MOLECULAR ASPECTS OF THE INTERACTIONS OF T4-CODED GENE 32-PROTEIN AND DNA POLYMERASE (GENE 43-PROTEIN) WITH NUCLEIC ACIDS[1]

John W. Newport[2], Stephen C. Kowalczykowski, Nils Lonberg[3], Leland S. Paul, and Peter H. von Hippel

Institute of Molecular Biology and Department of Chemistry
University of Oregon, Eugene, Oregon 97403 U.S.A

ABSTRACT The structure and overall function of the T4 DNA replication complex has been outlined by Alberts and coworkers. In order to further refine this picture we have been studying the molecular details of the interactions of some of the proteins of this replication system with relevant nucleic acid lattices, and with one another. We have shown that gene 32-protein binds to short (ℓ = 2→8 residues) *oligo*nucleotides essentially independently of base composition or sugar type; this binding is also relatively independent of salt concentration. In contrast, the cooperative binding of gene 32-protein to *poly*nucleotides shows an appreciable dependence on base composition and sugar-type, and a large dependence on salt concentration. This salt concentration dependence resides in the binding constant to the nucleic acid lattice (K), and not in the cooperativity parameter (ω); it has been shown that this salt concentration dependence involves a significant anion, as well as a cation, displacement reaction on binding. These results are interpreted in terms of an explicit two-conformation model of the interaction of this protein with nucleic acid lattices. In addition, the results provide a quantitative molecular interpretation

[1]This work was supported in part by USPHS Research Grant GM-15792, by USPHS Training Grants GM-00715 and GM-07759 (which provided predoctoral stipends for JWN and LSP), and by American Cancer Society Postdoctoral Fellowship PF-1301 (to SCK).

[2]Present address: Department of Biochemistry and Biophysics, University of California San Francisco, CA 94143.

[3]Present address: Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.

ISBN 0-12-048850-7

of the autogenous regulation by this protein of its own synthesis, and lead to general principles for the development of binding specificity via cooperative (cluster) protein binding.
T4 DNA replication *in vivo* involves at least seven proteins. We have characterized the binding properties of the T4 DNA polymerase to various DNA substrates, and present a model which describes how T4 DNA polymerase might bind to the primer-template substrate during replication based on these results. At *in vivo* salt concentrations we have shown that T4 DNA polymerase, by itself, synthesizes DNA "dispersively"; synthesis in the "progressive" mode requires T4 DNA polymerase, gene 32-protein, the proteins encoded by T4 genes 44, 62 and 45, and ATP. ATP hydrolysis is required only for the initial assembly of these proteins into a multiprotein complex. This complex has a lifetime of less than 45 seconds. A model bearing on some molecular aspects of the structure and possible function of this "five-protein" system is discussed.

INTRODUCTION

The identification and characterization of the components of several viral and bacterial DNA replication complexes, as well as the assembly of these components into functional *in vitro* replication systems, are advancing rapidly in a number of laboratories (for recent summaries see refs. 1-4, as well as related articles in this volume). Alberts and coworkers, in particular, have led the way in defining the bacteriophage T4 system, and have shown that the equivalent of "leading strand" elongation of DNA primers can be conducted *in vitro*, at essentially *in vivo* rates, using an appropriate DNA template-primer, deoxyribonucleotide triphosphates, ATP, and five phage-T4-coded proteins (2,3). The proteins involved are the products of gene 43 (T4 DNA polymerase), gene 32 (T4 helix-destabilizing protein), and genes 44/62 and 45 (polymerase accessory proteins); these comprise the "five-protein" elongation system.

This five-protein system can be expanded to form a "complete" (seven-protein) *in vitro* replication system, capable of both "leading" and "lagging" strand synthesis at a replication fork at approximately physiological rates, by adding ribonucleoside triphosphates and the products of genes 41 and 61 (the RNA priming proteins). In addition to the cleavage of deoxyribonucleoside triphosphates to

monophosphates (dNTP → dNMP) implicit in polymerization, functioning of the five-protein system requires the concommitant hydrolysis of ATP (to ADP) by gene 44/62 protein, and the seven-protein system involves the hydrolysis of GTP (to GDP) by gene 41 protein as well (5).

In order to further refine this picture we have been studying the molecular details of the interactions of some of the proteins of the T4 replication system with relevant nucleic acid lattices, and with one another. In this paper we will summarize certain aspects of our recent studies on: (i) the structure and nucleic acid binding specificity of gene 32-protein; (ii) the binding to nucleic acid templates and processivity of polymerization of the T4 DNA polymerase; and (iii) the interaction of the polymerase with (as well as the modulation of its processivity by) the polymerase accessory proteins. Details of these studies have been (6-8) or will be (9-11; also Newport and von Hippel, manuscripts in preparation) published elsewhere.

GENE 32-PROTEIN - NUCLEIC ACID INTERACTIONS

Initial Measurements. Alberts and coworkers (12,13) first isolated and purified gene 32-protein, and demonstrated *in vitro* that its central functional feature is its ability to bind preferentially and cooperatively to single-strand DNA sequences. Subsequent studies from this laboratory (6-8) initiated a quantitative thermodynamic examination of this property of the protein. It was shown that gene 32-protein can bind short (ℓ = 2→8 residues) oligonucleotides with an apparent association constant of $\sim 10^5$ M^{-1}, and that this binding is essentially independent of oligonucleotide base composition and sugar type. In addition, measurements of the cooperative binding of gene 32-protein to polynucleotide lattices revealed the binding site size of the protein (n) to be $\sim$ 7 nucleotide residues, the binding constant (K) to be $\sim 10^4$-10^6 M^{-1} in 0.1 M NaCl, and the cooperativity parameter (ω) to be $\sim 10^3$. [See Figure 1 and ref. (14) for definitions and descriptions of these binding parameters.] Even though gene 32-protein is thermodynamically defined as a "melting protein" as a consequence of its preferential binding to single-stranded nucleic acid sequences, it has been shown (13,6) that this protein is "kinetically" blocked from actually melting native DNA. However it *can* melt double-stranded poly[d(A-T)] to equilibrium.

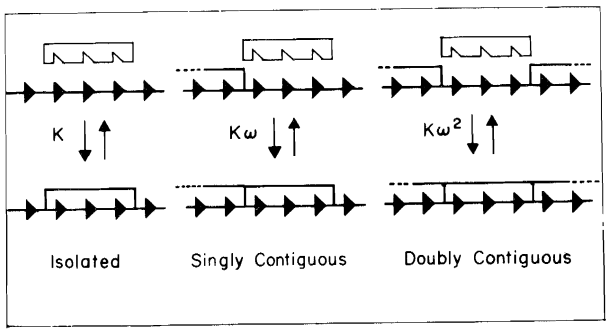

FIGURE 1. Definitions of thermodynamic parameters for the cooperative binding of proteins to single-stranded nucleic lattices. K is the intrinsic equilibrium constant for the isolated binding of a protein to the lattice, and ω is the equilibrium constant for the shifting of a protein molecule from an isolated to a singly contiguous binding site. In cooperative binding to a polynucleotide lattice each protein molecule (except that at the end of the cluster) is bound with a net affinity constant of Kω. n is the site size of the bound protein, in units of nucleotide residues covered (here 3 per protein).

In Vivo Titration of Single-Stranded DNA Sequences and Autogeneous Regulation of Gene 32-Protein Synthesis. Concurrently with the above studies, Gold and coworkers (15,16) and Krisch *et al.* (17) demonstrated *in vivo* that the amount of gene 32-protein synthesized in a T4 infection is proportional to the amount of single-stranded DNA present. Subsequently Lemaire *et al.* (18) showed, using an *in vitro* translation system, that gene 32-protein synthesis is autogenously regulated. They found that after all of the single-stranded DNA sequences present in the solution have been complexed with gene 32-protein, the free concentration of the protein increases to a critical level. At this point the protein binds specifically to its *own* (homologous) mRNA, shutting off further synthesis of gene 32-protein *without* interfering with the synthesis of other T4 proteins. These findings, demonstrating binding specificity to various nucleic acid substrates *in vivo*, seemed incompatible with our earlier demonstration of apparent nonspecificity of binding at the oligonucleotide level, and led us to a further examination of the binding specificity of gene 32-protein.

Oligonucleotide Titrations. To this end we (9) repeated and extended the earlier oligonucleotide titrations of Kelly *et al.* (7), monitoring the quenching of the intrinsic (tryptophan) fluorescence of gene 32-protein on oligonucleotide binding. Working with higher precision techniques we showed unequivocally (within a factor of $\sim$ 3 in K) that the binding (K_{oligo}) of gene 32-protein to oligonucleotides of length 2 to 8 residues is independent of base composition and oligonucleotide length, and that binding to RNA oligonucleotides is at most 3-fold weaker than binding to DNA oligonucleotides of the same length and composition. In addition we showed that K_{oligo} for gene 32-protein binding to oligonucleotides is approximately independent of salt concentration ($\partial \log K_{oligo}/\partial \log [NaCl] \simeq -0.3$; see ref. 19 for general treatment and interpretation of such data). The lack of specificity in oligonucleotide binding seemed indeed to suggest that the demonstrated *physiological* binding specificity must reflect either an enhancement of small differences in affinity as a consequence of protein monomer binding in cooperative clusters (20), or that binding to polynucleotide lattices must involve different gene 32-protein-nucleic acid interactions (or, of course, both of the above).

Polynucleotide Titrations. To investigate this aspect we undertook a comprehensive series of titrations of various homopolynucleotides with gene 32-protein, monitoring either the quenching of intrinsic protein fluorescence or the UV absorbance change of the nucleic acid due to the base unstacking which accompany binding (6). Titrations were carried out as a function of salt concentration, and typical results (here with poly rA) at different salt concentrations are presented in Figure 2. Clearly binding is cooperative and salt concentration dependent.

The *free* ligand (protein) concentration at the midpoint of titrations such as those of Figure 2 is equal to $(K\omega)^{-1}$ for that polynucleotide at that salt concentration. Values of $K\omega$ measured in this way on a variety of polynucleotides are presented as a plot of log $K\omega$ *versus* log [NaCl] in Figure 3. Clearly the effective binding constant ($K\omega$) for gene 32-protein (at constant salt concentration) is dependent on both the base composition and sugar type of the polynucleotide. Quantitative analysis of data such as that of Figure 3 shows that $K\omega$ for a random copolymer containing sveral types of bases (or for a natural DNA) is approximately equal to the compositionally-weighted sum of the $K\omega$ values for the individual homopolynucleotides, indicating that the specificity depends on

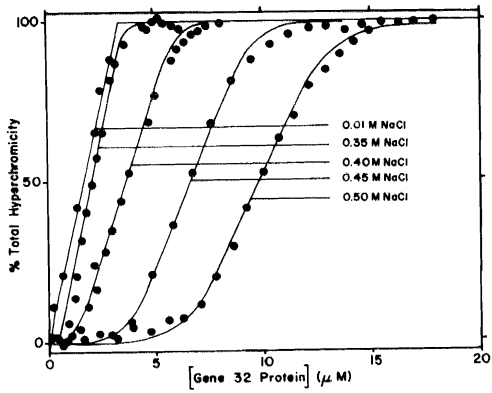

FIGURE 2. Salt concentration dependence of the binding of gene 32-protein to poly rA in 10 mM Hepes, 0.1 mM EDTA, pH 7.7, plus added NaCl as indicated. The solid lines represent calculated theoretical curves, using n = 7 nucleotide residues per protein monomer and $\omega = 2 \times 10^3$. K was determined (using this ω) from the experimental value of Kω.

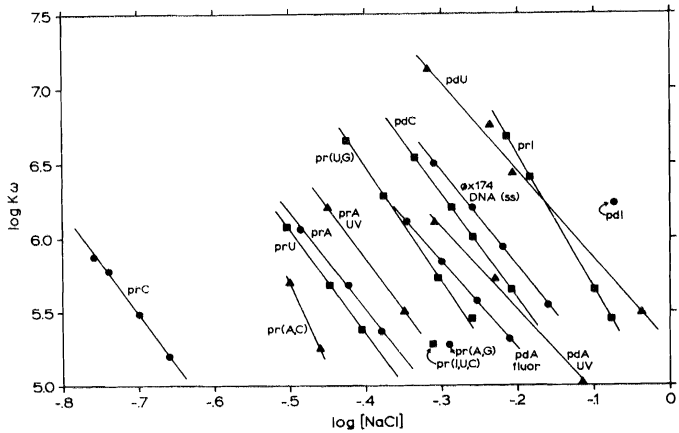

FIGURE 3. Plot of log Kω *versus* log [NaCl] for the cooperative binding of gene 32-protein to various polynucleotides; buffer conditions as in Figure 1. Lines marked "fluor" or "UV" represent duplicate sets of measurements carried out by quenching of intrinsic protein fluorescence, or changes in UV absorbance, respectively. The rest of the data all represent fluorescence quenching determinations.

differential binding of individual bases along the chain. In addition, $K\omega$ for a given polydeoxyribonucleotide is always greater than that for the homologous polyribonucleotide (Figure 3).

Both Figures 2 and 3 also show that, unlike oligonucleotide binding, cooperative polynucleotide binding is *very* salt concentration dependent ($\partial \log K\omega / \partial \log [NaCl] \simeq -6 \pm 1$). In addition, binding measurements performed with mono- and divalent salts carrying different anions suggest that approximately two-thirds of the above salt dependence reflects displacement of protein-bound *anions* as a consequence of nucleic acid interactions (9). Studies involving the fit of theoretical binding equations to experimental titration data, and measurements at very low binding densities (ν), show that all the salt dependence of $K\omega$ is in K, with ω remaining constant (salt-independent) at $\sim 10^3$.

These results indicate that gene 32-protein can bind to nucleic acids in two very different ways, probably representing two distinct protein conformations. We term these binding conformations the oligonucleotide and the polynucleotide binding modes, respectively. Figure 4 illustrates and summarizes some of the inferred molecular

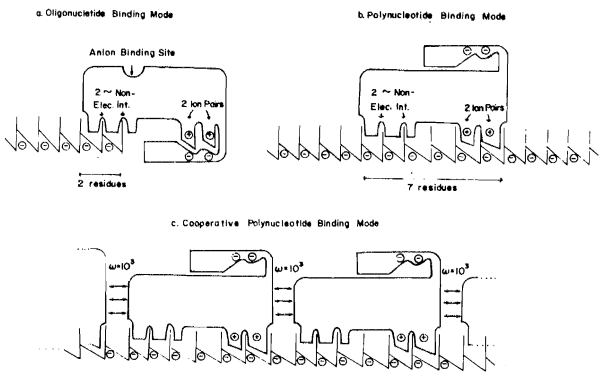

FIGURE 4. Schematic representation of the binding modes (conformations) of gene 32-protein to nucleic acid lattices. Note that binding in the polynucleotide mode involves unmasking of a largely electrostatically-binding sub-site, with the concommitant removal of the block to statistical ("shuffling") binding seen in the oligonucleotide mode, and also the disruption of the "anion binding site". Cooperative binding in the polynucleotide mode involves lattice distortion and protein-protein interaction (see text and, for further details, ref. 9).

features of these two binding modes in schematic form (for further details see ref. 9).

Binding Properties of GP32*I and GP32*III. Hosoda and coworkers (21,22) have shown that brief treatment of native gene 32-protein with proteolytic enzymes results in cleavage of an ∿ 60 residue peptide from the C-terminus, and an ∿ 20 residue peptide from the N-terminus of the original protein. These cleavage products can be isolated from such a digest; the resulting proteins have been termined GP32*I (C-terminal peptide removed), GP32*II (N-terminal peptide removed) and GP32*III (both peptides removed). These cleaved proteins show changes in both apparent binding affinity and cooperativity of binding to nucleic acid lattices, relative to undigested gene 32-protein (21,22,23). In order to further our molecular understanding of the interactions responsible for gene 32-protein complex formation with nucleic acids, we have measured the thermodynamic parameters characterizing the binding of GP32*I and GP32*III to nucleic acid lattices (11).

GP32*I binds to short oligonucleotides with approximately the same K as gene 32-protein (at ∿ 0.1 M NaCl), but with a somewhat increased salt dependence. Binding of GP32*I to the various polynucleotides of Figure 3 follows the same order of binding affinity, and shows the same overall salt dependence as gene 32-protein; n and ω are unchanged from the gene 32-protein values, and K is increased ∿ 2 to 4-fold. These results are all consistent with a partial proteolytic removal of the negatively charged "shuffling block" in gene 32-protein (see Figure 4). Thus thermodynamically, GP32*I is very similar to the native protein; the small change in K is not sufficient of itself to explain why this derivative *can* melt native DNA to equilibrium (11,24). Clearly kinetic explanations of this difference must be sought.

GP32*III binds nucleic acids very differently. This derivative shows no trace of binding cooperativity ($\omega = 1$), and a much lower net binding affinity. The dependence of binding affinity on salt concentrations is also decreased. On the other hand GP32*III deforms the single-stranded nucleic acid lattice on binding (as evidenced by CD and UV absorbance studies) to the same extent as gene 32-protein and GP32*I. This suggests (see refs. 6,11,20) that lattice deformation alone cannot be responsible for the cooperativity of binding of gene 32-protein; there must be substantial involvement of direct protein-protein interactions (Figure 4).

Competitive Cooperative "Non-Specific" Binding as a Genome Regulatory Mechanism. As described above, gene 32-protein binding in the polynucleotide mode does indeed exhibit some differences in net affinity ($K\omega$) for different nucleic acid lattices at constant salt concentration (Figure 3), depending on base composition and sugar type. In addition, at higher salt concentrations [such as those characterizing the *in vitro* translation system of Lemaire *et al.* (18)] we see a finite "lag phase" in the cooperative titration curve (Figure 2), prior to the onset of lattice binding. Thus the *free* protein concentration must reach a certain critical level before binding begins. These properties provide a quantitative explanation of the preferred order of binding of this protein to different nucleic acid lattices, and thus also of the mechanism by which the protein autoregulates its own synthesis. The protein "binding cycle" underlying this control system is represented schematically in Figure 5.

How does such a system work? Because binding is cooperative (Figure 2) and competitive for protein (i.e., generally there is an excess of lattice binding sites available), control depends on which of these sets competing binding sites are saturated first. Figure 3 suggests that for lattices of comparable size and base composition, binding to DNA sequences should precede RNA binding. [In addition (data not shown in Figure 3) binding of gene

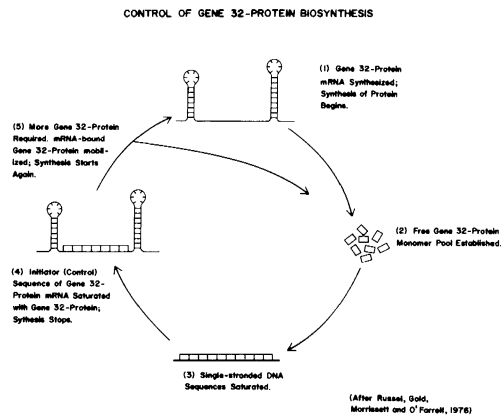

FIGURE 5. Schematic representation of the sequence of events involved in the autogenous regulation of gene 32-protein synthesis, as elucidated by Gold and coworkers (16,18). The various competitive binding equilibria involved are indicated (see text).

32-protein to poly dT is anomalously tight, further favoring the initial saturation of DNA lattices.] Cooperative binding can amplify small differences in intrinsic binding affinity (K) or cooperativity (ω); differences in binding lattice *size* can also modulate the binding competition, again because binding is cooperative (20). Thus for two competing lattices of identical base compositon and sugar type, the larger lattice will saturate first. At equal levels of lattice saturation the net binding free energy of the system is rendered more unfavorable (by RTlnω relative to an "unbroken" cluster) whenever a contiguously bound protein cluster is terminated and a new cluster is initiated at an isolated binding site.

An example of the use of this principle might be the original suggestion of Russel *et al.* (16) that gene 32 mRNA may be characterized by a critical unstructured control sequence (e.g., flanked by two stable hairpins) which is *longer* than the analogous sequences on other T4 mRNAs. This could account very simply for the binding preference of the protein for its own message. Studies of the base sequences of putative mRNA control regions (e.g., see Krisch *et al.*, 25) should soon permit assessment of the applicability of such notions to real systems.

Clearly such models can, in principle, account for sequential regulation processes such as that outlined in Figure 5. Elsewhere (10) we have computer-simulated some possible models, and have shown that these notions can be made quantitatively consistent with the gene 32-protein autoregulation data. In addition, these approaches can be used to construct a variety of genome regulatory systems based on the general principle that specificity of binding to competing nucleic acid sequences can arise through the *cooperative* binding of proteins that, binding individually at isolated sites, show only marginal sequence-dependent binding specificity.

INTERACTIONS OF T4 DNA POLYMERASE WITH PRIMER AND TEMPLATE

In order to function in either a polymerization or a 3'→5' exonuclease (editing) mode, DNA polymerase must bind at the 3'-(primer) terminus of a primer-template complex (see ref. 1 for many functional details and binding models). Thus elements of both single- and double-stranded DNA binding could be involved in this interaction. We have carried out a number of experiments to learn more about the molecular details of the formation of this central component of the T4 DNA replication complex.

Binding to Single-Stranded DNA. Sedimentation velocity and DNA-cellulose column techniques (26,27) have been used to measure the binding of T4 DNA polymerase to single-stranded, circular ΦX174 DNA. The results show that polymerase binds fairly tightly to single-stranded DNA at low ionic strength ($K \simeq 10^5\ M^{-1}$ at 0.05 M NaCl), and that binding decreases markedly with increasing salt concentration.

We find, as has been shown with many other nucleic acid-protein interaction systems (see above and ref. 19), that a log-log plot of the apparent binding constant versus salt concentration is linear ($\partial \log K / \partial \log [NaCl] \simeq -5$; data not shown), and that, unlike for gene 32-protein, binding shows no anion dependence. These data can be interpreted quantitatively (19) to suggest that the interaction of T4 polymerase with single-stranded DNA involves 7 to 8 charge-charge (DNA phosphate with basic protein side-chain) contacts.

These single-stranded DNA-polymerase binding data can also be extrapolated to estimate a binding constant for this interaction at physiological salt concentrations. Based on salt effects with DNA-repressor interactions, it has been shown (28) that the effective cation concentration (for such interactions within an *E. coli* cell is at least 0.20 M (in Na^+ equivalents). In 0.2 M NaCl we estimate the polymerase-single-stranded DNA binding constant to be $\leq 10^2\ M^{-1}$. This low value of K suggests that few, if any, polymerase molecules bind in isolation to (e.g.) single-stranded parts of DNA replication forks *in vivo*.

Salt-Dependence of Polymerase Function. The above data suggest that the free energy of binding of T4 DNA polymerase to single-stranded DNA has a major electrostatic component involving $\sim$ 8 charge-charge interactions. We may now ask whether the salt dependence of the binding of the T4 polymerase to other nucleic acid structures (including the primer-template complex) is similar. To approach this question we determined the effect of varying salt concentration on the rate of primer extension by T4 polymerase at an oligo dT - poly dA primer-template junction. In addition, we examined the salt dependence of the rate of digestion of (linear) single-stranded DNA by the polymerase acting in the 3'→5' exonuclease mode. In both cases the effect of NaCl on the rate of catalysis suggested the involvement of $\sim$ 7-8 charge-charge interactions per binding event. These results suggest that the *same* types of electrostatic interactions are involved in polymerase binding to: (i) "interior" single-stranded sites on

circular DNA; (ii) the primer-template junction in the polymerization mode; and (iii) the 3' termini of single-stranded DNA in the exonuclease mode.

Mapping the Interactions of T4 DNA Polymerase with the Primer-Template Complex. 3'→5' exonuclease activity of T4 polymerase on oligo dT - poly dA templates has been utilized to determine which particular residues of the primer make important contacts with the polymerase. Using 5'-^{32}P-labelled template-bound primers of initially unique length, we found (in the absence of deoxynucleoside triphosphates) that the polymerase digests these parameters at a uniform rate until the primer length is reduced to 10 nucleotide residues. At that point digestion is slowed, indicating that residue 11 (from the 3' end) makes an important contact with the polymerase. Further analysis involving application of the same technique to primers of shorter initial length indicates that T4 DNA polymerase interacts strongly with nucleotide residues 1, 2, 3, 4 and 8, 9, 10, and 11, counting from the 3'-OH primer terminus. Residues 5, 6 and 7 do not seem to interact significantly, based on this assay. (For details see ref. 29, and manuscript in preparation.) We note that the eight primer-polymerase contacts deduced by this technique are numerically consistent with the 7-8 charge-charge interactions elucidated from salt-dependence data for the interaction between polymerase and various nucleic acid "substrates".

In Figure 6 we present a model of the primer-template-polymerase complex which is consistent with all evidence presented above. Having no information to the contrary, the primer-template complex is shown in the usual double-helical form. We note that this permits contacts with residues 1, 2, 3, 4, 8, 9, 10 and 11 along one "side" of the double-helix; residues 5, 6 and 7, which apparently do not interact with the polymerase, are located on the opposite side of the DNA structure. These results suggest that polymerase may contain two (presumably positively charged) primer-binding sub-sites; one reacting with primer residues 1 through 4, and the other with residues 8 through 11.

Processivity of T4 Polymerase-Catalyzed DNA Synthesis. Several groups have shown that T4 DNA polymerase is capable of processive synthesis under some conditions (30,31). To investigate the processivity of this polymerase under a variety of conditions we have used the assay scheme outlined in Figure 7 (a similar procedure has been developed independently by McClure and Chow, 32). In analyzing the data obtained from this system we assume that processive

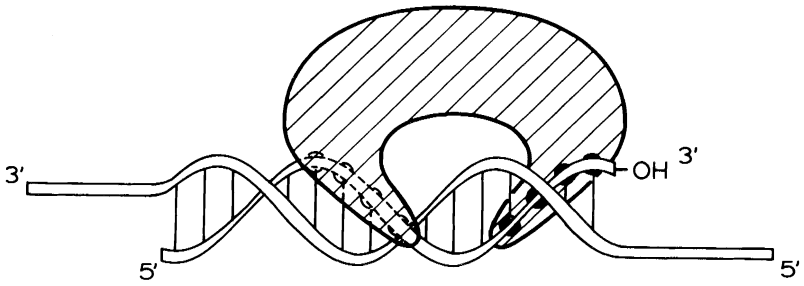

FIGURE 6. Proposed model for T4 DNA polymerase binding to the primer-template complex. Contacts between polymerase and primer occur at residues 1-4 (●) and residues 8-11 (○), respectively counting from the 3'-OH end of the primer (see text).

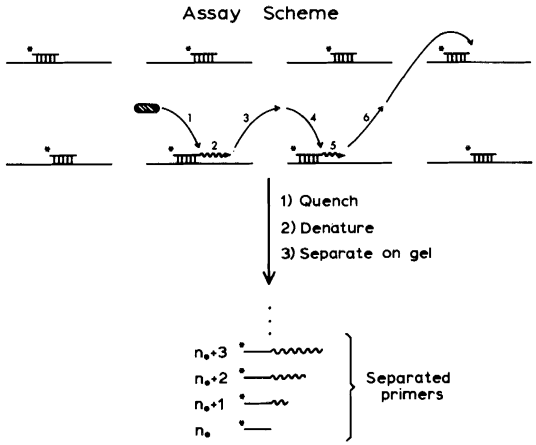

FIGURE 7. Model of the assay system for determining polymerase processivity (see text for conditions). The elongated primers are separated by polyacrylamide gel electrophoresis at the end of the experiment, and the size distributions are determined by quantitative autoradiography (see Figure 9).

synthesis involves a minimum of three steps: (i) binding of a (template complementary) nucleoside triphosphate unit to the polymerase-primer-template complex; (ii) covalent linking of this unit (as the nucleoside monophosphate) to the 3'-primer terminus; and (iii) translocation of the polymerase one nucleotide residue forward along the template to achieve correct alignment at the new primer terminus. The "processivity" of this synthesis depends on the number of such three-step cycles the polymerase can complete before it dissociates from the primer-template complex.

Assuming dissociation is a first-order process, we can assign a probability to each three-step cycle, where P is the probability that a particular number of cycles will be completed prior to a dissociation event. Thus for a given set of primers which have been associated with T4 DNA polymerase only once during the experiment (this is assured by maintaining a high primer to polymerase ratio), the fraction of primers that have had exactly *n* nucleotides added to the 3'-terminus is:

$$\frac{n_x}{n_{total}} = P^{(n-1)}(1 - P) \qquad (1)$$

or

$$\log\left(\frac{n_x}{n_{total}}\right) = (n - 1)\log P + \log(1 - P) \qquad (2)$$

where *n* is the number of nucleotides added (i.e., the number of cycles completed) prior to dissociation, and n_x/n_{total} is the fraction of elongated primers with exactly *n* additions.

For example, if P = 0.9, the probability of adding at least one nucleotide to the 3'-terminus of a particular primer is 0.9, that of adding at least two nucleotides is $P^2 = 0.81$, that of adding at least three nucleotides is $P^3 = 0.729$, etc. The probability of adding *exactly* one nucleotide is $P^1 - P^2 = 0.09$; that of adding *exactly* two nucleotides is $P^2 - P^3 = 0.081$, and so on. A plot of $\log (n_x/n_{total})$ as a function of $(n - 1)$ should yield a straight line with a slope = log P and y-intercept = log (1 - P), from which P can be determined. Such a plot is shown in Figure 8, utilizing data obtained from the analysis of the bands in Lane I of Figure 9. We find under these conditions (a 30 second reaction in 0.095 M NaCl) that synthesis is processive ($P = 0.84 \pm 0.01$); this value of P means that 50% of the elongated primers will have undergone at least four cycles of nucleotide addition prior to polymerase dissociation. The linearity of this plot confirms our initial assumption that during processive synthesis

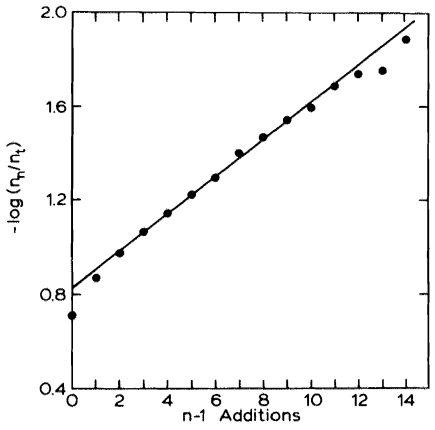

FIGURE 8. Data from Lane I of Figure 9, plotted according to eq. (2). A value of P = 0.833 is obtained from the slope, and of P = 0.848 from the y-intercept (see text).

T4 DNA polymerase dissociates from the primer-template complex in a first-order fashion.

Figure 9 shows a series of experiments in which the processivity of T4 DNA polymerase was determined as a function of NaCl. Clearly, as salt concentration is increased the processivity of the polymerase decreases. This finding is consistent with the results presented above, indicating a decreasing binding constant for polymerase to nucleic acid substrates (including the primer-template complex) at increasing salt concentrations. These data strongly suggest that DNA synthesis *by polymerase alone* under salt conditions comparable to the *in vitro* situation ($\sim$ 0.2 M NaCl) is totally dispersive.

PROPERTIES OF THE FIVE PROTEIN T4 DNA REPLICATION SYSTEM

Addition of T4 Polymerase Accessory Proteins Makes DNA Synthesis Processive at *in vitro* Salt Concentrations. The results described above show that T4 DNA polymerase, acting by itself under roughly *in vitro* salt conditions, binds only weakly to single-stranded DNA (or to primer-template complexes), and synthesizes new DNA in a totally dispersive manner [P = 0; eq. (1) and (2), above]. Using the assay system and conditions shown in Figures 7 and 9, we find that

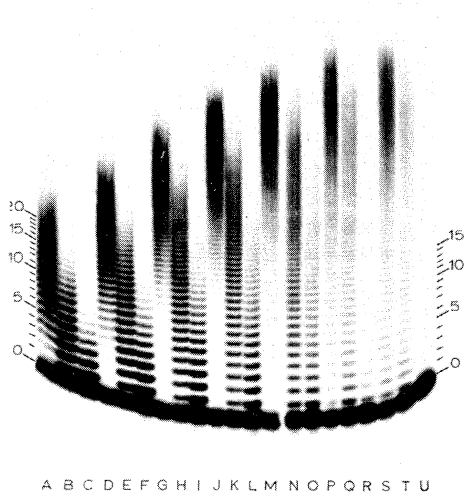

FIGURE 9. Autoradiogram of experiment to determine the processivity of T4 DNA polymerase (see text). Reaction mix contained 10 mM Tris (pH 7.8), 0.5 mM DTT, 200 μg/ml BSA, 0.1 mM EDTA, 2.5 mM $MgCl_2$, 3.1 x 10^{-7} $(dT)_{16}$ (molarity of nucleotide residues) and 1.16 x 10^{-8} M T4 DNA polymerase. The reactions were incubated at various concentations of added NaCl for the times indicated. Time points represent 30 (C, F, I, L, O, R, U), 120 (B, E, H, K, N, Q, T) and 300 (A, D, G, J, M, P, S) seconds. NaCl concentrations were 135 mM (A, B, C); 115 mM (D, E, F); 95 mM (G, H, I), 75 mM (J, K, L); 55 mM (M, N, O); 35 mM (P, Q, R) and 25 mM (S, T, U).

the addition of gene 32, 44/62 and 45 proteins, plus ATP, results in a tremendous increase in the rate and processivity of synthesis of new DNA (poly dT) under *in vivo* salt conditions. The omission of any of these four accessory proteins, or of ATP, from the reaction mix reduces the processivity back to approximately the level characteristic of T4 DNA polymerase alone (Figure 9). These results are in good accord with the earlier observations of Alberts and coworkers (2,4) and Nossal and Peterlin (33) that the "five protein system" is capable of efficient leading strand synthesis on nicked double-stranded DNA.

ATP Hydrolysis is Required for Assembly of the Five Protein Replication Complex, Rather Than for Elongation. Piperno and Alberts (34) have demonstrated that ATP *hydrolysis* is required to demonstrate the stimulation of DNA synthesis by gene 44/62 and 45 proteins in the five protein system; ATP *binding* alone does not suffice. Furthermore, these workers showed that less than one ATP is hydrolyzed per ten deoxyribonucleotides incorporated into DNA. These results suggest that ATP hydrolysis may control *assembly* of the replication complex, rather than being required as an energy source for polymerase translocation in chain elongation. This model predicts that once the replication complex is assembled, the length of the DNA chain synthesized should be independent of ATP hydrolysis. Furthermore, the *number* of replication complexes formed (or the number of chain initiation events per unit time) *should* be a function of ATP concentration. This hypothesis was tested in an experiment of the sort depicted in Figure 9, where DNA synthesis by the five protein system was monitored as a function of ATP concentration.

The results obtained (data not presented here; for details see ref. 29 and Newport and von Hippel, in preparation) show that as the ATP concentration is decreased, fewer primers are elongated. However, once initiated the *length* of the extended primers is *independent* of ATP concentration. This strongly suggests that ATP hydrolysis is required only for the establishment (assembly) of a functional five protein replication complex, and not for elongation *per se*.

Further analysis of such experiments also shows that the number of primers elongated increases with time. This suggests that a replication complex, once assembled, is not infinitely stable; a maximum complex life-time of $\sim$ 45 seconds is obtained under our experimental conditions. This life-time is essentially independent of the rate of primer elongation. These results suggest that ATP hydrolysis may serve as a timing mechanism to control the temporal stability of the functional (five protein) replication complex.

Stoichiometry of the Five Protein Replication Complex. Gene 44/62 proteins are isolated as a tight complex composed of approximately four gene 44- and two gene 62-coded polypeptide chains; the total molecular weight of the complex is $\sim$ 180,000 daltons (35,36). In experiments in which the rate of *initiation* of DNA synthesis (primer extension) was studied as a function of gene 44/62 protein concentration in the presence of excess gene 32 and 45 proteins, we find that maximal initiation activity occurs at a level of one gene

44/62 protein complex per primer. In similar experiments in which the concentration of gene 45 protein was varied, we were unable to achieve saturation of initiation activity, suggesting that gene 45 protein binding is much weaker.

Model of the Assembly of the T4-Coded Five Protein DNA Replication Complex. These results and others (4,5,29,33) are consistent with a preliminary working model for the five protein system on an oligo dT - poly dA primer-template which is schematized in Figure 10. We suggest that in the presence of gene 32- and 45-proteins, and ATP, one gene 44/62 protein moiety binds tightly to the primer-template complex. T4 DNA polymerase (gene 43 protein) then binds onto this complex, interacting mostly with the primer strand and resulting in a complex capable of rapid and high processive "leading strand" DNA synthesis at *in vivo* salt concentations. Additional experimental support for some aspects of this model will be presented elsewhere (Newport and von Hippel, in preparation).

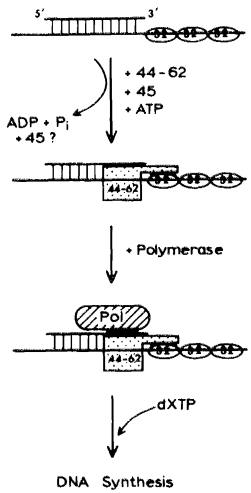

FIGURE 10. Proposed assembly scheme for the five protein T4 DNA replication complex. Dark areas represent strong interactions. For further details see text.

ACKNOWLEDGEMENTS

We are very grateful to Dr. Bruce Alberts and his colleagues for their generous gifts of gene 44/62 and gene 45 proteins and for advice, to Dr. Larry Gold for many discussions and for mutant T4 strains, Dr. Junko Hosoda for advice on the preparation of GP32*I and GP*III, and to Dr. Otto Berg for helpful discussions of the mathematics of processive and dispersive chain elongation.

REFERENCES

1. Kornberg, A. (1980). "DNA Replication" W. H. Freeman, San Francisco.
2. Alberts, B., Barry, J., Bittner, M., Davies, M., Hama-Inaba, H., Liu, C.-C., Mace, D., Moran, L., Morris, C. F., Piperno, J., and Sinha, K. K. (1977). In "Nucleic Acid-Protein Recognition" (J. Vogel, ed.), pp. 31-63. Academic Press, New York.
3. Kornberg, A. (1979). Cold Spring Harbor Symp. Quant. Biol. 43, 1-9.
4. Liu, C. C., Burke, R. L., Hibner, U., Barry J., and Alberts B. (1979). Cold Spring Harbor Symp. Quant. Biol. 43, 469-487.
5. Alberts, B., Barry, J., Bedinger, P., Burke, R. L., Hibner, O., Liu, C.-C. and Sheridan, R. (1980). This volume.
6. Jensen, D. E., Kelly, R. C., and von Hippel, P. H. (1976). J. Biol. Chem. 251, 7215-7228.
7. Kelly, R. C., and von Hippel, P. H. (1976). J. Biol. Chem. 251, 7229-7239.
8. Kelly, R. C., Jensen, D. E., and von Hippel, P. H. (1976). J. Biol. Chem. 251, 7240-7250.
9. Kowalczykowski, S. C., Lonberg, N., Newport, J. W., and von Hippel, P. H. (1970). J. Mol. Biol. (submitted).
10. Newport, J. W., Lonberg, N., Kowalczykowski, S. C., and von Hippel, P. H. (1980). J. Mol. Biol. (submitted).
11. Lonberg, N., Kowalczykowski, S. C., Paul, L. S., and von Hippel, P. H. (1980). J. Mol. Biol. (submitted).
12. Alberts, B. M., Amodio, F. J., Jenkins, M., Gutmann, E. D., and Ferris, F. L. (1968). Cold Spring Harbor Symp. Quant. Biol. 33, 289-305.
13. Alberts, B. M., and Frey, L. (1970). Nature 227, 1313-1317.
14. McGhee, J.D., and von Hippel, P. (1974). J. Biol. Chem. 86, 469-489.
15. Gold, L., O'Farrell, P. Z., and Russel, M. (1976). J. Biol. Chem. 251, 7251-7262.

16. Russel, M., Gold, L., Morrissett, H., and O'Farrell, P. Z. (1976). J. Biol. Chem. 7263-7270.
17. Krisch, H. M., Bolle, A., and Epstein, R. H. (1974). J. Mol. Biol. 88, 89-104.
18. Lemaire, G., Gold, L., and Yarus, M. (1978). J. Mol. Biol. 126, 73-90.
19. Record, T. M., Jr., Lohman, T. M., and de Haseth, P. (1976). J. Mol. Biol. 107, 145-158.
20. von Hippel, P. H., Jensen, D. E., Kelly, R. C., and McGhee, J. D. (1977). In "Nucleic Acid-Protein Recognition" (H. J. Vogel, ed.), pp. 65-89. Academic Press, New York.
21. Hosoda, J., Takacs, B., and Brack, C. (1974). FEBS Letters 47, 338-342.
22. Hosoda, J., and Moise, H. (1978). J. Biol. Chem. 253, 2464-2470.
23. Williams, K. R, and Konigsberg, W. (1978). J. Biol. Chem. 253, 2464-2470.
24. Kowalczykowski, S. C., Lonberg, N., Newport, J. W., Paul, L. S., and von Hippel, P. H. (1980). Biophys J. (in press).
25. Krisch, H. M., Duvoisin, R. M., Allef, B., and Epstein, R. H. (1980). Proc. Natl. Acad. Sci. USA (in press).
26. Draper, D. E., and von Hippel, P. H. (1979). Biochemistry 18, 753-760.
27. de Haseth, P. L., Gross, C. A., Burgess, R. R., and Record, M. T. Jr. (1977). Biochemistry 16, 47777-4782.
28. Kao-Huang, Y., Revzin, A., Butler, A. P., O'Conner, P., Noble, D. W., and von Hippel, P. H. (1977). Proc. Nat. Acad. Sci. USA 74, 4228-4232.
29. Newport, J. W. (1980). Ph.D. Thesis, University of Oregon, Eugene, Oregon.
30. Alberts, B.M., Morris, C.F., Mace, D., Sinha, N., Bittner, M., and Moran, L. (1975) in "DNA Synthesis and Its Regulation", ICN-UCLA Symposia on Molecular and Cellular Biology (Goulian, M., Hanawalt, P., and Fox, C. F. eds.), Vol. 3, pp. 2441-269, W. H. Benjamin, Menlo Park, California.
31. Das, S. K., and Fujimura, R. K. (1979). J. Biol. Chem. 254, 1227-1232.
32. McClure, W. R. and Chow, Y. (1980). "Methods in Enzymology" (Grossman, L., and Moldave, K., eds.), in press.
33. Nossal, N. G., and Peterlin, B. M. (1979). J. Biol. Chem. 254, 6032-6037.
34. Piperno, J. R., and Alberts, B. M. (1978). J. Biol. Chem. 253, 5174-5179.

35. Barry, J. and Alberts, B. M. (1972). Proc. Nat. Acad. Sci. USA 69, 2717-2721.
36. Morris, C. F., Hama-Inaba, H., Mace, D., Sinha, N. K., and Alberts, B. M. (1979). J. Biol. Chem. 254, 6787-6796.