THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 42, Issue of October 19, pp. 39340–39349, 2001

Printed in U.S.A.

Building a Replisome Solution Structure by Elucidation of Protein-Protein Interactions in the Bacteriophage T4 DNA Polymerase Holoenzyme*[S]

Received for publication, May 30, 2001, and in revised form, August 9, 2001
Published, JBC Papers in Press, August 14, 2001, DOI 10.1074/jbc.M104956200

Stephen C. Alley‡§, Michael A. Trakselis‡, M. Uljana Mayer‡, Faoud T. Ishmael‡¶, A. Daniel Jones‡, and Stephen J. Benkovic‡‖

From the ‡Department of Chemistry, the Pennsylvania State University and ¶Department of Biochemistry and Molecular Biology, Hershey Medical Center, the Pennsylvania State University, University Park, Pennsylvania 16802

Assembly of DNA replication systems requires the coordinated actions of many proteins. The multiprotein complexes formed as intermediates on the pathway to the final DNA polymerase holoenzyme have been shown to have distinct structures relative to the ground-state structures of the individual proteins. By using a variety of solution-phase techniques, we have elucidated additional information about the solution structure of the bacteriophage T4 holoenzyme. Photocross-linking and mass spectrometry were used to demonstrate interactions between I107C of the sliding clamp and the DNA polymerase. Fluorescence resonance energy transfer, analytical ultracentrifugation, and isothermal titration calorimetry measurements were used to demonstrate that the C terminus of the DNA polymerase can interact at two distinct locations on the sliding clamp. Both of these binding modes may be used during holoenzyme assembly, but only one of these binding modes is found in the final holoenzyme. Present and previous solution interaction data were used to build a model of the holoenzyme that is consistent with these data.

Many biochemical processes require the formation of multiprotein complexes. Replicative DNA polymerases employ processivity factors (or sliding clamps) that are loaded onto primer-template DNA by ATP hydrolysis-dependent loading proteins (or clamp loading complexes). Further interactions with a variety of other proteins, including additional DNA replication, recombination, and repair proteins, demonstrate that a complicated network of often transient interactions is required for rapid and accurate copying of genetic information (1, 2). In bacteriophage T4, the core of this DNA replication system is the DNA polymerase (gp43),[1] the sliding clamp (gp45), and the clamp loading complex (gp44/62) (reviewed in Refs. 3–5). The torroidal, trimeric gp45 increases the processivity of gp43 by encircling the DNA and provides a topological link between the polymerase and the DNA. The 4:1 gp44/62 complex sequentially hydrolyzes four molecules of ATP (6–7) to load gp45 onto DNA and to chaperone gp43 onto the gp45·DNA complex (8). Likewise, prokaryotic systems use the γ-complex to load the β-clamp onto DNA and assemble the DNA polymerase III holoenzyme, whereas eukaryotic systems use the replication factor C complex to load PCNA onto DNA and assemble the DNA polymerase δ and ϵ holoenzymes (9–12).

Assembly of the bacteriophage T4 DNA polymerase holoenzyme is a very dynamic process. Several studies employing a variety of fluorescent techniques have helped to elucidate a 10-step kinetic holoenzyme assembly mechanism (7, 13–16). These investigations as well as photocross-linking experiments (17) have shown that the holoenzyme and intermediate complexes leading to the holoenzyme exist in a variety of conformations that are structurally distinct from the ground state structures of the individual proteins. Similar studies on the *E. coli* DNA polymerase III holoenzyme have shown that the proteins of this holoenzyme also undergo a variety of conformations during assembly of the holoenzyme (18–20). An accurate and informative structure of a DNA polymerase holoenzyme will require careful integration of high resolution x-ray and/or NMR data as well as solution data that provide protein-protein contact points and distances measured during and after holoenyzme assembly. Because of the highly dynamic nature of the assembly process, an accurate solution structure must start with rather than end with high resolution data of ground-state structures.

The x-ray crystal structures of the sliding clamps from bacteriophages T4 (21) and RB69 (22), *Escherichia coli* (23), and yeast (24) have been solved as well as the C terminus of the cell cycle checkpoint protein p21 bound to human PCNA (25). Despite extremely limited primary sequence similarity, all of these sliding clamps are circular, multimeric proteins with a central hole large enough to accommodate duplex DNA (26). Trimeric gp45 and PCNA and the dimeric β-clamp each have six domains with flexible connecting loops tethering these domains. In all of the x-ray crystal structures, the sliding clamps

* This work was supported by National Institutes of Heath Grants GM13306 (to S. J. B.), GM19492 (to S. C. A.), and DK19691 (to F. T. I.), and a fellowship from the Jane Coffin Childs Memorial Fund for Medical Research (to M. U. M). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at *http://www.jbc.org*) contains a movie of the Fig. 7 model in motion.

§ Current address: Chiron Corp., 201 Elliott Ave. West, Suite 150, Seattle, WA 98119.

‖ To whom correspondence should be addressed: Dept. of Chemistry, 414 Wartik Laboratory, the Pennsylvania State University, University Park, PA 16802. Tel.: 814-865-2882; Fax: 814-865-2973; E-mail: sjb1@psu.edu.

[1] The abbreviations used are: gp, gene product; EDC, (1-(3-dimethylamino)propyl)-3-ethyl carbodiimide); ESI, electrospray ionization; F_A, fluorescence of the acceptor in the absence of the donor; F_{AD}, fluorescence of the acceptor in the presence of the donor; F_D, fluorescence of the donor in the absence of the acceptor; F_{DA}, fluorescence of the donor in the presence of the acceptor; FPLC, fast protein liquid chromatography; FRET, fluorescence resonance energy transfer; HPLC, high performance liquid chromatography; IAEDANS, 5-((((2-iodoacetyl)amino)ethyl))amino)naphthalene-1-sulfonic acid; MS, mass spectrometry; PCNA, proliferating cell nuclear antigen; DMF, *N*,*N*-dimethylformamide; PCR, polymerase chain reaction.

are closed at all of the subunit interfaces. However, we have shown that in solution gp45 adopts a conformation where one subunit interface is open by about 20 Å (15), a gap that is almost large enough for duplex DNA to pass through. In contrast, the *E. coli* β-clamp appears to be closed in solution since the interior of the subunit interface is resistant to chemical modification (20). The β-clamp binds to DNA much tighter than does gp45 (the γ-complex is required to unload the β-clamp (27), whereas gp45 has an off rate from the DNA of 0.3 and 0.002–0.03 s^{-1} in the absence (7) and presence (8, 13–14) of gp43, respectively), perhaps reflecting a difference in the holoenzyme stability required for processive DNA replication of the much larger *E. coli* chromosome.

The x-ray crystal structure of the RB69 DNA polymerase has also been solved (28). The T4 and RB69 DNA polymerases are 63% identical (29), with only one region, recently demonstrated to be important for dimerization (30), having a lack of significant homology. The C terminus of gp43 is a largely unstructured extension and is absolutely essential for interaction with gp45; removal of the last six C-terminal amino acids had no effect on the polymerase activity or DNA binding ability, but interactions with gp45 and processivity are abolished (31). We have shown that a peptide corresponding to the C terminus of gp43 inhibits the holoenzyme assembly reaction by blocking gp45 and gp44/62 interaction, binds to gp45 with a K_D of 7 μM and a 1:1 peptide to gp45 monomer stoichiometry, and forms photocross-links with Ala-159 of gp45, an amino acid that lies on the gp45 subunit interface (31–32). This final result coupled with the observation that gp45 is not completely closed in the final holoenzyme but that one subunit interface is still open by 11 Å led us to propose that the C terminus of gp43 is inserted into the open subunit interface of gp45 (15–16, 33), consistent with a wide variety of solution evidence.

To refine our previous model of the bacteriophage T4 DNA polymerase holoenzyme, we have employed a variety of methods to provide additional solution data on the structure of this multiprotein complex. 1) Fluorescence-resonance energy transfer measurements between the C-terminal peptide of gp43 and several gp45 mutants demonstrate that there are two distinct locations on gp45 where the C-terminal peptide of gp43 can bind. 2) Biophysical characterization of a monomeric mutant of gp45 demonstrates that one of these locations is not found in the final holoenzyme, but this does not exclude the other location from participating in intermediate holoenzyme assembly steps. 3) Isothermal titration calorimetry measurements demonstrate that a gp43 mutant lacking the six C-terminal residues still interacts with gp45, but with an affinity 2 orders of magnitude weaker than with wild-type gp43. 4) Finally, photocross-linking followed by mass spectrometry was used to demonstrate that there are additional contact points between gp43 and gp45 beyond our previously demonstrated contact between the C terminus of gp43 and the subunit interface of gp45 (32).

EXPERIMENTAL PROCEDURES

Materials and Their Sources—DNA primers and substrates were synthesized and purified as described previously (33). Peptide 1 was synthesized by Anne Stanley (Hershey Medical School), purified by HPLC before use, and labeled with IAEDANS as described previously (32). 4-Azido-2,3,5,6-tetrafluorobenzyl amine, iodoacetic acid, succinimidyl ester, IAEDANS, *N*-(4-azidophenylthio)phthalimide, streptavidin, and β-octyl glucoside were from Molecular Probes (Eugene, OR). *N*-(4-Azidophenylthio)phthalimide is no longer available from this source, although 4-azidophenyl disulfide can place the same photocross-linking moiety on the protein. All restriction enzymes were from New England Biolabs (Beverly, MA). T4 genomic DNA and L-tryptophan were from Sigma. Trypsin (L-1-tosylamide-2-phenylethyl chloromethyl ketone-treated) was from U. S. Biochemical Corp. 1-(3-Dimethylamino)propyl)-3-ethyl carbodiimide, *N*-hydroxysuccinimide, and triethylamine were obtained from Aldrich. Streptavidin-horseradish peroxidase conjugate was from Life Technologies, Inc. All other chemicals were of analytical grade or better.

Proteins and Labeling—Wild-type gp45, gp44/62, and exonuclease-deficient gp43 were purified as described previously (34–36). The gp45 mutants were cloned into pET26b as described previously (33) using the following PCR primers: A, 5′-GCG GAA TTC CAT ATG AAA CTG TCT AAA GAT; B, 5′-GCG GAA TTC GGA TTC CTA TTA AAA ATC GTG GGT; C, 5′-CGC TCA ACA ATT TTT TTC CCG GCC GCC GAT CCG AGT ACA; D, 5′-TGT ACT CGG ATC GGC GGC CGG GAA AAA AAT TGT TGA GCG; E, 5′-GCA ATT TAC GAT TTG AAC GGT TTT CTC GG; F, 5′-AAA CTT CTG CTT TGG GCA AAA GGT AAA; G, 5′-CGC GCA GTT AAT GGT ACC ACG TGG GCG GAA GCA AAT ATT TCT GAC; H, 5′-GTC AGA AAT ATT TGC TTC CGC CCA CGT GGT ACC ATT AAC TGC GCG; I, 5′-CCT AAT AAA CCA ATT CCA TGG CCG GTA GCA TCT GC; J, 5′-GCA GAT GCT ACC GGC CAT GGA ATT GGT TTA TTA GG; K, 5′-GAA GAT TCT GCT CTG ACA CGC GTT AAA TGG TCT TTG ACT CTT GGT G; L, 5′-CAC CAA GAG TCA AAG ACC ATT TAA CGC GTG TCA GAG CAG AAT CTT C; M, 5′-ATC GTA ATT AAC GGT TGG AAT AAA GTA GAA GAT AGT GCA CTG ACC CGT GTT AAA; N, 5′-TTT AAC ACG GGT CAG TGC ACT ATC TTC TAC TTT ATT CCA ACC GTT AAT TAC GAT; O, 5′-CCT AAT AAA CCA TGT CCC TTC CCG GTA GCA; and P, 5′-TGC TAC CGG GAA GGG ACA TGG TTT ATT AGG. By using T4 genomic DNA as a template, overlap extension PCR of fragments AD and BC yielded the mutant *45* gene W199F that included a new *Nru*I site, and fragments AF and BE yielded the mutant *45* gene W92F that included a new *Bsi*EI site. By using pET26b-W199F as a template, overlap extension PCR of fragments AF and BE yielded the mutant *45* gene W92F/W199F. By using pET26b-W92F/W199F as a template, overlap extension PCR of fragments AH and BG yielded the mutant *45* gene Y39W/W92F/W199F that included a new *Bsa*AI site; fragments AJ and BI yielded the mutant *45* gene W92F/F109W/W199F that included a new *Nco*I site; fragments AL and BK yielded the mutant *45* gene W92F/Y165W/W199F that included a new *Mlu*I site, and fragments AN and BM yielded the mutant *45* gene W92F/F152W/W199F that included a new *Apa*LI site. Overlap extension PCR of fragments AP and BO using T4 genomic DNA as a template yielded the mutant *45* gene I107C that included a new *Bsm*FI site. Mutations were confirmed by DNA sequencing. The gp45 mutants were purified as described previously (33). The ATPase activity assay of gp44/62 in Table I was performed as described previously (33). Protein concentrations were determined optically based on extinction coefficients at 280 nm predicted from the primary amino acid sequence (37).

Photocross-linking and Mass Spectrometry—The trifunctional photocross-linker was prepared as follows.

Step 1, the hydrochloride salt of 4-azidotetrafluorobenzylamine was dissolved in 1 ml of DMF and added to *N*-*t*-boc-biocytin (23 mg; 4.9×10^{-5} mol; 1.0 eq), 1-(3-dimethylamino)propyl)-3-ethyl carbodiimide (11 mg; 5.8×10^{-5} mol; 1.2 eq), and *N*-hydroxysuccinimide (1.0 mg; 8.7×10^{-6} mol; 0.20 eq) in a small round bottom flask equipped with a stir bar. Triethylamine (17 μl; 1.2×10^{-4} mol; 2.5 eq) was added. The reaction was monitored by TLC (10% MeOH/CH_2Cl_2) and allowed to stir at room temperature overnight in the dark. The solvent was evaporated *in vacuo*, and the sample was taken up in ethyl acetate and washed twice with 5% trifluoroacetic acid (pH 1). The organic layer (one major compound by TLC) was dried over magnesium sulfate, filtered, and concentrated (MS (ESI+), *m*/*z* 675.2 ($[M + H]^+$, theoretical value, 675.3)).

Step 2, methylene chloride (2 ml) and trifluoroacetic acid (1 ml) were added to the residue, which dissolved immediately upon addition of trifluoroacetic acid. The reaction was monitored by TLC (10% MeOH/CH_2Cl_2 with a drop of triethylamine, stained with ninhydrin (0.1% in isopropyl alcohol)) and found to be complete at 2.5 h. The solvent was reduced *in vacuo*. The product was 95% pure by HPLC performed on a Waters (Milford, MA) HPLC system with a C18 column using gradient A at 1 ml/min: solvent A, 0.1% trifluoroacetic acid in water; solvent B, 0.1% trifluoroacetic acid in acetonitrile; 100% solvent A for 2 min, a 20-min linear gradient to 50% solvent B, and a 5-min linear gradient to 100% solvent B. The product eluted at 19 min (MS (ESI+), *m*/*z* 575.2 ($[M + H]^+$, theoretical value, 575.2)).

Step 3, half of the product of step 2 was dissolved in DMF. Iodoacetic acid (succinimidyl ester; 6.9 mg; 2.4×10^{-5} mol; 1.0 eq) and triethylamine (7 μl; until basic) were added. The reaction was complete after 4 h as monitored by HPLC (elution time of the product was 23 min using gradient A). The product was purified by preparative HPLC on a C18 column on a Beckman (Palo Alto, CA) HPLC system, using the following gradient at 4 ml/min: solvent A, 0.1% trifluoroacetic acid in water; solvent B acetonitrile; a 2-min linear gradient from 0 to 50% solvent B,

a 20-min linear gradient to 70% solvent B, and then 70% solvent B. The compound eluted between 25 and 28 min (MS (ESI+), m/z 743.1 ($[M + H]^+$, theoretical value, 743.1)).

This product was then used to label gp45. The gp45 mutant I107C (1 ml of 81 μM trimers) was dialyzed *versus* 20 mM HEPES, pH 7.0, 50 mM NaCl, 1 mM EDTA, and 10% glycerol (storage buffer) to remove β-mercaptoethanol used for storage. The trifunctional photocross-linker (100 μl of 4.5 mM) dissolved in DMF was added, and the mixture nutated in the dark for 6 h. Conjugated I107C was purified away from unconjugated photocross-linker by FPLC using the gradient described below. The trifunctional photocross-linkers we have constructed are difficult to dialyze away from the proteins we have investigated (at least 3 buffer changes are often required).

Analytical scale photocross-linking reactions (20 μl total volume) contained 1 μM conjugated I107C in the presence or absence of 1 μM gp44/62, 1 μM gp43, 1 μM DNA (a self-concatenating primer-template), and 1 mM ATP in 20 mM Tris, pH 7.5, 150 mM sodium acetate, and 10 mM magnesium acetate. Photocross-linking reactions were performed and analyzed as described previously (32, 38).

Alternatively, gp45 I107C (500 μl of 50 μM trimers) was dialyzed *versus* storage buffer, and *N*-(4-azidophenylthio)phthalimide (50 μl of 10 mM) dissolved in DMF was added and the mixture nutated in the dark at 4 °C for 1.5 h. Following conjugation, the mixture was passed through Sephadex G-25 (Sigma) to remove any unincorporated photocross-linker.

Preparative scale photocross-linking reactions (1 ml total volume) contained 20 mM Tris, pH 7.5, 150 mM sodium acetate, 10 mM magnesium acetate, 1 mM ATP, 2.5 μM 4-azidophenylthio-I107C, 2 μM gp44/62, 2 μM gp43, 2 μM DNA (biotinylated forked primer-template (33)), and 2.5 μM streptavidin. All components except 4-azidophenylthio-I107C and ATP were premixed in a microcentrifuge tube. Under safe light illumination, 4-azidophenylthio-I107C and ATP were sequentially added. Small aliquots (20 μl) of the reaction mixture were then each placed into 50 microcentrifuge tubes in a standard 80-hole microcentrifuge tube rack to improve photocross-linking yields. Photocross-linking was carried out as described previously (32, 38) for 12 min. The individual aliquots were recombined, and β-octyl glucoside was added to a final concentration of 0.1%. The buffer was then changed to storage buffer plus 0.1% β-octyl glucoside using an Amicon Centricon-10 (Millipore, Bedford, MA), which also removed ATP and ADP. This mixture was then chromatographed on an Amersham Pharmacia Biotech FPLC system with a 5/5 Mono-Q anion-exchange column using the following gradient at 1 ml min^{-1}: solvent A, 20 mM Tris, pH 7.5, and 1 mM EDTA; solvent B, 20 mM Tris, pH 7.5, 1 M NaCl, and 1 mM EDTA; 10% solvent B for 5 min, a 23-min linear gradient to 35% solvent B, a 4-min linear gradient to 100% solvent B, 100% solvent B for 3 min, and a 2-min linear gradient to 10% solvent B. gp44/62 eluted in the void volume, uncross-linked gp43 eluted at 14 min, uncross-linked gp45 eluted at 22 min, and gp43-gp45 cross-links eluted at 24 min. Both uncross-linked and cross-linked gp43 were collected, and β-octyl glucoside and dithiothreitol were immediately added to a final concentration of 0.1% and 2.5 mM, respectively. The sample with reduced, cross-linked gp43 and gp45 was then concentrated using a Centricon-10, rechromatographed on the FPLC using the same gradient, and reduced, cross-linked gp43 was collected. The buffer of the gp43 mixtures was then changed to 0.1 M ammonium bicarbonate and 0.02% β-octyl glucoside using a Centricon-10, and the volume was reduced to about 200 μl. L-1-Tosylamide-2-phenylethyl chloromethyl ketone-treated trypsin (5 μg) was then added, and the mixtures were incubated at 37 °C for 16 h. Formic acid (2 μl) was added to stop the reaction.

HPLC electrospray ionization-mass spectrometry analyses were performed using a model 1100 HPLC (Hewlett-Packard, Palo Alto, CA) interfaced to a Mariner mass spectrometer (Perspective Biosystems, Framingham, MA). Spectra were obtained in a positive ion mode. Aliquots of the above tryptic digest were injected onto a 1 × 50-mm BetaBasic C18 column (Keystone Scientific, Bellefonte, PA) and eluted using a pre-injection split (pump flow rate = 0.5 ml/min; column flow rate was ~0.05 ml/min) and a solvent gradient based upon solvent A = 0.15% formic acid in H_2O and solvent B = 0.15% formic acid in acetonitrile. The gradient consisted of 95% solvent A, 5% solvent B from 0 to 2 min followed by a linear gradient to 30% solvent A, 70% solvent B from 22 to 40 min.

Fluorescence and Distance Measurements—Steady-state fluorescence measurements were made using an ISA (Edison, NJ) FluoroMax-2 spectrofluorimeter using the following parameters: 280 nm excitation wavelength, 5 (acceptor sensitization) or 3.75 nm (donor quenching) slit width, 0.5-s integration time, 5-mm excitation path length, and 10-mm emission path length at 25 °C. Emission was measured at the maximum emission wavelength of the gp45 mutants (given in Table II) for donor quenching or at 492 nm for acceptor sensitization and was corrected for dilution and background fluorescence. The extinction coefficient of the IAEDANS acceptor (ϵ_A) was found to be 1287 M^{-1} cm^{-1} at 280 nm. The extinction coefficient of the tryptophan donor (ϵ_D) was assumed to be 5700 M^{-1} cm^{-1} at 280 nm, yielding an ϵ_A/ϵ_D value (used below) of 0.2258.

To measure quantum yield, the following formula was used: $\phi_D = \phi_R(F_D/F_R)(A_R/A_D)$, where ϕ, F, and A are quantum yield, fluorescence, and absorbance, respectively. The subscripts designate reference (R) and donor (D); L-tryptophan was used as the reference (15). The quantum yields for the various gp45 mutants are given in Table III.

To determine R_0, the following equation was used: $R_0 = 0.211(\phi_D\kappa^2\eta^{-4}J)^{1/6}$, where κ^2 is an orientation factor assumed to be 2/3 (39); η is the refractive index of the medium (assumed to be 1.4), and J is the overlap integral between the fluorescence emission spectrum of the donor and absorption spectrum of the acceptor according to the formula $J = \int F_n(\lambda) \times \epsilon_A(\lambda) \times \lambda^4 \times \delta\lambda$, where $F_n(\lambda)$ is the fluorescence intensity of the donor (in the absence of the acceptor) as a fraction of the total integrated intensity; $\epsilon_A(\lambda)$ is the extinction coefficient of the acceptor; and λ is the wavelength of overlap. The overlap integrals and R_0 values for the various gp45 mutants are given in Table III.

Fluorescence titrations were performed as described previously (32). The equation $F_{obs} = F_i - ((T \times \Delta F)/(T + K_D))$ was used to fit the data (see Fig. 2 for example), and the relative fluorescence value achieved upon saturation ($F_{sat} = F_i - \Delta F$) was then used to determine energy transfer efficiency using the formulas $E = 1 - (F_{DA}/F_D)$ and $E = (\epsilon_A/\epsilon_D)(F_{AD}/F_A - 1)$ for donor quenching and acceptor sensitization, respectively. The distance between the gp45 tryptophan and IAEDANS on peptide 1 can then be calculated using the formula $R = R_0\,(1/E - 1)^{1/6}$.

Analytical Ultracentrifugation—Analytical ultracentrifugation measurements were made on the monomeric gp45 mutant W92F/Y165W/W199F using a Beckman Instruments (Fullerton, CA) XL-I Analytical Ultracentrifuge in absorbance mode at 25 °C (33). Samples were thoroughly dialyzed into 25 mM potassium phosphate, pH 7.4, and 200 mM KCl immediately before analysis. For velocity measurements, the rotor speed was 36,000 rpm. For equilibrium measurements, the rotor speeds were 22,000, 29,000, and 34,000 rpm. Data acquisition and processing were performed as described previously (33).

Isothermal Titration Calorimetry—The binding of the monomeric gp45 mutant W92F/Y165W/W199F to gp43 was investigated using a VP-ITC MicroCalorimeter (Microcal, Northampton, MA). The two protein solutions were dialyzed in the same chamber *versus* 20 mM Tris, pH 7.5, 150 mM potassium acetate, and 10 mM magnesium acetate and degassed immediately before analysis. The gp43 solution (1.4 ml of 20 μM solution) was placed into the sample cell, and the gp45 solution (300 μl of a 123 μM solution, expressed as monomers) was placed into the syringe. Alternatively, dialysis buffer was placed in the sample cell, and the above gp45 solution was titrated against buffer to obtain the heat of dilution. The following parameters were used in the titration: 25 °C, 10-μl injections, and 4 min between injections with stirring at 310 rpm. These titration data were fit using Microcal Origin version 5.0 SR2 to yield thermodynamic parameters. The equation used is $Q = (nM\Delta HV/2)(A - (A^2 - 4X/nM)^{1/2})$, where A is the quantity $1 + X/nM + 1/nKM$; Q is the heat content of the solution; n is the number of binding sites; M is the concentration of gp43; ΔH is the molar heat of gp45 binding; V is the cell volume; X is the concentration of gp45; and K is the association constant.

RESULTS AND DISCUSSION

Photocross-linking and Mass Spectrometry—Mapping protein-protein interaction locations by photocross-linking and mass spectrometry provides an opportunity to establish solution contacts that can be used to develop models of multiprotein complexes. We have shown previously (32) by mass spectrometry that the C terminus of gp43 (peptide 1, NH_2-CSLDFLFG-COOH, labeled with a photocross-linking moiety) interacts with the subunit interface of gp45 by demonstration of a cross-link to gp45 A159. To establish another point of interaction, we conjugated two different photocross-linkers to the I107C position of gp45, a location that has been shown previously (38) to cross-link to gp44. Our previous model of the holoenzyme places this amino acid close to many amino acids of gp43 (two of the three gp45 monomers can contact gp43 from I107C). Wild-type gp45 contains no cysteines, so mutagenesis can be

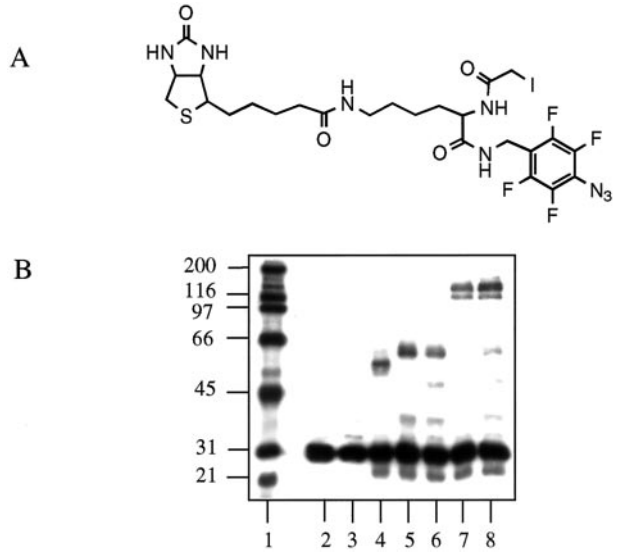

FIG. 1. **Photocross-linking reaction of the holoenzyme detected with biotin-specific chemiluminescence.** gp45 was labeled at the I107C position with the biotinylated trifunctional photocross-linker (structure shown in *A*) and cross-linked to gp44/62 and gp43. *B*, lanes contain the following components: *1,* biotinylated molecular weight markers; *2,* labeled gp45 alone not exposed to light; *3,* labeled gp45, gp44/62, gp43, ATP, and DNA not exposed to light; *4,* labeled gp45 alone exposed to light; *5,* labeled gp45 and gp44/62 exposed to light; *6,* labeled gp45, gp44/62, ATP, and DNA exposed to light; *7,* labeled gp45 and gp43 exposed to light; and *8,* labeled gp45, gp44/62, gp43, ATP, and DNA exposed to light. Following photocross-linking, mixtures were separated by SDS-PAGE, proteins blotted onto a nitrocellulose membrane, and the membrane probed with a streptavidin-horseradish peroxidase conjugate to visualize biotinylated species (32, 38).

used to place a single cysteine at any location in a gp45 monomer.

We first conjugated a trifunctional photocross-linker (photocross-linker structure shown in Fig. 1*A*) to the I107C position of gp45. This photocross-linker is a modification of a previous version (38) containing the following functional groups: 1) a cysteine-reactive iodoacetamide, 2) a perfluorinated arylazide, and 3) the affinity probe biotin. Our previously reported version had a disulfide linkage between the cysteine-reactive group and the biotin, which after photocross-linking and reduction would effect biotin transfer from the original protein to the photocross-linked protein. The present version cannot be cleaved and allows for rapid investigation of all photocross-linked species by use of a Western blot that is probed for the presence of biotin with a streptavidin-horseradish peroxidase conjugate.

As shown in Fig. 1*B*, photocross-links to gp43 and gp44/62 can be observed depending on the holoenzyme components present. *Lane 2* contains labeled gp45 in the absence of other proteins and without exposure to light, whereas *lane 3* also contains gp44/62, gp43, ATP, and DNA (a self-concatenating primer-template). No photocross-links are observed in either of these cases. In *lane 4*, labeled gp45 alone is exposed to light, and a small amount of photocross-linked dimers is observed. gp44/62 is added in *lane 5*, and the primary species observed is a gp45-gp44 photocross-link. Addition of ATP and DNA to gp44/62 in *lane 6* does not change the composition appreciably. In *lane 7*, labeled gp45 and gp43 are exposed to light, whereas in *lane 8* gp44/62, ATP, and DNA were included to favor formation of the holoenzyme. In both cases, two distinct photocross-linked species are observed in about equal intensity corresponding to the mobility of gp43 plus one monomer of gp45. Running the gel longer resolves the lower mobility band into two bands (data not shown), suggesting that as many as three photocross-linked species are present. Cross-links between two proteins will have different mobilities if the cross-linking points are separated by many amino acids; cross-links that form near the end of the proteins will have slower mobilities than those forming near the middle of the proteins. Since gp45 is labeled at only one location, the points of photocross-linking on gp43 must be different in the two bands.

Such a distribution of photocross-linked species has been observed before between gp45 and gp44/62 (17). As many as four photocross-linked species (three gp44 species and one gp62 species) were observed depending on the location of the photocross-linker on gp45 and whether the photocross-linking reaction was performed in the presence or absence of ATP and DNA. The asymmetry between a homotrimer (gp45) and heteropentamer (there are four gp44 subunits and one gp62 subunit per complex) was suggested to allow photocross-links to form at different sites on individual gp44 subunits and yield photocross-links that each possessed unique mobility (17). The formation of multiple gp45-gp43 photocross-link species is fully consistent with our holoenzyme model that places two gp45 subunits in contact with gp43. Observation of the same set of photocross-links in *lanes 7* and *8* (in the absence and presence of gp44/62, respectively) is also consistent with our previous observations that gp45 bound to gp43 alone is in the same conformation as gp45 bound to gp43 in the holoenzyme (with DNA running through the gp45 ring), with the subunit interface open by about 11 Å (15, 16). However, the gp45 conformation in these two complexes differs from the gp45 ground-state structure.

To map the location of one or more of the photocross-linked species, we turned to the very simple photocross-linker *N*-(4-azidophenyl)phthalimide. When conjugated to gp45 at I107C, yields of gp45-gp43 photocross-links up to ~50% were obtained under conditions favoring holoenzyme formation (gp45 in the presence of gp44/62, gp43, ATP, and a forked primer-template DNA). Separation of a preparative scale photocross-linking reaction by FPLC yielded a new, late-eluting peak that upon reduction with dithiothreitol could be rechromatographed to isolate a modified gp43 species and gp45. Modified gp43 so derived from reduced photocross-links contains a chemical modification due to the photocross-linking reaction that would increase the mass of the protein by 123 Da. Proteolytic digestion with trypsin followed by analysis by HPLC electrospray ionization mass spectrometry yielded one ion consistent with such a photocross-link, the triply charged 758.63 m/z ion identical to the mass of gp43 amino acids 675–693 plus 123 (calculated 758.67; Fig. 2). This fragment contains one missed tryptic cleavage, which may be the result of photocross-linking at or near Arg-682 such that this site is no longer recognized by trypsin. This ion was not observed in a tryptic digest of gp43 that was present in the photocross-linking reaction but was not incorporated into gp45-gp43 photocross-links. This ion is not consistent with the mass of any wild-type gp43 tryptic fragment with or without missed tryptic cleavages in either protonated or sodiated form. The location of modification is rotationally consistent with the previous model of the holoenzyme we have proposed where the C terminus of gp43 is inserted into the gp45 subunit interface (Fig. 3).

Activity of gp45 Mutants Used in Fluorescence Studies—Previously, we have demonstrated (33) an interaction between the C terminus of gp43 and the subunit interface of gp45 by photocross-linking and mass spectrometry. To improve the resolution of this contact information, we employed fluorescence-resonance energy transfer (FRET) measurements to measure the distance between the C-terminal peptide of gp43 and several mutants of gp45. FRET is a powerful spectroscopic tool

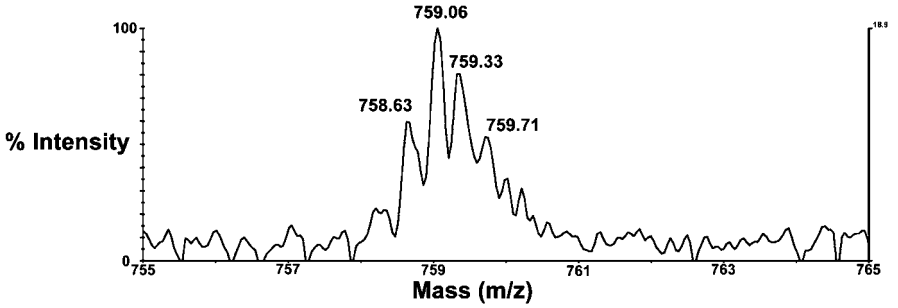

FIG. 2. **Electrospray ionization-mass spectrum of photocross-linked gp43 tryptic fragment.** gp45 was labeled with *N*-(4-azidophenylthio)phthalimide and preparatively photocross-linked to gp43 in the presence of gp44/62, ATP, and DNA. gp45-gp43 photocross-links were isolated by FPLC and then reduced with dithiothreitol to liberate modified gp43, which was then isolated by a second FPLC separation. Following a tryptic digest of modified gp43, peptides were separated by HPLC and analyzed by electrospray ionization mass spectrometry. The triply charged 758.63 m/z ion is consistent with the remains of the reduced photocross-linker attached to the gp43 tryptic fragment of amino acids 675–693 (calculated m/z 758.67).

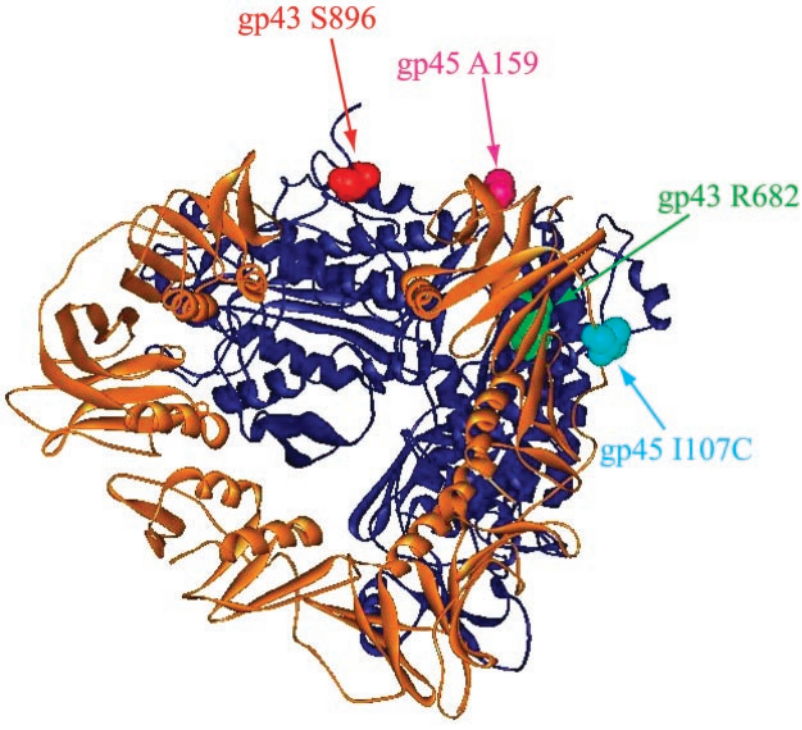

FIG. 3. **Model of holoenzyme showing locations of interactions as determined by mass spectrometry.** gp45 is shown in *orange,* and gp43 is shown in *blue.* The previously determined interaction (32) between the C terminus of gp43 (S896 in *red*) and the subunit interface of gp45 (A159 in *pink*) is shown along with the interaction between the interdomain loop of gp45 (I107C in *cyan*) and a peptide (675–693, with Arg-682 in *green*) of gp43 containing Arg-682 as the likely site of interaction.

that allows atomic resolution distance measurements to be made in solution under physiologically relevant conditions with an effective range of 10–75 Å (40). Peptide 1 was labeled with the fluorescence acceptor IAEDANS, a conjugate we have used previously to measure binding affinity and stoichiometry for the interaction with gp45 (32). The fluorescence donor was tryptophan from gp45, providing an effective range of about 10–33 Å. Because gp45 contains two wild-type tryptophans (Trp-92 and Trp-199), mutagenesis was used to construct a series of single-tryptophan mutants (Table I) to simplify analysis and provide distance measurements from a variety of locations on gp45 (see Fig. 4). These mutants were first tested for their ability to stimulate the ATPase activity of gp44/62, a measure of their ability to form a productive holoenzyme complex.

With the exception of W92F/Y165W/W199F, all gp45 mutants were found to stimulate the ATPase activity of gp44/62 upon addition of DNA (Table I). The rate of ATP hydrolysis by gp44/62 in the presence of gp45 provides a basal rate of ATPase activity. Addition of DNA stimulates the rate of ATP hydrolysis, with the absolute rate limited by the off rate of gp45 from DNA (7). Mutations that increase the DNA off rate (such as P108G (32)) cause increases to as much as 800 nM s^{-1}, whereas mutations that decrease the DNA off rate result in lower rates. Only the complete lack of stimulation of the ATPase activity indicates the absence of a functional interaction with gp44/62 and failure in loading gp45 onto DNA. Confirmation that gp45 mutations are not deleterious for holoenzyme formation comes from subsequent addition of gp43; formation of a functional holoenzyme will cause a shut down in the ATPase rate due to the slower off rate of the holoenzyme complex and a return to

TABLE I
ATPase activities of gp45 mutants

Mutant	gp45 + gp44/62	gp45 + gp44/62 + DNA	gp45 + gp44/62 + DNA + gp43
		nM s^{-1}	
Wild type	20.4	273	22.7
W199F	13.9	224	9.6
W92F	12.9	215	13.4
Y39W/W92F/W199F	20.3	109	28.0
W92F/F109W/W199F	19.3	332	21.7
W92F/Y165W/W199F	15.5	18	14.8
W92F/F152W/W199F	37.5	354	34.8

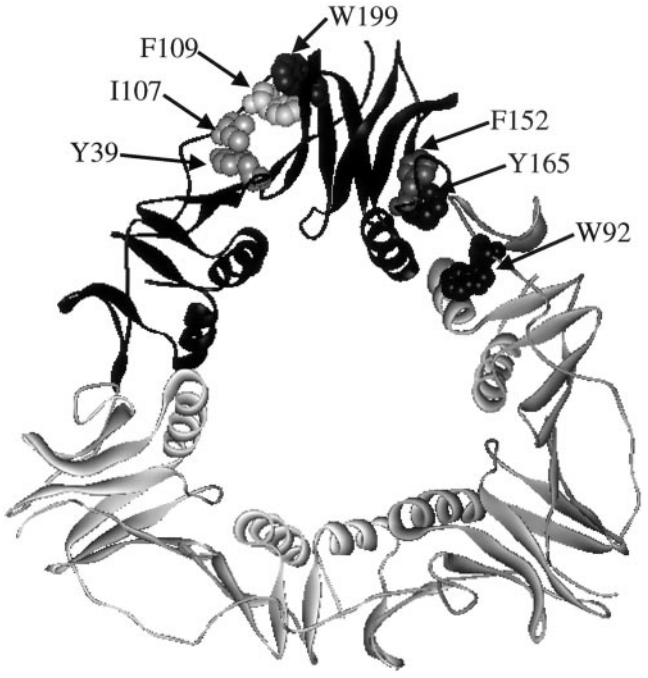

FIG. 4. **X-ray crystal structure of bacteriophage T4 gp45 (21) and location of mutations.** Photocross-linkers were attached to the I107C mutant, and fluorescent measurements were made using mutants with single tryptophans placed at Trp-92, Trp-199, Tyr-39, Phe-109, Tyr-165, or Phe-152 (wild-type tryptophans were mutated to phenylalanine).

the basal rate. The W92F/Y165W/W199F mutant does not display either a stimulation of gp44/62 ATPase activity upon addition of DNA nor a shut down upon addition of gp43. As shown below, the reason for this behavior is due to this mutant being monomeric.

FRET and Distance Measurements—To measure distances between gp45 tryptophans and IAEDANS on peptide 1, two series of measurements were made. In the first, 1 μM gp45 (concentration expressed as monomers) was placed in a fluorescence cuvette and peptide 1 (with and without conjugation to IAEDANS) titrated into this solution. By measuring the quenching of gp45 tryptophan fluorescence (at the fluorescence emission maxima given in Table II) upon excitation at 280 nm, a binding curve of relative fluorescence *versus* titrant concentration was constructed (Fig. 5*A*). Alternatively, 1 μM peptide 1-IAEDANS was placed in the fluorescence cuvette, and IAEDANS sensitization at 492 nm (excitation at 280 nm) monitored as gp45 was titrated into this solution (Fig. 5*B*). These titration curves were fit to the equation $F_{obs} = F_i - ((T \times \Delta F)/(T + K_D))$, where F_{obs} is the observed relative fluorescence; F_i is the initial relative fluorescence; T is the concentration of the titrated species in the cuvette; ΔF is the total change in relative fluorescence achieved upon saturation; and K_D is the dissociation constant. The value $F_i - \Delta F$ yields F_{sat}, the relative fluorescence value achieved upon saturation. A third measurement in which peptide 1 without IAEDANS was titrated into a solution of gp45 was made to observe the behavior of the donor tryptophan in the absence of the acceptor. K_D values for these titrations are reported in Table II as well. As shown in Table II, all of the mutants bind the peptide with micromolar affinities with K_D values that are in reasonable agreement between the donor quenching (in the presence and absence of acceptor) and acceptor sensitization methods. The concentration of gp45 is reported in monomers to achieve agreement in the K_D values and demonstrates the 1:1 peptide to gp45 monomer stoichiometry that was observed for wild-type gp45 (32).

The F_{sat} values were then used to calculate the individual tryptophan-IAEDANS energy transfer efficiencies and distances. For donor quenching, the fluorescence transfer efficiency (E) is defined as $E = 1 - (F_{DA}/F_D)$, where F_{DA} and F_D are the relative fluorescence values achieved upon saturation of the donor (D) in the presence and absence, respectively, of the acceptor (A). For acceptor sensitization, $E = (\epsilon_A/\epsilon_D)(F_{AD}/F_A - 1)$, where ϵ_A and ϵ_D are the extinction coefficients of the acceptor and donor, respectively, at the donor excitation wavelength, and F_{AD} and F_A are the relative fluorescence values achieved upon saturation of the acceptor in the presence and absence, respectively, of the donor. The F_A measurement was made using a gp45 mutant containing no tryptophans (W92F/V163C/W199F) that we have used previously (15). The energy transfer values are related to the distance (R) between donor and acceptor according to the relationship $R = R_0\,(1/E - 1)^{1/6}$, where R_0 is the distance at which the transfer efficiency is 50% (given in Table III; see "Experimental Procedures" for calculation of R_0 values). The distances obtained for donor quenching and acceptor sensitization are in very close agreement for all mutants except Y39W/W92F/W199F, where donor quenching (<11 Å) and acceptor sensitization (20 Å) yield significantly different values.

These measured distances were compared with the calculated distances for our holoenzyme model as shown in Table III. To make this model, three sets of FRET measurements were obtained between the donor Trp-92 and S158C, V163C, and T168C (independently labeled with the acceptor coumarin) during and after the holoenzyme assembly process (15, 16). These measurements allowed the triangulation of the gp45 conformation in the holoenzyme (16). Docking gp43 onto gp45 using known points of interaction (32) completed the model (16). In this model, the distances were calculated between the Cα atoms of Ala-159 of gp45 and the amino acids corresponding to the location of the tryptophan in the mutants. The Ala-159 location corresponds to the amino acid that was found to cross-link to a conjugate of peptide 1 and a photocross-linking moiety placed at the same location as IAEDANS in the present study (32).

Unfortunately, the measured and calculated distances were not in good agreement (see Table III). For example, the measured donor quenching distance for the Y39W/W92F/W199F mutant (<11 Å) was in worse agreement with the model (calculated distance of 27 Å) than the measured acceptor sensitization distance (20 Å). Attempting a global fit to these measured distances to a single peptide-binding site yielded high residual errors for any solution (data not shown).

One explanation for this behavior is that there are two binding sites on gp45 for peptide 1 and that the relative affinities for the two sites are similar and can be modified by mutagenesis to favor one site over the other. The Y39W/W92F/W199F mutant provides support for this suggestion; in the presence of excess gp45, peptide 1 partitions to the site with the highest affinity and yields a long distance (20 Å), whereas in the presence of excess peptide 1, the secondary binding site with a slightly weaker affinity is occupied as well and yields a short distance (<11 Å). The energy transfer efficiency of this mutant in the presence of excess peptide (0.9911) is nearly greater than 1 (theoretically not possible) for a model with three sites per gp45 trimer but would be much more consistent for a model with six sites per gp45 trimer. The W92F/Y165W/W199F mutant, which we show below to be monomeric, yielded distances that were very close to one another by both methods (20–21 Å) and with an affinity identical to that for the other mutants. Since this mutant would not have a subunit interface, we would not expect it to bind peptide 1 with high affinity. The presence

TABLE II
Donor quenching and acceptor sensitization values for single—tryptophan gp45 mutants

Mutant	Emission maximum	Donor quenching				Acceptor sensitization	
		F_{sat} (F_{DA})	K_D[a]	F_{sat} (F_D)	K_D[b]	F_{sat} (F_{AD})	K_D[c]
	nm		*μM*		*μM*		*μM*
W199F	333	0.5684	3.36 ± 0.11	0.8400	4.68 ± 0.44	6.6422	2.65 ± 0.23
W92F	346	0.1868	6.11 ± 0.20	1.2316	3.25 ± 0.19	9.6619	4.32 ± 0.47
Y39W/W92F/W199F	341	0.0063	6.20 ± 0.07	0.7017	10.47 ± 0.62	8.3686	10.30 ± 1.16
W92F/F109W/W199F	356	0.1416	1.11 ± 0.05	0.4669	0.94 ± 0.03	7.8404	2.35 ± 0.12
W92F/Y165W/W199F	331	0.4971	6.41 ± 0.28	0.9989[d]	NA[a,e]	6.3796	2.03 ± 0.12
W92F/F152W/W199F	332	0.4691	5.59 ± 0.18	1.0083[d]	NA[d,e]	8.4267	2.73 ± 0.13
W92F/V163C/W199F	NA[f]	NA[f]	NA[f]	NA[f]	NA[f]	2.2525[g]	3.12 ± 0.16

[a] K_D values were determined from titrating peptide I-IAEDANS into a gp45 solution.
[b] K_D values were determined from titrating peptide 1 into a gp45 solution.
[c] K_D values were determined from titrating gp45 into a peptide 1-IAEDANS solution.
[d] Since these F_{sat} values are very close to 1, excess peptide 1 was added to these mutants rather than performing a full titration; a K_D value was therefore not determined.
[e] NA, not applicable.
[f] This mutant does not contain any tryptophan and is therefore nonfluorescent. It was only used to determine F_A and K_D.
[g] F_A.

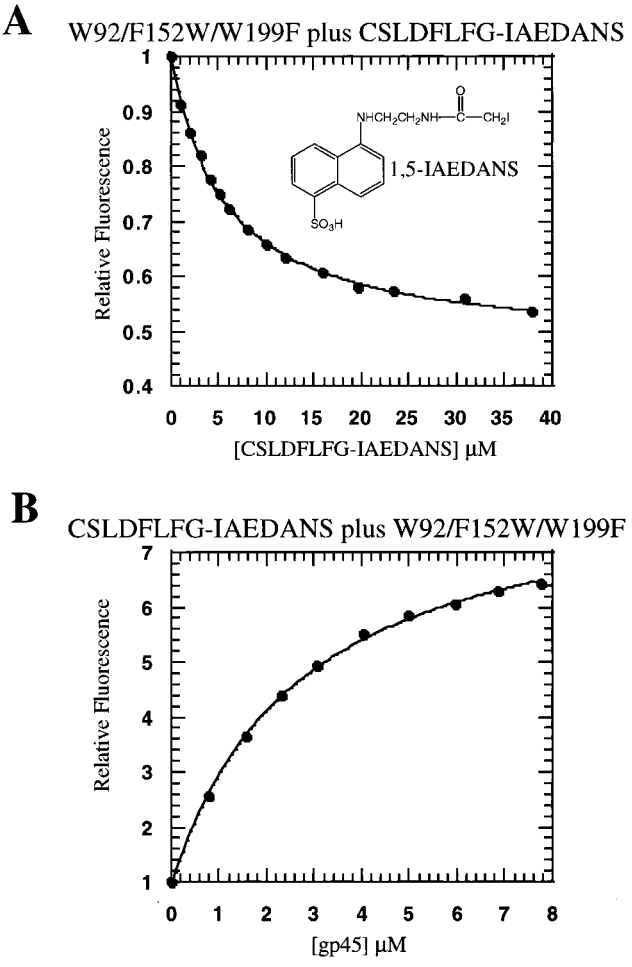

FIG. 5. **Fluorescence titrations of gp45 and peptide 1-IAEDANS and structure of IAEDANS.** *A*, 1 μM gp45 mutant W92F/F152W/W199F was placed in a fluorescence cuvette and peptide 1-IAEDANS added. The gp45 tryptophan was excited at 280 nm and donor quenching measured at 332 nm. *B*, 1 μM peptide 1-IAEDANS was placed in a fluorescence cuvette and the gp45 mutant W92F/F152W/W199F added. The gp45 tryptophan was excited at 280 nm and acceptor sensitization measured at 492 nm. Normalized fluorescence *versus* the concentration of the titrated species was fit to yield K_D and saturating fluorescence values as reported in Table II.

of a second binding site that does not depend on gp45 oligomerization and has an affinity close to that of the site at the subunit interface is consistent with these observations. The measured distance places this second site near the interdomain connecting loop or the interdomain interface.

Characterization of the Monomeric gp45 Mutant—The W92F/Y165W/W199F mutant previously shown above to be unable to stimulate the ATPase activity of gp44/62 upon addition of DNA was investigated by analytical ultracentrifugation to determine its oligomeric state. Concentrations between 2 and 13 μM monomers were found to have sedimentation coefficients of 2.67 ± 0.15 S (average ± S.D. for eight measurements at four concentrations with no systematic data trend) as measured by velocity sedimentation. Equilibrium sedimentation measurements were consistent with a monomer; however, nonideal behavior caused the observed molecular mass to decrease below the monomer value at high rotor speed and prevented the global calculation of a molecular mass at several rotor speeds and protein concentrations. Both of these results are consistent with a monomeric state; wild-type gp45 has a concentration-dependent sedimentation coefficient of 3.9–3.4 S (decreasing with increasing protein concentration) in this concentration range and an equilibrium sedimentation molecular mass of 69,000 Da (33). Mutation of the PCNA location analogous to Tyr-165 has been shown to result in monomerization as well (41).

By using fluorescence measurements, we have shown previously (15) that gp43 and gp45 interact with a K_D of 50 nM. Because these measurements were made by observing FRET across the gp45 subunit interface, we were unable to use this method to determine the affinity of gp43 for the monomeric gp45 mutant. Instead, isothermal titration calorimetry was used. Placing 123 μM W92F/Y165W/W199F monomers in the syringe and titrating into a solution of 20 μM gp43, the binding curve shown in Fig. 6 was obtained. Titration of this same W92F/Y165W/W199F solution into buffer was used to determine the heat of dilution (−1.4 kcal/mol injectant) that was then subtracted from each injection. The dilution-subtracted titration curve was fit (see "Experimental Procedures") to yield a stoichiometry of 0.781 ± 0.003, a binding enthalpy (ΔH) of −14.0 ± 0.1 kcal/mol, a binding entropy (ΔS) of −19.8 cal/mol K, and a K_D of 1.10 (1.06–1.13) μM. The error limits on the K_D measurement are unsymmetrical because the equations used for fitting solve for K_A, the inverse of K_D. The analogous experiment to determine the K_D for wild-type gp45 by this method was unsuccessful because an extremely large heat of dilution (about −22 kcal/mol injectant) overwhelmed the much smaller heat of interaction (about −3 kcal/mol injectant). Because of the high gp45 concentrations, dissociation of gp45 trimers into monomers was not observed.

TABLE III
Tryptophan to IAEDANS distance parameters and model distances

Mutant	Donor	Φ_D	J ($\cdot 10^{13}$)	R_0	Measured R		Calculated distance
					Donor quenching	Acceptor sensitization	
			M^{-1} cm^{-1} $(nm)^4$	Å	Å		Å
W199F	Trp-92	0.1437	5.71	22	25	23	18
W92F	Trp-199	0.1617	5.68	23	17	19	22
Y39W/W92F/W199F	Y39W	0.1316	5.67	22	<11	20	27
W92F/F109W/W199F	F109W	0.1052	5.38	21	18	20	29
W92F/Y165W/W199F	Y165W	0.0866	5.50	20	20	21	15
W92F/F152W/W199F	F152W	0.0488	5.33	18	18	17	9

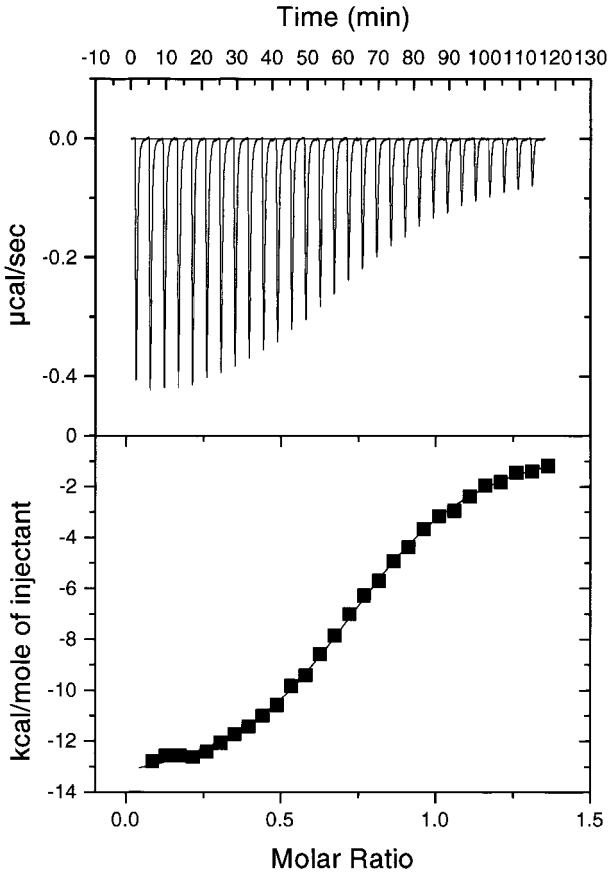

FIG. 6. **Isothermal titration calorimetry measurement of the interaction of the monomeric gp45 mutant W92F/Y165W/W199F and gp43.** gp45 was titrated into gp43 to yield the upper curve (*inset* shows gp45 titrated into buffer to yield the heat of dilution). Data were converted to heat per mol injectant (*lower curves*) and fit to yield thermodynamic parameters.

The affinity of the monomeric gp45 mutant for gp43 (1.1 μM) is very similar to the affinity for peptide 1 (2.7–5.6 μM), suggesting that gp43 can make no additional contacts with this mutant beyond the C-terminal interactions in this binding mode. This is in contrast to the 2 orders of magnitude stabilization gained by gp43 over the C-terminal peptide for interaction with trimeric gp45 (50 nM *versus* 7 μM), suggesting that there are significant additional contacts made beyond the C terminus of gp43 in the binding mode found in the final holoenzyme complex. The peptide-binding site on the monomeric mutant will therefore not lead to formation of the holoenzyme.

CONCLUSIONS

By using the solution interaction data described above, we have refined our original holoenzyme model to yield the one shown in Fig. 7 (and see the Supplemental Material for movie) (16). The important features of this model as well as the solution evidence for their support are as follows. 1) One subunit interface of gp45 is open by about 11 Å, consistent with our previous result that measured the distance across the gp45 subunit interface as it is assembled into the holoenzyme (15, 16). 2) The C terminus of gp43 is inserted into the gp45 subunit interface, consistent with previous photocross-linking data (32) as well as providing a steric block to complete gp45 ring closure. 3) A contact point is made between I107C on gp45 and near R682 on gp43, consistent with the present photocross-linking results and with the rotational orientation of gp45 relative to gp43 required for an interaction between the C terminus of gp43 and the subunit interface of gp45. 4) gp45 and gp43 make contact on two subunits of gp45, as required by the observation of multiple photocross-linked species with differing gel mobility. This mode of contact allows additional stabilizing interactions, in accord with the present isothermal titration calorimetry and previous fluorescence observation (15) of a 2 orders of magnitude stabilization of gp43-gp45 interaction over the C terminus of gp43 alone. 5) Contact on two subunits of gp45 rationalizes the rate of gp45 subunit exchange dropping to one-third of its original value upon formation of the holoenzyme (42). 6) Finally, the holoenzyme model is not built from ground-state structures but reflects the conformationally dynamic nature of the holoenzyme assembly process (7, 13–17, 32, 33, 42).

Recently, the C-terminal peptide of gp43 was cocrystallized with gp45 and the structure of the complex solved by x-ray crystallography (22). Surprisingly, the peptide was found to bind near the interdomain connecting loop of gp45. The ground state x-ray crystal structure of gp43 was then docked onto this structure to make a model of the holoenzyme (referred to hereafter as the interdomain connecting loop model). This interdomain connecting loop model presents several inconsistencies with previous and present solution data. 1) The bound and unbound x-ray crystal structures of gp45 were found to be very similar, in contrast to the large differences in gp45 conformations that were observed in solution for gp45 alone and in the holoenzyme (15, 16). 2) We have demonstrated that gp45 has an open subunit interface in solution (33) and that upon formation of the holoenzyme this subunit interface remains open by about 11 Å (15, 16). In the interdomain connecting loop model, all gp45 subunit interfaces are closed. It is possible that crystal packing forces push all of the subunit interfaces closed or that gp45 in solution equilibrates between open and closed forms with the closed form more readily crystallized. Regardless of the reason behind this phenomenon, the solution and crystal structures are not in agreement. 3) The x-ray crystal structure of the C terminus of gp43 bound to gp45 yielded a stoichiometry of 1 peptide per gp45 trimer (22), whereas fluorescence measurements in solution yielded a stoichiometry of 1 peptide per gp45 monomer (32). It was observed that crystal packing forces occluded one site but that the remaining two were not occluded. Given the high concentration used during crystallization (reported to be 7.5 mg/ml or 300 μM monomers

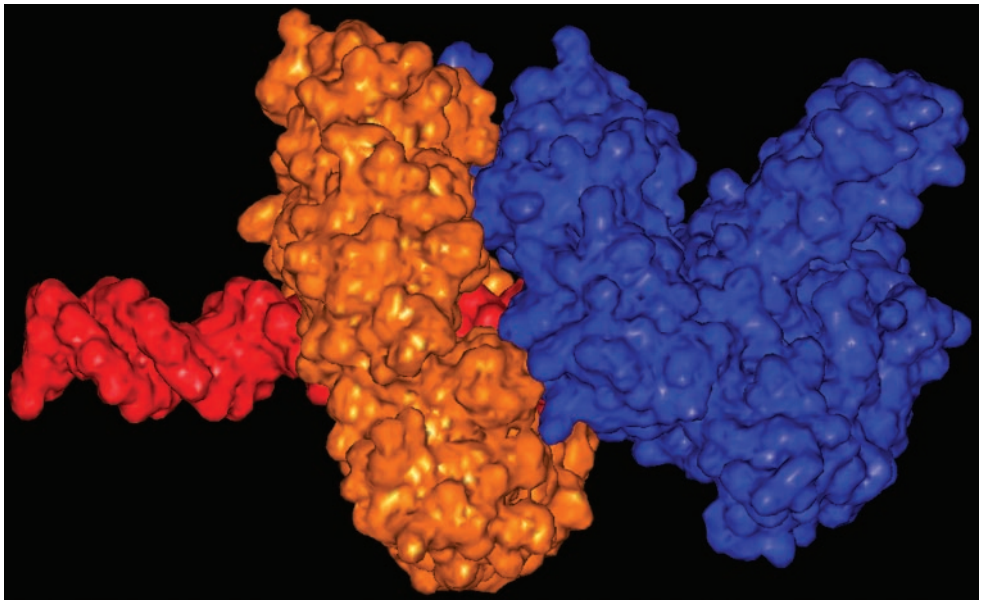

FIG. 7. **Solution structure of the bacteriophage T4 DNA polymerase holoenzyme.** Protein-protein interactions characterized on this work as well as previous investigations (7, 13–17, 32–33, 42) were used to build this model.

(22)), this site should be completely saturated since it has a measured K_D of 7 μM. Since gp45 crystallizes in a closed conformation, binding of the peptide at the subunit interface may be occluded, and binding may instead occur at a secondary site with weaker affinity resulting in partial saturation. 4) There is a significant difference between the affinity of peptide 1 for gp45 (7 μM) and gp43 for gp45 (50 nM), requiring additional contacts beyond the C terminus of gp43 to provide this stabilization. However, binding of gp43 to the interdomain connecting loop site of the monomeric gp45 mutant does not result in significant stabilization beyond peptide 1 alone (1.1 *versus* 2.7–5.6 μM). From a thermodynamic standpoint alone, gp43 will bind at the site where there is the greatest opportunity for stabilization, the gp45 subunit interface. Significantly, when proposing the interdomain connecting loop model, the authors point out that there are no "obvious areas of shape, charge, or hydrophobic complementarity that would lead to additional interactions (22)" between gp43 and gp45 in this orientation. The observation of significant stabilization beyond the C terminus of gp43 requires additional interactions of some kind, which we demonstrated in this work by finding cross-links between I107C of gp45 and near Arg-682 in gp43. 5) Finally, it has been observed that the rate of gp45 subunit exchange drops to one-third of its original value upon formation of the holoenzyme (42). This suggests that gp45 is contacted on *two* subunits and that only the third is free to exchange. The interdomain connecting loop model provides contact to only one gp45 subunit, which would lead to a drop in the rate of subunit exchange to two-thirds of its original value. We reiterate that models of multienzyme complexes that are assembled by dynamic processes to yield structures that are highly distinct from their ground state structures must include solution data as well as high resolution structural data to yield meaningful and physiologically relevant models.

Although we have demonstrated that the final holoenzyme complex has an interaction between the C terminus of gp43 and the subunit interface of gp45, an interaction between the C terminus of gp43 with the interdomain connecting loop of gp45, as found in the x-ray crystal structure, may be important for intermediary steps in holoenzyme formation. The binding of gp43 to the gp45/gp44/62/DNA complex has three steps observable by fluorescence (15, 16), and these intermediates in the holoenzyme assembly process are likely to have distinct structures from the final holoenzyme. The distance across the subunit interface in the initial gp45/gp44/62/DNA complex is greater than the final holoenzyme, and significant structural rearrangements take place during gp43 binding (15, 16). A holoenzyme assembly model in which initial binding of the C terminus of gp43 to the interdomain connecting loop of gp45 is followed by a rearrangement of the C terminus of gp43 to the subunit interface of gp45 is entirely consistent with both the x-ray crystal structure as well as our current and previous solution data. The geometry of the subunit interface in the gp45/gp44/62/DNA complex is not optimized for gp43 interactions, and the C terminus of gp43 would therefore bind to the interdomian connecting loop region of gp45 since it would provide the greatest initial stabilization. The thermodynamic driving force to move from a micromolar K_D (interdomain connecting loop) to a nanomolar K_D (subunit interface) would force a rearrangement of the gp45/gp44/62/gp43/DNA complex and may in the process force the ejection of gp44/62 from this intermediate complex, thereby completing the holoenzyme assembly process. Because only the C terminus of gp43 was used in the x-ray crystal structure, there would not be a driving force to rearrange the gp43/gp45 complex into the final holoenzyme structure since the additional contacts of gp43 outside of the C terminus that further stabilize the holoenzyme cannot be made. Further experiments will be required to determine whether both modes of gp43/gp45 interaction lie on the pathway to formation of the holoenzyme or that the interaction of the C terminus of gp43 and the subunit interface of gp45 that we have demonstrated in the final holoenzyme complex is found during the entire assembly process.

The holoenzyme model presented in Fig. 7 is novel in respect to the mode of interaction between the polymerase and sliding clamp; the C terminus of the polymerase is inserted into the open subunit interface of the sliding clamp. Alternatively in eukaryotes, the N terminus of DNA polymerase δ is apparently required for interaction with PCNA and formation of a functional holoenzyme (43–45), and the *E. coli* β-clamp appears to be completely closed in the DNA polymerase III holoenzyme (20). Both prokaryotic and eukaryotic replisomes require higher levels of processivity due to their much larger genomes, and closure of all clamp subunit interfaces could lead to the

greater holoenzyme stability required for this increase in processivity. The interaction of the β-clamp and PCNA with their cognate polymerases appears to be mediated by contacts at the C termini of the sliding clamps; it has been noted (46) that the final amino acids at the C termini of both the β-clamp and PCNA from a wide variety of species are highly conserved. Deletion or mutagenesis of these C-terminal amino acids abolishes interaction with the polymerases (27, 46), consistent with the conservation of these amino acids being important in maintaining protein-protein interactions. No study has addressed the importance of the C terminus of gp45, which could help contribute to the additional stabilization observed beyond the interaction of the C terminus of gp43 and the subunit interface of gp45. The interaction of DNA polymerase δ with the interdomain connecting loop of PCNA appears to be shared with gp43 and gp45 since photocross-links form between I107C of gp45 (at the midpoint of this loop) and gp43. Whether this interaction is only due to spacial proximity or is involved in stabilizing interactions as well remains to be tested.

The recently determined x-ray crystal structure of the C terminus of the herpes simplex virus DNA polymerase (36 amino acids) bound to its processivity factor UL42 (47) presents an interesting combination of the various modes of protein-protein interactions employed in the above-described sliding clamp-polymerase complexes. UL42 is a monomer, unlike the trimeric or dimeric sliding clamps, and is similar in structure but not amino acid sequence to PCNA and gp45 in that it has two domains tethered by an interdomain connecting loop (47). UL42 also has significant affinity for DNA, unlike the sliding clamps. As with bacteriophages T4 and RB69 but unlike prokaryotes or eukaryotes, the C terminus of the herpes simplex virus DNA polymerase interacts with the processivity factor, although it forms an αβα C-terminal structure not observed in any other species. The C-terminal herpes simplex virus DNA polymerase peptide forms an extensive network of interactions with UL42 that includes an antiparallel β-sheet with the interdomain connecting loop of UL42, similar to the interaction observed for the C terminus of p21, the cell cycle checkpoint protein, with PCNA (25). This peptide appears to contain all of the contacts necessary for UL42 interaction (48), similar to p21 but in contrast to gp43.

This mixing and matching of the various modes of polymerase processivity factor interactions between species suggests that there are a variety of potential protein-protein interaction combinations available to yield processive holoenzymes, rather than one uniform mode of interaction. Within the same species, the mode of interaction is apparently conserved. PCNA interacts with a number of other proteins besides p21, and all of these proteins contain a common PCNA-binding motif (49). Likewise, gp45 interacts with gp33 and gp55 during late RNA transcription through a motif shared with gp43 (31, 50). The diversity of sliding clamp-polymerase protein-protein interaction modes between species is consistent with the observation of functional conservation despite poor or nonexistent amino acid homology for the DNA replication systems across the divisions of life (51) and suggests that the protein-protein interaction modes between the other components of the replisome will show an interspecies diversity as well.

Acknowledgments—We thank Ismail Moarefi and John Kuriyan for providing the x-ray crystal structure coordinates of gp45 before publication. We also thank Ernesto Abel-Santos, Ann Valentine, and Vince Shier for helpful discussions. The mass spectrometer was purchased in part with funds from National Institutes of Health Grant RR11318.

REFERENCES

1. Baker, T. A., and Bell, S. P. (1998) *Cell* **92,** 295–305
2. von Hippel, P. H. (2000) *Trends Biochem. Sci.* **25,** 155
3. Young, M. C., Reddy, M. K., and von Hippel, P. H. (1992) *Biochemistry* **31,** 8675–8690
4. Nossal, N. G. (1992) *FASEB J.* **6,** 871–878
5. Sexton, D. J., Berdis, A. J., and Benkovic, S. J. (1997) *Curr. Opin. Chem. Biol.* **1,** 316–322
6. Berdis, A. J., and Benkovic, S. J. (1996) *Biochemistry* **35,** 9253–9265
7. Sexton, D. J., Kaboord, B. F., Berdis, A. J., Carver, T. C., and Benkovic, S. J. (1998) *Biochemistry* **37,** 7749–7756
8. Kaboord, B. F., and Benkovic, S. J. (1995) *Curr. Biol.* **5,** 149–157
9. Kuriyan, J., and O'Donnell, M. (1993) *J. Mol. Biol.* **234,** 915–925
10. Stillman, B. (1994) *Cell* **78,** 725–728
11. Herendeen, D. R., and Kelly, T. J. (1996) *Cell* **84,** 5–8
12. Waga, S., and Stillman, B. (1998) *Annu. Rev. Biochem.* **67,** 721–751
13. Latham, G. J., Pietroni, P., Dong, F., Young, M. C., and von Hippel, P. H. (1996) *J. Mol. Biol.* **264,** 426–439
14. Sexton, D. J., Carver, T. E., Berdis, A. J., and Benkovic, S. J. (1996) *J. Biol. Chem.* **271,** 28045–28051
15. Alley, S. C., Abel-Santos, E., and Benkovic, S. J. (2000) *Biochemistry* **39,** 3076–3090
16. Trakselis, M. A., Alley, S. C., Abel-Santos, E., and Benkovic, S. J. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 8368–8375
17. Pietroni, P., Young, M. C., Latham, G. J., and von Hippel, P. H. (1997) *J. Biol. Chem.* **272,** 31666–31676
18. Bloom, L. B., Turner, J., Kelman, Z., Beechem, J. M., O'Donnell, M., and Goodman, M. F. (1996) *J. Biol. Chem.* **271,** 30699–30708
19. Bertram, J. G., Bloom, L. B., Turner, J., O'Donnell, M., Beechem, J. M., and Goodman, M. F. (1998) *J. Biol. Chem.* **273,** 24564–24574
20. Hingorani, M. M., and O'Donnell, M. (1998) *J. Biol. Chem.* **273,** 24550–24563
21. Moarefi, I., Jeruzalmi, D., Turner, J., O'Donnell, M., and Kuriyan, J. (2000) *J. Mol. Biol.* **296,** 1215–1223
22. Shamoo, Y., and Steitz, T. A. (1999) *Cell* **99,** 155–166
23. Kong, X. P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992) *Cell* **69,** 425–437
24. Krishna, T. S. R., Kong, X. P., Gary, S., Burgers, P. M., and Kuriyan, J. (1994) *Cell* **79,** 1233–1243
25. Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. (1996) *Cell* **87,** 297–306
26. Kelman, Z., and O'Donnell, M. (1995) *Nucleic Acids Res.* **23,** 3613–3620
27. Naktinis, V., Turner, J., and O'Donnell, M. (1996) *Cell* **84,** 137–145
28. Wang, J., Sattar, A. K. M. A., Wang, C.-C., Karam, J. D., Konigsberg, W. H., and Steitz, T. A. (1995) *Cell* **89,** 1087–1099
29. Wang, C.-C., Yeh, L.-S., and Karam, J. D. (1995) *J. Biol. Chem.* **270,** 26558–26564
30. Salinas, F., and Benkovic, S. J. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 7196–7201
31. Berdis, A. J., Soumillion, P, and Benkovic, S. J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 12822–12827
32. Alley, S. C., Jones, A. D., Soumillion, P., and Benkovic, S. J. (1999) *J. Biol. Chem.* **274,** 24485–24489
33. Alley, S. C., Shier, V. K., Abel-Santos, E., Sexton, D. J., Soumillion, P., and Benkovic, S. J. (1999) *Biochemistry* **38,** 7696–7709
34. Nossal, N. G. (1979) *J. Biol. Chem.* **254,** 6026–6031
35. Rush, J., Lin, T.-C., Quinones, M., Spicer, E. K., Douglas, I., Williams, K. R., and Konigsberg, W. H. (1989) *J. Biol. Chem.* **264,** 10943–10953
36. Frey, M. W., Nossal, N. G., Capson, T. L., and Benkovic, S. J. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 2579–2583
37. Gill, S. C., and von Hippel, P. V. (1989) *Anal. Biochem.* **182,** 319–326
38. Alley, S. C., Ishmael, F. T., Jones, A. D., and Benkovic, S. J. (2000) *J. Am. Chem. Soc.* **122,** 6126–6127
39. Hass, E., Katchalski-Katzir, E., and Steinberg, I. Z. (1978) *Biochemistry* **17,** 5064–5070
40. Selvin, P. R. (1995) *Methods Enzymol.* **246,** 300–334
41. Jonsson, Z. O., Podust, V. N., Podust, L. M., and Hübscher, U. (1995) *EMBO J.* **14,** 5745–5751
42. Soumillion, P., Sexton, D. J., and Benkovic, S. J. (1998) *Biochemistry* **37,** 1819–1827
43. Zhang, S.-J., Zeng, X.-R., Zhang, P., Toomey, N. L., Chuang, R. Y., Chang, L. S., and Lee, M. Y. W. T. (1995) *J. Biol. Chem.* **270,** 7988–7992
44. Zhang, P., Mo, J.-Y., Perez, A., Leon, A., Liu, L., Mazloum, N., Xu, H., and Lee, M. Y. W. T. (1999) *J. Biol. Chem.* **274,** 26647–26653
45. Schumacher, S. B., Stucki, M., and Hübscher, U. (2000) *Nucleic Acids Res.* **28,** 620–625
46. Kelman, Z., Zuo, S., Arroyo, M. P., Wang, T. S.-F., and Hurwitz, J. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 9515–9520
47. Zuccola, H. J., Filman, D. J., Coen, D. M., and Hogle, J. M. (2000) *Mol. Cell* **5,** 267–278
48. Bridges, K. G., Hua, Q., Brigham-Burke, M. R., Martin, J. D., Hensley, P., Dahl, C. E., Digard, P., Weiss, M. A., and Coen, D. M. (2000) *J. Biol. Chem.* **274,** 472–478
49. Reynolds, N., Warbrick, E., Fantes, P. A., and MacNeill, S. A. (2000) *EMBO J.* **19,** 1108–1118
50. Sanders, G. M., Kassavetis, G. A., and Geiduschek, E. P. (1997) *EMBO J.* **16,** 3124–3132
51. Leipe, D. D., Aravind, L., and Koonin, E. V. (1999) *Nucleic Acids Res.* **27,** 3389–3401