Chemical Modifications of Functional Residues of fd Gene 5 DNA-Binding Protein†

Richard A. Anderson,‡ Yasutsugu Nakashima, and Joseph E. Coleman*

ABSTRACT: The binding of gene 5 protein from bacteriophage fd to poly[d(A-T)], fd DNA, and poly(A) is accompanied by a dramatic reversal in the signs of the large ellipticity bands of the nucleic acid chromophores from 250 to 290 nm. The change in the circular dichroism of the DNA induced by the protein, which reaches a maximum at a protein to nucleotide molar ratio of 1:4, has been used as an assay of the alterations in binding of gene 5 protein to DNA accompanying changes in the ionic environment and subsequent to chemical modification of the protein. Divalent cations completely dissociate the gene 5 protein–fd DNA complex at 0.1 *M*, while 0.5 *M* monovalent cations are required. All cations are more effective in dissociating the complex with poly[d(A-T)] commensurate with the accompanying stabilization of the double helix to which gene 5 protein does not bind. Acetylation of all six lysyl residues and three of the five tyrosyl residues of the protein with *N*-acetylimidazole prevents complex formation. Removal of the three tyrosyl *O*-acetyl groups with hydroxylamine does not restore the binding of gene 5 protein to DNA. Tetranitromethane nitrates the same three tyrosyl residues (Tyr-26, Tyr-41, and Tyr-56 as determined by peptide mapping) and reduces the binding affinity of the protein for fd DNA by ~100-fold. The ^{19}F NMR spectrum of gene 5 protein labeled with *m*-fluorotyrosine shows three surface and two buried fluorotyrosyl residues. All tyrosyl residues are protected from nitration in the complex with fd DNA, but acetylimidazole acetylates surface lysyl residues in the complex and dissociates it. The intrinsic circular dichroism of the acetylated and nitrated gene 5 proteins is not significantly altered. In contrast maleic anhydride reacts with the seven amino groups of the protein and changes the secondary structure to one similar to that present in 6 *M* guanidine · HCl. The single SH group of the native protein does not react with Ellman's reagent, but it reacts rapidly with one Hg^{2+} ion which unfolds the protein; fd DNA prevents reaction with Hg^{2+}. Electrostatic forces may be as important as hydrogen bonding in maintaining the native structure of this protein. The lysyl groups of the protein, exposed in both the free protein and the DNA complex, appear to be of prime importance in DNA binding, probably through electrostatic interactions with the DNA phosphate groups. Three tyrosyl residues also contribute to binding affinity through hydrogen bonding or intercalation. A model of gene 5 protein structure in relation to interactions with a tetranucleotide is presented.

The production of single-stranded viral DNA by replicating F-specific filamentous bacteriophages including f1, fd, and M13 has been shown to require the participation of two phage-coded proteins, the products of phage genes 2 and 5 (see Marvin and Hohn, 1969, for review). The single-stranded DNA of the infecting phage can be converted to a single parental double-stranded replicative form (RF)[1] by host cell enzymes but the gene 2 product, probably a nuclease which specifically nicks the viral strand of the RF, is necessary for the production of about 200 progeny RF (Pratt and Erdahl, 1968; Tseng and Marvin, 1972; Mazur and Model, 1973; Fidanian and Ray, 1972). In addition to the gene 2 protein, the product of gene 5 is necessary for the conversion of DNA synthesis from net RF production to asymmetric displacement synthesis of new single-stranded viral DNA using the complementary strand of the RF as a template (Salstrom and Pratt, 1971; Mazur and Model, 1973).

In normal filamentous phage infections the gene 5 protein accumulates in about 10^5 copies per cell (Alberts et al., 1972). In vivo complexes of gene 5 protein with fd DNA have been isolated from infected cells and appear in the electron microscope as cigar-shaped rods thick enough to contain two protein-covered strands of DNA (Pratt et al., 1974). With the exception of some branching, gene 5 protein–fd DNA complexes made by mixing purified protein and DNA in vitro look similar in the electron microscope to the in vivo complexes (Alberts et al., 1972). In both complexes there is 1 mol of protein for every 4–5 mol of DNA base. The addition of gene 5 protein to an in vitro reconstituted RF synthesizing system in stoichiometric amounts halts DNA synthesis (Geider and Kornberg, 1974). The function of gene 5 protein appears to be stoichiometric involving complex formation with single-stranded DNA and thereby preventing its use as a template for synthesis of the complementary strand until the complex moves to the cell membrane where the gene 5 proteins are replaced by virion coat proteins as mature virions are extruded from the intact bacterium. Work by Staudenbauer and Hofschneider (1973) suggests an additional positive role for gene 5 protein in the synthesis of single-stranded viral DNA.

Gene 5 protein is a single polypeptide chain of mol wt 9689 containing 87 amino acid residues whose sequence has recently been determined by Nakashima et al. (1974a,b).

† From the Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510. *Received October 7, 1974.* This work was supported by Grant No. AM 09070-10 from the National Institutes of Health and Grant No. GB-13344 from the National Science Foundation. This work forms part of a dissertation submitted by R. A. Anderson 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 No. GM 02044 from the National Institutes of Health.

[1] Abbreviations used are: CD, circular dichroism; NMR, nuclear magnetic resonance; RF, replicative form; $C(NO_2)_4$, tetranitromethane.

The protein contains 7 acidic residues, 11 basic residues including 6 lysines, 1 cysteine, and 5 tyrosines, and no tryptophan. The pure protein has been shown to bind tightly and cooperatively to single-stranded DNA without regard to base sequence, but not to double-stranded DNA (Oey and Knippers, 1972; Alberts et al., 1972). The preferential binding to single-stranded DNA lowers the T_m of a variety of double-stranded DNAs and synthetic deoxyribonucleotide polymers by an average of 40°. Gene 5 protein induces the melting of poly[d(A-T)] at room temperature. Day (1973) has examined the absorption and circular dichroism (CD)[1] spectra of gene 5 protein and its complex with fd DNA and found characteristic changes in the spectra of the bound molecules vs. their uncombined states. The near-ultraviolet CD of native gene 5 protein appears to reflect primarily the ellipticity of the tyrosyl chromophores whose dissymmetry changes significantly on complex formation (Day, 1973). Large changes in the CD of double-stranded nucleic acid polymers as well as of single-stranded native DNAs accompany the formation of complexes with gene 5 protein. The present study has utilized the CD changes occurring in the polymers as well as changes occurring in the CD spectrum of the protein on complex formation to monitor the interaction of gene 5 protein with DNA as a function of cation concentration and species, pH, and chemical modification of specific amino acid side chains of the protein.

Experimental Section

Bacterial and Viral Strains. Escherichia coli K37 (Hfr, su_I^+) was kindly supplied by Dr. Max Oeschger; *E. coli* AT2471 (Hfr, tyr^-, thi^-, rel^-, λ^-) was kindly supplied by Dr. Brooks Low (Taylor and Trotter, 1967). Bacteriophage fd, wild type, was kindly supplied by Dr. Don Marvin.

Chemicals. Poly[d(A-T)], salmon sperm DNA, tetranitromethane, *N*-acetylimidazole, and maleic anhydride were obtained from Sigma Chemical Co. (St. Louis, Mo.). Poly(A) and oligo[d(pA)$_4$] were obtained from P-L Biochemicals, Inc. (Milwaukee, Wis.). Guanidine · HCl and ammonium sulfate were Schwarz/Mann (Orangeburg, N.Y.) Ultra Pure and enzyme grade, respectively. 1-[^{14}C]Acetylimidazole was obtained from Calatomic (Los Angeles, Calif.). Trypsin (TPCK), chymotrypsin (CDI), pancreatic deoxyribonuclease I (D), and carboxypeptidase A (COADFP) came from Worthington Biochemical Corp. (Freehold, N.J.). *m*-Fluoro-DL-tyrosine was purchased from Aldrich Chemical Co. (Cedar Knolls, N.J.), polyethylene glycol from Union Carbide (New York, N.Y.), and cesium chloride from Kerr-McGee Chemical Co. (West Chicago, Ill.) All other chemicals were reagent grade.

Preparation and Purification of Gene 5 Protein. Twenty liters of *E. coli* K37 was grown to a cell density of 2×10^8 cells/ml under conditions of high aeration in a medium containing 32 g/l. of Difco Bacto-tryptone, 20 g/l. of Difco yeast extract, and 5 g/l. of NaCl at 37°. The culture was inoculated with wild-type fd bacteriophage at a multiplicity of 50. After 6 hr bacteria were harvested by continuous-flow centrifugation; the yield was generally 200 g of cells. Gene 5 protein was extracted and purified using DNA-cellulose chromatography according to the procedures of Nakashima and Konigsberg.[2] In addition a final purification step involving exclusion chromatography on a Sephadex G-75 (fine) column equilibrated with 10 m*M* Tris-HCl (pH 8) was necessary to remove high molecular weight material. Purity was monitored by electrophoresis on 10% acrylamide slab gels using the procedure of Weber and Osborn (1969) modified by the addition of 5 *M* urea to the gels. Yields of gene 5 protein approached 1 mg of protein per original gram of cells, as determined by optical density at 276 nm, $E_{276}^{0.1\%} = 0.73$ (Day, 1973).

Chemical modifications of gene 5 protein were carried out as described in the literature and details will be given in the Results section. Reactions were terminated by dialysis of the reaction mixture or by gel filtration of the reaction mixture over a column of Sephadex G-25 (medium) in 10 m*M* Tris-HCl (pH 8). The modified proteins were concentrated by ultrafiltration using an Amicon UM2 membrane.

Preparation of m-Fluorotyrosyl Gene 5 Protein. E. coli strain AT 2741 were grown to early exponential phase in a synthetic medium containing low tyrosine levels (4 μg/ml): M9 media (Miller, 1972) supplemented with 20 μg/ml of the 19 other common amino acids. *m*-Fluoro-DL-tyrosine, 30 μg/ml, was added at early log phase followed 10 min later by the addition of wild-type fd at a multiplicity of 100. The culture was harvested 6 hr later and the gene 5 protein purified in the usual manner.

Purification of fd DNA. Cells were removed by centrifugation from a standard culture of *E. coli* infected with fd as described above. Poly(ethylene glycol), 2% w/v (Carbowax 6000), and NaCl, 0.5 *M*, were added to the supernatant. The precipitated phage were allowed to settle for 2–3 days at 4° (Wiseman et al., 1972). The precipitate was collected at 8000*g* and the phage extracted 3× with cold H_2O. Cesium chloride was added to a density of 1.3 g/ml and the fd were banded by ultracentrifugation in a Spinco 30 rotor at 27,000*g* for 24 hr at 4°. After dialysis to remove CsCl, the fd DNA was extracted with redistilled phenol and quantitated by the method of Marvin and Schaller (1966). Electron microscopy revealed that greater than 95% of the DNA was in the form of intact covalent circles (mol wt 2×10^6).

Amino Acid and Peptide Analyses. Amino acid analyses were performed by the method of Spackman et al. (1958) on a Beckman 120B amino acid analyzer. Samples were hydrolyzed in 6 *N* HCl for 22 hr at 110°. Half-cystine was determined as cysteic acid by the method of Moore (1963). For peptide analysis of nitrated gene 5 protein, the protein, 10 mg/ml, in 0.05 *M* NH_4HCO_3 (pH 8.5) was digested with chymotrypsin at 0.1 mg/ml for 1 hr at 25° after which another aliquot of chymotrypsin (final ratio of gene 5 protein:chymotrypsin of 50:1, w/w) was added and incubation continued for another hour. The digestion was stopped by boiling for 2 min. The peptides were initially separated using a 0.9 × 100 cm Sephadex G-50 (fine) column eluted with 0.2 *M* ammonium hydroxide. The yellow material obtained in the first peak from this elution was digested with trypsin using the same procedure as with chymotrypsin, applied to a 0.9 × 50 cm Sephadex G-50 (fine) column, and eluted with a 1:1 mixture of 0.4 *M* ammonium hydroxide–1-propanol. The material in the major peak from this chromatography was applied to Whatman 20 chromatography paper and subjected to high voltage electrophoresis in 0.1 *M* sodium bicarbonate (pH 9.2). Other peptides from the original chymotryptic digest were separated and purified by means of electrophoresis in 10% pyridine–acetate buffer at pH 6.5. Yellow spots were detected by exposing the dried paper to ammonia fumes and other spots were visualized with fluorescamine (Bohlen et al., 1973). All spots were eluted with 0.2 *M* NH_4OH. Amino-terminal residues were

[2] Y. Nakashima and W. Konigsberg, manuscript in preparation.

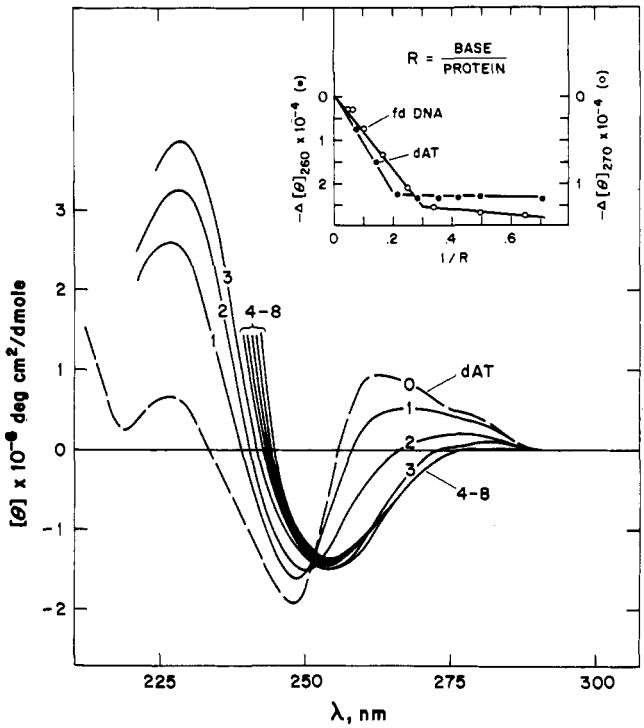

FIGURE 1: Circular dichroism of poly[d(A-T)], 3.5×10^{-5} *M*, in 2.85 ml of 10 m*M* Tris-HCl-1 m*M* Na_2EDTA (pH 8, 25°), during titration with 25-μl aliquots of gene 5 protein (5p), 2.8×10^{-4} *M*: (- - -) poly[d(A-T)] alone; (——) (1-8) poly[d(A-T)] + 5p; protein:DNA ratio = 1:14 (1), 1:7 (2), 1:5 (3), 1:3.5 (4), 1:2.8 (5), 1:2.4 (6), 1:2 (7), 1:1.4 (8). Insert: Change in DNA ellipticity is a function of the molar ratio of 5p:DNA bases: (●) poly[d(A-T)] which shows the maximum change in ellipticity at 260 nm; (○) fd DNA which shows the maximum ellipticity change at 270 nm.

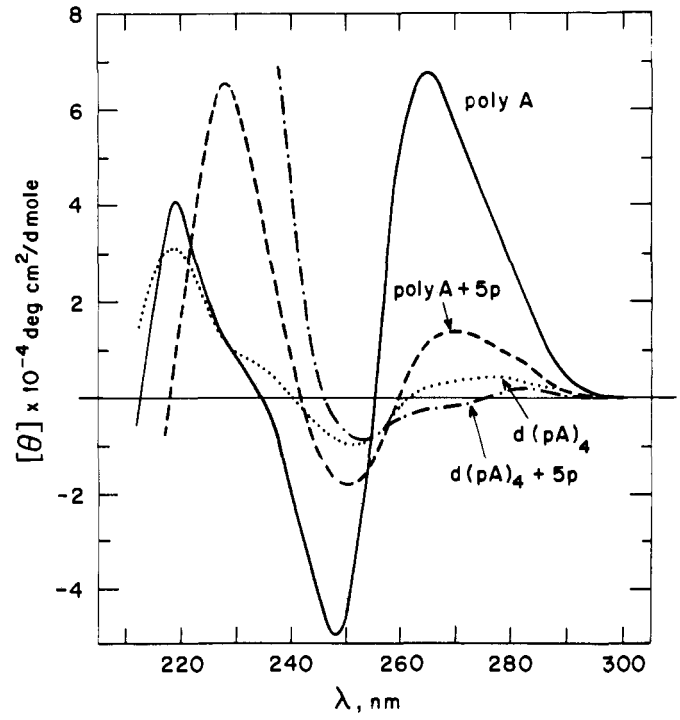

FIGURE 2: Circular dichroism changes on the binding of gene 5 protein (5p) to $d(pA)_4$ and poly(A): (——) poly(A), 2.4×10^{-5} *M*, in 2.5 ml; (- - -) same sample of poly(A) plus 50 μl of 5p, 30×10^{-5} *M*, to give a protein:base ratio of 1:4; (· · ·) $d(pA)_4$, 2.1×10^{-5} *M*, in 2.5 ml; (· - ·) $d(pA)_4$ plus 5p at a molar ratio of 3.5 protein molecules per tetranucleotide; conditions: 10 m*M* Tris-HCl-10 m*M* Na_2EDTA (pH 7, 25°).

determined by dansylation and thin-layer chromatography on polyamide plates as described by Gray (1972). Carboxyl-terminal residues were determined by digestion with carboxypeptidase A (Ambler, 1972).

Absorption spectra were measured on a Cary 15 spectrophotometer. Circular dichroism (CD) was measured on a Cary 61 spectropolarimeter. CD is expressed in terms of molecular ellipticity, $[\theta] = 2.303(4500/\pi)(\epsilon_L - \epsilon_R)$ in units of deg cm^2/dmol.

^{19}F nuclear magnetic resonance (NMR)[1] spectra were obtained on a Varian XL-100-15 NMR spectrometer operating at 94.1 MHz for ^{19}F, and locked on the ^2H resonance of the solvent (H_2O). The ^{19}F chemical shifts were determined relative to an external capillary of CF_3COOH. Data obtained by Fourier transform were processed as described by Sykes et al. (1974).

Mercury analyses were performed by atomic absorption spectroscopy (Duckworth and Coleman, 1970).

Results

CD of Gene 5 Protein-Nucleic Acid Complexes. Circular dichroism is a convenient method of monitoring changes in DNA conformation on the binding of gene 5 protein, since the major ellipticity changes in the DNA occur in the region of 260 nm and gene 5 protein shows very little ellipticity above 240 nm (Day, 1973; see below). The addition of aliquots of gene 5 protein to a constant amount of poly[d(A-T)] ($\epsilon_{260} = 6.7 \times 10^3$; Green and Mahler, 1971) shifts the positive ellipticity at 260 nm to negative values (Figure 1). A plot of the ellipticity change at 260 nm as a function of the amount of gene 5 protein present shows that no further spectral change occurs after approximately 0.25 mol of protein/mol of nucleotide has been added (Figure 1, insert). This ratio of one protein monomer to four nucleotides at the maximum ellipticity change is also observed with fd DNA (Figure 1, insert). The maximum ellipticity change occurs at 270 nm for fd DNA. The CD spectra for fd DNA (see Figure 3) are identical with those published by Day (1973). The CD assay can detect at least a 10% change in the degree of binding of gene 5 protein with little ambiguity. The concentration of protein used, 5×10^{-6} *M*, and the similar concentrations of DNA present (calculated in terms of the molarity of gene 5 protein binding sites) coupled with the lack of further change in the CD spectrum of the complex above a protein to nucleotide ratio of 1:4 (Figure 1) suggest that the dissociation constant, K_d, of the native gene 5 protein-fd DNA complex must be less than 5×10^{-8} *M*.

Significant changes in the CD of $d(pA)_4$ in the 260-nm region are induced by the addition of gene 5 protein (Figure 2). A protein to nucleotide ratio of greater than 1:1 is required to produce maximal ellipticity changes, commensurate with a much reduced affinity for the tetranucleotide. An estimate of K_d for the gene 5 protein-tetranucleotide complex pictured in Figure 2 is 3×10^{-6} *M*, at least two orders of magnitude greater than the constant estimated for the complex with intact fd DNA. The difference may be a measure of the increase in stability of the complex conferred by protein-protein interactions between adjacent molecules of gene 5 protein on the DNA. Cooperativity of binding of gene 5 protein to fd DNA has been suggested by other observations (Alberts et al., 1972). Gene 5 protein shows no demonstrable affinity for trinucleotides. However, the protein has a significant affinity for single-stranded

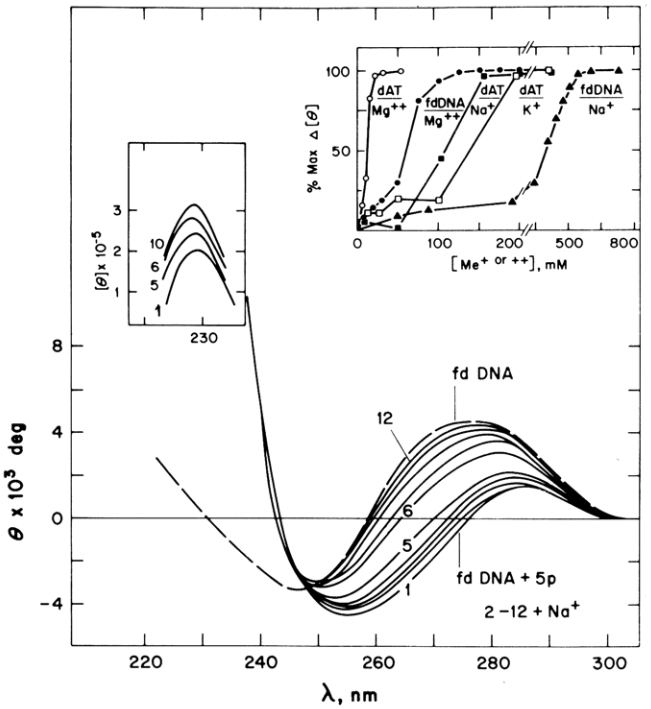

FIGURE 3: Circular dichroism changes during the titration of the gene 5 protein (5p)–fd DNA complex with NaCl: (1) no NaCl; (2) 0.05 *M*, (3) 0.10 *M*, (4) 0.19 *M*, (5) 0.28 *M*, (6) 0.38 *M*, (7) 0.42 *M*, (8) 0.46 *M*, (9) 0.51 *M*, and (10–12 superimposed) 0.55–0.80 *M* NaCl; (- - -) fd DNA alone at the same concentration as in spectrum 12, 3.9×10^{-5} *M*; conditions: molar ratio of 5p:fd DNA = 1:3; 10 m*M* tris-HCl–1 m*M* Na_2 EDTA (pH 8). Ellipticity is plotted in degrees due to the concentration changes accompanying the salt additions. The insert at the left shows the restoration of the magnitude of the ellipticity maximum of the protein at 228 nm as the complex dissociates (the numbers correspond to the spectra in the main figure). The insert at the right summarizes the salt titration data by plotting the percent of the maximum ellipticity change (at 260 nm for the 5p–poly[d(A-T)] complex and at 270 nm for the 5p–fd DNA complex) as a function of the molarity of the added cation: (○) Mg^{2+}–d(AT); (●) Mg^{2+}–fd DNA; (■) Na^+–d(AT); (□) K^+–d(AT); (▲) Na^+–fd DNA.

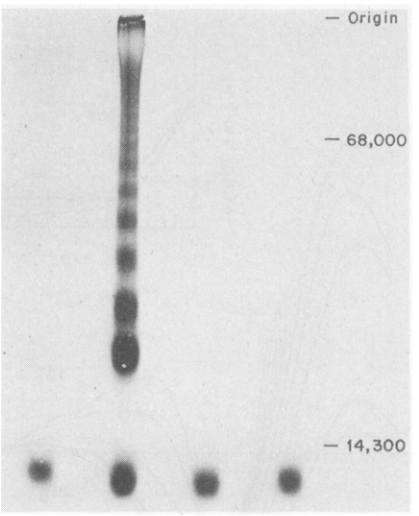

FIGURE 4: Polyacrylamide gel (10%) showing (from left to right) purified native gene 5 protein, nitrated gene 5 protein, separated monomers of nitrated gene 5 protein, and gene 5 protein protected from nitration by prior binding to fd DNA. Positions of marker proteins are indicated at the right by their molecular weights.

RNA as shown by the induction of large changes in the CD spectrum of poly(A) (Figure 2). As in the case of DNA the maximal CD change occurs at a protein to ribonucleotide molar ratio of 1:4.

Effects of Cations on Gene 5 Protein–DNA Complexes. The addition of various cations to the gene 5 protein–fd DNA complex reverses the CD of the complex to that observed for the DNA alone under the same salt concentrations, suggesting that the complex has dissociated. The complete CD spectra during the dissociation of the gene 5 protein–fd DNA complex by Na^+ are shown in Figure 3. The protein shows a prominent CD maximum at 228 nm apparently due to the dissymmetry of the tyrosyl chromophores. On complex formation with fd DNA the magnitude of this peak falls significantly (Day, 1973). On dissociation of the gene 5 protein–fd DNA complex with Na^+ the ellipticity peak of the protein at 228 nm returns to its normal magnitude (Figure 3, left inset).

Similar effects of other cations on the gene 5 protein complexes with both poly[d(A-T)] and fd DNA are shown in the right inset to Figure 3 by plotting the ellipticity change of the complex against the salt concentration. The chloride salts of K^+, Li^+, and Cs^+ dissociate the fd DNA–gene 5 protein complex at the same concentration as Na^+. Divalent cations, Mg^{2+} and Ca^{2+}, are much more effective than monovalent cations in reversing the binding of gene 5 protein to poly[d(A-T)]. This is consistent with the known effectiveness of magnesium in stabilizing the double-helical conformation.

Effect of Chemical Modification of Gene 5 Protein on the Formation of DNA Complexes. The CD data summarized above suggest that the alteration in the CD of poly[d(A-T)] or fd DNA induced by gene 5 protein appears to be a sensitive means of detecting complex formation between the protein and DNA. We have used this CD assay to monitor the effect of several chemical modifications of gene 5 protein on its ability to form complexes with DNA and the results are presented in the remainder of this paper.

Reaction of Gene 5 Protein with N-Acetylimidazole. *N*-Acetylimidazole has been shown to selectively O-acetylate tyrosyl phenolic hydroxyl groups and to N-acetylate the ϵ-amino groups of lysl residues in intact proteins (Riordan et al., 1965). Gene 5 protein, 3×10^{-4} *M*, in 0.02 *M* barbital buffer (pH 7.5) was allowed to react with a 100-fold molar excess of *N*-acetylimidazole for 30 min, 25°. Reaction for longer times resulted in precipitation of the protein. Monitoring of the absorption spectrum of the modified protein at 276 nm before (ϵ_{276} = 3500 M^{-1} cm^{-1}) and after (ϵ_{276} = 7300 M^{-1} cm^{-1}) treatment with 1 *M* NH_2OH showed that 3.3 tyrosyl residues of the protein had been O-acetylated with *N*-acetylimidazole and removed with hydroxylamine (Simpson et al., 1963). Acetylation of the protein with [^{14}C]-*N*-acetylimidazole resulted in the incorporation of 3200 cpm/100 nmol of protein. Treatment of the ^{14}C-labeled protein with 1 *M* NH_2OH removed 1140 cpm/100 nmol from the protein. The optical density increase at 276 nm accompanying the removal of the ^{14}C-acetyl groups showed that 3.4 tyrosyl residues were deacylated. This suggests that the 2069 cpm/100 nmol remaining with the protein following deacylation represent 5.96 *N*-acetyl groups per protein monomer, probably on the ϵ-amino groups of the six lysyl residues of the protein. Neither the acetylated protein nor the acetylated protein treated with hydroxylamine alter the CD spectrum of poly[d(A-T)] or fd DNA when incubated with the the polymers in large excess. Thus the dissociation constant must be increased by at

Table I: Amino Acid Analysis Data for Nitrated Gene 5 Protein and Proteolytic Fragments.

Amino Acid[a]	Expected from Sequence Data	Total Nitrated Protein	Nitrated Monomers	DNA-Protected Protein after $C(NO_2)_4$ Reaction	Peptides[b]			
					C2	C4	C4T1	C4T2
Lys	6	5.7	5.7	6.4	1.1 (1)	0.9 (1)	0.8 (1)[d]	
His	1	0.9	1.0	1.3				
Arg	4	3.8	3.9	3.1	1.7 (2)			
Asp	5	5.3	5.2	5.0		2.8 (3)	1.7 (2)	1.0 (1)
Thr	4	3.8	3.9	3.7	1.5 (2)	1.1 (1)		0.9 (1)
Ser	7	6.4	6.6	5.9	1.7 (2)			
Glu	10	9.5	9.8	9.9	1.1 (1)	3.6 (3)	1.3 (1)	2.3 (2)
Pro	6	6.5	6.9	6.0	0.8 (1)	2.8 (3)	1.3 (1)	1.8 (2)
Gly	7	7.3	7.3	6.7	2.1 (2)	2.6 (3)	0.5 (1)	2.0 (2)
Ala	4	4.0	4.1	4.0		1.5 (2)		1.9 (2)
Val	8	8.1	8.0	7.2	1.1 (1)	2.6 (3)	2.4 (3)[c]	
Met	2	2.1	1.9	1.6				
Ile	4	4.0	3.8	4.0		0.8 (1)		0.9 (1)[c]
Leu	10	9.9	10.0	10.2		3.8 (4)	1.8 (2)	2.0 (2)
Tyr	5	1.8	1.7	4.0	(1)	0.9 (3)	(1)	0.5 (1)[d]
Phe	3	3.2	3.0	2.9				
NO_2-Tyr		2.4	2.9	0.1	0.7	1.9	1.4	0.5 (1)
Recovery[e]					40% (S, E)	70% (S)	27% (S, E)	30% (S, E)

[a] Half-cystine not regularly determined. [b] Numbers in parentheses refer to expected number of residues based on Nakashima et al. (1974a,b). [c] N-Terminal residue. [d] C-Terminal residue. [e] Percent recovery after purification procedures: gel filtration on Sephadex (S); electrophoresis on paper (E).

least four orders of magnitude by acetylation of the lysyl residues of the protein.

Formation of the complex between gene 5 protein and fd DNA does not protect the protein from acetylation. Acetylation of the complex under the conditions described above, followed by passage of the product through a G-25 Sephadex column to remove the reagent, results in a CD spectrum for the products identical with that for uncomplexed fd DNA. The unmodified complex passed through the same column retains the CD spectrum of the gene 5 protein-fd DNA complex (Figure 3, curve 1). Reaction of fd DNA with *N*-acetylimidazole does not affect its ability to complex with native gene 5 protein. While both lysyl and tyrosyl residues are involved in the reaction with *N*-acetylimidazole, it may be primarily the lysyl residues that are accessible to the reagent in the complex (see below).

Reaction of Gene 5 Protein with Tetranitromethane ($C(NO_2)_4$).[1] $C(NO_2)_4$ specifically nitrates tyrosyl residues of proteins at the 3 position of the ring (Sokolovsky et al., 1966). Gene 5 protein, 8×10^{-4} *M*, in 0.05 *M* Tris-HCl–0.15 *M* NaCl (pH 8) was allowed to react for 20 min at room temperature with a 64-fold molar excess of $C(NO_2)_4$. Reaction for longer times or with a greater molar excess of $C(NO_2)_4$ leads to precipitation of the protein. Acrylamide gel electrophoresis (Figure 4) indicates that more than 50% of the nitrated gene 5 protein is present as dimers or higher oligomers, necessitating the isolation of a monomeric fraction using Sephadex G-100 gel chromatography in 0.2 *M* KCl–10 m*M* Tris-HCl (pH 8).

Characterization of Nitrated Gene 5 Protein. Amino acid analysis of nitrated gene 5 protein monomers shows the presence of three mononitrotyrosyl residues out of the total of five tyrosyl residues in the protein. Two of the tyrosyl residues are unmodified. The mole ratio of the other amino acids in the protein was not altered (Table I). No material appears in the amino acid analysis in the position of cysteic acid, the expected product of the reaction of cysteine with $C(NO_2)_4$ (Sokolovsky et al., 1966), unless the modified protein is oxidized with performic acid before hydrolysis. Using a series of amino acid analyses correlated with spectral data to determine the concentration of nitrated protein, extinction coefficients for the nitrated protein at pH 8 have been calculated: $\epsilon_{428} = 10.2 \times 10^3\ M^{-1}\ cm^{-1}$, $\epsilon_{381} = 6.2 \times 10^3\ M^{-1}\ cm^{-1}$, and $\epsilon_{276} = 16.9 \times 10^3\ M^{-1}\ cm^{-1}$. The extinction at 381 nm, the isosbestic point for the hydrogen ion titration of the hydroxyl group of 3-nitrotyrosine (Sokolovsky et al., 1966), is consistent with the presence of three nitrated tyrosyl residues per protein monomer, in agreement with the amino acid analysis. Amino acid analysis indicates that all five tyrosines are nitrated if the protein is denatured and reacted in 6 *M* guanidine · HCl.

Addition of nitrated gene 5 protein monomers to fd DNA produces little change in the 270-nm ellipticity band of fd DNA until protein to nucleotide molar ratios significantly greater than the 1:4 seen for the native protein are obtained. A 45% change was seen at a ratio of 1:1. This suggests that the nitrated monomers retain a slight affinity for DNA. The dissociation constant estimated from the change in the 270-nm ellipticity band as a function of the nitrated monomer concentration is not less than 6×10^{-6} *M*, a 100-fold increase over the maximum dissociation estimated for the native protein.

Spectrophotometric hydrogen ion titration curves of native gene 5 protein show that the molar extinction coefficient at 293 nm increases over a broad pH range from 8 to 13 (Figure 5). The final absorption indicates the presence of five ionized tyrosyl residues. Since there are no tryptophan residues in the molecule the interpretation of the spectral

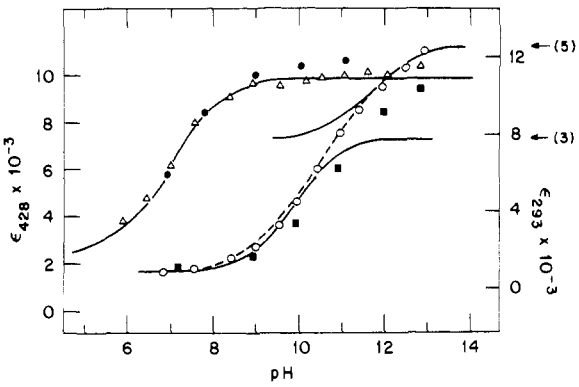

FIGURE 5: Spectrophotometric hydrogen ion titration of native gene 5 protein (5p) and nitrated gene 5 protein. Aliquots of 1 *M* NaOH (1 *M* HCl for the reverse titrations) sufficient to change the pH by 0.5 unit were added to protein samples using a glass syringe driven by a micrometer; conditions: 0.01 m*M* Tris-HCl-0.15 *M* NaCl-0.001 *M* Na_2EDTA. Above pH 11.5 and below pH 5.5 the protein solutions began to show irreversible turbidity. The data shown are corrected for concentration changes and light scattering artifacts (Beavan and Holiday, 1952): (○) native 5p, 4.5×10^{-5} *M*, titration followed at 293 nm; (■) back titration of the same sample; similar data were obtained at 283 nm; (△) nitrated 5p, 2.2×10^{-5} *M*; (●) back titration of the same sample; the titration was followed at 428 nm; (- - -) smooth curves manually drawn to fit the titration points; (——) theoretical curves drawn by assuming the three tyrosyl hydroxyl groups of native 5p have $pK_a = 9.95$ and two tyrosyl hydroxyl groups have $pK_a = 11.7$. $\Delta\epsilon_{293}$ assumed to accompany tyrosyl ionization was $2.5 \times 10^3\ M^{-1}\ cm^{-1}$.

data is quite straightforward. Taking account of a slight change in slope of the line through the points above pH 11, two sigmoid titration curves can be fitted to the data corresponding to three tyrosyl residues with apparent pK_a vales of 10.0 and two tyrosyl residues with apparent pK_a values of 11.7 (Figure 5).

Spectrophotometric hydrogen ion titration of the nitrated gene 5 protein at 428 nm indicates that three nitrotyrosyl residues ionize with apparent pK_a values of 7 (Figure 5), close to the pK_a value reported for the hydroxyl group of free nitrotyrosine (Sokolovsky et al., 1966).

Protection of Gene 5 Protein against Nitration by $C(NO_2)_4$. If gene 5 protein is complexed with excess fd DNA before reaction with $C(NO_2)_4$, no tyrosyl residues are modified as judged by the lack of change in the absorption spectrum and amino acid analysis indicating less than 0.1 modified residue per protein (Table I). The CD of the product after removal of the reagent by gel filtration is identical with that of the gene 5 protein–fd DNA complex; the CD undergoes the usual changes when the complex is dissociated with Mg^{2+} as illustrated in Figure 2. Acrylamide gel electrophoresis of the "protected" gene 5 protein shows no evidence of cross-linking (Figure 4), nor is there any apparent cross-linking of protein to DNA as the DNA separated from the reacted complex has normal absorption and CD spectra and complexes with a normal stoichiometry of native gene 5 protein. There is no reaction when $C(NO_2)_4$ is added to a solution of fd DNA, as monitored by absorption changes at 350 nm. This is in contrast to results found with calf-thymus DNA by Hugli and Stein (1971). Gene 5 protein which has been previously acetylated also shows no spectral evidence of nitration with $C(NO_2)_4$. Thus the same three tyrosyl residues appear to be modified by both acetylation with acetylimidazole and nitration with $C(NO_2)_4$.

Location of the Nitrated Tyrosyl Residues in the Primary Structure of Gene 5 Protein. Two peptides (C2 and C4) isolated from a chymotryptic digest of nitrated gene 5 protein contained all three nitrated tyrosyl residues (Table I). Peptide C2, in the nomenclature of Nakashima et al. (1974a,b), extending from residue 14 through 26 (Figure 8) contained only one tyrosine, no. 26, and it is nitrated. Peptide C4, residues 35–61, has two of its three tyrosines modified. This peptide was further digested with trypsin yielding two major fragments, C4T1 and C4T2, with compositions and amino and carboxyl termini corresponding to residues 35–46 and 47–61, respectively. The former contained one nitrated tyrosine, no. 41. A carboxypeptidase A digestion of C4T2, which contains one unmodified and one nitrotyrosyl, yielded equimolar amounts of unmodified tyrosine (Tyr-61) and leucine (Leu-60). Thus, residue 56 is the third modified tyrosyl. It is concluded that tyrosyl residues at positions 26, 41, and 56 are specifically nitrated by reaction of the protein with $C(NO_2)_4$.

Characterization of Cross-Linked Gene 5 Nitrated Protein Oligomers. Using concentrations determined by amino acid analysis, the extinction coefficients of the oligomeric fraction of nitrated gene 5 protein are found to be lower than those of the monomers: $\epsilon_{428} = 8.2 \times 10^3\ M^{-1}\ cm^{-1}$, $\epsilon_{381} = 5.6 \times 10^3\ M^{-1}\ cm^{-1}$, and $\epsilon_{278} = 14.6 \times 10^3\ M^{-1}\ cm^{-1}$. Again the amino acid analysis shows approximately 3 mol of nitrotyrosine/mol of protein and the normal molar ratios of other residues (Table I). The oligomeric proteins produce no change in the CD spectra when added to fd DNA even in great excess. The CD spectrum of the oligomeric proteins shows a significantly decreased ellipticity at 228 nm (not shown).

The lack of measurable cross-linking of gene 5 protein complexed to DNA by reaction with $C(NO_2)_4$ is consistent with the conclusions of Vincent et al. (1970) implicating the nitrated tyrosyls themselves in the bridging of proteins. This conclusion also explains the changes in the tyrosine regions of the CD and absorption spectra of the modified proteins.

NMR Spectrum of Fluorotyrosyl Gene 5 Protein. A fluorotyrosyl derivative of gene 5 protein was prepared by infecting a tyrosine auxotroph of *E. coli* growing on minimal media supplemented with *m*-fluorotyrosine with fd bacteriophage. The yields of active gene 5 protein were about 15% of those normally obtained. However, no difference between the DNA-binding properties of the native and fluorotyrosyl gene 5 protein could be detected by the CD assay for complex formation as depicted in Figures 1 and 2 for the native protein. Both the CD and absorption spectra of the fluorotyrosyl protein at pH 8 were identical with the normal protein except for a slight shoulder at the long-wavelength side of both spectra reflecting a slightly larger proportion of ionized fluorotyrosyl residues at this pH. This would be expected from the moderate decrease in pK_a of the hydroxyl group induced by the meta fluoro substitution (Sykes et al., 1974). A ^{19}F NMR spectrum of the fluorotyrosyl-substituted protein is shown in Figure 6. Five resonances corresponding to the five tyrosyl residues of the protein are resolved.[3] Three of the resonances (3, 4, and 5 in Figure 6) are grouped around the resonance position of ^{19}F in free *m*-fluorotyrosine and close to the resonance position shown by the fluorines in denatured fluorotyrosyl-substituted alkaline phosphatase (Hull and Sykes, 1974). The other two ^{19}F resonances (1 and 2 in Figure 6) are shifted downfield as

[3] A complete analysis of the ^{19}F NMR of the fluorotyrosyl derivative of 5p and the changes in chemical environment of the fluorotyrosyl residues induced by nucleotide binding is in preparation (R. A. Anderson, W. E. Hull, J. E. Coleman, and B. D. Sykes).

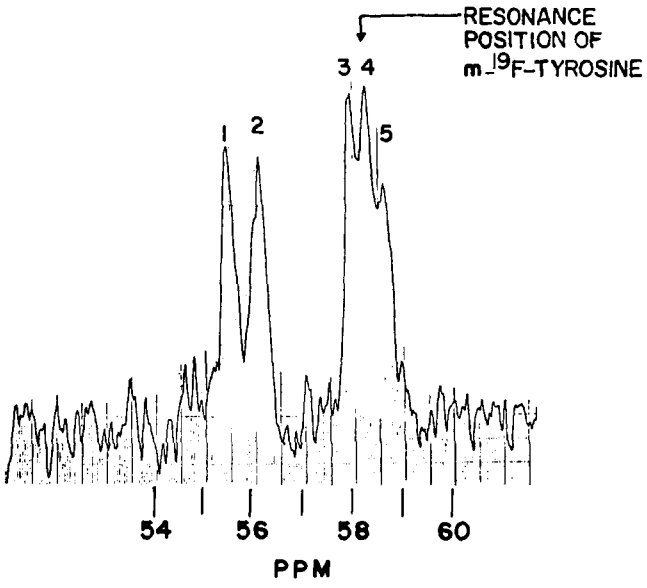

FIGURE 6: ^{19}F NMR spectrum of 7.3×10^{-4} *M* of gene 5 protein containing *m*-fluorotyrosyl residues; conditions: 10 m*M* Tris-HCl (pH 8), in H_2O, 25°. (See Experimental Section and Hull and Sykes (1974) for further details.)

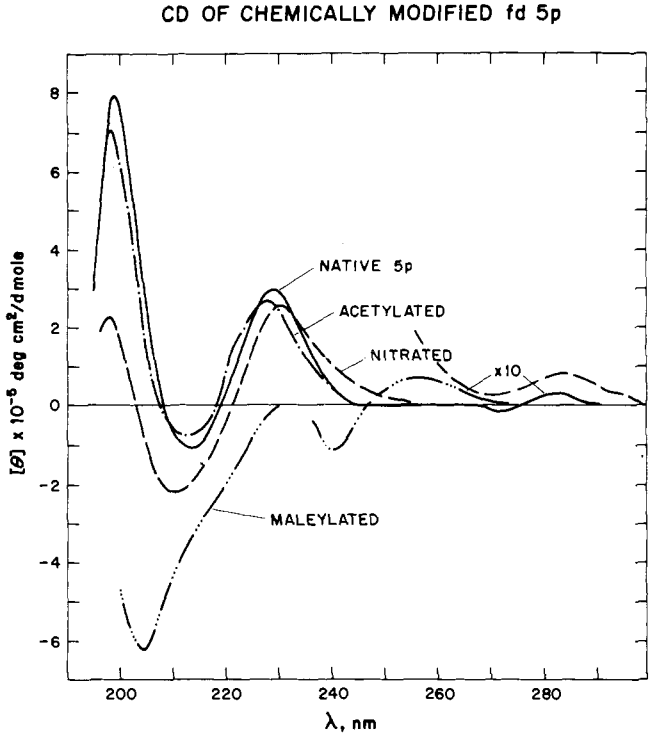

FIGURE 7: Circular dichroism of native and chemically modified gene 5 protein (5p); conditions: protein, 2×10^{-5} *M*, in 0.01 *M* Tris-HCl (pH 8). Ellipticity at wavelengths greater than 240 nm is shown magnified 10×: (—) native 5p; (- · -) fully acetylated 5p; (- - -) nitrated 5p; (- · · · -) maleyated 5p.

would be expected of fluorines on tyrosyl residues buried in the interior of the protein (Sykes et al., 1974). This result is consistent with the accessibility of three of the tyrosyl residues to modification by *N*-acetylimidazole and $C(NO_2)_4$.

Maleylation of Amino Groups in Gene 5 Protein. The reaction of proteins with maleic anhydride introduces negatively charged maleyl groups at α and ϵ amino moieties (Freedman et al., 1968). Gene 5 protein, 1.8×10^{-4} *M*, in 0.15 *M* borate buffer–0.15 *M* NaCl (pH 9) was allowed to react with a 40-fold molar excess (calculated per lysyl residue present) of solid maleic anhydride for 5 min at 25°, followed by treatment with 1 *M* NH_2OH to remove any maleylated hydroxyl groups. The value of *n*, the number of maleyl groups introduced, was calculated from the formula (Freedman et al., 1968):

$$n = \frac{A_{280}\epsilon_{250}^{5p} - A_{250}\epsilon_{280}^{5p}}{A_{250}\epsilon_{280}^{\text{mal-NH}} - A_{280}\epsilon_{250}^{\text{mal-NH}}} \quad (1)$$

Using the following values of the parameters in eq 1, $A_{280} = 0.142$, $A_{250} = 0.410$, $\epsilon_{280}^{5p} = 6900$, $\epsilon_{250}^{5p} = 3150$, $\epsilon_{280}^{\text{mal-NH}} = 310$, and $\epsilon_{250}^{\text{mal-NH}} = 3360\ M^{-1}\ \text{cm}^{-1}$, $n =$ 6.8 mol of maleyl groups/mol of protein. Apparently the ϵ-amino groups of all six lysyl residues of the protein and the α-amino groups of the N-terminal methionyl residue are maleylated. There is no binding of the maleylated protein to fd DNA as determined by the CD assay.

Chemical Modification of the Single Sulfhydryl Group of Gene 5 Protein. Ellman's reagent, 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman, 1959), is known to react specifically with accessible protein sulfhydryl groups producing a characteristic absorbance at 412 nm. The reagent does not react with native gene 5 protein. In contrast, Ellman's reagent reacts rapidly with 1 mol of sulfhydryl group/mol of protein when the protein is first denatured in 6 *M* guanidine · HCl.

Dialysis of native gene 5 protein against equimolar (50 μM) mercuric chloride in 0.05 *M* Tris-HCl–0.15 *M* NaCl (pH 8) results in the binding of 1 mol of Hg(II)/mol of protein as determined by atomic absorption. The mercurated protein does not form a complex with fd DNA and undergoes major conformational changes (see below). Prior formation of the complex with fd DNA, however, completely prevents the binding of Hg^{2+} to the protein. Gel filtration of the mercury-treated gene 5 protein–fd DNA complex on a 1 × 30 cm column of G-25 Sephadex completely removes the Hg^{2+}. The CD of the excluded material is the same as that of the native gene 5 protein–fd DNA complex.

Conformation of the Chemically Modified Gene 5 Proteins. The CD spectrum of the chemically modified species of gene 5 protein can be taken as a measure of the integrity of the protein secondary structure. Since the large positive ellipticity band at 228 nm in the native protein is probably a reflection of the particular environment of the tyrosyl chromophores conferred by the tertiary folding of the protein (Day, 1973), the CD in this case may also be considered to reflect the intactness of the tertiary folding.

The CD spectra of the several modified proteins are summarized in Figure 7. The CD spectrum of unmodified gene 5 protein dissolved in 6 *M* guanidine · HCl is that expected for a polypeptide in the random coil conformation. A similar spectrum is obtained for gene 5 protein reacted with maleic anhydride; the 228-nm band has disappeared and the CD spectrum is dominated by a large negative ellipticity band at 205 nm (Figure 7). A similar spectrum (not shown) is observed for the Hg^{2+} protein. The CD spectrum of fully acetylated gene 5 protein is very slightly altered from that of the native protein. A slight shift is reversed on treatment of the protein with 1 *M* NH_2OH.

Nitration only slightly changes the contours of the CD spectrum of gene 5 protein. The ellipticity band at 228 nm is shifted 3 nm to longer wavelength and the long-wavelength side of this band is broadened (Figure 7). The nitrated protein also shows increased ellipticity in the region around 280 nm consistent with the increased absorption of nitrotyrosine in this region. Thus, the nitrated protein does

not appear to exhibit major changes in secondary or tertiary structure.

Discussion

The CD assay used for protein–DNA binding in this study appears to be a reliable measure of the interactions of gene 5 protein with single-stranded DNA. The relative sensitivity of the gene 5 protein–fd DNA complex to dissociation by various cations as measured by the CD changes (Figure 3) correlates with the known affinities of cations for nucleotides (Ross and Scruggs, 1964). The change in the CD of the complex as a function of the concentration and species of cation is also consistent with the observed elution behavior of gene 5 protein on DNA cellulose columns (Alberts et al., 1972). The gene 5 protein:nucleotide stoichiometry in the complex as indicated by the CD assay (Figure 1) is in agreement with the stoichiometry determined by other methods (Alberts et al., 1972; Pratt et al., 1974). Electron micrographs of complexes formed from our preparations of gene 5 protein and fd DNA appear exactly the same as those previously published by Alberts et al. (1972).

While definitive statements about the structure of the DNA in the complex with gene 5 protein cannot be made from the CD, the dramatic reversal of the 260–270-nm ellipticity band is similar to the CD changes seen when DNA is complexed with polylysine (Chang et al., 1973). Similar changes, however, are observed when DNA is dissolved in concentrated alcohol (Green and Mahler, 1971; Girod et al., 1973) or in very concentrated salt solutions (Studdert et al., 1972). The basis of these CD changes could include dehydration in the vicinity of the DNA molecule, charge neutralization of the phosphate groups of the DNA (Chang et al., 1973), aggregation or condensation as occurs in alcohol (Girod et al., 1973), as well as changes in the secondary structure of the DNA. That change in the secondary structure of the single-stranded DNA underlies the CD changes accompanying complex formation with gene 5 protein is supported by the electron microscopy which shows that in the complex the circular fd DNA collapses such that two protein covered strands are pulled into a helical rod-like structure (Alberts et al., 1972; Pratt et al., 1974). It is not known whether gene 5 protein interacts with four bases on one strand or with two bases on each of two strands.

The amino acid sequence of gene 5 protein reveals an uneven distribution of charged residues along the polypeptide chain, with the basic residues clustered near the amino and carboxyl termini and the acidic residues located mainly in the central part of the chain. At physiological pH gene 5 protein has a net positive charge (Alberts et al., 1972). A crude calculation of the average degree of hydrophobicity of the gene 5 protein using the methods of Bigelow (1967) gives a number significantly higher than that for all the common globular proteins, with the most markedly hydrophobic residues distributed uniformly over the sequence except for a relative dearth over the amino-terminal one-fourth of the protein.

It is thought that the formation of salt linkages between protein lysyl ϵ-amino groups and the DNA backbone phosphates provides a major part of the energy for many stable DNA–peptide interactions, particularly those involving histones (Malchy and Kaplan, 1974). The dissociation of the gene 5 protein–fd DNA complex induced by salt and the prevention of complex formation by N-acetylation of the lysyl ϵ-amino groups suggest that the lysyl residues of gene 5 protein play a major role in formation of the DNA complex. As might be expected, all the ϵ-amino groups of the lysyl residues appear to be on the surface of the molecule and are susceptible to acetylation and maleylation. The introduction of acetyl groups at the ϵ-amino groups does not appear to produce changes in the secondary or tertiary structure of the protein, since no change occurs in the CD of the acetylated protein in which the *O*-acetyl groups have been removed (Figure 7). The rapid disruption of preformed gene 5 protein–DNA complexes by mild acetylation of the complex suggests that the lysyl ϵ-amino-phosphate linkages may occur on the exterior of the complex.

The single sulfhydryl group of gene 5 protein seems to be centrally involved in the maintenance of tertiary structure as well as centrally located in the primary structure. The SH group is probably not exposed enough to interact directly with the DNA, considering its lack of reactivity with Ellman's reagent. Yet addition of two positive charges at this site by Hg^{2+} complexation completely destroys the folding of the protein. Protection of the -SH from reaction with Hg^{2+} by the DNA implies that the DNA prevents access from the solvent to this residue.

It is remarkable that gene 5 protein is stable in 6 *M* urea and in 3 *M* guanidine · HCl, but its native structure appears to be very sensitive to alterations in charge distribution over the molecule. The introduction of negative charges on the α- and ϵ-amino groups by maleylation denatures the protein (Figure 7). This result, along with the sensitivity of the structure to the manipulation of the charge on the single sulfhydryl group, suggests that electrostatic interactions may be as important, if not more so, than hydrogen bonding in maintaining the tertiary structure of the protein.

Tyrosyl residues have been implicated in the active sites of two enzymes involved in the processing of DNA: *Staphylococcus aureus* extracellular nuclease, by X-ray diffraction (Arnone et al., 1971), and pancreatic deoxyribonuclease I, by chemical modification (Hugli and Stein, 1971). Fluorescence studies have indicated an interaction between the tyrosyl residues of histones and DNA (Matsuyama and Nagata, 1970). The tyrosyl-rich amino-terminal portion of the lac repressor apparently contains the amino acid sequence which binds to the lac operator (Adler et al., 1972; Beyreuther et al., 1973). Proton magnetic resonance studies of oligopeptide–nucleic acid polymer complexes have demonstrated the tendency of tyrosyl, phenylalanyl, and tryptophanyl residues to partially intercalate between nucleotide bases (Helene et al., 1971; Gabbay et al., 1973; Dimicoli and Helene, 1974).

The intercalation mechanism involves partial overlapping of the aromatic ring of the amino acid with the heterocyclic ring of the nucleotide bases and induces changes in the nucleic acid conformation (Helene et al., 1971; Gabbay et al., 1973; Sellini et al., 1973; Dimicoli and Helene, 1974). A preference of tyrosine for intercalation with A-T over G-C sites has been observed (Gabbay et al., 1973). Infrared and proton magnetic resonance studies in organic solvents have demonstrated the ability of analogs of tyrosine to also hydrogen bond with the DNA bases *via* the phenolic hydroxyl (Sellini et al., 1973). (Carboxylic acids also hydrogen bond to DNA bases.)

The protection by DNA of the three accessible tyrosyls of gene 5 protein from nitration with $C(NO_2)_4$ indicates that these tyrosyls, initially on the surface of the isolated protein, become buried in the gene 5 protein–DNA complex. This is consistent with the decrease in ellipticity of the pro-

tein at 228 nm accompanying formation of the DNA complex which has been interpreted as reflecting a shift of the tyrosyl residues to a more hydrophobic environment (Day, 1973). These three exterior tyrosyl residues could interact either with the DNA by hydrogen bonding or intercalation between bases or with the adjacent gene 5 protein molecules.

Nitro substitution on the phenoxazone ring of actinomycin does not interfere with its ability to intercalate with DNA as reflected in the similar dissociation constants of the DNA complex with the substituted and unsubstituted antibiotic (Müller and Crothers, 1968). However, the situation with the nitrated tyrosines of gene 5 protein and single-stranded DNA may not be entirely analogous to the models for intercalation with double-stranded DNA (Sobell, 1972). The spectral characteristics of model compounds suggest that the nitro group ortho to the ionized phenolic oxygen is forced to deviate from a coplanar alignment with the phenolate ring. This could increase the effective ring thickness from 3.4 to greater than 5 Å. Thus, the loss of binding affinity of nitrated gene 5 protein for DNA might involve interference of the nitrotyrosine ring with intercalation. The bulkier nitrophenol ring might also be a steric hindrance to complex formation by coming into contact with the nucleic acid backbone.

Alternatively, the fall in the pK_a of the phenolic hydroxyl group of the 3-nitrotyrosyl residue to ~7 (Figure 5) might introduce negative charges which interfere with complex formation or which prevent hydrogen bonding with the DNA bases. The latter two alternatives are less likely since the gene 5 protein containing fluorotyrosyl residues exhibits no detectable loss of binding affinity despite the fact that the pK_a of the tyrosyl hydroxyl groups is lowered by 2 pH units. The intercalation mechanism might also explain the apparent changes in DNA conformation suggested by the CD spectra (Figure 1) and the electron micrographs. Upfield chemical shifts of the ^{19}F resonances of the three surface tyrosyls observed on the binding of tetranucleotides to fluorotyrosyl gene 5 protein are compatible with the intercalation mechanism.[3]

The results of the chemical modification of gene 5 protein can be summarized in terms of a model of the secondary structure of the protein (Figure 8) constructed from calculations of the probability of the occurrence of α, β, or random coil structure based on amino acid sequence as outlined by Chou and Fasman (1974). We have attempted to use the predicted secondary structure along with the accumulated information on the reactivities and functional roles of individual residues in the free protein and in the protein-DNA complex to determine the relationships between particular chemical groups of the protein and the DNA by constructing a speculative model of the gene 5 protein-DNA complex. The most striking features of the protein secondary structure predicted by the conformational analysis are two long regions of β structure (residues 31-50 and 61-84) symmetrically broken by β turns (Figure 8). Several other nucleic acid binding proteins also appear to contain a significant proportion of β structure. *Staphylococcus aureus* nuclease and ribonuclease S contain 25 and 44% β structure, respectively, as shown by X-ray diffraction (Chou and Fasman, 1974). It has also been hypothesized on the basis of various conformational parameters that both the lac repressor protein (Chou et al., 1974) and pancreatic deoxyribonuclease I (Cheng, 1966) contain large amounts of β structure. A model for the interaction of polypeptides in the β

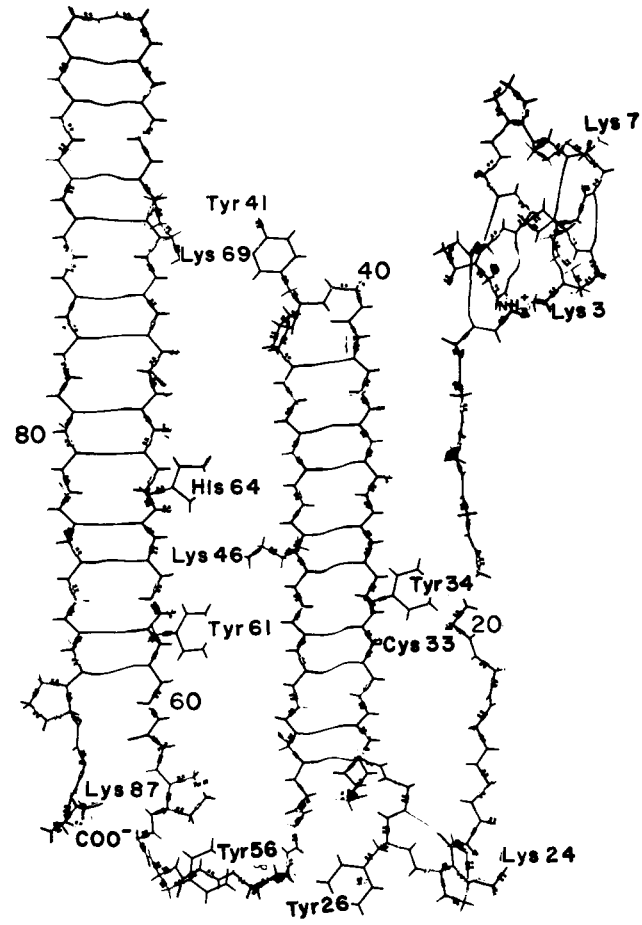

FIGURE 8: Secondary structure of gene 5 protein as predicted from the conformational parameters for amino acids in α-helical, β-sheet, and random coil regions as determined by the method of Chou and Fasman (1974). The sequence of gene-5 protein as determined by Nakashima et al. (1974a,b) is given below and the three tyrosyl residues modified by $C(NO_2)_4$ are underlined.

Met-Ile-Lys-Val-Glu-Ile-Lys-Pro-Ser-Gln-Ala-Gln-
10
Phe-Thr-Thr-Arg-Ser-Gly-Val-Ser-Arg-Gln-Gly-Lys-
20
Pro-Tyr-Ser-Leu-Asn-Glu-Gln-Leu-Cys-Tyr-Val-Asp-
30
Leu-Gly-Asn-Glu-Tyr-Pro-Val-Leu-Val-Lys-Ile-Thr-
40
Leu-Asp-Glu-Gly-Gln-Pro-Ala-Tyr-Ala-Pro-Gly-Leu-
50 60
Tyr-Thr-Val-His-Leu-Ser-Ser-Phe-Lys-Val-Gly-Gln-
70
Phe-Gly-Ser-Leu-Met-Ile-Asp-Arg-Leu-Arg-Leu-Val-
80
Pro-Ala-Lys

conformation with nucleic acid polymers has recently been proposed (Carter and Kraut, 1974).

Arrangement of the predicted secondary structure of gene 5 protein within the limited dimensions of the stacked bases of a tetranucleotide with a maximum number of lysyl residues in contact with the phosphate groups, and with the surface tyrosyls 26, 41, and 56 in close contact with the DNA, severely limits the tertiary configurations of large segments of the peptide chain. This model allows the three tyrosines which can be modified to be "stacked" in a manner suggested by the prominent ellipticity at 228 nm. However, it is impossible to bring Tyr-26 and Tyr-56, in the

"coil regions," closer than about 20 Å to Tyr-41 which is in the midst of a sheet region and maintain the long regions of β-pleated sheet. It may be that Tyr-41 is involved in a protein-protein contact if the other two accessible tyrosyl residues are involved with the DNA in the complex. If the lysine-rich, relatively nonhydrophobic amino-terminal region of the peptide chain is wound around to the tyrosyl region, four lysyl residues (Lys-3, -7, -24, and -69) can be manipulated to interact with tetranucleotide phosphates. This leaves the cysteine (Cys-33) "buried" behind the DNA interacting Tyr-26. The elongated structure of the protein and its possible mechanisms of binding to DNA as deduced from the above simple assumptions are reminiscent of the overlapping "fish scale" structure suggested by Marvin et al. (1974) for the binding of the largely helical coat protein to DNA in intact filamentous phage. The speculative nature of this structure is emphasized, however, by a recent report by Schulz et al. (1974) demonstrating the tendency of the Chou and Fasman system to overpredict β structure.

There is already a long list of known nonenzymatic DNA-binding proteins similar to gene 5 protein including those from bacteriophage T4 (Alberts and Frey, 1970) and T7 (Reuben and Gefter, 1974), *E. coli* (Sigal et al., 1972), adenovirus (Van der Vliet and Levine, 1973), prophase *Lilium* nuclei (Hotta and Stern, 1971), and mammalian spermatocytes (Hotta and Stern, 1971). The gene 5 protein may be the smallest and simplest of the DNA unwinding proteins. The spectroscopic and chemical techniques used in the present work may be employed in the investigation of general mechanisms of the binding of the DNA unwinding proteins.

Knowledge of the molecular details of this type of interaction should be useful in attempting to understand the basis for the affinity exhibited by other, more specific DNA binding proteins such as the bacteriophage λ and lac operon repressor proteins, and the RNA polymerases.

Acknowledgment

We thank Dr. W. Konigsberg for many helpful suggestions and discussions. We thank Mr. E. Anderson and Dr. D. Marvin for the gift of bacteriophage and for much helpful advice. We thank Dr. G. Bourgignon for performing the electron microscopy.

References

Adler, K., Beyreuther, K., Fanning, E., Geisler, N., Gronenborn, B., Klemm, A., Müller-Hill, B., Pfahl, M., and Schmitz, A. (1972), *Nature (London) 237,* 322–327.

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

Alberts, B., and Herrick, G. (1971), *Methods Enzymol. 21,* 198–217.

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

Ambler, R. P. (1972), *Methods Enzymol. 25,* 143–154.

Beavan, G. H., and Holiday, E. R. (1952), *Adv. Protein Chem. 7,* 319–386.

Beyreuther, K., Adler, K., Geisler, N., and Klemm, A. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 3576–3580.

Bigelow, C. C. (1967), *J. Mol. Biol. 16,* 187–211.

Bohlen, P., Stein, S., Dairman, W., and Udenfriend, S. (1973), *Arch. Biochem. Biophys. 155,* 213–220.

Carter, C. A., Jr., and Kraut, J. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 283–287.

Chang, C., Weiskopf, M., and Li, H. J. (1973), *Biochemistry 12,* 3028–3032.

Cheng, P.-Y. (1966), *Proc. Natl. Acad. Sci. U.S.A. 55,* 1535–1538.

Chou, P. Y., Adler, A. J., and Fasman, G. D. (1974), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 33,* 1383 Abstr.

Chou, P. Y., and Fasman, G. D. (1974), *Biochemistry 13,* 222–245.

Day, L. A. (1973), *Biochemistry 12,* 5329–5339.

Dimicoli, J. L., and Helene, C. (1974), *Biochemistry 13,* 724–730.

Duckworth, H. W., and Coleman, J. E. (1970), *Anal. Biochem. 34,* 382–386.

Ellman, G. L. (1959), *Arch. Biochem. Biophys. 82,* 70–77.

Fidanian, H. M., and Ray, D. S. (1972), *J. Mol. Biol. 72,* 51–63.

Freedman, M. H., Grossberg, A. L., and Pressman, D. (1968), *Biochemistry 7,* 1941–1950.

Gabbay, E. J., Sanford, K., Baxter, C. S., and Kapicak, L. (1973), *Biochemistry 12,* 4021–4029.

Geider, K., and Kornberg, A. (1974), *J. Biol. Chem. 249,* 3999–4005.

Girod, J. C., Johnson, W. C., Jr., Huntington, S. K., and Maestre, M. F. (1973), *Biochemistry 12,* 5092–5096.

Gray, W. R. (1972), *Methods Enzymol. 25,* 121–138.

Green, G., and Mahler, H. R. (1971), *Biochemistry 10,* 2200–2216.

Helene, C., Montenay-Garestier, T., and Dimicoli, J. L. (1971), *Biochim. Biophys. Acta 254,* 349–365.

Hotta, Y., and Stern, H. (1971), *Nature (London), New Biol. 234,* 83–86.

Hugli, T. E., and Stein, W. H. (1971), *J. Biol. Chem. 246,* 7191–7200.

Hull, W. E., and Sykes, B. D. (1974), *Biochemistry 13,* 3431–3437.

Malchy, B., and Kaplan, H. (1974), *J. Mol. Biol. 82,* 537–545.

Marvin, D. A., and Hohn, B. (1969), *Bacteriol. Rev. 33,* 172–209.

Marvin, D. A., and Schaller, H. (1966), *J. Mol. Biol. 15,* 1–7.

Marvin, D. A., Wiseman, R. L., and Wachtel, E. J., (1974), *J. Mol. Biol. 82,* 121–138.

Matsuyama, A., and Nagata, C. (1970), *Biochim. Biophys. Acta 224,* 588–596.

Mazur, B. J., and Model, P. (1973), *J. Mol. Biol. 78,* 285–300.

Miller, J. H. (1972), Experiments in Molecular Genetics, Appendix I, Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.

Moore, S. (1963), *J. Biol. Chem. 238,* 235–237.

Müller, W., and Crothers, D. M. (1968), *J. Mol. Biol. 35,* 251–290.

Nakashima, Y., Dunker, A. K., Marvin, D. A., and Konigsberg, W. (1974a), *FEBS Lett. 40,* 290–292.

Nakashima, Y., Dunker, A. K., Marvin, D. A., and Konigsberg, W. (1974b), *FEBS Lett. 43,* 125.

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

Pratt, D., and Erdahl, W. S. (1968), *J. Mol. Biol. 37,* 181–200.

Pratt, D., Laws, P., and Griffith, J. (1974), *J. Mol. Biol. 82,* 425–439.

Reuben, R. C., and Gefter, M. L. (1974), *J. Biol. Chem. 249,* 3843–3850.

Riordan, J. F., Wacker, W. E. C., and Vallee, B. L. (1965), *Biochemistry 4,* 1758–1765.
Ross, P. D., and Scruggs, R. L. (1964), *Biopolymers 2,* 79–89.
Salstrom, J. S., and Pratt, D. (1971), *J. Mol. Biol. 61,* 489–501.
Schultz, G. E., Barry, C. D., Friedman, J., Chou, P. Y., Fasman, G. D., Finkelstein, A. V., Lim, V. I., Ptitsyn, O. B., Kabat, E. A., Wu, T. T., Levitt, M., Robson, B., and Nagano, K. (1974), *Nature (London) 250,* 140–142.
Sellini, H., Maurizot, J. C., Dimicoli, J. L., and Helene, C. (1973), *FEBS Lett. 30,* 219–224.
Sigal, N., Delius, H., Kornberg, T., Gefter, M. L., and Alberts, B. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 3537–3541.
Simpson, R. T., Riordan, J. F., and Vallee, B. L. (1963), *Biochemistry 2,* 616–622.
Sobell, H. M. (1972), *Prog. Nucleic Acid Res. Mol. Biol. 13A,* 153–190.
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.
Staudenbauer, W. L., and Hofschneider, P. H. (1973), *Eur. J. Biochem. 34,* 569–576.
Studdert, D. S., Patroni, M., and Davis, R. C. (1972), *Biopolymers 11,* 761–779.
Sykes, B. D., Weingarten, H. I., and Schlesinger, M. J. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 469–473.
Taylor, A. L., and Trotter, C. D. (1967), *Bacteriol. Rev. 31,* 332–353.
Tseng, B. Y., and Marvin, D. A. (1972), *J. Virol. 10,* 384–391.
Van der Vliet, P. C., and Levine, A. J. (1973), *Nature (London), New Biol. 246,* 170–174.
Vincent, J. P., Lazdunski, M., and Delaage, M. (1970), *Eur. J. Biochem. 12,* 250–257.
Weber, K., and Osborn, M. (1969), *J. Biol. Chem. 244,* 4406–4412.
Wetlaufer, D. B. (1962), *Adv. Protein Chem. 17,* 303–390.
Wiseman, R. L., Dunker, A. K., and Marvin, D. A. (1972), *Virology 48,* 230–244.

Regulatory Proteins of Lobster Striated Muscle†

Joe M. Regenstein‡ and Andrew G. Szent-Györgyi*

ABSTRACT: The regulatory proteins of lobster muscles consist of tropomyosin and of troponin. Troponin contains a 17,000 chain weight component, two closely related components of about 30,000 and a 52,000 chain weight component. In addition to troponin, tropomyosin is required for the inhibition of the magnesium activated actomyosin ATPase activity in the absence of calcium and for the reversal of this inhibition by calcium. Lobster tropomyosin interacts with rabbit actin and lobster troponin interacts with rabbit tropomyosin. The 30,000 doublet component corresponds to the troponin-I of rabbit and inhibits the ATPase activity of actomyosin both in the presence and in the absence of calcium. The 17,000 component corresponds to the troponin-C of rabbit; it binds calcium and reverses the inhibition of the ATPase activity by troponin-I in the presence of calcium. No more than 1 mol of calcium is bound by a mole of troponin-C or by troponin. The 52,000 component interacts with tropomyosin and has been tentatively identified as troponin-T; however, it has not been demonstrated as yet that this component had a role in the regulation of lobster actomyosin.

The mechanism of troponin action in the calcium regulation of contraction has been studied in detail only in vertebrate striated muscles (*cf.* Ebashi and Endo, 1968; Weber and Murray, 1973). Rabbit troponin was shown to contain three different subunits (Greaser and Gergely, 1971). Troponin-I (TN-I)[1] inhibits the actin activated ATPase both in the presence and in the absence of calcium (Perry *et al.*, 1972; Ebashi *et al.*, 1972; Greaser *et al.*, 1972; Hartshorne and Dreizen, 1972). Troponin-C (TN-C) is the calcium binding protein, and in the presence of calcium it removes the TN-I induced inhibiton of the ATPase (Hartshorne and Mueller, 1968). Troponin-T (TN-T) binds to tropomyosin, and it is required together with tropomyosin for calcium regulated control (Greaser and Gergely, 1971). The components of chicken troponin, although less extensively studied, are similar to rabbit (Hitchcock *et al.*, 1973).

The presence of troponin has also been shown in insect muscles (Maruyama *et al.*, 1968; Meinrenken, 1969; Bullard *et al.*, 1973), and in *Limulus* (Lehman *et al.*, 1972). An actin-linked regulation requiring tropomyosin has been demonstrated in a number of invertebrate muscles, and it may coexist with the myosin-linked regulatory system, as in insects and annelids, or alternatively it functions alone as in most decapod muscles (Lehman *et al.*, 1972, 1974).

These studies also indicate that invertebrate troponin is a multicomponent protein system (Lehman *et al.*, 1972; Bul-

† From the Department of Biology, Brandeis University, Waltham, Massachusetts 02154. *Received May 28, 1974.* This research was supported by grants from the U.S. Public Health Service (GM 14675 and AM 15963 to A.G.S.G.) and a Gillette Corporation Fellowship (to J.M.R.). A preliminary account of these studies has been published (Regenstein and Szent-Gyorgyi, 1973).

‡ Present address: Department of Poultry Science, Cornell University, Ithaca, New York.

[1] Abbreviations used are: TN, troponin; EGTA, ethylene glycol bis(β-aminoethyl ether)-*N,N'*-tetraacetic acid; SDS, sodium dodecyl sulfate.