Gene V Protein Dimerization and Cooperativity of Binding to Poly(dA)[†]

Thomas C. Terwilliger*

Mail Stop M888, Structural Biology Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Received May 1, 1996; Revised Manuscript Received October 8, 1996[⊗]

ABSTRACT: Gene V protein of bacteriophage f1 is a dimeric protein that binds cooperatively to single-stranded nucleic acids. In order to determine whether a monomer−dimer equilibrium has an appreciable effect upon the thermodynamics of gene V protein binding to nucleic acids, the dissociation constant for the protein dimer was investigated using size-exclusion chromatography. At concentrations ranging from 5×10^{-10} to 1.2×10^{-5} M, the Stokes radius of the protein was that expected of the dimer of the gene V protein. The Stokes radius of the protein was also independent of salt concentration from 0.2 to 1.0 M NaCl in a buffer containing 10 mM Tris-HCl, pH 7.4, and 1 mM EDTA. The binding of the dimeric gene V protein to poly(dA) was studied using a simplified lattice model for protein−protein interactions adapted for use with a dimeric protein that binds simultaneously to two strands of nucleic acid. Interpretation of the salt dependence, $C = [\mathrm{d}\log(K_{\mathrm{int}}\omega)]/[\mathrm{d}\log([\mathrm{NaCl}])]$, of binding of such a dimeric protein to nucleic acid using the theory of Record et al. (Record, M. T., et al. (1976) *J. Mol. Biol. 107*, 145−158) indicates that C is a function of the numbers of cations and anions released from protein and nucleic acid upon binding of the dimer, not of the monomer. Cooperativity of gene V protein binding to poly(dA) was studied with titration experiments that are sensitive to the degree of cooperativity of binding. The cooperativity factor ω, defined as the ratio of the binding constant for a site adjacent to a previously bound dimer to that for an isolated site, was found to be relatively insensitive to salt, with a value in the range of 2000−7000 for binding to poly(dA) at 3 °C and at 23 °C. This high cooperativity factor supports the suggestion that protein−protein contacts play a major role in the formation of the superhelical gene V protein−single-stranded nucleic acid complex.

The gene V protein of bacteriophage f1 (fd, M13) is a small dimeric protein that binds single-stranded nucleic acids with relatively little sequence specificity (Baas, 1985; Kowalczykowski et al., 1981a). This protein and its interactions with nucleic acids have been intensively studied by many laboratories, and gene V protein, along with the *Escherichia coli* SSB protein and bacteriophage T4 gene 32 protein, serves as a general model for nonspecific protein−single-stranded nucleic acid interactions. *In vivo*, gene V protein is required for replication of bacteriophage f1 DNA (Alberts et al., 1972; Baas, 1985; Chase & Williams, 1986; Fulford & Model, 1988; Kowalczykowski et al., 1981a; Pratt & Ehrdahl, 1968; Salstrom & Pratt, 1971). The protein binds cooperatively to the single-stranded intermediate in f1 DNA replication, preventing conversion of single-stranded DNA to replicative form DNA, and instead allowing packaging of the single-stranded DNA into new virus particles.

The DNA-binding and structural properties of gene V protein have been examined in considerable detail. The binding affinities and cooperativity of binding of gene V protein to many polynucleotides and oligonucleotides have been measured, and the binding properties of a large number of mutants with amino acid substitutions have been studied as well (Alma et al., 1983; Bulsink et al., 1985, 1988; Kansy et al., 1986; Stassen et al., 1992a,b). The structure of the complex formed by gene V protein bound to polynucleotides has been examined by electron microscopic techniques (Gray, 1989), and the structure of the dimeric form of the protein has been determined by X-ray crystallographic and multidimensional NMR techniques (Folkers et al., 1994; Skinner et al., 1994). Many of the protein side chains involved in binding to nucleic acids have been identified, so that the protein−single-stranded DNA interaction is beginning to be understood in detail for this protein (de Jong et al., 1989; Folkers et al., 1991; King & Coleman, 1987; Stassen et al., 1992a; van Duynhoven, 1992).

Oligomeric State of Gene V Protein. Despite the many biophysical studies carried out on the gene V protein, two of the important properties of the protein, the oligomeric state of the protein at micromolar concentrations and the cooperativity of binding to single-stranded nucleic acids, are still not precisely defined. The gene V protein is known to exist predominantly as a dimer when it is present at high concentrations ($>10^{-5}$ M) in solutions containing low to moderate concentrations of salt (0.15 M NaCl or lower) (Cavalieri et al., 1976; Oey & Knippers, 1972; Porschke & Rauh, 1983; Pretorius et al., 1975), and the crystal and NMR structures of the protein clearly show a dimer with extensive interactions between subunits (Skinner et al., 1994). The concentration of the gene V protein *in vivo* inside an infected *E. coli* cell is high enough (10^{-4} M) that the functional form of the protein is certainly the dimer (Alberts et al., 1972). Many *in vitro* experiments on the binding of gene V protein to nucleic acids, on the other hand, are carried out at concentrations of $(1-10) \times 10^{-6}$ M protein, and the interpretation of these experiments depends considerably on whether the protein is monomeric or dimeric under these conditions (Liang & Terwilliger, 1991; Stassen et al., 1992a; Zabin & Terwilliger, 1991). The dissociation constant of the dimer has been reported to be about 10^6 M^{-1} in 0.1 M

[†] Supported by the NIH (GM38714) and the LDRD program of Los Alamos National Laboratory (94214).

* Tel: (505) 667-0072.

[⊗] Abstract published in *Advance ACS Abstracts,* December 1, 1996.

S0006-2960(96)01050-1 CCC: $12.00

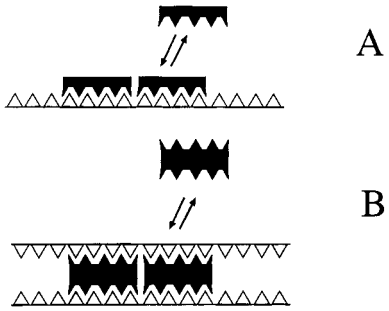

FIGURE 1: Independent and simultaneous binding of protein monomers to nucleic acid. (A) Independent binding model. Each subunit of protein in solution can bind separately from any others to a strand of nucleic acid. (B) Simultaneous binding model. Each subunit of protein is part of a dimer, and the two subunits in a dimer bind simultaneously to two separate strands of nucleic acid.

NaCl using measurements based on equilibrium ultracentrifugation, but this estimate is uncertain as the average concentration used in each of these centrifugation experiments was always well above 10^{-6} M, the concentration where the dimer would be 50% dissociated (Porschke et al., 1983). Other equilibrium ultracentrifugation experiments carried out at protein concentrations of 7×10^{-6} M and above showed little dependence of average molecular mass with protein concentration or salt concentration from 0 to 0.2 M KCl (Cavalieri et al., 1976). Sucrose gradient sedimentation velocity experiments gave a somewhat different result, indicating that the protein was monomeric in 0.2 M KCl in the presence of sucrose (Oey et al., 1972). High concentrations of NaCl (0.68 M) were found in another equilibrium ultracentrifugation study to dissociate the gene V protein dimer (Pretorius et al., 1975).

As the knowledge of this dissociation constant is important for interpretation of *in vitro* binding experiments on the gene V protein, we felt that it was necessary to obtain more experimental information on it. Reported analytical ultracentrifugation experiments have been limited by sensitivity of detection to gene V protein concentrations above about 10^{-6} M of protein, so here we have used size-exclusion chromatography with unlabeled and radioactively labeled gene V protein to obtain a lower bound estimate for the value of the K_D of the gene V protein dimer.

Cooperativity of Gene V Protein Binding to Single-Stranded Nucleic Acids. The cooperativity of binding of gene V protein and related proteins to single-stranded nucleic acids is thought to be due to protein−protein contacts made as the protein binds to nucleic acids. In the case of the T4 gene 32 protein, for example, some of the cooperative interactions appear to reside within the N-terminal domain, as cooperativity is lost when this part of the protein is removed (Lonberg et al., 1981), and is affected by amino acid substitution in this domain (Villemain & Giedroc, 1993). In the case of gene V protein, specific protein−protein contacts are also involved in cooperativity. The residues that affect cooperativity are not identified with certainty, but one residue that has been implicated in protein−protein contacts is tyrosine 41 (Folkers et al., 1991; King & Coleman, 1988). Substitution of the tyrosine at position 41 with histidine, for example, decreases the cooperativity of binding to poly(dA) by at least 5-fold (Stassen et al., 1992a). The closely related protein from phage IKe shows cooperative binding, but there is little or no cooperativity between the two different gene V proteins, indicating that specific interactions are responsible for cooperativity (de Jong et al., 1987).

A useful model for describing the cooperative binding of proteins in this class of single-stranded nucleic acids was developed by McGhee and von Hippel (1974). In this model, the nucleic acid is treated as a one-dimensional lattice of nucleotides with a molecule of protein (the ligand) occupying a certain number of lattice sites on the nucleic acid (see Figure 1A and Theory section). The binding site size n, the affinity of binding K_{int}, and the cooperativity factor ω for ligand binding adjacent to an already-bound ligand relative to isolated binding are the parameters in this model. As illustrated in Figure 1A, it is imagined that a protein ligand binds to the polynucleotide lattice, effectively removing n nucleotides from further binding. The binding of gene V protein to polynucleotides is highly dependent on the concentrations of ions. The theory of Record and co-workers (see Theory section) (Newport et al., 1981; Record et al., 1978, 1976, 1981) has been used to describe this dependence for gene V protein (Alma et al., 1983; Bulsink et al., 1985; Stassen et al., 1992a).

Though this model has been used to analyze experiments carried out on gene V protein, it is not entirely appropriate because gene V protein is thought to bind to nucleic acids as a dimer and presumably interacts at once with two separate strands or segments of a single strand (Brayer & McPherson, 1984; Cavalieri et al., 1976). That is, the binding of a dimer of gene V protein to polynucleotides is more likely to look like Figure 1B, in which a dimer of gene V protein interacts simultaneously with two strands of nucleic acid, than Figure 1A, in which only a monomer of protein is considered. The view of gene V protein−polynucleotide interactions depicted in Figure 1B gives a different interpretation of the value of the salt dependence of ssDNA binding than the model illustrated in Figure 1A because the value of C (see eq 1, below) is a measure of the numbers of ions released for one binding unit. This means that, depending on whether the gene V protein binding is considered to be the binding of a monomer or a dimer, the estimate of the numbers of ions released per polypeptide chain will differ by a factor of 2. A second aim of this work was to examine whether a simple extension of the theories of McGhee and von Hippel and of Record (McGhee & von Hippel, 1974; Record et al., 1976) could be reasonably applied to the binding a dimeric protein to polynucleotides, and if so, to evaluate the resulting value of C.

The final aim of this work was to measure the cooperativity of binding of gene V protein to poly(dA), its salt dependence. As described above, the cooperativity of binding of gene V protein and related proteins to single-stranded nucleic acids may be described by a parameter ω, essentially the ratio of the binding constant for binding of a protein to a site adjacent to a previously bound protein, to the binding constant for an isolated location (McGhee & von Hippel, 1974). The cooperativity factor for T4 gene 32 protein has been shown to be over 1000 (Chase et al., 1986; Kowalczykowski et al., 1981b). *E. coli* SSB protein binds to single-stranded DNA in several different modes of binding, with cooperativity factors ranging from 50 to 10^5 (Lohman & Ferrari, 1994). A number of measurements of the value of ω for gene V protein have been made using the effect of the cooperativity parameter on the shapes of several types of binding titrations (Kowalczykowski et al., 1986). Salt

titrations of gene V protein−poly(dA) complexes were interpreted by Bulsink et al. (1985) in terms of a value of ω of about 500. Titrations of poly(dA) with gene V protein at concentrations of salt ranging from 0.11 to 0.22 M NaCl were carried out by Alma et al. (1983), yielding estimates of ω ranging from 50 to 290, with no clear dependence on salt. Similar titrations of poly(dA) with gene V protein by Bulsink et al. (1985) and by Stassen et al. (1992a) yielded estimates of 130 and 600, respectively for ω. Porschke and Rauh (Porschke et al., 1983) used titrations of gene V protein with poly(dA) at a salt concentration of 0.11 M NaCl to estimate a value of ω of about 800. Estimates of the cooperativity factor ω therefore vary over the rather wide range of 50−800 under similar conditions. In this work, we have attempted to make a set of measurements of this cooperativity factor under conditions where the shapes of binding titration curves are as sensitive as possible to the value of ω. In addition, we have estimated the cooperativity parameter at two temperatures and as a function of salt concentration.

MATERIALS AND METHODS

Gene V Protein Purification. Gene V protein was purified as described earlier (Zabin, 1991) using ssDNA−agarose and size-exclusion chromatography as purification steps. The protein had a final purity of at least 98% as determined by inspection of Coomassie blue-stained SDS−polyacrylamide gels (Ito et al., 1980; Zabin, 1991).

Purification of ^{35}S-Labeled Gene V Protein. *E. coli* strain K561 containing plasmids pTT15 and pTT18 (Terwilliger, 1988) was grown in minimal A media (Miller, 1972) supplemented with 14 μg/mL methionine containing 300 μCi of ^{35}S, and 14 μg/mL of each of the other 19 amino acids. The cells were induced to express gene V protein with isopropyl β-D-thiogalactoside and were grown overnight at 37 °C with vigorous shaking. The ^{35}S-labeled cells were washed and broken by sonication, and ^{35}S-labeled gene V protein was purified in essentially the same way as the unlabeled wild-type protein. The purified ^{35}S-labeled protein was stored at −20 °C at a concentration of about 0.25 mM in 1 M NaCl, 1 mM EDTA, 20 mM Tris-HCl, pH 8.0. Radioactive gene V protein concentrations were determined using the known specific activity of the [^{35}S]methionine and two methionine residues in the gene V protein monomer as a basis.

Size-Exclusion Chromatography. A Bio-Rad TSK-125 column (300 × 7.5 mm) was used at room temperature (23 °C) with a flow rate of 1 mL/min for all size-exclusion chromatography. The column was calibrated with thyroglobulin (670 kDa), IgG (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa), cytochrome *c* (12.4 kDa), and cyanocobalamin (1.4 kDa). Buffers used were buffer C (0.2 M NaCl, 10 mM Tris-HCl, pH 7.4, 1 mM EDTA) and buffer D (1.0 M NaCl, 8 mM Tris-HCl, Ph 7.4, 0.8 mM EDTA). Samples were brought to the composition of the mobile phase using stock solutions before injection, except that samples of gene V protein at 1.2×10^{-5} and 1.5×10^{-6} M run in buffer C and the sample at 1.0×10^{-6} M in buffer D contained 100%, 10%, and 10%, respectively, of a solution of 0.02 M Tris-HCl, pH 8.1, 1 mM 2-mercaptoethanol, 5 mM EDTA, 1 M NaCl, and 10% glycerol. Samples were injected without further equilibration. Proteins were detected by absorbance at 276 nm, and 0.25 mL fractions were counted for radioactivity by scintillation counting.

Molecular mass estimates were obtained by assuming a linear relationship between retention time on the TSK-125 column and the logarithm of the molecular mass, and using interpolation between the myoglobin and ovalbumin molecular mass standards (Corbett & Roche, 1984). Final concentrations were taken to be the maximum concentration in the peak of gene V protein after separation on the TSK-125 column and were determined from the radioactivity in fractions or from the absorbance of the peak of gene V protein, as appropriate. Starting concentrations of gene V protein were the concentrations in the injected sample and were calculated from the composition of the injected sample or the total amount of gene V protein recovered after separation divided by the sample volume. The ratio of initial to final concentrations estimated in this way was less than 4:1 in all cases. The average concentration during the separation was taken to be the average of the initial and final concentrations.

Extinction Coefficients. Gene V protein concentration was determined using an extinction coefficient of $\epsilon_{276} = 7100$ M^{-1} cm^{-1} (Day, 1973) and is expressed in units of M of subunits. Poly(dA) was obtained from Sigma, and its concentration (in nucleotide units) was determined using an extinction coefficient of $\epsilon_{260} = 9650$ M^{-1} cm^{-1} at 260 nm in 1 mM sodium phosphate, pH 8.0 (Bollum, 1966). The lengths of the poly(dA) strands were reported by the manufacturer to range from approximately 400 to 1700 nucleotides as estimated by gel electrophoresis.

The extinction coefficients used in absorption corrections for gene V protein were obtained by measuring the absorbance of a stock solution of protein at various wavelengths. For gene V protein alone, the extinction coefficient at 303 nm was $\epsilon_{303} = 160$ M^{-1} cm^{-1}. The extinction coefficient for poly(dA) is sensitive to the components of the buffer (Kansy et al., 1986), decreasing at 260 nm by about 3% between 1 mM Tris, pH 8.0, and 25 mM MOPS, pH 7.0, 50 mM NaCl, and 0.5 mM EDTA. Approximate values of extinction coefficients at various wavelengths were obtained in 25 mM MOPS, pH 7.0, 50 mM NaCl, and 0.5 mM EDTA, based on an extinction coefficient of $\epsilon_{260} = 9300$ M^{-1} cm^{-1}, and were used for absorption corrections in all experiments reported here. The absorption correction is complicated somewhat in our experiments by the fact that the absorbance of the single-stranded DNA−gene V protein complex is not equal to the sum of the absorbances of the components. It was assumed that this additional absorbance is due to partial unstacking of the nucleotide bases when bound to gene V protein (Kansy et al., 1986). To estimate the extent of this additional absorbance, the absorbances at 260, 276, and 303 nm of 141 μM poly(dA) and 32 μM gene V protein were measured separately and together in a buffer containing 25 mM MOPS, pH 7.0, 50 mM NaCl, and 0.5 mM EDTA. Based on previously measured values of binding parameters $K\omega$ and C (Zabin et al., 1991), and using a value of $n = 4$ (Kansy et al., 1986), it was expected that, in the solution containing both protein and poly(dA), 100% of the gene V protein will be bound to the poly(dA) and 91% of the poly(dA) will be bound to gene V protein. The additional absorbance of the mixture relative to its components was divided by the concentration of nucleotides bound to protein to yield an estimate of the increase in extinction coefficient of the poly(dA) when bound to protein. The values of extinction coefficients for poly(dA) alone at 276 and 303

nm were $\epsilon_{276} = 4100\ M^{-1}\ cm^{-1}$ and $\epsilon_{303} = 20\ M^{-1}\ cm^{-1}$, respectively. When poly(dA) was bound to protein, the effective increase in absorbance of the complex over that of the components, expressed as if it were due to an increase in poly(dA) absorbance, was $\Delta\epsilon_{276} = 1600\ M^{-1}\ cm^{-1}$ and $\Delta\epsilon_{303} = 100\ M^{-1}\ cm^{-1}$, respectively. In the modeling procedures described below, the parameters in the models were used to estimate the fraction of the gene V protein and of poly(dA) that are bound to each other. The absorption coefficients for the complex were then used for these portions of the protein and DNA in calculating an overall absorption correction.

Fluorescence Titrations and Absorption Corrections. Titrations were carried out either at 3 °C using a Perkin-Elmer fluorimeter or at room temperature (approximately 23 °C) using a Photon Technologies International fluorimeter. In each case, the fluorimeters were equipped with a magnetic stirrer. An excitation wavelength of 276 nm and an emission wavelength of 303 nm were used. An empirically determined effective path length was obtained for each fluorimeter using samples with known absorption and fluorescence at varying concentrations. Corrections for absorption due to the sample at the excitation and emission wavelengths were made using these path lengths (Lakowicz, 1983). The buffers used for all titration experiments contained 10 mM MOPS (2-(*N*-morpholino)propanesulfonic acid), 0.1 mM EDTA, and varying concentrations of NaCl. Buffer A contained 14 mM NaCl.

Measurement of Parameters in the Simultaneous Binding Model for Gene V Protein Binding to Poly(dA). The parameters in the simultaneous-binding model described below for gene V protein binding to poly(dA) were obtained using a set of three different types of experiments, each of which has been applied to gene V protein previously (Alma et al., 1983; Bulsink et al., 1985, 1988; Kansy et al., 1986; Porschke et al., 1983; Stassen et al., 1992a; Zabin et al., 1991). The binding of gene V protein to ssDNA is monitored in each case by the decrease in intrinsic tyrosyl fluorescence that accompanies binding. The molar fluorescence of the protein alone (F_{prot}) was calculated directly from the fluorescence of the highest-concentration sample used in an experiment, corrected for absorption.

In the first experiment, gene V protein is titrated with poly(dA) in a low ionic strength buffer. The binding of gene V protein to ssDNA becomes increasingly strong as the [NaCl] is reduced, so that in buffer A at $[Na^+] = 39\ \mu M$ the effective binding constant $K_{int}\omega$ is greater than $10^8\ M^{-1}$ (Zabin, 1991). A solution of gene V protein at a concentration of 0.1−8 μM is titrated with 300 μM poly(dA) up to a nucleotide: protein monomer ratio of approximately 6:1 in buffer A. This experiment yields information on the molar fluorescence of the free protein (F_{prot}) and of the protein bound to poly(dA) (F_{bound}). This experiment also yields an estimate of the ratio of nucleotides to protein subunits when the protein just saturates the ssDNA lattice (n).

The second experiment is a salt titration of protein−ssDNA complexes formed in the first experiment using a solution of 2.5 or 5 M NaCl in buffer A. For situations in which the cooperativity parameter is high, as it is for the gene V protein, this titration is very insensitive to n, depending rather on the effective binding constant, $K_{int}\omega$ (Bulsink et al., 1985), the dependence of $K_{int}\omega$ on salt, C, and to a lesser extent on the value of the cooperativity, ω. A value of $n = 4$ is used based on previous work (Alma et al., 1983; Bulsink et al., 1986; Kansy et al., 1986; Zabin et al., 1991). A value of C was estimated by carrying out salt titration of gene V protein−poly(dA) complexes at varying concentrations of the complex and determining for each the salt concentration where half the protein is bound (Bulsink et al., 1985; Stassen et al., 1992b). In the analysis of this experiment, a value of the molar fluorescence of the free protein after salt dissociation of the complexes, F_{prot}, is estimated separately from the value obtained in the first experiment to compensate for any quenching of the fluorescence by the ssDNA or any incomplete dissociation from ssDNA by high salt. The value obtained in this way was within 5% of that calculated for the first experiment, except for the two experiments in Figure 4C at the lowest protein concentrations, where the values differed by 14% and 20%, respectively. For analysis each of the experiments in Figure 3, two analyses were carried out. First, a simultaneous fitting using 6 parameters of F_{bound}, F_{prot} for each of the two parts of the experiment, n, K_{int}, and ω was carried out by least squares. Then, after it was noted that ω did not vary substantially with salt, the mean value of $\omega = 2250$ was used for all the experiments and was not further refined. In all the least-squares fitting used here, the value of χ^2 calculated from the fit to the raw data was minimized.

The third type of experiment was a titration of poly(dA) with gene V protein under conditions where the effective dissociation constant $1/K\omega$ is comparable to the total protein concentration (Kowalczykowski et al., 1986; Lohman & Mascotti, 1992). Under these conditions, the concentration of free protein (L_f) is very sensitive both to the cooperativity ω and to the binding constant K_{int}. In each experiment, a solution containing gene V protein is titrated with a second solution that is identical except that it contains 300 μM poly(dA) and no protein. The buffer is buffer A plus a concentration of NaCl chosen so that $1/K_{int}\omega$ is approximately 1/2 the total protein concentration.

As the cooperativity of gene V protein binding to poly(dA) is quite high, it is important to consider whether the finite length of the poly(dA) strands used substantially affects our analysis. One way to address this is to examine the mean cluster length of gene V proteins bound to poly(dA) (estimated from the model) and compare this to the lengths of the strands of poly(dA). The poly(dA) we use is reported by the manufacturer (Sigma) to have a size range of approximately 400−1700 nucleotides. The mean cluster length estimated from the model for the experiments reported here varies quite substantially with the degree of saturation of the poly(dA), but the length that is most relevant in evaluating the need to include end effects is the cluster length in the portion of the titrations where the fit is most sensitive to ω. In salt titrations, this sensitive portion of the titrations occurs when the degree of saturation of the poly(dA) is low, and where the cluster length is consequently relatively short. In Figure 4D, for example, this is near 0.165 M NaCl, where the average cluster length is estimated from the theory of McGhee and von Hippel (1974) to be just 6. Using a ratio of 8 nucleotides bound to each gene V protein dimer, this corresponds to just 48 nucleotides. In Figure 4E, the most sensitive portion of the titration is in the region where the concentration of poly(dA) is near 40 μM, where the average cluster length is estimated to be about 12, and the corresponding number of nucleotides covered by a cluster of gene

V protein dimers about 100. In each of these cases, the length of ssDNA bound by a typical cluster of gene V protein is much shorter than the poly(dA) length, suggesting that end effects do not have a substantial effect on our measurements. These conclusions are similar to those arrived at by others using Monte Carlo methods (Epstein, 1978).

THEORY

Salt Dependence of Gene V Protein Binding to ssDNA. Record and co-workers have described a model useful for analyzing ligand binding to a polyelectrolyte (Newport et al., 1981; Record et al., 1978, 1976, 1981). In this model, the logarithm of the effective binding constant, $\log(K_{int}\omega)$, is linearly related to the logarithm of the salt concentration, log([NaCl]), with a slope (C) depending on the effective number of cations released from the DNA upon complex formation, $m'\psi$, and on the number of anions released from the protein ligand upon binding, k:

$$\frac{\mathrm{d}\log(K_{int}\omega)}{\mathrm{d}\log([\mathrm{NaCl}])} = C = -(m'\psi + k) \qquad (1)$$

In eq 1 m' is the number of ion pairs formed with the nucleic acid per molecule of ligand and ψ is the fractional binding of cations to negative charges on the nucleic acid. The value of the slope (C) for gene V protein binding to a variety of polynucleotides has been found to be approximately $C = -4$ (Alma et al., 1983; Bulsink et al., 1985; Stassen et al., 1992a; Zabin et al., 1991). This has been interpreted for gene V protein in terms of $m' = 4$ ion pairs formed per protein monomer during complex formation (one for each of 4 nucleotides bound per protein monomer), a fractional binding of cations to negative charges on the nucleic acid of $\psi = 0.75$, and k = 1 anion displaced per protein molecule bound (Alma et al., 1983; Bulsink et al., 1985).

Cooperative Binding of a Monomeric Protein to a ssDNA Lattice. McGhee and von Hippel (1974) have developed a detailed model describing the cooperative binding of a monomeric protein to a lattice such as ssDNA. This will be briefly described to provide the background for an extension to one particular case for a dimeric protein. Designating the total concentration of oligonucleotide (in units of nucleotides) by D_t, the fraction of sites that are free as B, and the concentrations of ligand bound and unbound to polynucleotide as L_b and L_f, respectively, the binding constant K_{int}, can be written as,

$$K_{int} = L_b/BD_tL_f \qquad (2)$$

Next, defining the binding density ν as the ratio of the concentration of ligand bound to polynucleotide L_b to the total concentration of oligonucleotide D_t, expressed in units of nucleotide monomers,

$$\nu = L_b/D_t \qquad (3)$$

and letting the total ligand concentration be L_t, this model leads to the relation (McGhee & von Hippel, 1974),

$$\nu = K_{int}B(\omega,n,\nu)(L_t - \nu D_t) \qquad (4)$$

where the fraction of lattice sites on the polynucleotide that are free is now written as $B(\omega,n,\nu)$ to emphasize its dependence on ω, n, and ν, and is given by,

$$B(\omega,n,\nu) = (1 - n\nu) \times ([(2\omega - 1)(1 - n\nu) + \nu - \{[1 - (n + 1)\nu]^2 + 4\omega\nu(1 - n\nu)\}^{1/2}]/[2(\omega - 1)(1 - n\nu)])^{n-1} \times ([1 - (n + 1)\nu + \{[1 - (n + 1)\nu]^2 + 4\omega\nu(1 - n\nu)\}^{1/2}]/[2(1 - n\nu)]) \qquad (5)$$

Given the total concentration of ligand (L_t) and polynucleotide (D_t), and the parameters K_{int}, ω, and n, eqs 2−5 may be used to determine the binding density ν and therefore the concentration (L_b) of ligand bound to the polynucleotide under this specified set of conditions.

Cooperative Binding of a Dimeric Protein to a ssDNA Lattice: Simultaneous Binding Model. Previous analyses of the binding of gene V protein to single-stranded nucleic acids have been based on the model of McGhee and von Hippel (1974) and the assumption that the functional binding unit of the protein is a monomer (Figure 1A). The gene V protein, however, is a dimeric protein that forms a relatively well-characterized helical structure when it binds to ssDNA. This structure has been analyzed by electron microscopy and solution scattering methods, and consists of a regular left-handed superhelix in which the gene V protein dimers are arrayed on the outside of the superhelix and antiparallel ssDNA strands are inside (Gray, 1989; Gray et al., 1982). Several modeling studies have indicated that in this superhelix adjacent dimers of gene V protein are packed closely against each other (Folmer et al., 1994; Guan et al., 1993; Skinner et al., 1994).

One way to explicitly incorporate the dimeric nature of the gene V protein and the essential elements of the superhelical structure formed when it binds to ssDNA into a model for its interactions with ssDNA is to assume that the two subunits of the dimer bind simultaneously to two antiparallel, separate, strands of ssDNA. In the simultaneous model, any time a gene V protein subunit binds to a strand, the other subunit within the dimer binds to another strand (Figure 1B). In essence, this is equivalent to the assumption of infinite positive cooperativity of binding for the two binding sites in a gene V protein dimer, and to ignoring any partially bound dimers. For simplicity, it is assumed here that these two strands are in fixed register with respect to one another. This is not necessarily the case, but if the binding is very cooperative (as it is for gene V protein) then most binding events will be adjacent to a dimer already bound and the register of the ssDNA in the two strands will be essentially fixed.

The model pictured in Figure 1B can be analyzed in almost the same way as McGhee and von Hippel (1974) analyzed the model in Figure 1A. The key to the approach used here is to note that the gene V protein is always a dimer and every time one monomer binds to a site on one strand of ssDNA the other monomer in the dimer binds to a second strand of ssDNA. As long as this is the case, we can ignore exactly half of the (free or bound) monomers of protein and half of the (free or bound) nucleotides of ssDNA, because whatever happens to the first half of the protein monomers and nucleic acid will happen at the same time to the other half. The equilibrium between bound and free ligand (monomers) in Figure 1B can be written in much the same way as in eq 2, except that only half the protein monomers and half the ssDNA present are considered. Once again, define the affinity of binding as K_{int} (this time of the dimer to 2 binding sites of n nucleotides each), and the cooperativity factor as

ω. Next, designate the total concentration of oligonucleotide by D_t, and the concentration to be considered in the analysis as $(1/2)D_t$. As before, the fraction of sites that are free is B. The concentrations of ligand bound and unbound to polynucleotide are L_b and L_f, respectively, but the concentrations to be considered in Figure 1B are just half of these, $(1/2)L_b$ and $(1/2)L_f$, respectively. The binding constant K_{int}, for the model in Figure 1b can now be written as,

$$K_{int} = \frac{(^1/_2 L_b)}{B(^1/_2 D_t)(^1/_2 L_f)} \quad (6)$$

Next, define the binding density ν as the ratio of the concentration of bound ligand L_b to the total concentration of oligonucleotide D_t (eq 3), and, let the total ligand (protein subunit) concentration be L_t, and the ligand concentration considered be $(1/2)L_t$. Using exactly the same mathematics as was used to derive eq 4, this leads to,

$$\nu = {}^1/_2 K_{int} B(\omega,n,\nu)(L_t - \nu D_t) \quad (7)$$

Except for the factor of $^1/_2$ on the right side of the equation this is identical to eq 4. Equation 7 describes the simultaneous binding of a dimeric protein to two strands of ssDNA where the total concentrations of protein monomers and ssDNA nucleotides in solution are L_t, and D_t, respectively.

If the same data were analyzed with eqs 4 and 7, the values of all the parameters (K_{int}, ω, n, ν) would be identical except that the value of K_{int} for the simultaneous binding model (eq 7) would be twice that in the independent binding model (eq 4). This difference in K_{int} reflects the fact that the binding unit in the simultaneous binding model is the dimer. The interpretations of the values of ω and n obtained from the analysis of a series of experiments according to eqs 4 or 7 are essentially identical. In either case, ω is the ratio of binding affinity at a site adjacent to an occupied site and n is the number of nucleotides covered by one monomer of protein. In the analysis using eq 7, of course, the number of nucleotides covered by one protein dimer is $2n$.

As the simultaneous binding model and the independent binding model yield values of the binding constant K_{int} that differ by exactly a factor of 2, the analysis of the salt dependence of binding using the two models will yield the same values of the parameter C in eq 1. This parameter C reflects the effective number of cations and anions released from the nucleic acid and protein ligand upon binding of one protein binding unit to nucleic acid (Newport et al., 1981; Record et al., 1978, 1976, 1981; Zhang et al., 1996). The independent and simultaneous binding models differ in an important way in the interpretation of this parameter C. In the independent model, C corresponds to displacement of cations and anions from one monomer of the protein and the n bound nucleotides, while in the simultaneous binding model, C corresponds to the displacement of cations and anions from one protein dimer and the $2n$ bound nucleotides.

The simultaneous binding model is not expected to describe the binding to isolated sites on a pair of strands in complete detail, as this binding can occur in many ways that are not considered. For example, this binding could consist of two steps. In the first step, one subunit of a dimer might bind transiently to a strand of ssDNA, followed in a second step by bending back of the strand and binding by the second subunit. The loop formed in this way could be of essentially any length. Despite the complexity of the binding process, however, it seems likely that once one subunit of a dimer binds to ssDNA, the other subunit will have an effective binding constant far greater than it had when the dimer was free in solution, simply due to the proximity of the ssDNA to the second subunit. In describe the binding of a dimer to two strands of ssDNA as simultaneous, then, the initial binding of one subunit to one strand and the subsequent and relatively rapid binding of the second subunit to another part of the same strand are treated as a single process.

RESULTS AND DISCUSSION

HPLC Size-Exclusion Chromatography of Gene V Protein. HPLC size-exclusion chromatography has been demonstrated to be an effective means of measuring the Stokes radii of proteins and to be a reliable method for estimation of molecular mass (Corbett et al., 1984). As it is possible to detect gene V protein by absorbance at concentrations as low as 2×10^{-7} M and by scintillation counting of ^{35}S-labeled protein at concentrations as low as 10^{-10} M, this technique seemed well-suited for determining the oligomerization state of the protein at low concentrations. Figure 2A shows the results of a size-exclusion separation of radioactively labeled gene V protein at concentrations of 5×10^{-10} and 1.2×10^{-5} M. For the separation of 1.2×10^{-5} M protein, the ^{35}S-labeled protein was simply mixed with unlabeled gene V protein. Figure 2A illustrates that the retention time of radioactively labeled gene V protein was identical at these two concentrations. Comparison with molecular mass standards (included in the run at 5×10^{-10} M gene V protein) shows that the gene V protein has an effective molecular mass of 18.2 kDa in both cases. As the molecular mass of one monomer of the gene V protein is 9.7 kDa, the gene V protein is evidently a dimer at both concentrations.

The effective molecular mass of the gene V protein was measured at two different salt concentrations and at several protein concentrations. Figure 2B illustrates that the effective molecular mass of the gene V protein is independent of NaCl concentration from 0.2 to 1.0 M and is independent of protein concentration from 5×10^{-10} to 1.2×10^{-5} M. To be certain that the radioactively labeled gene V protein had the same retention time as the bulk of the unlabeled protein in samples containing both, the absorbance at 276 nm was used to monitor the unlabeled gene V protein. As illustrated in Figure 2, the effective molecular mass estimated from these absorbance measurements was the same as that estimated from the retention time of labeled protein.

These results demonstrate that the gene V protein is a dimer at concentrations as low as 10^{-9} M in 0.2 or 1 M NaCl solutions. In practical terms, this means that for *in vitro* experiments carried out in the micromolar concentration range the gene V protein is always a dimer, and monomer−dimer equilibria may safely be neglected. Although this conclusion is, as described in the introduction, at variance with some earlier work using equilibrium ultracentrifugation (Cavalieri et al., 1976; Oey et al., 1972; Porschke et al., 1983; Pretorius et al., 1975), it is in good agreement with more recent work on the properties of gene V protein in solutions

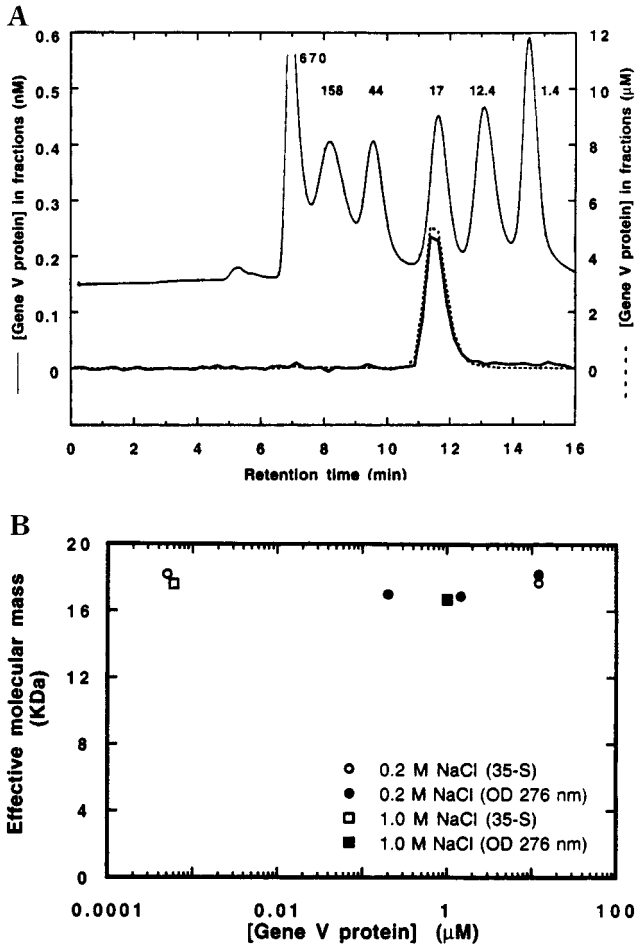

FIGURE 2: Size-exclusion chromatography of ^{35}S-labeled gene V protein. (A) Average concentrations during separation were 5×10^{-10} M (solid line) and 1.2×10^{-5} M (dotted line). Radioactivity in fractions are shown in the lower part of the figure, and elution positions of molecular mass standards, monitored by absorbance at 276 nm, are shown in the upper portion. (B) Estimates of average molecular mass of gene V protein under various conditions (see text). Open symbols are detection by ^{35}S; closed symbols are detection by absorbance. Circles are in 0.2 M NaCl; squares are in 1.0 M NaCl.

containing guanidine hydrochloride (Liang et al., 1991). The concentration dependence of folding of the gene V protein in solutions containing 2.7 M guanidine hydrochloride demonstrates that the folded form of the protein is a dimer even at this high concentration of denaturant. A folded, monomeric form of the protein was not detected and probably makes up less than 10% of the protein at all concentrations monitored ($(1-40) \times 10^{-6}$ M protein).

Binding Model-Independent Analysis of Gene V Protein Binding to Poly(dA). Gene V protein binding to polynucleotides can be readily monitored by the decrease in protein tyrosyl fluorescence that occurs upon complex formation. The analyses of gene V protein binding to poly(dA) can be greatly simplified by the assumption that quenching of protein fluorescence upon binding to the nucleic acid is proportional to the fraction of protein that is bound. Such an assumption is not necessarily justified, as binding of the protein to an isolated site on the nucleic acid might affect the fluorescence differently than binding at the end or in the middle of a cluster of gene V protein molecules. We therefore carried out an experiment to determine whether fluorescence quenching was proportional to the fraction of protein bound to poly(dA), at least over the range of binding densities that occur in the experiments carried out here. A binding model-independent procedure for carrying out this analysis was followed (Lohman & Bujalowski, 1991) using the fluorescence data corrected for absorption effects shown in Figure 3A.

In the first step of this model-independent analysis, the ligand binding density function (LBDF) is plotted as a function of added poly(dA) (Figure 3B). Here the LBDF is defined as,

$$\mathrm{LBDF} = (F_{obs} - F_f L_t)/(F_f L_t)(L_t/D_t) \qquad (8)$$

where F_{obs} is the corrected fluorescence of a sample, F_f is the molar fluorescence in the absence of added poly(dA), L_t is the concentration of protein monomers in the sample, and D_t is the concentration of poly(dA) in the sample. If the LBDF is defined in this way (Lohman et al., 1991), sets of points with the same values of the LBDF in Figure 3B have the same values of the binding density $\sum\nu_i$, the ratio of bound ligand to total poly(dA) concentration ($\sum\nu_i = L_b/D_t$). This means that sets of values of (L_t, D_t) that all lead to the same value of $\sum\nu_i$ can be found along horizontal lines in Figure 3B. For values of the LBDF between 0.08 and 0.11 in Figure 3B, two sets of values of (L_t, D_t) are present for each value of the LBDF, and for values from 0.11 to 0.18, there are three. For values of the LBDF greater than about 0.18, the LBDF is not very sensitive to values of L_t and D_t, so the analysis is restricted to values less than this (Lohman et al., 1991).

Given two or more combinations of total protein and poly(dA) concentrations, (L_t, D_t), that yield the same value of the binding density $\sum\nu_i$, it is possible to estimate both this value of $\sum\nu_i$ and the concentration of free ligand, L_f, by fitting the linear conservation-of-mass equation,

$$L_t = D_t\sum\nu_i + L_f \qquad (9)$$

where $D_t\sum\nu_i = L_b$ is the total amount of protein bound to poly(dA) and L_f is the amount that is free (Lohman et al., 1991). In Figure 3C, the total protein concentration L_t is plotted as a function of total poly(dA) concentration D_t for seven values of the binding density $\sum\nu_i$, obtained over a range of LBDF values from 0.11 to 0.17. As expected from the theoretical treatment of this type of binding (Lohman et al., 1991), the relationship between D_t and L_t is linear in each case. Each set of points with constant binding density ν was fit to eq 9 in order to obtain the value of the binding density $\sum\nu_i$ and of the free protein concentration, L_f.

The purpose of this analysis was to determine whether the quenching of fluorescence was proportional to the fraction of protein bound to poly(dA), and therefore whether the quenching could be used as an accurate measure of this binding. Figure 3D illustrates that the relationship between the fraction of protein bound to poly(dA), given by L_b/L_t, and the fluorescence quenching, given by $Q = (F_{obs} - F_0)/F_0$, is indeed linear. Each of the three clusters of points in Figure 3D is derived from one of the three binding curves in Figure 3A. The points are clustered in each case because the fraction of protein bound to nucleic acid is almost constant in each case over the range that is examined (compare Figures 3A and 3B). Extrapolating to the case

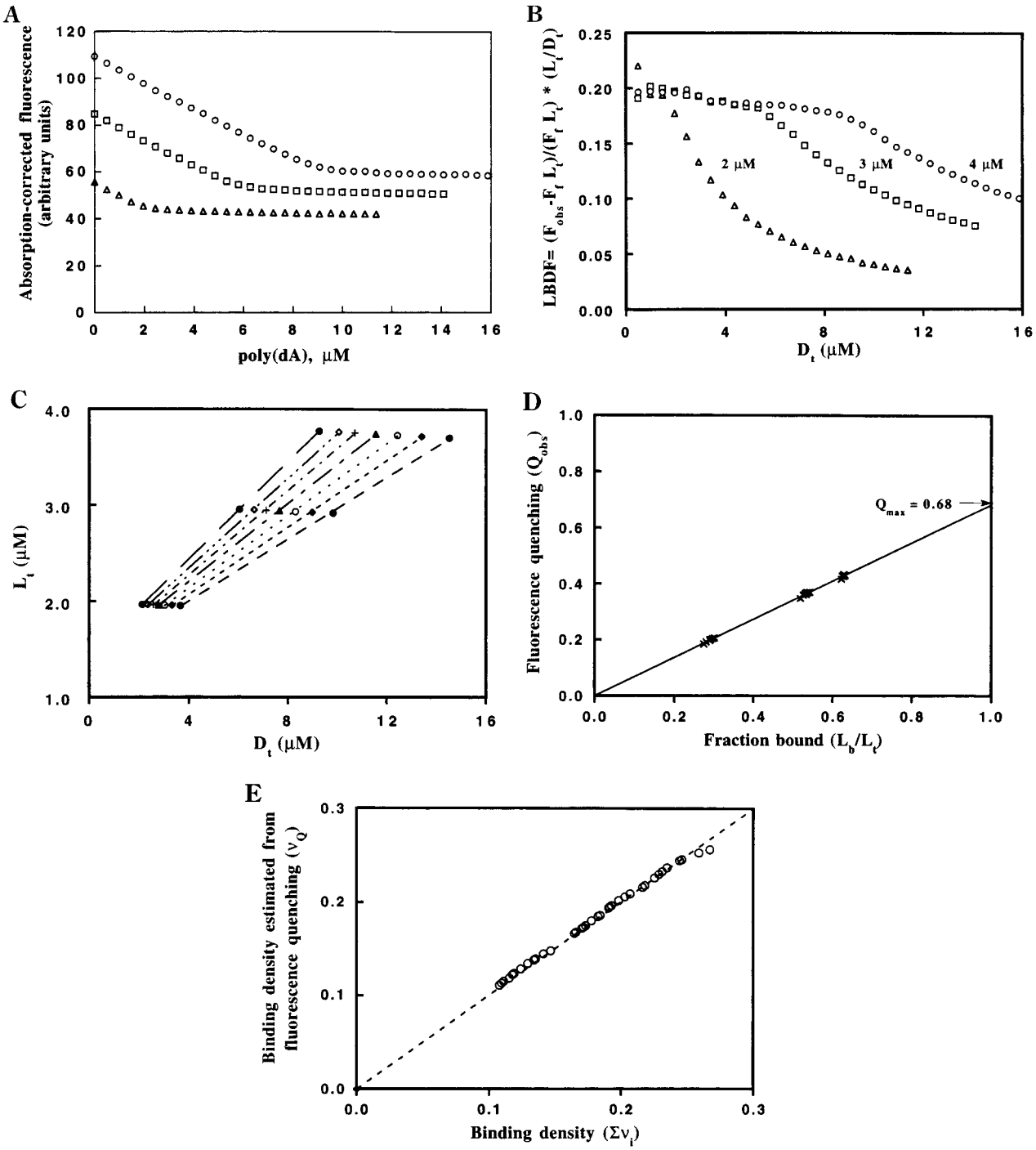

FIGURE 3: Titration of gene V protein with poly(dA) in 0.139 M NaCl at 23 °C and ligand binding density function (LBDF) analysis. (A) Absorption-corrected raw fluorescence. Total initial protein monomer concentrations were 2.0, 3.0, and 3.9 μM in the three experiments shown. (B) Ligand binding density function (LBDF) analysis. LBDF values (eq 8) are plotted as a function of the added nucleotide concentration for the 3 experiments shown in panel A. (C) Pairs of values of (L_t, D_t) corresponding to constant values of the LBDF. Pairs corresponding to LBDF values (right to left in the figure) of 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, and 0.17 are shown along with linear least-squares fits. (D) Fluorescence quenching as a function of fraction of protein bound to poly(dA). Values of L_t, D_t, and $\sum\nu_i$ are from the regressions shown in panel C. The value of L_b/L_t is given by: $L_b/L_t = \sum\nu_i D_t/L_t$. Q_{obs} is given (see eq 8) by: $(F_{obs} - F_f L_t)/(F_f L_t)$. (E) Comparison of binding density obtained from LBDF analysis with that obtained using the assumption that fluorescence quenching is proportional to the fraction of protein bound to poly(dA). Values of L_t, D_t, and $\sum\nu_i$ were obtained from regressions similar to those shown in panel D, except the range of values of the LBDF considered was from 0.08 to 0.18. Values of the binding density were estimated from fluorescence quenching using the relation, $\nu_Q = (Q_{obs}/Q_{max})(L_b/D_t)$.

where all of the protein is bound to DNA ($L_b/L_t = 1$), the maximum fluorescence quenching is found to be $Q_{max} = 0.68$.

The linearity of the data in Figure 3D suggests that the fractional fluorescence quenching of gene V protein is equal to the fraction of gene V protein bound to poly(dA), i.e., that $Q_{obs}/Q_{max} = L_b/L_t$. This assumption can be checked more thoroughly by comparing the binding density estimated using this assumption, $\nu_Q = (Q_{obs}/Q_{max})\,(L_b/D_t)$, with that obtained from the more rigorous LBDF analysis above, $\sum\nu_i$ (Lohman et al., 1991, 1992). This comparison is presented in Figure 3E and shows that the binding density obtained using the assumption that fluorescence quenching is proportional to ssDNA binding is in good agreement with that obtained from

the LBDF analysis over the range of accessible binding densities from 0.10 to 0.25.

Analysis of Gene V Protein Binding to Poly(dA) Using the Simultaneous Binding Model. We carried out measurements of the binding of gene V protein to poly(dA) that can be analyzed in terms of the simultaneous binding model. In making these measurements, particular attention was paid to experiments that yielded information on ω, the cooperativity of binding. These experiments were carried out in several steps. First, a set of gene V protein−poly(dA) complexes were formed by titrating gene V protein at various concentrations with poly(dA) in low salt. These complexes were then dissociated using salt titrations. The salt dependence (C) of binding of gene V protein to poly(dA) was estimated in a model-independent fashion from this group of titrations. The poly(dA) and salt titrations were then initially individually analyzed, fitting 6 parameters, including the cooperativity parameter, ω, to each pair of ssDNA and salt titrations. After finding that ω was independent of salt, a single average value of $\omega = 2250$ was used and 5 parameters were fitted for each experiment. Additionally, a separate experiment to determine the salt dependence of ω was carried out.

The salt dependence of $K_{int}\omega$, the effective binding constant for gene V protein−poly(dA) association, can be obtained without reference to a detailed model for the binding equilibrium. For proteins with highly cooperative binding to nucleic acid, the effective binding constant ($K_{int}\omega$) at a particular salt concentration is approximately equal to the reciprocal of the free protein concentration under conditions where half the protein is bound. The salt dependence of binding (C) can therefore be measured by carrying out salt titrations of protein−nucleic acid complexes at varying concentrations of the complex and determining for each the salt concentration where half the protein is bound (Bulsink et al., 1985; Stassen et al., 1992b). The results of such an analysis of 6 experiments in which gene V protein−poly(dA) complexes at varying concentrations are titrated with salt are illustrated in the upper portion of Figure 4A. In these experiments, gene V protein was titrated first with poly(dA) in a buffer containing 39 mM Na^+ to a nucleotide:protein monomer ratio of about 6:1, followed by titration of the protein−nucleic acid complexes with NaCl to dissociate them. The six experiments differed only in the concentrations of protein and nucleic acid used, with the relative amounts remaining constant. The initial protein concentrations ranged from 0.13 to 8 μM. As expected from eq 1, the logarithm of the effective binding constant is linearly related to the logarithm of the concentration of Na^+. The value of the slope in this case is $C = -4.65$, in the range of values of -3.5 to -4.7 reported earlier (Alma et al., 1983; Bulsink et al., 1986; Stassen et al., 1992a).

The initial titration of protein with nucleic acid in one of these experiments, corrected for background, dilution, and absorption effects, is presented in Figure 4B as a function of the ratio of nucleotides:protein monomers. The protein fluorescence is reduced as poly(dA) is added until the ratio of nucleotides:protein monomers is about 4.9:1, after which it remains relatively constant with a value about 47% of the initial fluorescence. For the 6 experiments in this set, the ratio of nucleotides:protein monomers at saturation ranged from 4.6 to 5.4. The experiment carried out at the highest concentration of protein and nucleic acid, which is likely to have yielded the most accurate estimate of this ratio, showed a ratio of 4.6, slightly higher than those obtained previously under similar conditions which range from 3.6 to 4.3 (Alma et al., 1983; Bulsink et al., 1986; Kansy et al., 1986; Stassen et al., 1992a; Zabin et al., 1991). As we have not attempted to eliminate all systematic sources of error in measurement of n such as nonspecific binding to the fluorescence cuvette, these values are not likely to be significantly different.

The titrations of the gene V protein−DNA complexes at 6 different concentrations with NaCl that were analyzed in Figure 4A are shown in Figure 4C. For each titration, the observed fluorescence is corrected for dilution and absorption and is then used along with the fluorescence of free and bound gene V protein obtained in Figure 4B (F_{prot} and F_{bound}) to estimate the percentage of gene V protein that is free in solution. Fits to each titration were initially calculated as described in Materials and Methods using a fixed value of $C = -4.65$, and adjusting ω to optimize the least-squares fit. The values of ω obtained in this fashion for the 6 experiments shown in Figure 4C are plotted as a function of the salt concentration in the lower portion of Figure 4A. There is very little dependence of ω on salt, and all the values are in the range of 1700−4000, with an average value of 2250. As there appeared to be little salt dependence of ω, the analysis was then repeated using the average value of 2250 for all 6 experiments. The fits of the model with $\omega = 2250$ to these 6 experiments are shown in Figure 4C. The estimates at 0.15 M NaCl of the product $K_{int}\omega$ obtained from the 6 experiments ranged from 12.7×10^5 M^{-1} to 14.4×10^5 M^{-1}, close to values reported earlier (after extrapolation to 0.15 M NaCl and multiplication by a factor of 2 for the simultaneous binding model) in the range of 11.8×10^5 M^{-1} to 14.5×10^5 M^{-1} (Alma et al., 1983; Bulsink et al., 1986; Stassen et al., 1992a).

Two additional analyses were carried out to estimate the cooperativity parameter ω and its uncertainty. The first consisted of fits of the data presented in Figure 4C using several fixed values of ω. An analysis of this kind is shown in Figure 4D for the experiment carried out at 2 μM gene V protein. For this experiment, when the value of ω was obtained directly by least-squares, its value was $\omega = 3200 \pm 1000$. The value of ω was sequentially fixed at 500, 1000, 3000, and 10 000, and the remaining parameters were adjusted as in Figure 4C so as to minimize the discrepancy between observed and calculated fluorescence values. Figure 4D indicates that the optimal value of the cooperativity parameter ω is approximately 3000 and that values lower than 1000 are clearly incompatible with the observed fluorescence values.

The second analysis carried out to estimate ω involved additional experiments in which gene V protein was titrated with poly(dA) in moderate salt concentrations where the reciprocal of the effective binding constant ($1/K_{int}\omega$) was comparable to the protein concentration. Under these conditions the shape of the binding curve is sensitive to the cooperativity parameter (Kowalczykowski et al., 1986; Lohman et al., 1992). An experiment of this type, carried out at a concentration of 0.214 M Na^+, where the reciprocal of the effective binding constant is 4×10^{-6} M, and at a protein concentration of 8×10^{-6} M, is illustrated in Figure 4E. For the solid curve in Figure 4E, the values of n, the product $K_{int}\omega$, and ω were adjusted to optimally fit the observed data. The values obtained for n and ω were 4.8

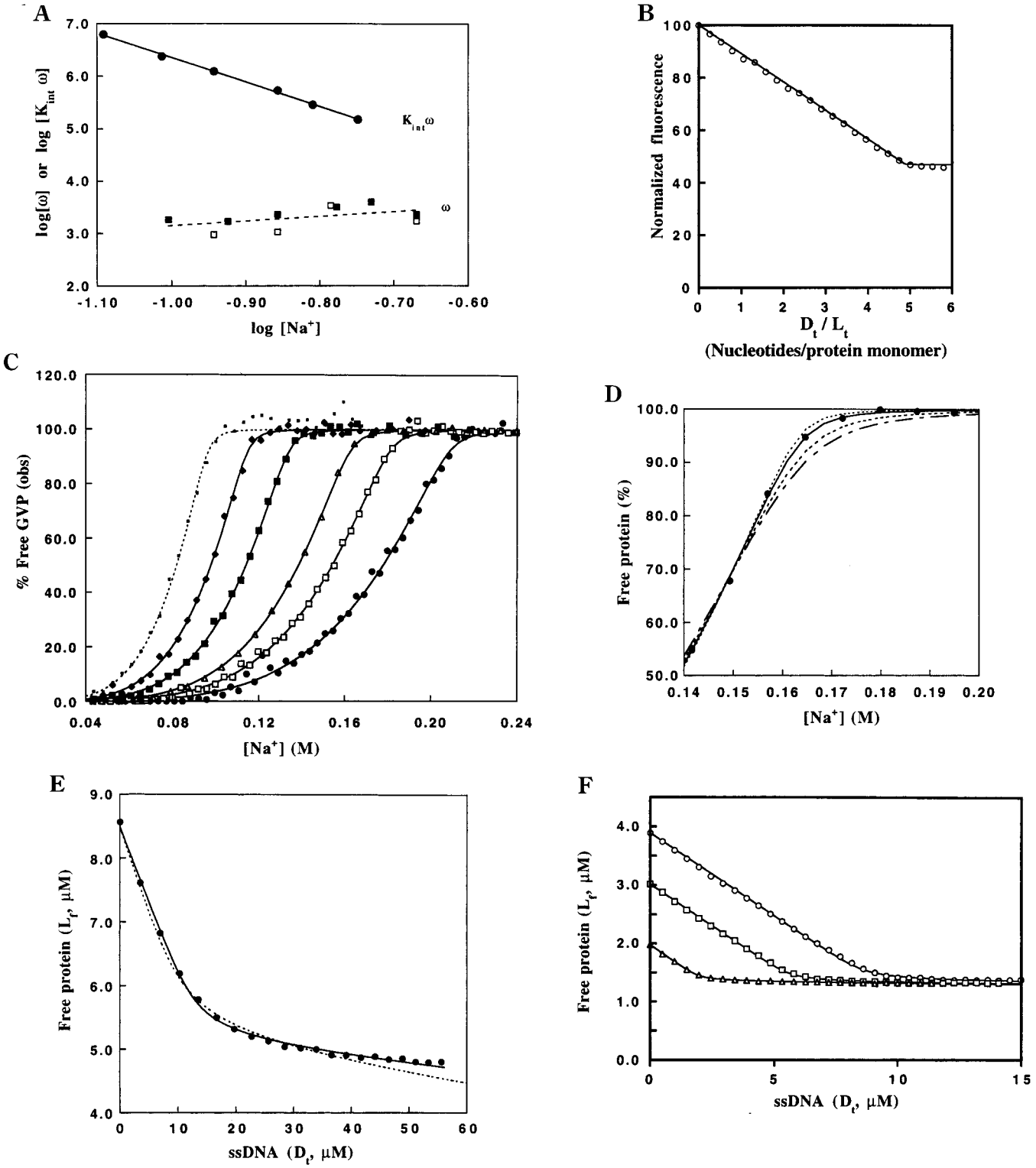

FIGURE 4: Sequential titrations of gene V protein with poly(dA) and NaCl at 3 °C and analysis of cooperativity of binding of gene V protein to poly(dA). (A) Logarithms of effective binding constants, log($K_{int}\omega$), and of cooperativity parameters, log(ω), as a function of log([Na^+]). Closed circles indicate estimates of obtained log($K_{int}\omega$) obtained from the midpoints of the titrations shown in panel C as discussed in the text. Closed squares indicate estimates of log(ω) obtained by least-squares analysis of the full titrations in panel C. Open squares show estimates of log(ω) obtained by analysis of titrations such as the one shown in panel E. (B) Titration of 2 μM gene V protein with poly(dA) to a nucleotide:protein monomer ratio of 6:1 in 39 mM [Na^+]. The normalized fluorescence, corrected for background, absorption, and dilution is plotted as a function of the nucleotide:protein ratio. (C) Titrations of protein−poly(dA) complexes with NaCl. The starting ratio of nucleotides:protein was approximately 6:1 for each experiment, and the starting concentrations of protein were (left to right), 0.16, 0.42, 0.83, 1.9, 3.6, and 6.9 μM. (D) A part of the titration of 1.9 μM gene V protein and 12 μM poly(dA) shown in panel C is illustrated along with least-squares fits (of the entire titration) obtained with ω fixed at 500 (broken line), 1000 (dashed line), 3000 (solid line), and 10 000 (dotted line). (E) Titration of 8 μM gene V protein with poly(dA) in 0.214 M Na^+ at 3 °C. The fluorescence data are fitted allowing ω to vary (solid line, $\omega = 1700$) and fixing ω at 500 (dotted line). (F) Poly(dA) titrations of gene V protein at concentrations of 2, 3, and 3.9 μM in 0.114 M Na^+ at 23 °C, based on the data in Figure 3A. The solid lines indicate free protein calculated from the simultaneous binding model as a function of added poly(dA). The fitted values of ω are 5000, 7000, and 3000 for protein at concentrations of 2, 3, and 3.9 μM, respectively.

and 1800, respectively. For the dotted line, the value of ω was fixed at 500, and the other two parameters were adjusted to fit the data. It is evident that the slope of the observed data at concentrations of poly(dA) above 30 μM is not compatible with a value of the cooperativity parameter of 500, but is quite consistent with the value of 1800. Figure 4A shows the values of ω estimated from a series of 5 experiments carried out at Na^+ concentrations ranging from 0.119 to 0.214 M. These estimates of ω are close to those obtained from the analysis of the salt titration experiments.

The data in Figure 4A show that while the effective binding constant $K_{int}\omega$ is strongly dependent on [Na^+], the cooperativity parameter ω is not. This means that, as has been suggested earlier, the salt dependence of binding resides largely in the intrinsic binding constant K_{int} (Alma et al., 1983; Porschke et al., 1983).

The temperature dependence of the cooperativity parameter for gene V protein binding to poly(dA) was briefly examined by carrying out an experiment similar to that shown in Figure 4E at 23 °C instead of 3 °C. A set of experiments at a single salt concentration (0.114 M Na^+) was carried out in which the starting protein concentration was varied from 2 to 4 μM (Figure 4F, these are the same data that are the basis of Figure 3A). Each of the three titration curves in Figure 4F was fit by optimizing the values of n, the product $K_{int}\omega$, and ω as in Figure 4E. The resulting estimates of n varied from 3.6 to 4.0, the estimates of $K_{int}\omega$ varied from 10.0×10^5 M^{-1} to 10.5×10^5 M^{-1}, and the estimates of ω varied from 3000 to 7000, with an average of 5000. As previously reported, the effective binding constant $K_{int}\omega$ is relatively independent of temperature (Bulsink et al., 1986). We find that the cooperativity parameter changes by only about a factor of 2 between 3 and 23 °C. This small increase is only slightly larger than the 30% uncertainty estimated above in these measurements.

The estimates of the cooperativity parameter ω obtained here (2000−7000) are from 2 to 100 times higher than those previously reported, which range from 50 to 800 (Bulsink et al., 1985; Porschke et al., 1983; Stassen et al., 1992a). The reason for this difference is not entirely clear, but one factor may be that, without very accurate measurements of fluorescence, it is difficult to discriminate between cooperativity values above about 500 (e.g., see Figures 4D and 4E). This may have resulted in lower estimates of the cooperativity factor in previous analyses. The present conclusion that the cooperativity of gene V protein binding to poly(dA) is in the range of 2000−7000 is supported by the agreement between the two methods of estimating this factor (Figure 4A).

As discussed above, the interpretation of the salt dependence of binding affinity in the simultaneous binding model differs from that for the independent binding model used previously. In the simultaneous binding model, gene V protein−nucleic acid binding is in terms of the numbers of anions and cations displaced upon binding of a dimer of protein to 2n nucleotides on the nucleic acid. According to the theory of Record and co-workers (eq 1), the value $-C$ is equal to the sum of the number of anions displaced from the protein and the number of cations displaced from the nucleic acid upon complex formation (Newport et al., 1981; Record et al., 1978, 1976, 1981). The value of C of approximately −4.7 for gene V protein, and the number of nucleotides bound per monomer of approximately 4 under the conditions used here therefore indicates that, for 1 dimer binding to a total of approximately 8 nucleotides, the total number of cations and anions displaced is no more than about 4.7. This is somewhat surprising as it might have been expected that each of the 8 phosphates bound per dimer would have had a corresponding amino acid side chain that would displace any counterions associated with the nucleic acid phosphates when free in solution. Assuming that the average number of cations bound per phosphate group (ψ) is about 0.75 (Record et al., 1978), if each of 8 phosphates formed a salt bridge with protein side chains, the value of $-C$ would be at least 6. This interpretation of C suggests that not all of the nucleic acid phosphates form salt bridges with the protein.

Several mutations of gene V protein are known that strongly affect the salt dependence of binding, though so far none of the basic residues thought to be directly involved in binding has been tested for changes in salt dependence. The mutation Arg-82 → Ser decreases $-C$ to 2.5, and the mutation of Tyr-26 → Cys decreases $-C$ to 2.0 (Stassen et al., 1992a). Arg-82 is probably not located close to bound nucleic acid in gene V protein (Folmer et al., 1994; Guan et al., 1995), and the tyrosine-to-cysteine mutation does not change the charge of the protein. It seems likely that each of these mutations affects the mode of binding in some way that indirectly affects the counterion release upon binding. Consistent with this, the value of n for the Tyr-26 → Cys mutant decreases to 2 and the value for the Arg-82 → Ser mutation decreases to 3.1 (Stassen et al., 1992a). In contrast to these mutations that lower the value of $-C$, mutants Glu-30 → Met and Glu-30 → Phe at Glu-30, an acidic side chain located close to bound nucleic acid in one proposed model of the gene V protein−nucleic acid complex (Guan, et al., 1995), show much higher values of $-C$ of 5.4−5.9 (Zabin et al., 1991). It seems possible that there are cations that remain associated with the bound nucleic acid in the wild-type protein with Glu-30 present, whereas these cations become dissociated from the nucleic acid when it binds to the mutants lacking the Glu-30 side chain.

Significance of Gene V Protein Cooperativity in Binding. The high cooperativity of gene V protein binding to ssDNA has important implications both for the nature of the protein−protein interactions that account for this cooperativity and for the binding site for ssDNA on the protein. As discussed earlier, there is considerable evidence that protein−protein interactions play a major role in gene V protein binding to ssDNA. When bound to ssDNA, the gene V protein forms a fairly regular left-handed superhelical structure (Gray, 1989). It seems likely that the protein−protein interactions between adjacent gene V protein dimers in the structure are at least partially responsible for defining the structure itself (de Jong et al., 1987; Folkers et al., 1991; King et al., 1988). Additionally, recent work has shown that mutants of gene V protein that have substitutions at Glu-40, Tyr-41, and Lys-69 have cooperativities that are reduced from that of the wild-type by factors of 4-fold to 100-fold, and that at least the Tyr-41 → Phe mutant has a greatly reduced propensity for formation of the superhelical structure formed by the wild-type protein (Terwilliger et al., in preparation; C. W. Gray, personal communication).

According to the analysis of gene V protein binding to poly(dA) presented here, the binding of a gene V protein dimer to a site on a pair of ssDNA strands that is adjacent

to a previously bound dimer proceeds with a binding constant that is about 2000−7000 times stronger than does the binding to an isolated site. It is possible that this large factor is partially due to the fact that the site adjacent to a previously bound dimer has two strands of ssDNA already in an arrangement suitable for binding to another dimer of protein, while the binding at an isolated site presumably requires a reorientation of the ssDNA to form two binding sites with the necessary spatial relationship. This entropic factor is not the principal reason for cooperativity, however. As mentioned earlier, the gene V protein from bacteriophage f1 and the related protein from phage IKe do not show mutual cooperativity, though both show cooperative binding (de Jong et al., 1987). This indicates that the observed cooperativity is due to specific interactions between protein dimers bound to ssDNA, not to the proximity effect of bringing two antiparallel ssDNA strands close together. Additional evidence for specific protein−protein interactions comes from the observation described above that mutants of gene V protein can show as much as 100-fold reduced cooperativity.

A corollary of the observation of very high cooperativity for the gene V protein is that binding affinity of the dimer for an isolated pair of strands of ssDNA is very weak. Under physiological salt conditions ([NaCl] = 0.15 M), the effective dissociation constant for gene V protein ($1/K\omega$) for poly(dA) is about 1 μM. If the cooperativity factor ω is about 5000; however, then the intrinsic dissociation constant K for this substrate is only about 5 mM under these conditions. This binding is so weak that it would ordinarily be considered not relevant to processes requiring specificity in binding. It is only because this very weak binding is multiplied 5000-fold by binding adjacent to an existing bound dimer that the binding becomes specific for ssDNA at all.

In a sense, then, a large part of the specificity of gene V protein for ssDNA, as opposed to its binding to any other molecule, resides in the ability of gene V protein to form strong contacts with other dimers and with the DNA when bound to the ssDNA lattice. This view differs somewhat from the view that could be taken in which gene V protein binds to ssDNA simply through strong contacts with the bases of the ssDNA. Based on the present work, the protein−protein interactions may in fact be quite substantial relative to the ssDNA−protein contacts that are found in the complex. It is known that some polynucleotides bind more weakly than others to gene V protein, where poly(dA) is among the weakest (Bulsink et al., 1985). It is possible that, for these weakly-binding polynucleotides, the role of ssDNA in formation of a superhelical complex is little more than to serve as a template upon which the complex is built. Gene V protein at high concentration does form aggregates readily by itself (Folkers et al., 1991); possibly these involve interactions similar to those involved in formation of the superhelical complex.

REFERENCES

Alberts, B., Frey, L., & Delius, H. (1972) *J. Mol. Biol. 68*, 139−152.

Alma, N. C. M., Harmsen, B. J. M., de Jong, E. A. M., van der Ven, J., & Hilbers, C. W. (1983) *J. Mol. Biol. 163*, 47−62.

Baas, P. D. (1985) *Biochim. Biophys. Acta 825*, 111−139.

Bollum, F. J. (1966) in *Procedures in Nucleic Acid Research* (Cantoni, G. L., & Davies, D. R., Eds.) p 577, Harper & Row, New York.

Brayer, G. D., & McPherson, A. (1984) *Biochemistry 23*, 340−349.

Bulsink, H., Harmsen, B. J. M., & Hilbers, C. W. (1985) *J. Biomol. Struct. Dyn. 3*, 227−247.

Bulsink, H., Wijnaendts van Resandt, R. W., Harmsen, B. J. M., & Hilbers, C. W. (1986) *Eur. J. Biochem. 157*, 329−334.

Bulsink, H., Harmsen, B. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 176*, 597−608.

Cavalieri, S. J., Neet, K. E., & Goldthwait, D. A. (1976) *J. Mol. Biol. 102*, 697−711.

Chase, J. W., & Williams, K. R. (1986) *Annu. Rev. Biochem. 55*, 103−136.

Corbett, R. J. T., & Roche, R. S. (1984) *Biochemistry 23*, 1888−1894.

Day, L. A. (1973) *Biochemistry 13*, 5329−5339.

de Jong, E. A. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1987) *Biochemistry 26*, 2039−2046.

de Jong, E. A. M., van Duynhoven, J. P. M., Harmsen, B. J. M., Tesser, G. I., Konings, R. N. H., & Hilbers, C. W. (1989) *J. Mol. Biol. 206*, 133−152.

Epstein, I. R. (1978) *Biophys. Chem. 8*, 327−339.

Folkers, P. J. M., Stassen, A. P. M., van Duynhoven, J. P. M., Harmsen, B. J. M., Konings, R. N. H., & Hilbers, C. W. (1991) *Eur. J. Biochem. 200*, 139−148.

Folkers, P. J. M., Nilges, M., Folmer, R. H. A., Konings, R. N. H., & Hilbers, C. W. (1994) *J. Mol. Biol. 236*, 229−246.

Folmer, R. H., Nilges, M., Folkers, P. J., Konings, R. N., & Hilbers, C. W. (1994) *J. Mol. Biol. 240*, 341−357.

Fulford, W., & Model, P. (1988) *J. Mol. Biol. 203*, 39−48.

Gray, C. W. (1989) *J. Mol. Biol. 208*, 57−64.

Gray, D. M., Gray, C. W., & Carlson, R. D. (1982) *Biochemistry 21*, 2707−2713.

Guan, Y., Zhang, H., Konings, R. N. H., Hilbers, C. W., Terwilliger, T. C. & Wang, A. H.-J. (1993) *Biochemistry 33*, 7768−7778.

Guan, Y., Zhang, H., & Wang, A. H. (1995) *Protein Sci. 4*, 187−197.

Ito, K., Date, T., & Wickner, W. (1980) *J. Biol. Chem. 255*, 2123−2130.

Kansy, J. W., Clack, B. A., & Gray, D. M. (1986) *J. Biomol. Struct. Dyn. 3*, 1079−1109.

King, G. C., & Coleman, J. E. (1987) *Biochemistry 26*, 2929−2937.

King, G. C., & Coleman, J. E. (1988) *Biochemistry 27*, 6947−6953.

Kowalczykowski, S. C., Bear, D. G., & von Hippel, P. H. (1981a) in *The Enzymes* (Boyer, P. D., Ed.) pp 373−444, Academic Press, New York.

Kowalczykowski, S. C., Lonberg, N., Newport, J. W., & von Hippel, P. H. (1981b) *J. Mol. Biol. 145*, 75−104.

Kowalczykowski, S. C., Paul, L. S., Lonberg, N., Newport, J. W., McSwiggin, J. A., & von Hippel, P. H. (1986) *Biochemistry 25*, 1226−1240.

Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.

Liang, H., & Terwilliger, T. C. (1991) *Biochemistry 30*, 2772−2782.

Lohman, T. M., & Bujalowski, W. (1991) *Methods Enzymol. 208*, 258−290.

Lohman, T. M., & Mascotti, D. P. (1992) *Methods Enzymol. 212*, 424−458.

Lohman, T. M., & Ferrari, M. E. (1994) *Annu. Rev. Biochem. 63*, 527−570.

Lonberg, N., Kowalczykowski, S. C., Paul, L. S., & von Hippel, P. H. (1981) *J. Mol. Biol. 145*, 123−138.

McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469−489.

Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Newport, J. W., Lonberg, N., Kowalczykowski, S. C., & von Hippel, P. H. (1981) *J. Mol. Biol. 145*, 105−121.

Oey, J. L., & Knippers, R. (1972) *J. Mol. Biol. 68*, 125−138.

Porschke, D., & Rauh, H. (1983) *Biochemistry 22*, 4737−4745.

Pratt, D., & Ehrdahl, W. S. (1968) *J. Mol. Biol. 37*, 181−200.

Pretorius, H. T., Klein, M., & Day, L. A. (1975) *J. Biol. Chem. 250*, 9262−9269.

Record, M. T., Lohman, T. M., & de Hasteth, P. (1976) *J. Mol. Biol. 107*, 145−158.

Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103−178.

Record, M. T., Jr., Mazur, S. J., Melancon, P., Roe, J.-H., Shaner, S. L., & Unger, L. (1981) *Annu. Rev. Biochem. 50*, 997−1024.
Salstrom, J. S., & Pratt, D. (1971) *J. Mol. Biol. 61*, 489−501.
Skinner, M. M., Zhang, H., Leschnitzer, D. H., Bellamy, H., Sweet, R. M., Gray, C. W., Konings, R. N. H., Wang, A. H.-J., & Terwilliger, T. C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2071−2075.
Stassen, A. P. M., Harmsen, B. J. M., Schoenmakers, J. G. G., Hilbers, C. W., & Konings, R. N. H. (1992a) *Eur. J. Biochem. 206*, 605−612.
Stassen, A. P. M., Zaman, G. J. R., van Deursen, J. M. A.; Schoenmakers, J. G. G., & Konings, R. N. H. (1992b) *Eur. J. Biochem. 204*, 1003−1014.
Terwilliger, T. C. (1988) *Gene 71*, 41−47.
van Duynhoven, J. (1992) NMR studies of the single-stranded DNA binding proteins encoded by the filamentous bacteriophages M13 and IKe, Ph.D. Thesis, University of Nijmegen.
Villemain, J. L., & Giedroc, D. P. (1993) *Biochemistry 32*, 11235−11246.
Zabin, H. B. (1991) Isolation and in vitro characterization of mutants of the Escherichia coli bacteriophage f1 gene V protein, Ph.D. Thesis, The University of Chicago.
Zabin, H. B., & Terwilliger, T. C. (1991) *J. Mol. Biol. 219*, 257−275.
Zhang, W., Bond, J. B., Anderson, C. F., Lohman, T. M., & Record, M. T., Jr. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2511−2516.

BI961050C