Proton Nuclear Magnetic Resonance (500 MHz) of Mono-, Di-, Tri-, and Tetradeoxynucleotide Complexes of Gene 5 Protein†

Thomas P. O'Connor‡ and Joseph E. Coleman*

ABSTRACT: The 500-MHz proton spectra of the complexes of gene 5 protein from bacteriophage fd (M13) with dpA, $d(pA)_2$, $d(pA)_3$, $d(pA)_4$, and $d(pA)_8$ show that oligonucleotides carrying a 5′-phosphate form 1:1 complexes involving the same specific binding site for the 5′-phosphate dianion. Formation of the dpA mononucleotide complex induces small upfield shifts in the resonances of one Phe and one Tyr residue. The magnitudes of these shifts increase as the length of the nucleotide increases. Phe protons shift 0.05 ppm for dpA, 0.15 ppm for $d(pA)_2$, 0.25 ppm for $d(pA)_4$, and 0.6 ppm for $d(pA)_8$, while the associated shifts for the Tyr protons are 0.03, 0.14, 0.15, and 0.32 ppm. Thus, the 5′ base of the bound nucleotide must rotate closer to the pocket formed by Tyr and Phe, or the stacking distance of the Tyr(5′-A)Phe must decrease as the distal residues of an oligonucleotide are bound. This movement may be induced by conformational changes in a mobile part of the DNA binding domain which contains Tyr-26, a domain clearly delineated in the crystal structure. As the nucleotide is lengthened, protons from two additional Tyr residues shift upfield. The uncomplexed gene 5 protein exists in solution as the dimer, and the proton NMR spectra show that no change in quaternary structure occurs until the oligonucleotide length exceeds four bases. The sequential ring-current shifts induced in the aromatic residues of the protein as the complexing nucleotide is lengthened from one to four residues support the sequential arrangement of three tyrosyl and one phenylalanyl rings along the DNA binding groove. The ring-current shifts in the protons of the phenylalanyl and one of the tyrosyl residues appear to be closely coupled and associate these two residues with the 5′ end of the DNA binding groove. Assignments of the proton resonances to specific amino acid residues located sequentially in the DNA binding groove can be made on the basis of the refined crystal structure as follows: Tyr-26, Phe-73, Tyr-34, and Tyr-41, with Tyr-26 and Phe-73 forming the 5′ end of the DNA binding domain while Tyr-41 is located at the 3′ end.

Gene 5 protein from the filamentous bacteriophage fd (M13), M_r 9689 (87 amino acids), has served as a model for studies of the binding mode of single-stranded DNA binding proteins to single-stranded DNA and oligodeoxynucleotides (Anderson et al., 1975; Garssen et al., 1977; Coleman & Armitage, 1978; Alma et al., 1981). The in vivo function of this protein is to complex the emerging single-stranded daughter virions and prevent their use as templates for the synthesis of the complementary strand [Alberts et al., 1972; see Kornberg (1980 and 1982) for a review]. Hence, gene 5 is responsible for the shift from double-stranded to single-stranded DNA synthesis during phage infection. Nitration studies of the gene 5 protein originally implicated three of the five tyrosyl residues of the protein as involved in DNA binding (Anderson et al., 1975). Subsequent ^{19}F and ^{1}H studies of the binding of oligonucleotides (tetra- and octaoligonucleotides) showed the protons of one phenylalanyl and two tyrosyl residues to undergo upfield shifts upon binding of nucleotides to the protein (Coleman et al., 1976; Coleman & Armitage, 1978; Garssen et al., 1980). A model has been proposed in which nucleotide binding to gene 5 protein involves intercalation of the bases with the tyrosyl and phenylalanyl residues, while much of the binding energy is supplied by electrostatic interactions of the Lys and Arg side chains with the phosphate backbone (Coleman & Armitage, 1978; Alma et al., 1981).

More recently, O'Connor & Coleman (1982) have shown by using ^{31}P and ^{19}F NMR of dpA, FdUMP, $d(pA)_2$, $d(pA)_3$, $d(pA)_4$, and $d(pA)_8$ complexes of gene 5 protein that the 5′-phosphate dianions of the mono-, di-, and oligonucleotides make a major contribution to the binding affinity, so much so in fact that only a single equivalent of each of these nucleotides (even the mononucleotide) binds at a specific site on gene 5 protein, a binding mode directed by the 5′ dianion. The first equivalent is in slow exchange on the ^{31}P NMR time scale, while any additional equivalents are in rapid exchange, if bound at all.

The present 500-MHz proton NMR confirms a single predominant mode of binding for these 5′-oligonucleotides with the exception of $d(pA)_2$ and allows a systematic exploration of the base-binding lattice of gene 5 protein by following the incremental changes in the ^{1}H spectrum of the aromatic residues of the protein as the nucleotide is extended from one to four residues along the binding groove from the specific binding site for the 5′-phosphate. The latter site appears the same for all of them.

Materials and Methods

Gene 5 Protein. Homogeneous gene 5 protein was prepared by DNA–cellulose chromatography as previously described (Alberts & Herrick, 1971; Anderson et al., 1975). Protein concentration was determined by using $E^{0.1\%}_{276nm} = 0.73$.

Nucleotides. $d(pA)_n$'s (n = 1–4 and 8) were purchased from Collaborative Research (Waltham, MA). Nucleotides were dissolved in H_2O and extracted with a CCl_4 solution of dithizone to remove paramagnetic metals. Purity of the nucleotides was monitored by the proton NMR spectra of the free nucleotides. Nucleotides were added to protein solutions by lyophilizing the appropriate amount of nucleotide from D_2O and adding the protein solution. All solutions were in 0.01 M DPO_4^{2-} and 0.1 M NaCl, pH 8.0, 25 °C, unless otherwise noted. Protein concentration was ~1.5 mM for all samples.

^{1}H NMR spectra were obtained at 500 MHz on a Bruker WM-500 spectrometer. D_2O present in the sample served as the field-frequency lock. Samples were 0.5 mL contained in 5-mm tubes. Chemical shifts are reported in parts per million

† From the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510. *Received January 19, 1983.* This work was supported by Grant GM 21919 from the National Institutes of Health. The 500-MHz NMR facility is supported by Grant CHE-7916210 from the National Science Foundation.

‡ Present address: Makari Research Laboratories, Englewood, NJ 07631.

0006-2960/83/0422-3375$01.50/0

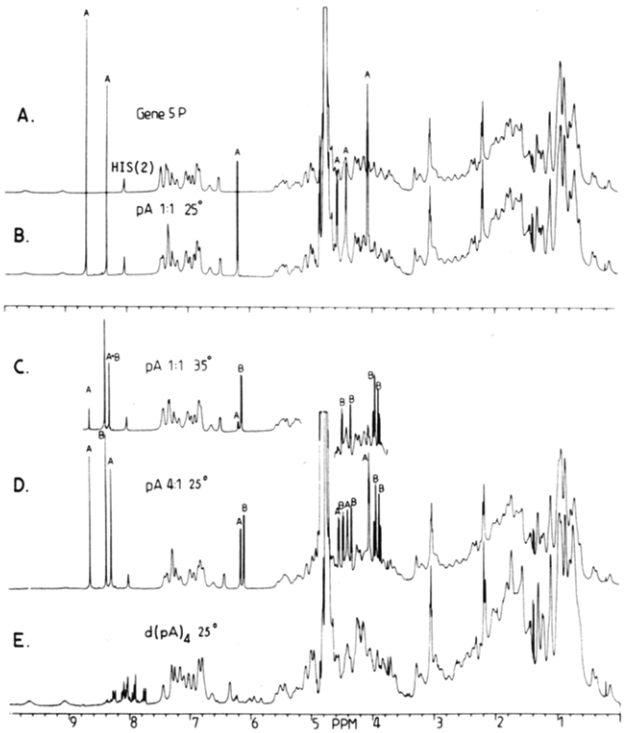

FIGURE 1: 500-MHz ^{1}H NMR spectra of the aromatic and aliphatic proton regions of gene 5 protein and oligodeoxynucleotide complexes. (A) Gene 5 protein; (B) 1:1 complex with pA at 25 °C; (C) 1:1 complex with pA at 35 °C; (D) 4:1 mixture of pA to gene 5 protein at 25 °C; (E) 1:1 complex with $d(pA)_4$ at 25 °C. At the level of the 5′-mononucleotide, gene 5 protein does not distinguish between ribose and deoxyribose. The phenomena pictured here for pA are also observed for dpA, except that H(1′) is a triplet and the most upfield sugar resonance is that of the H(2′) protons.

(ppm) downfield from sodium 4,4-dimethyl-4-silapentanesulfonate.

Results

Aromatic and Aliphatic Proton Spectra of Gene 5–$d(pA)_n$ Complexes. The complete proton spectrum of the deuterium-exchanged gene 5 protein is shown in Figure 1A. Detailed assignments of the aromatic protons are given in Figure 2A on the basis of extensive ^{1}H NMR of gene 5 protein (Coleman & Armitage, 1978; Alma et al., 1981). On formation of oligonucleotide complexes of the $d(pA)_n$ series, if n is under 4 there is relatively little change in the envelope of the aliphatic protons, suggesting that there are no general changes in protein conformation or in quaternary structure when the nucleotide is bound. Spectra of two examples, the complexes with pA and $d(pA)_4$, are shown in parts B and E, respectively, of Figure 1. On formation of the 1:1 dpA (or pA) complex, the aromatic protons undergo the maximum upfield shift observed with the 5′-mononucleotide (Figure 1B). Increasing the dpA (or pA):protein ratio to 4:1 results in no further change in the aromatic spectrum of the protein, and a set of resonances (B) from "free" nucleotide appears (Figure 1D). Hence, only a 1:1 stoichiometric complex of dpA is formed with gene 5 protein, and the bound nucleotide is in slow exchange on the ^{1}H NMR time scale. The amplitudes of the proton resonances of the bound and free nucleotides cannot be directly compared, since the T_1 values of the protons on the free ligand are much longer and these spectra are not fully relaxed. Thus, the first molecule of dpA (or pA) must be bound at a unique site as previously suggested by the ^{31}P data (O'Connor & Coleman, 1982).

The most upfield ribose resonance in both the bound and free pA is from the H(5′) resonance (two protons). The H(3′)

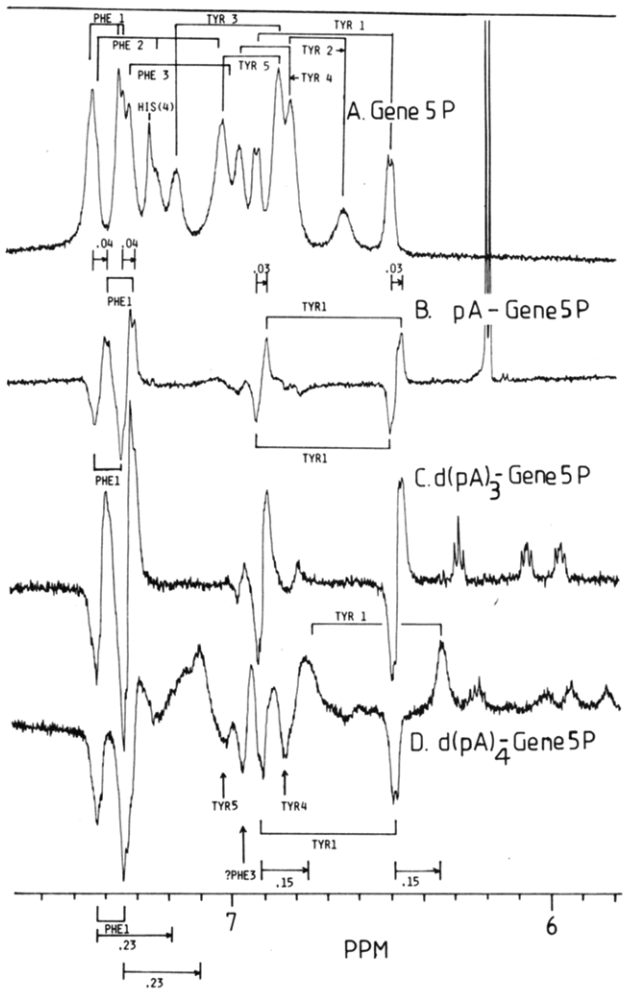

FIGURE 2: ^{1}H aromatic proton NMR difference spectra, spectra of complexes minus spectra of gene 5 protein. (A) Gene 5 protein alone. In this and the following figures, the dash is a minus sign, and the nucleotide symbol $d(pA)_n$ refers to the complex with gene 5 protein. (B) Spectrum of the pA complex (1:1) minus that of the unliganded protein. (C) Spectrum of the $d(pA)_3$ complex minus that of the unliganded protein. (D) Spectrum of the $d(pA)_4$ complex minus that of the unliganded protein.

resonance is lost in the HDO peak. In the free nucleotide, the two H(5′) protons have well-separated chemical shifts and give rise to a typical AB quartet. In the free nucleotide, the adenine H(8) is also shifted upfield by ~0.3 ppm. The relatively low amplitude of the peak at 8.25 ppm (Figure 1C,D) reflects the fact that the H(2) resonance of adenine in free and bound pA has only a slightly different chemical shift. This peak, marked A+B in Figure 1C, consists of both resonances as can be shown by expansion. The unique chemical shifts of the free nucleotide, different from pA in buffer, may relate to the high viscosity of the solution or nonspecific binding involving rapid exchange. Increased stacking may account for the upfield shifts of the adenine H(8) and the sugar H(2).

A dramatic demonstration that the B set of resonances in Figure 1D represents free nucleotide is provided by warming the 1:1 complex of gene 5 protein with pA from 25 to 35 °C. A large fraction of the nucleotide dissociates and gives rise to the B set of resonances (Figure 1C), identical with those observed at 25 °C in the presence of excess pA (Figure 1D). The aromatic proton resonances of the protein also shift back toward their positions in the free protein on temperature-induced dissociation of the nucleotide. The dissociation is readily reversible, and the original spectrum of the 1:1 complex is restored on lowering the temperature.

Difference Spectra, $d(pA)_n$ Complexes minus Gene 5 Protein. As an assay of the specific aromatic resonances shifting on complex formation, difference spectra were determined by assuming that the amplitude of the His H(2) resonance does not change on complex formation, an assumption supported by previous studies (Coleman et al., 1976; Coleman & Armitage, 1978). The only severe overlap between the His H(2) and base protons occurs in the $d(pA)_2$ complex, but expansion of the signal near 8.03 ppm separates the resonances sufficiently to do difference spectra.

Three of the difference spectra formed by subtracting the gene 5 protein spectrum from the spectra of the complexes are presented in Figure 2 with the aromatic ^{1}H spectrum of gene 5 protein shown at the top for reference. Formation of the 1:1 complex with dpA (or pA) results in resonance shifts assignable to two specific residues. Both the doublets assigned to Tyr-1 (the 2,6 and 3,5 protons, respectively) shift upfield by ~0.03 ppm. In addition, resonances assigned to Phe-1 shift upfield by ~0.04 ppm (Figure 2B). The 1:1 complex of $d(pA)_3$ shows almost the same difference spectrum, involving the signals from the same residues, Phe-1 and Tyr-1 (Figure 2C). The $d(pA)_4$ complex includes shifts of the same resonances, but the shifts are now ~0.25 ppm for the Phe-1 resonances and ~0.15 ppm for both the 2,6 and 3,5 resonances of Tyr-1. In addition to the upfield shifts involving Phe-1 and Tyr-1, shifts occur in the resonances assigned to Tyr-4 and Tyr-5 on formation of the $d(pA)_4$ complex (Figure 2D).

Difference Spectra, $d(pA)_2$ Complexes minus Gene 5 Protein. The difference spectrum for the $d(pA)_2$ complex was not included in Figure 2, since in contrast to the other oligonucleotides both 1:1 and 2:1 complexes of reasonable stability can be formed. The different difference spectra for the spectrum of the $d(pA)_2$ complex minus that of the gene 5 protein are shown in Figure 3 with the spectrum of the free protein for reference. The first involves the complex formed with a 1:1 ratio of nucleotide to protein (Figure 3B). The second involves the complex formed with a 2:1 ratio of $d(pA)_2$ to protein (Figure 3C).

The upfield shifts most easily identifiable in both difference spectra are those assignable to Phe-1 and Tyr-1. The pattern of shifts for the protons of Phe-1 and Tyr-1 is similar to the pattern observed for the dpA complex except larger, 0.15 ppm for the Phe-1 protons and 0.14 ppm for both sets of protons on Tyr-1 in the case of the 1:1 complex. In the presence of the 2:1 ratio of $d(pA)_2$ to gene 5 protein, these same protons shift further upfield, a total of 0.27 ppm for Phe-1 and 0.20 ppm for Tyr-1. The additional incremental shifts on formation of the 2:1 complex are illustrated by the difference spectrum between the two complexes (Figure 3D).

In addition to the shifts of Phe-1 and Tyr-1 protons, both the $d(pA)_2$ complexes show upfield shifts of resonances assigned to Tyr-4 and Tyr-5. These additional changes are most developed in the 2:1 complex and coupled with the larger upfield shifts of Phe-1 and Tyr-1 make the difference spectrum of the 2:1 complex almost identical with that of the 1:1 $d(pA)_4$ complex (Figure 2D). This suggests that $d(pA)_2$ has occupied two sequential binding sites, thus achieving a conformation nearly equivalent to that of the $d(pA)_4$ complex. The difference spectra in the central region when Tyr-4 and -5 are represented are difficult to interpret precisely because of the multiple overlapping resonances. In the difference spectrum of the two $d(pA)_2$ complexes (Figure 3D), however, the negative peaks marked A and A′ (occurring at the position of resonances assigned to Tyr 4) suggest that the major change in this region between the 1:1 and 2:1 complexes of $d(pA)_2$

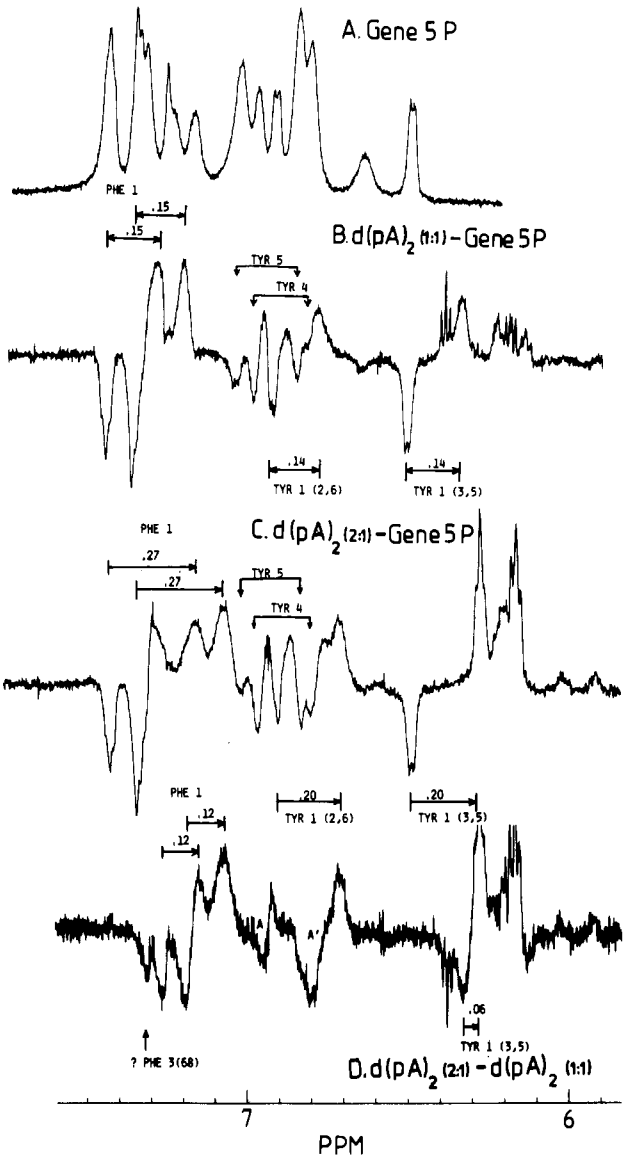

FIGURE 3: ^{1}H aromatic proton difference spectra of the 1:1 and 2:1 complexes of $d(pA)_2$ with gene 5 protein. (A) Spectrum of gene 5 protein for reference. (B) Spectrum of the $d(pA)_2$ complex (1:1) minus that of gene 5 protein. (C) Spectrum of the $d(pA)_2$ complex (2:1) minus that of gene 5 protein. (D) Spectrum of the $d(pA)_2$ complex (2:1) minus that of the $d(pA)_2$ complex (1:1).

may be largely due to upfield shifts of the protons of Tyr-4. A portion of A′ of course represents the further upfield shift of the 2,6 protons of Tyr-1. A small negative peak at the most downfield position of this difference spectrum may reflect a perturbation of a second Phe by the second mole of $d(pA)_2$, a perturbation that also appears to be brought about by $d(pA)_4$ (see below).

Difference Spectra, $d(pA)_n$ Complex minus $d(pA)_{n-1}$ Complex. Since each oligonucleotide from dpA to $d(pA)_4$ appears to form a specific 1:1 complex with gene 5 protein (Figures 1–3), the aromatic protons coming under the influence of the bound bases as each base-binding pocket is successively filled can be determined by forming the following set of difference spectra, dpA minus gene 5 protein, $d(pA)_2$ minus $d(pA)_1$, $d(pA)_3$ minus $d(pA)_2$, and $d(pA)_4$ minus $d(pA)_3$, where the nucleotide letter codes represent the complexes with gene 5 protein. The first member of this set of difference spectra was presented in Figure 2B, while $d(pA)_2$ (1:1) minus $d(pA)_1$, $d(pA)_3$ minus $d(pA)_2$ (1:1), and $d(pA)_4$ minus $d(pA)_3$ are presented in spectra B, C, and D, respectively, of Figure 4.

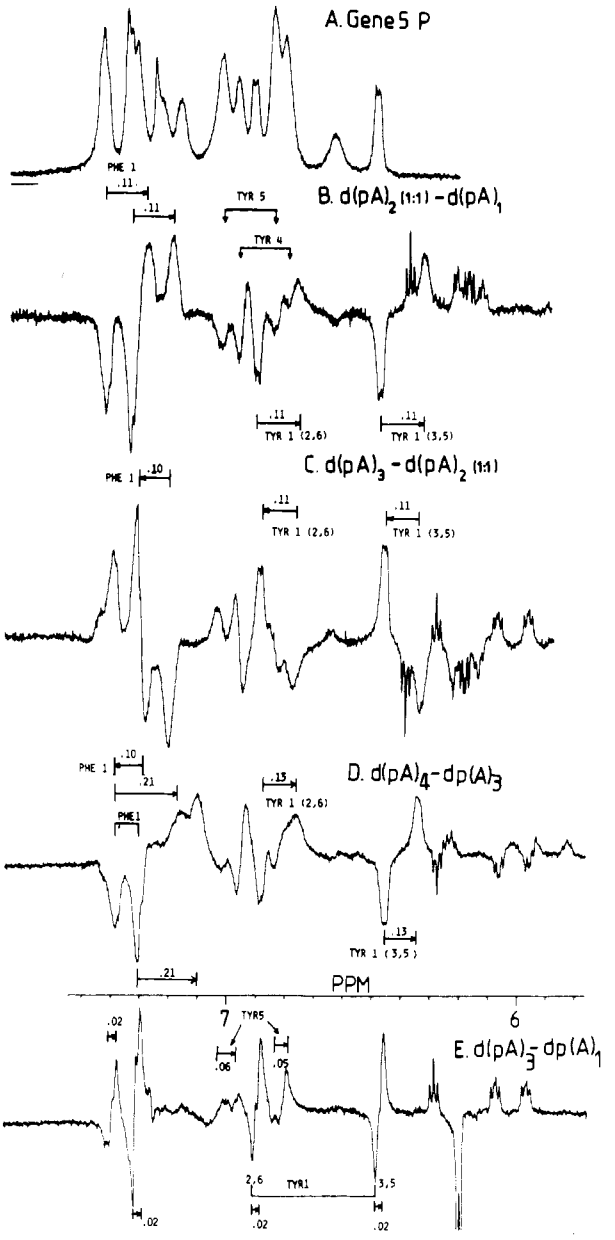

FIGURE 4: ^{1}H aromatic proton NMR difference spectra between successively longer nucleotide complexes, $d(pA)_n - d(pA)_{n-1}$ where the symbols refer to the protein complexes. (A) Spectrum of the gene 5 protein for reference. (B) Spectrum of the dinucleotide complex (1:1) minus that of the mononucleotide complex. (C) Spectrum of the trinucleotide complex minus that of the dinucleotide complex (1:1). (D) Spectrum of the tetranucleotide complex minus that of the trinucleotide complex. (E) Spectrum of the trinucleotide complex minus that of the mononucleotide complex.

The $d(pA)_2$ minus $d(pA)_1$ difference spectrum is similar to the $d(pA)_2$ minus gene 5 protein spectrum, reflecting the very small initial shifts of the dpA (or pA) complex, 0.03–0.04 ppm (Figure 2B). The shifts of Phe-1 and Tyr-1 in Figure 4B are actually reduced by this amount.

As could be inferred from the simple difference spectrum [$d(pA)_3$ minus gene 5 protein] in Figure 2C, the addition of a third base has an unexpected result. As shown in the $d(pA)_3$ minus $d(pA)_2$ (1:1) difference spectrum (Figure 4C), the shifts of Tyr-4 and Tyr-5 are pretty well reversed, and the Phe-1 and Tyr-1 protons are shifted less upfield by $d(pA)_3$ than by $d(pA)_2$ (1:1) (Table I). With addition of the fourth base, however, maximum upfield shifts of the proton resonances from all four residues (Phe-1, Tyr-1, -4, and -5) are observed (Figure 4D). The upfield shifts of the $d(pA)_1$ and $d(pA)_3$ complexes are

Table I: Upfield Shifts (in ppm) of the ^{1}H Resonances Phe-1 and Tyr-1 on Oligonucleotide Binding

nucleotide	Phe-1	Tyr-1	
		2,6 protons	3,5 protons
dpA (or pA)	0.04	0.03	0.03
$d(pA)_2$ (1:1)	0.15	0.14	0.14
$d(pA)_2$ (2:1)	0.27	0.20	0.20
$d(pA)_3$	0.06	0.05	0.05
$d(pA)_4$	0.23	0.15	0.15
$d(pA)_8$	0.64	~0.30	0.32
$d(pT)_8$	0.30[a]		~0.20
$d(A)_8$	0.74[b]	~0.35	0.4

[a] Coleman & Armitage (1978). [b] Alma et al. (1981).

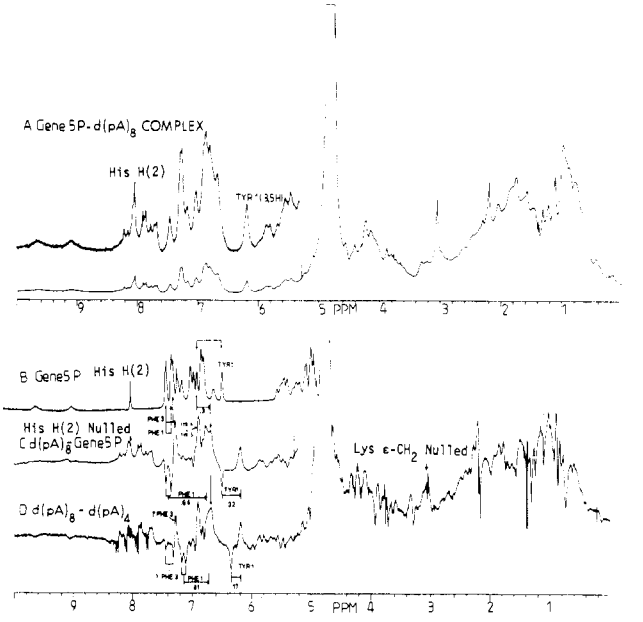

FIGURE 5: 500-MHz ^{1}H NMR spectra of the aromatic and aliphatic protons of the (A) 1:1 complex of gene 5 protein with $d(A)_8$. (B) Aromatic proton spectrum of gene 5 protein alone for reference. (C) Aromatic proton difference spectrum, spectrum of the $d(pA)_8$ complex minus that of the unliganded gene 5 protein. (D) Aromatic and aliphatic proton difference spectrum, spectrum of the $d(pA)_8$ complex minus that of the $d(pA)_4$ complex.

similar and involve predominantly Phe-1 and Tyr-1, although they are larger in the $d(pA)_3$ complex and also appear to involve Tyr-5 as shown in the $d(pA)_3$ minus $d(pA)_1$ difference spectrum in Figure 4E.

^{1}H NMR at 500 MHz of the Gene 5 Protein–$d(pA)_8$ Complex. The complete proton NMR spectrum of the 1:1 molar ratio of $d(pA)_8$ to gene 5 protein is shown in Figure 5A. The resonances of both the aromatic and aliphatic protons of the protein are broader than those in the complexes of pA and $d(pA)_4$. This broadening is most adequately explained by a shift from a dimer of gene 5 protein to a tetramer induced by the octanucleotide (see Discussion). The broadening of the lines makes interpretation of the difference spectra less precise than in the highly resolved spectra presented above. By formation of two difference spectra, i.e., the $d(pA)_8$ complex minus gene 5 protein (Figure 5C) and the $d(pA)_8$ complex minus the $d(pA)_4$ complex (Figure 5D), several shifts of specific residues can be identified. The aromatic ^{1}H spectrum of gene 5 protein is shown in Figure 5B for orientation.

In both difference spectra, the resonance most clearly identifiable is that of the 3,5 protons of Tyr-1 (Figure 5C,D). These protons undergo an upfield shift of 0.32 ppm in the $d(pA)_8$ complex, 0.17 ppm further upfield than their position in the $d(pA)_4$ complex (Figure 5D). The disappearance of the

resonances of Phe-1 from their original position as induced by $d(pA)_8$ is clear in Figure 5C (negative peaks near 7.5 ppm). Likewise, their disappearance from their new position in the $d(pA)_4$ complex on going to the $d(pA)_8$ complex is clear from Figure 5D (negative peaks near 7.15 ppm). The latter, by the way, confirms their assignment in the difference spectra from the $d(pA)_4$ complex (Figures 2D and 4D). The new upfield location of the Phe-1 peaks in the $d(pA)_8$ complex is not so clear, since a large overlapping positive group of peaks from 6.5 to 7.0 ppm represents the shifted protons of Tyr-4, Tyr-5, and Phe-1 and the 2,6 protons of Tyr-1. This group of protons has been assigned in what appears to be the most self-consistent manner, assignments for $d(pA)_8$ which also correspond most closely to assignments made by Alma et al. (1981) in the $d(A)_8$ complex on the basis of NOE data. The most upfield peak near 6.65 ppm has been assigned to the 2,6 protons on Tyr-1, which suggests an upfield shift of $\sim$0.3 ppm, rather close to the $\Delta\delta$ of the 3,5 protons (0.32). The positive area between 6.7 and 6.8 ppm in both difference spectra is then assigned to Phe-1, while the large positive area near 6.9 ppm has been assigned to Tyr-4 and Tyr-5. This assignment requires that $d(pA)_8$ induce an upfield shift of $\sim$0.65 ppm for the protons of Phe-1 and maximum upfield shifts of the unresolved protons on Tyr-4 or Tyr-5 of $\sim$0.2 ppm.

Nulling of the His H(2) in formation of the difference spectrum between the octa- and tetranucleotide complexes automatically nulls the Lys δ-CH_2 resonances (Figure 5D), supporting the notion that complex formation does little to the chemical shift of the Lys protons despite the fact that acetylation suggests that the ϵ-NH_3^+ protons are closely associated with complex formation (Anderson et al., 1975). The rest of the peaks in the aliphatic difference spectrum between the two complexes are difficult to interpret. While the sharp lines probably reflect some conformational changes in aliphatic amino acid side chains characterizing the formation of the octanucleotide vs. the tetranucleotide complex, much of this difference simply reflects the broadening of the resonances in the octanucleotide complex. There may in fact be less conformational change between the tetra- and octanucleotide complex than we earlier concluded (Coleman et al., 1976). The $d(pA)_8$ complex minus gene 5 protein difference spectrum shows additional loss of intensity in the area of the Phe-3 resonances (Figure 5D) which apparently reappears $\sim$0.1 ppm upfield. From the crystal structure, Phe-68 is near the binding groove, and formation of the $d(pA)_8$ complex could perturb Phe-68.

Discussion

In vivo, gene 5 protein binds to a single-stranded DNA of length 6407 bases (Schaller et al., 1978; Beck & Zink, 1981). Each gene 5 monomer covers approximately four bases, and tight binding involves cooperative interactions between adjacent monomers on the chain (Alberts & Frey, 1970; Alberts et al., 1972). The binding is independent of base sequence, although moderate variation in the magnitude of the binding constant has been observed between T- and A-containing oligonucleotides (Coleman et al., 1976). Protein–protein cooperative interactions between three or more sequentially bound gene 5 monomers must account for the great stability of the DNA complexes, since oligonucleotides like $d(A)_4$ and $d(A)_8$, lacking the 5′ dianion and interacting only via the monoionic diester links, bind relatively weakly to gene 5 protein and are in fast exchange with free nucleotide on the ^{31}P NMR time scale (O'Connor & Coleman, 1982). This weak binding occurs despite the fact that the number of residues is sufficient to fill the nucleotide binding lattice on one and two monomers, respectively. In contrast, oligonucleotides carrying a phosphate dianion at the 5′ end bind very tightly to gene 5 protein, are in slow exchange on the ^{31}P NMR time scale (O'Connor & Coleman, 1982), and have dissociation constants of $\sim 10^{-7}$ M for $d(pA)_4$ and $< 5 \times 10^{-8}$ M for $d(pA)_8$ as measured by titration following circular dichroism of the complexes (Coleman et al., 1976). On the basis of ^{31}P chemical shifts of the diester phosphates, the conformation of the diester backbone appears to be identical in the tetra- and octanucleotide complexes (O'Connor & Coleman, 1982).

The crystal structure of the protein shows the molecules to exist as tightly coupled dimers (McPherson et al., 1979, 1980a,b). Hence, the minimal quaternary structure observed in the NMR spectrum must be the dimer with a rotational correlation time, τ_r, of $\sim 7 \times 10^{-8}$ s (Coleman & Armitage, 1978). Since the two DNA binding grooves of the dimer are related by a 2-fold rotation axis, the $d(pA)_4$ complex must involve the interaction of two nucleotides per dimer, one on each monomer. The bound nucleotide should result in little change in τ_r, an expectation borne out by both the aromatic and the aliphatic proton spectra (Figure 1). The spectra in Figures 1 and 3 make it highly unlikely that species larger than the dimer are present in significant amounts. This is at variance with some previous reports that even dinucleotides can induce oligomer formation [see discussion at the end of McPherson et al. (1980b)]. Since the dimer of 2-fold symmetry is the basic unit in solution, octanucleotides must bind two dimers together to form a tetramer, and indeed complex formation with octanucleotides broadens many of the proton lines (Coleman et al., 1976; Alma et al., 1981) (Figure 5), despite the identical configuration of the phosphate diesters in the tetra- and octanucleotide complexes (O&Connor & Coleman, 1982). Thus, the nucleotide must be larger than four residues to induce cooperative side to side interactions of gene 5 monomers.

It is probable that the reason for the favored 1:1 stoichiometry of all the oligonucleotides of four residues or less is that once a 5′-phosphate dianion is bound, the interactions of additional bases in the open-lattice positions are by themselves not strong enough for slow exchange binding. This circumstance does allow the unique opportunity to examine the changes in the proton NMR signals as the base-binding pockets are occupied one at a time as the nucleotide lengthens from one to four residues (Figures 2–4). The binding of one molecule of dpA (or pA) to each gene 5 monomer induces upfield shifts in the resonances assigned to just two aromatic residues, Tyr-1 and Phe-1 (Figure 2). The sensitivity of pA binding to temperature (Figure 1) may reflect a temperature-dependent conformational change in the protein, possibly a change in the mobile domain of the binding groove (see below).

From the chemical modification data (Anderson et al., 1975) and the crystal structure at 2.3-Å resolution (McPherson et al., 1980b; Brayer & McPherson, 1983), the relationship of the aromatic residues to the nucleotide binding domain can be outlined as follows. The only structural feature of the folded polypeptide fitting the dimensions of a DNA binding surface is a groove formed by three short segments of chain constituting a three-stranded antiparallel β-pleated sheet. Side chains of three tyrosyl residues are associated with this DNA binding surface. Tyrosyl-41 defines one end of the groove, while Tyr-26 defines the other. The latter residue is part of a small flexible domain whose residues show large temperature factors in the crystal of the unliganded protein (Brayer & McPherson, 1983). Tyrosyl-34 is located near the center of

the groove, exposed to solvent, but packed against hydrophobic residues in the bottom of the groove. Phe-13 is close to Tyr-34 but points away from the groove and does not appear available for nucleotide interaction. On the other hand, Phe-73, located on one end of a two-stranded β sheet which forms half of the domain which ties the two monomers together, swings over into the DNA binding groove of the opposite monomer and falls between Tyr-26 and -34 (Brayer & McPherson, 1983). In addition, the groove contains near its center Cys-33, whose reaction with Hg^{2+} is prevented by nucleotide binding (Anderson et al., 1975). Cys-33 also cross-links with T-containing nucleotides when their complexes are subject to UV irradiation (Paradiso et al., 1979). The DNA binding groove is lined with a number of Lys and Arg side chains, suggesting that the phosphate backbone is placed in the deepest part of the groove via electrostatic interactions with the positive charges. Acetylation of the Lys residues does destroy nucleotide binding (Anderson et al., 1975).

The phenyl rings of Tyr-26 and -41 are both subject to nitration in the unliganded protein and are protected from nitration by DNA binding (Anderson et al., 1975), confirming that both residues are surface tyrosyls apparently in close contact with the nucleotide. Tyrosyl-34, however, is not subject to nitration. Instead, the third tyrosyl residue subject to nitration is Tyr-56 (Anderson et al., 1975). In the crystal structure, Tyr-56 is located some distance from the groove and is relatively buried, but its edge is exposed to solvent (Brayer & McPherson, 1983). The NMR difference spectra reported here clearly show that only three tyrosyl side chains are involved directly in nucleotide binding. Inspection of the crystal structure suggests that the best candidates for these residues are Tyr-26, -34, and -41.

From extensive comparative studies of nitration of proteins in solution vs. location of tyrosyl residues in the crystal structures of the same proteins, nitration with $C(NO_2)_4$ does not always include all "surface" tyrosyls and may in fact nitrate some that are relatively "buried" (Glazer, 1976). Tyr-34 may escape nitration because of its relatively protected position. On the other hand, Tyr-56, nitrated in the unliganded protein, may be protected by fd DNA binding because of the cooperative "superhelical" complex formation with fd DNA involving side to side interactions of gene 5 oligomers. ^{19}F NMR of the fluorotyrosyl derivative of gene 5 protein shows the ^{19}F resonances to occur in two groups, two signals downfield of that for "free" fluorotyrosine and three relatively upfield near the resonance expected for free fluorotyrosine (Coleman et al., 1976). The ^{19}F NMR spectrum is probably a good gauge of the separation of two buried and three relatively accessible tyrosyl residues but does not of course identify the three whose ^{19}F resonances occur upfield as those nitratable.

The base adjacent to the 5′-phosphate in these oligonucleotides is responsible for the upfield shifts of the protons on one Tyr and one Phe ring, since they are shifted by dpA and all other oligonucleotides (Table I). This particular tyrosyl is the one tyrosyl in gene 5 that must initially have relatively complete rotational freedom, since its protons resonate as a well-defined doublet of doublets, even in the complexes with the shorter nucleotides (Figures 2 and 4). On the basis of the refined crystal structure, the best candidate for this residue is Tyr-26. The small mobile domain containing this residue also involves Lys-24, Arg-16, and Arg-21 (Brayer & McPherson, 1983). In solution, this domain may not be completely fixed until the nucleotide binds. The other aromatic residue that the ring of Tyr-26 can closely approach is Phe-73 from the opposite monomer. This may account for the relatively upfield positions of the 3,5 protons of Tyr-26 in the unliganded protein (Figure 1). The coupled upfield movement of the protons on Tyr-1 and Phe-1 accompanying nucleotide complex formation would thus suggest that these protons are assignable to Tyr-26 and Phe-73. The progressive increase in the magnitude of these upfield shifts as the nucleotide is lengthened (Table I) might relate to conformational changes as the mobile domain containing Tyr-26 is anchored to the rest of the monomer by the longer oligonucleotides. The ^1H NMR spectra of the dpA (or pA) complex (Figure 2) then identify Tyr-26 as forming the 5′ end of the binding groove, while Tyr-41 is at the 3′ end. The positive charges of Lys-24, Arg-16, and Arg-21 could form the high-affinity site for the 5′-phosphate dianion.

Among this set of oligodeoxynucleotides, $d(pA)_2$ is the only one that forms a 2:1 complex of sufficient stability for the second nucleotide to cause changes in the aromatic proton NMR of the protein (Figure 3). This appears to relate to the simultaneous occupancy of all four base-binding pockets by two sequential dinucleotides. The ^{31}P NMR identifies a 1:1 complex of $d(pA)_2$ in slow exchange ($\Delta\delta$ of the 5′-phosphate $\simeq$ 1.5), while a second mole of $d(pA)_2$ is bound in the fast exchange condition with respect to the ^{31}P chemical shift (O'Connor & Coleman, 1982). The latter shows an upfield shift of the diester phosphate with little change in the chemical shift of the 5′-phosphate. The third base-binding pocket alone may have little affinity for the nucleotide, since $d(pA)_3$ unexpectedly does not appear to be as tightly tied down as either the 1:1 or the 2:1 $d(pA)_2$ complexes (Figures 2 and 3). Occupancy of the second, third, and fourth base-binding pockets progressively shifts the protons of Tyr-4 and -5 upfield, Tyr-5 before Tyr-4 (Figures 3 and 4). This suggests that Tyr-4 is residue 41 which defines the opposite end of the DNA binding groove.

Giessner-Prettre et al. (1976) have presented detailed contour plots of the ring-current shifts expected for protons located at various distances above adenine rings and at various latitudes relative to the centers of the five- and six-membered rings. If one assumes a stacking distance of $\sim$3.5 Å and varies the position of Tyr or Phe protons relative to the adenine ring, the upfield shifts of Table I would suggest that the average positions of the aromatic protons in the dpA complex are as far as 3–4 Å from the center of the adenine ring; i.e., the rings do not directly overlap. In the $d(pA)_4$ complex, the distance of the tyrosyl protons from the center of the adenine ring must be closer to 2.5 Å, depending on whether the overlap is primarily with the five- or the six-membered ring; i.e., the rings are just barely overlapping. In the $d(pA)_8$ complex, however, the aromatic protons must be overlapped perhaps within 1.5 Å of the center to achieve a shift of $\sim$0.6 ppm. If the stacking distance were larger than 3.5 Å, the overlap could be considerably greater, since the ring-current shift is steeply dependent on stacking distances below 4 Å. The increasing shifts as the nucleotide lengthens could also be explained by a progressive decrease in the stacking distance between the bases and Tyr-26 and Phe-73 as the nucleotide lengthens. The H(2) and H(8) protons of the base rings are subject to the relatively smaller fields from the tyrosyl and phenylalanyl ring currents and may also be some distance from the aromatic rings because of their 2–3-Å distance from the centers of the base rings. These protons undergo only slight chemical shifts on binding to gene 5 protein (Figure 1).

The present proton NMR of the nucleotide complexes provides strong support for the sequential distribution of Tyr-26, Phe-73, Tyr-34, and Tyr-41 along the DNA binding

groove as suggested from the crystal structure of the unliganded protein. Hydrophobic interactions between the incoming bases and suitably placed aromatic rings appear to be an attractive way of stabilizing nucleotide–protein complexes. In the gene 5 protein complexes with the longer oligonucleotides, this interaction appears to become actual intercalation of the ring which probably also characterizes the DNA complexes. The crystal structure of ribonuclease T_1, specific for guanylic acid, shows the ring of Tyr-45 at the active center of this enzyme to be stacked over the six-membered ring of guanine at 3.5 Å (Heinemann & Saenger, 1982). Whether intercalation as observed in the gene 5 protein–nucleotide complexes forms a general feature of the binding of helix-destabilizing proteins to single-stranded DNA requires exploration of other examples by NMR techniques.

Acknowledgments

We thank Dr. Gary Brayer and Professor Alex McPherson for supplying us with their most recent fitting of the electron density map of gene 5 protein prior to publication.

References

Alberts, B. M., & Frey, L. (1970) *Nature (London) 227*, 1313–1318.
Alberts, B. M., & Herrick, G. (1971) *Methods Enzymol. 21*, 198–217.
Alberts, B. M., Frey, L., & Delius, H. (1972) *J. Mol. Biol. 68*, 139–152.
Alma, N. C. M., Harmsen, B. J. M., Hull, W. E., van der Marel, G., van Boom, J. H., & Hilbers, C. W. (1981) *Biochemistry 20*, 4419–4428.
Anderson, R. A., Nakashima, Y., & Coleman, J. E. (1975) *Biochemistry 14*, 907–917.
Beck, E., & Zink, B. (1981) *Gene 16*, 35–58.
Brayer, G. D., & McPherson, A. (1983) *J. Mol. Biol.* (in press).
Coleman, J. E., & Armitage, I. M. (1977) in *NMR in Biology* (Dwek, R. A., Campbell, I. D., Richards, R. E., & Williams, R. J. P., Eds.) p 171, Academic Press, New York.
Coleman, J. E., & Armitage, I. M. (1978) *Biochemistry 17*, 5038–5045.
Coleman, J. E., Anderson, R. A., Ratcliffe, R. G., & Armitage, I. M. (1976) *Biochemistry 15*, 5419–5430.
Garssen, G. J., Hilbers, C. W., Schoenmakers, J. G. G., & van Boom, J. H. (1977) *Eur. J. Biochem. 81*, 453–463.
Garssen, G. J., Tesser, G. I., Schoenmakers, J. G. G., & Hilbers, C. W. (1980) *Biochim. Biophys. Acta 607*, 361–371.
Giessner-Prettre, C., Pullman, B., Borer, P. N., Kan, L.-S., & Ts'o, P. O. P. (1976) *Biopolymers 15*, 2277–2286.
Glazer, A. N. (1976) *Proteins (3rd Ed.) 2*, 1–103.
Heinemann, U., & Saenger, W. (1982) *Nature (London) 299*, 27–31.
Kornberg, A. (1980) *DNA Replication*, pp 479–496, W. H. Freeman, San Francisco, CA.
Kornberg, A. (1982) *Supplement to DNA Replication*, p 571, W. H. Freeman, San Francisco, CA.
McPherson, A., Jurnak, F. A., Wang, A. H. J., Molineux, I., & Rich, A. (1979) *J. Mol. Biol. 134*, 379–400.
McPherson, A., Wang, A. H. J., Jurnak, F. A., Molineux, I., Kolpak, F., & Rich, A. (1980a) *J. Biol. Chem. 255*, 3174–3177.
McPherson, A., Jurnak, F. A., Wang, A., Kolpak, F., Rich, A., Molineux, I., & Fitzgerald, P. (1980b) *Biophys. J. 32*, 155–173.
O'Connor, T. P., & Coleman, J. E. (1982) *Biochemistry 21*, 848–854.
Paradiso, P., Nakashima, Y., & Konigsberg, W. (1979) *J. Biol. Chem. 254*, 4739–4744.
Schaller, H., Beck, E., & Takanami, M. (1978) in *The Single-Stranded DNA Phages*, pp 139–163, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.

Heterogeneity of Albumin As Detected by Its Interactions with Deuteroporphyrin IX†

Gregory A. Moehring, Amy H. Chu, Leon Kurlansik, and Taffy J. Williams*

ABSTRACT: The interaction of deuteroporphyrin IX with human serum albumin was examined by absorption and fluorescence methods. Changes in the porphyrin fluorescence polarization were utilized in the determination of the association constant of deuteroporphyrin IX binding to albumin. The results were consistent with a model where human albumin exists in at least two conformers, with 37% having an association constant of 4.85×10^8 M^{-1} and 63% a value of 1.47 $\times 10^7$ M^{-1}. Analysis of the temperature dependence of binding yielded enthalpy terms found to be of opposite sign, high affinity having a value of $\Delta H = -4.77$ kcal/mol and the lower affinity having a value of $\Delta H = 2.70$ kcal/mol. The comparison of thermodynamic values, spectral shifts, and iodide quench curves suggests that the porphyrin binding site is hydrophobic in character and the bound porphyrin has limited access to at least some of the ions in solution.

In order to better understand the role of human serum albumin (HSA) in the in vivo clearance of porphyrins and the relationship of albumin to other serum proteins involved in the in vivo clearance of porphyrins, many studies have been reported that attempt to characterize the interactions between porphyrins and HSA (Muller-Eberhard & Morgan, 1975; Beaven et al., 1974; Lamola et al., 1981; Morgan et al., 1980; Reddi et al., 1981; Rosenfeld & Surgenor, 1950; Adams & Berman, 1980; Little & Neilands, 1960). Most of these investigations have utilized absorption and fluorescence spectroscopy measurements and the method of Halfman & Nishida

† From the Naval Medical Research Institute, Bethesda, Maryland 20814. *Received October 26, 1982; revised manuscript received March 4, 1983.* This research was supported by the Naval Medical Research and Development Command, Research Work Unit No. MF 58524013.1044.