Assembly of the Bacteriophage T4 Helicase

ARCHITECTURE AND STOICHIOMETRY OF THE \text{gp41-gp59}\ COMPLEX

Faoud T. Ishmael, Stephen C. Alley\textsuperscript{‡}, and Stephen J. Benkovic\textsuperscript{§}

From the Department of Chemistry, Pennsylvania State University, University Park, Pennsylvania 16802

DNA replication is a multistaged reaction that requires the recruitment of proteins, assembly of replication forks at an origin, and template-directed replication of DNA. The replication machinery in the T4 bacteriophage is composed of three units, the primosome (a complex of the primase (gp61) bound to the helicase (gp41)), the holoenzyme (composed of a DNA polymerase (gp43), a sliding clamp (gp45), and a clamp loader (gp44/62)), and the single-stranded binding protein gp32.

An initial event in the replication process is the recruitment and assembly of the primosome at the DNA replication fork. Formation of the primosome on the lagging strand most likely begins with assembly of the helicase. The bacteriophage T4 helicase plays an essential role in DNA replication and recombination during phage infection of Escherichia coli. The helicase is needed to unwind duplex DNA ahead of the holoenzyme to allow nucleotide incorporation during replication and is required to unwind duplex DNA for pairing of single-strand DNA with new partners in recombination (1). gp41 forms a hexameric ring that encircles the lagging strand of the replication fork and unwinds the duplex in a 5'–3' direction by hydrolyzing ATP (1). During DNA replication, the gp41 helicase is intimately associated with the gp61 primase, which synthesizes the RNA primers used in discontinuous lagging strand synthesis. The formation of this primosome is needed for efficient unwinding of duplex DNA during replication as well as productive primer synthesis (2, 3).

Loading of the gp41 replicative helicase at the bacteriophage T4 DNA replication fork requires the displacement of single-stranded DNA-binding proteins (gp32) coating the lagging strand. The helicase assembly protein, gp59, is required to load the helicase under these conditions, thus assuming a critical role in bacteriophage T4 DNA replication as well as recombination (4–6).

The gp59 protein is a 26-kDa, basic (pI = 10.18) protein that exists as a monomer in solution (7, 8). It has a high affinity for DNA and can bind duplex DNA, ssDNA,\textsuperscript{1} and forked DNA substrates (7, 9). The protein is composed of almost all \( \alpha \)-helices and contains a C-terminal and an N-terminal domain (10). Mueser et al. (10) have shown that gp59 binds to forked DNA substrates with higher affinity than ssDNA. Furthermore, it has been proposed that duplex DNA binds to the N-terminal domain, whereas ssDNA binds to the C-terminal domain based on models generated from the gp59 crystal structure and its similarity to high mobility group proteins (10).

gp59 acts via direct contacts with gp41 and has been shown to interact with the helicase both on and off DNA (8, 11). gp59 also binds to gp32 in both the presence and absence of DNA (8, 12). Morrical et al. (13) have suggested that the C-terminal, acidic domain (A-domain) of gp59 mediates the interaction between the two proteins. More recent studies have demonstrated that the DNA binding core domain of gp59 also interacts with gp59 (14), and the combination of the two sites of interaction may be necessary for gp32 displacement. Furthermore, gp59 contains distinct sites for gp32 and gp41 binding (11). As such, the gp59-mediated assembly of gp41 onto gp32-coated DNA most likely involves a ternary complex between the three proteins.

Although gp59 exists as a monomer in solution, it can oligomerize in the presence of gp32 or DNA. Cross-linking experiments demonstrated that gp59 could form at least pentamers

\textsuperscript{1} The abbreviations used are: ssDNA, single-stranded DNA; DTT, dithiothreitol; gp41c, gp41 with a cysteine residue added to the C terminus (Cys-476); DMF, dimethylformamide; SA-HRP, streptavidin linked to horseradish peroxidase; Sulfo-SBED, sulfosuccinimide-2-(6-amino-4-methylcoumarin-3-acetamido) hexanoamide[ethyl-1,3-di-thiopropionate]; BMH, 1,6 bis-maleimidohexane; AMCA, N-6-(2,5-di-thiophosphoryl)-1,6-bis-maleimidohexane; AMCA-HPPD, \( N \)-[6-(7-maleimidoethyl)amino; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; MALDI, matrix-assisted laser desorption ionization; HPLC, high pressure liquid chromatography; ATP\( \gamma \)-S, adenosine 5'-O-(thiotriphosphate).
in the presence of either of these molecules (14). Furthermore, Raney et al. (15) showed that the enhancement of the ATP-dependent DNA unwinding activity of gp41 was maximal when gp59 was present in a 1:1 molar ratio with gp41 (hexamer of gp59 to hexamer of gp41). As such, gp59 is expected to form higher multimeric states when bound to the helicase.

In *E. coli*, the DnAB helicase is loaded onto DNA by a protein analogous to gp59, DnaC. Cryoelectron microscopy studies indicated that DnaC forms a hexameric ring associated with the hexameric DnaB ring (16). By analogy, we predict a similar oligomeric state and arrangement of gp59 and gp41.

To gain a better understanding of the mechanism of helicase assembly, the architecture and stoichiometry of the gp41-gp59 complex was investigated. We have demonstrated that 1) both the N and C terminus of gp41 lie in the gp41-gp59 subunit interface and interact with gp59; 2) the site of interaction of gp41 on gp59 is close to Cys-215 of gp59; 3) conformational changes in the gp41-gp59 complex occur upon gp41 loading on DNA; and 4) gp59 forms a hexamer upon binding gp41, and the gp59 subunits are arranged in a head to head orientation.

The importance of these findings is discussed herein.

**Experimental Procedures**

**Cloning and Purification of gp41**—The 41 gene was isolated from bacteriophage T4 genomic DNA (Sigma) by PCR amplification using the primers 5’-GGCGAACATATGTTGAAATATTCTTCTT and 5’-GGCG- GAATTGCTCTTCCGCAAAATTTTAATTC. The product was digested with *NdeI* and *SapI* and ligated into a custom IMPACT vector placed under the control of a T7 promoter. This vector places a self-cleaving site at the 3’ end of gp41 which is cleaved from the chitin column with 100 mM cysteine dissolved in 20 mM Tris, pH 7.5, 1 M NaCl to remove uncleaved peptides. The labeled peptide bound to the column by virtue of the Sulfo-SBED label was then treated with 1 M 2-mercaptoethanol for 30 min to cleave the disulfide bond in the cross-linker, thus eluting the peptide. The eluted peptide was dried in a speed vacuum, and the sample was desalted with a zip-tip pipette tip containing C-18 resin (Millipore Corp.) according to the manufacturer’s instructions. The peptide was eluted from the zip-tip C-18 resin using 60% acetonitrile, 40% H2O, 0.1% trifluoroacetic acid. The isolated peptide (500 μM) was then mixed with 500 μM of α-cyano-4-hydroxy cinnamic acid (15 mg/ml dissolved in 50% acetic trifle, 50% water, 0.1% trifluoroacetic acid), and the mixture was spotted and air-dried on a MALDI target plate. The sample was analyzed by a Voyager MALDI mass spectrometer (Perseptive) in linear mode to confirm that only the N terminus was labeled. Positive ions were detected using a delayed extraction time of 300 ns and 250 laser scans.

**Thiol-Thiol Cross-linking between gp41c and gp59**—gp41c was labeled at the N terminus with biotin succinimide (to allow visualization by Western blot using SA-HRP) and mixed with gp59 (to a final concentration of 1 μM) in the presence of 10 μM 1,6-bis-maleimidohexane (BMH; Pierce) with or without 3 mM ATP and/or 5 mM M13mp18 DNA. After 7 min, the reaction was quenched with 5 μl of gel loading buffer containing 60 mM Tris, pH 6.8, 25% glycerol, 2% SDS, 14.4 mM 2-mercaptoethanol, and 0.1% bromphenol blue. The products were separated by 10% polyacrylamide gel and subjected to SDS-PAGE. The gel was stained with Gel Code Blue (Pierce).

**Mapping the Point of Contact between gp41c and gp59**—Thiol-thiol cross-linking between gp41c and gp59 could occur between the C-terminal cysteine on gp41c and one of the two cysteine residues on gp59 (Cys-42 or Cys-215). It was predicted that the site of cross-link on gp59 is Cys-215. To determine whether cross-linking occurred between the C-terminus of gp41c and gp59, gp41c was cross-linked to investigate higher oligomer formation by mixing a C-terminal cysteine added (Cys-476, denoted as gp41c) with gp59, which was cleaved from the chitin column with 100 mM M13 cleavage buffer containing 60 mM Tris, pH 8.0, 500 mM NaCl, 10% glycerol, 1 μM EDTA. The cells were lysed by sonication and centrifuged for 30 min at 20,000 × g, and the supernatant was applied to an 10-ml chitin column. The column was washed with 1 liter of chitin column buffer, and gp41c was cleaved from the column by the addition of 100 mM DTT dissolved in chitin column buffer for 24 h. The protein was eluted from the chitin column and subjected to MALDI mass spectrometry to confirm that the N-terminal cysteine was present as an intact protein. The protein was concentrated to ~10 μM with an Amicon stirred cell concentrator equipped with a YM30 membrane. For purification of gp41 with a C-terminal cysteine added (Cys-476, denoted as gp41c), the protein was cleaved from the chitin column with 100 mM M13 cleavage buffer in the presence and absence of 10 μM 1,6-bis-maleimidohexane (BMH; Pierce) with or without 3 mM ATP and/or 5 mM M13mp18 DNA. After 7 min, the reaction was quenched with 5 μl of gel loading buffer containing 60 mM Tris, pH 6.8, 25% glycerol, 2% SDS, 14.4 mM 2-mercaptoethanol, and 0.1% bromphenol blue. The products were separated by 10% polyacrylamide gel and subjected to Western blot analysis using SA-HRP as a probe.
of label was confirmed by digesting labeled protein with trypsin, separating the fragments by HPLC (14), and identifying the labeled peptide by MALDI mass spectrometry, as described above.

Isothermal Titration Calorimetry of gp41 and gp59—Binding of gp41 to gp59 was assessed by isothermal titration calorimetry (ITC) on a MicroCal VP-ITC system (according to the protocol published previously) (18). Proteins of interest were dialyzed into complex buffer (20 mM Tris, pH 7.5, 150 mM KCl, 10 mM Mg(OAc)₂) for at least 8 h to reach dialysis equilibrium. Protein solutions were centrifuged to remove particles before beginning the experiment and quantitated by UV-visible absorbance. All solutions were degassed prior to use. This degassing step was determined in control experiments not to affect protein activity.

gp59 (21 μM) was in the syringe and injected in 7-μl volumes into a 2.2 μM solution of gp41 containing 1 mM ATP/γS. Data were analyzed by Origin 5.0 (MicroCal). Heats of dilution were subtracted from all data before fitting. Heats of dilution were determined in a separate control experiment by injecting the protein into buffer. Data analysis requires a binding model, and fitting returns values for the binding constant K, ΔH of binding, ΔS of binding, and the binding stoichiometry N.

Oligomerization of gp41 in the Presence of gp59—To determine whether gp59 induced higher oligomer formation of gp41, the two proteins were cross-linked in the absence or presence of ATP. A 100-μl solution of gp59 (10 μM) was labeled with biotin by the following procedure. Biotin-maleimide (Sigma) was dissolved in DMF to make a 3,3′-dithio-bis(propanoic acid N-hydroxysuccinimide ester (Sigma)) was added to each sample for 10 min. The products were separated by SDS-PAGE, and gp41 cross-link bands were detected with a Fluor-S multiimager, whereas the gp59 cross-link bands were detected with SA-HRP as a probe. gp41 cross-linked in the presence of ATP was used as a control to assess the size of the cross-linked species.

Mapping the Point of Contact between gp59 Subunits—The fluorescently labeled thiol-thiol cross-linker shown in Fig. 1C was synthesized from the trifunctional maleimide cross-linker tris(2-maleimidoethyl)amine (TEMA; Pierce) and the fluorescent probe AMCA-HDP. AMCA-HDP dissolved in DMF was reduced with 3 equivalents of tris(2-carboxylethyl)phosphine hydrochloride (TCEP; Molecular Probes) and HPDP dissolved in DMF was reduced with 3 equivalents of tris(2-maleimidoethyl)amine. The two labeled proteins were mixed to a final concentration of 4 μM each in the presence or absence of 1 mM ATP. A solution of 0.25% (final concentration) 3,3′-dithio-bis(propanoic acid N-hydroxysuccinimide ester (Sigma)) was added to each sample for 10 min. The products were separated by SDS-PAGE, and gp41 cross-link bands were detected with a Fluor-S multiimager, whereas the gp59 cross-link bands were detected with a Western blot procedure using SA-HRP as a probe. gp41 cross-linked in the presence of ATP was used as a control to assess the size of the cross-linked species.

RESULTS

Labeling of gp41c at the C Terminus and Cross-linking to gp59—gp41 contains 4 cysteine residues (amino acids 208, 233, 316, and 319). A DTNB assay determined that one of these cysteines is solvent-accessible, but labeling at this site with a cross-linker yielded no cross-links to gp59, suggesting that it is not close to the interface of interaction between the two proteins.² To determine whether the C terminus of gp41c interacted with gp59, cross-linking between the C terminus of gp41c and gp59 was conducted. A cysteine residue was added to the C terminus of gp41 (Cys-476, with the modified protein denoted as gp41c) by cleavage of a gp41-intein fusion protein with cysteine. The protein was then labeled with the thiol-reactive trifunctional cross-linker shown in Fig. 1A. The labeled gp41c was subsequently mixed with gp59 in the presence or absence of DNA and/or ATP and subjected to ultraviolet light to initiate cross-linking. When the products of the cross-linking reaction were separated by SDS-PAGE and subjected to a Western blot using SA-HRP as a probe, cross-link bands corresponding to gp41c dimers were seen in the absence of gp59, signifying that the C terminus of gp41c is close to or in the homodimeric interface (Fig. 2A, lane 2). When gp41c was mixed with gp59 and exposed to light, in the presence or absence of ATP and/or DNA, high molecular weight species were seen, consistent with gp41c-gp59 cross-links (Fig. 2A, lanes 3–6). Upon the addition of DTT, the cross-link was reduced and the biotin label was transferred to gp59, confirming that a gp41c-gp59 cross-link had occurred (Fig. 2A, lanes 9–12).

Cross-linking between the N Terminus of gp41 and gp59—in order to elucidate a second point of contact between gp41 and gp59, the N terminus of gp41 was labeled as described under “Experimental Procedures” with the amine-reactive trifunctional cross-linker Sulfo-SBED (Fig. 1B). The proteins were mixed in the absence or presence of DNA and/or ATP and then separated by SDS-PAGE followed by a Western blot using SA-HRP as a probe to identify the biotinylated cross-links. When gp41c was cross-linked by itself, a band corresponding to a gp41c-gp41 homodimer was observed (Fig. 2B, lane 2). When gp41 and gp59 were mixed, cross-links were visualized in both

---

² F. T. Ishmael, S. C. Alley, and S. J. Benkovic, unpublished data.
the absence and presence of ATP and/or DNA (Fig. 2B, lanes 3–6). Reduction of the cross-link with DTT caused transfer of the biotin to gp59, validating that the cross-link occurred between gp41 and gp59 (Fig. 2B, lanes 7 and 8).

Thiol-Thiol Cross-linking between gp41c and gp59—Morris et al. (13) suggested that gp59 has distinct binding sites for gp41c and gp32. Experiments described previously demonstrated that Cys-42 of gp59 is in close proximity to gp32 (14). gp59 contains a second cysteine residue, Cys-215, that lies on the opposite side of the protein relative to Cys-42. Since it is likely that gp41 interacts on this side of the protein, away from the gp32 binding site, thiol-thiol cross-linking was attempted to cross-link the C terminus of gp41c to gp59. gp41c was N-terminally labeled with biotin to allow easier detection of cross-links by Western blot using SA-HRP as a probe. When mixed with the thiol-thiol cross-linker BMH, gp41c cross-linked to form a homodimer linked at the C-terminal cysteine residues (Fig. 3A). In the presence of gp59, with or without ATP added, bands were seen that correspond to gp41c-gp59 cross-links (Fig. 3A). Typically, two cross-link bands were observed, one that migrated with the expected mobility of a gp41c-gp59 cross-link (molecular mass of about 75 kDa) and a band that migrated with a faster mobility (65 kDa). The intensity of the higher mobility band varies from experiment to experiment, whereas the lower mobility band does not vary. In experiments using a cleavable thiol-thiol cross-linker, this pattern of cross-linking between gp41c and gp59 is also seen. Excision of these bands followed by cleavage of the cross-linker and separation of the products on a second gel indicate that both bands are composed of gp41c-gp59 cross-links. Most likely, the faster migrating band is due to an intramolecular cross-link of gp41c as well as an intermolecular cross-link to gp59, resulting in higher mobility of the more compact structure than the gp41c-gp59 intermolecular cross-link alone. When DNA was added, no cross-links between gp41c and gp59 were visualized, indicating that some conformational change occurs upon loading of gp41 on DNA (Fig. 3A, lanes 1 and 2).

To confirm that gp41c cross-linked to Cys-215 of gp59, a cross-linking experiment was conducted in which the Cys-42 of gp59 was blocked by covalent modification. Previously, we showed that incubation of gp52 labeled with a probe via a mixed disulfide with gp59 resulted in complete transfer of the probe from gp32 to gp59 by a disulfide exchange (Fig. 4A) (14). As described under “Experimental Procedures,” a label (AMCA) was placed selectively on Cys-42 of gp59 using this method. The modified gp59 protein contained only one free cysteine residue, Cys-215, as determined by MALDI mass spectrometry. Thiol-thiol cross-linking using BMH showed that gp41c cross-linked to both the modified (Fig. 4B, lanes 2 and 3) gp59 as well as the unmodified gp59 (Fig. 4B, lanes 6 and 7), demonstrating that Cys-42 was not involved in the cross-linking.
The interaction between gp41c and gp59 was also analyzed in the presence of gp32. Preincubation of gp59 with gp32 prior to the addition of gp41c led to the same pattern of cross-linking observed above (Fig. 3B), indicating that gp32 bound to gp59 does not affect gp41 binding. As before, cross-links between gp41 and gp59 disappear in the presence of DNA. Experiments conducted by Morrical et al. (13) indicated that under similar conditions to those used for the cross-linking, gp32, gp59, and gp41 (in the absence of DNA) form a ternary complex. As such, it is expected that gp41 does not displace gp32 in these experiments but rather binds independently. In a separate experiment in which gp32 labeled with a biotin handle was preincubated with gp59, mixed with gp41, and then subjected to cross-linking (Fig. 3C), cross-links were observed between gp32 and gp59, indicating that gp41 did not affect gp32-gp59 binding. These data support a model in which gp32 and gp41 bind to distinct sites on gp59, such that gp32 interacts at a point on gp59 near Cys-42, while gp41 interacts with gp59 on the opposite side, close to Cys-215 (Fig. 4C).

**The Oligomeric State of gp59 in the Presence of gp41**—In the presence of gp32 or DNA, gp59 forms a multimer composed of at least five subunits (14). Moreover, kinetic experiments by Raney et al. (15) indicated that gp41 and gp59 formed a 1:1 complex (six subunits of gp59 and six subunits of gp41). Thiol-thiol cross-linking was performed to investigate the oligomeric state of gp59 in the presence of gp41. gp59 was subjected to cross-linking with BMH in the presence of gp41 (without a C-terminal cysteine added) in the presence and absence of DNA and/or ATP. Since a cysteine residue placed at the C terminus of gp41 is required for gp41-gp59 thiol-thiol cross-linking, only gp59-gp59 cross-links are observed. Cross-links were seen that correspond to higher multimeric states of gp59 (Fig. 5A) regardless of the presence of ATP or DNA. The largest observable species migrated as a gp59 pentamer. It is likely that the protein forms oligomers higher than pentamers, but limitations on the yield of cross-linking and resolution by gel electrophoresis make observation of this multimeric state difficult.

The association of gp41 and gp59 was further examined by isothermal titration calorimetry (ITC). ITC can provide accurate measures of stoichiometry, binding constant, and enthalpy and entropy of binding, making it an excellent technique for the detection of macromolecular interactions. The direct interaction of the untagged, unmodified molecules is monitored. In the ITC experiment, gp59 was titrated into a gp41 solution in the presence of ATP or DNA. The data were fit to a one-site model, and the estimated K_d for the process was 450 ± 120 nM, with a stoichiometry of 0.95 ± 0.06 gp59 hexamers/1 gp41 hexamer. C, the oligomerization of gp41 was investigated in the presence of gp59. gp59 was labeled with biotin, and gp41 was labeled with rhodamine to determine complexes involving each protein by using Western blot with SA-HRP as a probe and fluorescence, respectively. Lanes 1–5 contain gp59, gp41, and DSS, visualized by Western blot (lane 1); gp59, gp41, and DSS visualized by fluorescence (lane 2); gp41, ATP, and DSS (lane 3); gp41 and DSS (lane 4); and gp41 (lane 5).
with respect to a reference cell. As the binding sites are saturated, each injection peak becomes smaller until the heat reflects only dilution of the titrant molecule. An independent control experiment with only buffer in the cell provides a measure of this background. The data in this experiment were fit to a model describing a single set of identical binding sites (single site model). Binding of gp41 to gp59 is an exothermic process, with a dissociation constant of 450 ± 120 nM (Fig. 5B). The stoichiometry of gp59 to gp41 was calculated to be 0.95:1 ± 0.06, indicating that a hexamer of gp59 bound to the gp41 hexamer.

Oligomerization of gp41 in the Presence of gp59—gp41 was labeled with rhodamine and gp59 was labeled with biotin to distinguish between cross-links involving gp41 and those with gp59. The proteins were then cross-linked in the absence or presence of ATP to determine the effect of gp59 on the oligomerization of gp41. In the absence of ATP and gp59, only a minimal number of gp41 higher order cross-links are seen (Fig. 5C, lane 4). When ATP is added, the majority of gp41 is cross-linked to hexamer (Fig. 5C, lane 3). In the absence of ATP but in the presence of gp59, hexameric cross-links of gp41 are also observed, although not to the same extent as in the presence of ATP (Fig. 5C, lane 2). These higher order bands are not gp59-gp59 cross-links or gp59-gp41 cross-links, since a Western blot using SA-HRP to visualize biotinylated gp59 protein did not show these high molecular weight bands (Fig. 5C, lane 1).

Mapping the Site of gp59-gp59 Interaction—gp59 contains only two cysteine residues. Cross-linking of gp59 with a thioldiol cross-linker in the presence of gp32, DNA, or gp41 yields higher molecular weight species. To form cross-links of trimers or higher, both cysteine residues must be involved and be close to or in the gp59-gp59 subunit interface. Two simple models can be generated that describe gp59-gp59 oligomer formation. The subunits could be oriented head to head, as shown in Fig. 6A, such that Cys-42 cross-links to Cys-42 and Cys-215 cross-links to Cys-215. Alternatively, the subunits could be in a head to tail orientation, such that Cys-215 cross-links to Cys-42 (Fig. 6A).

A fluorescently labeled thioldiol cross-linker (TMEA-AMCA) was utilized to map protein-protein interactions. TMEA-AMCA (Fig. 1C) was synthesized from the trifunctional thiol-reactive cross-linker TMEA and the fluorescent probe AMCA-HPDP (as described under “Experimental Procedures”). AMCA-HPDP was reduced with the phosphine TCEP to yield the fluorescence probe with a free sulfhydryl group, which was then mixed with TMEA. The mixture was separated by HPLC, and the product was verified by MALDI mass spectrometry (data not shown).

gp59 was cross-linked with this reagent in the presence of DNA. Some of the mixture was separated by SDS-PAGE, and the cross-links were visualized with a Fluor-S multimager. In transilluminator mode, the fluorescence of the probe, AMCA, was detected (Fig. 6B) and demonstrated that higher cross-linked species are present. The remaining solution (not subjected to SDS-PAGE) was digested with trypsin, and the resulting peptides were separated by HPLC. Fractions were collected and placed on the stage of the Fluor-S to determine which contained fluorophore and thus the cross-linked peptides (Fig. 6C). Fraction 17 contained the most fluorescence and was analyzed by MALDI mass spectrometry. Masses of 2550.93 (Fig. 7A) and 3668.38 (Fig. 7B) were observed, consistent with tryptic fragments containing Cys-42 cross-linked to Cys-42 plus the cross-linker (M + H)⁺ = 2550.91 and Cys-215 cross-linked to Cys-215 plus the cross-linker (M + H)⁺ = 3668.20. The fragment containing Cys-215 has two missed cleavages (sequence KSCKY), which is not unexpected, considering that the labeled cysteine residue is so close to the tryptic cleavage.

![Fig. 6. Mapping the point of contact between gp59 subunits. A. possible orientations of the gp59 multimer. The protein subunits could be oriented head to head with Cys-42 in close proximity to Cys-42 and Cys-215 close to Cys-215. Alternatively, the subunits could be arranged “head to tail” in which Cys-42 and Cys-215 are in close proximity. B, gp59 was cross-linked with the fluorescent cross-linker shown in Fig. 1C and separated by SDS-PAGE, and the cross-link bands were visualized with a Fluor-S multimager. C, cross-linked gp59 was digested with trypsin and subjected to HPLC as described under “Experimental Procedures.” Cross-linked peptides were identified by detecting the fluorescent fractions with a Fluor-S multimager.](image)
to Cys-215. However, as the helicase is loaded onto DNA, a conformational change in the gp41-gp59 complex occurs, such that thiol-thiol cross-linking is no longer effective.

The ability of gp59 to interact at the gp41 subunit interface could have important implications for gp41 assembly. The thermodynamics of gp41 hexamer formation has not been well characterized. At physiologic concentrations of gp41 and ATP, the protein may exist in an equilibrium between oligomeric states. Cross-linking experiments indicated that the presence of gp59 caused formation of gp41 hexamers even in the absence of ATP. Binding of gp59 to the gp41 subunit interface could strengthen the association between gp41 subunits, facilitating formation of the hexamer, and promote loading of the helicase.

**Oligomeric State and Subunit Arrangement of gp59**

Previous studies showed that gp59 was capable of self-association on DNA or in the presence of gp32 and could form at least pentameric species (14). Thiol-thiol cross-linking of gp59 in the presence of gp41 demonstrated that conformational changes induced by gp41 allow self-association. The largest observable species migrates as a gp59 pentamer, although higher multimers could be formed. ITC was employed to accurately assess the multimeric state of gp59 in the presence of the helicase and showed that the gp59 bound as a hexamer to the gp41 hexamer.

Given the functional hexameric state of gp41, we speculate that this multimeric state may also be the active conformer of gp59. As such, the architecture of the gp59 complex in the presence of gp41 was investigated.

The site of gp59-gp59 interaction was mapped via a novel methodology involving cross-linking of the protein in the presence of gp41 with a fluorescently labeled cross-linker, digestion of the protein with trypsin, isolation of the cross-linked peptides by virtue of the fluorophore, and identification of the site of cross-linking by MALDI mass spectrometry. The subunits were found to be oriented in a head to head orientation, with Cys-42 on one subunit being in close proximity to Cys-42 in an adjacent subunit and Cys-215 on one subunit being in close proximity to Cys-215 in a second subunit, as represented in the speculative model in Fig. 8B.

**A Model for Helicase Assembly**

A sequential pathway for primosome assembly is outlined in Fig. 9. It is likely that the gp41-gp59 complex forms after gp59 is bound to gp32-coated DNA. The $K_d$ value for a gp59 binding to a gp32-DNA complex was determined to be $-2$ nm (19), whereas the $K_d$ for the binding of gp41 to gp59 in the absence of DNA or gp32 is $\approx 450$ nm. Hence, we favor as the initial step binding of gp59 to gp32-coated DNA.
The predicted binding site for gp41 on gp59 bound to gp32-DNA is close to Cys-215, based on the site of gp32-gp59 interaction and thiol-thiol cross-linking between gp41 and gp59 in the absence of DNA. Thiol-thiol cross-linking between gp41 and gp59 is maintained even when gp32 is preincubated with gp59, consistent with previous results that indicate that gp59 has distinct sites for gp32 and gp41. However, this cross-linking is abolished when DNA is added, although the N and C terminus of gp41 remain in contact with gp59. It is possible that the site of thiol-thiol cross-linking in the absence of DNA represents the initial site of gp41 binding to the gp59-gp32-DNA complex. In the second step, loading of the helicase on DNA is mediated by conformational changes in gp59 and the architecture of the gp41-gp59 complex, resulting in the loss of thiol-thiol cross-linking (which is slow relative to helicase loading), although the proteins remain in contact. gp59 binds to both the A-domain of gp32 and the core domain, potentially destabilizing gp32-DNA interactions to facilitate the removal of gp32 from DNA. The end result is the removal of gp32 in the third step followed by replacement by the helicase.

It is not known if gp59 departs or remains associated with the gp41-DNA complex. Jones et al. (9) demonstrated by a gel shift assay that the gp41-gp59-DNA complex is stable. Furthermore, our cross-linking experiments in which the N and C termini of gp41 were labeled with a trifunctional cross-linker also showed cross-linking between gp41 and gp59 in the presence of DNA, suggesting that the tripartite complex is stable. Once the helicase is assembled, the primase is likely to be loaded next. Binding of the primase to helicase is weak in the absence of DNA (2, 20); thus, loading of the helicase onto DNA is fast relative to helicase loading, although gp41-gp59 complex, resulting in the loss of thiol-thiol cross-linking in the absence of DNA represents the initiation site of gp41 binding to the gp59-gp32-DNA complex. In the second step, loading of the helicase on DNA is mediated by conformational changes in gp59 and the architecture of the gp41-gp59 complex, resulting in the loss of thiol-thiol cross-linking (which is slow relative to helicase loading), although the proteins remain in contact. gp59 binds to both the A-domain of gp32 and the core domain, potentially destabilizing gp32-DNA interactions to facilitate the removal of gp32 from DNA. The end result is the removal of gp32 in the third step followed by replacement by the helicase.

Acknowledgment—We gratefully acknowledge Dr. Ann Valentine for assistance with the isothermal titration calorimetry experiments.