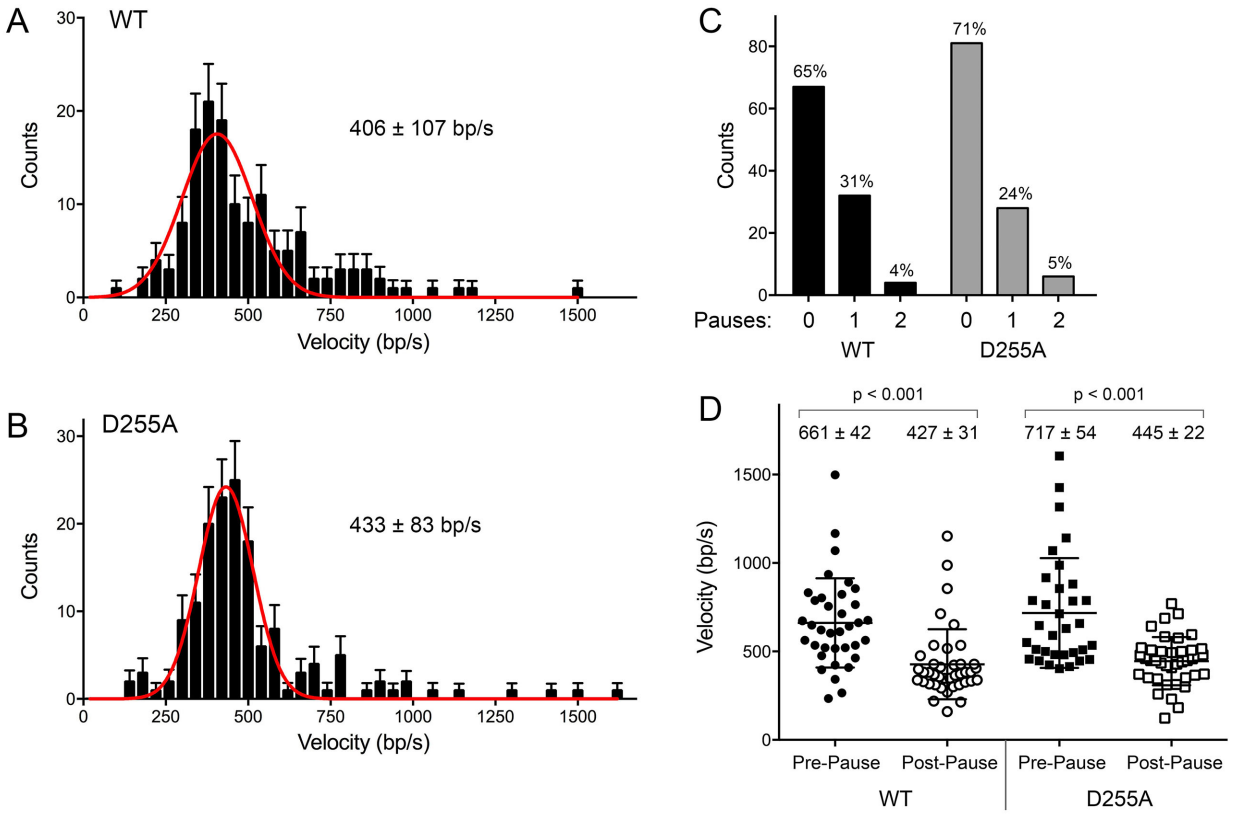


Figure S4. Single-molecule analysis of DSB resection by wild-type AdnAB and the AdnA ATPase active site mutant AdnA(D255A)-AdnB. (A,B) Velocity distributions of individual resection tracts at 1 mM ATP for wild-type AdnAB (panel A; N=143) and the AdnA(D255A)-AdnB mutant (panel B; N=155). Black error bars represent 95% confidence intervals calculated from bootstrap analysis. The solid red lines represent Gaussian fits to the data. The mean velocities (±SD) derived from the Gaussian fit are shown in each panel. (C) The numbers of individual DNA resection events at 1 mM ATP that had no detectable pause (0) or paused once (1) or twice (2) before terminating are plotted for wild-type AdnAB (N = 103 DNAs) and the AdnA(D255A)-AdnB mutant (N = 115 DNAs). The percent distributions are indicated above the bars. The difference in pause frequency between wild-type and mutant AdnAB was not significant (p = 0.082 by t test). (D) Scatter plots of resection velocities prior to and after pausing by wild-type AdnAB (N = 36 pauses) and the AdnA(D255A)-AdnB mutant (N = 34 pauses) at 1 mM ATP. Black lines indicate the mean and 68% confidence intervals. The mean velocity values (±SEM) are shown above the scatter plots along with the p values (unpaired t test).

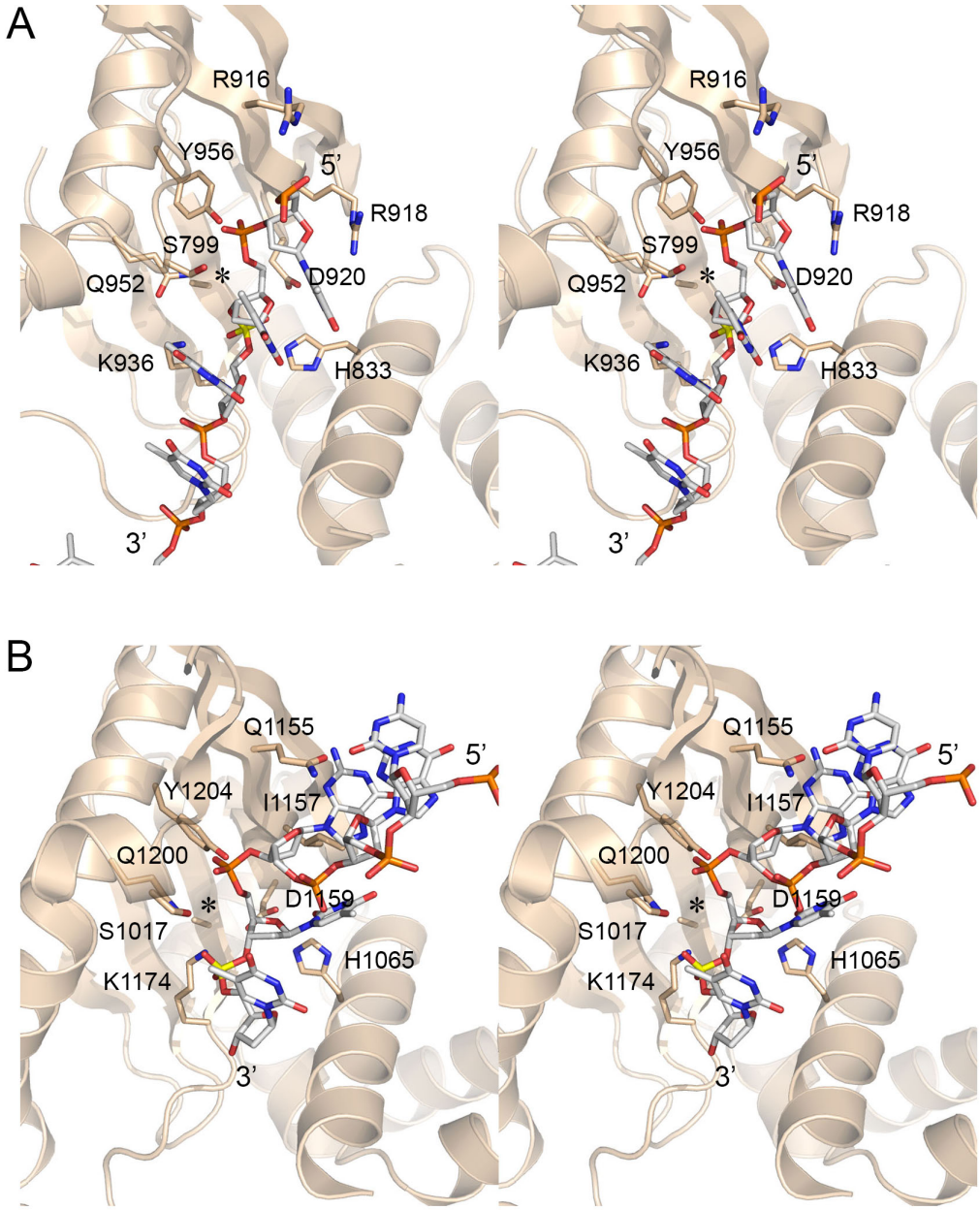


Figure S5. Comparison of the AdnAB 5' DNA-processing AdnA nuclease domain with the AddAB 3' DNA-processing nuclease domain in their respective DNA complexes. Stereo views of the AdnA nuclease (panel A) and AddA nuclease (panel B; ref. 11) aligned with respect to their conserved structural elements. Selected amino acids that engage the DNA are shown a stick models with beige carbons. In both nuclease structures, an essential metal-binding aspartate is replaced by alanine (indicated by the asterisks). DNA is rendered as a stick model with gray carbons; the scissile phosphodiester is indicated by a yellow phosphorus atom. The 5' ssDNA end enters the AdnA active site and the 3' ssDNA end enters the AddA active site.

References for SI citations

1. Sinha KM, Unciuleac MC, Glickman MS, Shuman S. (2009) AdnAB: a new DSB-resecting motor-nuclease from Mycobacteria. *Genes Development* 23: 1423-1437.
2. Zheng SQ, et al. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Meth* 14: 331-332.
3. Rohou A, Grigorieff N. (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192(2):216-221.
4. Scheres SH. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180: 519-530.
5. Kucukelbir A, Sigworth FJ, Tagare HD. (2014) Quantifying the local resolution of cryo-EM density maps. *Nature Meth* 11: 63-65.
6. Emsley P, Lohkamp B, Scott WG, Cowtan K. (2010) Features and development of Coot. *Acta Crystallogr* D66: 486-501.
7. Pettersen EF, et al. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605-1612.
8. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr* D66: 213-221.
9. Xue C, Wang W, Crickard B, Moevus C, Kwon Y, Sung P, Greene E. (2019) Regulatory control of Sgs1 and Dna2 during eukaryotic DNA end resection. *Proc Natl Acad Sci USA* 116: 6091-6100.
10. Lee JY, Yang W. (2006) UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* 127: 1349-1360.
11. Krajewski WW, Fu X, Wilkinson M, Cronin NB, Dillingham MS, Wigley DB. (2014) Structural basis for translocation by AddAB helicase-nuclease and its arrest at χ sites. *Nature* 508: 416-419.