Proc. Nat. Acad. Sci. USA
Vol. 69, No. 9, pp. 2741–2744, September 1972

Self-Association of Gene-32 Protein of Bacteriophage T4

(stable dimer/sedimentation equilibrium centrifugation/gel electrophoresis/replication fork model/recombination/*E. coli*)

ROBERT B. CARROLL, KENNETH E. NEET, AND DAVID A. GOLDTHWAIT*

Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106

Communicated by Harland G. Wood, June 26, 1972

ABSTRACT **The self-association of gene-32 protein has been studied by sedimentation equilibrium centrifugation and polyacrylamide gel electrophoresis, in order to better understand its role in DNA replication and genetic recombination. The monomer molecular weight of gene-32 protein is 38,000 in guanidine hydrochloride and 34,000 in sodium dodecyl sulfate, in agreement with the results of Alberts and coworkers. Stable dimers of gene-32 protein occur under various conditions, among which are high ionic strength and pH 10. The occurrence of stable dimers under some conditions and higher aggregates under others indicates there are two types of protein–protein interactions occurring in gene-32 protein self-association. The association that occurs above about 0.1 mg/ml concentration of protein produces at least decamers.**

A model for the DNA replication fork is postulated that requires the two different interactions that occur in gene-32 protein aggregation. In the model, gene-32 protein holds the two strands of the DNA duplex in a conformation that prevents their reannealing and, therefore, facilitates replication and recombination.

The protein product of gene 32 of phage T4, isolated and characterized by Alberts and his coworkers, is required for DNA synthesis (1–4) and genetic recombination (4–6) in T4-infected *Escherichia coli*. Gene-32 protein at high concentration has a much greater affinity for single-stranded DNA than at low concentrations (2, 3, 7). This phenomenon indicates that a protein–protein interaction also occurs. The protein–protein interaction may allow the gene-32 protein to hold the DNA in a rigid linear conformation, thus accounting for its ability to increase the rate of renaturation of DNA (7).

The gene-dosage experiments of Sinha and Snustad and of Snustad (9, 10) have demonstrated a stoichiometric requirement for gene-32 protein in DNA synthesis *in vivo*. Since a large proportion of the intracellular protein produced during T4 infection is gene-32 protein (2), it has been proposed (7) that a certain number of interacting gene-32 protein molecules may be required for each replication fork. Alberts and coworkers have proposed a model of the replication fork involving the linear aggregation of gene-32 protein along the phosphodeoxyribose backbone of each of the two separated DNA strands (8).

The *in vitro* aggregation of gene-32 protein, first reported by Alberts and coworkers (7, 8) has been studied in our laboratory by sedimentation equilibrium centrifugation and polyacrylamide gel electrophoresis. We have been able to separate the aggregation of this protein into two different types of bonding; an association of monomers to give dimers, and a further association of dimers to large aggregates. We propose a variation of Alberts' model for the replication fork based on this mechanism of aggregation of gene-32 protein.

* To whom reprint requests should be addressed.

MATERIALS AND METHODS

Materials. E. coli B strains SW1485 (nonpermissive for T4M41) and CR63 (permissive for T4M41 as a result of an amber suppressor) were the gifts of L. Astrachan. *E. coli* SW1485 was infected with phage T4M41 (lysozyme minus).

Growth of Bacteriophage and Preparation of T4-Infected E. coli. A 45-liter culture of *E. coli* CR63 (permissive strain) was grown in Frazier–Jarrel medium to 3.5×10^8 cells per ml at 37° and infected with phage T4M41 at a multiplicity of infection of 0.001. The infected culture was incubated until lysis occurred. The phage were then harvested by the method of Salser (details available from Winston Salser, personal communication).

E. coli SW1485 infected with bacteriophage T4 M41 were prepared at the New England Enzyme Center by addition to 500 liters of *E. coli* SW1485 grown to 10^9 cells per ml in Frazier–Jarrel medium (11), with 5×10^{15} phage (multiplicity of infection of 10) and incubation for 1 hr at 37°. The cells were then harvested in a Sharples centrifuge and frozen.

Preparation of Gene-32 Protein. The initial steps of the preparation were identical to the complete procedure of Alberts and Frey (7), except that the preparation was scaled up to 200 g of infected cells. An additional step was introduced, because a difficulty was encountered in obtaining a homogeneous protein with only the preceding steps. A 46 × 2 cm Sephadex G-100 column was equilibrated with 2 M NaCl–buffer B [buffer B of the Alberts & Frey procedure (7)]. The 0.2 M NaCl–buffer B eluate of the preceeding DEAE-cellulose column was adjusted to 2 M in NaCl and applied to the column, which was then eluted with 2 M NaCl–buffer B. The gene-32 protein was eluted shortly after the breakthrough volume and homogeneous as determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Several contaminating protein bands were eliminated in this way, apparently by preventing their binding to gene-32 protein.

Sedimentation Equilibrium Centrifugation of Gene-32 Protein. The aggregation of gene-32 protein was studied in a Spinco model E analytical ultracentrifuge equipped with interference optics. The protein was dialyzed overnight against a standard "sedimentation equilibrium buffer": 0.02 M Tris·HCl (pH 8.1)–1 mM 2-mercaptoethanol–0.10 M KCl, unless

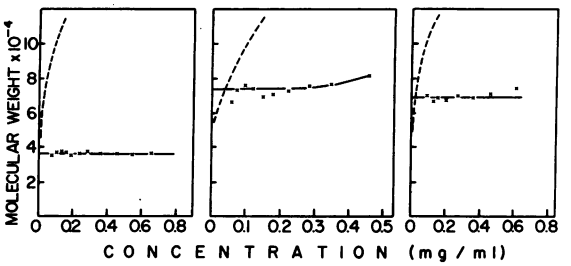

FIG. 1. Monomer and dimer molecular weights of gene-32 protein. The molecular weight of gene-32 protein was determined by sedimentation equilibrium centrifugation in the presence of: (*left*) 6.0 M guanidine hydrochloride, (*center*) sedimentation equilibrium buffer adjusted to pH 10, (*right*) 2.0 M KCl. The control, without perturbant, is represented by *dashed lines* (– – –, see Fig. 3), the experimental sample by *solid lines* (——).

otherwise noted. The dialyzed protein was centrifuged in a 6-mm or 30-mm double sector cell or in a 12-mm, 3-channel Yphantis cell. The meniscus depletion method of Yphantis (12) was used for most experiments, but some experiments were performed at lower speeds utilizing the white-light fringe (13). The attainment of equilibrium was followed as a function of time. The interference plates were measured on a Nikon microcomparator, and the molecular weight was determined as a function of protein concentration, expressed in fringe displacement, by use of the equation:

$$\bar{M}_w(r) = \frac{2RT}{(1-\bar{v}\rho)\omega^2}\,\frac{2.303\,(d\log c)}{d(r^2)}$$

$\bar{M}_w$ is the weight-average molecular weight; R, gas constant; T, temperature; $\bar{v}$, partial specific volume; ρ, density of the solution; ω, angular velocity; c, concentration of protein; and r, distance from center of rotation. A program for a Hewlett–Packard Calculator 9100 was used to calculate the least-squares line through each set of five successive points of the log c and r^2 data; $\bar{v}$ was calculated from the aminoacid composition (14).

Polyacrylamide Gel Electrophoresis. Electrophoresis was in cylindrical gels by published techniques: polyacrylamide gel electrophoresis (15) (pH changed to 9); sodium dodecyl sulfate–polyacrylamide gel electrophoresis (16), and urea–gel electrophoresis (17). All gels were stained with 0.25% Coomassie Blue in methanol–acetic acid–water (5:1:5) and were destained electrophoretically in 7% acetic acid.

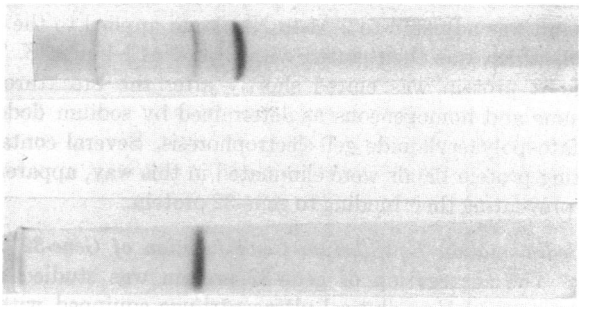

FIG. 2. Sodium dodecyl sulfate– and urea–polyacrylamide gel electrophoresis of gene-32 protein. 5 μg of gene-32 protein was applied to sodium dodecyl sulfate (*lower*) and 6 M urea (pH 9) (*upper*) gels.

EXPERIMENTAL RESULTS

Resolution of monomeric and dimeric forms of gene-32 protein

The molecular weight of gene-32 protein has been determined in the presence of KCl, triethylamine hydrochloride, urea, guanidine hydrochloride, glycerol, and sedimentation buffer at pH 10. The molecular weight of gene-32 protein in 6 M guanidine hydrochloride at 20° is 38,000 (Fig. 1 *left*). The $\bar{v}$ of gene-32 protein in this solvent was determined by variation of the density of the solution by the addition of D_2O. The apparent partial specific volume $\bar{v}$ was obtained by the method of Edelstein and Schachman (18); it was 0.72. Preferential solvation appears to be minimal as the same molecular weight was obtained in 8 M guanidine hydrochloride (the same $\bar{v}$ was assumed).

Adjustment of the buffer to pH 10 or centrifugation in the presence of high salt allows dimerization but prevents the formation of higher aggregates, as the molecular weights obtained are (Fig. 1 *center* and *right*) 74,000 and 69,000. The partial specific volume, $\bar{v}$, in aqueous solution was estimated from the aminoacid composition (unpublished observations) as 0.74. The molecular weight obtained in KCl is uncorrected for any preferential solvation, which would tend to decrease the apparent molecular weight. The aqueous value of $\bar{v}$ was assumed for calculations in KCl. The degree of association as a function of salt concentration was determined; the values illustrated (Fig. 1) represent a minimum aggregation. This experiment indicates that high salt or high pH prevents the formation of higher aggregates of gene-32 protein, but do not dissociate the dimer (perhaps even favor the dimer), while guanidine hydrochloride prevents all association of monomers. 5% Glycerol (not shown) inhibits aggregation of the protein at low concentrations, and favors the monomer. This effect is noted in older preparations, which have lowered association constants. In fresh preparations, with high association constants, no monomer is seen in glycerol, but some inhibition of aggregation does occur.

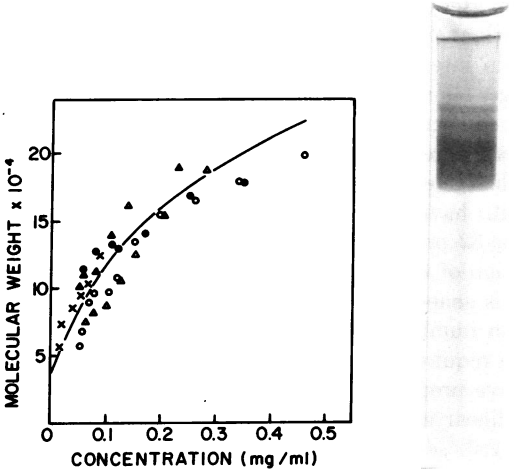

FIG. 3. The self-association of gene-32 protein (*left*) in an analytical ultracentrifuge or (*right*) on polyacrylamide gels. (*Left*) Gene-32 protein at an initial concentration of 1.06 mg/ml was centrifuged at 18,000 rpm in a 30-mm double-sector cell (×). Gene-32 protein at initial concentrations of 1.00 mg/ml (○), 0.75 mg/ml (●), and 0.50 mg/ml (△) was centrifuged in a 12-mm, three-channel Yphantis cell at 22,000 rpm. Gene-32 protein at an initial concentration of 1.00 mg/ml was centrifuged in a 12-mm double-sector cell at 26,000 rpm (▲). (*Right*) 200 μg of gene-32 protein was electrophoresed in an 8% polyacrylamide gel.

The behavior of gene-32 protein on sodium dodecyl sulfate- and urea-polyacrylamide gels

Gene-32 protein prepared by the procedure of Alberts and Frey (7), then eluted from Sephadex G-100 in 2 M NaCl, migrated in sodium dodecyl sulfate–polyacrylamide gels (Fig. 2) as a single band, indicating a homogeneous molecular weight. A molecular weight of 34,000–35,000 was estimated by the method of Weber and Osborn (16), in agreement with the results of Alberts (3).

This same preparation migrated in urea gels at 4° as two bands. The faster band migrated just slightly slower than chymotrypsinogen, while the slower band migrated with a mobility expected for a dimer of the fast band, according to the relationship: log $M \propto$ mobility. As the porosity of the urea gels was increased between 10 and 5% acrylamide, the mobility of the slow band increased relative to that of the fast band until they converged, indicating that the slow band is a dimer and not a species of different charge. When 10% urea gels were made 0.1% in sodium dodecyl sulfate, only one band of gene-32 protein was found.

Self-association of gene-32 protein

The aggregation of gene-32 protein, previously demonstrated by Alberts in sucrose gradients (8), was studied by sedimentation equilibrium centrifugation and polyacrylamide gel electrophoresis (Fig. 3). The increase in weight-average molecular weight, $\bar{M}_w$, (Fig. 3, *left*) indicates that the protein is in an association–dissociation equilibrium. The overlap of the data from 6-mm, 12-mm, and 30-mm cells, at various angular velocities and initial concentrations, indicates that there is no dependence of the aggregation upon pressure and that the association is readily reversible. The slope of the molecular weight versus concentration curve is quite steep, indicating a marked sensitivity of the association to protein concentration in this range. The pH optimum of the aggregation is about 8 (unpublished observations); the ionic strength optimum is about 0.1 (unpublished observations). Higher ionic strength produces a reversible decrease in the association, while very low ionic strength produces an irreversible decrease, and the denatured protein exists in the dimeric form. We have estimated the association constant for each step of an indefinite association (at pH 8.1 and $\mu = 0.1$) to be $K = 6 \times 10^5\ M^{-1}$, which would mean the free energy of association is about $\Delta G = -7.8$ Cal/mol. The gene-32 protein used in the experiment illustrated in Fig. 3 (*left*) was freshly prepared. The ability of the protein to associate to higher aggregates decreases upon storage, while association to dimers is undiminished. The upper molecular weight limit of the association is currently under investigation, but at least decamers have been observed.

When polyacrylamide gels are loaded with a high concentration of gene-32 protein, the various molecular weight aggregates of gene-32 protein are separated (Fig. 3, *right*). This same sample gives a single band on sodium dodecyl sulfate–polyacrylamide gels (Fig. 2), indicating the presence of a single polypeptide chain. The bands decrease in concentration and are closer together as the molecular weight of the aggregates increases. This indicates that the aggregates follow the relationship: log $M \propto$ mobility. The presence of discrete protein bands instead of a smear indicates that interconversion between the various aggregates must be rather slow under these conditions.

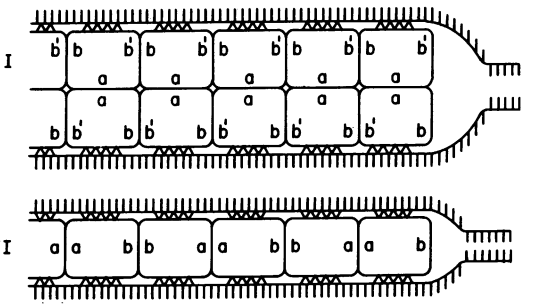

FIG. 4. "Bases up" or "bases out" model of gene-32 protein and DNA at a replication fork. The *rectangular bodies* represent gene-32 protein; the *short vertical lines* represent the plane of the DNA bases, and the *long lines* represent the phosphoribose backbones in the helical DNA regions and at the replication fork. The two protein–protein interactions are represented by *aa* and *bb*. The protein–DNA interaction is indicated by a *wavy line*.

DISCUSSION

The determination of a molecular weight of 34,000–38,000 for gene-32 protein in sodium dodecyl sulfate or guanidine hydrochloride on the one hand, and the determination of a molecular weight twice as large at high salt concentrations on the other hand, indicate that gene-32 protein can occur as either a monomer or as a dimer in a stable, unaggregated form.

Under appropriate conditions, the dimers further associate to at least decamers. The fact that association of dimers to higher aggregates is prevented by high concentrations of salt, conditions under which dimers are not dissociated, indicates that the bonds between dimers, necessary for aggregation, do not have the same properties as those that hold two monomers together as a dimer. Dimers may only be dissociated under the very different conditions of treatment with either sodium dodecyl sulfate or guanidine hydrochloride, perhaps indicating that this is primarily a hydrophobic interaction.

We propose a model for the replication fork (Fig. 4, model I) based upon the monomer–dimer–higher aggregation form of the association of gene-32 protein. The model requires that there be two types of interaction, as we have demonstrated, between gene-32 protein monomers. The model previously proposed by Alberts and coworkers (2, 3) requires that only one type of bonding exist between monomers. Model I (Fig. 4) requires the following postulates: (*i*) gene-32 protein associates as monomer–dimer–higher aggregates on the DNA replication fork, (*ii*) there are two or four DNA-binding sites per dimer, (*iii*) the bonds between monomers in the dimer are perpendicular to the axis of the large linear aggregate and, thus, to the bonds between dimers. Although it has not been demonstrated that gene-32 protein aggregates are linear in the absence of DNA, Alberts and coworkers (7, 8) have demonstrated that the aggregation occurs linearly on binding to the DNA molecule. As it has been indicated that gene-32 protein binds along the phosphodeoxyribose backbone of DNA (7), we propose that the bases may be rotated up or outward, "bases-up conformation," and that the phosphodeoxyribose backbones of the two strands are held on the inside of the structure by the dimer of gene-32 protein. Since gene-32 protein is unable to denature GC-rich regions of DNA, we suggest that there is a protein(s) responsible for separating base pairs ahead of the gene-32 protein, allowing

the bases to rotate into the "bases-up conformation." An alternative model (Fig. 4, model II) would differ in postulate *iii*, in that the two types of protein–protein bonds would lie on the same axis.

The efficacy of this model for DNA synthesis would be that the two parental strands would be held partially separated from each other, with no tendency to reform hydrogen bonds and would, therefore, be free to base-pair with incoming nucleotide precursors. The efficacy for recombination would be that the two strands would be held in a conformation in which they could not re-establish base pairing, and this conformation would facilitate the search for complementary strands, perhaps in a similar structure, with which to form a recombinant (synapsis). This structure could thus be common to both replication and recombination. This "bases-up" model of the replicative fork requires the unique properties of the association of gene-32 protein that have been described in this paper. Additional regulatory features of DNA replication and recombination could result from effects on the bonds involved in monomer–dimer association or those involved in the dimer–higher aggregation.

We thank Jean Bleistein for technical assistance, Charles Starling and Sam Zito for operating the ultracentrifuges, and Dr. Howard Bensusan for the aminoacid analysis. This work was supported by grants from the NIH (CA-11322 and 5-K06 GM 21444) and from the American Cancer Society (NP-13A, In-57-J, & PRA-55).

1. Epstein, R. H., Bolle, A., Steinberg, C. M., Kellenberger, E., Boy de la Tour, E., Chevalley, R., Edgar, R. S., Susman, M., Denhardt, G. H. & Lielausis, A. (1963) *Cold Spring Harbor Symp. Quant. Biol.* **28**, 375–394.
2. Alberts, B. M., Amodio, F. J., Jenkins, M., Gutmann, E. D. & Ferris, F. L. (1968) *Cold Spring Harbor Symp. Quant. Biol.* **33**, 289–305.
3. Alberts, B. M. (1970) *Fed. Proc.* **29**, 1154–1163.
4. Kozinski, A. N. & Felgenhauer, Z. Z. (1967) *J. Virol.* **1**, 1193–1202.
5. Tomizawa, J., Annaku, N. & Iwama, Y. (1966), *J. Mol. Biol.* **21**, 247–253.
6. Tomizawa, J. (1967) *J. Cell. Physiol.* **70**, Suppl. 1, 201–214.
7. Alberts, B. M. & Frey, L. (1970) *Nature* **227**, 1313–1318.
8. Huberman, J. A., Kornberg, A. & Alberts, B. M. (1971) *J. Mol. Biol.* **62**, 39–52.
9. Sinha, N. K. & Snustad, D. P. (1971) *J. Mol. Biol.* **62**, 267–271.
10. Snustad, D. P. (1968) *Virology* **35**, 550–563.
11. Fraser, D. & Jerrel, E. A. (1953) *J. Biol. Chem.* **205**, 291–295.
12. Yphantis, D. A. (1964) *Biochemistry* **3**, 297–317.
13. Goldberg, M. E. & Edelstein, S. J. (1969), *J. Mol. Biol.* **46**, 431–440.
14. Cohn, E. J. & Edsall, J. T. (eds) (1943) in *Proteins, Amino Acids & Peptides* (American Chemical Society, Washington, D.C.), p. 370.
15. Jovin, T., Chrambach, A. & Naughton, M. A. (1964) *Anal. Biochem.* **9**, 351–369.
16. Weber, K. & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406–4412.
17. Walter, W. G., Scifert, W. & Zillig, W. (1968) *Biochem. Biophys. Res. Commun.* **30**, 240–247.
18. Edelstein, S. J. & Schachman, H. K. (1967) *J. Biol. Chem.* **242**, 306–311.