Mapping Protein−Protein Interactions in the Bacteriophage T4 DNA Polymerase Holoenzyme Using a Novel Trifunctional Photo-cross-linking and Affinity Reagent

Stephen C. Alley,[†] Faoud T. Ishmael,[†,‡] A. Daniel Jones,[†] and Stephen J. Benkovic*[,†]

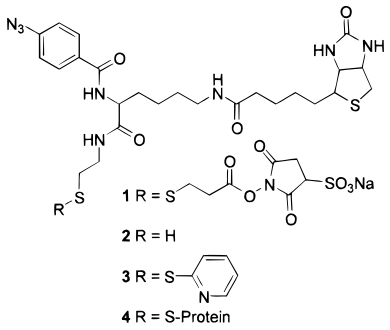
Department of Chemistry and Department of Biochemistry and Molecular Biology Hershey Medical Center The Pennsylvania State University, 415 Wartik Laboratory University Park, Pennsylvania 16802

Received February 18, 2000

DNA replication requires the formation of multiprotein complexes, whose functions[1] and structures[2] have been conserved during evolution. In bacteriophage T4, the DNA polymerase holoenzyme forms the core of the DNA replication system, consisting of the DNA polymerase (gp43), a trimeric, circular processivity factor called the sliding clamp (gp45), and the clamp loader (a 4:1 complex of gp44 and gp62) that hydrolyzes ATP to assemble the holoenzyme.[3] The holoenzymes of prokaryotic and eukaryotic DNA replication systems contain analogous proteins.[1a,4]

The structures of gp45[5] and gp43[6] have been individually solved by X-ray crystallography, but neither the gp44/62 complex nor any of the other multiprotein complexes such as the final holoenzyme have been solved. Models of the holoenzyme based on the individual gp45 and gp43 X-ray crystal structures have been proposed.[5a,6] However, these models have not taken into account solution evidence that points to a holoenzyme that has undergone ATP hydrolysis-dependent conformational changes relative to the individual ground-state structures.[7]

Site-specific incorporation of photo-cross-linking reagents has been extensively used to investigate the protein−protein and protein−nucleic acid interactions in solution in the bacteriophage T4 holoenzyme, including the organization of the holoenzyme,[8] the conformationally dynamic nature of the holoenzyme assembly process,[7c] and the location of the interaction between the C-terminus of gp43 and gp45.[7g] To streamline investigations such as these and make them amenable to even more complex multiprotein systems, we have synthesized **3**, a novel, trifunctional photo-cross-linking and affinity reagent.[9] The three functional groups in **3** are (1) a thiol-reactive 2-thiopyridine mixed disulfide for conjugation to a bait protein, (2) a photoactivatable aryl azide for photo-cross-linking to a target protein (the target protein may be the same protein as the bait, resulting in intra- or intersubunit photo-cross-links), and (3) the affinity probe biotin, which can be used for purification and visualization of photo-cross-linked proteins. The succinimide **1** conjugates to lysine,[10] but it does not allow site-specific attachment, a requirement for mapping protein−protein interactions. The mixed disulfide **3** conjugates to cysteine, allowing site-specific attachment on many proteins. The resulting conjugate (following the loss of 2-thiopyridine; e.g., **4**) has a mixed disulfide at the surface of the bait protein that is exchangeable in the presence of free thiols and allows the transfer of biotin from the bait to the target upon reduction of photo-cross-links.



The utility of **3** was demonstrated with the gp45 mutant I107C,[11] which contains a single cysteine per monomer. I107C is at the midpoint of the interdomain connecting loop of gp45, a region of the protein suggested to play a role in DNA loading by gp44/62.[7f] The point of attachment of **3** was verified by digesting the I107C-**3** conjugate with trypsin and then purifying biotinylated peptides on a monomeric avidin column.[12] Analysis by matrix-assisted laser desorption ionization (MALDI) mass spectrometry (Figure 1) resulted in an ion at *m*/*z* 3889.0, corresponding to the mass of the predicted gp45 tryptic fragment containing I107C

* To whom correspondence should be addressed. E-mail: sjb1@psu.edu. Phone: (814) 865-2882. Fax: (814) 865-2973.
† The Pennsylvania State University.
‡ Hershey Medical Center.

(1) (a) Baker, T. A.; Bell, S. P. *Cell* **1998**, *92*, 295−305. (b) Leipe, D. D.; Aravind, L.; Koonin, E. V. *Nucleic Acids Res.* **1999**, *27*, 3389−3401.
(2) (a) Kelman, Z.; O'Donnell, M. *Nucleic Acids Res.* **1995**, *23*, 3613−3620. (b) Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395−17398.
(3) Reviewed in: (a) Young, M. C.; Reddy, M. K.; von Hippel, P. H. *Biochemistry* **1992**, *31*, 8675−8690. (b) Nossal, N. G. *FASEB J.* **1992**, *6*, 871−878. (c) Sexton, D. J.; Berdis, A. J.; Benkovic, S. J. *Curr. Opin. Chem. Biol.* **1997**, *1*, 316−322.
(4) Reviewed in: (a) Kuriyan, J.; O'Donnell, M. *J. Mol. Biol.* **1993**, *234*, 915−925. (b) Stillman, B. *Cell* **1994**, *78*, 725−728. (c) Herendeen, D. R.; Kelly, T. J. *Cell* **1996**, *84*, 5−8. (d) Waga, S.; Stillman, B. *Annu. Rev. Biochem.* **1998**, *67*, 721−751.
(5) (a) Shamoo, Y.; Steitz, T. A. *Cell* **1999**, *99*, 155−166. (b) Moarefi, I.; Jeruzalmi, D.; Turner, J.; O'Donnell, M.; Kuriyan, J. *J. Mol. Biol.* **2000**, *296*, 1215−1223.
(6) Wang, J.; Sattar, A. K. M. A.; Wang, C.-C.; Karam, J. D.; Konigsberg, W. H.; Steitz, T. A. *Cell* **1997**, *89*, 1087−1099.
(7) (a) Sexton, D. J.; Carver, T. E.; Berdis, A. J.; Benkovic, S. J. *J. Biol. Chem.* **1996**, *271*, 28045−28051. (b) Latham, G. J.; Pietroni, P.; Dong, F.; Young, M. C.; von Hippel, P. H. *J. Mol. Biol.* **1996**, *264*, 426−439. (c) Pietroni, P.; Young, M. C.; Latham, G. J.; von Hippel, P. H. *J. Biol. Chem.* **1997**, *272*, 31666−31676. (d) Soumillion, P.; Sexton, D. J.; Benkovic, S. J. *Biochemistry* **1998**, *37*, 1819−1827. (e) Sexton, D. J.; Kaboord, B. F.; Berdis, A. J.; Carver, T. C.; Benkovic, S. J. *Biochemistry* **1998**, *37*, 7749−7756. (f) Alley, S. C.; Shier, V. K.; Abel-Santos, E.; Sexton, D. J.; Soumillion, P.; Benkovic, S. J. *Biochemistry* **1999**, *38*, 7696−7709. (g) Alley, S. C.; Jones, A. D.; Soumillion, P.; Benkovic, S. J. *J. Biol. Chem.* **1999**, *274*, 24485−24489. (h) Alley, S. C.; Abel-Santos, E.; Benkovic, S. J. *Biochemistry* **2000**, *39*, 3076−3090.
(8) (a) Capson, T. L.; Benkovic, S. J.; Nossal, N. G. *Cell* **1991**, *65*, 249−258. (b) Latham, G. J.; Bacheller, D. J.; Pietroni, P.; von Hippel, P. H. *J. Biol. Chem.* **1997**, *272*, 31677−31684. (c) Latham, G. J.; Bacheller, D. J.; Pietroni, P.; von Hippel, P. H. *J. Biol. Chem.* **1997**, *272*, 31685−31692.

(9) To 2 mg of **1** (sulfosuccinimidyl-2-[6-(biotinamido)-2-(*p*-azidobenzamido)hexanoamido]ethyl-1,3′-dithiopropionate) (sulfo-SBED, Pierce, Rockford, IL) in 500 μL of DMF (4.5 mM) was added 50 μL of 100 mM DTT in DMF (9 mM final) in the dark at 25 °C, forming **2**. After 30 min, 11 mg of 2,2′-dithiodipyridine (Sigma, St. Louis, MO) was added (90 mM final) and the mixture incubated for an additional 30 min. To purify **3**, 175-μL aliquots of this mixture were separated on a 4.6 × 250 mm C18 HPLC column (Vydac, Hesperia, CA) using the following gradient at 1 mL min^{-1}: solvent A, 0.1% trifluoroacetic acid in H_2O; solvent B, 0.1% trifluoroacetic acid in CH_3CN; 100% solvent A for 2 min, a 20-min linear gradient from 100 to 50% solvent A, and a 5-min linear gradient from 50 to 0% solvent A, with **3** eluting at 19.5 min. HRMS (ESI$^+$): *m*/*z* 686.2348 ([M + H]$^+$, theor 686.2365).

(10) (a) Hermanson, G. T. *Bioconjugate Techniques*; Academic Press: San Diego, 1996; pp 289−291. (b) Ilver, D.; Arnqvist, A.; Ögren, J.; Frick, I.-M.; Kersulyte, D.; Incecik, E. T.; Berg, D. E.; Covacci, A.; Engstrand, L.; Borén, T. *Science* **1998**, *279*, 373−377.

(11) The I107C gene was cloned and the protein purified as previously described for other gp45 mutants (ref 7f). I107C (150 μM in monomers) was dialyzed into NR buffer (20 mM HEPES, pH 7.0, 50 mM NaCl, 1 mM EDTA, and 10% glycerol) and **3** added in 0.1 vol DMF (1.5 mM final). This mixture was nutated in the dark at 4 °C for 12 h and then purified on a MonoQ (Pharmacia, Piscataway, NJ) anion-exchange FPLC column. Although we have excluded light in both the synthesis of **3** and the formation of the I107C-**3** conjugate, we have found that the aryl azide of **3** is stable enough to produce satisfactory results without the rigorous exclusion of light. Exclusion of free thiols in the conjugation and photo-cross-linking reactions is essential.

10.1021/ja000591t CCC: $19.00
Published on Web 06/09/2000

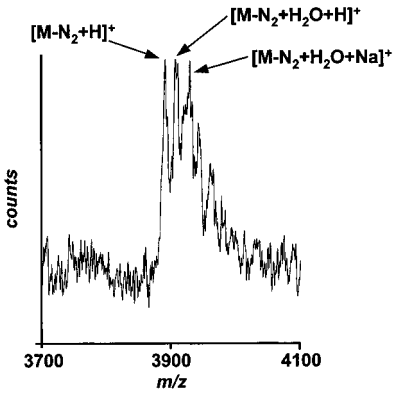


Figure 1. MALDI mass spectrum of tryptic digest of gp45 I107-**3** conjugate. The digest was applied to a monomeric avidin column to bind biotinylated tryptic fragments, the column washed to remove other tryptic fragments, and biotinylated tryptic fragments eluted with 2 mM biotin. The eluant was desalted and linear mode MALDI mass spectra obtained using an α-cyano-4-hydroxycinnamic acid matrix. Peaks corresponding to $[M - N_2 + H]^+$ (*m*/*z* 3889.0, theoretical 3889.6), $[M - N_2 + H_2O + H]^+$ (*m*/*z* 3904.2, theoretical 3907.6), and $[M - N_2 + H_2O + Na]^+$ (*m*/*z* 3927.6, theoretical 3929.6) were observed.

(amino acids 88−119), plus **3**, minus 2-thiopyridine, and minus N_2 (theoretical *m*/*z* 3889.6).

The interaction of the I107C-**3** conjugate with gp44/62 was then investigated using a procedure similar to a Western blot.[10] I107C-**3** alone or in the presence of gp44/62, ATP, and DNA was subjected to photo-cross-linking, separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; the gp45 trimer and gp44/62 complex are denatured into monomers in this gel), the gel blotted onto a nitrocellulose membrane, and proteins containing biotin from **3** visualized by chemiluminescence (Figure 2).[13] Lanes 1 and 2 contain the I107C-**3** conjugate with and without gp44/62, ATP, and DNA, respectively, without exposure to light. Only gp45 bears biotin in these two cases and is the only protein with a chemiluminescent signal. Lanes 6 and 7 are identical to lanes 1 and 2 except that 100 mM DTT was added before SDS-PAGE. DTT cleaves the disulfide bond in the I107C-**3** conjugate and releases biotin to solution, yielding no chemiluminescent signal. Lane 3 contains I107C-**3** that has been exposed to light. Intersubunit photo-cross-linked gp45 dimers are observed with slower mobility than gp45 monomers, and the smear with faster mobility may be intrasubunit photo-cross-links. Lane 8 is the DTT-treated version of lane 3, where reduction of photo-cross-links transfers biotin from the bait to the target and allows identification of the target (gp45; bait and target are the same) based on electrophoretic mobility. Lane 4 contains a mixture of I107C-**3**, gp44/62, ATP, and DNA that has been exposed to light. Gp45 inter- and intrasubunit photo-cross-links are observed as in lane 3, but another slow mobility band is observed. Reduction of this photo-cross-linked mixture (lane 9) identifies gp45 and gp44 as targets, demonstrating that gp44 but not gp62 comes in contact with gp45 at position I107C. Previously, gp45 has been observed to photo-cross-link to just gp44 at position S19C, while gp45 photo-cross-links to both gp44 and gp62 at position K81C.[7c] A light band with the mobility of gp44 can be seen in lane 4, possibly due to reduction of gp45/gp44 cross-links by exogenous thiols (likely from gp44/62).

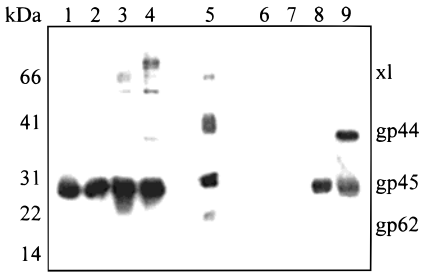


Figure 2. Photo-cross-linking reactions separated by SDS-PAGE and detected by biotin-specific chemiluminescence.[13] Lanes 1−4 contain the following components treated as described: (1) I107C-**3** not exposed to light, (2) I107C-**3**, gp44/62, ATP, and DNA not exposed to light, (3) I107C-**3** exposed to light, and (4) I107C-**3**, gp44/62, ATP, and DNA exposed to light. Lanes 6−9 are identical to lanes 1−4 except that DTT was added following photo-cross-linking to reduce the disulfide bond in **3**. Lane 5 contains standard biotinylated proteins (BioRad, Hercules, CA) as size markers (M_r shown at the left). The mobilities of gp62 (21 kDa), gp45 (25 kDa), gp44 (36 kDa), and cross-linked species (designated xl) are shown at the right.

This rapid procedure avoids traditional methods of visualizing a potentially complex mixture of proteins using nonspecific staining techniques (Coomassie or silver), multiple Western blots individually probing for all possible targets, or radioactive probes. Further verification of the identity of cross-linked proteins can be accomplished by isolating biotinylated targets by affinity chromatography followed by Edman sequencing or mass spectrometry. The precise location of protein−protein interactions can also be determined by proteolytic digestion of a photo-cross-linking mixture followed by affinity isolation of biotinylated target fragments and identification by mass spectrometry. The perfluoroaryl azide of **3** would facilitate this by increasing the yield of photo-cross-links while decreasing the possible number of adduct sites and therefore increasing specificity.[7c,g]

We are currently using **3** to map specific contact points of protein−protein interactions within the bacteriophage T4 holoenzyme to build a more accurate solution structure. We are also using **3** to identify the multiple protein−protein interactions found within the entire bacteriophage T4 replication fork (four additional proteins) and provide information on the organization of this complex biological system.[3b,c] Multifunctional reagents such as **3** will facilitate the investigation of important biological systems that require the formation of multiprotein complexes, for example DNA replication, transcription, recombination, DNA repair, signal transduction, and molecular recognition. In the investigation of highly complex biological systems, the use of affinity probes becomes increasingly essential.

Acknowledgment. This work was supported by National Institutes of Health grants GM13306 (S.J.B.), GM19492 (S.C.A.), and DK19691 (F.T.I.). The mass spectrometer was purchased in part with funds from National Institutes of Health grant RR11318.

JA000591T

(12) The I107C-**3** conjugate (250 μg) was digested with 25 μg of trypsin and applied to a 1.5-mL monomeric avidin column (Pierce). After washing to remove unbound material, biotinylated tryptic fragments were eluted with 2 mM biotin. These peptides were analyzed by MALDI mass spectrometry as previously described (ref 7g). Monomeric avidin and biotin form a reversible noncovalent interaction, unlike tetrameric avidin or streptavidin (K_a of 10^8 versus 10^{15}) (ref 14).

(13) Photo-cross-linking was performed as previously described (302 nm lamp with 1800 mW/cm^2 output at 15 cm; ref 7g) with 1 μM I107C-**3**, 1 μM gp44/62 (in NR buffer), and 1 μM primer-template DNA (a 30-mer/44-mer) in 20 μL of 20 mM Tris, pH 7.5, 150 mM sodium acetate, 10 mM magnesium acetate, and 1 mM ATP. Following photo-cross-linking, 5 μL of 150 mM Tris, pH 6.8, 4% sodium dodecyl sulfate, 0.1% bromophenol blue, 30% glycerol, 8 mM *N*-ethylmaleimide, and 20% DMF were added and the mixture allowed to stand at 25 °C for 5 min. DTT (2 μL of a 1 M solution) was then added to some samples as indicated, 10-μL aliquots separated by SDS-PAGE (12% gel), and the gel blotted onto a nitrocellulose membrane (MicronSeparations, Westborough, MA). The membrane was blocked for 15 min with 3% bovine serum albumin (BSA) in 20 mM Tris, pH 7.6, 140 mM NaCl, and 0.1% Tween-20 (TBST), probed for 45 min with a 1:3000 dilution of streptavidin-horseradish peroxidase conjugate (Life Technologies, Rockville, MD) in 3% BSA in TBST, washed three times for 5 min each with TBST, and developed with a luminol/hydrogen peroxide mixture (Pierce). Chemiluminescence was detected by autoradiography onto BioMax film (Kodak, Rochester, NY) with typical exposure times of 25−45 s.

(14) Henrikson, K. P.; Allen, S. G. H.; Maloy, W. L. *Anal. Biochem.* **1979**, *94*, 366−370.