13100–13113 *Nucleic Acids Research, 2022, Vol. 50, No. 22* *Published online 7 December 2022*
https://doi.org/10.1093/nar/gkac1142

Structural transitions during the cooperative assembly of baculovirus single-stranded DNA-binding protein on ssDNA

Jiayi Yin[1,2,†], Yan Fu[1,†], Guibo Rao[1,†], Zhiqiang Li[1,2], Kexing Tian[1,2], Tingting Chong[1,2], Kai Kuang[1,2], Manli Wang[3], Zhihong Hu[3] and Sheng Cao[1,*]

[1]CAS Key Laboratory of Special Pathogens, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan 430071, PR China, [2]University of Chinese Academy of Sciences, Beijing 100049, PR China and [3]State Key Laboratory of Virology, Wuhan Institute of Virology, Center for Biosafety, Mega-Science, Chinese Academy of Sciences, Wuhan 430071, PR China

Received September 13, 2022; Revised October 28, 2022; Editorial Decision November 09, 2022; Accepted November 16, 2022

ABSTRACT

Single-stranded DNA-binding proteins (SSBs) interact with single-stranded DNA (ssDNA) to form filamentous structures with various degrees of cooperativity, as a result of intermolecular interactions between neighboring SSB subunits on ssDNA. However, it is still challenging to perform structural studies on SSB–ssDNA filaments at high resolution using the most studied SSB models, largely due to the intrinsic flexibility of these nucleoprotein complexes. In this study, HaLEF-3, an SSB protein from *Helicoverpa armigera* nucleopolyhedrovirus, was used for *in vitro* assembly of SSB–ssDNA filaments, which were structurally studied at atomic resolution using cryo-electron microscopy. Combined with the crystal structure of ssDNA-free HaLEF-3 octamers, our results revealed that the three-dimensional rearrangement of HaLEF-3 induced by an internal hinge-bending movement is essential for the formation of helical SSB–ssDNA complexes, while the contacting interface between adjacent HaLEF-3 subunits remains basically intact. We proposed a local cooperative SSB–ssDNA binding model, in which, triggered by exposure to oligonucleotides, HaLEF-3 molecules undergo ring-to-helix transition to initiate continuous SSB–SSB interactions along ssDNA. Unique structural features revealed by the assembly of HaLEF-3 on ssDNA suggest that HaLEF-3 may represent a new class of SSB.

INTRODUCTION

Single-stranded DNA-binding proteins (SSBs) play essential roles in DNA replication, repair and recombination (1). SSBs have been found in all three kingdoms of life, and some of them have been extensively structurally and functionally characterized, such as *Escherichia coli* SSB (EcSSB) (2,3) and Replication Protein A (RPA; the primary SSB in eukaryotes) (4). In infected cells, many viruses express their own SSBs. Crystal structures of several viral SSBs have been determined, including gene 32 protein (gp32) from bacteriophage T4 (5), gene 2.5 protein (gp2.5) from bacteriophage T7 (6) and ICP8 from herpes simplex virus type 1 (7). Although these investigated SSBs show divergent structural characters and function in different oligomeric states, they contain at least one oligonucleotide/oligosaccharide-binding (OB) fold to interact with single-stranded DNA (ssDNA). The OB fold consists primarily of a five β-stranded β-barrel capped by an α-helix with a secondary structural arrangement of β1–β2–β3–α–β4–β5. The OB fold is also capable of interacting with other biological molecules such as RNA, oligosaccharides and proteins (8), and the general ligand-binding site on the OB fold is centered on β-strands 2 and 3 (9,10). The most studied EcSSB functions as homotetramers and each EcSSB molecule contains one OB fold (11,12). In contrast, human RPA is a heterotrimeric complex of RPA70, RPA32 and RPA14. RPA70 consists of four OB folds, whereas RPA32 and RPA14 each have one OB fold (4,13).

For non-specific binding of SSB to ssDNA, there are tandem arrays of binding sites on ssDNA, and one SSB molecule covers *n* nucleotides. As the binding density of SSB on ssDNA increases, the remaining vacant regions with fewer than *n* nucleotides in length cannot be used by new SSB molecules. SSBs have been reported to bind to ssDNA with various degrees of cooperativity (14). gp2.5 was found

*To whom correspondence should be addressed. Tel: +86 27 87198286; Fax: +86 27 87198286; Email: caosheng@wh.iov.cn
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

to bind to ssDNA in 'a non-cooperative way' in which gp2.5 molecules independently bind to ssDNA and occlude the binding sites to more SSB molecules (15). In contrast, the term 'cooperative binding' has been used to describe the situation for many other SSBs when binding of one SSB molecule is promoted by the previously bound neighboring protein, resulting in the continuous distribution of SSB molecules on ssDNA without gaps (1,16). EcSSB has been reported to show a higher level of binding cooperativity than human RPA (14), but lower than that of gp32 (17).

To study coordination between the neighboring SSB molecules on a long ssDNA molecule, it is necessary to solve the structures of SSB–ssDNA complexes. However, it is still challenging to build high-resolution SSB–ssDNA models with most studied SSBs, largely due to the extremely flexible nature of SSB–ssDNA filaments, which often lack a highly ordered arrangement. Early electron microscopy (EM) work has demonstrated that the cooperative binding of EcSSB organizes ssDNA into repeating chains of spherical beads and protein-free linkers (18), such as the classic 'beads on a string' arrangement of chromatin. With an increasing EcSSB:ssDNA mass ratio, the beaded complex changes into a more smoothly contoured appearance (19). EM and atomic force microscopy (AFM) have also revealed that RPA binding organizes ssDNA into nucleosome-like quaternary structures (20,21). To study SSB–ssDNA structures at higher resolution, investigators have focused on adjacent SSB functional units on short fragments of ssDNA, such as crystal packing analysis of *Bacillus subtilis* SsbA tetramers (22), and cryo-EM of dimeric RPA assembly on ssDNA (23).

Here, we report structural and biochemical studies of a baculovirus SSB, LEF-3, from *Helicoverpa armigera* nucleopolyhedrovirus (hereafter referred to as HaLEF-3). Homologs of LEF-3 are present in all alpha- and beta-baculovirus genomes (24,25). LEF-3, as an essential protein for viral replication, localizes to the virogenic stroma where baculoviral DNA replication and nucleocapsid assembly occur (26). In addition to its role as an SSB in the metabolism of viral DNA (24,27), LEF-3 is involved in the nuclear transport of viral helicase (28) and interacts with alkaline nuclease, which participates in the repair of double-strand breaks and in homologous recombination (29). Our studies showed that HaLEF-3 is structurally distinct from other previously reported SSBs in the absence or presence of ssDNA substrates, and helical HaLEF-3–ssDNA filaments are structurally more stable for cryo-EM studies at atomic resolution.

MATERIALS AND METHODS

Cloning, expression and purification of HaLEF-3

The gene encoding HaLEF-3 (GenBank no. NP_203620.1) was amplified by polymerase chain reaction (PCR) from the genomic DNA of *Helicoverpa armigera* nucleopolyhedrovirus C1 strain. The amplified products containing two restriction sites (BamHI and SalII) were cloned into the same sites of the pET-28a vector (Novagen). The resulting expression plasmid of HaLEF-3 with an N-terminal His_6-tag was transformed into *E. coli* BL21 (DE3) cells, which were grown in Luria–Bertani medium supplemented with kanamycin (100 μg/ml) at 37°C. The growth culture was induced by 0.5 mM isopropy-β-D-thiogalactopyranoside (IPTG; at OD_{600} = 0.6) for 18 h at 16°C. The cells were harvested by centrifugation at 4000 *g* and 4°C for 30 min and resuspended in buffer containing 150 mM NaCl, 20 mM imidazole, 50 mM Tris–HCl, pH 7.5, 5% (v/v) glycerol and 3 U/ml DNase I. A stock solution of 1 M imidazole (pH 7.5) was used for the preparation of buffer solutions. The cell suspension was then lysed by sonication and clarified by centrifugation at 10 000 *g* and 4°C for 60 min. The supernatant was purified by metal affinity chromatography with Ni-NTA agarose (TransGen), and extensive washing was performed using high-salt buffer containing 1.2 M NaCl before elution with the buffer containing 250 mM imidazole. Size-exclusion chromatography (SEC) was carried out for further purification with a Superdex 200 (10/300) GL column (GE Healthcare) equilibrated with 150 mM NaCl, 250 mM imidazole, 50 mM Tris–HCl, pH 7.5, 5 mM dithiothreitol (DTT) and 5% (v/v) glycerol. Purified wild-type HaLEF-3 protein was concentrated to ∼5 mg/ml in the SEC elution buffer with a 30 kDa MWCO spin concentrator (Millipore). Protein concentration was determined by the Bradford assay. DNA/RNA contamination for each HaLEF-3 preparation was evaluated by measuring the A_{260}/A_{280} ratio using NanoPhotometer N60 (IMPLEN). The purity of HaLEF-3 (>90%) was determined by denaturing 12% sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE) followed by Coomassie Blue staining. Site-directed mutagenesis of HaLEF-3 was carried out using the Mut Express II Fast Mutagenesis Kit V2 (Vazyme), in accordance with the manufacturer's instructions, and the resulting alanine substitution mutants (K164A, E184A, R268A, K271A, S292A, R294A, Y311A and N361A) were expressed and purified using the same procedures as for wild-type HaLEF-3.

Crystallographic studies of ssDNA-free HaLEF-3

Crystals were grown at 16°C using the hanging drop vapor diffusion method by mixing 1 μl of freshly prepared protein solution with 1 μl of well solution (1.2 M sodium phosphate monobasic, 800 mM potassium phosphate dibasic, 100 mM CAPS/NaOH, pH 10.5, 200 mM lithium sulfate). Crystals were cryoprotected using 15% (v/v) glycerol before cooling to liquid nitrogen temperature. The best crystals diffracted to 3.5 Å resolution on the BL19U1 beamline of the Shanghai Synchrotron Radiation Facility (SSRF), using a Pilatus3 6M detector. Crystals were in space group C2, with a = 250.059 Å, b = 117.738 Å, c = 128.473 Å and β = 108.119°. All diffraction data were processed using the program XDS (30).

The crystal structure of HaLEF-3 was solved by two rounds of molecular replacement (MR) with PHENIX.Phaser (31). We used different truncated forms of the HaLEF-3 structure model based on the cryo-EM density map (see below) as search templates for MR. We first located eight copies of the C-terminal fragment of HaLEF-3 (residues 127–379) in the asymmetric unit, which were then used as a fixed partial model for the second round of MR to locate the N-terminal fragment (residues 49–122). The models were improved by rounds of manual

adjustment using Coot (32) and crystallographic refinement with PHENIX.refinement. Model validation was performed using MolProbity (33). The crystallographic data collection and refinement statistics are summarized in Supplementary Table S1.

Assembly of the HaLEF-3–ssDNA complex *in vitro* and electrophoretic mobility shift assay (EMSA)

The synthetic oligo-deoxyadenosine sequences (dA_8, dA_{12}, dA_{16}, dA_{20}, dA_{30}, dA_{40}, dA_{50}, dA_{60} and dA_{80}) were purchased from Sangon Biotech. dA_{140} was purchased from Tsingke Biotechnology. Except for dA_{60} and dA_{140}, which were used for cryo-EM data collection, all of the synthesized oligonucleotides were labeled with 6-carboxy-fluorescein at both the 5′ and 3′ ends. Bacteriophage φX174 ssDNA (∼5400 nt in length) was purchased from NEB. ssDNA powder was dissolved in ddH_2O at a concentration of 0.6 mg/ml and stored at −20°C. In accordance with the manufacturer's instructions, λ exonuclease (NEB) treatments were performed on dsDNA fragments prepared by PCR to produce unlabeled ssDNA strands of various lengths ranging from 150 to 1100 nt. The HaLEF-3 expression plasmid was used as the dsDNA template in the PCR, with a pair of forward primer and 5′-phosphorylated reverse primer.

In a 10 μl reaction volume, different amounts of freshly prepared HaLEF-3 [150 mM NaCl, 250 mM imidazole, 50 mM Tris–HCl, pH 7.5, 5% (v/v) glycerol] were mixed with 1 μl of different ssDNA substrates, at mass ratios of 0:1, 1:1, 5:1, 10:1, 20:1, 40:1, 60:1 and 80:1 for SSB:ssDNA (μg:μg) (refer to Supplementary Tables S2–S4 for corresponding SSB:ssDNA molar ratios and nucleotide:SSB molar ratios). The mixtures were incubated at room temperature for 60 min. Separation of bound and unbound ssDNA species was carried out by electrophoresis (10 min, 200 V) at room temperature on a 1% agarose gel in 1× TAE buffer, unless otherwise specified. Fluorescein-labeled ssDNA and ssDNA–SSB complexes were detected under UV transillumination. EMSA results shown below are representative of at least three independent experiments.

Cryo-EM sample preparation and data collection

Freshly prepared wild-type HaLEF-3 was dialyzed to the binding buffer containing 150 mM NaCl, 150 mM sodium citrate, pH 6.3, and 5 mM DTT, and was concentrated to 0.8 mg/ml. SSB–ssDNA complexes were formed at a mass ratio of 20:1 for SSB:ssDNA (μg:μg), and were first negatively stained and examined under a transmission electron microscope to evaluate sample quality. A 3 μl aliquot of the sample at a protein concentration of ∼0.05 mg/ml was applied to a glow-discharged 400 mesh copper grid coated with a thin carbon film. After 30 s of adsorption at room temperature, the grid was negatively stained with 2% (w/v) uranyl acetate and allowed to air-dry. Samples were examined using a Tecnai G2 20 TWIN microscope (FEI) (operating at 200 kV) equipped with a Cantega G2 CCD camera (Olympus).

For cryo-EM specimen preparation, 3 μl of SSB–ssDNA sample at a concentration of ∼0.5 mg/ml was applied onto a glow-discharged 300 mesh holey carbon copper grid (Quantifoil R 1.2/1.3). After excess sample had been removed by blotting with filter paper for 3.5–4.5 s, the grid was plunge-frozen in liquid ethane using a Vitrobot Mark IV plunger (Thermo Fisher Scientific), with the chamber maintained at 25°C and 85% humidity.

Cryo-EM single-particle data were collected on a CRYO ARM 300 electron microscope (JEOL) operated at 300 kV with an in-column energy filter (slit width 30 eV). Movies were recorded at a nominal magnification of ×50 000 in super-resolution mode by a K3 direct electron detector (Gatan) using Serial-EM (34) through the beam-image shift method, corresponding to a physical pixel size of 0.495 Å. Each movie stack was dose-fractionated to 40 frames with an accumulated dose of ∼40 $e^-/Å^2$ and a total exposure time of 2 s.

Cryo-EM data processing

Recorded movie stacks were 2 × 2 binned (resulting in a pixel size of 0.95 Å) and subjected to beam-induced motion correction using MotionCor2 implemented in RELION (35). After manual selection, dose-weighted micrographs were then imported into cryoSPARC (36) for subsequent data processing. The contrast-transfer function (CTF) parameters were estimated by the patch-based CTF estimation procedure. Micrographs with a CTF fit resolution better than 4.5 Å and a detected defocus range of −0.5 to −2.5 μm were retained for structural reconstructions.

For the HaLEF-3–dA_{60} complex, 631 201 particles were auto-picked in a subset (220 micrographs) of 2268 micrographs by the blob picker in cryoSPARC and subjected to reference-free two-dimensional (2D) classification. Eight representative classes were selected as references for template-based particle picking for the whole dataset. After multiple rounds of 2D classification, 387 121 particles were selected from 10 good classes for *ab initio* reconstruction of four maps, which were used as starting models for heterogeneous refinement. The best three-dimensional (3D) class, containing 149 026 particles, yielded a map at an overall resolution of 4.7 Å using non-uniform refinement and CTF refinement, without applying any symmetry. The resolution was calculated based on the gold-standard Fourier shell coefficient (FSC) = 0.143 criterion.

For the HaLEF-3–dA_{140} filamentous complex, 1 193 748 overlapping segments were picked from 3123 micrographs using the filament tracer within cryoSPARC with a separation distance between segments of 40 Å. These particles were subtracted with a box size of 320 pixels and were subjected to 2D classification. To obtain more particles for calculations, a template-based auto-picking strategy was performed with the best 2D averages as templates; a total of 5 784 262 particles were picked. After 2D classification, 770 471 straight helical particles were selected for particle subtraction with a box size of 220 pixels. After a new round of 2D classification, 605 825 particles were retained for *ab initio* reconstruction and heterogeneous refinement without applying helical symmetry. The generated 4.1 Å helical map based on 530 580 particles was used for a global search for helical symmetry parameters. Particles were re-subtracted using different box sizes (200–260 pixels) for further helical refinement, while helical parameters were iteratively refined

locally. The best helical reconstruction (box size = 240 pixels) has a global resolution of 3.5 Å, with an axial rise of 15.105 Å and a helical twist of −108.281°. Particles were then symmetry expanded with the helical parameters, and local refinement without imposing any symmetry was performed to yield a 3.1 Å map using a mask encompassing only one HaLEF-3 molecule and nearby nucleotides.

Model building of HaLEF-3–ssDNA complexes

A starting atomic model for HaLEF-3 was built into the local refinement map using the phenix.map_to_model tool. An initial model for the nucleotide fragment was built using RCrane (37). The model of the HaLEF-3–dA_{140} complex was improved by iterative cycles of real-space refinement implemented with phenix.real_space_refinement (38) and interactive modifications in Coot. The model refinement statistics are summarized in Supplementary Table S5. To better visualize the molecular organization of HaLEF-3 molecules in the helical assembly, the model was then docked back into the 3.5 Å helical map to generate four HaLEF-3 molecules and longer ssDNA along the helical path, followed by real-space refinement. Figures of 3D volume maps and the atomic models were prepared with either UCSF ChimeraX (39) or PyMOL (The PyMOL Molecular Graphics System, version 1.8, Schrödinger, LLC), and local resolution estimation was performed in cryoSPARC.

RESULTS

Expression of HaLEF-3

Full-length HaLEF-3 was expressed in *E. coli* cells with a His_6-tag at the N-terminus. DNase I (in the cell lysis buffer) and a high-salt wash buffer (containing 1.2 M NaCl) were used to minimize DNA contamination, and the A_{260}/A_{280} ratio of purified protein was measured for each preparation (lower than 0.5). Full-length HaLEF-3 was purified to near homogeneity based on SDS–PAGE (molecular mass of ∼47 kDa). On SEC, HaLEF-3 was eluted (∼11 ml) close to the position of the exclusion volume (∼8 ml) of the column (Supplementary Figure S1A), suggesting that HaLEF-3 in solution was in highly oligomeric states.

Crystallization of ssDNA-free HaLEF-3

To investigate how HaLEF-3 molecules organized into high-order oligomers in the absence of ssDNA, we crystallized HaLEF-3 under conditions featuring a high concentration of salt. The 3.5 Å crystal structure of HaLEF-3 in the ssDNA-free state was solved by molecular replacement using our cryo-EM structures of HaLEF-3 as initial models (see below). In the asymmetrical unit, there are eight HaLEF-3 molecules arranged into an octamer with D4 symmetry (Figure 1).

Most residues for the eight molecules are clearly defined in the electron density map, with the exceptions of 47 residues at the N-termini and some residues between E248 and Q259. The lack of obvious degradation for the crystalized HaLEF-3 (Supplementary Figure S1B) suggested that the missing N-terminal residues are largely due to the flexibility of this region. Previous studies demonstrated that N-terminal tail residues may contain the nuclear localization signal to mediate the transport of LEF-3 and associated helicase into the nucleus (40). Because K48 is located on the outer surface of the HaLEF-3 octamer (Figure 1D), the missing N-terminal residues are expected to be accessible from the exterior of the octamer, promoting their interactions with host proteins.

Structure of the HaLEF-3 protomer in the octamer

The HaLEF-3 molecule exhibits a rod-shaped structure (∼80 Å in length), which is mainly composed of two structured domains (Figure 1A): domain I (DI) containing one OB fold (OB1, from K48 to Q119) and a C-terminal domain II (DII) consisting of two OB folds (OB2 and OB3) (Figure 1A, B). The two domains are connected by a highly flexible hinge (FH) located between E118 and K131. A search of the Protein Data Bank with the online Dali server (41) did not identify any close overall structural homologs to HaLEF-3, and the most similar partial structure was found for some OB folds from proteins showing high affinity for telomeric ssDNA, including *Tetrahymena* telomerase processivity factor Teb1 (PDB 3U58) (42) and *Oxytricha nova* telomere end-binding protein (PDB 1K8G) (43).

Within DII, OB2 and OB3 are connected by a rigid linker (RL, amino acids 247–271), in which one segment (N265–K271) is highly conserved within alpha-baculoviruses and beta-baculoviruses (Supplementary Figure S2). A hydrophobic core is formed by residues from OB2 (Y152, F154 and V166) and OB3 (I273 and Y351) (Supplementary Figure S3A). Meanwhile, the side chain of Y157 is inserted into a hydrophobic pocket made by residues from OB2 (F295, I310, I327, L334 and L379) (Supplementary Figure S3B). These hydrophobic interactions at the OB2–OB3 interface make OB2 and OB3 act as a single rigid domain (DII).

Structure of the HaLEF-3 octamer

HaLEF-3 molecules are arranged into a homo-octameric assembly comprising two C4-symmetric rings, which are distinct from homo-tetrameric EcSSB (D2 symmetry) (11). Viewed along the C4 axis (Figure 1C), the overall profile of the HaLEF-3 octamer can be likened to a square with a central hole (∼35 Å of diameter) and four sides (∼110 Å in length). Four extended HaLEF-3 molecules assemble into a tetrameric ring in an end-to-end arrangement in which OB1 interacts with OB3 from the neighboring molecule (Figure 1C). DI (OB1) is located at the vertex, while DII (OB2 and OB3) forms the side of the square. The intermolecular OB1–OB3 contacts are dominated by hydrophobic and electrostatic interactions. Two hydrophobic cores are formed by the residues from OB1 (L52, M64, Y86, Y107 and L112) and the residues from neighboring OB3 (L329, V332 from α6, L340 from α7, and nearby L345) (Supplementary Figure S3C). Electrostatic interactions occur between positively charged K110 on OB1 and a negatively charged patch on OB3 contributed to by E371 and E377 (Supplementary Figure S3D).

The two rings stack in a face-to-face mode, forming a D4-symmetric octamer to bury ∼20 600 $Å^2$ of surface area between the two tetramers, compared with ∼109 100 $Å^2$ of

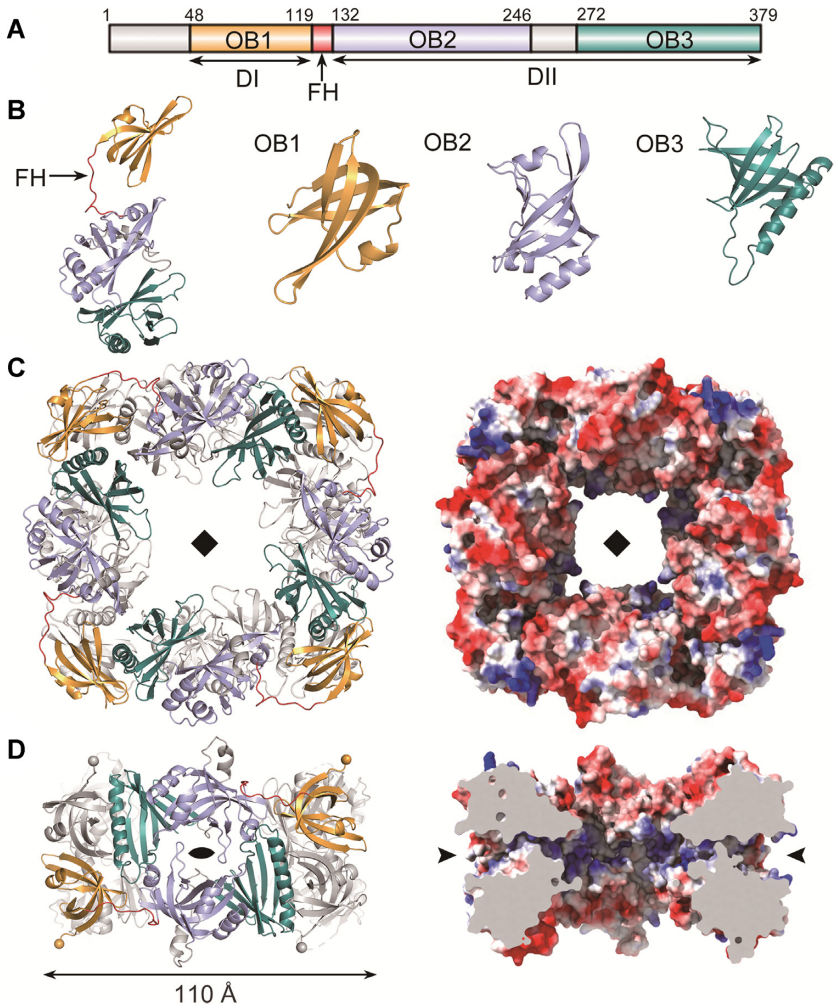

Figure 1. Crystal structure of HaLEF-3. (**A**) Domain organization of the HaLEF-3 sequence. The full-length HaLEF-3 containing 379 amino acids is mainly divided into the DI domain containing one OB fold (OB1), a flexible hinge (FH) and the C-terminal DII domain consisting of two OB folds (OB2 and OB3) and a rigid linker (RL). (**B**) Overall structure of a HaLEF-3 protomer (left) and three isolated OB folds (right). FH (in red) is denoted by an arrow. The characteristic β-barrel and α-helix of isolated OBs are roughly placed in similar orientations. (**C**) The octameric assembly of ssDNA-free HaLEF-3 viewed down the C4 axis. (Left) Ribbon diagram of the octamer highlighting four protomers colored as in (B). Other protomers are colored in gray. (Right) Electrostatic surface potential of the HaLEF-3 octamer [electropositive (blue) and electronegative (red)]. (**D**) The HaLEF-3 octamer viewed down one C2 axis. (Left) Ribbon diagram of the octamer highlighting two protomers colored as in (B). The N-terminal K48 residues are denoted as spheres. (Right) A cutaway view through the center of the octamer to show the inner surface electrostatic distribution of HaLEF-3 octamer. The stacking of the two tetrameric HaLEF-3 rings results in positively charged side channels along the C2 axes (indicated by arrowheads).

solvent-accessible surface area of one octamer. The contacts between the two rings mainly involve antiparallel interactions between β1 strands from two neighboring DIs (Supplementary Figure S3E). Two DII domains are related by a C2 axis, along which there is a narrow channel connected to the central hole (Figure 1D, left). A number of positively charged residues (K164, K271, R268 and K360) near the central hole are distributed along the channel (Figure 1D, right), most of which are critical for binding ssDNA (see below).

Binding of HaLEF-3 to ssDNA

To evaluate the binding activity of *E. coli*-expressed HaLEF-3 to ssDNA, EMSA experiments were performed in the presence of dA_{30}. Agarose gels were used instead of polyacrylamide gels to resolve high molecular weight protein–DNA complexes (44). To increase the signal, fluorescent tags were attached to the two ends of dA_{30}. When a saturating concentration of SSB (mass ratios of SSB:dA_{30} >20:1) was incubated with dA_{30}, there was only one band for the complex (Figure 2A). When the SSB:ssDNA ratio was 10:1 or 20:1, there were two sharp bands indicative of cooperative binding (see below): the faster migrating one representing free ssDNA and the slower one representing the SSB–ssDNA complex. The number of nucleotides bound by one SSB molecule was estimated to be between 7.2 (SSB:ssDNA mass ratio = 20:1) and 3.6 (SSB:ssDNA mass ratio = 40:1) (Supplementary Table S4). Previous studies showed that the association of SSBs with ssDNA is influenced by solution variables, notably the ionic strength (45). When a range of buffer con-

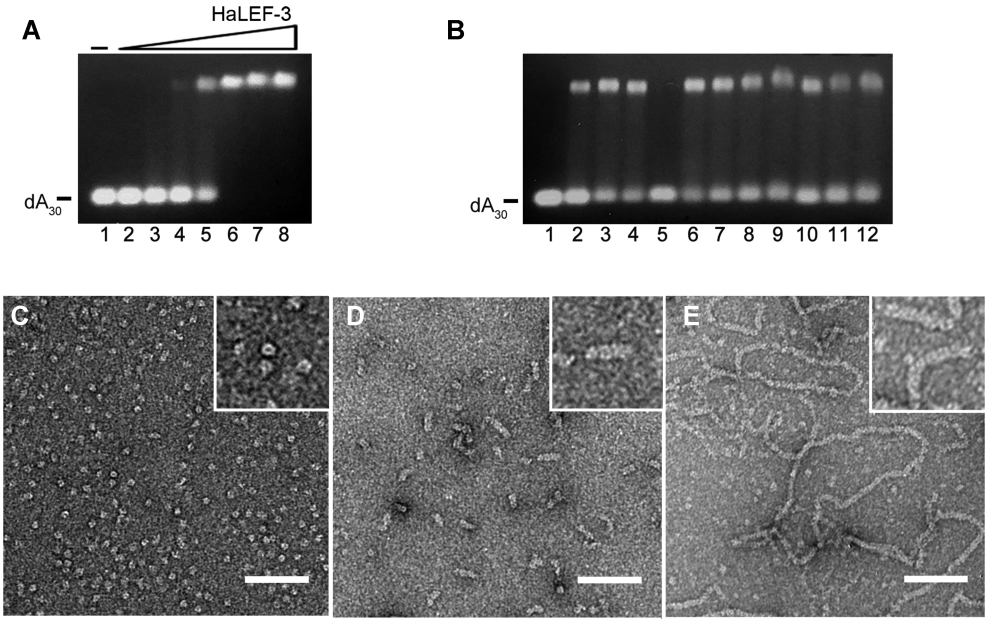

Figure 2. ssDNA binding capacity of wild-type HaLEF-3 analyzed by EMSA and negative staining TEM. (**A**) Titration of dA_{30} with increasing concentrations of HaLEF-3. SSB:ssDNA mass ratios were 0:1 (lane 1), 1:1 (lane 2), 5:1 (lane 3), 10:1 (lane 4), 20:1 (lane 5), 40:1 (lane 6), 60:1 (lane 7) and 80:1 (lane 8). (**B**) Highly cooperative SSB–ssDNA complexes were formed in different binding buffers. Each reaction was carried out in a 10 μl reaction mixture containing 1 μl of SSB (2 mg/ml, 150 mM NaCl, 250 mM imidazole, 50 mM Tris–HCl, pH 7.5, 5 mM DTT), 1 μl of dA_{30} (0.1 mg/ml, in ddH_2O) and 8 μl of binding buffer: PBS (lane 2), 150 mM NaCl, 250 mM imidazole, 50 mM Tris–HCl, pH 7.5 (lane 3), 150 mM NaCl, 25 mM CHES, pH 9.0 (lane 4), 150 mM NaCl, 25 mM sodium acetate, pH 5.5 (lane 5), 150 mM NaCl, 50 mM Tris–HCl, pH 7.5 (lane 6), 300 mM NaCl, 50 mM Tris–HCl, pH 7.5 (lane 7), 500 mM NaCl, 50 mM Tris–HCl, pH 7.5 (lane 8), 1000 mM NaCl, 50 mM Tris–HCl, pH 7.5 (lane 9), 150 mM NaCl, 150 mM sodium citrate, pH 6.3 (lane 10), 500 mM NaCl, 150 mM sodium citrate, pH 6.3 (lane 11), 1000 mM NaCl, 150 mM sodium citrate, pH 6.3 (lane 12). Lane 1 shows a control reaction containing only dA_{30} resolved in ddH_2O. (**C–E**) Negatively stained HaLEF-3 in complex with different lengths of ssDNA, dA_{60} (C), dA_{140} (D) and the φX174 ssDNA genome (∼5400 nt) (E). Typical particles are shown in the insets. The scale bars represent 100 nm.

ditions were tested, HaLEF-3 interacted with ssDNA in similar ways at a SSB:ssDNA ratio of 20:1 (Figure 2B), except for when an acidic buffer (pH 5.5) was applied, suggesting that HaLEF-3 is more tolerant of the changes in binding conditions than other SSBs.

Negative staining of the SSB–ssDNA complex

Assemblies of HaLEF-3 with various lengths of ssDNA were visualized by the negative staining technique under transmission electron microscopy (TEM). Undersaturated amounts of HaLEF-3 were used (input SSB:ssDNA mass ratio of 20:1) to reduce background noise caused by unbound HaLEF-3 molecules. After testing for different assembly buffers, the assemblies with the best contrast were observed with a mildly acidic buffer containing an intermediate concentration of salt (see the Materials and Methods). Most HaLEF-3–dA_{60} particles appeared near-spherical (∼10 nm diameter) with a central hole (filled by dark stains) (Figure 2C). HaLEF-3 binds to ssDNA of 140 nt or longer to form structurally different mixtures: HaLEF-3–dA_{60}-like spherical particles and short flexible filaments with different lengths but similar diameters to near-spherical particles (Figure 2D; Supplementary Figure S4A–F). When HaLEF-3 interacts with φX174 ssDNA, large closed filament rings were formed with the same diameter as shorter filaments (Figure 2E).

Cryo-EM of HaLEF-3–ssDNA complexes

HaLEF-3–ssDNA complexes were further analyzed using cryo-EM techniques. Two-dimensional averages of the near-spherical particles identified in negatively stained TEM revealed that they are actually short helices with a central channel (Figure 3A). The structure was solved to 4.7 Å resolution (Figure 3B), and applying helical symmetry did not improve the resolution. To take advantage of the large amount of symmetry information present in the helical assembly, we used longer dA_{140} to assemble SSB–ssDNA complexes. On the basis of cryo-EM images and 2D averages, HaLEF-3–dA_{140} filaments are more or less bent (Figure 3C). To pick straight helical particles, long segments were initially picked by a filament-tracing algorithm and extracted (box size = 320 pixels) (Supplementary Figure S5). After 2D classification, the best straight classes were selected as the templates for template-based particle auto-picking for 2D classification. The best particles were re-extracted with a smaller box size (220 pixels) for further classification. In this way, sufficiently straight helical particles from the curved filaments were isolated for helical reconstructions. The overall resolution for the final helical reconstruction was 3.5 Å (Figure 3D) and local resolution ranged from 3.0 to 3.8 Å (Supplementary Figure S5). HaLEF-3 binds to dA_{140} to form a left-handed filament that has ∼3.3 HaLEF-3 subunits per turn and a pitch height of 50.2 Å. Local-refine strategies were applied to further improve the density of one HaLEF-3 molecule and nearby nu-

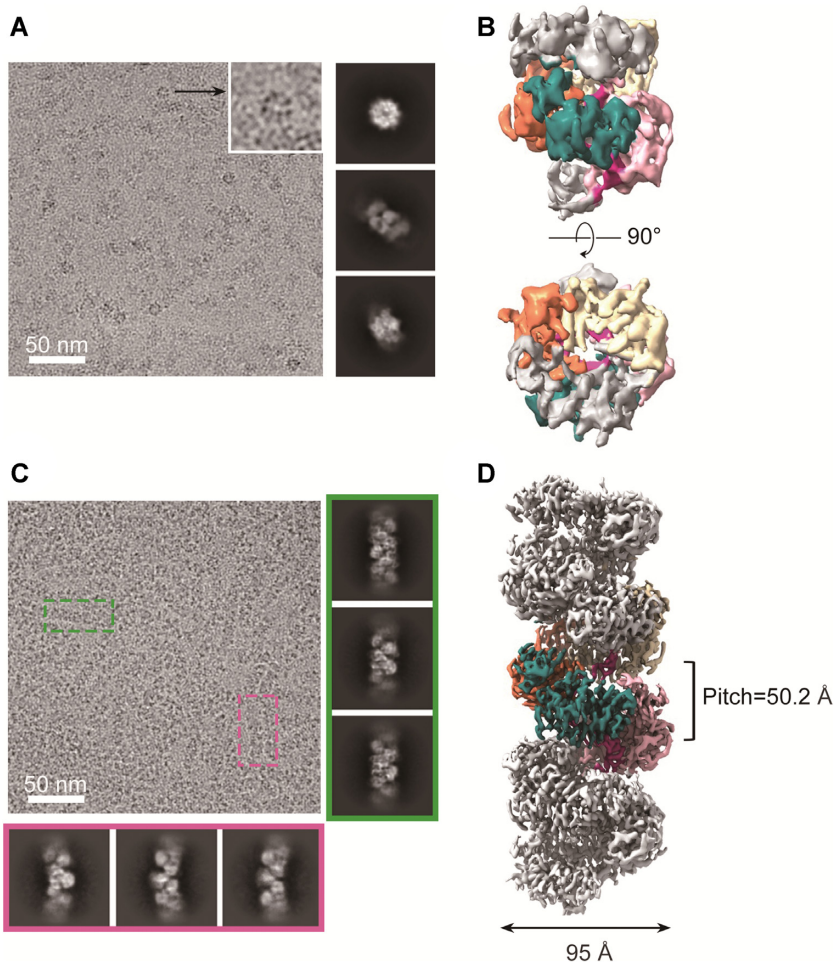

Figure 3. Cryo-EM analysis of HaLEF-3–ssDNA complexes. (**A**) (Left) A typical cryo-EM micrograph of HaLEF-3–dA_{60} particles. A round particle with a central hole is shown in the inset. (Right) Three representative 2D-class averages. In the second class, the particles clearly have short helical features. The scale bar represents 50 nm. (**B**) Two orthogonal views of the 3D reconstruction of HaLEF-3–dA_{60} particles (contoured at 11.1 σ above average) without applying helical symmetry. Four HaLEF-3 subunits in distinct colors are arranged along the left-handed helical path, and the ssDNA strand is highlighted in magenta. (**C**) A typical cryo-EM micrograph of HaLEF-3–dA_{140} filaments. Two-dimensional-class averages of straight or bent helical segments are highlighted by green or magenta boxes, respectively. (**D**) Helical reconstruction of HaLEF-3–dA_{140} filaments (contoured at 4.8 σ) calculated based on straight 2D-class averages. HaLEF-3 subunits and ssDNA are colored as in (B).

cleotides. Clear electron density of HaLEF-3 and ssDNA in the resulting 3.1 Å local map allowed direct model building. Although full-length HaLEF-3 was used to form the HaLEF-3–ssDNA complexes, the N-terminus of HaLEF-3 was still invisible, like its counterpart in the nucleotide-free HaLEF-3 crystal structure. Docking four HaLEF-3 subunits into the 3D reconstructions of HaLEF-3–dA_{60} and HaLEF-3–dA_{140} suggested that helical features of HaLEF-3–dA_{60} and HaLEF-3–dA_{140} are basically identical (Figure 3B, D).

Non-specific interactions between HaLEF-3 and ssDNA

Successive HaLEF-3 molecules bind ssDNA directionally with the C-terminal DII domain at the 3′ end of the HaLEF-3–ssDNA complex, whereas the DI domain is not directly involved in SSB–ssDNA interactions (Figure 4A). The ssDNA chain runs within a basic groove mainly contributed to by K164, R268, K271, R294, R313 and K360 from DII (Figure 4B), spiraling along the inner channel of the tubular scaffold assembled by HaLEF-3. The positively charged groove appears to bind the phosphate backbone while a set of six stacked bases face the central channel (Figure 4C). Each HaLEF-3 subunit interacts with six nucleotides: $dA_{(1)}dA_{(2)}dA_{(3)}dA_{(4)}dA_{(5)}dA_{(6)}$ (dA nucleotides are numbered in parentheses), in agreement with our EMSA data (above). Base-stacking interactions between $dA_{(3)}$ and $dA_{(4)}$ are disrupted by hydrophobic interactions between the base of $dA_{(3)}$ and the phenol group of Y311 (Figure 4C, D), which divides the six nucleotides into two groups [$dA_{(1)}dA_{(2)}dA_{(3)}$ and $dA_{(4)}dA_{(5)}dA_{(6)}$] and each forms continuous base stacking with the nucleotides associated with the adjacent HaLEF-3 subunits. Y311A mutants showed similar binding to dA_{60} but decreased binding to dA_{30}, compared with wild-type HaLEF-3 (Figure 5A), suggesting that disruption of base-stacking interactions by Y311 may be compensated for by other molecular forces between HaLEF-3 and long ssDNA fragments.

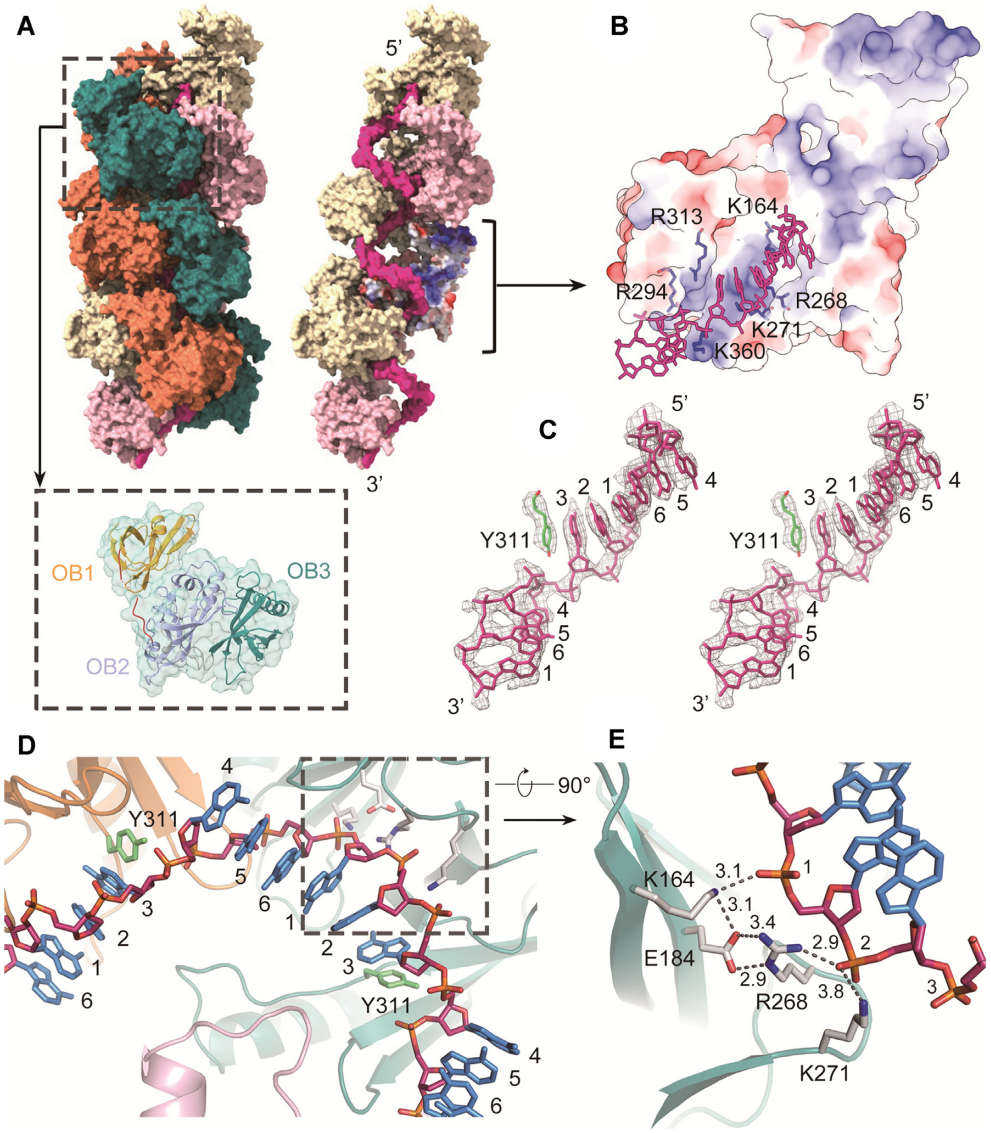

Figure 4. Atomic model of HaLEF-3 in complex with ssDNA. (**A**) (Left) Surface representation of the HaLEF-3–ssDNA filament, with each HaLEF-3 subunit along the helical path colored teal, orange, beige and pink. (Right) Half of the subunits are removed to show ssDNA (colored in magenta) bound to the inner surface of the HaLEF-3 tube. The inset shows the arrangement of OB1, OB2 and OB3 in one HaLEF-3 subunit (as a transparent surface). (**B**) Inner surface electrostatic map of one HaLEF-3 subunit. ssDNA binds to a positively charged groove (blue). (**C**) Stereo view of an isolated electron potential map (7.3 σ) of ssDNA superimposed with the atomic model of ssDNA (in the same orientation as in B). The ssDNA bases are numbered sequentially in the 5′ to 3′ direction. The side chain density for Tyr311 stacking against the base of $dA_{(3)}$ is shown. (**D**) The isolated HaLEF-3 subunits in the ribbon representation depict the SSB-binding sites on HaLEF-3 viewed roughly along the long axis of the filament. There are six nucleotides associated with one HaLEF-3 subunit. (**E**) Key residues involved in the SSB–ssDNA interaction network. Residues are labeled and represented as stick models. Distances (dashed lines) are given as Å.

K164, R268 and K271 form hydrogen bonds and salt bridges with phosphate oxygen atoms of nucleotides (Figure 4E; Supplementary Table S6). K164A and R268A mutants completely lost their DNA binding capability for dA_{30} and dA_{60} (Figure 5A), whereas K271A showed weaker binding for dA_{30} and dA_{60} than wild-type HaLEF-3. Importantly, the negatively charged residue E184 forms hydrogen bonds and salt bridges with K164 and R268 (Figure 4E). Similar to K164A and R268A, the E184A mutant also resulted in complete loss of binding for dA_{30} and dA_{60}, suggesting that E184–K164 and E184–R268 interactions are required for the positioning of positively charged K164 and R268 for ssDNA binding. In addition, main chain O or N atoms of S267, S269 and Q290 and side chains of S292, R294 and N361 are involved in hydrogen bonding interactions with ssDNA (Supplementary Table S6). Mutagenesis studies of S292, R294 and N361 showed that they had no obvious impact on ssDNA binding (Supplementary Figure S6A). Taking these findings together, the non-specific affinity of HaLEF-3 for ssDNA is mainly mediated by electrostatic interactions between phosphate oxygens and lysine and arginine residues (K164, R268 and K271).

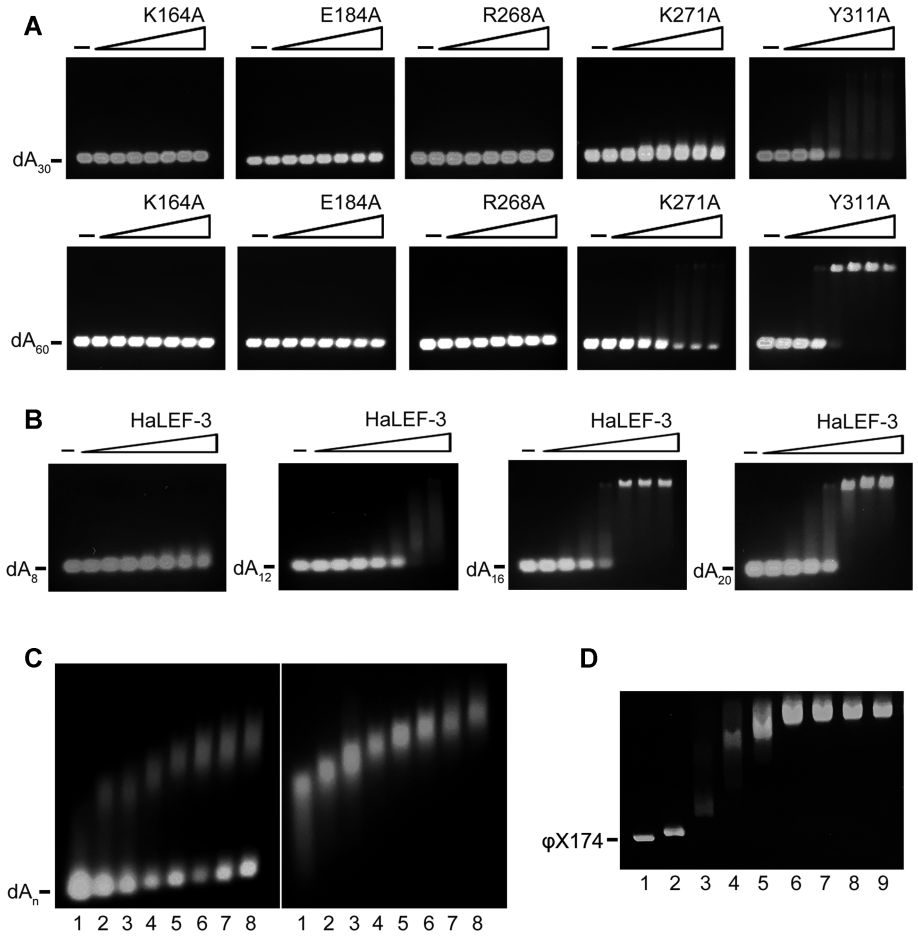

Figure 5. ssDNA binding activity analyzed by EMSA. (**A**) Titration of dA_{30} (upper panel) or dA_{60} (lower panel) with increasing concentrations of HaLEF-3 variants. Protein was mixed with ssDNA at the same SSB:ssDNA mass ratios as in Figure 2A. (**B**) Wild-type HaLEF-3 was titrated agaist ssDNA of various lengths (dA_8, dA_{12}, dA_{16} and dA_{20}). (**C**) EMSA was performed with ssDNA of different lengths on 0.6% agarose gels: dA_{12} (lane 1), dA_{16} (lane 2), dA_{20} (lane 3), dA_{30} (lane 4), dA_{40} (lane 5), dA_{50} (lane 6), dA_{60} (lane 7) and dA_{80} (lane 8). The assay was performed under conditions when ssDNA was present in excess (left, SSB:ssDNA = 20:1) or HaLEF-3 was oversaturated (right, SSB:ssDNA = 80:1). (**D**) Titration of φX174 ssDNA with HaLEF-3 at different SSB:ssDNA mass ratios: 0:1 (lane 1), 1:1 (lane 2), 5:1 (lane 3), 10:1 (lane 4), 15:1 (lane 5), 20:1 (lane 6), 25:1 (lane 7), 30:1 (lane 8) and 35:1 (lane 9). The reactions were analyzed by 0.6% agarose gel electrophoresis.

Local cooperative binding for short ssDNA

EMSA is a classical technique to qualitatively detect cooperative binding of SSB to ssDNA when there is an excess of ssDNA (45). For non-cooperative binding, ssDNA with an increasing amount of bound SSB gradually migrates more slowly, which produces a single band that migrates faster than the fully bound ssDNA. In contrast, cooperative binding would yield two sharp bands: the fastest one for free ssDNA and the other one representing the ssDNA–SSB complex in which ssDNA is bound in a saturated manner by multiple SSB molecules. As shown in our EMSA results, HaLEF-3 binds to dA_{30} with evident cooperativity (Figure 2A, lane 5). To study the binding cooperativity of HaLEF-3, we examined how different ssDNA fragments (dA_8, dA_{12}, dA_{16}, dA_{20}, dA_{30}, dA_{40}, dA_{50}, dA_{60} and dA_{80}) are involved in SSB–ssDNA interactions. Although dA_8 was expected to provide binding sites for one SSB molecule, its affinity for SSB was quite low (Figure 5B). The diffuse band for HaLEF-3–dA_{12} at high SSB:ssDNA mass ratios suggested that ssDNA fragments accommodating at least two SSBs were required for SSB–ssDNA binding. Compared with the sharp band for HaLEF-3–dA_{30} (Figure 2A, lane 5), less cooperative binding was observed for HaLEF-3–dA_{16} and HaLEF-3–dA_{20} (diffuse bands at SSB:ssDNA = 20:1, Figure 5B). With increasing ssDNA length, more SSB molecules are cooperatively bound to ssDNA (Figure 5C), and similar mobility patterns showing stronger cooperativity were observed for dA_{30}, dA_{40}, dA_{50}, dA_{60} and dA_{80} (Supplementary Figure S6B), indicating that ssDNA fragments accommodating at least four HaLEF-3 molecules are more desirable for efficient cooperative SSB–ssDNA binding than shorter nucleotide sequences.

Interestingly, at a low SSB:ssDNA mass ratio (1:1), HaLEF-3 formed one shifted band with φX174 ssDNA on the agarose gel (Figure 5D, lane 2), suggesting that HaLEF-3 randomly binds to ssDNA when only a small amount of SSB is present. In contrast, a broad diffusion migration pattern was found when more HaLEF-3 molecules were associated with φX174 ssDNA (lanes 3–5), suggesting that HaLEF-3 binds to long ssDNA less cooperatively than to

short ssDNA. We proposed that the local cooperative binding involved in short ssDNA tracts described above would be the major contributor to the non-random distribution of HaLEF-3 on long ssDNA molecules.

Structural transitions and rearrangement of HaLEF-3

Structural comparison of HaLEF-3 in the absence or presence of ssDNA showed that the architecture of individual OB folds and structure domains (DI and DII) is highly stable. The DI (OB1) domains in the two structures were superimposed very well with a root-mean-square deviation (rmsd) value of 0.55 Å when Cα positions of 67 residues aligned. Similarly, the DII (OB2 and OB3) domains were also well superimposed with an rmsd of 1.35 Å when 234 Cα atoms were aligned. In the ssDNA-free structure, HaLEF-3 self-assembles into octamers and the ssDNA-binding sites in DII reside in the narrow channel along the C2 axis (Figure 6A). To expose these binding sites to ssDNA, the two tetrameric rings must separate from each other. During the ring-to-helix transition, dramatic conformational changes in HaLEF-3 are mediated by the rotation of DI relative to DII around the FH region (Figure 6A).

In both crystal and cryo-EM structures, DI (OB1) of one HaLEF-3 molecule interacts with DII (OB3) of the adjacent molecule. Superimposition of a pseudo-molecule consisting of DII from one HaLEF-3 molecule and DI from another HaLEF-3 molecule revealed that the OB1–OB3 junction between two successive molecules laterally packed along the helical path was almost unchanged (Figure 6B), indicating that not all of the connectivity of HaLEF-3 molecules in the tetrameric ring may necessarily be disrupted during ring-to-helix transition. Based on buried surface analysis of neighboring HaLEF-3 molecules (Supplementary Table S7), there are >80 residues that are likely to be involved in the intersubunit interactions, and ∼50% of these residues contribute to the OB1–OB3 connectivity both before and after ssDNA binding. Mutagenesis studies of L340, a conserved residue at the OB1–OB3 interface (Supplementary Figures S2 and S3C), showed smaller SSB–ssDNA complexes for ssDNA of various lengths (Supplementary Figure S7), suggesting that the OB1–OB3 interactions are indeed critical for cooperative SSB–ssDNA binding.

DISCUSSION

In this study, we present structural and biochemical analyses of HaLEF-3, a long identified viral SSB in many baculoviruses. This SSB shows distinct structural features from well-studied SSBs (bacterial EcSSB, eukaryotic RPA and other viral SSBs), and thus may represent a new class of SSB. First, HaLEF-3 is assembled as octamers (two stacked tetramers) in the substrate-free state, in contrast to EcSSB (homotetramers), RPA (heterotrimers), gp2.5 (dimers) and gp32 (monomers) (5,6,11,13). The biological roles of HaLEF-3 octamers or tetramers need further investigation. Second, as shown in the cryo-EM reconstruction of HaLEF-3–ssDNA complexes, ssDNA is fully coated by HaLEF-3 subunits in a relatively rigid conformation, making high-resolution cryo-EM characterization feasible. The 3.5 Å helical reconstruction of the ssDNA–HaLEF-3 complex enables us to directly analyze SSB–SSB interactions along the ssDNA chain at better resolution than before.

As an SSB containing three OB folds, HaLEF-3 has shown unique structural dynamics for cooperative binding to ssDNA, which exhibits a distant resemblance to RPA. The RPA complex has a stable hetero-trimerization core consisting of three domains from RPA70, RPA32 and RPA14. These three domains have an OB fold followed by a C-terminal helix, which associates the three molecules into a stable heterotrimer by forming a three-helix bundle (13). The three other domains (DBD-A, DBD-B and DBD-F) on RPA70, each containing one OB fold, are flexible relative to the core. Similarly, HaLEF-3 has a relatively stable DII 'core' consisting of two OB folds (OB2 and OB3), while the DI domain (OB1) is connected to DII by a flexible hinge. Upon ssDNA binding, HaLEF-3 undergoes dramatic internal conformational changes, while the intermolecular interface between OB1 and OB3 from adjacent molecules remains almost intact for cooperative activities. However, unlike HaLEF-3, intermolecular interactions between the DBD-A domain of one RPA and the DBD-E domain of the adjacent RPA are less stable in cooperative ssDNA binding (23). It would be interesting to determine the functional roles of ssDNA–HaLEF-3 filaments in the eukaryotic cell when host SSBs (more discrete RPA heterotrimers or other eukaryotic SSBs) are also involved in DNA replication. In contrast to HaLEF-3 and RPA, other studied SSBs usually have a limited flexible region responsible for inter-SSB interactions during cooperative binding to ssDNA, like the intrinsically disordered linker of EcSSB (46) and the C-terminal 60 amino acid residues of ICP8 (47).

On the basis of our structural and biochemical observations, we propose a two-step local cooperative binding model for the assembly of ssDNA–HaLEF-3 helical filaments: initiation and extension (Figure 7). When an ssDNA fragment is presented to substrate-free HaLEF-3 octamers, the two tetrameric rings of the octamers separate from each other to expose once-hidden positively charged residues (K164, R268 and K271), which subsequently bind to the phosphate backbone of ssDNA. This process is accompanied by large conformational changes within HaLEF-3 (the movement of DI relative to DII), which drive multiple SSB molecules to organize into a spiral conformation to load onto ssDNA. Although two SSB molecules can initiate the ssDNA binding, efficient cooperative binding requires at least four HaLEF-3 molecules (Figure 5), which may be structurally favorable because four SSB subunits can arrange into a full turn in the HaLEF-3 helical scaffold. The four HaLEF-3 molecules derived from one ssDNA-free tetramer are therefore likely to undergo ring-to-helix transition to bind to ssDNA of different lengths to form the initial SSB–ssDNA cores (Figure 7).

For short ssDNA, straight helical architecture may grow on one spiral core by the incorporation of more HaLEF-3 molecules (Figure 7), and ssDNA in this case is fully coated by successive SSB molecules stabilized by the quaternary OB1–OB3 interactions. Because the electrostatic interactions that keep SSB bound to ssDNA are also essential for helical architecture, the extension of SSB molecules on ssDNA is expected to be dependent on the length of input ssDNA, as shown in Figure 5C, although HaLEF-3 tetramers

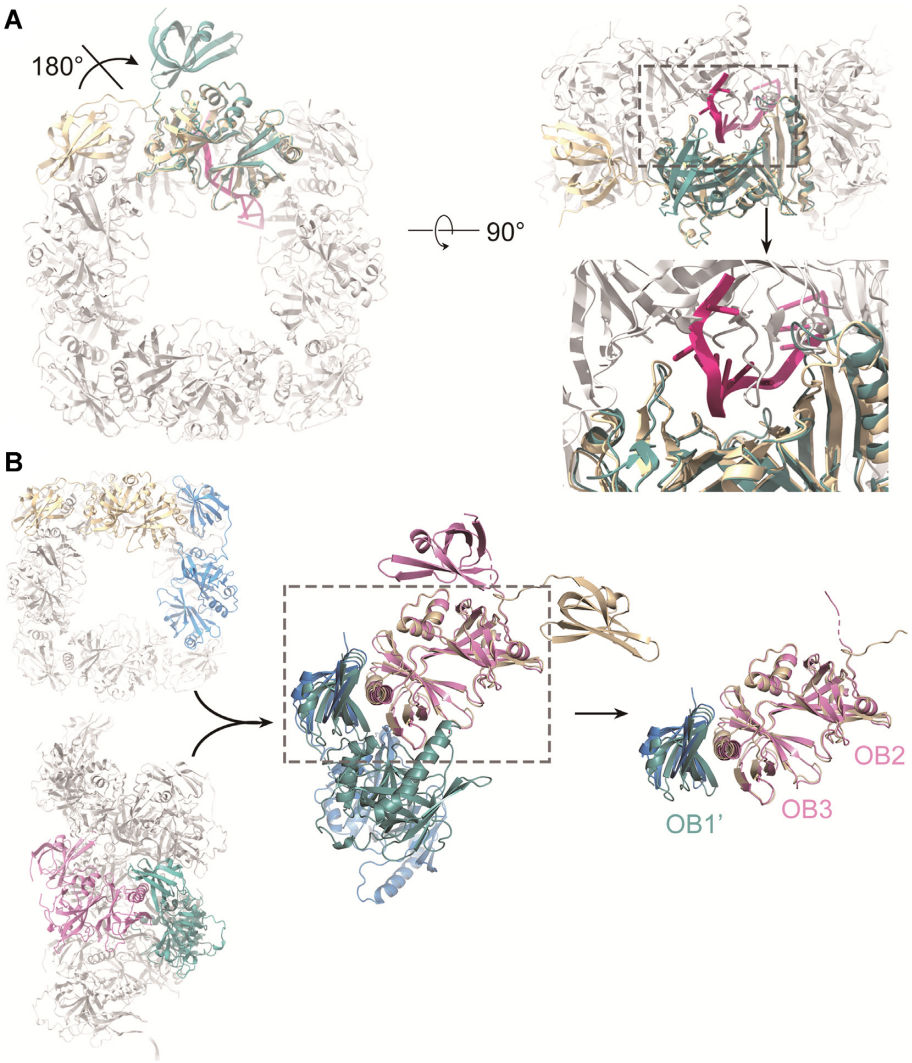

Figure 6. Comparison of ssDNA-free and ssDNA-bound HaLEF-3. (**A**) Orthogonal views of superimposition of an ssDNA-bound HaLEF-3 molecule (colored in teal) onto an ssDNA-free HaLEF-3 subunit (in yellow) in the context of the octameric architecture (in gray). (Left) DIIs from the two structures are well aligned, while the conformational change can be observed for DIs, which are related by ∼180° rotation about an axis roughly parallel to the diagonal line of the tetramer square. (Right) ssDNA is shown as magenta ribbon models, residing in the C2 channel. (**B**) Superimposition of DI (OB1) and DII (OB2 and OB3) from two adjacent HaLEF-3 molecules in ssDNA-bound and ssDNA-free states, highlighting that the interactions between DI and adjacent DII remain almost fixed.

could also load onto ssDNA as in the initiation step. For long ssDNA, helical nucleoprotein regions would grow simultaneously from multiple spiral cores that are randomly distributed on ssDNA, and gaps could be introduced between SSB-coated regions, when SSB is not sufficient to cover all of the binding sites or the vacant sites are insufficient to accommodate more SSB molecules. These gaps would result in local bending of the filaments as observed in Figure 3C, and the flexibility of the ssDNA–SSB filaments would be more pronounced with longer ssDNA.

HaLEF-3 may bind much longer ssDNA tracts (many kilobases) in the cell, and the extension of SSB–ssDNA filament would be more complicated than our assembled SSB–ssDNA complexes. In our cryo-EM models, ssDNA with two free ends follows the same spiral path as helically arranged HaLEF-3 subunits. However, it would be expected that ssDNA is not always free to twist at the ends of binding regions, such as closed circular φX174 ssDNA, and thus the growth of the SSB–ssDNA helical structure would inevitably stop at some sites. Diffusion of short core structures has been reported to occur along ssDNA for EcSSB (48,49) and RPA (50,51). Although a few bound HaLEF-3 molecules might use this strategy to adjust their distribution on ssDNA and form longer nucleoprotein tracts by merging two short complex regions, the strong interactions between ssDNA and multiple HaLEF-3 molecules are likely to limit the diffusion of helical HaLEF-3 stretches, which may keep the gap regions on ssDNA relatively stable.

In this assembly model, the flexibility that results from the discontinuous arrangement of HaLEF-3 on ssDNA may be biologically important. In the cellular context, SSB–ssDNA complexes usually interact with specific protein partners to form ssDNA intermediates, leading to different processing pathways. For instance, nucleoprotein filament structures

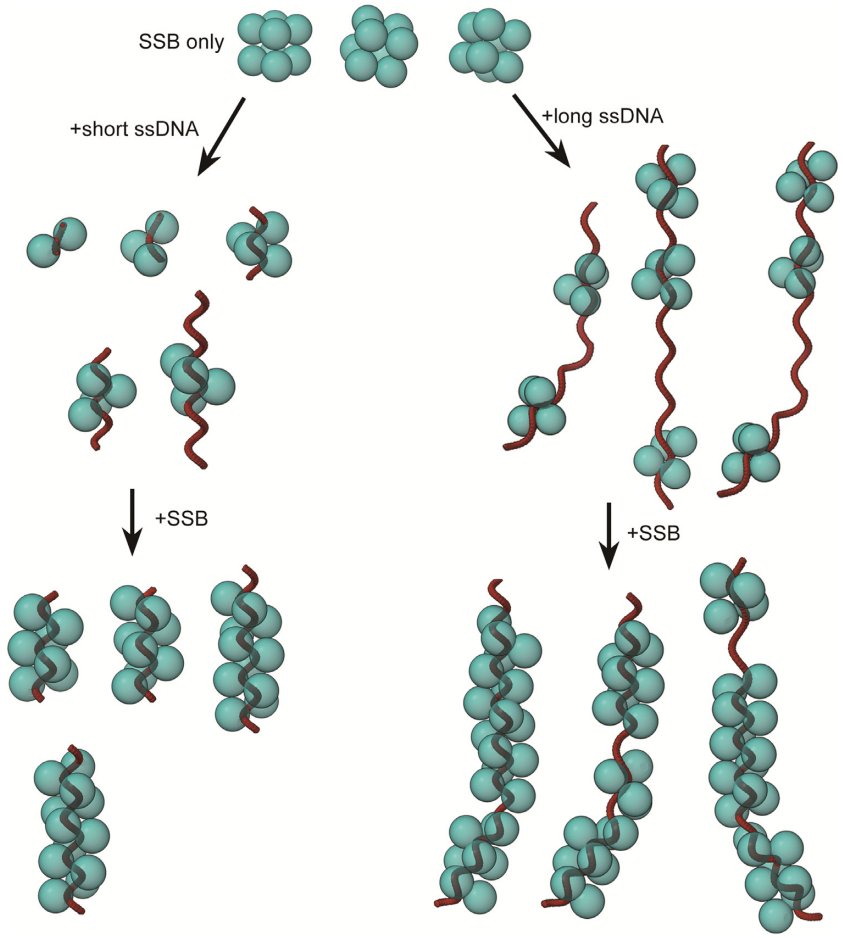

Figure 7. Proposed binding strategies of HaLEF-3 in association with ssDNA of various lengths. Before exposure to ssDNA, HaLEF-3 molecules are organized as octamers. For simplicity, HaLEF-3 subunits are depicted as transparent cyan spheres. Exposure to ssDNAs (shown as red curved lines), which would accommodate at least two SSBs (>12 nt), drives the separation of two tetramers in the HaLEF-3 octamers to expose binding sites. The tetrameric rings then turn into a spiral core to bind onto ssDNA. In the initialization step, only one spiral core is formed on 'short' ssDNA, while there are multiple spiral cores on 'long' ssDNA. In the extension step, with more HaLEF-3 molecules cooperatively binding to ssDNA, straight nucleoprotein filaments are formed for 'short' ssDNA. In contrast, 'long' ssDNAs are not fully coated by HaLEF-3 molecules and the helical conformation is broken at the gap regions, resulting in structural heterogeneity along the filament.

have been extensively studied for RecA-family recombinase (bacterial RecA or eukaryotic Rad51), which replace SSB for homologous recombination (52,53). In our HaLEF-3–ssDNA model (Figure 7), although most ssDNA is well protected in the highly ordered helical architecture, the flexibility in ssDNA–HaLEF-3 filaments may help to expose the binding sites on HaLEF-3 and/or ssDNA molecules for non-SSB molecules, which may initiate their polymerization to displace HaLEF-3 and accomplish their biological function.

In addition to their ability to bind and stabilize transiently exposed ssDNA during genomic metabolism, many SSBs have been found to facilitate the annealing of complementary strands of DNA, such as gp2.5 (54,55) and ICP8 (56). ICP8 binds to ssDNA to form regular helical nucleoprotein filaments or ring-shaped particles, some of which have been proposed to be intermediate states for the ICP8-catalyzed annealing reactions (57–59). Formation of an SSB-coated ssDNA filament usually results in fewer secondary structures in ssDNA, which is generally thought to promote annealing activity. A previous study has reported that, when LEF-3 was undersaturated compared with ssDNA, LEF-3 also promoted the annealing of complementary ssDNA (60). In this case, more SSB-free gap regions on ssDNA could provide more sites for the initiation of annealing reactions. With the extension of annealing product, SSB with lower affinity for dsDNA would then dissociate from the newly formed dsDNA. Interestingly, studies have also found that oxidized or thermally unfolded LEF-3 stimulated the annealing of complementary ssDNA (27,60). One likely explanation for this is that disruption of the highly ordered helical conformation by non-specific cross-linking of SSB or mild thermal treatment of SSB, even partially, could expose more active ssDNA regions for initiating annealing reactions. It will also be interesting to investigate whether the helical conformation of HaLEF-3–ssDNA would contribute to the annealing activity like that for ICP8–ssDNA filaments.

In summary, HaLEF-3 shows unique tertiary and quaternary architecture compared with previously reported SSB

structures. Within flexible HaLEF-3–ssDNA nucleoprotein filaments, there are stable helical stretches desirable for high-resolution cryo-EM characterization. On the basis of structural analysis, the three OB folds of HaLEF-3 play different roles in forming SSB–ssDNA complexes: intermolecular OB1–OB3 interactions are a major contributor of cooperative interactions, whereas OB2 and OB3 function as a rigid domain that is responsible for non-specific SSB–ssDNA interactions. In our proposed assembly model, cooperative binding of HaLEF-3 onto ssDNA may occur at local regions, and SSB-free gap regions can introduce significant flexibility to the long SSB–ssDNA filaments. The ring-to-helix transition of HaLEF-3 is important for initiation of the cooperative binding of SSB onto ssDNA. HaLEF-3 may serve as a new SSB model system to dissect the dynamic mechanism of SSB–ssDNA and SSB–SSB interactions, through further structure-based mutagenesis analysis.

DATA AVAILABILITY

The atomic coordinates and structure factors for the nucleotide-free HaLEF-3 crystal structure have been deposited in the Protein Data Bank under accession number 7YNY. Cryo-EM density maps have been deposited in the Electron Microscopy Data Bank with accession codes EMD-34010 (helical reconstruction of HaLEF-3–dA_{140}) and EMD-34011 (local reconstruction), and the corresponding atomic models have been deposited in the Protein Data Bank under accession codes 7YPO and 7YPQ.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the staff of the BL19U1 beamline at the National Facility for Protein Science (NFPS) and Shanghai Synchrotron Radiation Facility (SSRF) for assistance in data collection. We also thank the Center for Instrumental Analysis and Metrology of Wuhan Institute of Virology for providing technical assistance.

FUNDING

This work was supported by the National Key Research and Development Program of China [2018YFA0507200] and the Key Research Projects of Frontier Science, Chinese Academy of Sciences [QYZDJ-SSW-SMC021]. Funding for open access charge: National Key Research and Development Program of China [2018YFA0507200].
Conflict of interest statement. None declared.

REFERENCES

1. Kowalczykowski,S.C., Bear,D.G. and Von Hippel,P.H. (1981) In: Boyer,P.D. (ed). *The Enzymes*. Academic Press, Inc., NY, Vol. **14**, pp. 373–444.
2. Bianco,P.R. (2017) The tale of SSB. *Prog. Biophys. Mol. Biol.*, **127**, 111–118.
3. Antony,E. and Lohman,T.M. (2019) Dynamics of *E. coli* single stranded DNA binding (SSB) protein–DNA complexes. *Semin. Cell Dev. Biol.*, **86**, 102–111.
4. Byrne,B.M. and Oakley,G.G. (2019) Replication protein A, the laxative that keeps DNA regular: the importance of RPA phosphorylation in maintaining genome stability. *Semin. Cell Dev. Biol.*, **86**, 112–120.
5. Shamoo,Y., Friedman,A.M., Parsons,M.R., Konigsberg,W.H. and Steitz,T.A. (1995) Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature*, **376**, 362–366.
6. Hollis,T., Stattel,J.M., Walther,D.S., Richardson,C.C. and Ellenberger,T. (2001) Structure of the gene 2.5 protein, a single-stranded DNA binding protein encoded by bacteriophage T7. *Proc. Natl Acad. Sci. USA*, **98**, 9557–9562.
7. Mapelli,M., Panjikar,S. and Tucker,P.A. (2005) The crystal structure of the herpes simplex virus 1 ssDNA-binding protein suggests the structural basis for flexible, cooperative single-stranded DNA binding. *J. Biol. Chem.*, **280**, 2990–2997.
8. Arcus,V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr. Opin. Struct. Biol.*, **12**, 794–801.
9. Murzin,A.G. (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, **12**, 861–867.
10. Theobald,D.L., Mitton-Fry,R.M. and Wuttke,D.S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **32**, 115–133.
11. Raghunathan,S., Ricard,C.S., Lohman,T.M. and Waksman,G. (1997) Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength x-ray diffraction on the selenomethionyl protein at 2.9-A resolution. *Proc. Natl Acad. Sci. USA*, **94**, 6652–6657.
12. Raghunathan,S., Kozlov,A.G., Lohman,T.M. and Waksman,G. (2000) Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat. Struct. Biol.*, **7**, 648–652.
13. Bochkareva,E., Korolev,S., Lees-Miller,S.P. and Bochkarev,A. (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J.*, **21**, 1855–1863.
14. Kim,C., Paulus,B.F. and Wold,M.S. (1994) Interactions of human replication protein A with oligonucleotides. *Biochemistry*, **33**, 14197–14206.
15. Kim,Y.T., Tabor,S., Bortner,C., Griffith,J.D. and Richardson,C.C. (1992) Purification and characterization of the bacteriophage T7 gene 2.5 protein. A single-stranded DNA-binding protein. *J. Biol. Chem.*, **267**, 15022–15031.
16. Ruyechan,W.T. and Wetmur,J.G. (1975) Studies on the cooperative binding of the *Escherichia coli* DNA unwinding protein to single-stranded DNA. *Biochemistry*, **14**, 5529–5534.
17. Kowalczykowski,S.C., Lonberg,N., Newport,J.W. and Hippel,P.H. (1981) Interactions of bacteriophage T4-coded gene 32 protein with nucleic acids. I. Characterization of the binding interactions. *J. Mol. Biol.*, **145**, 75–104.
18. Chrysogelos,S. and Griffith,J. (1982) *Escherichia coli* single-strand binding protein organizes single-stranded DNA in nucleosome-like units. *Proc. Natl Acad. Sci. USA*, **79**, 5803–5807.
19. Griffith,J.D., Harris,L.D. and Register,J. 3rd (1984) Visualization of SSB–ssDNA complexes active in the assembly of stable RecA–DNA filaments. *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 553–559.
20. Alani,E., Thresher,R., Griffith,J.D. and Kolodner,R.D. (1992) Characterization of DNA-binding and strand-exchange stimulation properties of y-RPA, a yeast single-strand-DNA-binding protein. *J. Mol. Biol.*, **227**, 54–71.
21. Hamon,L., Pastre,D., Dupaigne,P., Le Breton,C., Le Cam,E. and Pietrement,O. (2007) High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein–DNA complexes. *Nucleic Acids Res.*, **35**, e58.
22. Dubiel,K., Myers,A.R., Kozlov,A.G., Yang,O., Zhang,J., Ha,T., Lohman,T.M. and Keck,J.L. (2019) Structural mechanisms of cooperative DNA binding by bacterial single-stranded DNA-binding proteins. *J. Mol. Biol.*, **431**, 178–195.
23. Yates,L.A., Aramayo,R.J., Pokhrel,N., Caldwell,C.C., Kaplan,J.A., Perera,R.L., Spies,M., Antony,E. and Zhang,X. (2018) A structural and dynamic model for the assembly of replication protein A on single-stranded DNA. *Nat. Commun.*, **9**, 5447.
24. Hang,X., Dong,W. and Guarino,L.A. (1995) The *lef-3* gene of *Autographa californica* nuclear polyhedrosis virus encodes a single-stranded DNA-binding protein. *J. Virol.*, **69**, 3924–3928.

25. Garavaglia,M.J., Miele,S.A.B., Iserte,J.A., Belaich,M.N. and Ghiringhelli,P.D. (2012) The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family baculoviridae. *J. Virol.*, **86**, 12069–12079.
26. Okano,K., Mikhailov,V.S. and Maeda,S. (1999) Colocalization of baculovirus IE-1 and two DNA-binding proteins, DBP and LEF-3, to viral replication factories. *J. Virol.*, **73**, 110–119.
27. Mikhailov,V.S., Okano,K. and Rohrmann,G.F. (2006) Structural and functional analysis of the baculovirus single-stranded DNA-binding protein LEF-3. *Virology*, **346**, 469–478.
28. Wu,Y. and Carstens,E.B. (1998) A baculovirus single-stranded DNA binding protein, LEF-3, mediates the nuclear localization of the putative helicase P143. *Virology*, **247**, 32–40.
29. Mikhailov,V.S., Okano,K. and Rohrmann,G.F. (2003) Baculovirus alkaline nuclease possesses a 5′→3′ exonuclease activity and associates with the DNA-binding protein LEF-3. *J. Virol.*, **77**, 2436–2444.
30. Kabsch,W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
31. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
32. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
33. Chen,V.B., Arendall,W.B. 3rd, Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
34. Mastronarde,D.N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.*, **152**, 36–51.
35. Zivanov,J., Nakane,T., Forsberg,B.O., Kimanius,D., Hagen,W.J., Lindahl,E. and Scheres,S.H. (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife*, **7**, e42166.
36. Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
37. Keating,K.S. and Pyle,A.M. (2012) RCrane: semi-automated RNA model building. *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 985–995.
38. Afonine,P.V., Poon,B.K., Read,R.J., Sobolev,O.V., Terwilliger,T.C., Urzhumtsev,A. and Adams,P.D. (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.*, **74**, 531–544.
39. Goddard,T.D., Huang,C.C., Meng,E.C., Pettersen,E.F., Couch,G.S., Morris,J.H. and Ferrin,T.E. (2018) UCSF chimerax: meeting modern challenges in visualization and analysis. *Protein Sci.*, **27**, 14–25.
40. Chen,Z. and Carstens,E.B. (2005) Identification of domains in *Autographa californica* multiple nucleopolyhedrovirus late expression factor 3 required for nuclear transport of P143. *J. Virol.*, **79**, 10915–10922.
41. Holm,L. and Rosenstrom,P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.*, **38**, W545–W549.
42. Zeng,Z., Min,B., Huang,J., Hong,K., Yang,Y., Collins,K. and Lei,M. (2011) Structural basis for *Tetrahymena* telomerase processivity factor Teb1 binding to single-stranded telomeric-repeat DNA. *Proc. Natl Acad. Sci. USA*, **108**, 20357–20361.
43. Classen,S., Ruggles,J.A. and Schultz,S.C. (2001) Crystal structure of the N-terminal domain of *Oxytricha nova* telomere end-binding protein alpha subunit both uncomplexed and complexed with telomeric ssDNA. *J. Mol. Biol.*, **314**, 1113–1125.
44. Williams,T.L. and Levy,D.L. (2013) Assaying cooperativity of protein–DNA interactions using agarose gel electrophoresis. *Methods Mol. Biol.*, **1054**, 253–265.
45. Lohman,T.M., Overman,L.B. and Datta,S. (1986) Salt-dependent changes in the DNA binding co-operativity of *Escherichia coli* single strand binding protein. *J. Mol. Biol.*, **187**, 603–615.
46. Kozlov,A.G., Weiland,E., Mittal,A., Waldman,V., Antony,E., Fazio,N., Pappu,R.V. and Lohman,T.M. (2015) Intrinsically disordered C-terminal tails of *E. coli* single-stranded DNA binding protein regulate cooperative binding to single-stranded DNA. *J. Mol. Biol.*, **427**, 763–774.
47. Mapelli,M., Muhleisen,M., Persico,G., van Der Zandt,H. and Tucker,P.A. (2000) The 60-residue C-terminal region of the single-stranded DNA binding protein of herpes simplex virus type 1 is required for cooperative DNA binding. *J. Virol.*, **74**, 8812–8822.
48. Zhou,R.B., Kozlov,A.G., Roy,R., Zhang,J.C., Korolev,S., Lohman,T.M. and Ha,T. (2011) SSB functions as a sliding platform that migrates on DNA via reptation. *Cell*, **146**, 222–232.
49. Roy,R., Kozlov,A.G., Lohman,T.M. and Ha,T. (2009) SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature*, **461**, 1092–1097.
50. Nguyen,B., Sokoloski,J., Galletto,R., Elson,E.L., Wold,M.S. and Lohman,T.M. (2014) Diffusion of human replication protein A along single-stranded DNA. *J. Mol. Biol.*, **426**, 3246–3261.
51. Mishra,G., Bigman,L.S. and Levy,Y. (2020) ssDNA diffuses along replication protein A via a reptation mechanism. *Nucleic Acids Res.*, **48**, 1701–1714.
52. Bell,J.C., Plank,J.L., Dombrowski,C.C. and Kowalczykowski,S.C. (2012) Direct imaging of RecA nucleation and growth on single molecules of SSB-coated ssDNA. *Nature*, **491**, 274–278.
53. Short,J.M., Liu,Y., Chen,S., Soni,N., Madhusudhan,M.S., Shivji,M.K. and Venkitaraman,A.R. (2016) High-resolution structure of the presynaptic RAD51 filament on single-stranded DNA by electron cryo-microscopy. *Nucleic Acids Res.*, **44**, 9017–9030.
54. Rezende,L.F., Willcox,S., Griffith,J.D. and Richardson,C.C. (2003) A single-stranded DNA-binding protein of bacteriophage T7 defective in DNA annealing. *J. Biol. Chem.*, **278**, 29098–29105.
55. Hernandez,A.J. and Richardson,C.C. (2019) Gp2.5, the multifunctional bacteriophage T7 single-stranded DNA binding protein. *Semin. Cell Dev. Biol.*, **86**, 92–101.
56. Weerasooriya,S., DiScipio,K.A., Darwish,A.S., Bai,P. and Weller,S.K. (2019) Herpes simplex virus 1 ICP8 mutant lacking annealing activity is deficient for viral DNA replication. *Proc. Natl Acad. Sci. USA*, **116**, 1033–1042.
57. Makhov,A.M., Boehmer,P.E., Lehman,I.R. and Griffith,J.D. (1996) Visualization of the unwinding of long DNA chains by the herpes simplex virus type 1 UL9 protein and ICP8. *J. Mol. Biol.*, **258**, 789–799.
58. Makhov,A.M. and Griffith,J.D. (2006) Visualization of the annealing of complementary single-stranded DNA catalyzed by the herpes simplex virus type 1 ICP8 SSB/recombinase. *J. Mol. Biol.*, **355**, 911–922.
59. Tolun,G., Makhov,A.M., Ludtke,S.J. and Griffith,J.D. (2013) Details of ssDNA annealing revealed by an HSV-1 ICP8–ssDNA binary complex. *Nucleic Acids Res.*, **41**, 5927–5937.
60. Mikhailov,V.S., Okano,K. and Rohrmann,G.F. (2005) The redox state of the baculovirus single-stranded DNA-binding protein LEF-3 regulates its DNA binding, unwinding, and annealing activities. *J. Biol. Chem.*, **280**, 29444–29453.