Science
AAAS

Supplementary Materials for

Structures and operating principles of the replisome

Yang Gao, Yanxiang Cui, Tara Fox, Shiqiang Lin, Huaibin Wang, Natalia de Val, Z. Hong Zhou and Wei Yang *
*Correspondence to: wei.yang@nih.gov

This PDF file includes:

Materials and Methods

Purification of gp4 and gp5 proteins

DNA encoding E343Q gp4 was synthesized and subcloned into a modified pET28a vector, with histidine tags and PreScission cleavage sites on its N-terminus (pWY2909). *E. coli* HMS174 (DE3) cells transfected with pWY2909 were grown in LB medium to OD_{600} of 0.5 at 37 °C. Protein expression was then induced by addition of 0.5 mM IPTG and incubation at 20 °C overnight. Cells were collected by centrifugation at 4000 rmp for 30 minutes and lysed by sonication. The clarified supernatants were loaded onto 5 ml HisTrap Fastflow column (GE Healthcare). The column was washed with 300 ml of buffer containing 20 mM Tris (pH 7.5), 1 M NaCl, 5 mM β-mercaptoethanol (BME), 0.1 mM PMSF and 50 mM imidazole, and gp4 protein was eluted with the same buffer containing 300 mM imidazole. The gp4 was then mixed with PreScission protease at 50:1 mass ratio at 4°C for 3 hours. After dilution 20-fold with a low-salt buffer containing 20 mM Tris (pH 7.5), 50 mM KCl, and 3 mM DTT, the cleaved gp4 was loaded onto a MonoQ 10/100 GL column (GE Healthcare) equilibrated with the same low-salt buffer and eluted by a gradient of increasing KCl to 500 mM. Fractions containing gp4 were pooled, concentrated and buffer exchanged to storage buffer containing 20 mM Tris (pH 7.5), 300 mM KCl, 3 mM DTT, and 50% glycerol, and stored at -80 °C.

DNAs encoding D5A/E7A (exo^-) gp5 in pET15b (pWY2911) and Trx in pRSFDuet vector (pWY2912) were co-transfected into *E. coli* BL21 (DE3) cells. The transfected cells were first grown to OD_{600} of 0.5 at 37 °C and then induced by addition of 0.5 mM IPTG. The gp5-Trx complex was purified by sequentially passing through a Heparin HP and MonoQ 10/100 GL column (both from GE Healthcare). For both columns, a low salt buffer containing 20 mM Tris (pH 7.5), 50 mM KCl, 3 mM DTT, 0.1 mM EDTA and 0.1 mM PMSF was used for sample loading and column equilibration, and a high salt buffer with 20 mM Tris (pH 7.5), 1000 mM KCl, 3 mM

DTT, 0.1 mM EDTA and 0.1 mM PMSF was used for gradient elution. The gp5-Trx was concentrated and loaded onto a Superdex200 column (GE Healthcare), and the eluted peak fractions were stored at -80 °C in 20 mM Tris (pH 7.5), 200 mM KCl, 3 mM DTT, and 50% glycerol.

Assemble of gp4-DNA complexes

Gp4 was mixed with DNA (5′-TGG TCT TTT TTT TTT TTT TTT TTT T-3′) at 1:1.5 molar ratio in a buffer containing 50 mM Tris (pH 7.5), 150 mM KCl, 3 mM DTT and 0.1 mM dTTP. Following incubation on ice for 10 minutes, the gp4-DNA complex was loaded onto a Superdex200 column (GE Healthcare) equilibrated in the same buffer. The eluted gp4-DNA complex was concentrated to 1 μM (~0.4 mg/ml), as determined by Bradford assay. For primer synthesis, $MgCl_2$, ATP, and 3′-dCTP were added to the gp4-DNA complex to final concentrations of 2 mM, 0.1 mM and 0.2 mM respectively, and the sample was reacted at room temperature for 30 minutes. After addition of 0.5 mM CTP, the sample was diluted to 0.2 mg/ml for cryo-EM grid preparation. The gp4-DNA/$pppAC_d$-CTP-dTTP sample was kept at room temperature, as the sample tends to precipitate at lower temperature. 4 μl of gp4-DNA sample was deposited onto a freshly glow-discharged Quantifoil R1.2/1.3 300 mesh grid and blotted using a Vitrobot (FEI) with the standard Vitrobot filter paper, Ø55/20mm (Ted Pella). The blotting time was set to 4 s, blotting force was set to 4 and the blotting was done under 100% humidity at 20°C. The grids were flash-frozen in liquid ethane and stored in liquid nitrogen.

Primase activity and gp4-DNA/$pppAC_d$-CTP assembly assay

For primase activity analysis, 1 μM gp4-DNA-dTTP complex, 2 mM $MgCl_2$, 0.1 mM ^{32}P-ATP, and 0.2 mM 3′-dCTP (or CTP) were mixed and incubated at room temperature for 30 min. The reaction was terminated by adding an equal volume of 80% formamide and 20 mM EDTA.

The products were resolved on a 30% acrylamide electrophoresis gel and analyzed using a Typhoon phosphor-imager.

To determine whether the short RNA primer was bound to the gp4-DNA complex, 50 µl of gp4-DNA complex samples with ATP, CTP, $pppAC_d$ (3′-deoxy), or $pppAC_d$+CTP (either ATP or CTP was ^{32}P-labeled) were loaded onto a home-made size-exclusion column of NAP-25 beads (GE Healthcare). The eluate was collected in 100-µl fractions. 1 µl of each eluted fraction was dot-blotted onto Whatman™ filter paper and the ^{32}P was quantitated using a Typhoon phosphor-imager. The elution profile of gp4 was determined by visual inspection of Bradford reactions.

Assemble of gp4-gp5 on the DNA-DNA/RNA hybrid

Gp4, gp5, DNA (5′-TTT TTA GCT GGT CAT TTT TTT TTT TTT TTT TTT TTT TTT TT-3′), and RNA primer (5′-$_rA_rC_rC_rA_rGC_{dd}$-3′) were mixed at 1:1:1.2:2 molar ratio in a buffer containing 50 mM Tris (pH 7.5), 150 mM KCl, 3 mM DTT, 0.1 mM dTTP and 2 mM $MgCl_2$. The sample was incubated at room temperature for 10 minutes and then concentrated in a 0.5 ml Amicon centrifugal filter concentrator (Millipore) to 0.4 mg/ml (determined by Bradford assay). Cryo-EM grids were prepared using the same procedure as for the gp4-DNA complexes.

Assemble of replisome and gp4-gp5 leading-strand complexes

DNA template (5′-TTT GGT CAT TTT TTT TTT TTT TTT TTT TTC GGA GTC GTT TCG ACT CCG ATT ATC ACG CTA TGT CGT CAA GTT GTA CC-3′) and primer (5′-GGT ACA ACT TGA CGA CAT AGC GTG-3′) were annealed by first incubating at 95 °C for 5 minutes and then slowly cooling to 4 °C in a PCR machine. The gp5 and DNA were mixed at 1:1.1 molar ratio in a buffer containing 50 mM Tris (pH 7.5), 150 mM KCl, and 3 mM DTT. To produce the 3′-dideoxy primer end, 2 mM $MgCl_2$ and 10 µM dideoxy-ATP were added and incubated at room temperature for 10 minutes. Following addition of gp4 (equal molarity to gp5) and 0.1 mM dTTP, 0.1 mM ATP and 0.2 mM 3′-dCTP were added and incubated at room temperature for 30 minutes.

Finally, 0.5 mM CTP was added. The sample were concentrated to 0.5 mg/ml and cryo-EM grids were prepared as described above.

Cryo-EM data collection

2722 micrographs of gp4-DNA complexes were collected on a Titan Krios electron microscope operated at 300 kV (UCLA) using the super-resolution mode with a nominal magnification of 130K (calibrated pixel size of 1.07Å on the sample level, corresponding to 0.535Å in super-resolution mode). Movies were recorded with a K2 Summit camera, with the dose rate at the detector set to 6.0 $e^{-}s^{-}Å^{-2}$. The total exposure time for each video was 12 s, which was fractionated into 60 frames of sub-images. The defocus values ranged between 0.7 and 2.5 μm.

1731 micrographs of gp4-gp5 complexes on lagging-strand DNA and 5250 micrographs of gp4-gp5 complexes on the fork were collected on a Titan Krios electron microscope operated at 300 kV (Frederick National Laboratory, Frederick) using the counting mode with a nominal magnification of 175K (calibrated pixel size of 0.86Å). The movies were recorded with K2 Summit camera and dose rate of 2.63 $e^{-}s^{-}Å^{-2}$. The total exposure was 40 $e^{-}Å^{-2}$ over 15.2 s, fractionated into 38 frames. The defocus values ranged between 0.6 and 2.8 μm.

Image processing

MotionCor2 (*57*) was used for drift-correction and electron-dose-weighting for all movies. The defocus values were estimated on non-dose-weighted micrographs with Gctf (*58*). The dose-weighted micrographs were phase-flipped using Bsoft (*59*). Around 5000 particles were manually picked for each dataset and 2D-classification was performed with RELION (version 2.1) (*60*) to generate a template for autopicking with Gautomatch (developed by Dr. K. Zhang; https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/).

For gp4-DNA complexes, 816251 particles were picked from 2616 manually screened micrographs. After 2D classification with RELION (*60*), 613222 particles were used to generate the initial model using cryoSPARC (*61*). The particles were separated into six classes by 3D-classification in RELION (*60*) and two classes with high-resolution features were selected. 3D-refinement with a soft mask covering the entire volume of gp4 yielded a map of 3.2 Å resolution ($gp4_5$-DNA), in which five gp4 subunits were well resolved. Further 3D-classification with finer angular sampling intervals (3.7 degrees) in RELION (*60*) produced three additional maps with all six helicase subunits traceable (gp4-DNA I-III).

For gp4-gp5 complexes on the DNA/RNA hybrid and lagging strand DNA, 363927 particles were extracted from 1694 manually screened micrographs. After 2D classification, 323690 particles were supplied to cryoSPARC (*61*) for 3D-classification and heterogeneous refinement. Two classes out of four showed clear features of gp4 helicase and gp5 polymerase and were further 3D-classified in RELION (*60*). Particles in cryoSPARC class1 were separated to two classes and the dominant class was refined to 4.0 Å (LagS1). Particles in cryoSPARC class3 were separated into four classes, three of which were subsequently refined to generate LagS2-4. The LagS1 was locally refined with a soft mask covering only the gp5-DNA and primase parts, and a 3.7 Å map was generated with improved local densities.

For the leading-strand complexes, 976106 particles were selected from 5096 micrographs after reducing the resolution to 5.16 Å by 6-fold binning of pixels. 684218 particles were retained after eliminating bad classes identified in 2D-classification. Further 2D classification to 200 classes in cryoSPARC revealed the existence of gp4-gp5 complexes with different stoichiometries and configurations. Among them, 4 classes (38615 particles) showed features of a replisome with gp5 polymerases on both sides of gp4. They were combined and further 2D classified into 50 classes, 14 of which (28695 particles) are shown in Extended Data Fig. 6b. All 684218 particles

from RELION 2D classification were used for 3D-classification in cryoSPARC (*61*). The gp4-DNA and gp5-DNA-alone complexes were discarded and only complexes with both gp4 and gp5 were used for further classification in RELION (*60*). Two classes with features of leading-strand complexes were combined and divided into fifteen classes. Five of them, containing complete gp4 and gp5, were re-extracted with pixel size of 1.72 Å and refined (Lead1-5). Particles from classes containing complete gp4 were combined, re-extracted and refined with a soft-mask covering only gp4, yielding a Lead-gp4-DNA structure at 3.8 Å. Similarly, Lead-gp5-DNA structure was determined at 4.5 Å. Another 3D class produced by cryoSPARC showed features similar to the gp4-gp5 complex on lagging-strand DNA, with gp5 polymerase attached to adjacent primase domains. The particles were separated into eight classes, and four of them were refined to generate LagL1-4 structures.

All reported resolutions are based on the "gold standard" refinement procedures and the 0.143 Fourier Shell Correlation (FSC) criterion. Local resolution was estimated using Resmap (*62*).

Model building and refinement

Gp4 helicase, primase and gp5 polymerase structures from PDB 1E0J (*19*), 1NUI (*20*), and 1T8E (*63*) were used as templates for model building. Standard ssDNA or double stranded DNA (generated in COOT (*64*)) were used as starting models. Monomeric helicase, primase, polymerase and DNA were manually docked into the cryo-EM density maps in Chimera (*65*). The models were first manually adjusted in COOT (*64*) and then refined in Phenix (*66*) with real-space refinement and secondary structure and geometry restraints. For the lower resolution LagS4, Lead1-4, and LagL1-4, the refined gp4-DNA complex and gp5-DNA complex were docked into the cryo-EM density with Chimera (*65*) and the connections were adjusted in COOT (*64*).

Statistics of all cryoEM data collection and structure refinement are shown in Table S1.

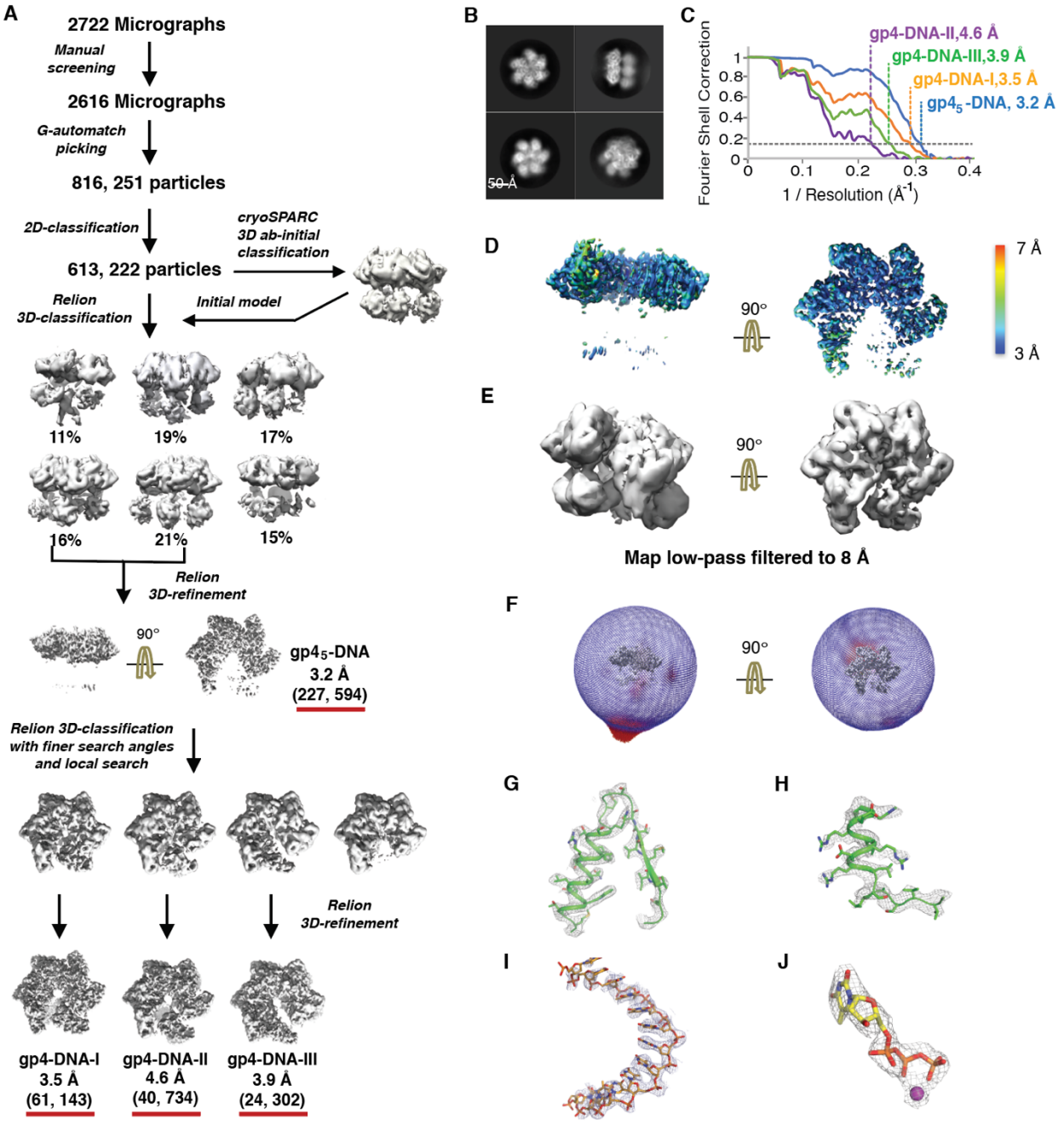

Fig. S1. Cryo-EM data processing of gp4-DNA complexes. (**A**) Workflow of cryo-EM data processing. The numbers of particles used in final refinement are shown in parentheses. (**B**) Representative 2D-classification images of gp4-DNA complexes. (**C**) FSC curves of the final refined gp4-DNA structures. (**D**)Two views of the local resolution of the density for helicase domains in gp4$_5$-DNA structure. (**E**) Density map of gp4$_5$-DNA structure after low-pass-filtering to 8 Å. (**F**) Two views of angular distribution of particles used for gp4$_5$-DNA structure reconstruction. (**G** to **J**) Local densities of representative regions for HelC (G), helicase-primase linker in HelC (H), DNA (I) and dTTP at HelBC interface (J).

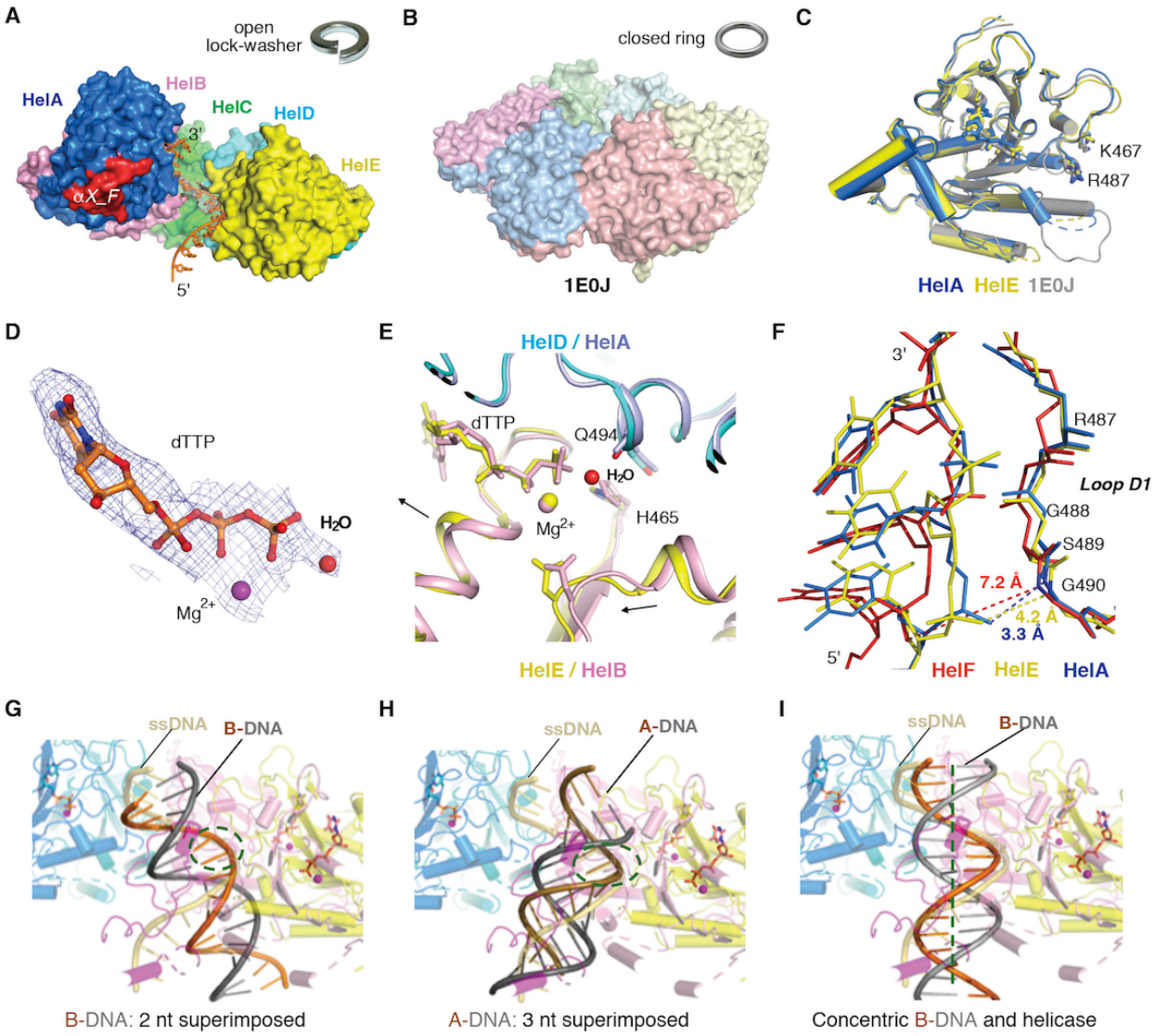

Fig. S2. Structural analysis of gp4-DNA complex. (**A**) When bound to ssDNA, gp4 helicase (in the same color scheme as in Fig. 1E) assumes a lock-washer-like conformation. For clarity, HelF (except for the linker) is omitted. (**B**) The crystal structure of apo gp4 (1E0J (*19*)) is a closed ring. (**C**) Superposition of HelA (blue, residues 285-540), HelE (yellow) and a subunit from the crystal structure (grey). TRMSD between HelA and HelE is 0.85 Å, and between HelA and apo-gp4 is 0.78 Å. DNA-binding residues R487 and K467 are shown in sticks. (**D**) Density map of dTTP-Mg^{2+}-H_2O at the D-E interface. (**E**) Comparison of the NTPase active sites at HelAB (light blue and pink) and HelDE (cyan and yellow) interfaces by superimposing HelA and HelD. Conformational changes from HelB to HelE are indicated by black arrows. (**F**) Comparison of DNA-binding LoopD1 in HelA (blue), HelE (yellow) and HelF (red) after superimposing the helicase domains. The distances between G490 and DNA backbones are indicated (HelF>>HelE>HelA). (**G**) B-DNA can be superimposed with 2 nt of the ssDNA complexed with gp4, resulting in a 30° angle to the central channel and bad clashes with the helicase. (**H**) A-DNA can be superimposed with 3 nt of the ssDNA, resulting in a tile and many clashes. (**I**) B-DNA can be best fit into gp4 helicase concentrically with the central channel.

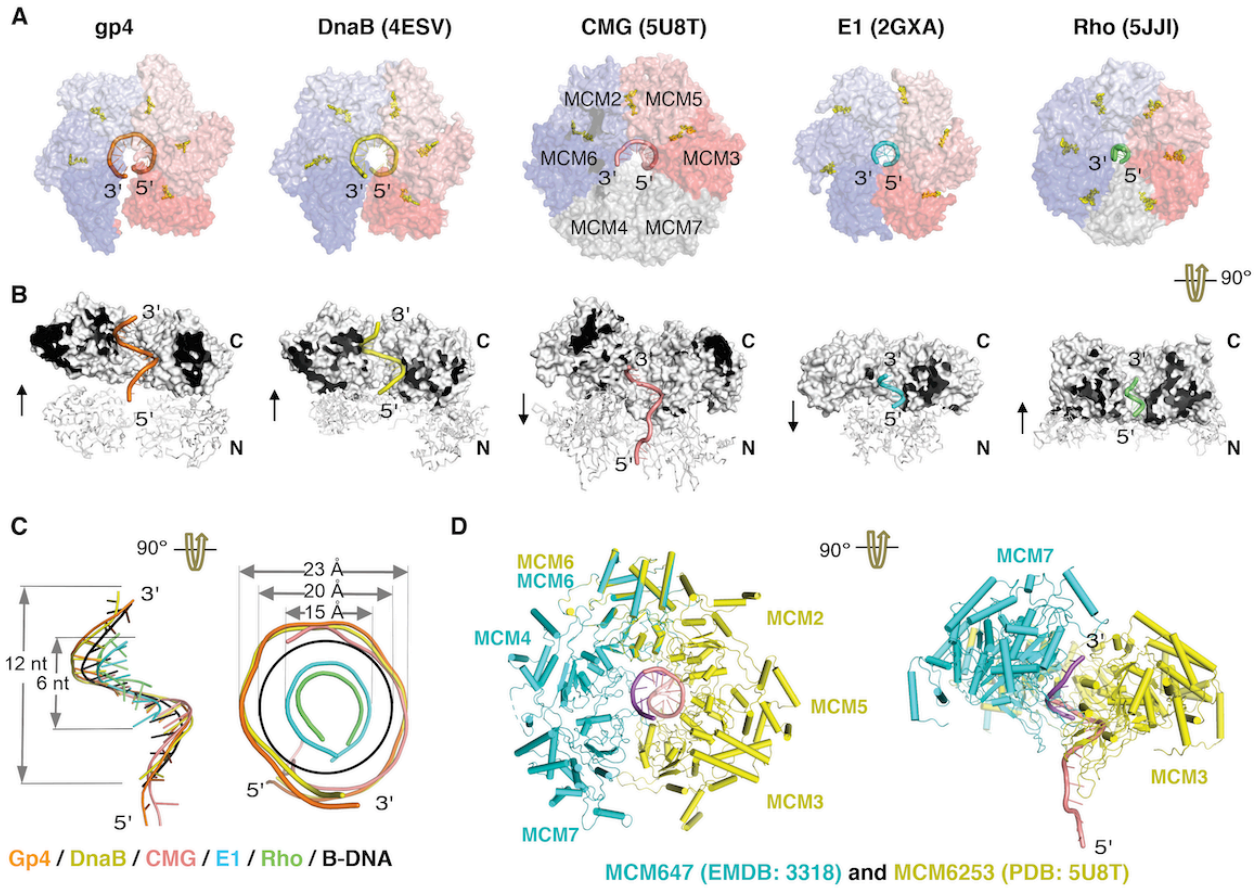

Fig. S3. Structural comparison of hexameric helicase-substrate complexes. (**A**) Structures of gp4, DnaB (4ESV (*14*)), CMG (5U8T (*16*)), E1 (2GXA (*11*)) and Rho (5JJI (*67*)) helicases viewed from the 3′-end of the DNA or RNA. Protein, DNA/RNA, and NTP are shown as surface, cartoon, and sticks, respectively. The subunits are colored blue to red following their positions at the 3′- to 5′-end of DNA/RNA. Subunits not contacting DNA/RNA are colored grey. (**B**) An orthogonal view with cut-through to expose the central channel of each helicase. The C-terminal helicase domains (upper half) are shown as surface and the N-terminal domains (lower half) as cartoons. The direction of translocation for each helicase is indicated by an arrow on the left side. (**C**) Superposition of DNAs and RNA in these helicase complexes with B-DNA (black line). (**D**) Overlay of two CMG-DNA complexes. In one (yellow, PDB: 5U8T (*16*)) 8 nucleotides contact helicase domain of MCM6, 2, 5 and 3, and in the other (cyan, EMDB: 3318 (*15*)) 6 nucleotides are held by MCM6, 4, 7. When MCM6 subunits in the two structures are superimposed, DNA forms a complete turn within CMG just as with gp4 and DnaB, and CMG adopts a lock-washer-like structure.

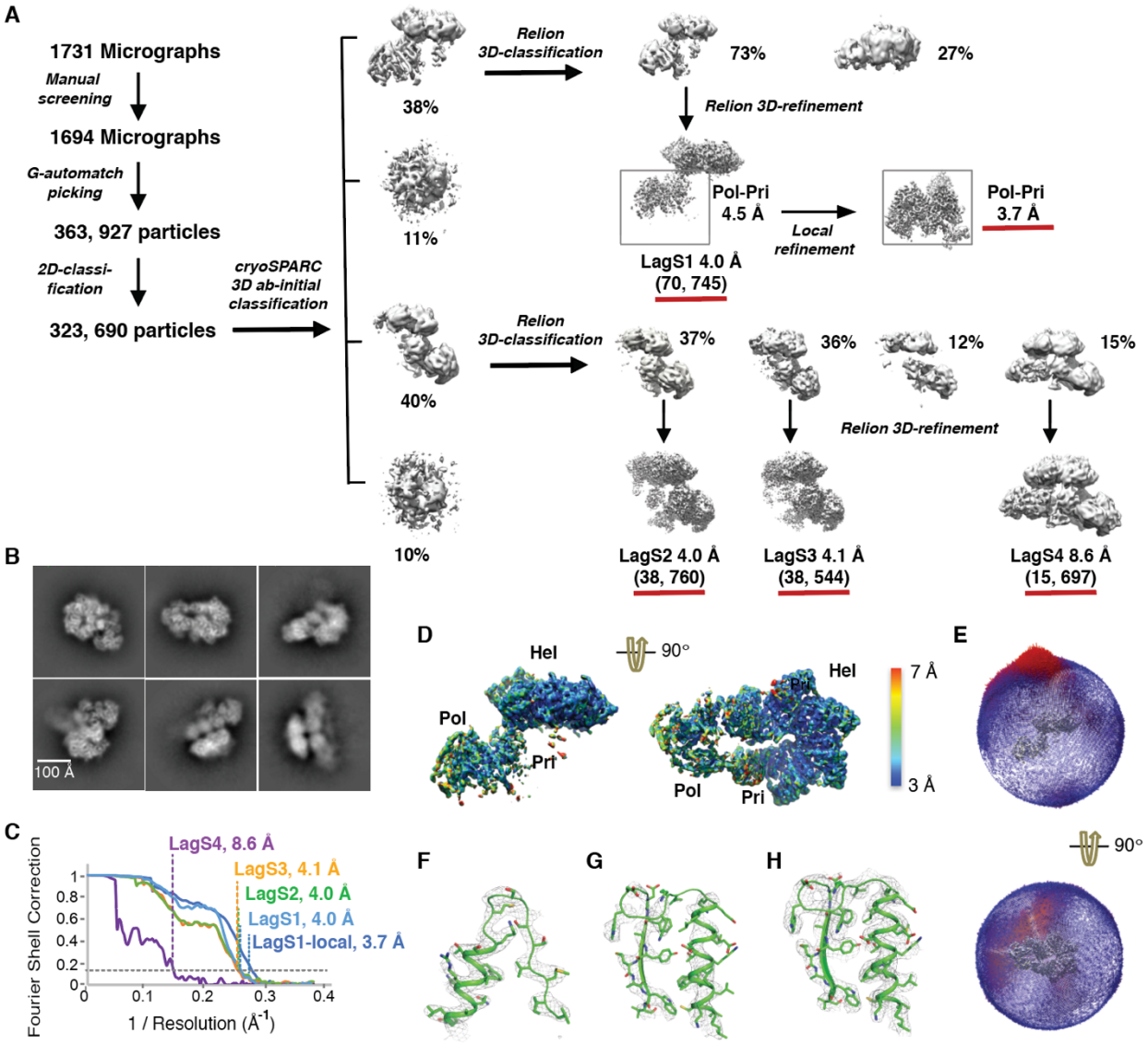

Fig. S4. Cryo-EM data processing of the lagging-strand gp4-gp5 complexes. (**A**) Workflow of cryo-EM data processing. The numbers of particles used in final refinement are shown in parentheses. (**B**) Representative 2D-classification images. (**C**) FSC curves of the refined structures. (**D**) Two views of the local resolution of the density for LagS1 structure. (**E**) Two views of angular distribution of particles used for LagS1 structure reconstruction. (**F** to **H**) Densities of representative regions in LagS1, HelF (F), gp5 before (G) and after local refinement (H).

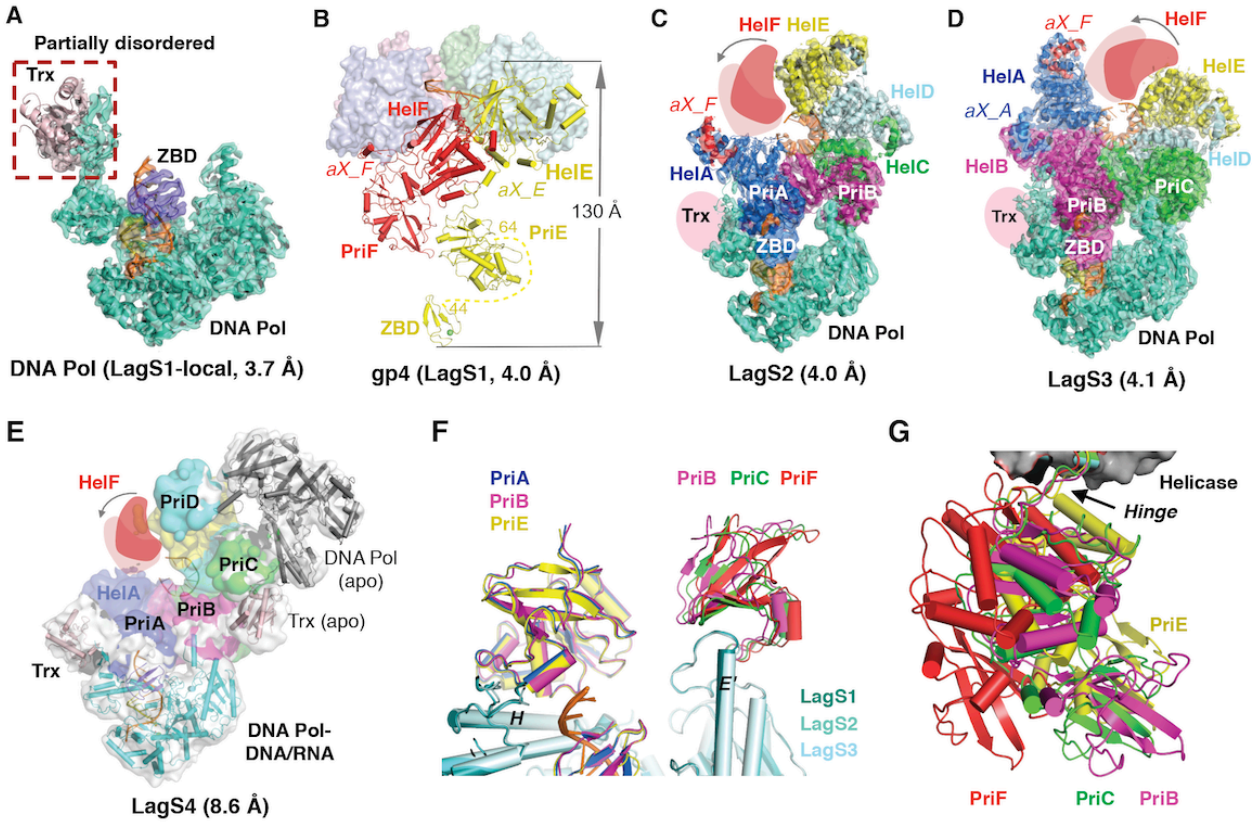

Fig. S5. Structure and flexibility of the lagging-strand gp4-gp5 complexes. (**A**) Structure of gp5 polymerase-Trx-DNA/RNA-dTTP fitted into the cryo-EM densities of LagS1. Trx (boxed in red) is partially disordered and the density is incomplete because of the lack of contact with the short DNA/RNA hybrid. (**B**) Complete gp4-E and gp4-F subunits in LagS1. (**C** and **D**) Structures of LagS2 (B) and LagS3 (C) complexes. In both complexes, HelF is disordered. (**E**) LagS4 complexes with two gp5 polymerases bound to gp4, one containing the DNA/RNA substrate (green) and the other without any substrate (grey). (**F**) Comparison of polymerase-primase interactions in LagS1-3. With polymerases superimposed, the primase domains that interface with gp5's Helix H are immobile, but at the other interface with gp5's Helix E′ the primase domains are flexible. (**G**) The hinge region between primase and helicase (indicated by an arrow) is very flexible. Locations of primase domains are shown after superimposing their contacting helicase domains.

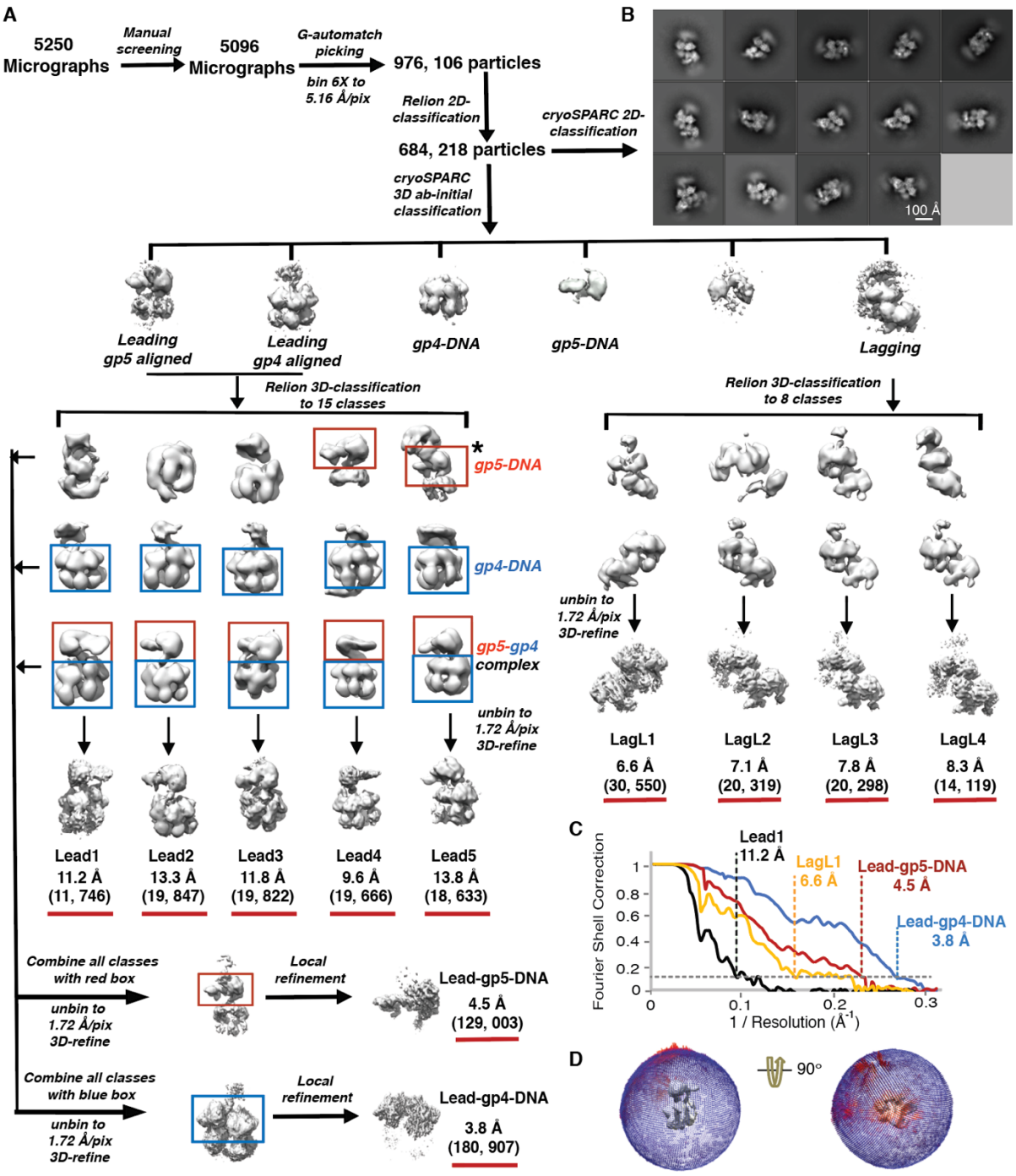

Fig. S6. Cryo-EM data processing of gp4 and gp5 bound to a DNA fork. (**A**) Workflow of cryo-EM data processing. gp4 and gp5 in the leading-strand complexes are boxed blue and red, respectively. The numbers of particles used in final refinement are shown in parentheses. The asterisk (*) denotes that the 3D class of an intact gp5 also contains parts of two gp4 bound to gp5. Residual densities corresponding to a leading-strand gp5 polymerase are also visible in the 3D classes of LagL3 and LagL4. (**B**) 2D classes of the replisome, which consists of one gp4 and two gp5. (**C**) FSC curves of the final refined gp4-gp5-DNA complexes. (**D**) Two views of angular distribution of particles used for Lead1 structure reconstruction.

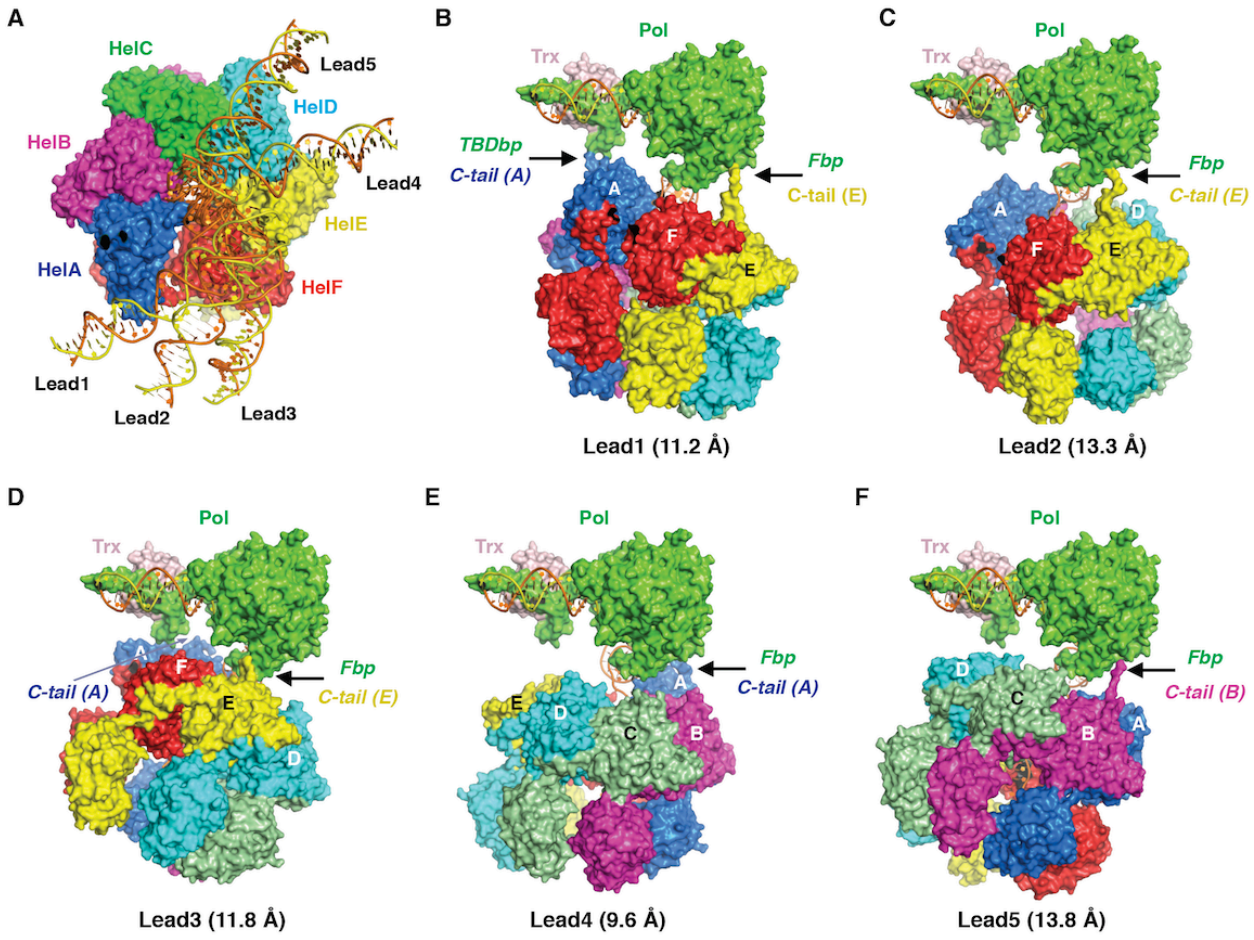

Fig. S7. Varied helicase-polymerase interface in the leading-strand complexes. (**A**) When Lead1-5 structures were aligned by their gp4 helicase domains, DNA (yellow primer and orange template) bound to gp5 polymerase points in five different directions, differing by up to 180°. The helicase is shown as colored molecular surface; polymerase (Pol) and Trx are omitted for clarity. (**B** to **F**) Varied gp4-gp5 interface in Lead1-5. Polymerases from Lead1-5 complexes are aligned and shown in the same orientation. In all five Lead complex structures, one C-tail (HelE, HelA or HelB) contacts the basic patches near the separation pin at the DNA fork, which is known as the front basic patch, or Fbp (*45*). In Lead1 structure, a second C-tail from HelA interacts with the basic patch near the Trx-binding domain (known as TBDbp (*44*)). These interactions are indicated by black arrows.

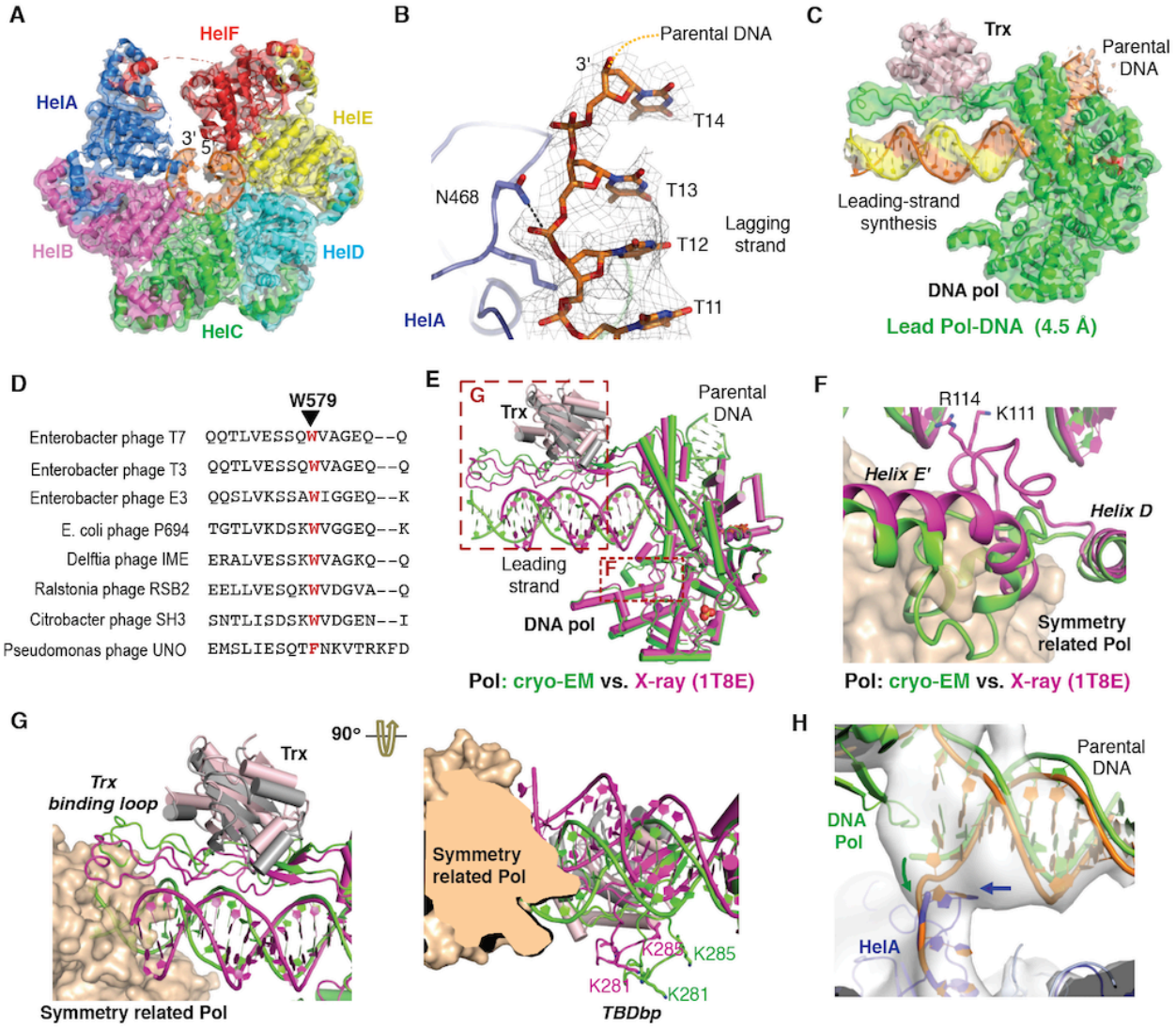

Fig. S8. Structure of the leading-strand gp4-gp5 complexes. (**A**) All six helicase domains of gp4-DNA in the leading-strand complex are ordered (colored as in Fig. 1E). (**B**)The cryo-EM densities define two additional nucleotides (T13 and T14) beyond the 3′ end that interact with the gp4 helicase. (**C**) Structure of gp5-Trx-DNA fitted into cryo-EM densities. T7 polymerase, Trx, and DNA are colored green, light pink, and orange/yellow, respectively. (**D**) Sequence alignment of the separation pin in DNA polymerases from different bacteriophages reveals the conservation of W579. (**E**) Superposition of the cryo-EM (green) and the crystal structure of gp5-Trx-DNA complexes (*63*) (magenta). The most different parts are boxed in red dashes. (**F**) LoopDE′ of Exo domain (gp5) in the leading-strand complex assumes a different conformation and is detached from DNA. The symmetry-related molecule in the crystal structure (brown surface) may stabilize the LoopDE′-DNA interface. (**G**) Two views of the DNA distal end. Sidechains in the TBDbp, which likely interact with the gp4 C-tail, are shown as sticks. (**H**) Docking of the refined Lead-gp4-DNA (blue, 3.8 Å) and Lead-gp5-DNA (green, 4.5 Å) into Lead1 density (grey). The connections in the DNA fork in Lead1 (orange) was slightly adjusted after rigid-body docking.

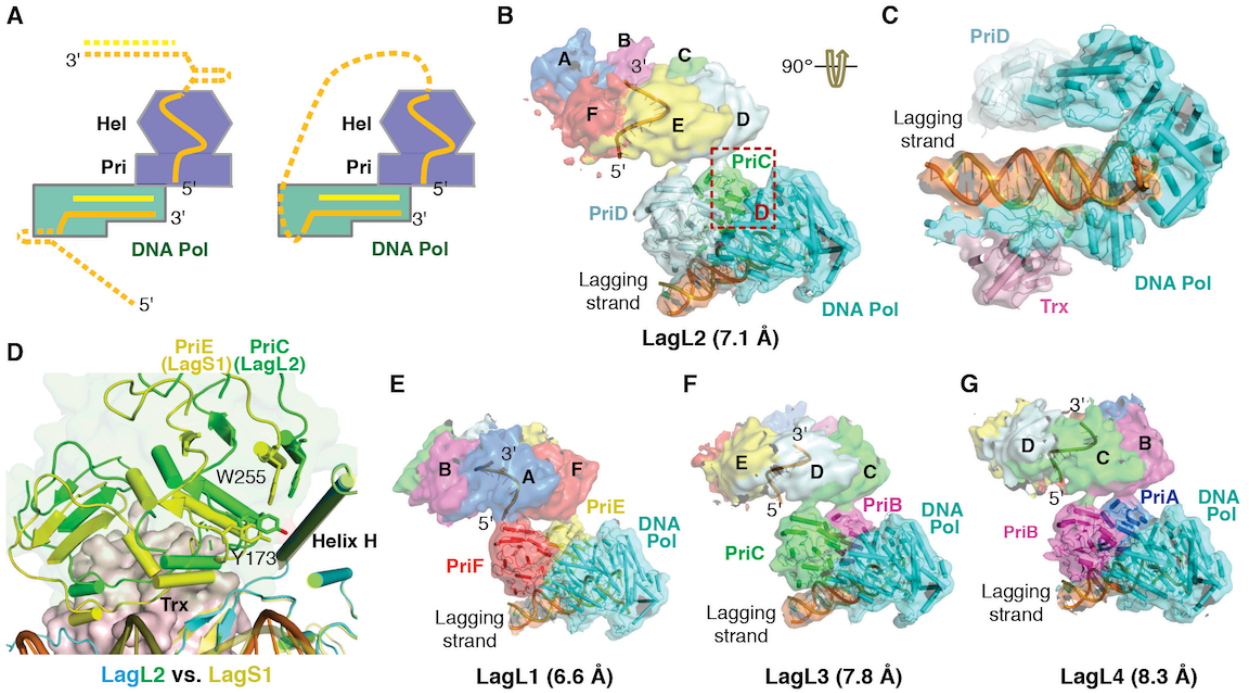

Fig. S9. Complexes of gp4-gp5 on a DNA fork. (A) Diagrams of two alternative ways of forming LagL complexes. gp4 and gp5 may each bind to a different DNA fork substrate (left), or alternatively the leading strand DNA folds back to contact the primase domains of gp4 (right). (**B**) Overall structure of LagL2 (Fig. S6). gp4 subunits are colored as in Fig. 1E. The gp5-Trx-DNA complex of Lead-gp5 (4.5 Å) and two primase domains from LagS1 (4.0 Å) were docked into the cryo-EM density map and shown as cartoons. (**C**) gp4 helicase was omitted and a zoom-in view of the gp5-Trx-DNA and PriD domain is shown. (**D**) The interface between the primase domain with Helix H of gp5 is strengthened in LagL2 comparing to that of LagS1. The gp5 polymerase from LagL2 and LagS1 are superimposed. PriC from LagL2 is shown as green cartoon and transparent surface while the PriE from LagS1 is shown as yellow cartoon. (**E** to **G**) Structures of LagL1, LagL3, and LagL4 with gp5 shown in the same view as LagL2 in panel B.

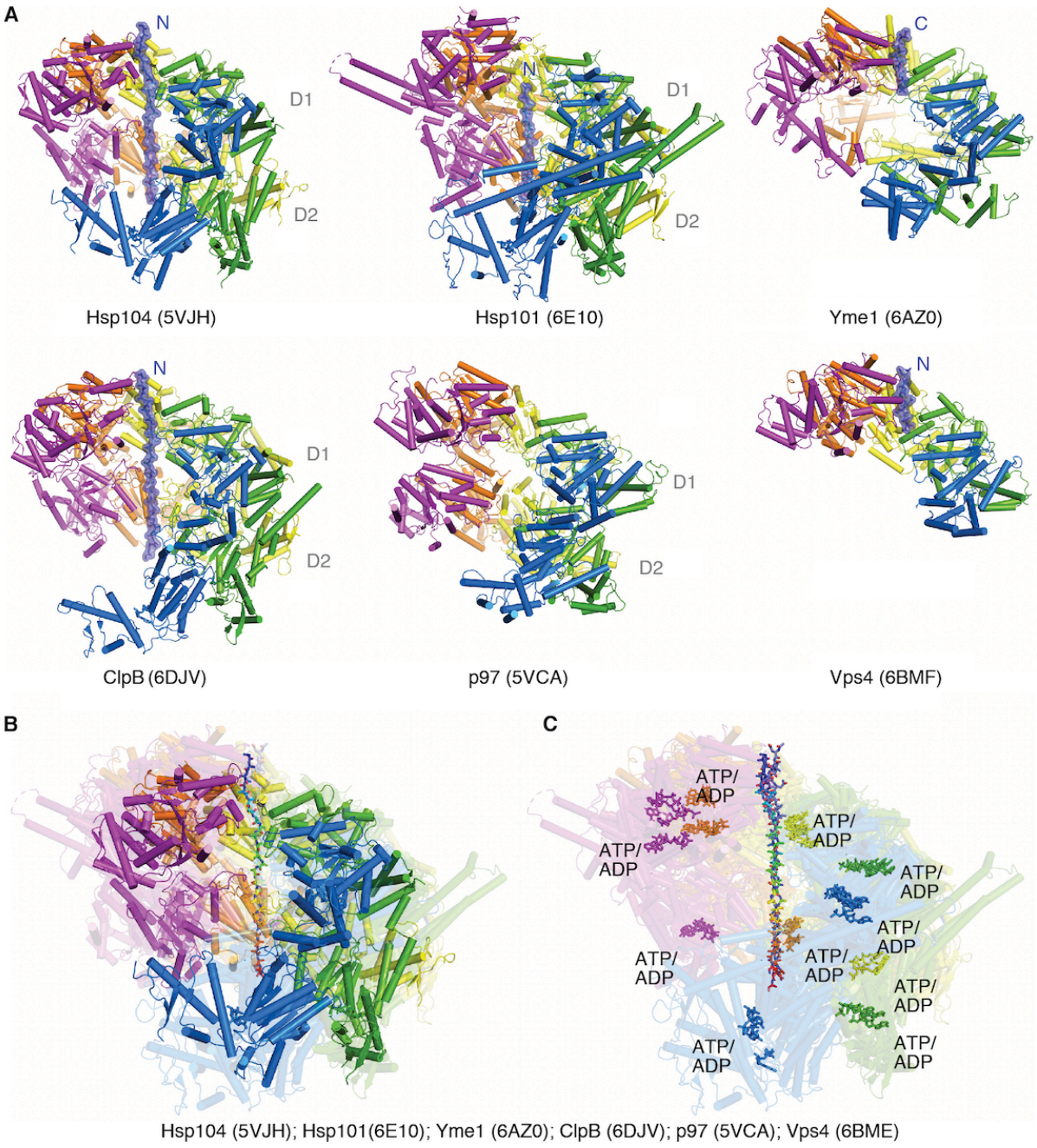

Fig. S10. Structural comparison of AAA+ protein translocase-substrate complexes. (**A**) AAA+ ATPases, Hsp104 (5VJH, (52)), Hsp101 (6E10 (51)), Yme1 (6AZ0 (56)), ClpB (6DJV (54)), p97 (5VCA (53)) and Vps4 (6BMF (*30*), all have a right-handed spiral structure(lock-washer like). Five subunits are shown in magenta, orange, yellow, green and blue cartoons, and the 6th subunit is omitted to reveal the peptide substrate (covered with blue-purple surface, N- and C-terminus labeled) in the central channel. ATP, ADP or ATP γS at the subunit interfaces are shown as sticks. (**B**) Superposition of the six AAA+ ATPases confirms the conserved lock-washer shape. Hsp104 is shown in solid colors and others are semi-transparent. **(C)** The ATPases are shown in semi-transparent cartoons, and ATP/ADP and ATP γS are shown as sticks. The peptide substrates share a conserved binding mode with two amino acids contacting each ATPase subunit.

Table S1. Cryo-EM data collection, refinement and validation statistics

	Gp4$_{5}$-DNA	Gp4-DNA-I	Gp4-DNA-II	Gp4-DNA-III
PDB Code	6N7I	6N7N	6N7S	6N7T
EMDB Code	EMD-0357	EMD-0359	EMD-0362	EMD-0363
Data collection and processing				
Magnification	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300
Electron exposure (e/Å^2)	72	72	72	72
Defocus range (μm)	0.7-2.5	0.7-2.5	0.7-2.5	0.7-2.5
Pixel size (Å)	1.075	1.075	1.075	1.075
Number of micrographs				
Collected	2722	2722	2722	2722
Used	2616	2616	2616	2616
Symmetry imposed	C1	C1	C1	C1
No. particles used	227594	61143	40734	24302
Map resolution (Å)				
Corrected, FSC=0.143	3.2	3.5	4.6	3.9
Masked, FSC=0.143	3.2	3.4	4.1	3.8
Masked, FSC=0.5	3.7	3.9	6.7	6.7
Unmasked, FSC=0.143	3.6	4.0	6.8	6.8
Unmasked, FSC=0.5	4.4	7.2	9.8	9.3
Refinement				
Initial model (PDB code)	1E0J	1E0J	1E0J	1E0J
Map sharpening B (Å^2)	-110	-95	-120	-91
Cross-correlation				
CC_mask	0.73	0.67	0.61	0.74
CC_volumes	0.71	0.66	0.62	0.74
CC_peaks	0.57	0.53	0.47	0.63
Model composition				
Non-hydrogen atoms	10842	12667	12844	12859
Protein / DNA	10538 / 300	12362 / 300	12540 / 300	12554 / 300
Ligands / Mg^{2+}	116 / 4	145 / 5	116 / 4	145 / 5
B-factors				
Non-hydrogen atoms	49.2	40.4	39.1	57.4
Protein / DNA	48.8 / 64.8	40.1 / 52.4	38.4 / 69.3	57.2 / 65.9
Ligands / Mg^{2+}	31.2 / 33.2	28.6 / 29.2	16.4 / 17.0	38.1 / 37.6
R.m.s deviations				
Bond lengths (Å)	0.006	0.011	0.020	0.017
Bond angles (°)	1.16	0.90	1.11	1.05
Ramachandran				
Favored (%)	96.3	94.9	88.0	91.1
Allowed (%)	3.7	5.1	12.0	8.9
Outlier (%)	0.1	0	0.07	0
Validation				
MolProbity score	1.37	1.59	2.09	1.90
Clashscore	3.28	4.66	9.1	5.57
Poor rotamers (%)	0.79	0.98	0.52	1.26
EMRinger	2.33	1.88	1.16	1.25

	LagS1 (Pol-Pri-Hel)	LagS1-gp5-DNA (Pol-Pri)	LagS2	LagS3	LagS4
PDB Code	6N9V	6N9U	6N9W	6N9X	
EMDB Code	EMD-0380	EMD-0379	EMD-0381	EMD-0382	EMD-0386
Data collection and processing					
Magnification	130,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e/Å^2)	40	40	40	40	40
Defocus range (μm)	0.6-2.8	0.6-2.8	0.6-2.8	0.6-2.8	0.6-2.8
Pixel size (Å)	0.86	0.86	0.86	0.86	0.86
Number of micrographs					
Collected	1731	1731	1731	1731	1731
Used	1694	1694	1694	1694	1694
Symmetry imposed	C1	C1	C1	C1	C1
No. particles used	70745	70745	38760	38544	15697
Map resolution (Å)					
Corrected, FSC=0.143	4.0	3.7	4.0	4.1	8.6
Masked, FSC=0.143	3.9	3.7	4.0	4.0	7.9
Masked, FSC=0.5	4.5	4.2	4.6	4.6	10.1
Unmasked, FSC=0.143	4.6	4.2	4.7	4.7	13.3
Unmasked, FSC=0.5	8.6	7.9	8.6	8.8	20.0
Refinement					
Initial model (PDB code)	1E0J, 1T8E, 1NUI	1T8E, 1NUI	1E0J, 1T8E, 1NUI	1E0J, 1T8E, 1NUI	
Map sharpening B (Å^2)	-105	-123	-118	-119	
Cross-correlation					
CC_mask	0.80	0.80	0.79	0.77	
CC_volumes	0.77	0.77	0.77	0.75	
CC_peaks	0.62	0.66	0.59	0.57	
Model composition					
Non-hydrogen atoms	21527	8822	19579	19661	
Protein / DNA	20678 / 668	8423 / 368	18820 / 608	18882 / 628	
Ligands / Metal ions	173 / 8	28 / 3	144 / 7	144 / 7	
B-factors					
Non-hydrogen atoms	108.1	46.2	178	70.9	
Protein / DNA	108.1 / 117	45.9 / 51.3	178 / 165	70.9 / 74	
Ligands / Metal ions	73.4 / 101.6	42.9 / 54.0	162 / 164	50.0 / 60.4	
R.m.s deviations					
Bond lengths (Å)	0.008	0.010	0.012	0.010	
Bond angles (°)	0.72	0.80	0.69	0.75	
Ramachandran					
Favored (%)	96.4	95.1	96.2	96.1	
Allowed (%)	3.6	4.9	3.8	3.9	
Outlier (%)	0	0	0	0	
Validation					
MolProbity score	1.27	1.48	1.36	1.39	
Clashscore	2.46	3.54	3.13	3.29	
Poor rotamers (%)	0.18	0.34	0.30	0.35	
EMRinger	2.0	3.1	1.30	1.75	

	Lead-gp4-DNA	Lead-gp5-DNA	Lead1	Lead2	Lead3	Lead4	Lead5
PDB Code	6N7V	6N7W					
EMDB Code	EMD-0364	EMD-0365	EMD-0391	EMD-0392	EMD-0393	EMD-0394	EMD-0395
Data collection and processing							
Magnification	130,000	130,000	130,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300	300	300
Electron exposure (e/Å^2)	40	40	40	40	40	40	40
Defocus range (μm)	0.6-2.8	0.6-2.8	0.6-2.8	0.6-2.8	0.6-2.8	0.6-2.8	0.6-2.8
Pixel size (Å)	0.86	0.86	0.86	0.86	0.86	0.86	0.86
Number of micrographs							
Collected	5250	5250	5250	5250	5250	5250	5250
Used	5096	5096	5096	5096	5096	5096	5096
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1
No. particles used	180,907	129,003	11,746	19,847	19,822	19,666	18,633
Map resolution (Å)							
Corrected, FSC=0.143	3.8	4.5	11.2	13.3	11.8	9.6	13.8
Masked, FSC=0.143	3.8	4.4	10.6	11.2	10.9	9.6	13.3
Masked, FSC=0.5	5.4	8.1	15.3	18.7	14.2	14.7	18.8
Unmasked, FSC=0.143	4.4	7.5	14.7	14.7	14.7	14.2	18.7
Unmasked, FSC=0.5	8.8	10.0	21.7	20.6	20.6	21.7	21.7
Refinement							
Initial model (PDB code)	1E0J	1T8E					
Map sharpening B (Å^2)	-110	-95					
Cross-correlation							
CC_mask	0.73	0.80					
CC_volumes	0.72	0.80					
CC_peaks	0.62	0.74					
Model composition							
Non-hydrogen atoms	12894	7837					
Protein / DNA	12569 / 320	6374 / 1434					
Ligands / Mg^{2+}	145 / 5	28 / 1					
B-factors							
Non-hydrogen atoms	86.6	139.6					
Protein / DNA	86.6 / 105.2	118.3 / 235.2					
Ligands / Mg^{2+}	68.7 / 68.8	89.1 / 90.7					
R.m.s deviations							
Bond lengths (Å)	0.006	0.003					
Bond angles (°)	0.76	0.69					
Ramachandran							
Favored (%)	96.0	95.5					
Allowed (%)	4.0	4.4					
Outlier (%)	0	0.1					
Validation							
MolProbity score	1.39	1.36					
Clashscore	3.22	2.62					
Poor rotamers (%)	0.18	0.15					
EMRinger	1.44	1.92					

	LagL1	LagL2	LagL3	LagL4
EMDB Code	EMD-0387	EMD-0388	EMD-0389	EMD-0390
Data collection and processing				
Magnification	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300
Electron exposure (e/Å^2)	40	40	40	40
Defocus range (μm)	0.6-2.8	0.6-2.8	0.6-2.8	0.6-2.8
Pixel size (Å)	0.86	0.86	0.86	0.86
Number of micrographs				
Collected	5250	5250	5250	5250
Used	5096	5096	5096	5096
Symmetry imposed	C1	C1	C1	C1
No. particles used	30,550	20,319	20,298	14,119
Map resolution (Å)				
Corrected, FSC=0.143	6.6	7.1	7.8	8.3
Masked, FSC=0.143	6.0	7.0	7.7	8.1
Masked, FSC=0.5	9.2	9.6	10.9	13.8
Unmasked, FSC=0.143	8.8	9.2	9.6	11.2
Unmasked, FSC=0.5	14.2	15.3	19.6	20.6