Evidence of Novel Secondary Structure in DNA-Bound Protamine Is Revealed by Raman Spectroscopy†

Nicholas V. Hud,*,‡ Fred P. Milanovich,§ and Rod Balhorn‡

Biology and Biotechnology Research Program and Physics Division, Lawrence Livermore National Laboratory, Livermore, California 94551

Received December 27, 1993; Revised Manuscript Received March 25, 1994®

ABSTRACT: Raman spectroscopy studies of protamine–DNA complexes are reported for samples in the solid state at 98% relative humidity. Previous reports utilizing other physical techniques have indicated the presence of B-form DNA in protamine–DNA complexes. The present Raman data support the assignment of a modified B-form which is characterized by appreciable unstacking of the bases. The quality of the present spectra has made it possible, for the first time, to obtain the Raman spectrum of DNA-bound protamine by digital spectral subtraction. The difference spectrum indicates that protamine adopts an unusual secondary structure upon binding to DNA. A dominant amide I band is observed at 1683 cm^{-1} which is indicative of neither an α-helix or β-sheet conformation. An amide I band at this position has been associated with the 1→3 hydrogen bond that occurs within a γ-turn [Bandekar, J., & Krimm, S. (1985) *Int. J. Pept. Protein Res. 26,* 158–165]. On the basis of this assignment, as well as preliminary results obtained by computer modeling, we propose a new model for the secondary structure of DNA-bound protamine that is rich in 1→3 hydrogen bonding. Spectral data demonstrate that this structure is absent in protamine molecules in solution. Analyses of spectra of polyarginine–DNA complexes suggest that polyarginine, although similar to protamine in primary structure, assumes a conformation when bound to DNA that is distinct from that adopted by protamine.

The sperm cells of many vertebrates utilize small arginine-rich proteins, known as protamines, for the packaging of DNA (Ando et al., 1973). Protamine protects DNA from enzymatic degradation and suppresses genetic activities, such as transcription, replication, and DNA repair (Chandley & Kofman-Alfaro, 1971; Sega, 1974; Kierszenbaum & Tres, 1975). During sperm cell maturation, the replacement of histones by protamines yields the most condensed form of chromatin observed in eukaryotic cells. Volume measurements of mammalian sperm cell nuclei, for example, have shown the level of DNA condensation to be approximately 40-fold greater than that in somatic cell nuclei (Pogany et al., 1981). Studies of sperm chromatin have revealed that this high level of compaction is made possible by the charge neutralization that occurs when protamines form a complex with DNA (Felix, 1953).

Biochemical studies of sperm cell nuclei began over a century ago (Miescher, 1874, 1897) and resulted in the identification of protein and nucleic acid as the predominant nuclear components (Kossel, 1928). These early studies and many others have focused on the protamines of fish. The amino acid sequence of salmon protamine (salmine A1; Ando & Watanabe, 1969)

NH_3^+-PRRRRSSSRPVRRRRRPRVSRRRRRR-
GGRRRR-COO^-

is representative of these widely studied protamines. While the protamines of other species, such as man, differ significantly in length and sequence, the most distinguishing characteristic of protamines (regions of consecutive arginine residues, or "arginine clusters") is common throughout all species.

Physical investigations of the protamine–DNA complex were initiated almost 40 years ago by Wilkins and collaborators (Feughelman et al., 1955; Wilkins, 1956). These X-ray fiber diffraction studies determined that DNA complexed with protamine remains in the B-form and provided evidence for a model of protamines in an extended conformation bound within the minor groove. Subsequent diffraction studies have reinforced the principle aspects of this model, along with providing a compelling argument for the localization of protamines within the major groove (Fita et al., 1983).

The periodic structure of DNA, particularly the regularity in positioning of phosphates along the sugar–phosphate backbone, has inspired many investigators to hypothesize that the similarly periodic arginine residues in protamine bind DNA in a regular repeating secondary structure. However, due to the limited information provided by fiber diffraction patterns and the lack of a protamine–DNA complex crystal, the secondary structure of protamines while bound to DNA has not been resolved. The only single crystal of a protamine–nucleic acid complex reported to date is a protamine–tRNA complex (Warrant & Kim, 1978). Electron density maps of this crystal have been interpreted to suggest that protamines bind to DNA in the α-helix conformation. This assignment, however, is in opposition to infrared spectroscopy and deuterium-exchange studies of the complex that do not indicate the presence of α-helix or β-sheet secondary structure in DNA-bound protamine (Bradbury et al., 1962; Herskovits & Brahms, 1976).

Raman spectroscopy is a physical technique that can be used to identify the secondary structure of proteins and nucleic acids (Tu, 1982; Spiro, 1987). In this paper we present the Raman spectrum of salmine (salmon protamine) complexed

† This work was performed at the Lawrence Livermore National Laboratory under the auspices of the U.S. Department of Energy and supported by Contract W-7405-ENG-48 and Laboratory Directed Research and Development Award 93-DI-003.

* To whom correspondence should be addressed.
‡ Biology and Biotechnology Research Program.
§ Physics Division.

® Abstract published in *Advance ACS Abstracts,* June 1, 1994.

0006-2960/94/0433-7528$04.50/0

with DNA in the solid state at 98% rh.[1] The quality of the present Raman spectra and the use of digital spectral subtraction have allowed us to analyze the spectrum of DNA-bound salmine and obtain information about its conformation. While the Raman spectrum of a "soluble" protamine–DNA complex has already been reported in the literature (Herskovits & Brahms, 1976), the natural insolubility of sperm chromatin and a thorough examination of the reported spectra have caused us to question the validity of conclusions drawn from studies of soluble complexes.

Perhaps the simplest peptide analog to protamine is polyarginine. Complexes formed by the binding of arginine-rich peptides (including polyarginine) to DNA have been studied for some time as a model of the protamine–DNA complex (Suwalsky & Traub, 1972; Mansy et al., 1976; Fita et al., 1983). While these studies have provided some insight into the nature of the protamine–DNA complex, the secondary structure of polyarginine while bound to DNA has not been resolved. In this study we also report the Raman spectrum of the polyarginine–DNA complex at 98% rh and, through digital spectral subtraction, the DNA-bound polyarginine spectrum.

MATERIALS AND METHODS

Sample Preparation for Raman Spectroscopy. Salmon protamine (protamine sulfate, grade X) and polyarginine (MW 10 000) were purchased from Sigma Chemical Co. (St. Louis, MO) and dialyzed extensively against 150 mM NaCl, followed by 10 mM HCl, to remove sulfate counterions. Samples were then lyophilized and resuspended in appropriate buffers. To obtain solution spectra, salmine and polyarginine were prepared at approximately 30 mg/mL in dH_2O and brought to neutral pH by titration with 100 mM NaOH.

The salmine–DNA complex was prepared by the mixing of 1 mg/mL salmine and calf thymus DNA (Worthington Biochemical Co., Freehold, NJ) solutions in the presence of 1 M NaCl and 10 mM Tris buffer (pH 8) at a ratio of 1.5 mg of protamine/mg of DNA. At this ionic strength, no precipitate formed upon mixing. The samples were subsequently transferred into dialysis tubing (MW 1000; Spectrum Medical Industries, Los Angeles, CA) and dialyzed against 150 mM NaCl and 10 mM Tris (pH 8) for at least 4 h. After removal from dialysis, the precipitate (salmine–DNA complex) was washed with distilled water to remove excess salt and unbound protamine. This step is required to eliminate the possible generation of high salt concentrations during the drying process (Suau & Subirana, 1977). The polyarginine–DNA complex was prepared in an identical manner. The sodium and lithium salts of calf thymus DNA (NaDNA and LiDNA) were prepared by dialyzing DNA extensively against the appropriate chloride salts, followed by ethanol precipitation. The NaDNA sample was rehydrated with a NaCl solution to produce a sample containing 12% salt by weight. In the case of the LiDNA sample, a low-salt sample was desired; no additional salt was added prior to spectral collection. Solid samples were dried and pressed into pellets to provide a uniform surface for laser excitation.

Raman Spectral Collection. Samples were irradiated with the 488-nm line of an argon-ion laser (Spectra Physics, Model 160), and Raman spectra were collected by counting amplified and discriminated events from a photomultiplier tube (RCA, Model 2805). Wavelength selection was accomplished by a double monochromator (Spex, Model 1408). The spectral resolution of this system, with the slits used, is estimated to be 8 cm^{-1}. Photons were counted at every 1 cm^{-1} with an integration time of 2 s. Up to 12 spectra were collected consecutively for each sample and inspected to verify sample and system stability before averaging.

During spectral collection, the salmine–DNA, polyarginine–DNA, and NaDNA samples were sealed in a closed chamber

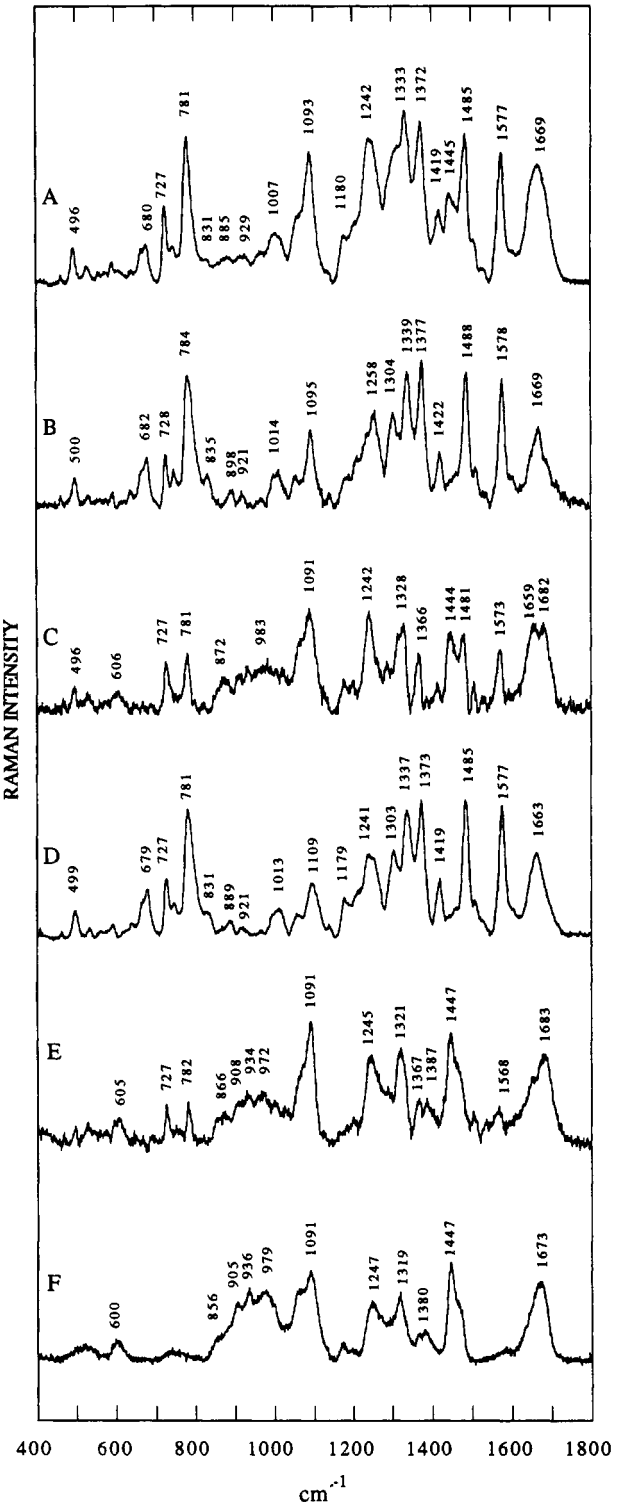

FIGURE 1: Raman spectra in the 400–1800-cm^{-1} region of (A) the salmine–DNA complex at 98% rh, (B) NaDNA at 98% rh, (C) the salmine–DNA complex – NaDNA difference spectra [3-fold amplification of (B) subtracted from (A)], (D) low-salt LiDNA at 95% rh, (E) the salmine–DNA – low-salt LiDNA difference spectrum [3-fold amplification of (D) subtracted from (A)], and (F) salmine in solution (3%) at neutral pH. All spectra were taken at 20 °C and are baseline corrected.

[1] Abbreviations: rh, relative humidity; dH_2O, distilled water.

at 20 °C and equilibrated to 98% rh using a saturated solution of $CaSO_4$. The LiDNA spectrum was taken at 20 °C and 95% rh using a saturated solution of Na_2HPO_4. Line excitation was employed by focusing the laser beam onto the solid samples with a cylindrical lens. Total laser power at the sample surface was limited to approximately 30 mW. Solution samples were placed in a sealed cuvette and excited with 70 mW of laser power.

Spectral Analysis. Solvent effects were subtracted out of the salmine and polyarginine solution spectra using reference spectra. Fluorescent backgrounds from solid samples were removed using a simple linear fit. For subtraction of the DNA spectra from the salmine- and polyarginine–DNA spectra, the 1577-cm^{-1} band served as an internal intensity standard, because it is in general conformationally insensitive and free of underlying salmine and polyarginine bands. Spectra are presented without smoothing.

RESULTS AND DISCUSSION

Raman Spectroscopy of the Salmine–DNA Complex

The Raman spectrum of the salmine–DNA complex at 98% rh is shown in Figure 1 for the region 400–1800 cm^{-1}. It is clear that this spectrum is dominated by the vibrational bands of DNA. This is to be expected because DNA is a more intense Raman scatterer than protein and the percentage of protein (by weight) in this complex is less than 40% (Ando et al., 1973). While many of the DNA bands can be analyzed quite well in the nucleic acid–protein complex spectrum, the only clearly visible protein band is the methylene deformation band at 1445 cm^{-1}. The absence of this methylene band in the Raman spectra of soluble protamine–DNA complexes reported earlier in the literature indicates that these "complexes" do not contain sufficient protamine to be considered similar to sperm chromatin.

Analysis of the Salmine–DNA Complex Spectrum. A vibrational band of significant importance in the determination of DNA secondary structure is the guanine ring breathing mode in the 600–700-cm^{-1} vicinity of DNA spectra. It has been shown that the energy of this vibration is sensitive to the angle of the glycosidic bond between the base and the deoxyribose sugar and to the conformation of the sugar ring itself (Nishimura et al., 1983). In particular, the appearance of this band at 625, 665, and 682 cm^{-1} has been assigned to the C3′ endo-syn (Z-form), C3′ endo-anti (A-form), and C2′ endo-anti (B-form) sugar–phosphate backbone conformations (Nishimura et al., 1986; Benevides & Thomas, 1983), respectively. Thus, the 680-cm^{-1} band in the salmine–DNA complex spectrum and its dominance over the thymine band at 668 cm^{-1} suggest that the sugars of the guanine bases have puckers and glycosidic bond geometries similar to those of B-form DNA. This band does, however, appear to be less intense than in the Raman spectrum of B-form DNA (Figure 1B). A decrease in the intensity of this band has been observed in melting studies of calf thymus DNA (Erfurth & Peticolas, 1975). In the present study this indicates the possible unstacking of the guanine bases upon formation of the salmine–DNA complex.

Comparison of the 727-cm^{-1} adenine vibrational band in the complex spectrum with the respective band in a B-form DNA spectrum reveals that a significant enhancement of this vibrational mode occurs when DNA is bound by salmine. A similar intensity increase accompanying the helix to coil transition (Raman hypochromism; Tomlinson & Peticolas, 1970) has been well documented (Erfurth & Peticolas, 1975; Benevides et al., 1991a). The quantitative study by Benevides et al. has shown an 84% increase in the area of this band upon complete melting of calf thymus DNA. We calculate a 47% increase in the case of the salmine–DNA complex, with respect to B-form DNA. Assuming that the sample is homogeneous, that is, with no local regions of complete melting, the hypochromism of the 727-cm^{-1} band would then indicate that the adenine bases are in contact with neighboring bases but with a considerable degree of unstacking.

The symmetric and antisymmetric O–P–O diester stretching modes have been shown to be greatly influenced by the secondary structure of DNA. In particular, normal mode analysis has shown that the antisymmetric O–P–O diester stretch exhibits a broad low-intensity band at 834 cm^{-1} when the sugars are in the C2′ endo conformation (Brown & Peticolas, 1975; Lu et al., 1977). Alternately, when the sugars are in the C3′ endo conformation, the symmetric stretch exhibits a narrow intense band at 807 cm^{-1}. The salmine–DNA complex spectrum contains a small band at 831 cm^{-1}, which indicates the presence of B-form DNA. It should be noted that this band is, however, shifted slightly and is of considerably reduced intensity in comparison to spectra of B-form DNA. Perhaps of equal significance to the 831-cm^{-1} band is the lack of an 807-cm^{-1} band, which rules out the presence of appreciable A-form backbone conformation.

In the spectra of both A- and B-form DNA the thymine vibrational band normally observed near 1242 cm^{-1} is dominated by the nearby adenine–cytosine vibrational band at 1259 cm^{-1}. In the spectrum of the salmine–DNA complex, however, the 1242-cm^{-1} band appears to be significantly enhanced. This is an additional indication of base unstacking (Erfurth & Peticolas, 1975; Benevides et al., 1991a) and is consistent with the observed enhancement of the 727-cm^{-1} adenine band. Quantization of the enhancement of the 1242-cm^{-1} band in the salmine–DNA complex would be tenuous because the salmine amide III band lies within this same region.

The 1333-cm^{-1} band is assigned to a vibrational mode of adenine. In this position it has been cited as a marker band for C3′ endo-anti (Thomas, 1986). To be an indicator of C2′ endo-anti, it should appear closer to 1339 cm^{-1}. The absence of an 807-cm^{-1} vibrational band suggests that the position of the 1333-cm^{-1} band in the salmine–DNA complex is not due to C3′ endo-anti backbone conformation but rather to alterations in the environment or stacking of the adenine bases. Near this band, an additional adenine band is typically well resolved at 1303 cm^{-1}. In this spectrum it appears as a shoulder on the 1333-cm^{-1} band. This may be due to the presence of the spectral band from the CH_2 twisting mode of the protein side chains observed near 1320 cm^{-1}.

Analysis of Salmine–DNA Complex Subtraction Spectra. Subtraction of a B-form DNA spectrum (NaDNA at 98% rh) from the salmine–DNA complex spectrum allows further identification of alterations in DNA structure upon complex formation (Figure 1C). In the difference spectrum the previously discussed increases in the 727- and 1240-cm^{-1} bands are clearly visible. Additionally, the positive 1480- and 1572-cm^{-1} subtraction bands, which receive contributions from both of the purine bases, are again indicative of base unstacking. In B-form DNA spectra the 784-cm^{-1} band receives contributions to its intensity from both the cytosine bases and from the sugar–phosphate backbone (O–P–O symmetric stretch). Thus, while the positive band at 781 cm^{-1} in the subtraction spectrum is consistent with cytosine base unstacking, alterations in the intensity of this band could also be due to perturbations of backbone vibrational modes resulting from

arginine–phosphate salt bridge formation. The 1657-cm^{-1} band represents a decrease in energy of the thymine carbonyl stretch (normally at 1673 cm^{-1}) which, although consistent with unstacking of the thymine bases, could also be indicative of alterations in thymine hydrogen bonding (Erfurth & Peticolas, 1975).

While several protein bands were revealed when the B-form DNA spectrum was subtracted from the salmine–DNA complex spectrum, the appearance of DNA bands in the amide I and amide III regions of the subtraction spectrum makes it difficult to analyze the conformation of DNA-bound salmine. This problem is obviously due to the fact that the DNA in the salmine–DNA complex is not exactly B-form. A comparison with spectra of DNA samples of known structure suggests that the DNA within the salmine–DNA complex exhibits features common to both the B- and C-forms. Under low-salt and/or low-humidity conditions certain salts of DNA are known to assume conformations which are intermediate between the B- and C-forms (Marvin et al., 1961; Portugal & Subirana, 1985). We have recorded a Raman spectrum of a low-salt LiDNA sample at 95% (Figure 1D). This spectrum is not identical to the salmine–DNA complex spectrum, but it is more similar to the complex spectrum than either B- or C-form DNA spectra. The spectrum resulting from the subtraction of the LiDNA spectrum from the salmine–DNA spectrum is shown as Figure 1E. The degree to which this subtraction spectrum represents the spectrum of DNA-bound salmine can be appreciated by a comparison with the salmine solution spectrum (Figure 1F).

In Raman spectroscopy studies of proteins, the amide I band serves as the most significant indicator of secondary structure (Williams & Dunker, 1981; Williams, 1983; Krimm, 1987). Comparison of the salmine–DNA subtraction spectra which result from using the NaDNA and the LiDNA spectra as the subtrahends (spectra C and E of Figure 1, respectively) shows that the amide I band of DNA-bound salmine is at 1683 cm^{-1}, whereas the 1659-cm^{-1} shoulder is apparently a residual of the shifted thymine band. The sharpness of the amide I component suggests the presence of a considerable amount of a regular secondary structure. However, while amide I bands near 1655 and 1670 cm^{-1} have been correlated with the α-helix and β-sheet structures, respectively, a band at 1683 cm^{-1} has not been assigned to any known regular secondary structure. Thus, secondary structure assignment for DNA-bound salmine is not straightforward. It is clear from the position of the DNA-bound salmine amide I band, when compared to the salmine solution spectra amide I band at 1673 cm^{-1}, that the secondary structure of salmine changes considerably when it binds to DNA. Significant changes in the secondary structure of transcription proteins upon binding to DNA have been previously reported in the literature (Patel et al., 1990; Weiss et al., 1990); here we observe a similar result for a protein that binds to DNA in a non-sequence-specific manner.

The Raman spectra of bovine insulin and the cyclic peptide D-Phe-L-Pro-Gly-D-Ala-L-Pro both contain amide I bands at 1685 cm^{-1} (Yu et al., 1972; Bandekar & Krimm, 1985). We have examined the data available for each of these in an effort to assign the secondary structure of DNA-bound salmine. In the case of insulin, the origin of the 1685-cm^{-1} band has not been satisfactorily determined. Yu et al. first associated the 1685-cm^{-1} band with random coil, while later analysis suggested its assignment to non-hydrogen-bonded β-sheet or α-helix capping turns (Williams, 1983). The non-hydrogen-bonded β-sheet would be consistent with the extended structure

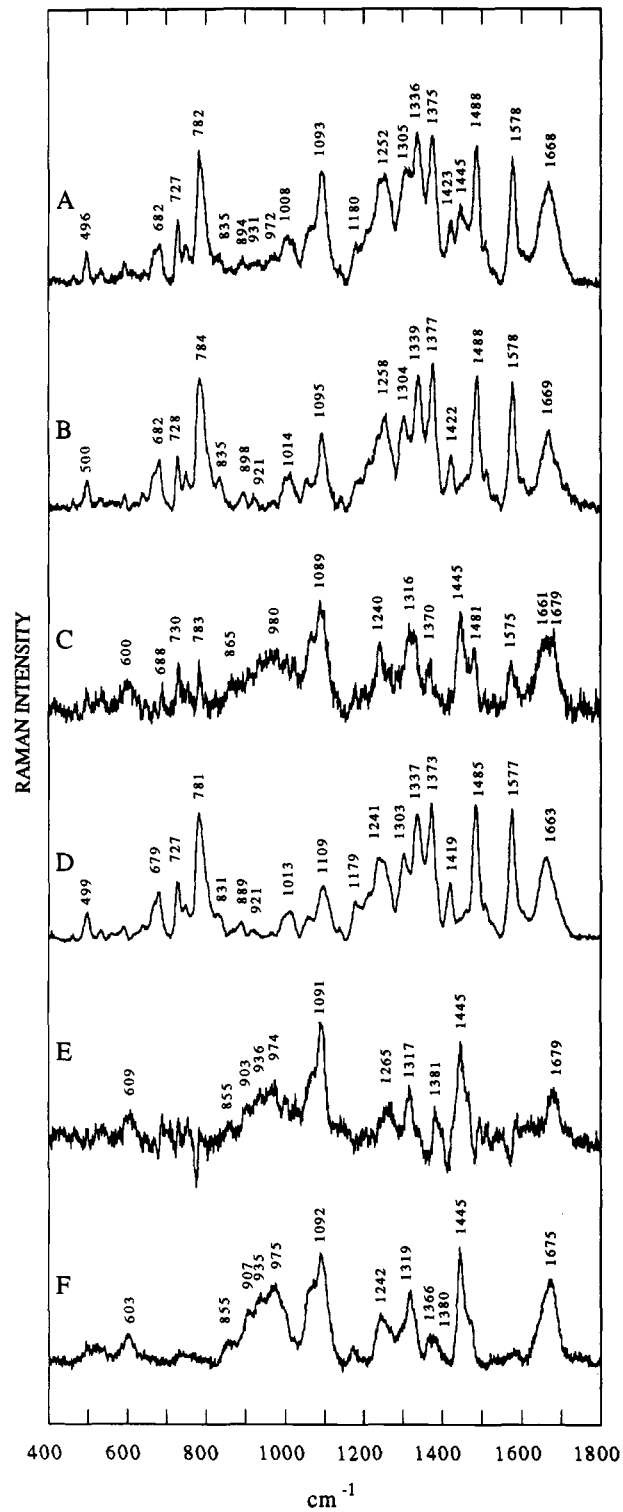

FIGURE 2: Raman spectra in the 400–1800-cm^{-1} region of (A) the polyarginine–DNA complex at 98% rh, (B) NaDNA at 98% rh, (C) the polyarginine–DNA complex – NaDNA difference spectra [3-fold amplification of (B) subtracted from (A)], (D) low-salt LiDNA at 95% rh, (E) the polyarginine–DNA – low-salt LiDNA difference spectrum [3-fold amplification of (D) subtracted from (A)], and (F) polyarginine in solution (3%) at neutral pH. All spectra were taken at 20 °C and are baseline corrected.

of protamine as first suggested by Feughelman et al. However, we are not presently convinced that the 1685-cm^{-1} band should be assigned to this structure, primarily because a comparison of the secondary structure and hydrogen bonds of insulin with proteins which lack such an energetic amide I band does not appear to support this assignment (Hud, unpublished data). In the case of the cyclic peptide, the origin of the 1685-cm^{-1}

band has been determined by normal mode analysis to originate from a peptide group involved in forming the 1→3 hydrogen bond that stabilizes a γ-turn in this peptide (Bandekar & Krimm, 1985). The possible presence of 1→3 hydrogen bonds in the secondary structure of DNA-bound protamine is intriguing because a regular secondary structure containing multiple 1→3 hydrogen bonds has never been observed.

The amide III band, which typically appears between 1230 and 1290 cm^{-1}, has also been used for assigning protein secondary structure. However, this band is sensitive to side-chain composition (Krimm & Bandekar, 1986), and it is not generally considered to be as reliable for structure assignment as the amide I band. In the solution and DNA-bound salmine spectra, the amide III band appears at 1247 and 1245 cm^{-1}, respectively. Neither position, though minimally separated, suggests a particular secondary structure.

The arginine side-chain vibrational band at 1091 cm^{-1} in the salmine solution spectrum appears to be enhanced in the DNA-bound salmine subtraction spectrum (Figure 1E). This is most likely an artifact resulting from the subtraction of the low-salt LiDNA spectrum, which contains a significantly reduced PO_2 symmetric stretch at 1097 cm^{-1}, representing the partial B- to C-form transition. While one might expect to observe a change in the 1091-cm^{-1} arginine band upon complex formation, this would be difficult to quantify since the DNA phosphate band at the same position would also be expected to experience a reduction in intensity, as has been observed in the studies of phage DNA packaged with high local concentrations of Mg^{2+} (Aubrey et al., 1992).

The N–C–C_α–C_β skeletal stretch of polypeptides typically exhibits a band in the 900–1000-cm^{-1} region of the Raman spectrum. The energy of this vibrational mode has been found to depend on secondary structure for polypeptides with side chains larger than CH_3 (Sengupta & Krimm, 1987). In the salmine solution spectrum (Figure 1F) the 979-cm^{-1} band is assigned to the N–C–C_α–C_β skeletal stretch. Three smaller bands are also apparent in the region between 850 and 936 cm^{-1}. These originate from arginine side-chain vibrational modes and are conserved in the DNA-bound salmine spectrum (Figure 1E). The 979-cm^{-1} band, however, appears to be significantly broadened or of reduced intensity in the DNA-bound state. This may not reflect a loss of regular secondary structure, because the amide I band for DNA-bound salmine is approximately as sharp as it is in the solution-state spectrum. Benevides et al. (1991b) suggested that changes in this band, without corresponding changes in the amide I band, may be the result of conserved secondary structure with changes in N–C–C_α–C_β linkages. In the present case, this could indicate that van der Waals contacts or the extensions of the arginine side chains of DNA-bound salmine are heterogeneous.

Raman Spectroscopy of the Polyarginine–DNA Complex

The Raman spectrum of the polyarginine–DNA complex at 98% rh is shown in Figure 2. As in the spectrum of the salmine–DNA complex, the structure of DNA bound by polyarginine can be assessed directly from the complex spectrum, but spectral subtraction is required to reveal the polypeptide spectrum.

Analysis of the Polyarginine–DNA Complex Spectrum. The guanine vibrational mode at 682 cm^{-1} is the dominant band in the 600–700-cm^{-1} region; however, it is reduced in intensity as compared to that of a B-form DNA spectrum. This again indicates that the guanine residues are primarily in the B-form conformation, but with some degree of base unstacking. The adenine 727-cm^{-1} band is enhanced with respect to the B-form NaDNA spectrum, but not as drastically as in the salmine–DNA spectrum. Similarly, the 1242-cm^{-1} band is less prominent. At least two other bands in the polyarginine–DNA spectrum are closer to those found in the B-form DNA spectra than they are in the salmine–DNA spectrum. Perhaps the most important is the 835-cm^{-1} band which is more pronounced than in the salmine–DNA spectrum, further supporting the presence of the C2′ endo conformation. Also of interest is the 1336-cm^{-1} adenine band (1333 cm^{-1} in the salmine–DNA spectrum), which is indicative of C2′ endo-anti. All of these indicate that the structure of B-form DNA is less perturbed when bound by polyarginine than when bound by salmine.

The principal difference between the X-ray fiber diffraction patterns of the protamine–DNA complex and B-form DNA is an increase in the first-layer line (Feughelman et al., 1955; Suwalsky & Traub, 1972). This was originally interpreted by Feughelman et al. as an indication that protamines bind within the minor groove. The lack of an increase in the first-layer line in diffraction patterns of the polyarginine–DNA complex has given rise to an alternate model in which protamines are bound in the major groove (Fita et al., 1983). In such a model, the increase in the first-layer line in protamine–DNA diffraction patterns is attributed to the protrusion of the neutral amino acids of protamines into the minor grooves of neighboring DNA molecules. This is physically possible because individual DNA molecules are known to be packed within 25 Å of each other (center-to-center distance between protamine-complexed DNA molecules) when DNA is condensed by protamine (Fita et al., 1983). If this model is correct, some of the differences in the two Raman spectra may result from the insertion of the neutral amino acids of salmine into the minor groove, which disrupt otherwise well-ordered water molecules (Kopka et al., 1983). This would be in accord with the differences observed in the Raman spectra of B-form DNA crystals and B-form DNA in solution, which have been suggested to be caused by slight differences in the hydration of the two states (Brahms et al., 1992).

Analysis of Polyarginine–DNA Complex Subtraction Spectra. Subtraction of the B-form NaDNA spectrum from the polyarginine–DNA spectrum has also been performed (Figure 2C). This reveals DNA difference bands at 730, 782, 1241, 1481, 1575, and 1660 cm^{-1}, all of which indicate base unstacking with respect to B-form DNA, but to a lesser degree than in the salmine–DNA complex. Again the conformationally sensitive amide I and amide III bands are obscured by DNA bands. Use of the low-salt LiDNA spectrum as the subtrahend yields a more interpretable DNA-bound polyarginine subtraction spectrum (Figure 2E). In this spectrum the amide I band is observed at 1679 cm^{-1}, a position also not readily assigned to a regular form of protein secondary structure. The symmetry of the amide I band in the bound subtraction spectrum, as compared to that of the unbound spectrum (Figure 2F), demonstrates that the structure of polyarginine when bound to DNA is more uniform than in solution. Although the shift in the amide I band upon binding to DNA is not extensive (4 cm^{-1}), a substantial shift in the amide III band from 1242 cm^{-1} in the solution state to 1265 cm^{-1} when bound to DNA indicates a significant change in secondary structure. As in the spectrum of DNA-bound salmine, a decrease in the N–C–C_α–C_β skeletal stretch (which occurs at 975 cm^{-1} in the solution spectrum) is observed in the DNA-bound polyarginine spectrum. The sharpness of the bound polyarginine amide bands further suggests that the

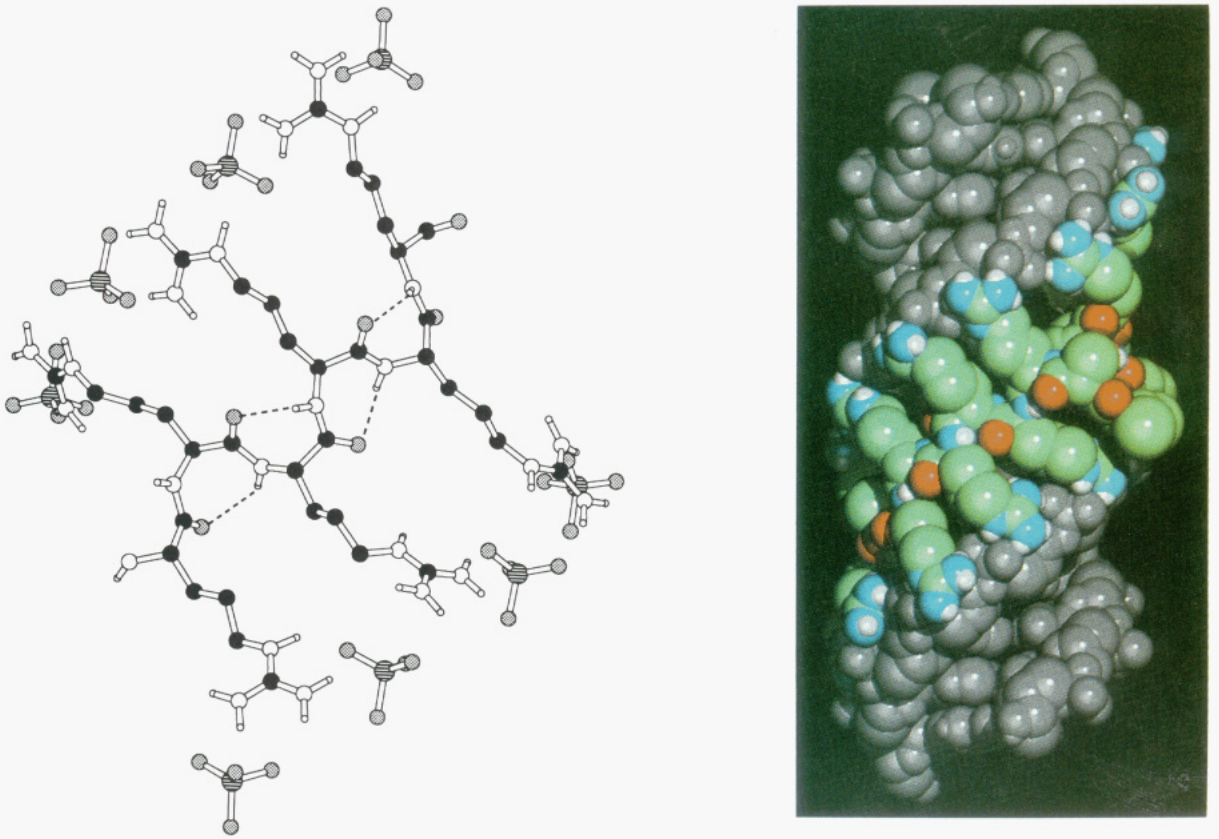

FIGURE 3: (A, left) Ball and stick representation of six arginine residues bound to DNA in the conformation proposed in the text for the arginine clusters of DNA-bound protamine. For clarity of presentation, only the phosphate groups of DNA are shown. ϕ, ψ angles of the peptide backbone are –81° and 74°, respectively. 1→3 hydrogen bonds are depicted by dashed lines. Coordinates are taken from a model of a complete protamine bound to DNA (Hud, 1992). C, N, O, and P atoms are represented by black, white, striped, and hatched spheres, respectively. (B, right) Space-filling model of Arg-12 through Arg-32 of a model of salmine bound to B-form DNA. Preliminary computer models exhibit no steric hindrance in the proposed structure. Note the excellent fit of protamine within the major groove demonstrated by the minimal increase in the diameter of the complex beyond that of the DNA. The C, N, O, and H atoms of protamine are represented by green, blue, red, and light gray spheres, respectively. DNA is shown in gray.

reduction (or broadening) of the N–C–C_α–C_β skeletal stretch band is caused by heterogeneous arginine side-chain geometries or environments rather than disordering of the peptide backbone structure. The observed decrease in the intensity of the amide I band in DNA-bound polyarginine spectrum is not presently understood.

The secondary structure of polypeptides with charged side chains in solution has been studied extensively using spectroscopic techniques, with the general conclusion that these polymers exist in an "extended-helix" conformation (Sengupta & Krimm, 1987). Polyarginine is likely to assume such a conformation. The similarity of the amide I, amide III, and N–C–C_α–C_β skeletal stretch vibrational bands observed in the solution spectra of salmine and polyarginine indicates that the majority of salmine in solution also adopts the extended-helix conformation.

Significant differences are observed in the positions and shapes of the amide I and III bands in the DNA–polyarginine and DNA–salmine subtraction spectra. This indicates that the DNA-bound secondary structures of polyarginine and salmine are not identical. The asymmetry of the salmine amide I and III bands may be due to the presence of the neutral amino acids which intersperse the arginine clusters. These are obviously absent from polyarginine, which exhibits symmetric amide I and III bands when bound to DNA. The neutral amino acids of salmine have long been hypothesized to aid in aligning consecutive arginine residues with the DNA phosphate groups (Feughelman et al., 1955). The potential for alignment of arginine residues and the suggested interaction of these amino acids with the minor groove of neighboring DNA within a protamine–DNA condensate (Fita et al., 1983) may be the origin of the greater perturbation experienced by DNA when bound by protamine rather than polyarginine.

CONCLUSIONS

Most X-ray diffraction and infrared spectroscopy studies reported to date support the existence of B-form DNA in the protamine–DNA complex (Feughelman et al., 1955; Fita et al., 1983; Bradbury et al., 1962). Our analysis of the salmine–DNA Raman spectrum indicates the presence of a modified B-form DNA which retains, as far as we can determine, the same sugar–phosphate backbone structure but with some degree of base unstacking. Similar results are observed for DNA bound by polyarginine, though the degree of base unstacking appears to be less severe. The deviations of the salmine–DNA complex spectrum from a B-form DNA spectrum are similar to, yet not as drastic as, those in the Raman spectrum of C-form DNA (Eurfurth et al., 1972; Fish et al., 1983). This suggests that the level of base unstacking experienced by DNA bound by salmine is less than would be the result of tightening the B-form double helix to $9^1/_3$ base pairs per turn, which is approximately the structure of C-form DNA (Marvin et al., 1961; Arnott & Selsing, 1975). These results are consistent with previous infrared spectroscopy investigations of the salmine–DNA complex which also indicated the presence of a modified B-form DNA (Herskovits & Brahms, 1976), based upon a measured change in the

bisector angle of the O–P–O (as compared to B-form DNA) by 4° with respect to the helical axis.

The apparent partial B- to C-form transition observed in a DNA sample complexed with butylamine at 92% rh (25 °C) by Fish et al. (1983) is similar to the results presented here for the salmine–DNA complex. Fish et al. suggested the C-form character of the spectrum might reflect structural changes induced by an increased temperature sensitivity of the sample. While this may be true, we favor an alternate explanation in which the changes toward a low-humidity-type DNA spectrum are a result of the displacement of water molecules normally associated with DNA at high relative humidity by the binding of protamine or butylamine.

The secondary structure of DNA-bound protamine has previously been proposed as extended (Feughelman et al., 1955), β-sheet (Cid & Arellano, 1982) or α-helix (Warrant & Kim, 1978). The Raman spectra presented here reveal a dominant amide I band of DNA-bound salmine at 1683 cm^{-1}. Such an energetic amide I band is not indicative of α-helix or β-sheet conformations but suggests the presence of a novel secondary structure. These studies also demonstrate that salmine and polyarginine undergo conformational changes upon binding to DNA. While the solution structures of these polypeptides appear to be very similar, their DNA-bound secondary structures may be quite different. This has significant implications regarding attempts to predict the structure of DNA-bound protamine and protamine-like proteins in the absence of DNA.

PROPOSAL OF A NEW MODEL

We have developed a new model of the protamine–DNA complex which is consistent with our Raman spectra obtained for DNA-bound salmine and with other data in the literature. A particularly attractive feature of this model is the implicit description of hydrogen bonding along the peptide backbone of protamine molecules bound to DNA, a feature that is absent in all past models, except for those which support the α-helix assignment. In our model, arginine side chains extend outward from alternate sides of the peptide backbone similar to the extended model first proposed by Feughelman et al. (1955). Protamine is bound within the major groove, as suggested by Fita et al. (1983), not in the minor groove as we proposed earlier (Balhorn, 1982),[2] and the ϕ, ψ angles of the peptide backbone within the arginine regions are approximately –81° and 74°, respectively. This backbone conformation facilitates 1→3 intramolecular hydrogen bonding between neighboring arginine residues in each cluster (Figure 3A). The existence of this type of hydrogen bonding is consistent with the presence of a prominent amide I band at 1683 cm^{-1} in the difference spectrum of DNA-bound salmine.

A poly 1→3 hydrogen-bonded structure, similar to that described here for the backbone of DNA-bound protamine, was proposed as early as 1942 (Huggins, 1942) and later designated the 2_7a-ribbon (where 2 refers to the two residues per helical turn and 7 refers to the formation of a seven-membered ring). Although seven-membered rings appear along the peptide backbone in γ-turns and as kinks in peptide chains (Baker & Hubbard, 1984), the longest stretch of 1→3 hydrogen bonding observed previously was a double ring in a synthetic peptide (Pysh & Toniolo, 1977). The geometry of 1→3 hydrogen bonds has been suggested to be at the very limit of what is considered to be an effective hydrogen bond (Baker & Hubbard, 1984), and this could explain their infrequency in proteins.

Preliminary computer modeling studies of the salmine–DNA complex (Figure 3B) have demonstrated that salt-bridge formation between the guanidinium groups of protamine and the phosphate backbone of DNA places the arginine side chains in the proper orientation and spacing that are characteristic of a 2_7a-ribbon (Hud, 1992). Thus, the formation of 1→3 hydrogen bonds may not be the driving force that defines the secondary structure adopted by protamine but rather an interaction favored by a geometry that is promoted when consecutive arginine residues bind to DNA. Models of the 2_7a-ribbon suggest that such a structure should be quite linear, a feature that would at first seem unfavorable for a protein suspected to wrap around DNA. However, this linearity could be disrupted at regular intervals by the insertion of neutral amino acids into the salmine sequence, residues which function as hinges to redirect the arginine-rich, 1→3 hydrogen-bonded, segments of the peptide chain around the double helix.

[2] Previous actinomycin D binding studies were interpreted to demonstrate that the antibiotic could not bind to DNA complexed with protamine because the protein was positioned in the minor groove, blocking its site of binding. If, as we and others have proposed, the arginine residues within the DNA-binding domain of individual protamine molecules interact with phosphate groups along both strands of DNA, protamines bound in the major groove should also inhibit actinomycin D binding. This is because, in order for actinomycin D to intercalate between base pairs of DNA, the two DNA strands must shift in opposite directions to allow the stacked bases to separate enough for the planar ring of the antibiotic to slip between them. The binding of protamine would lock the two phosphodiester strands in place relative to each other, preventing the conformational change in DNA required for actinomycin D binding.

ACKNOWLEDGMENT

We thank Dr. James M. Benevides and the reviewers of the manuscript for their valuable comments. We also thank Drs. William Conaway, August Droege, and Chuck Stevens for use of laboratory space and equipment.

REFERENCES

Ando, T., & Watanabe, S. (1969) *Int. J. Protein Res. 1*, 221–224.

Ando, T., Yamasaki, M., & Suzuki, K. (1973) *Protamines: Isolation, Characterization, Structure and Function*, 114 pp, Springer-Verlag, New York.

Arnott, S., & Selsing, E. (1975) *J. Mol. Biol. 98*, 265–269.

Aubrey, K. L., Casjens, S. R., & Thomas, G. J., Jr. (1992) *Biochemistry 31*, 11835–11842.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–179.

Balhorn, R. (1982) *J. Cell Biol. 93*, 298–305.

Bandekar, J., & Krimm, S. (1985) *Int. J. Pept. Protein Res. 26*, 158–165.

Benevides, J. M., & Thomas, G. J., Jr. (1983) *Nucleic Acids Res. 11*, 5747–5761.

Benevides, J. M., Stow, P. L., Ilag, L. L., Incardona, N. L., & Thomas, G. J., Jr. (1991a) *Biochemistry 30*, 4855–4863.

Benevides, J. M., Weiss, M. A., & Thomas, G. J., Jr. (1991b) *Biochemistry 30*, 4381–4388.

Bradbury, E. M., Price, W. C., & Wilkinson, G. R. (1962) *J. Mol. Biol. 4*, 39–49.

Brahms, S., Fitrsch, V., Brahms, J. G., & Westhof, E. (1992) *J. Mol. Biol. 223*, 455–476.

Brown, E. B., & Peticolas, W. L. (1975) *Biopolymers 14*, 1259–1271.

Chandley, A. C., & Kofman-Alfaro, S. (1971) *Exp. Cell Res. 69*, 45–48.

Cid, H., & Arellano, A. (1982) *Int. J. Biol. Macromol. 4*, 3–8.

Erfurth, S. C., & Peticolas, W. L. (1975) *Biopolymers 14*, 247–264.

Erfurth, S. C., Kiser, E. J., & Peticolas, W. L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 938–941.
Felix, K. (1953) in *The Chemical Structure of Proteins*, pp 151–164, J. and A. Churchill, London.
Feughelman, M., Langridge, R., Seeds, W. E., Stokes, A. R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., Barclay, R. K., & Hamilton, L. D. (1955) *Nature 175*, 834–838.
Fish, S. R., Chen, C. Y., Thomas, G. J., Jr., & Hanlon, S. (1983) *Biochemistry 22*, 4751–4756.
Fita, I., Campos, J. L., Puigjaner, L. C., & Subirana, J. A. (1983) *J. Mol. Biol. 167*, 157–177.
Herskovits, T. T., & Brahms, J. (1976) *Biopolymers 15*, 687–706.
Hud, N. V. (1992) Ph.D. Thesis, University of California, Davis, Davis, CA.
Huggins, M. L. (1942) *Annu. Rev. Biochem. 11*, 195–218.
Kierszenbaum, A. L., & Tres, L. L. (1975) *J. Cell Biol. 65*, 258–270.
Kopka, M. L., Fratini, A. V., Drew, H. R., & Dickerson, R. E. (1983) *J. Mol. Biol. 163*, 129–146.
Kossel, A. (1928) *The Protamines and Histones*, 107 pp, Longmans, Green & Co., London.
Krimm, S. (1987) in *Biological Applications of Raman Spectroscopy* (Spiro, T. G., Ed.) pp 1–45, John Wiley & Sons, New York.
Krimm, S., & Bandekar, J. (1986) *Adv. Protein Chem. 38*, 181–364.
Lu, K. C., Prohofsky, E. W., & van Zandt, L. L. (1977) *Biopolymers 16*, 2491–2506.
Mansy, S., Engstrom, S. K., & Peticolas, W. L. (1976) *Biochem. Biophys. Res. Commun. 4*, 1242–1247.
Marvin, D. A., Spencer, M., Wilkins, M. H. F., & Hamilton, L. D. (1961) *J. Mol. Biol. 3*, 547–565.
Miescher, F. (1874) *Ber. Dtsch. Chem. Ges. 7*, 376–379.
Miescher, F. (1897) in *The Histochemical and Physiological Work of F. Meischer*, pp 50–107, Vogel, Leipzig.
Nishimura, Y., Masamichi, T., Nakano, T., Higuchi, S., Sato, T., Shida, T., Uesugi, S., Ohtsuka, E., & Ikehara, M. (1983) *Nucleic Acids Res. 11*, 1579–1588.
Nishimura, Y., Masamichi, T., Sato, T., & Aoki, K. (1986) *J. Mol. Struct. 146*, 123–153.
Patel, L., Abate, C., & Curran, T. (1990) *Nature 347*, 572–575.
Pogany, G. C., Corzett, M., Weston, S., & Balhorn, R. (1981) *Exp. Cell Res. 136*, 127–136.
Portugal, J., & Subirana, J. A. (1985) *EMBO J. 4*, 2403–2408.
Pysh, E. S., & Toniolo, C. (1977) *J. Am. Chem. Soc. 99*, 6211–6219.
Sega, G. A. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4955–4959.
Sengupta, P. K., & Krimm, S. (1987) *Biopolymers 26*, S99–S107.
Spiro, T. G., Ed. (1987) *Biological Applications of Raman Spectroscopy*, Vol. 1, 352 pp, John Wiley & Sons, New York.
Suau, P., & Subirana, J. A. (1977) *J. Mol. Biol. 117*, 909–926.
Suwalsky, W., & Traub, W. (1972) *Biopolymers 11*, 2223–2231.
Thomas, G. J., Jr. (1986) in *Spectroscopy of Biological Systems* (Clark, R. J. H., & Hester, R. E., Eds.) pp 233–309, John Wiley & Sons, New York.
Tomlinson, B. L., & Peticolas, W. L. (1970) *J. Chem. Phys. 52*, 2154–2156.
Tu, A. T. (1982) *Raman Spectroscopy in Biology: Principles and Applications*, 448 pp, John Wiley & Sons, New York.
Warrant, R. W., & Kim, S.-H. (1978) *Nature 271*, 130–135.
Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., & Struhl, K. (1990) *Nature 347*, 575–578.
Wilkins, M. H. F. (1956) *Cold Spring Harbor Symp. Quant. Biol. 21*, 75–90.
Williams, R. W. (1983) *J. Mol. Biol. 166*, 581–603.
Williams, R. W., & Dunker, A. K. (1981) *J. Mol. Biol. 152*, 783–813.
Yu, N.-T., Liu, C. S., & O'Shea, D. C. (1972) *J. Mol. Biol. 70*, 117–132.