Models for the Binary Complex of Bacteriophage T4 Gp59 Helicase Loading Protein

GP32 SINGLE-STRANDED DNA-BINDING PROTEIN AND TERNARY COMPLEX WITH PSEUDO-Y JUNCTION DNA*

Received for publication, December 13, 2011, and in revised form, March 15, 2012 Published, JBC Papers in Press, April 5, 2012, DOI 10.1074/jbc.M111.333476

Jennifer M. Hinerman1,2, J. David Dignam‡, and Timothy C. Mueser‡1
From the ‡1Department of Chemistry, University of Toledo, College of Natural Sciences and Mathematics, Toledo, Ohio 43606 and the ‡2Department of Biochemistry and Cancer Biology, University of Toledo, College of Medicine, Toledo, Ohio 43614

**Background:** Accessory proteins assist replicative helicases single-stranded binding protected DNA.
**Results:** Small angle x-ray scattering molecular envelopes allows modeling of the gp59-gp32-DNA complex.
**Conclusion:** The core and A-domains of gp32 protein interact with the C-terminal domain of gp59 protein near the identified binding site for gp41 helicase.
**Significance:** Directional loading of replicative helicases requires structure-specific recognition of the DNA replication fork.

Bacteriophage T4 gp59 helicase assembly protein (gp59) is required for loading of gp41 replicative helicase onto DNA protected by gp32 single-stranded DNA-binding protein. The gp59 protein recognizes branched DNA structures found at replication and recombination sites. Binding of gp32 protein (full-length and deletion constructs) to gp59 protein measured by isothermal titration calorimetry demonstrates that the gp32 protein C-terminal A-domain is essential for protein-protein interaction in the absence of DNA. Sedimentation velocity experiments with gp59 protein and gp32ΔB protein (an N-terminal B-domain deletion) show that these proteins are monomers but form a 1:1 complex with a dissociation constant comparable with that determined by isothermal titration calorimetry. Small angle x-ray scattering (SAXS) studies indicate that the gp59 protein is a prolate monomer, consistent with the crystal structure and hydrodynamic properties determined from sedimentation velocity experiments. SAXS experiments also demonstrate that gp32ΔB protein is a prolate monomer with an elongated A-domain protruding from the core. Fitting structures of gp59 protein and the gp32 core into the SAXS-derived molecular envelope supports a model for the gp59 protein-gp32ΔB protein complex. Our earlier work demonstrated that gp59 protein attracts full-length gp32 protein to pseudo-Y junctions. A model of the gp59 protein-DNA complex, modified to accommodate new SAXS data for the binary complex together with mutational analysis of gp59 protein, is presented in the accompanying article (Dolezal, D., Jones, C. E., Lai, X., Brister, J. R., Mueser, T. C., Nossal, N. G., and Hinton, D. M. (2012) J. Biol. Chem. 287, 18596–18607).

The organization of the DNA replication complex at the replication fork is important for accurate, efficient DNA synthesis (1, 2). During bacteriophage T4 infection, coordination between the replisome and the primosome is essential for successful replication of the phage genome (3). Our review of the structural analysis of T4 DNA replication proteins discusses in more detail the individual components and their purported interactions (4). The gp43 DNA polymerase and the gp45 clamp comprise a replisome that synthesizes the leading and lagging DNA strands, whereas the gp41 helicase and gp61 primase comprise a primosome that unwinds double-stranded DNA ahead of the replisome. The gp41 helicase loads onto fork substrates in the absence of gp32 protein (5). However, in the presence of gp32 protein, the gp59 helicase assembly protein is required for efficient gp41 helicase loading (6). Understanding how T4 replication complexes assemble begins with determining how the gp59 protein recognizes the replication fork.

Bacteriophage T4 initiates replication in an origin-dependent and recombination-dependent manner; the former occurs in the early stage of infection and the latter prevails in the late stage (7). Both origin-dependent and recombination-dependent replication initiation require gp59 protein, gp32 protein, and gp41 helicase (8). The gp59 protein is involved in both types of replication initiation, but known mutants in the gp59 protein are most often defective in recombination-dependent replication (5, 9).

The gp59 protein is a 26-kDa basic protein with little structural or functional similarity to other prokaryotic DNA-binding proteins, but with structural similarity to eukaryotic HMG proteins (10). The gp59 protein recognizes single- and double-stranded DNA, but binds preferentially to fork (pseudo-Y junctions), cruciform, and recombination-like DNA structures (5, 10–13). Interaction with the gp32 protein, gp41 helicase, and
gp43 DNA polymerase occur while binding to the DNA fork (11, 12, 14). In the presence of the gp32 protein, gp59 protein is crucial for the loading of the gp41 helicase on the lagging strand (5). The gp59 protein interacts with gp32 protein in the presence or absence of DNA and promotes gp32 protein binding to DNA pseudo-Y junctions designed to be too short for cooperative binding of multiple gp32 proteins (5, 15, 16). The gp32 protein has three domains: the N-terminal B-domain (residues 1–21) responsible for cooperative self-assembly, a core domain (residues 22–253), which contains an OB-fold motif and binds with high affinity to single-stranded DNA, and a C-terminal A-domain (residues 254–301) that contributes to interactions with other replication proteins (5, 17–19). Cross-linking studies reveal that the gp32 core domain interacts with the C-terminal domain of the gp59 protein (20). Binding studies employing fluorescence anisotropy indicate that the A-domain interacts with gp59 protein (15).

Electron microscopy (EM) studies show that the gp59 protein remains at the DNA fork after loading the gp41 helicase and after replication initiates (21). The exact role of the gp59 protein–gp32 protein interaction at this DNA branch point is not clear, but evidence suggests that gp59 protein prevents the gp43 DNA polymerase from interacting with the gp41 helicase, perhaps to prevent early progression (5). FRET studies (22) indicate that gp43 DNA polymerase-gp41 helicase interaction occurs only if gp59 protein is removed or progresses from the fork. EM studies also show the gp59 protein localized at or near the DNA fork during replication fork progression, suggesting that the gp59 protein remains associated with the primosome but releases the DNA allowing replication to proceed (21, 23). The gp32 protein inhibits the loading of gp41 helicase in the absence of gp59 protein, yet the gp32 protein is required for leading strand synthesis if the gp41 helicase is loaded by gp59 protein (5). This complementarity of functions may indicate that gp32 protein facilitates removal of gp59 protein from the DNA fork to allow for replication fork progression (5).

In the present study, we investigate the interaction between the gp59 and gp32 proteins using small angle x-ray scattering (SAXS). These structural studies are complemented by sedimentation velocity analysis of the gp59 protein–gp32 protein heterodimer complex and with isothermal titration calorimetry (ITC) measurements of the formation of this complex. We verify that the interaction between gp59 protein and gp32 protein requires the C-terminal A-domain of gp32. We utilized a truncated form of the gp32 protein lacking the N-terminal B-domain, gp32ΔB, for subsequent experiments. This form of gp32 protein does not self-assemble and is less prone to aggregation, which is essential for small angle x-ray scattering experiments. The gp59 protein interacts with the gp32ΔB protein as a 1:1 complex with moderate affinity. The low-resolution SAXS envelopes of the gp59 protein–gp32 protein heterodimer indicate that it is a bent prolate structure. Models of the protein binary complex and the protein-DNA ternary complex have been generated. It appears that the gp32ΔB protein becomes more compact upon interaction with the gp59 protein. This model of the ternary complex combines the protein structure derived from the SAXS envelope with the mutational analysis of gp59 protein provided in our accompanying article (38).

**EXPERIMENTAL PROCEDURES**

**Protein Expression**—The expression vectors for T4 gp59 (pNN2859, full-length), gp32 (pAS6, full-length), gp32ΔA (pEK1, C-terminal deletion), 32ΔB (pEK2, N-terminal deletion), and gp32ΔAB (pKC30, N-terminal and C-terminal domain deletion) proteins were gifts from Drs. Nancy Nossal (NIH), Charles Jones (NIH), Richard Karpel (University of Maryland), and Yousif Shamoos (Rice University), respectively. The expression hosts used were: *Escherichia coli* BL21(DE3) pLysS for the gp59 protein, *E. coli* AR120 for the gp32ΔA, and *E. coli* AR120 for gp32ΔB with plasmids under control of the T7 promoter. Full-length gp32 protein (in *E. coli* N4830) and 32ΔAB (in *E. coli* OR1265) utilize the heat inducible λ<sub>P</sub> promoter. Transfected cells were grown in Luria broth (6 liters, 28 °C for heat inducible and 37 °C for chemical inducible promoters, 45 mg/ml of ampicillin) to an A<sub>600</sub> = 0.6 prior to induction, followed by growth for 3 h. The gp59 protein was induced with 1 mM isopropyl β-d-thiogalactopyranoside. The gp32 protein truncation constructs in AR120 cells were induced with 100 μM naldixic acid. Vectors under λ<sub>P</sub> control were induced at 42 °C for 1 h, followed by incubation at 37 °C for 2 h. Cells were harvested by centrifugation and the pellets were stored at −20 °C.

**Protein Purification**—Modifications to purification protocols (5, 10) are described here. Frozen cells with expressed gp59 protein were suspended in 10 ml of lysis buffer/g of cell pellet (25 mM BisTris, pH 6.5, 10 mM MgCl<sub>2</sub>, 150 mM NaCl, 2 mM β-mercaptoethanol (β-ME), 0.3% polyethyleneimine, and 1 mg of hen egg white lysozyme/100 ml of lysis buffer). The cells were stirred at room temperature until thawed, sonicated (70% duty cycle, 2 min), and centrifuged (18,000 × g, 30 min, 4 °C) to clarify. The pellet from the centrifugation was resuspended in a high salt extraction buffer (25 mM Tris-HCl, pH 7.5, 1 mM NaCl, 10 mM MgCl<sub>2</sub> and 2 mM β-ME, 2 ml/g of cells) and centrifuged (18,000 × g, 30 min, 4 °C). The supernatant fraction was diluted with 25 mM Tris-HCl, pH 7.5, until conductivity matched that of HA buffer A (25 mM Tris-HCl, pH 7.5, 100 mM NaCl), loaded onto ceramic hydroxyapatite (HA, Applied Biosystems, 10-ml bed volume, 8 ml/min) equilibrated with HA buffer A, and eluted with a 10-column volume linear gradient of HA buffer B (HA buffer A + 1 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>). The flowthrough was dialyzed into HS buffer A (25 mM Tris-HCl, pH 6.5, 75 mM NaCl, 10 mM MgCl<sub>2</sub>, 2 mM β-ME) and applied to a Poros HS (Applied Biosystems, 20-ml bed volume, 8 ml/min) equilibrated with HS buffer A and eluted with a 10-column volume linear gradient of HS buffer B (HS buffer A + 1 mM (NH<sub>4</sub>)Cl). Preparations were concentrated with a Millipore Amicon ultra-
Complexes of T4 Gp59 Helicase Loading Protein

centrifugation concentrator (MWCO 10,000) and stored frozen in 20% glycerol at −80 °C; yields were 100–150 mg/6 liters of culture.

Lysates for gp32 constructs were described above for gp59 protein with variations in salt and buffer; 25 mM BisTris, pH 6.5, 50 mM NaCl, 2 mM CaCl2 for gp32 and gp32AB proteins, 25 mM BisTris, pH 6.5, 100 mM NaCl, 2 mM CaCl2 for 32ΔA; and 25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 2 mM CaCl2 for gp32ΔB protein. No extraction step was necessary as the proteins were in the low salt lysates. For gp32 protein constructs (except as noted for gp32AB), clarified lysates were applied to Q-Septarose (QS, GE Healthcare, 60-ml bed volume, flow rate 1 ml/min) equilibrated with QS buffer A (25 mM BisTris/HCl, pH 6.5, 50 mM NaCl, 1% glycerol, and 2 mM β-ME; 25 mM Tris-HCl, pH 7.5, 50 mM trisodium citrate, 1% glycerol, and 2 mM β-ME for 32ΔB protein) and eluted with a 10-column volume linear gradient with QS buffer B (QS buffer A plus 1 M NaCl). To remove endogenous nucleases, the QS fractions were adjusted to 1 M NaCl and applied to Poros PE (8-ml bed volume, 10 ml/min) equilibrated with QS buffer B. The Poros PE eluents were diluted with buffer (25 mM BisTris, pH 6.5, 25 mM Tris, pH 7.5, for 32ΔB) to reduce the conductivity and applied to Poros HQ (Applied Biosystems; 20-ml bed volume, flow rate 10 ml/min) equilibrated with QS buffer A, and eluted with a 10-column volume linear gradient of QS buffer B. Preparations were concentrated with a Millipore Amicon ultracentrifugation concentrator (MWCO 10,000) and stored frozen in 20% glycerol at −80 °C; yields were 100–150 mg/6 liters of culture. Solubility screens (24) were performed on protein preparations to identify optimal buffer conditions favorable for the experiments to reduce aggregation and polydispersity and increase the solubility of the proteins.

Agarose Gel Electrophoresis—Protein (250 nmol of gp59, gp32, gp32ΔA, gp32ΔB, and gp32ΔNC) and protein complexes (250 nmol of each protein) were loaded on a 0.6% agarose gel. Electrophoresis was performed at 4 °C at 50 V for 3.5 h using 50 mM Tris borate, pH 6.5, 50 mM NaCl, 1 mM EDTA running buffer. The gel Electrophoresis was performed at 4 °C at 50 V for 3.5 h using 50 mM Tris borate, pH 6.5, 1 mM EDTA running buffer. The gel was prepared for visualization by incubation in 0.05% SDS for 1 h, followed by incubation in SYPRO Orange (Invitrogen) overnight. Protein bands were detected using an Illumatool BLS 470 (Lightools Research).

Isothermal Titration Calorimetry—The gp32ΔB protein titrations into gp59 protein were performed with a MicroCal VP-ITC microcalorimeter (GE Healthcare, MicroCal, Northampton, MA) at 20 °C. Proteins were dialedyzed against binding buffer A (25 mM BisTris, pH 6.5, 150 mM NH4Cl, 10 mM MgCl2, and 2 mM β-ME). The injection syringe contained gp32ΔB protein (883 μM) or buffer. The sample cell contained 1.4 ml of buffer for blank titrations or 25 μM gp59 protein. Control experiments (buffer into buffer, buffer into protein, and protein into buffer) were performed to determine the heats of dilution, mixing, and injection (25, 26). Forty 5-μl injections (10 s each) were made with a stir rate of 310 rpm; injections were made at 300-s intervals. Experiments were performed in triplicate. A one-site binding model was fit to the data using Microcal Origin software.

Fluorescence Anisotropy Using Labeled Pseudo-Y DNA—For these studies, proteins were titrated into 5’-hexachloro-6-carboxyl fluorescein (5’-Hex)-labeled DNAs. As diagrammed, a 30/30 DNA pseudo-Y junction had a 15-mer duplex with two 15-mer single-stranded overhangs and a 20/30 DNA pseudo-Y junction had a 15-mer duplex with a 5-mer 5’ overhang and a 15-mer 3’ overhang.

Anisotropy was determined using a PTI QM4 fluorimeter (Photon Technology Inc., Birmingham, NJ) with excitation and emission film polarizers. The excitation wavelength was 535 nm with a band pass of 8 nm; emission was measured at 556 nm with a band pass of 2 nm. Proteins and the DNA substrates were in buffer A. Proteins were titrated into 50, 200, and 400 nm DNA (425 μl). Protein additions were extended to concentrations where the signal plateaued. Anisotropy was calculated from the relationship,

$$ r = \frac{l_{VY} - G l_{VV}}{l_{VY} + 2G l_{VV}} $$

where G is a factor to correct for the difference in the sensitivity of the system to vertically and horizontally polarized light (27).

Dynamic Light Scattering—The DynaPro Titan (Wyatt Technologies, Santa Barbara, CA) with Dynapro temperature-controlled sampler was employed to determine the polydispersity and oligomeric state of the proteins and protein complexes. Samples (~1 mg/ml of protein) were centrifuged (20,000 × g) or filtered (0.45 μm) just prior to measurements. Data were collected at 20 and 4 °C in a 12-μl masked cuvette. Protein samples were equilibrated in binding buffer A at the experimental temperature and 10–15, 10-s acquisitions were collected. Data were analyzed with the Dynamics instrument software.

Analytical Ultracentrifugation—Sedimentation velocity experiments were performed with a Beckman XL-1 analytical ultracentrifuge (Beckman Coulter, Palo Alto, CA) at 20 °C, 48,000 rpm, 295 nm using absorbance optics. Proteins were in binding buffer A. Data were collected for the gp59 and gp32ΔB proteins individually at 15 μM. The gp59 protein-32ΔB protein complex was investigated over a range of concentrations (59:32ΔB): 1:1 (20:40 μM), 1:2 (20:40 μM), 1:4 (20:80 μM), 2:1 (40:20 μM), and 4:1 (80:20 μM). Data were deconvoluted to determine sedimentation coefficient distributions using SEDFIT and SEDANAL (28–30). The equilibrium constant for formation of the gp59-gp32B protein complex was determined by fitting a model for the formation of a heterodimer (A + B → A:B) to the scans using Sedanal. Data sets at three ratios of protein (20:20, 40:40, and 20:40 μM) were used for this analysis.
**Small Angle X-ray Scattering**—Data were collected at the Advanced Photon Source beamline 15-ID (ChemMatCARS, Argonne, IL) with the Bruker 6000 CCD on the gp59-gp32ΔB protein complex, gp59 protein, gp32ΔB protein, and ovalbumin as a control. Data were collected at 1.5 Å, with a camera distance of 0.56 m, Q range of 0.02–0.84 Å⁻¹ (7.6–312.7 Å) in a 1-mm capillary. Multiple protein (3 mg/ml, 94 μM gp32ΔB, 115 μM gp59, 94 μM gp59-gp32ΔB complex in binding buffer A) and buffer measurements were taken at 20- and 40-s intervals with the sample flow rate of 0.20 μl/s. The measurements were averaged, followed by subtraction of the buffer data from the protein data using the beamline program (SAXS/WAXS version 3.294). Software from the ATSAS 2.2 program suite was used to analyze the data and generate molecular envelopes. The radius of gyration (Rg) and forward scattering (I(0)) were determined using PRIMUS and GNOM (χ₂ gp59 protein = 0.99; χ₂ gp32ΔB = 1.00). The molecular weights of the complex and individual proteins were calculated from forward scattering values using the relation in Equation 2.

\[
\frac{I_{\text{overall}}}{MW_{\text{overall}}} = \frac{I_{\text{unk}}}{MW_{\text{unk}}} = c_{\text{unk}} \frac{c_{\text{unk}}}{c_{\text{unk}}} \quad \text{(Eq. 2)}
\]

Molecular weights were determined using the Porod volume (Porod volume × 1.2)/2 and molecular volume was determined with Primus Porod calculation. A Fourier transform of the scattering data generated the pair distribution function (P(r) plot) (GNOM). The ab initio program GASBOR generated molecular envelopes for the gp59 and gp32ΔB proteins using chain-like dummy atom residues. The GASBOR default parameters (no symmetry, unknown shape) were used in each computation, with the program searching for 217 residues for the gp59 protein and 301 residues for 32ΔB. Six to 10 GASBOR models were aligned and averaged using SUPCOMB and DAMAVER. MASSHA generated five heterodimer complexes using the atomic coordinates of gp59 and gp32ΔB proteins. The output complex was subjected to rigid body refinements against the experimental scattering data. To model in the A-domain of the gp32ΔB protein, the program CHADD was utilized to build in the missing A-domain into the output MASSHA model. For each output model, a theoretical scattering curve was calculated using the program CRYSOL, and compared with the experimental scattering curve.

**RESULTS**

*Interactions between T4 Gp59 Protein and Gp32 Protein Truncations*—Aggregation of gp59 and gp32 protein complexes has been problematic for investigations using structural methods (31, 32). In an effort to discern solvent conditions that improve the homogeneity of the protein samples, we performed solubility screens on gp59 protein, gp32 protein, and gp32 protein truncations (24). The optimal solution (25 mM BisTris, pH 6.5, 150 mM NH₄Cl, 10 mM MgCl₂, and 2 mM β-ME) decreased the polydispersity from >25 to <10%, as determined by dynamic light scattering. We then used this solution in all subsequent experiments (isothermal titration calorimetry, analytical ultracentrifugation, fluorescence anisotropy, and small angle x-ray scattering).

**Complexes of T4 Gp59 Helicase Loading Protein**

| Parameter | Gp32 protein | gp32ΔB protein |
|-----------|-------------|---------------|
| N'        | 0.53 ± 0.01 | 1.34 ± 0.03   |
| Kₜ        | 2.72 ± 10² ± 0.2 × 10⁵ | 2.8 ± 10⁵ ± 0.3 × 10⁶ |
| ΔH        | −7.24 ± 0.2 kcal mol⁻¹ | −4.12 ± 0.15 kcal mol⁻¹ |
| ΔS        | 0.16 kcal mol⁻¹ K⁻¹ | 0.109 kcal mol⁻¹ K⁻¹ |

We verified that gp59 and gp32 proteins formed complexes under our new solution conditions using native agarose gel electrophoresis and results were in accordance with earlier results (15, 16). No interactions were detected between gp59 protein and the gp32ΔA or gp32ΔB proteins (20). The interactions of gp59 protein with full-length gp32 and gp32ΔB proteins were confirmed. These complexes were then further analyzed by dynamic light scattering and ITC to determine thermodynamic parameters of binding (Table 1). Using DLS, an equimolar mixture of the gp59 protein and full-length gp32 protein displayed an ~1:1 heterodimeric complex but with significant polydispersity. An equimolar mixture of gp59 and gp32ΔB proteins was also ~1:1 and the polydispersity was only slightly higher than the individual proteins. Titration of gp59 protein into full-length gp32 protein (Fig. 1, panel A) was exothermic (ΔH = −7.24 kcal/mol) with a dissociation constant of 3.7 μM, a stoichiometry of 0.538 to 1 (59:32), and an entropy (ΔS) of 0.164 kcal/mol-K. The titration of gp59 protein into gp32ΔB protein (Fig. 1, panel B) was also exothermic (ΔH = −4.12 kcal/mol) with a dissociation constant of 3.6 μM, a stoichiometry of 1.34:1 (59:32ΔB), and an entropy (ΔS) of 0.109 kcal/mol-K. Titration of gp32ΔB into gp59 protein did not result in significant heat change suggesting no interaction between the gp59 protein and gp32ΔB, consistent with the results of native agarose gel electrophoresis. No precipitation was detected before or after the ITC experiments, indicating that the proteins remain soluble.

The individual proteins and different ratios of gp59 and gp32ΔB proteins in the optimal buffer were examined using analytical ultracentrifugation (AUC) sedimentation velocity experiments. The results indicate that the gp59 protein and gp32ΔB protein alone are monomeric (Fig. 2A). The estimated molecular masses were 23.4 (25.9 kDa calculated) and 36.8 kDa (31.8 kDa calculated), respectively (Table 3). These results are comparable with early AUC studies (15). To determine stoichiometry of the gp59 protein-gp32ΔB protein complex, we examined mixtures of gp59 protein and the gp32ΔB protein. The peak broadening indicates that dissociation occurs during centrifugation. Increasing the concentration of either gp59 protein or gp32ΔB protein relative to the other, shifts the major peak toward the molecular mass expected for a 1:1 complex (57.7 kDa).

The analysis of the gp59-gp32ΔB protein complex by analytical ultracentrifugation indicates that it is a 1:1 complex. Global analysis of the gp59-gp32ΔB AUC experiments gave an association constant of 1.07 × 10⁵ (Kₜ ~ 9 μM) and sedimentation coefficient of 4.38 for s₄ using a hetero-association model A + B → A:B. Fixing s₄ for the complex at 4.4 gave a value of 1.02 × 10⁵ for Kₜ. These values are consistent with the results of the ITC experiments discussed above. Using
the gp32ΔB protein deletion, which is less prone to aggregation, and solution conditions that minimize aggregation allowed measurements of a wider range of protein concentrations in AUC experiments than employed in the earlier studies (15).

**Gp59-Gp32ΔB Protein Interactions with Fork DNA Substrates**—The interaction of the gp59 protein and gp32ΔB proteins with fork DNA substrates were characterized by titrating each protein with solutions of hexachlorofluorescence-labeled fork substrates and measuring the change in anisotropy and total fluorescence to detect protein-DNA complexes. Gp59 protein and gp32ΔB protein exhibited hyperbolic binding with stoichiometries consistent with the formation of 1:1 complexes with pseudo-Y DNA substrates. The gp59 protein had similar affinity for 30/30 fork DNA ($K_d = 55.4 \pm 0.1$ nM) and for 20/30 fork DNA ($40.5 \pm 0.5$ nM). Gp32ΔB protein bound relatively weakly to 30/30 fork DNA ($K_d = 6.9 \pm 0.1$ μM) and weaker to 20/30 fork DNA ($K_d = 29.2 \pm 0.2$ μM). Results of the binding studies are summarized in Table 2.

As described previously, binding of the gp32ΔB protein is directional with the gp59 protein inhibiting gp32ΔB protein binding when the 5' arm is shorted (10, 33). Gp59 protein binding to fork DNAs significantly reduces the affinity of gp32ΔB for these DNA structures as shown in Table 2. Our data agrees with earlier studies indicating that gp59 protein binds to fork DNA and recruits gp32 protein to DNA structures with at least 9 nucleotides on the lagging strand arm (5). Also in agreement is the effect of gp32 and gp32ΔB on the affinity of the gp59 protein for fork DNA (13, 34). The relative weak level of incorporation of gp32ΔB protein into the gp59-DNA complex and the propensity of these ternary complexes to aggregate at higher concentrations precluded further analysis.

**Small Angle X-ray Scattering**—The $K_d$ for the gp59-gp32ΔB binary complex was $\sim 4$ μM by ITC and $\sim 9$ μM by AUC. Based on these values, gp59 and gp32ΔB were combined at 94 μM, each to ensure that this mixture would be greater than 90% gp59-gp32ΔB complex in SAXS experiments. The scattering curves, Guinier plots, and pair distribution function plots ($P(r)$ plot) for the gp59 protein, gp32ΔB protein, and gp59-gp32ΔB protein complex are shown in Fig. 3. The radius of gyration ($R_g$) was calculated from the Guinier region and the pair distribution function ($P(r)$ plot) are shown in Table 3.

Scattering data for the gp59-gp32ΔB complex consistently generated elongated models for this binary complex. $R_g$ for the gp59 protein ($\sim 22$ Å) and gp32ΔB protein ($\sim 29$ Å) are consistent in both methods of evaluation (PRIMUS and GNOM), whereas the $R_g$ for the complex varies slightly ($\sim 33$ to $\sim 38$ Å). The Guinier region for each protein and protein complex is linear indicating the absence of significant aggregation in the samples during data collection. The pair distribution function suggested an elongated, prolate envelope for the proteins. The maximum particle distance (maximum width) increased significantly with the formation of the protein complex.

The maximum particle distance of the gp59-gp32ΔB complex is 130 Å (over the $q$ range 0.02–0.2 Å$^{-1}$), which is significantly larger than the individual proteins (gp59 protein, 80 Å; gp32ΔB protein, 118 Å) (Fig. 3C). The molecular mass of the 59/32ΔB complex determined from the forward scatter of the protein complex ($I_0 = 20,460 \pm 81.42$) is 67.7 kDa, whereas the molecular mass estimation determined by Porod volume is...
58.4 kDa. Both calculated molecular masses of the protein complex are consistent with the theoretical molecular mass of 57.7 kDa for a 1:1 gp59 protein-gp32ΔB protein complex and the results of analytical ultracentrifugation.

*Ab initio* models of the individual proteins were generated with GASBOR and averaged and aligned using DAMAVER and SUBCOMB. The molecular envelope of the gp59 protein is elongated ($\chi^2 = 1.46, \text{NSD} = 0.88$) (Fig. 4A); the theoretical scattering curve has a $\chi^2$ value of 2.51. Gp32ΔB protein has an overly extended, elongated molecular envelope ($\chi^2 = 1.76, \text{NSD} = 0.97$) (Fig. 4B); unlike the gp59 protein, the theoretical scattering curve of the gp32ΔB protein has a higher $\chi^2$ value of 18.2, suggesting differences in the scattering between the full-length protein and the truncated version. Kratky plots were calculated to estimate the degree of protein folding. The gp59 protein and the gp59 protein-gp32ΔB protein complex had bell-shaped curves that increased with the increasing $q$ range, suggesting that the proteins are mostly folded (data not shown). The plots for the gp32ΔB protein had a prominent increase with increasing $q$ range revealing that a significant percentage of the protein is extended or unfolded (data not shown).

Rigid body refinements of the individual high-resolution structures of the gp59 and gp32ΔB proteins were performed using the program MASSHA to generate a quaternary structure of the complex (Fig. 4C). In this model, the gp32ΔB protein is oriented diagonally relative to the long axis of the gp59 protein ($\chi^2 = 6.1, \text{NSD} = 0.95$). MASSHA places the gp32ΔB in proximity of the C-terminal domain of the gp59 protein. The missing A-domain of the gp32ΔB protein was built into the averaged MASSHA complex model using the program CHADD based on the experimental scattering curve ($\chi^2 = 5.5, \text{NSD} = 0.5$). CHADD places the missing A-domain between the gp32ΔB and gp59 proteins (Fig. 4D), in agreement with native gel and ITC results that show the requirement of the A-domain for the formation of the complex. Our data supports the previous AUC experimental elongated profile for the gp59, gp32-gp32ΔB, and the binary complex (15).

**DISCUSSION**

The bacteriophage T4 gp59 helicase loading protein is required for the loading of the gp41 helicase onto gp32 single strand DNA-binding protein-protected DNA. The interaction between gp59 protein and gp32 protein is mitigated by the core and A-domains of gp32 protein (5, 6, 12, 15, 16, 20, 35).

Because attempts to obtain detailed structural information using x-ray crystallography were unsuccessful, we decided to generate a molecular model of the complexes by docking the known high-resolution structures into molecular envelopes determined by small angle x-ray scattering. Both the earlier crystallization efforts and these studies were hindered by high polydispersity of the protein samples. By reformulating a new buffer system (based on a crystallization solubility screen) we significantly reduced polydispersity and were able to re-examine the binding and hydrodynamic properties of the gp59 protein, the gp32 protein (including gp32 protein truncations), and the gp59-gp32 protein complexes.

We report here a molecular model of the gp59-gp32-DNA ternary complex derived from fitting to the experimental molecular envelopes of gp59 protein, gp32ΔB protein, and their binary complexes. The ternary complex models are consistent with both affinity analysis and mutagenesis studies of gp59 protein presented in a accompanying article (38). Native agarose gel electrophoresis results were in qualitative agreement with other studies (15, 16), indicating binding of the gp59 protein to both the gp32 full-length and the gp32ΔB-truncated proteins, but not to gp32ΔA and gp32ΔAB proteins.

A dissociation constant of 2 nm for the gp59-gp32ΔB protein complex and the gp59:32A-domain peptide reported earlier suggests extremely tight binding (15). However, and in contrast,
Complexes of T4 Gp59 Helicase Loading Protein

![Graph A: Normalized Scattering Curves](image)

![Graph B: Guinier Plot](image)

![Graph C: Pair Distribution Function Plot](image)

**Figure 3. SAXS data analysis.** A, normalized scattering curves for the gp59 protein, gp32ΔB protein, and gp59-gp32ΔB protein complex (buffer subtracted experimental scattering curves). B, plot of the Guinier region of the scattering data for each protein and the protein complex indicating there is no aggregation in either of the protein samples. The radius of gyration ($R_g$) of each sample is calculated from the slope of the line; gp59 protein, $R_g = 21.5 \pm 0.2$ Å; gp32ΔB protein, $R_g = 29.9 \pm 0.2$ Å; gp59-gp32ΔB protein complex, $R_g = 33.2 \pm 0.1$ Å. C, pair distribution function plot ($P(r)$ plot) determines the overall shape and maximum particle distance of each protein and protein complex. All curves are characteristic of elongated proteins.

**Table 3**

| Protein                  | Radius of gyration ($R_g$) (SAXS-Guinier) | Radius of gyration ($R_g$) (SAXS-GNOM) | Hydrodynamic radius (DLS) |
|--------------------------|-------------------------------------------|----------------------------------------|---------------------------|
| Gp59 protein             | $21.5 \pm 0.2$ Å                          | $21.6 \pm 0.2$ Å                      | $23.0$ Å                  |
| Gp32ΔB protein           | $29.9 \pm 0.2$ Å                          | $30.8 \pm 0.4$ Å                      | $28.0$ Å                  |
| Gp59-gp32ΔB complex      | $33.2 \pm 0.1$ Å                          | $38.2 \pm 0.2$ Å                      | $31.0–34.0$ Å             |

Our analysis of the gp59-gp32ΔB complex using dynamic light scattering and size exclusion chromatography does not support this notion, as we note significant dissociation regardless of concentration adjustments.

The extremely low dissociation constant previously reported (15) is based on rhodamine labeling of cysteine residues in gp59 protein and an observed decrease in anisotropy upon addition of the gp32 proteins. The authors attribute decreased anisotropy to increased mobility of the fluorophor upon protein complex formation, although one might typically expect an increase in anisotropy with the formation of a larger complex. The two cysteines in gp59, Cys-42 located in the HMG domain, and Cys-215 near the C terminus, form intramolecular disulfides and the resulting oligomers decrease solubility and increase polydispersity. While determining the gp59 protein crystal structure via multiple heavy atom isomorphous replacement phasing, we observed that the cysteine residues could not be successfully modified by a number of small, thiol-specific labels without disruption of the structure of the protein. Therefore, it would seem unlikely that rhodamine derivatives would be any more successful.

To examine further if the extremely tight binding claimed by Xu et al. (15) can be supported, we employed noninvasive techniques of isothermal titration calorimetry and analytical ultracentrifugation. Using ITC, we measured a $K_d$ of ~4 μM between gp59 and gp32ΔB. Furthermore, global analysis of the AUC data using varying ratios of protein indicate a $K_d$ of ~9 μM for the same complex. These values, consistent within experimental error, are 3 orders of magnitude less than previously reported.

Intuitively, we feel that a system such as DNA replication, characterized by myriad dynamic interactions between different protein species, would be unlikely to involve extremely tight complex formation. Therefore, these new estimated $K_d$ values are more reasonable physiologically. Given the estimated number of 10,000 gp32 molecules expressed per bacterium during T4 infection (36), the gp32 concentration is ~20 μM, which is likely an underestimation where specific activity is significantly greater due to macromolecular crowding (37).

Given this relatively weak interaction, could it be that binding to DNA substrate increases protein-protein affinity? We have previously reported that gp59 protein preferentially binds to branched DNA substrates. The $K_d$ of gp59 protein binding to the short pseudo-Y junction DNA is ~50 nm, whereas gp32ΔB protein binding to the same minimal substrates is significantly weaker in the low micromolar range (Table 2). Titration of the gp59 protein-DNA binary complex with the gp32ΔB protein indicated an even lower affinity than that measured between the two proteins alone. These results further support the ideas that gp32 protein binds cooperatively and that gp59 protein binds preferentially at the replication fork. The gp59 protein prevents progression of the gp43 polymerase until the gp41 helicase is loaded onto the gp32 protein-protected lagging strand.

Two structural models have been proposed for gp32 protein interaction with gp59 protein. Based on the crystal structure, the core of the gp32 protein was modeled to interact with the
C-domain of gp59 and the A-domain of gp32 to interact with the N-domain of gp59 (10). In an alternate model, the A-domain of the gp32 protein interacts with the C-domain of the gp59 protein, whereas the core of the gp32 protein interacts with the N-domain of the gp59 protein (20). The models have gp59 protein binding to fork DNA with gp32 protein protecting the lagging strand. This arrangement suggests a model for oligomerized gp32 to weaken gp59 affinity for the fork DNA, allowing replication to progress (13, 21).

We analyzed gp59 protein, gp32ΔB protein, and the gp59-gp32ΔB protein complex using the ChemMatCARS 15-ID small angle x-ray scattering device at the Advanced Photon Source, Argonne National Laboratories. Molecular envelopes of the individual proteins are consistent with the crystal structures (gp59, Protein Data Bank code 1C1K; and gp32ΔAB, Protein Data Bank code 1GPC). The model of the core of gp32 protein has been positioned into the SAXS-generated molecular envelope (Fig. 4B). The large empty region is likely due to the presence of the disordered A-domain in the gp32ΔB protein, which is absent in the crystal structure used as the model. This result is entirely consistent with our crystal structure of gp32ΔB protein (data not shown) that revealed only the core of the protein with the A-domain disordered. Analysis of the molecular envelope for the gp59-gp32ΔB protein complex revealed an elongated structure the size of a heterodimer. The loss of extended regions in the complex suggests that the A-domain of the gp32 protein adopts a more compact structure in the presence of the gp59 protein. Our AUC and SAXS results indicate that the gp32ΔB protein is monomeric at concentrations up to 94 μM (Table 4). In contrast to previous reports, we do not observe weak dimerization of gp32ΔB; we attribute this difference to our improved solution conditions (15). The ab initio envelope from CHADD (Fig. 4D) and the MASSHA docking of two crystal structures monomers (Fig. 4C) are consistent. The MASSHA best-fit model has the core of gp32 interacting with the C-domain of gp59. The binary model was used to generate a model of the ternary complex of gp59-gp32-DNA (Fig. 5).

The mutational analysis, described in the accompanying article (18), confirms the significance of the HMG-like domain previously used to generate the model for gp59 protein binding to pseudo-Y junction DNA, which proposed binding sites for the gp32 protein and the gp41 helicase. The mutational analysis indicates that gp41 helicase binds to the C-terminal region of the gp59 protein. We have generated a model of the gp59-gp32-DNA ternary complex by superimposing the gp59 protein models of the gp59-DNA model and the new gp59-gp32 model from the SAXS described here (Fig. 5). Remarkably, the only adjustment required was repositioning the 5′ single-stranded arm, which now lies diagonally across the gp59 protein with gp32 protein docking to the C terminus of gp59.
The gp41 helicase protein allows for a handoff of DNA from the gp32 protein to the binding site. Within this complex, the gp59 DNA lies diagonally across the top face of gp59 adjacent to the binding site for the gp41 helicase. whereas gp32 protein protects the single-stranded DNA, a section of DNA lies across the surface of gp59 adjacent to the binding site for the gp41 helicase. In the gp59 protein-DNA binary model, the HMG domain of the gp59 protein interacts with the branch point of the pseudo-Y junction DNA. For the ternary complex, the gp32 protein-DNA binary model, the HMG domain of the gp59 protein interacts with the branch point of the pseudo-Y junction DNA. For the ternary complex, the gp59 protein-DNA model, the HMG domain of the gp59 protein interacts with the branch point of the pseudo-Y junction DNA. In the gp59 protein-DNA binary model, the HMG domain of the gp59 protein interacts with the branch point of the pseudo-Y junction DNA. For the ternary complex, the gp59 protein-DNA model, the HMG domain of the gp59 protein interacts with the branch point of the pseudo-Y junction DNA.

Acknowledgments—We thank Drs. Joan and Craig Hyde and Dr. Leif Hanson for helpful discussions and editorial assistance. We thank Dr. Andrew Herr, Michelle Gomes, and Deborah Conrady for the use of the JOURNAL OF BIOLOGICAL CHEMISTRY

REFERENCES

1. Aves, S. J. (2009) DNA replication initiation. Methods Mol. Biol. 521, 3–17
2. Hübscher, U. (2009) DNA replication fork proteins. Methods Mol. Biol. 521, 19–33
3. Karam, J. D., and Drake, J. W. (1994) Molecular Biology of Bacteriophage T4, American Society for Microbiology, Washington, D.C.
4. Mueser, T. C., Hinerman, J. M., Devos, J. M., Boyer, R. A., and Williams, K. J. (2010) Structural analysis of bacteriophage T4 DNA replication. A review in the Virology journal series on bacteriophage T4 and its relatives. Virology 1, 7, 359
5. Jones, C. E., Mueser, T. C., and Nossal, N. G. (2004) Bacteriophage T4 DNA replication is required for helicase-dependent leading strand synthesis when the helicase is loaded by the T4 59 helicase-loading protein. J. Biol. Chem. 279, 12067–12075
6. Barry, E. R., and Bell, S. D. (2006) DNA replication in the archaea. Microbiol. Mol. Biol. Rev. 70, 876–887
7. Kreuzer, K. N. (2000) Recombination-dependent DNA replication in phage T4. Trends Biochem. Sci. 25, 165–173
8. Kong, D., Nossal, N. G., and Richardson, C. C. (1997) Role of the bacteriophage T7 and T4 single-stranded DNA-binding proteins in the formation of joint molecules and DNA helicase-catalyzed polar branch migration. J. Biol. Chem. 272, 8380–8387
9. Shah, D. B. (1975) Replication and recombination of gene 59 mutant of bacteriophage T4D. J. Virol. 17, 175–182
10. Mueser, T. C., Jones, C. E., Nossal, N. G., and Hyde, C. C. (2000) Bacteriophage T4 gene 59 helicase assembly protein binds replication fork DNA. The 1.45-Å resolution crystal structure reveals a novel α-helical two-domain fold. J. Mol. Biol. 296, 597–612
11. Barry, J., and Alberts, B. (1994) Purification and characterization of bacteriophage T4 gene 59 protein. A DNA helicase assembly protein involved in DNA replication. J. Biol. Chem. 269, 33049–33062
12. Yonesaki, T. (1994) The purification and characterization of gene 59 protein from bacteriophage T4. J. Biol. Chem. 269, 1284–1289
13. Jones, C. E., Green, E. M., Stephens, J. A., Mueser, T. C., and Nossal, N. G. (2004) Mutations of bacteriophage T4 gene 59 helicase loader defective in binding fork DNA and in interactions with T4 32 single-stranded DNA-binding protein. J. Biol. Chem. 279, 25721–25728
14. Xi, J., Zhuang, Z., Zhang, Z., Selter, T., Spiering, M. M., Hammes, G. G., and Benkovic, S. J. (2005) Interaction between the T4 helicase-loading protein (gp59) and the DNA polymerase (gp43). A locking mechanism to delay replication during replisome assembly. Biochemistry 44, 2305–2318
15. Xu, H., Wang, Y., Bleuit, J. S., and Morrical, S. W. (2001) Helicase assembly protein gp59 of bacteriophage T4. Fluorescence anisotropy and sedimentation studies of complexes formed with derivatives of gp32, the phage ssDNA-binding protein. Biochemistry 40, 7651–7661
16. Lefebvre, S. D., Wong, M. L., and Morrical, S. W. (1999) Simultaneous interactions of bacteriophage T4 DNA replication proteins gp59 and gp32 with single-stranded (ss) DNA. Co-modulation of ssDNA binding activities in a DNA helicase assembly intermediate. J. Biol. Chem. 274, 22830–22838
17. 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
18. Spicer, E. K., Williams, K. R., and Konigsberg, W. H. (1979) T4 gene 32 protein tryptophan-generated fragments. Fluorescence measurement of DNA-binding parameters. J. Biol. Chem. 254, 6433–6436
19. Waindner, L. A., Flynn, E. K., Wu, M., Li, X., and Karpel, R. L. (2001) Domain effects on the DNA-interactive properties of bacteriophage T4 gene 32 protein. J. Biol. Chem. 276, 2519–2526
20. Ishmael, S. T., Alley, S. C., and Benkovic, S. J. (2001) Identification and mapping of protein-protein interactions between gp32 and gp59 by cross-linking. J. Biol. Chem. 276, 25236–25242
21. Nossal, N. G., Makhov, A. M., Chastain, P. D., 2nd, Jones, C. E., and Griffith, J. D. (2007) Architecture of the bacteriophage T4 replication complex revealed with nanoscale bioprinters. J. Biol. Chem. 282, 1098–1108
22. Xi, J., Zhang, Z., Zhuang, Z., Yang, J., Spiering, M. M., Hammes, G. G., and Benkovic, S. J. (2005) Interaction between the T4 helicase loading protein (gp59) and the DNA polymerase (gp43). Unlocking of the gp59-gp43 DNA complex to initiate assembly of a fully functional replisome. Biochemistry 44, 7747–7756
23. Chastain, P. D., 2nd, Makhov, A. M., Nossal, N. G., and Griffith, J. (2003) Architecture of the replication complex and DNA loops at the fork generated by the bacteriophage T4 proteins. J. Biol. Chem. 278, 21276–21285
24. Izaac, A., Schall, C. A., and Mueser, T. C. (2006) Assessment of a preliminary solubility screen to improve crystallization trials. Uncovering crystal condition searches. Acta Crystallogr. D Biol. Crystallogr. 62, 833–842
25. Leavitt, S., and Freire, F. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. Curr. Opin. Struct. Biol. 11, 560–566
26. Izaac, A., Schall, C. A., and Mueser, T. C. (2006) Assessment of a preliminary solubility screen to improve crystallization trials. Uncovering crystal condition searches. Acta Crystallogr. D Biol. Crystallogr. 62, 833–842
27. Janelidze, J., and Bossard, H. R. (1998) Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. J. Mol. Recognit. 12, 3–18
28. Lakowicz, J. R. (1999) Principles of Fluorescence Spectroscopy, Second Ed., Kluwer Academic/Plenum Publishers, New York
29. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biochem. Biophys. J. 78, 1606–1619
30. Laue, T. M., and Stafford, W. F., 3rd. (1999) Modern applications of ana-
lytical ultracentrifugation. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 75–100
30. Stafford, W. F. (2000) Analysis of reversibly interacting macromolecular systems by time derivative sedimentation velocity. *Methods Enzymol.* **323**, 302–325
31. Ishmael, F. T., Alley, S. C., and Benkovic, S. J. (2002) Assembly of the bacteriophage T4 helicase. Architecture and stoichiometry of the gp41-gp59 complex. *J. Biol. Chem.* **277**, 20555–20562
32. Prigodich, R. V., Shamoo, Y., Williams, K. R., Chase, J. W., Konigsberg, W. H., and Coleman, J. E. (1986) $^1$H NMR (500 MHz) identification of aromatic residues of gene 32 protein involved in DNA binding by use of protein containing perdeuterated aromatic residues and by site-directed mutagenesis. *Biochemistry* **25**, 3666–3672
33. Jones, C. E., Mueser, T. C., Dudas, K. C., Kreuzer, K. N., and Nossal, N. G. (2001) Bacteriophage T4 gene 41 helicase and gene 59 helicase-loading protein. A versatile couple with roles in replication and recombination. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8312–8318
34. Ma, Y., Wang, T., Villemain, J. L., Giedroc, D. P., and Morrical, S. W. (2004) Dual functions of single-stranded DNA-binding protein in helicase loading at the bacteriophage T4 DNA replication fork. *J. Biol. Chem.* **279**, 19035–19045
35. Morrical, S. W., Beernink, H. T., Dash, A., and Hempstead, K. (1996) The gene 59 protein of bacteriophage T4. Characterization of protein-protein interactions with gene 32 protein, the T4 single-stranded DNA-binding protein. *J. Biol. Chem.* **271**, 20198–20207
36. Alberts, B. M., and Frey, L. (1970) T4 bacteriophage gene 32. A structural protein in the replication and recombination of DNA. *Nature* **227**, 1313–1318
37. Zimmerman, S. B., and Trach, S. O. (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J. Mol. Biol.* **222**, 599–620
38. Dolezal, D., Jones, C. E., Lai, X., Brister, J. R., Mueser, T. C., Nossal, N. G., and Hinton, D. M. (2012) Mutational analysis of the T4 gp59 helicase loader reveals its sites for interaction with helicase, single-stranded binding protein, and DNA. *J. Biol. Chem.* **287**, 18596–18607

**Complexes of T4 Gp59 Helicase Loading Protein**