Structural and Enzymatic Studies of the T4 DNA Replication System

I. PHYSICAL CHARACTERIZATION OF THE POLYMERASE ACCESSORY PROTEIN COMPLEX*

(Received for publication, November 28, 1988)

Thale C. Jarvis‡, Leland S. Pauls, and Peter H. von Hippel

From the Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, Oregon 97405

In this study, we have investigated the structural and physical properties of the bacteriophage T4 DNA polymerase accessory proteins. We find that T4 gene 44 and 62 proteins associate to form a tight, highly homogeneous complex, containing four gene 44 protein subunits and one gene 62 protein subunit. The molecular mass of the complex is 163,700 daltons. Sedimentation results suggest that the complex is quite asymmetric, with a prolate ellipsoid axial ratio of about 6:1. This protein complex is known to carry a DNA-dependent ATPase activity; we show by photolabeling that the ATP-binding sites reside in the gene 44 protein subunits of the complex. Equilibrium sedimentation and chemical cross-linking studies indicate that the T4 gene 45 protein self-associates to form a trimer in solution. This trimer species also appears to be quite asymmetric, showing an axial ratio for a prolate ellipsoid of about 6:1, assuming normal hydration.

The bacteriophage T4 gene 43, 44, 62, 45, and 32 proteins, known to be required for DNA replication in vitro, can be formed into a reconstituted complex that is able to carry out leading strand DNA synthesis in vitro (1, 2). The "polymerase accessory proteins," encoded by genes 44, 62, and 45, show a DNA-dependent ATPase activity (3, 4), and in the presence of ATP these proteins greatly stimulate the enzymatic activity of the polymerase (5-7). Mechanistic studies of functional replication complexes require a structural understanding of the component proteins. Thus, we have undertaken a physical characterization of the T4 gene 44, 62, and 45 proteins.

In this paper we present structural and physical evidence for the association states and subunit compositions of the T4 polymerase accessory proteins. In the accompanying paper (8), we examine functional aspects of the complex as manifested by its DNA-dependent ATPase activity. Through a combination of physical and enzymatic characterization, we can begin to elucidate the molecular mechanisms whereby the accessory proteins interact with the DNA template and the other proteins within the replication complex.

T4 gene 44 and 62 proteins associate to form a tight complex, dissociable only under denaturing conditions (3, 9, 10). Both of the genes have been cloned (11, 12), and the monomer molecular masses are known to be 35,584 and 21,547 daltons, respectively. Estimates in the literature for the size and composition of the complex vary (9, 11, 13). Thus, in order to understand the mechanism we must first define the stoichiometry of the complex. We have approached this by molecular weight determination, using velocity sedimentation, dynamic laser light scattering, and equilibrium sedimentation techniques. In addition, we have determined subunit composition directly by using reverse-phase HPLC to separate the gene 44 and 62 protein subunits from the complex.

The gene 45 protein has also been cloned (14) and has a monomer molecular mass of 24,710 daltons. In order to understand how this protein interacts with the gene 44/62 protein complex, it is important to determine the association state of gene 45 protein both in solution and in association with the gene 44/62 protein complex. We have used equilibrium sedimentation and chemical cross-linking to determine the molecular weight of the associated species of gene 45 protein.

MATERIALS AND METHODS

Preparation of Proteins and Nucleic Acids—T4 gene 45 protein was prepared according to the method of Morris et al. (10) with the following modification. After the norelucine-Sepharose chromatography step, the protein was dialyzed into 50 mM NaCl, 1 mM EDTA, 10% glycerol, 1 mM β-mercaptoethanol, and 20 mM Tris-HCl, pH 8.1, and run over single-stranded DNA-cellulose in series with Bio-Rex-70 (Bio-Rad), followed by a second DEAE-cellulose column (Whatman DE52). Under these conditions, the gene 45 protein passes through the first two columns and binds to DE52. It was then batch eluted from DE52 with a similar buffer containing 400 mM NaCl.

T4 gene 44/62 protein was also purified according to Morris et al. (10), except that the protein was eluted from the hydroxypatite column with a linear salt gradient (13). In addition, the final pool was dialyzed into 1 buffer (as defined by Morris et al. (10)), loaded onto single-stranded DNA-cellulose, and eluted with a linear gradient from 1 buffer to 1 buffer plus 100 mM NaCl. The gene 44/62 protein complex eluted at about 40 mM NaCl.

T4 gene 32 protein was purified according to Bittner et al. (15), using T4 phage amN134, amBL292, and amE219 (33%, 53%, 58% -61%). The buffers for norelucine-Sepharose chromatography contained 0.8 M NaCl, instead of the 0.5 M specified by Bittner et al. (15). An additional chromatography step, DEAE-Sephasel, was performed according to Alberts and Frey (16).

T4 gene 43 protein was purified according to Morris et al. (10). Escherichia coli rho protein and NusA protein were gifts from the

* This work was supported in part by United States Public Health Service (USPHS) Research Grants GM-15792 and GM-29158 (to P. H. v. H.) and by a grant from the Lucille P. Markey Charitable Trust. 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.

† Predoctoral trainee supported by USPHS Institutional Research Service Award GM-07759. Submitted to the Graduate School of the University of Oregon in partial fulfillment of the requirements for the Ph.D. degree in Chemistry. Present address: Dept. of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309.

‡ Predoctoral trainee supported by USPHS Institutional Research Service Award GM-07759. Present address: Abbott Laboratories, Abbott Park, IL 60064.

§ American Cancer Society Research Professor of Chemistry.

1 The abbreviations used are: HPLC, high performance liquid chromatography; DMS, dimethylsulfate; NHEES, 4-(2-hydroxyethyl)-l-piperazineethanesulfonic acid; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; N,ATP, 8-azidoadenosine 5'-triphosphate.
nues Geiselman and Stanley C. Gill, respectively (this laboratory). All protein preparations were judged >98% pure based on SDS-PAGE and were shown to be free of contaminating endonuclease by incubation with supercoiled pBR322.

Poly(dA) was purchased from Pharmacia LKB Biotechnology Inc. Partial Specific Volume and Extinction Coefficient—Amino acid sequence information (11, 12, 14) allows us to calculate theoretical values for various physical properties of the accessory proteins, such as the molar extinction coefficient and the partial specific volume. In general, the implicit assumption in these calculations is that the physical characteristics of the individual amino acid residues will remain essentially the same in the native, folded protein as they are free in solution. While there are clearly examples of proteins for which this is not the case, the assumption proves surprisingly valid for the majority of proteins, and thus such calculations can provide valuable estimates of the physical properties of proteins of known sequence.

The theoretical partial specific volume, \( \bar{\rho} \), of gene 45 protein is 0.743 g/ml. This is based on known partial specific volumes of individual amino acid residues (17). Gene 44 and 62 proteins have calculated values of \( \bar{\rho} \) of 0.739 and 0.746 g/ml, respectively.

Theoretical molar extinction coefficients for the proteins were calculated based on the number of tryptophans and tyrosines they contain, and using the molar extinction coefficients determined by Edelhoch (19) for proteins of known sequence. The values for the molar extinction coefficients in water are given in Table I. The values for the molar extinction coefficients in water are given in Table I. The assumption that the extinction coefficients remain essentially the same in the native, folded protein as they are free in solution, although this assumption is not to be expected to hold for all cases, is often a valid simplification as judged by the agreement between calculated and observed values.

Sedimentation Equilibrium—Equilibrium ultracentrifugation was performed in a Beckman model E Ultracentrifuge, using the meniscus depletion method (21). The sample buffer for both gene 45 and 44/62 proteins contained 100 mM KCl, 25 mM Tris-HCl, pH 7.4, and 0.4 mM dithiothreitol. The sedimentation boundary was monitored using a UV scanner set at a wavelength of 280 nm, and the boundaries were analyzed as described by Van Holde and Weischat (19). The initial gene 45 protein concentration was 0.38 mg/ml, and the rotor speed was 36,000 rpm.

Velocity Sedimentation—Sedimentation coefficients for gene 45 and 44/62 proteins were determined by velocity sedimentation in a Beckman model E Ultracentrifuge. A 12-mm double sector cell with quartz windows was filled with 0.9 ml of sample and a slightly greater amount of reference buffer. The buffer for both proteins consisted of 200 mM KCl, 25 mM Tris-HCl, pH 7.4, and 0.4 mM dithiothreitol. The sedimentation boundary was monitored using a UV scanner set at a wavelength of 280 nm, and the boundaries were analyzed as described by Van Holde and Weischat (19).

Light Scattering—The diffusion coefficient for the gene 44/62 protein complex determined by dynamic laser light scattering performed in a Beckman model E Ultracentrifuge, using the meniscus depletion method (21). All protein preparations were judged >98% pure based on SDS-PAGE and were shown to be free of contaminating endonuclease by incubation with supercoiled pBR322.

Poly(dA) was purchased from Pharmacia LKB Biotechnology Inc. Partial Specific Volume and Extinction Coefficient—Amino acid sequence information (11, 12, 14) allows us to calculate theoretical values for various physical properties of the accessory proteins, such as the molar extinction coefficient and the partial specific volume. In general, the implicit assumption in these calculations is that the physical characteristics of the individual amino acid residues will remain essentially the same in the native, folded protein as they are free in solution. While there are clearly examples of proteins for which this is not the case, the assumption proves surprisingly valid for the majority of proteins, and thus such calculations can provide valuable estimates of the physical properties of proteins of known sequence. The theoretical partial specific volume, \( \bar{\rho} \), of gene 45 protein is 0.743 g/ml. This is based on known partial specific volumes of individual amino acid residues (17). Gene 44 and 62 proteins have calculated values of \( \bar{\rho} \) of 0.739 and 0.746 g/ml, respectively.

Velocity Sedimentation—Sedimentation coefficients for gene 45 and 44/62 proteins were determined by velocity sedimentation in a Beckman model E Ultracentrifuge. A 12-mm double sector cell with quartz windows was filled with 0.9 ml of sample and a slightly greater amount of reference buffer. The buffer for both proteins consisted of 200 mM KCl, 25 mM Tris-HCl, pH 7.4, and 0.4 mM dithiothreitol. The sedimentation boundary was monitored using a UV scanner set at a wavelength of 280 nm, and the boundaries were analyzed as described by Van Holde and Weischat (19). The initial gene 45 protein concentration was 0.38 mg/ml, and the rotor speed was 36,000 rpm.

Determination of the Molecular Weight and Subunit Composition of the Gene 44/62 Protein Complex

As noted above, the products of T4 genes 44 and 62 associate to form a tight complex. Subunit exchange studies (9) have failed to detect any dissociation and reassociation of the complex under native conditions, indicating that the complex, once formed, is quite stable. Although the molecular weights of the monomers are accurately known from the DNA sequence, the precise molecular weight of the total complex has only been estimated. Various stoichiometries have been reported for the subunits of this complex, ranging from a ratio of four gene 44 to two gene 62 subunits (9) based on Coomassie staining of the proteins eluted from denaturing gels, to a 5:1 ratio (13), obtained by densitometry of a Coomassie-stained gel. Most recently, Spicer et al. (11) have reported a 3.6 ± 0.6:1 ratio, based on quantitation of the size of the peaks obtained at each step of protein sequencing. This probably gives the best estimate currently available, but leaves some uncertainty as to the exact subunit stoichiometry, and hence as to the molecular weight of the complex. To firmly establish the composition of this multiprotein complex, we have re-examined the issue of subunit stoichiometry and determined the molecular weight of the complex by two different methods.

Molecular Weight Determination Using the Svedberg Equation—When a macromolecular experiences a centrifugal field, the rate at which its sedimentation becomes proportional to its reduced mass and inversely proportional to its frictional coefficient. The frictional coefficient is influenced by factors such as shape and hydration and these same frictional forces will modulate the rate at which the macromolecule diffuses. The Svedberg equation (for a derivation, see Ref. 24):

\[
M = \frac{RTS}{D(1 - \phi)} \tag{1}
\]

allows the calculation of molecular weight using both the sedimentation and diffusion coefficients. Here \( M \) equals the molecular weight of the solute, \( R \) is the gas constant, \( T \) is the temperature, \( S \) is the sedimentation coefficient, \( D \) is the diffusion coefficient, \( \phi \) is the partial specific volume, and \( \rho \) is the solute density. The frictional factors cancel, and we can
obtain a molecular weight that is independent of shape and hydration.

The molecular weight of the gene 44/62 protein complex was determined by the sedimentation and diffusion method. A sedimentation coefficient of 7.1 ± 0.2 S was obtained by velocity sedimentation in an analytical ultracentrifuge. A diffusion coefficient of 3.9 × 10^{-7} cm²/s was determined by dynamic laser light scattering. The sedimentation and diffusion constants are corrected to 20 °C in water. The data were not obtained over a wide concentration range and thus are not corrected to infinite dilution. The samples are already quite dilute, however, and as noted by Van Holde (25), the S and S' (sedimentation coefficient extrapolated to infinite dilution) values for most globular proteins differ by only about 0.5% at protein concentrations of 1 mg/ml.

In addition, several lines of evidence support the homogeneity of the species, including laser light scattering and sedimentation equilibrium. Therefore, we feel justified in combining the sedimentation and diffusion data using the Svedburg equation, to give a molecular mass of 170,000 daltons for the complex. This relies on theoretical values for the partial specific volumes of the proteins, based on amino acid sequence (see "Materials and Methods"). The sedimentation constant of 7.1 S for the gene 44/62 protein complex is in good agreement with that determined by Barry and Alberts (9) using sucrose density gradient sedimentation. Consistently low values of the "quality parameter", µ/γ², for the light-scattering data show that the autocorrelation function fits well to a single exponential decay, providing a sensitive indication of sample homogeneity (20).

Determination of Molecular Weight by Equilibrium Sedimentation—The molecular weight of the gene 44/62 protein complex was further studied by equilibrium sedimentation, a technique capable of high precision. At sufficiently low rotor speeds, the centrifugal force that results in transport of the macromolecule by sedimentation will be balanced by diffusion transport in the opposite direction. An equilibrium condition is established, generating a concentration gradient throughout the cell. A homogeneous species, in an ideal two component system (i.e. solute and solvent) will show a concentration distribution described by:

\[ \frac{d \ln c}{dr^2} = \frac{M(1 - \varphi_p)\omega^2}{2RT} \]  

(2)

where c is the concentration of the solute, r is the distance from the center of rotation, and ω is the angular velocity (for a review see Schachman (24)).

The molecular weight values for the gene 44/62 protein complex, obtained by equilibrium sedimentation at two different rotor speeds, are shown in Table I. The results are in excellent agreement with those obtained by the sedimentation and diffusion method (see above). Calculation of the molecular weights relies on theoretical values for the partial specific volumes of the proteins, obtained from the known amino acid sequences. Since the experimentally determined partial specific volume of most proteins falls within 2% of the value predicted from sequence,^3 this assumption is expected to contribute an uncertainty of up to ±5% in the final calculated molecular weight, thus constituting the largest single source of error in the molecular weight determination. Experimental error, such as uncertainties in rotor speed, temperature, and quantitation of the interference fringe patterns, can be estimated and brings the total uncertainty in molecular weight to a maximum of about ±10%.

The In c versus r² plots for equilibrium sedimentation experiments run at 14,000 rpm are shown in Fig. 1A. Since the fringe displacement is proportional to the solute concentration, the slope of the plot is proportional to the apparent molecular weight of the protein. A, gene 44/62 protein sedimented at 14,000 rpm. B, gene 44/62 protein sedimented at 20,000 rpm.

^3 The calculated partial specific volumes of the gene 44 and 62 proteins are close enough in value so that the choice of stoichiometry (3:1, 4:1, 5:1, etc.) for the complex makes little difference in the calculated molecular mass. A complex of four gene 44 protein subunits and one gene 62 subunit should have a partial specific volume of 0.740 g/ml.

^4 S. C. Gill, personal communication.

### Table 1

| Complex          | Molecular mass (daltons) |
|------------------|-------------------------|
| Gene 44/62 protein | 163,000, 177,000         |
| Gene 45 protein   | 76,600, 78,700, 79,900  |

### Figure 1

**A. Gene 44/62 Protein**

**B. Gene 45 Protein**

The molecular weights of the gene 44/62 protein complex and the gene 45 protein were determined by equilibrium ultracentrifugation as described (see "Materials and Methods"). The initial concentration of gene 44/62 protein was 0.48 mg/ml for both runs, and the initial concentration of gene 45 protein was 0.69 mg/ml for the 24,000 and 28,000 rpm runs, and 0.46 mg/ml for the 20,000 rpm run.
laser light-scattering experiments, however, are quite sensitive to heterogeneity and argue strongly in favor of a high degree of homogeneity for the gene 44/62 protein complex. In view of these results, the data of Fig. 1A can be most simply interpreted as representing a homogeneous complex behaving in a relatively ideal fashion. It should also be noted that the Yphantis (21) meniscus depletion method used here, although quite sensitive to low molecular weight contaminants, can entirely miss very high molecular weight species. Again, because the light scattering technique is very sensitive to high molecular contaminants, the results give us confidence that gene 44/62 protein in solution exists as a single species of molecular mass 170 ± 20 kDa.

**Determination of Subunit Ratios by Reverse-phase HPLC—** The molecular mass studies described above have sufficient error to make it difficult to distinguish, for example, a complex containing four gene 44 subunits and one gene 62 subunit (molecular mass 164 kDa) from one containing four gene 44 subunits and two gene 62 subunits (molecular mass 185 kDa). Therefore, we have also investigated the stoichiometry of the subunits by using reverse-phase HPLC to separate and quantify the individual protein components of the complex. The gene 44 and 62 proteins separate on a C4 column with a gradient of increasing acetonitrile concentration. The elution profiles are shown in Fig. 2. At least 90% separation of the subunits is achieved under these conditions. Gel electrophoresis of fractions collected from this column identify the species that elutes first to be gene 62 protein; the larger peak with the longer retention time corresponds to gene 44 protein.

The elution of the proteins from the column was monitored by a UV detector at two different wavelengths. The dominant chromophore at 220 nm is expected to be the peptide bond. Therefore, the relative areas at 220 nm have been corrected by the molecular masses of the gene 44 and 62 protein monomers (35,584 and 21,347 daltons, respectively) to give the calculated subunit stoichiometry. At 280 nm, the absorbance is principally due to tryptophan and tyrosine residues. Thus, the theoretical extinction coefficients for the gene 44 and 62 proteins at this wavelength were used to convert the relative areas into subunit stoichiometries. The resulting stoichiometries are shown in Table II. Note that higher subunit ratios show incrementally smaller changes in the relative peak areas, making it difficult to distinguish, for example, a 4:1 from a 5:1 complex. This accounts for the asymmetric errors shown in Table II. Although the errors are too large to allow us uniquely to specify the subunit stoichiometry, this technique is particularly good at distinguishing between complexes with low subunit ratios; the difference between a 2:1 and a 3:1 complex is quite large. Thus, we can feel confident in ruling out the 2:1 gene 44 to 62 protein ratio originally suggested by Barry and Alberts (9).

**The Gene 44 and 62 Proteins Form a 4:1 Complex—** The molecular weight and subunit stoichiometry data described above for the gene 44/62 protein complex are summarized in Table III. Several hypothetical subunit compositions are listed along with predicted molecular weights that span the range of compositions previously reported in the literature. The results of the HPLC experiments (determining subunit stoichiometry of the complex), and the sedimentation experiments (yielding complex molecular weight), are scored in terms of agreement with the proposed subunit composition. While the molecular weight determination could not unambiguously rule out the 3:2 and 4:2 stoichiometries, the HPLC data is clearly inconsistent with such low subunit ratios. Only the 4:1 ratio is corroborated by both the sedimentation and HPLC data. Therefore, the combined evidence indicates that T4 gene 44 and 62 proteins form a tight, homogeneous species consisting of four gene 44 subunits, and one gene 62 protein subunit, with a total molecular mass of 163,700 daltons.5

| Wavelength of detection | % Gene 62 protein peak (relative area) |
|-------------------------|---------------------------------------|
| 220 nm                  | 23.1                                  |
| 280 nm                  | 14.4                                  |

**Table II**

Subunit stoichiometry of the gene 44/62 protein complex determined by HPLC separation of the subunits

Native gene 44/62 protein complex was loaded onto the column and the gene 44 and 62 protein subunits were separated by reverse-phase HPLC. The effluent was monitored at both 220 and 280 nm. Integration of the absorbance peaks gives the relative areas corresponding to the gene 44 and 62 proteins.

**Fig. 2**

Chromatograms of HPLC separation of gene 44 and 62 proteins. The elution profile of the reverse-phase HPLC separation of gene 44 and 62 protein is shown, with the detection wavelength set at 220 nm (A) and 280 nm (B). Fractions were collected and analyzed by SDS-PAGE, identifying the faster eluting species to be gene 62 protein and the slower to be gene 44 protein (data not shown).

5 A molecular mass for this protein complex was reported by Barry et al. (26), using sucrose density gradient sedimentation and gel filtration. They obtained a sedimentation constant of 7.1 S, and a calculated molecular mass of 164,000 daltons, in excellent agreement with our current study. The stoichiometry of the complex was at the time, however, incorrectly thought to be 2:1 (44–62 subunits). Therefore Barry et al. (26) concluded that the actual molecular mass of the complex must be 176,000 daltons (i.e. a 4:2 subunit stoichiometry) based on apparent molecular masses of 34 and 20 kDa for gene 44 and 62 proteins, respectively.
Several hypothetical subunit ratios of the gene 44/62 protein complex are listed, along with predicted molecular weights. The list focuses on those ratios most consistent with the current data as well as estimates in the literature of the molecular weight of the native complex. The results from the HPLC separation of the subunits (Table II), which yields subunit stoichiometries, are compared with those from the sedimentation and light scattering experiments, which provides molecular weight information. For each type of experiment, (+) indicates that the results were consistent with the subunit composition in question, (–) indicates that the composition is unlikely based on the experiment, and (±) indicates that the results of the experiment clearly rule out the subunit composition. Thus, it can be seen that a complex consisting of four gene 44 protein subunits and one gene 62 protein subunit, having a molecular mass of 163,700 daltons, is the only one consistent with both the HPLC and the sedimentation data.

| Subunit composition (G44P:G62P) | Molecular mass | HPLC | Sedimentation |
|---------------------------------|----------------|------|---------------|
| 3:2                             | 150,400        | –    | +             |
| 4:2                             | 185,000        | –    | +             |
| 5:2                             | 220,600        | –    | –             |
| 3:1                             | 128,100        | ±    | –             |
| 4:1                             | 163,700        | +    | +             |
| 5:1                             | 199,500        | +    | +             |

Shape Factors Indicate an Asymmetric Complex—We can make some predictions about the shape of the gene 44/62 protein complex on the basis of the sedimentation coefficient. A frictional coefficient for the complex can be calculated using the relationship:

\[ s = \frac{M(1 - \frac{\omega^2}{\eta})}{Nf} \]  

(5)

We can also calculate the frictional coefficient expected for an anhydrous sphere of the same size. Using this approach we obtain a maximum value of 1.45 for the Perrin shape factor, \( F \) (for a general discussion see Ref. 27). If we assume that the hydration of the gene 44/62 protein complex falls in the range of 0.3-0.4 g of water/g of protein, as is typical for most proteins, then the shape factor, \( F \), is about 1.25. This corresponds to a prolate ellipsoid with an axial ratio of about 5:1. This agrees with the shape factor reported by Barry and Alberts (9) for this complex, based on sucrose density gradient sedimentation and gel filtration. Thus, it appears that the complex either has a very asymmetric shape or an unusually high degree of hydration. Since such a degree of hydration is very unlikely, a relatively asymmetric complex seems the most plausible explanation of this shape factor.

Gene 45 Protein Associates to Form a Trimer

Gene 45 protein has a monomer molecular mass, based on the DNA sequence, of 24.7 kDa (14). Previous literature reports had suggested that the protein associates to form a dimer in solution (10, 28) corresponding to a calculated molecular mass of 49 kDa. Gene 45 protein elutes with the void volume on a Bio-Gel P-60 column, for which the exclusion limit is approximately 60 kDa. This could be consistent with an asymmetric dimer model but is more likely to be indicative of higher states of gene 45 protein association. We have undertaken a more rigorous analysis of the physical state of gene 45 protein in order to ascertain whether the protein exists principally as a dimer, or as some higher associated species.

\[ ^a \] J. Hockensmith, unpublished results.
peaks of chemical cross-linking. Protein monomer (in the DMS cross-linked lanes) is probably the very low result for DMS cross-linking of gene theoretical structures that might yield the cross-linked products labeled monomer. Protein concentrations used in this study are low enough (4 mg/ml) to make the probability of collisional cross-linking to form collisional cross-links. Thus, we can feel fairly confident that under identical reaction conditions, NusA shows no tendency of gene species of gene pre-existing in the solution? Although it can be assumed that the overall trimer is quite compact. In fact, the actual trimer is quite asymmetric, and perhaps one simple way to visualize this in the context of Fig. 3B is to consider the cross-linking species as representing end-on views of quite asymmetric (cylindrical?) subunits.

The schematic cross-linking diagrams of Fig. 3B are merely intended to represent subunit cross-linking connectivity and not to suggest that the overall trimer is quite compact. In fact, (see following section) the actual trimer is quite asymmetric, and perhaps one simple way to visualize this in the context of Fig. 3B is to consider the cross-linking species as representing end-on views of quite asymmetric (cylindrical?) subunits.

Shape Information from the Sedimentation Coefficient—Velocity sedimentation of gene 45 protein yields a sedimentation coefficient, $S_{20,w}$, of 3.9 ± 0.1 S. Fig. 4 shows the apparent (uncorrected) S values plotted versus $t_w$ for seven values of $w = C(r)/C_m$, the ratio of the concentration of protein in the ultracentrifuge cell at radius r to the plateau concentration. This analysis, described by Van Holde and Weischet (19), results in a “fan plot” for a homogeneous sample, because at infinite time, the apparent $S_w$ values converge to the same limit, S. Therefore, it appears that the trimer of gene 45 protein exists as a fairly homogeneous species. If we assume

It should be noted that if the subunits of the trimer associate and dissociate on a rapid time scale, the resulting sedimentation boundary could appear to be homogeneous. In this case, the sedimentation constant obtained would be somewhat lower than that for the true trimer species, being an intermediate value between, for example, monomers and trimers. Since no indication of lower molecular weight species was seen in equilibrium sedimentation, however, we have no reason to believe that any significant fraction of the gene 45 protein exists as monomers or dimers under the solution conditions we have studied.

---

S. C. Gill and P. H. von Hippel, manuscript in preparation.

No change in the cross-linking pattern or efficiency was seen on addition of DNA or ATP, indicating that the association state of the protein is probably not affected by binding of either of these potential cofactors (data not shown).
that the hydration is between 0.3 and 0.4 g water/g protein, then the Perrin shape factor for gene 45 protein would be about 1.32, corresponding to an axial ratio of 61 for a prolate ellipsoid. Thus, it appears that the trimer gene 45 protein about 1.32, corresponding to an axial ratio of 61 for a prolate ellipsoid. The literature reports identifying gene 45 protein as a dimer may also be quite asymmetric in shape. Since the literature details and references, see (8)). Piperno et al. (31). Gene 45 protein treated in this manner is able to stimulate the ATPase activity of normal gene 44/62 protein, while modified gene 44/62 protein shows no ATPase activity in the complex. This agrees with the results of Lin et al., that gene 44 protein reacts specifically with N3ATP. Since the substrate for ATP hydrolysis is Mg-ATP, this lends credence to the idea that N3ATP is binding to the true ATP-binding site in a manner similar to the normal substrate. Gene 62 protein shows virtually no labeling. Thus, it appears that the ATP hydrolysis activity resides in the gene 44 subunits of the complex. This agrees with the results of Lin et al. (3), showing that gene 44 protein purified from the cloned T4 gene has a low level DNA-dependent ATPase activity, while gene 62 protein (also purified from a clone) does not. The only other protein to show specific photoaffinity labeling under these conditions is the polymerase (gene 43 protein). Although it lacks any detectable ATPase activity, the polymerase clearly must contain a deoxynucleoside triphosphate-binding site, and this site probably has considerable affinity for N3ATP. None of the other proteins show any significant labeling by the photoaffinity probe. In particular, the lack of labeling of gene 45 protein supports the conclusion of Piperno et al. (3), that the gene 44/62 protein complex is responsible for the ATPase activity of the accessory proteins.

The ATP-binding Site Resides in the Gene 44 Subunit

The T4 polymerase accessory proteins have been shown to have an ATPase activity that is stimulated by DNA (for details and references, see (8)). Piperno et al. (3), examined the identity of the ATPase protein by treating gene 44/62 protein and gene 45 protein in turn with 6-mercaptourine ribonucleoside 5'-triphosphate, which reacts at the ATP binding site of various ATPases and inhibits ATPase activity (31). Gene 45 protein treated in this manner is able to stimulate the ATPase activity of normal gene 44/62 protein, while modified gene 44/62 protein shows no ATPase activity in the presence of normal gene 45 protein. Thus, the functional site for ATP hydrolysis appeared to reside in the gene 44/62 protein complex.

We have further investigated the locations of potential ATP-binding sites on the replication proteins by covalently labeling the proteins with an ATP photoaffinity analog, N3ATP. Aromatic azido compounds such as N3ATP form highly reactive nitrines when exposed to ultraviolet light. Photoactivation is possible outside the normal absorption range of most biological macromolecules (i.e., above 300 nm). Thus, absorption of the incident light responsible for photoactivation is likely to be minimal and photochemical reaction of the macromolecules themselves to be slight. Covalent bond formation between the affinity probe and a macromolecule is likely if the probe is actually bound at the moment of photoactivation. Ideally, unbound affinity probe will react with solvent, although earlier estimates of solution lifetimes on the order of milliseconds (33) have been challenged by the work of Staros (34), suggesting that the lifetimes of reactive intermediates may be on a many millisecond to second time scale. However, under proper reaction conditions photoaffinity labeling can nonetheless be made quite specific.

**Gene 44 Protein Reacts Specifically with Azido-ATP**—Samples of each of the five T4 replication proteins (gene 44, 62, 32, 45, and 43 proteins), alone and in combination, were irradiated with 300 nm light in the presence of [γ-32P]N3ATP. The results are shown in Fig. 5. The autoradiogram shows that gene 44 protein reacts specifically with N3ATP. Since the gene 44/62 protein complex is the only one among the group for which ATPase activity has been demonstrated, it is not surprising that one of its subunits binds ATP. This specific labeling is notably absent when no Mg2+ is present. Since the substrate for ATP hydrolysis is Mg-ATP, this lends credence to the idea that N3ATP is binding to the true ATP-binding site in a manner similar to the normal substrate. Gene 62 protein shows virtually no labeling. Thus, it appears that the ATP hydrolysis activity resides in the gene 44 subunits of the complex. This agrees with the results of Lin et al. (3), that the gene 44 protein complex is responsible for the ATPase activity of the accessory proteins.

**Evidence That Photoaffinity Labeling Occurs at the Normal ATP-binding Site**—In order to demonstrate unequivocally that the observed labeling of gene 44 protein occurs specifically at (or near) the normal ATP-binding site, several criteria should be met. The most important criterion is that ATP can compete for the binding site, and reduce the amount of

---

11 Lin, T. C., Rush, J., McKim, I., and Konigsberg, W. (1985) Abstracts from the 1985 Evergreen International T4 Meeting, August, 1985, Olympia, WA.
binding sites reside in the gene 44 protein subunits. Thus, the complex has the potential capability to bind up to four ATP molecules at once.

The gene 45 protein is seen to exist primarily as a trimer in dilute solution, contrary to previous reports in the literature (10, 28). Since the association constant between the gene 44/62 protein complex and gene 45 protein is relatively weak (8) it has not been possible to isolate the complete accessory protein complex. Therefore, the stoichiometry of binding of gene 45 protein trimers to gene 44/62 protein complexes can only be surmised from enzymatic studies. This will be discussed in greater detail in the accompanying paper. In addition, the interaction of the accessory protein complex with primer-template junction DNA (the functional site for replication) will be discussed.

Acknowledgments—We are very grateful to Johannes Geiselman for extensive assistance with the dynamic laser light-scattering experiments and to Stanley C. Gill for useful discussions of the calculation of molar extinction coefficients and partial specific volumes from amino acid sequence information.

REFERENCES

1. Nossal, N. G., and Peterlin, B. M. (1979) J. Biol. Chem. 254, 6032-6036
2. Sinha, N. K., Morris, C. F., and Alberts, B. M. (1980) J. Biol. Chem. 255, 10099-10103
3. Piperno, J. R., Kalen, R. G., and Alberts, B. M. (1978) J. Biol. Chem. 253, 5180-5186
4. Mace, D. C., and Alberts, B. M. (1984) J. Mol. Biol. 177, 270-293
5. Newport, J. W., Kowalczykowski, S. C., Lomberg, N., Paul, L. S., and von Hippel, P. H. (1980) in Mechanistic Studies of DNA Replication and Genetic Recombination (ICN-UCLA Symp. Mol. Cell. Biol. (Alberts, B. M., ed) Vol. 19, pp. 485-505, Academic Press, New York
6. Piperno, J. R., and Alberts, B. M. (1978) J. Biol. Chem. 253, 5174-5179
7. Mace, D. C., and Alberts, B. M. (1984) J. Mol. Biol. 177, 315-327
8. Jarvis, T. C., Hockenmuth, J. W., Paul, L. S., and von Hippel, P. H. (1980) J. Biol. Chem. 254, 12717-12729
9. Barry, J., and Alberts, B. M. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 2727-2729
10. Morris, C. F., Hama-Imaba, H., Mace, D., Sinha, N. K., and Alberts, B. M. (1979) J. Biol. Chem. 254, 6787-6796
11. Spicer, E. R., Nossal, N. G., and Williams, K. R. (1984) J. Biol. Chem. 259, 16425-16432
12. Spicer, E. R., and Königberg, W. H. (1983) in Bacteriophage T4 (Mathews, C. K., Kutter, E. M., Mosig, G., and Berget, P. B., eds) pp. 291-301, American Society for Microbiology, Washington, DC
13. Nossal, N. G. (1979) J. Biol. Chem. 254, 6026-6031
14. Spicer, E. R., Noble, J. A., Nossal, N. G., Königberg, W. H., and Williams, K. R. (1982) J. Biol. Chem. 257, 8792-8797
15. Blümer, M., Burke, R. L., and Alberts, B. M. (1979) J. Biol. Chem. 254, 9660-9677
16. Alberts, B., and Frey, L. (1970) Nature 227, 1313-1318
17. Cohn, E. J., and Edsall, J. T. (1943) in Proteins Amino Acids and Peptides, pp. 370, Academic Press, New York
18. Edelhoch, H. (1967) Biochemistry 6, 1948-1954
19. Van Holde, K. E., and Weischat, W. O. (1978) Biopolymers 17, 1387-1403
20. Bloomfield, V. A., and Lim, T. C. (1978) Methods Enzymol. 56, 410-494
21. Yphantis, D. A. (1964) Biochemistry 3, 257-311
22. Laemmli, U. K. (1970) Nature 227, 680-685
23. Morrissey, J. H. (1981) Anal. Biochem. 117, 507-510
24. Schachman, H. K. (1959) Ultracentrifugation in Biochemistry, Academic Press, New York
25. Van Holde, K. E. (1975) in The Proteins (Neurath, H., and Hill, R. L., eds) 3rd Ed., Vol. 1, pp. 225-291, Academic Press, New York
26. Barry, J., Hama-Imaba, H., Muro, L., and Alberts, B. (1973) in DNA Synthesis in Vivo (Wells, R. D., and Inman, R. B., eds) pp. 196-214, University Park Press, Baltimore
27. Cech, T. R., and Szostak, J. W. (1980) Biochemical Physics, Part II, W. H. Freeman and Co., San Francisco
28. Alberts, B. M., Morris, C. F., Mace, D., Sinha, N., Bitter, M., and Moran, L. (1978) in DNA Synthesis and Its Regulation ICN-UCLA Symp. Mol. Cell. Biol. (Goulian, M., and Hanawalt, P., eds) Vol. 3, pp. 241-298, W. A. Benjamin, Inc., Menlo Park
29. Davies, G. E., and Stark, G. R. (1970) Proc. Natl. Acad. Sci. U.S.A. 66, 651-656
30. Finger, L. L., and Richardson, J. P. (1982) J. Mol. Biol. 156, 203-219
31. Patel-Wenzeler, S., Pauls, H., Erdmann, E., and Schoner, W. (1975) Eur. J. Biochem. 55, 301
32. Czarnecki, J., Uskulak, R., and Haley, B. (1979) Methods Enzymol. 56, 642-653
33. Nachman, M. W., and Fiorini, J. R. (1973)Anal. Biochem. 55, 338-339
34. Starus, J. V. (1980) Trends Biochem. Sci. 5, 320-322
35. Haley, B. E., and Hoffman, J. F. (1974) Proc. Natl. Acad. Sci. U.S.A. 71, 2367-2371

Table IV

Covalent incorporation of azido-ATP by the gene 44/62 protein complex

| Reaction components | Cpm* |
|---------------------|------|
| G4/62P              | 157  |
| G4/62P (no irradiation) | 23  |
| G4/62P + ATP        | 49   |
| G4/62P + poly(AD)   | 153  |

* Counts per min of 32P covalently incorporated into the gene 44/62 protein complex.