Falch, E. (1970), *Acta Chem. Scand. 24,* 137.
Gerchman, L. L., and Ludlum, D. B. (1973), *Biochim. Biophys. Acta 308,* 310.
Kramer, B. S., Fenselau, C. C., and Ludlum, D. B. (1974), *Biochem. Biophys. Res. Commun. 56,* 783.
Lawley, P. D. (1972), *Top. Chem. Carcinog. Proc. Int. Symp. 2nd,* 237.
Lawley, P. D., and Brookes, P., (1963), *Biochem. J. 89,* 127.
Liehr, J. G., Von Miden, D. L., Mattox, S. E., and McCloskey, J. A. (1974), *Biomed. Mass Spectrom. 1,* 281.
Ludlum, D. B. (1966), *Mol. Pharmacol. 2,* 585.
Ludlum, D. B. (1970a), *Biochim. Biophys. Acta 213,* 142.
Ludlum, D. B. (1970b), *J. Biol. Chem. 245,* 477.
Ludlum, D. B. (1971), *Biochim. Biophys. Acta 247,* 412.
Ludlum, D. B., and Wilhelm, R. C. (1968), *J. Biol. Chem. 243,* 2750.
McCloskey, J. A. (1974), in Basic Principles in Nucleic Acid Chemistry, Vol. 1, Ts'o, P. O. P., Ed., New York, N.Y., Academic Press, Chapter 3.
Ohashi, Z., Maeda, M., McCloskey, J. A., and Nishmura, S. (1974), *Biochemistry 13,* 2620.
Singer, B., and Fraenkel-Conrat, H. (1970), *Biochemistry 9,* 3694.
Ueda, T., and Fox, J. J. (1963), *J. Am. Chem. Soc. 85,* 4024.
Von Minden, D. L., Stillwell, R. N., Koenig, W. A., Lyman, K. J., and McCloskey, J. A. (1972), *Anal. Biochem. 50,* 110.
Wheeler, G. P., and Chumley, S. (1967), *J. Med. Chem. 10,* 259.

Physicochemical Properties of DNA Binding Proteins: Gene 32 Protein of T4 and *Escherichia coli* Unwinding Protein†

Richard A. Anderson‡ and Joseph E. Coleman*

ABSTRACT: The single-stranded DNA binding protein coded for by gene 32 of bacteriophage T4 and a similar protein isolated from uninfected *Escherichia coli* both induce characteristic changes in the circular dichroism (CD) of single-stranded nucleic acids. These CD changes have been adapted as an assay of protein–DNA complex formation. Far-ultraviolet CD spectra show the secondary structure of the two proteins to be similar with ~20% α helix, ~20% β structure, and 60% random coil. Both proteins show prominent Cotton effects arising from their aromatic chromophores. Nitration of five of the nine tyrosyl residues of gene 32 protein prevents DNA binding, while prior formation of the DNA complex protects all tyrosyl residues from nitration. The tyrosyl residues may participate in gene 32 protein–DNA binding by intercalation between bases of the single strand. In contrast, no tyrosyl residues can be nitrated in the *E. coli* protein suggesting that surface tyrosyls do not play a part in binding of the *E. coli* protein to DNA. Approximately 50 amino acids can be cleaved from the gene 32 protein with trypsin. This cleavage also occurs spontaneously in infected cell extracts. The remaining protein of mol wt 30000 has the same CD spectra and DNA binding properties as the native protein. The physicochemical properties can be correlated with previous work on the structures and functions of the group of DNA "unwinding proteins".

Since the introduction of DNA-cellulose affinity chromatography by Alberts et al. (1968), a number of proteins exhibiting specific affinity for single- or double-stranded nucleic acids have been isolated from viral, bacterial, and eukaryotic sources. One class of DNA-binding proteins has become known as "unwinding" proteins because of their tight, cooperative, and preferential binding to single-stranded nucleic acids without regard to base composition. This results in a decrease in the melting temperature of double-stranded nucleotide polymers mixed with the purified proteins. A number of proteins in this class, including the gene 5 protein coded for by bacteriophage fd (Alberts et al., 1972), the gene 32 protein coded for by bacteriophage T4 (Alberts and Frey, 1970), and the "unwinding" protein found in uninfected *Escherichia coli* (Molineaux et al., 1974; Weiner et al., 1975) can be isolated in a homogeneous form. The presence of many copies (800–100000) of these proteins per cell as well as the in vitro formation of complexes between these proteins and single-stranded DNA suggests a stoichiometric as opposed to catalytic role for these proteins in the cellular processing of nucleic acids (Mazur and Model, 1973; Sinha and Snustad, 1971; Geider and Kornberg, 1974).

The gene 32 protein has been shown to be necessary for both the replication and recombination of the T4 genome in vivo (Tomizawa et al., 1966). In vitro it facilitates renaturation of native DNA (Alberts and Frey, 1970) and specifically enhances the catalytic rate of purified T4 DNA polymerase (Huberman et al., 1971). The *E. coli* unwinding protein has been found to be a necessary component of in

† From the Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510. *Received June 30, 1975.* This work was supported by Grant GM 21919-01 from the National Institutes of Health and Grant GB-13344 from the National Science Foundation. This work forms part of a dissertation submitted by R.A.A. to the Graduate School of Yale University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

‡ Medical Scientist Training Program Trainee under Grant GM 02044 from the National Institutes of Health.

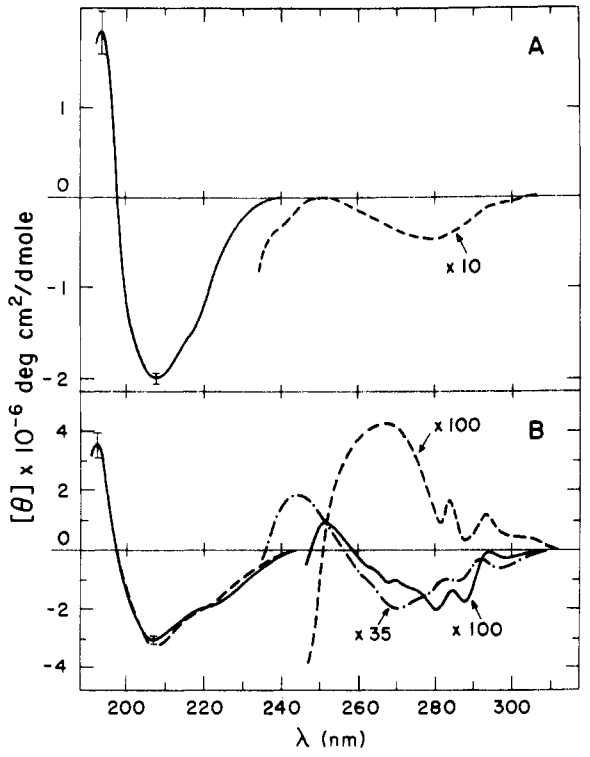

FIGURE 1: Circular dichroism of DNA binding proteins. (A) *E. coli* DNA unwinding protein (1.6×10^{-5} *M*) in 0.01 *M* Tris-HCl, 1×10^{-4} *M* dithiothreitol, 0.015 *M* $NaClO_4$, 10% glycerol, pH 8, 25°. (B) Gene 32 protein in 0.01 *M* Tris-HCl, pH 8, 25°. (—) Native protein (2.8×10^{-6} *M*; (- - - -) nitrated monomers (2.0×10^{-6} *M*); (- · -) human carbonic anhydrase C.

vitro DNA synthesizing systems using single-stranded DNA as templates and employing *E. coli* DNA polymerase III* (Geider and Kornberg, 1974). The unwinding protein has apparent roles in limiting and directing initiation as well as in stimulation of the rate of synthesis (Geider and Kornberg, 1974). The *E. coli* protein has also been reported to enhance the action of *E. coli* DNA polymerase II with single-stranded DNA templates (Molineaux et al., 1974). One can speculate that these proteins act by maintaining the template in a conformation which can be recognized and utilized by the catalytic proteins. In contrast the gene 5 protein of fd has been shown to function in keeping newly synthesized viral DNA single stranded until it can be packaged in the phage coat (Mazur and Model, 1973).

Investigations in this laboratory of the structure of the complex formed by the gene 5 protein of bacteriophage fd with single-stranded nucleic acids have demonstrated the utility of circular dichroism as an assay for complex formation and the power of chemical modification techniques to implicate specific amino acid residues in the binding mechanism (Anderson et al., 1975). Extension of these techniques to other single-stranded DNA-binding proteins, the gene 32 protein from phage T4 infected *E. coli* and the *E. coli* DNA-unwinding protein, reveals similarities and differences in physicochemical properties of their DNA complexes which may correlate with previously observed variations in complex formation and function of these three proteins.

Materials and Methods

Gene 32 protein was prepared from *E. coli* BW46 (xon A^-, su^-) (Kushner et al., 1972) infected with a T4 bacteriophage carrying amber mutations in genes 33 and 55 (kindly supplied by Dr. E. Niles) which results in the overproduction of the gene 32 protein (Gold et al., 1973). The mutant T4 phage was prepared in *E. coli* KL 239 (su_I^+, pro^-, met^-, trp^-, his^-) kindly provided by Dr. B. Low. The protein was purified by the method of Alberts and Frey (1970) modified by a final purification using phosphocellulose chromatography as per Wackernagel and Radding (1974). The DNA unwinding protein was prepared from *E. coli* BW46 using the methods described by Molineaux et al. (1974). Isolation of intact, homogeneous fd DNA was carried out as described in Anderson et al. (1975).

Amino acid analyses were performed by the standard methods of Spackman et al. (1958), determining half-cystine as cysteic acid according to Moore (1963), using a Beckman 121 M amino acid analyzer. Amino terminal residue determinations were made using the procedures of Gray (1972). Nitration of these proteins with tetranitromethane was carried out as previously described by Sokolovsky et al. (1966).

Slab gel electrophoresis (10% polyacrylamide with 0.1% sodium dodecyl sulfate) was performed using the system of Weber and Osborn (1969) modified by the addition of 5 *M* urea to the gels. Circular dichroism (CD) and its adaption to measuring DNA–protein binding were as described in Anderson et al. (1975). CD is expressed in terms of molar ellipticity, $[\theta] = 2.303(4500/\pi)(\epsilon_L - \epsilon_R)$ in units of deg cm^2/dmol. Magnetic circular dichroism was measured with a Cary 61 spectropolarimeter equipped with a superconducting magnet producing a field of 44.7 kG in the sample chamber. Tryptophan content was calculated as described by McFarland and Coleman (1972).

Poly[d(A-T)] and tetranitromethane were obtained from Sigma Chemical Co. Trypsin was purchased from Worthington Biochemical Corp. All other chemicals were reagent grade.

Results and Discussion

Protein Structures. The circular dichroism from 320 to 190 nm of the *E. coli* unwinding protein and gene 32 protein are shown in Figure 1. The general conformation of the polypeptide backbone must be similar in both proteins, since the far-ultraviolet CD spectra of the two proteins show similar band positions and mean residue ellipticities. Both have prominent minima at 208 nm and maxima of approximately equal amplitude at 192–194 nm. Secondary minima near 222 nm are less prominent. Mean residue ellipticities at representative wavelengths are nearly identical for the two proteins: $[\theta_m]_{208} = -9900$, $[\theta_m]_{217} = -7650$, $[\theta_m]_{222} = -6050$ for the gene-32 protein; and $[\theta_m]_{208} = -9700$, $[\theta_m]_{217} = -7300$, $[\theta_m]_{222} = -4850$ for the *E. coli* unwinding protein. Application of the graphical method of Greenfield and Fasman (1969) to fit these CD spectra between 208 and 240 nm with contributions from a combination of α, β, and random polypeptide structures gives the most adequate fit for the spectra of both proteins with the following approximate percentages of the three conformations considered: 20% α, 20% β, and 60% random coil.

Using a molecular weight of 22000 for the *E. coli* protein (Sigal et al. 1972, confirmed in this laboratory, by gel electrophoresis, not shown) and 35000 for the gene 32 protein (Figure 4 below) combined with the total amino acid compositions, it can be estimated that these proteins contain approximately 207 and 314 residues, respectively (Table I). The gene 32 protein contains 35 lysyl residues as well as a

Table I: Amino Acid Composition of Gene 32 Protein and *E. coli* "Unwinding" Protein.

Amino Acid	Number of Residues						
	Gene 32 Protein				AA Residue Difference from Native	*E. coli* "Unwinding" Protein	
	Native	Nitrated	fd DNA "Protected"	Trypsin Treated		Native	Nitrated
Tyr	8.6	3.6	8.3	9.4	0	5.2	4.4
NO_2-Tyr		5.1[b]	0.2				<0.1
Asp	55.2	52.8	49.1	39.5	−15	19.0	
Thr	12.6	12.5	13.5	10.4	−2	15.7	
Ser	21.3	21.4	22.4	19.5	−2	13.9	
Glu	28.6	29.0	31.0	29.5	0	30.6	
Pro	9.6	10.0	10.0	8.9	−1	11.8	
Gly	19.5	19.3	20.3	18.9	0	28.8	
Ala	26.7	22.1	26.5	14.8	−12	16.4	
1/2-Cys	4.1	[d]		4.0	0	1.0	
Val	20.7	21.1	19.9	19.1	−2	14.4	
Met	10.0[c]		8.9	5.5[c]	−4	5.8	
Ile	10.0	10.0	10.7	11.1	(+1)	6.2	
Leu	21.0	21.0	18.6	15.7	−5	10.6	
Phe	19.1	16.8	19.2	16.8	−2	4.6	
His	2.8	3.7	2.9	3.4	0	3.1	
Lys	35.2	33.0	35.2	30.6	−4	7.2	
Arg	4.1	3.7	4.4	3.6	0	9.5	
Trp[a]	5.0	(5)[d]		(5)	(0)	3.1	
Total	314			266	−48	207[e]	

[a] Determined using MCD. [b] Determined both spectrally and by amino acid analysis. [c] Amino terminal residue is methionine in both the native and the trypsin-treated protein. [d] The relative magnitudes and contours of the MCD bands corresponding to the tryptophanyl absorption bands at 293 and 286 nm are unchanged in the nitrated protein suggesting that no modification of tryptophanyl residues has occurred. Hydrolysis of the modified protein yields no additional cysteic acid suggesting that extensive oxidation of cysteine is not induced by tetranitromethane. [e] The value of 207 residues is based on a molecular weight of 22000 based on a series of calibrated sodium dodecyl sulfate-acrylamide gels in which the *E. coli* protein migrates consistently very near to a trypsin standard (mol wt 23000), in agreement with the value of the molecular weight published by Sigal et al. (1972). We have not observed migration suggesting a molecular weight as low as 19000 reported by Weiner et al. (1975).

relative abundance of aromatic residues. The tryptophan content as determined by magnetic circular dichroism is three residues per molecule for the *E. coli* protein and five residues per molecule for the gene 32 protein (Table I). Basic and aromatic residues have been implicated in DNA binding interactions using both model systems (Dimicoli and Helene, 1974; Brun et al., 1975) and intact proteins (Anderson et al., 1975; Fanning, 1975).

Optical activity of the aromatic chromophores makes a large contribution to the near-ultraviolet circular dichroism of both proteins (Figure 1). The magnitude of these Cotton effects is particularly striking in the *E. coli* unwinding protein with a negative band ($[\theta] = \sim -5 \times 10^4$ deg cm^2/dmol) centered at 280 nm. The aromatic Cotton effects of the gene 32 protein are similar to those observed previously in carbonic anhydrase C (Figure 1) (Coleman, 1968). From the extensive solution data and the crystal structure of the latter enzyme these Cotton effects appear to arise from an aromatic cluster near the N-terminus of the molecule containing two tryptophanyl and one tyrosyl residue stacked with several prolinyl and phenylalanyl residues (Lindskog et al., 1971).

CD Assay of the Binding of Gene 32 and E. coli Protein to DNA. As previously demonstrated for gene 5 protein (Anderson et al., 1975), the change in CD of fd DNA or poly[d(A-T)] above 250 nm induced by DNA–protein complex formation can be used to assay the formation of complexes of both gene 32 and the *E. coli* protein with the nucleic acids. Representative CD titrations of fd DNA with gene 32 protein and the *E. coli* protein are shown in Figure 2A and B. From these titration curves the protein to base ratio at saturation of the DNA can be determined for both the *E. coli* and gene 32 proteins by plotting the change in ellipticity, $\Delta[\theta]$, vs. the protein to nucleotide base ratio, $1/R$, (Figure 2, insets).

The titration with gene 32 protein is complete at a $1/R$ value of 0.09 indicating that each protein covers ~11 bases. This is close to the figure of 10 bases per protein reported from optical absorption measurements (Alberts and Frey, 1970). With the *E. coli* protein saturation is reached at $1/R = 0.07$ indicating that each protein covers approximately 14 bases, a number significantly different from the previously reported value of eight determined from optical spectroscopy (Sigal, 1972). However, the higher number approaches the stoichiometry observed using filter binding techniques (Weiner et al., 1975).[1] A significant feature of these CD titrations is that the contribution of the DNA to the CD of the complex (i.e., the portion of the final spectra above ~250 nm) is approximately the same (when corrected for a small contribution from the protein) whether the *E. coli*, gene 32, or gene 5 (see previous data, Anderson et al., 1975) proteins are used to form the complex. The same holds for the complexes with poly[d(A-T)]. Thus the final CD of the DNA does not appear particularly sensitive to subtle differences in the conformation of the DNA, differences which clearly must exist between the three complexes judging from the length measurements and the gross fea-

[1] In contrast to the other DNA binding proteins, the base/protein ratio determined for the *E. coli* protein-DNA complex by several methods appears to give a significant range of values, 8–14. This may reflect uncertainty in the method of determining protein concentration or possibly variation in the nature of the complex under different conditions (this complex does appear by electron microscopy to be the most highly supercoiled).

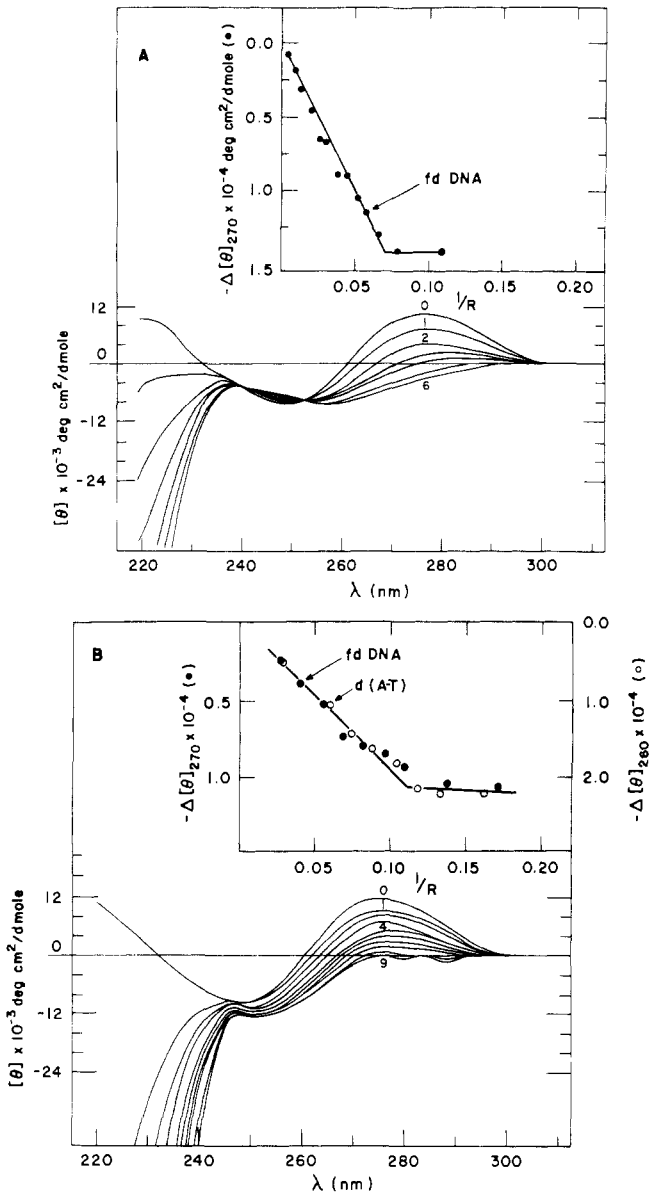

FIGURE 2: Circular dichroism spectra of fd DNA showing changes induced by the addition of DNA binding proteins. fd DNA (2.5×10^{-5} *M*) was initially present in 2.4 ml under conditions of Figure 1. (A) Changes in CD on addition of 50-μl aliquots of *E. coli* unwinding protein (1.6×10^{-5} *M*) in the same buffer to yield a final protein concentration of 1.8×10^{-6} *M* are indicated by curves 1–6. For curve 6, protein/DNA base molar ratio = 1:12.7. Insert: Change in DNA ellipticity (corrected for slight amount of ellipticity due to protein) as a function of molar ratio of *E. coli* unwinding protein/DNA bases, $1/R$; (●) fd DNA (270 nm). (B) Changes in CD on addition of 10- and 20-μl aliquots of gene 32 protein (8.4×10^{-5} *M*) in the same buffer to yield final protein concentration of 4.1×10^{-6} *M* are indicated by curves 1–9. For curve 9, the protein/DNA base molar ratio = 1:5.8. Insert: Change in DNA ellipticity as a function of molar ratio of gene 32 protein/DNA bases; $1/R$; (●) fd DNA (270 nm); (○) poly[d(A-T)] (260 nm).

tures of electron micrograph images of the complexes (see below). The ellipticity changes in the chromophores of the DNA on binding to the proteins may result primarily from dehydration of the DNA and complexation of the phosphates with positively charged groups on the protein. Dehydration and electrostatic bonding of the phosphates are known from model systems to cause similar changes in the CD of polynucleotides (Wolf and Hanlon, 1975).

There is no change in the far-ultraviolet CD of the proteins upon binding to DNA, suggesting that no gross changes in protein secondary structure are induced. The concentration of proteins and nucleic acids used and the sharp end point of the titrations (Figure 2) show that the dissociation constants for the complexes of both proteins with fd DNA must be less than 3×10^{-8} *M* at 0.015 *M* Na^+.

Gene 32 protein and the *E. coli* unwinding protein both bind tightly to the synthetic ribonucleotide polymer, poly(A), as indicated by the large reversible changes induced in the CD spectrum of the polymer by addition of the proteins. The CD changes on complex formation with poly(A) are similar to those reported previously for the gene 5 protein (Anderson et al., 1975).

Dissociation of DNA- and RNA-Protein Complexes by Cations. The reversal of the CD spectra of the DNA in the complexes to that of free DNA on dissociation of the binding protein provides a precise method for determining the effect of various cations on the stability of the protein-DNA complexes as well as demonstrating that the induced CD changes are a result of a specific protein-DNA interaction (Figure 3A and B). In general the reversal of the CD of the complexes by increasing salt concentrations follows the concentration dependence predicted from the elution profiles seen during DNA-cellulose chromatography (Alberts and Frey, 1970; Molineaux et al., 1974; Weiner et al., 1975). As salt concentration increases, the *E. coli* protein complex begins to dissociate earlier than the gene 32 protein complex, but dissociation is spread over a much wider concentration range (Figure 3A). The apparent wider salt concentration range required to bring about the complete dissociation of the *E. coli* protein-DNA complex compared to the gene 32 protein-DNA complex does not seem to be explained by a very slow dissociation rate for the complex. The observed CD spectrum of the mixture at intermediate salt concentrations does not change upon extended incubation at room temperature. Striking features of these dissociation profiles include the greater degree of dissociation caused by low concentrations of divalent metal ions as opposed to monovalent metal ions, especially with poly[d(A-T)]. The latter finding most probably reflects the additional stabilization of the double-stranded conformation of DNA by Me^{2+}.

The interaction of the *E. coli* unwinding protein with poly[d(A-T)] is completely abolished at less than 0.01 *M* Mg^{2+} (Figure 3A), while nearly 0.1 *M* is required to completely dissociate the gene 32 protein-poly[d(A-T)] complex (Figure 3B). These differences must reflect structural differences in the respective complexes. Electron microscopy suggests a different form of gross supercoiling for the complexes with the two proteins. The *E. coli* protein-fd DNA complex appears much less compact, more highly supercoiled (Sigal et al., 1972). These structural differences may determine how the DNA interacts with other proteins and enzymes in the cell. The gene 32 protein-poly(A) interaction is less stable than that with fd DNA, but it still requires ~0.4 *M* Na^+ to completely dissociate the complex (Figure 3B). The potential of these proteins to bind to cellular RNA as well as DNA may be of importance in considerations of their physiological functions in the cell.

Tyrosyl Residues Involved in DNA Binding. We have previously shown, based on CD and ^{19}F NMR data, that three of the five tyrosyl residues of the bacteriophage fd gene 5 DNA-binding protein are involved in its interaction with nucleic acids (Anderson et al., 1975). These tyrosyl

residues occupy a structure, probably some form of stacking, that markedly enhances the optical activity of their chromophores. Nitration of these tyrosyl residues by reaction with tetranitromethane hinders binding to the DNA and prior complexation of the protein with DNA prevents nitration of the three surface tyrosyl residues (Anderson et al., 1975). Likewise five of the nine tyrosyl residues in gene 32 protein can be nitrated without modifying other residues (Table I). Conditions of the nitration reaction, a 100-fold molar excess of reagent reacted for 30 min at room temperature, were the same as those used for the gene 5 protein modification. Their nitration radically alters the near-ultraviolet CD of the aromatic residues (Figure 1) and destroys binding of the protein to DNA as assayed by the CD binding assay. Nitration does not significantly alter the secondary structure of the polypeptide, since the far-ultraviolet CD is unchanged after the reaction (Figure 1). The alteration in the CD of the aromatic chromophores induced by nitration may represent some change in their conformation, but it may reflect a contribution from the intrinsic optical activity of the nitrotyrosyl chromophore which has additional strong absorption bands in the 270–280-nm region (Sokolovsky et al., 1966). Prior complexation of gene 32 protein with fd DNA prevents the nitration (Table I) and prevents the change in the CD of the aromatic chromophores. Thus the native conformation and electronic structure of five of the nine tyrosyl residues of gene 32 protein are required for DNA binding. In contrast to both gene 32 protein studied here and gene 5 protein of fd reported earlier, none of the tyrosyl residues in the *E. coli* DNA unwinding protein are available for nitration (Table I).

In control experiments (not shown) it can be demonstrated that tetranitromethane does not react with fd DNA and that the nitration of proteins which do not specifically bind to DNA, such as bovine serum albumin, is not altered by the presence of DNA.

Electron micrographs of the complexes of all three proteins with fd DNA have revealed major differences in the nature of the complexes of the three different DNA binding proteins with the same DNA. Gene 5 protein collapses circular fd DNA into a rod containing two oppositely directed DNA strands (Pratt et al., 1974). Length measurements show a linear length of complex corresponding to ~3.8 Å per base.[2] The detailed electron micrographs suggest considerable supercoiling of the DNA in the complex. Thus the actual distance between bases along the DNA strand itself must be considerably longer, perhaps as much as 6 Å between bases. The gene 32 protein–fd DNA complex leaves the circle open, but length measurements indicate a linear length of the complex of 5.3 Å/base, considerably greater than fd DNA alone (Delius et al., 1972). Again supercoiling of the DNA in the complex suggests a greater actual distance between bases along the extended DNA chain.

On the other hand, the *E. coli* protein contracts the circumference of the circular complex with fd DNA to 2.1 Å/base (Sigal et al., 1972). While supercoiling also appears to be present in this complex, the distance between bases along the DNA strand itself is probably not as great in this complex as in the other two, Intercalation of aromatic residues between bases in single-stranded DNA could account for some of the expansion of the DNA length observed in

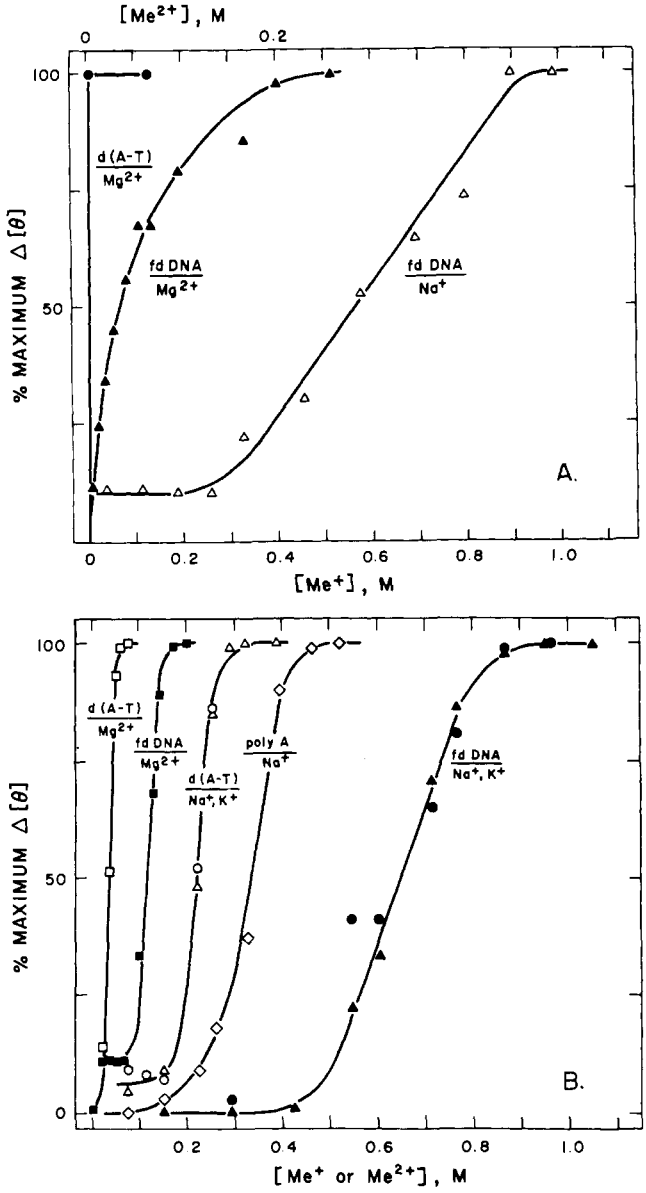

FIGURE 3: Dissociation of DNA–protein and RNA–protein complexes by cations as followed by circular dichroism. Saturating amounts of protein were added to solutions of nucleic acid (2.5×10^{-6} *M*) initially under salt concentrations and temperature as in Figure 1, mixed, and a CD spectrum taken. Aliquots of concentrated (4 *M*) solutions of the chloride salts of the cations of interest were added thoroughly mixed for 10 min, and a CD spectrum was taken. This process was repeated until no further change in the CD spectra could be noted. At the end point, the CD spectrum coincided with the spectrum of free nucleic acid at similar salt concentrations in each case. The percent maximum ellipticity change (at 260 nm for poly[d(A-T)], at 270 nm for fd DNA, and at 265 nm for poly(A) complexes) is plotted as a function of the molarity of added cation. (A) *E. coli* unwinding protein: (●) Mg^{2+}-polyd[(A-T)]; (▲) Mg^{2+}-fdDNA; (△) Na^+-fd DNA. (B) Gene 32 protein: (□) Mg^{2+}-poly[d(A-T)]; (■) Mg^{2+}-fd DNA; (◇) Na^+-poly(A); (△) K^+-poly[d(A-T)]; (○) Na^+-poly[d(A-T)]; (●) K^+-fd DNA; (▲) Na^+-fd DNA.

complexes with gene 5 and gene 32 proteins, both of which seem to have tyrosyl residues which participate in DNA binding. The *E. coli* protein apparently does not have accessible surface tyrosyl residues and would appear not to use tyrosyl intercalation as part of the binding mechanism.

Nitration of gene 32 protein using tetranitromethane results in cross-linking of the protein into dimers, trimers, and larger oligomers (Figure 4A), a reaction which also occurs

[2] Calculations of the linear length per base for the protein–fd DNA complexes are based on the recent revised data which show fd DNA to contain 5740 ± 210 nucleotides (Berkowitz and Day, 1974).

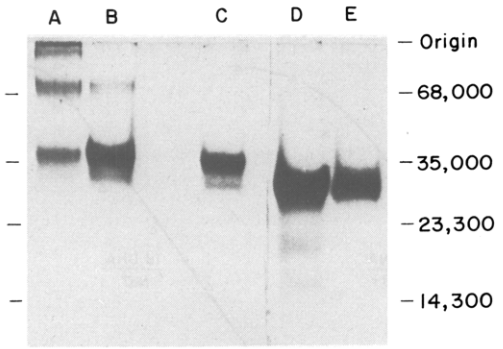

FIGURE 4: Dodecyl sulfate polyacrylamide gel (10%) showing: (A) nitrated gene 32 protein; (B) gene 32 protein bound to fd DNA before reaction with tetranitromethane; (C) purified gene 32 protein stored at 4°C in 10 m*M* Tris-HCl (pH 8) for 1 week; (D) gene 32 protein after treatment with trypsin as in text; (E) gene 32 protein bound to fd DNA before trypsin treatment and recovered in void volume with DNA after chromatography on Sephadex G-100 or after repurification by DNA-cellulose chromatography.

on nitration of gene 5 protein which we have discussed extensively before (Anderson et al., 1975). All experiments on DNA binding and physicochemical properties of the nitroprotein described here were performed on nitrated monomers separated from the other oligomers by Sephadex G-75 chromatography. Reaction of the protein with tetranitromethane in the presence of fd DNA results in minimal dimer formation (Figure 4B) as expected from the shielding of the reactive tyrosines (Table I).

Enzymatic Cleavage of Gene 32 Protein. Immediately after the final purification step, gene 32 protein migrates on polyacrylamide gels as a single band of mol wt 35000. However, throughout the earlier purification stages a smaller band consistently appears at a position indicating a mol wt of ~30000. This band reappears and increases with storage of the protein, particularly when stored prior to the final purification steps, and derives from the gene 32 protein itself. Hosoda et al. (1974) have independently found a similar phenomenon when they purified gene 32 protein from an overproducer strain of T4 similar to the one employed in these studies.

Treatment of the intact protein with trypsin (1:100 w/w for 2 hr at 25°C) converts almost all of the protein to the 30000 mol wt form with a few lower molecular weight fragments detectable (Figure 4D). This large fragment can be purified by binding to a column of DNA-cellulose; it can be eluted from the column with 0.6 *M* NaCl and the gel of the eluted protein is shown in Figure 4E. The CD of the 30000 mol wt peptide and its binding to fd DNA (by CD assay) are similar to these parameters as measured for native gene 32 protein, implying that the major structural properties of the protein are preserved after cleavage. There is, however, a significant increase in ellipticity of the positive CD band at 250 nm and a significant decrease in the magnitude of the negative band at 208 nm, as well as a red shift to 210 nm. These changes imply at least moderate conformational changes in the protein upon removal of the small peptide. Trypsin treatment of the gene 32 protein–fd DNA complex also results in the conversion of all the protein to the 30000 mol wt form. The modified protein–DNA complex migrates into the void volume of a Sephadex G-100 column (10 m*M* Tris-HCl, pH 8), again indicating that the binding function of the cleaved protein is intact.

Amino acid analysis of the 30000 mol wt peptide reveals that approximately 50 residues are lost from the native gene 32 protein. Fifteen of the cleaved residues are aspartic acid. Depending on the unknown amount of amidation present, this could represent a loss of considerable negative charge. There is also a loss of five positively charged lysyl residues, however. A majority of the cleaved residues are accounted for by two species, alanyl and aspartyl residues. Thus the cleaved region has an unusual amino acid composition and could consist of some type of repeated structure. The amino terminal residue of both the native gene 32 protein and the trypsin-cleaved protein is methionine suggesting, but not proving, that cleavage occurs at the carboxy end of gene 32 protein.

Carroll et al. (1975) have reported that the aggregation of their preparations of gene 32 protein decreases with storage at 4°C. Although they see no change in the gel electrophoretic properties of the protein accompanying the changes in self-association, such a change could relate to limited proteolysis of the molecule. The similarity between their published amino acid composition for the gene 32 protein and ours confirms that the same polypeptide is being studied (Carroll et al., 1975) (Table I). Purification of gene 32 protein from overproducer mutants of T4 may result in residual traces of trypsin-like proteolytic contamination not detectable by polyacrylamide gel electrophoresis in our preparations.

A new trypsin-like activity is induced in *E. coli* at about 15 min after infection with T4 phage (Poglazov and Levshenko, 1974). There is evidence that limited proteolysis of several of the structural proteins of the mature phage plays an important part in the maturation and assembly of T4 (Eiserling and Dickson, 1972).

The in vivo significance of the tryptic-like cleavage of gene 32 protein is not certain, but a reasonable postulate can be made. Loss of a large, highly charged fragment from the protein as well as the accompanying small conformational changes in the molecule (see above) might alter protein–protein interactions and the configuration of the DNA–protein complex recognized by enzymes such as the T4 DNA polymerase, which function in the replication and recombination of T4 DNA.

The fact that the modified gene 32 protein melts T4 DNA at physiological temperatures (Hosoda et al., 1974), while the parent molecule of 35000 mol wt does not (Alberts and Frey, 1970), suggests a physiological role for the proteolytic cleavage. A molecule participating in replication and recombination of T4 might reasonably be expected to melt the phage DNA at physiological temperatures. The failure of gene 32 to do so has been puzzling. Determination of the structure of these "unwinding" proteins and how they control the conformation of single-stranded DNA should reveal information about the basic mechanisms of DNA–protein interactions which are applicable to more sequence-specific proteins which control gene expression.

References

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.

Alberts, B. M., and Frey, L. (1970), *Nature (London) 227,* 1313–1318.

Alberts, B. M., Frey, L., and Delius, H. (1972), *J. Mol. Biol. 68,* 139–152.

Anderson, R. A., Nakashima, Y., and Coleman, J. E. (1975), *Biochemistry 14,* 907–917.

Berkowitz, S. A., and Day, L. A. (1974), *Biochemistry 13,* 4825–4831.
Brun, F., Toulme, J. J., and Helene, C. (1975), *Biochemistry 14,* 558–563.
Carroll, R. B., Neet, K., and Goldthwait, D. A. (1975), *J. Mol. Biol. 91,* 275–291.
Coleman, J. E. (1968), *J. Biol. Chem. 243,* 4574–4587.
Delius, H., Mantell, N. J., and Alberts, B. (1972), *J. Mol. Biol. 67,* 341–350.
Dimicoli, J. L., and Helene, C. (1974), *Biochemistry 13,* 724–730.
Eiserling, F. A., and Dickson, R. C. (1972), *Annu. Rev. Biochem. 41,* 467–502.
Fanning, T. G. (1975), *Biochemistry 14,* 2512–2520.
Geider, K., and Kornberg, A. (1974), *J. Biol. Chem. 249,* 3999–4005.
Gold, L. M., O'Farrell, P. Z., Singer, B., and Stormo, G. (1973), Virus Research 2nd ICN-UCLA Symposium on Molecular Biology, Fox, C. F., and Robinson, W. S., Ed., New York, N.Y., Academic Press, pp 205–225.
Gray, W. R. (1972), *Methods Enzymol. 25,* 121–138.
Greenfield, N., and Fasman, G. D. (1969), *Biochemistry 8,* 4108–4116.
Hosoda, J., Takacs, B., and Brack, C. (1974), *FEBS Lett. 47,* 338–342.
Huberman, J. A., Kornberg, A., and Alberts, B. M. (1971), *J. Mol. Biol. 62,* 39–52.
Kushner, S. R., Nagaishi, H., and Clark, A. J. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 1366–1370.
Lindskog, S., Henderson, L. E., Kannan, K. K., Liljas, A., Nyman, P. O., and Strandberg, B. (1971), *Enzymes, 3rd Ed. 5,* 587–665.
Mazur, B. J., and Model, P. (1973) *J. Mol. Biol. 78,* 285–300.
McFarland, T. M., and Coleman, J. E. (1972), *Eur. J. Biochem. 29,* 521–527.
Molineaux, I. J., Friedman, S., and Gefter, M. L. (1974), *J. Biol. Chem. 249,* 6090–6098.
Moore, S. (1963), *J. Biol. Chem. 238,* 235–237.
Poglazov, B. F., and Levshenko, M. T. (1974), *J. Mol. Biol. 84,* 463–467.
Pratt, D., Laws, P., and Griffith, J. (1974), *J. Mol. Biol. 82,* 425–439.
Sigal, N., Delius, H., Kornberg, T., Gefter, M., and Alberts, B. M. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 3537–3541.
Sinha, N. K., and Snustad, D. P. (1971), *J. Mol. Biol. 62,* 267–271.
Sokolovsky, M., Riordan, J. F., and Vallee, B. L. (1966), *Biochemistry 5,* 3582–3589.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30,* 1190–1206.
Tomizawa, J., Anraku, N., and Iwana, Y. (1966), *J. Mol. Biol. 21,* 247–253.
Wackernagel, W., and Radding, C. M. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 431–435.
Weber, K., and Osborn, M. (1969), *J. Biol. Chem. 244,* 4406–4412.
Weiner, J. H., Bertsch, L. L., and Kornberg, A. (1975), *J. Biol. Chem. 250,* 1972–1980.
Wolf, B., and Hanlon, S. (1975), *Biochemistry 14,* 1661–1670.