Mariam, Y. H., & Wilson, W. D. (1979) *Biochem. Biophys. Res. Commun. 88*, 861–866.
McDonald, C. C., Phillips, W. D., & Penman, S. (1964) *Science (Washington, D.C.) 144*, 1234–1237.
McGhee, Y. H., & von Hippel, P. H. (1975) *Biochemistry 14*, 1281–1296.
Mehring, M. (1976) in *High Resolution NMR Spectroscopy in Solids*, Springer-Verlag, New York.
Opella, S. J., Yesinowski, J. P., & Waugh, J. S. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3812–3815.
Opella, S. J., Frey, M. H., & DiVerdi, J. A. (1980) *J. Magn. Reson. 37*, 165–169.
Patel, D. J. (1979a) *Biopolymers 18*, 553–569.
Patel, D. J. (1979b) *Acc. Chem. Res. 12*, 118–125.
Patel, D. J., & Canuel, L. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 674–678.
Pines, A., Gibby, M. G., & Waugh, J. S. (1973) *J. Chem. Phys. 59*, 569–590.
Rill, R. L., Hilliard, P. R., Bailey, J. T., & Levy, G. C. (1980) *J. Am. Chem. Soc. 102*, 418–420.
Robinson, B. H., Lerman, L. S., Beth, A. H., Fusch, H. L., Dalton, L. R., & Auer, C. (1980) *J. Mol. Biol. 139*, 19–44.
Shindo, H. (1980) *Biopolymers 19*, 509–522.
Shindo, H., Wooten, J. B., Pheiffer, B. H., & Zimmerman, S. B. (1980) *Biochemistry 19*, 518–526.
Shindo, H., McGhee, J. D., & Cohen, J. S. (1980) *Biopolymers 19*, 523–537.
Sobell, H. M., Tsai, C. C., Gilbert, S. G., Jain, S. C., & Sakare, T. D. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3068–3072.
Teitelbaum, H., & Englander, S. W. (1975) *J. Mol. Biol. 92*, 55–78.
Terao, T., Matsui, S., & Akasaka, K. (1977) *J. Am. Chem. Soc. 99*, 6136–6138.
Urbina, J., & Waugh, J. S. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 5062–5067.
Wahl, P., Paoletti, J., & LePecq, J.-B. (1970) *Proc. Natl. Acad. Sci. U.S.A. 65*, 417–421.
Watson, J. D., & Crick, F. H. C. (1953) *Nature (London) 171*, 737–738.
Waugh, J. S., Maricq, M. M., & Cantor, R. (1978) *J. Magn. Reson. 29*, 183–190.
Yamada, A., Kaneko, H., Akasaka, K., & Hatano, H. (1978) *FEBS Lett. 93*, 16–18.
Yamane, T. (1971) in *Procedures in Nucleic Acid Research*, Vol. 2, pp 262–294, Harper and Row, New York.

Hydrogen-1 and Carbon-13 Nuclear Magnetic Resonance of the Aromatic Residues of fd Coat Protein†

T. A. Cross and S. J. Opella*

ABSTRACT: The aromatic residues of fd coat protein in sodium dodecyl sulfate micelles are characterized by ^1H and ^{13}C NMR. Resonances from both types of nuclei show structure-induced chemical shift dispersion and line widths indicative of a folded native structure for the protein. The two tyrosines were found to have pK_as of 12.3 and 12.5 by ^1H NMR and spectrophotometric titrations. ^{13}C relaxation measurements show that two of the three Phe rings have significant internal mobility, the two Tyr rings have moderate internal mobility, and the Trp side chain is completely immobilized. Qualitative comparisons are made between the intact virus and the isolated coat protein.

Studies of protein structure and dynamics have been greatly facilitated by the spectroscopic accessibility of the aromatic amino acid residues; this is particularly true for ^1H and ^{13}C NMR.[1] The relatively few aromatic amino acids present in a protein are usually distributed throughout the sequence; therefore they are selective monitors of various protein domains. Resonance frequencies of aromatic carbons and protons do not overlap with those from aliphatic sites, allowing ready discrimination. The aromatic groups are of substantial biochemical interest because of their hydrophobic interactions, acid–base chemistry, hydrogen bonding, and ability to intercalate between other planar groups.

The filamentous bacteriophages are an attractive system for study because of their small size, experimental tractability, and interesting properties. In particular, the major coat protein of these viruses is extraordinary in that it exists as a cytoplasmic protein, as a structural element of a nucleoprotein complex, and as an integral membrane protein at different stages of its life cycle (Denhart et al., 1978; Kornberg, 1974). When *Escherichia coli* is infected with fd, the cells produce very large amounts of the viral coat protein. This protein is synthesized as the water-soluble procoat protein which has a 23 amino acid amino-terminal leader sequence (Chang et al., 1978). After insertion into the cell membrane, the leader sequence is cleaved to form the mature coat protein which is stored in the membrane bilayer (Smilowitz et al., 1972). Virus assembly occurs at the cell membrane; as the DNA is extruded

† From the Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104. *Received March 28, 1980.* This work was supported by grants from the National Institutes of Health (GM-24266) and the American Cancer Society (NP-225A). The ^{13}C NMR experiments were performed on a spectrometer obtained with a National Science Foundation grant. ^1H NMR spectra were obtained at the Middle Atlantic NMR Facility. T.A.C. is supported by a Cell and Molecular Biology Training Grant. S.J.O. is a fellow of the A. P. Sloan Foundation (1980–1982).

[1] Abbreviations used: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; PFU, plaque-forming units; $NaDodSO_4$, sodium dodecyl sulfate; Me_4Si, tetramethylsilane; DSS, 4,4-dimethyl-4-silapentane-1-sulfonic acid.

0006-2960/81/0420-0290$01.00/0

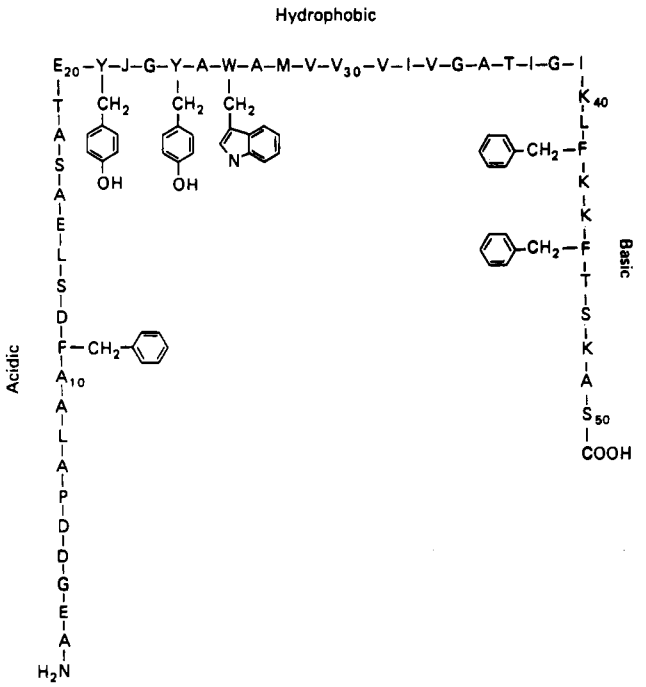

FIGURE 1: Amino acid sequence of fd coat protein (Nakashima & Konigsberg, 1974). Single letter abbreviation for the amino acids are used, and the aromatic side chains are drawn in for emphasis. The three regions of the polypeptide are represented by the three legs of the polypeptide in the figure.

from the cell interior through the membrane, coat proteins are wrapped around it, forming the filamentous structure. The assembly of fd may involve significant changes in the structure as well as the dynamics of the viral components. NMR studies of the virus and the isolated macromolecular subunits can give a description of the assembly process.

fd is 88% by weight major coat protein. A total of 2700 copies of this protein are wrapped in a helical array forming a hollow tube for packaging the viral DNA (Newman et al., 1977; Marvin et al., 1974). The fd coat protein has 50 amino acids in a single peptide chain, of which 6 are aromatic (Nakashima & Konigsberg, 1974). No other protein is present in the virus, except for about five copies of a minor coat protein at one end of the particle. Figure 1 contains the primary sequence of the major coat protein, emphasizing the three Phe, two Tyr, and one Trp of the protein. This sequence appears to be divided into three distinct domains, all of which contain aromatic residues. The amino-terminal region has a high percentage of acidic side chains while the carboxyl-terminal region has a high percentage of basic residues. The central section of 19 amino acids is very hydrophobic. The three phenylalanines are located in the hydrophilic terminal sections, two in the basic region (Phe-42, Phe-45) and one in the acidic region (Phe-11). The two tyrosines and the tryptophan are located in a cluster at one end of the hydrophobic midsection of the protein. One of the tyrosines (Tyr-21) is the first amino acid of the hydrophobic region and separated from the second tyrosine by only two residues (Tyr-24). The tryptophan follows one residue later (Trp-26).

The coat protein is stable in several quite different environments. When first synthesized, it is water soluble by virtue of the extra amino acid residues of the leader sequence. The mature coat protein is insoluble in water (Knippers & Hoffmann-Berling, 1966). It exists as a dimer with about 50% α helix in high concentrations of detertgents and 6 M guanidine hydrochloride; obviously it is very stable toward dissociation and denaturation (Nozaki et al., 1978; Cavalieri et al., 1976). Without detergents present, the protein subunits aggregate to form an intractable, insoluble polymer. In contrast, the virus itself, even though it is 88% coat protein, is highly water soluble. The coat protein is clearly stabilized by both protein–lipid and protein–protein interactions and possibly by protein–DNA interactions. A drastic change in solubility occurs when the coat protein goes from a lipid environment to the assembled virus. The extent of change of protein conformation accompanying viral assembly is not known. Circular dichroism suggests that the protein has in lipids or detergents about 50% α helix but close to 90% α helix in the intact virus (Williams & Dunker, 1977; Day, 1969). Predictions of secondary structure by Green & Flanagan (1976) based on the rules of Lim (1974) indicate that no β sheet is present, but there is about 70% α helix. NMR studies of the protein α carbons have shown that the backbone has about eight residues that have a great deal of flexibility while the rest are rigidly constrained (Cross & Opella, 1980).

This paper focuses on the aromatic residues of the fd coat protein in $NaDodSO_4$ micelles with a few qualitative comparisons to the intact virus. It is essential for these studies that the aromatic resonances are well separated in chemical shift from the aliphatic resonances since large amounts of detergent must be present to solubilize the protein and the detergent resonances obscure significant spectral regions. The aromatic region contains only protein signals, and, by use of appropriate data collection technqiues, dynamic range problems can be avoided. The detergent-solubilized coat protein can be studied in solution by high-resolution ^1H and ^{13}C NMR because the protein dimer is bound to $\sim$60 detergent molecules, forming a well-defined micelle of $\sim$27 000 daltons (Makino et al., 1975).

High-resolution ^1H NMR of proteins is most useful for empirically characterizing conformations and monitoring local and overall changes in structure. Because of its reduced sensitivity, natural abundance ^{13}C NMR is less useful for following protein conformational changes but is a much more powerful approach to the study of protein dynamics because of the well-understood relaxation properties of ^{13}C spins with attached protons.

Sykes and co-workers have reported complementary ^{19}F NMR studies of this coat protein where the tyrosines were substitued with 3-fluorotyrosine (Hagen et al., 1978). The two tyrosines were found to be similar in chemical shift and relaxation properties in phospholipid vesicles and detergent micelles. Their relaxation measurements were consistent with relatively rapid rotation occurring about the α,β and β,γ tyrosine bonds. Solvent perturbation chemical shift experiments showed that the rings were completely buried in phospholipid vesicles and partially exposed to solvent in micelles (Hagen et al., 1979).

Materials and Methods

Virus Preparation. A single colony of *E. coli* K3300 F+ is grown in a 50-mL culture. This is used to infect 500 mL of rich media which is grown to late log phase and subsequently used to inoculate 10 L of superrich media based on beef tryptone, yeast extract, and phosphate buffer. Oxygen is bubbled vigorously into the growth medium; 4 N KOH is added to maintain the pH between 7.2 and 7.5, and Antifoam B (Fisher Scientific) is added to control excessive foaming during the growth. When the culture reaches an optical density of $\sim$3 at 550 nm, fd is added at a multiplicity of at least 10, usually 5×10^9 PFU/mL. The growth is continued for $\sim$4 h when the virus yield plateaus at about 2×10^{13} PFU/mL.

Defined media growths, such as for the incorporation of [$^2H_5$]phenylalanine ([$^2H_5$]Phe), utilize M9 minimal salts supplemented with 100 μg/mL of each of the L amino acids, 10 μg/mL of thiamine hydrochloride, 1% glycerol, 1 mM $MgSO_4$, and 0.1 mM $CaCl_2$ (Miller, 1972). A 50-mL culture of *E. coli* grown on the defined media including the labeled amino acid is used to inoculate 1 L of growth media. The *E. coli* are grown to an optical density of about 1 at 550 nm and infected at a multiplicity of 10. As with the rich media, 4 M KOH is used to maintain the pH and Antifoam B to control foaming. The infected cells are grown for 6 h when the virus titer is 5×10^{12} PFU/mL.

The virus is purified from the growth media after pelleting the bacterial cells by centrifugation. The virus is precipitated with poly(ethylene glycol)-6000 (Sigma). The bulk supernatant is decanted, and the remaining volume is centrifuged to pellet the virus. The pellet is resuspended in distilled water, and residual cell debris is spun down. The solution is layered on cesium chloride (EM reagent) step gradients; blue bands of virus are recovered. The virus is then loaded on cesium chloride equilibrium gradients in tubes for the SW50.1 rotor. Blue bands are removed and then dialyzed exhaustively against distilled water. Virus is stored in distilled water with 0.02% sodium azide. Virus concentrations are determined by PFU counts or by absorbance at 268 nm by using an extinction coefficient of 3.84 mg^{-1} cm^2.

Coat Protein Solubilization. The virus is disrupted with $NaDodSO_4$ recrystallized from ethanol. Two grams of $NaDodSO_4$ per gram of fd and one drop of $CHCl_3$ are added to a solution of virus that is 40 mM in borate buffer at pH 9.0. The cloudy suspension is kept at 50 °C for 30 min or until the solution has been clear for 10 min. The solubilization process can be followed spectrophotometrically by a hyperchromic and hypsochromic shift in the λ_{max} of the ultraviolet spectrum.

The detergent-solubilized coat protein is separated from the single-stranded fd DNA by gel chromatography. By use of a Sephacryl S-200 (Pharmacia Fine Chemicals) column equilibrated with 10 mM $NaDodSO_4$, 40 mM borate at pH 9.0, and 8% glycerol, the DNA elutes in the void volume while the coat protein is retained. The coat-protein fractions are pooled and concentrated in an Amicon stirred cell Model 402 by using a 5000-dalton cutoff membrane, DM-5. The coat protein in $NaDodSO_4$ micelles is then washed a number of times in the cell with 40 mM borate buffer at pH 9.0. The concentrated detergent solubilized protein is then lyophilized for storage. Protein concentrations are determined spectrophotometrically by using an extinction coefficient of 1.65 mg^{-1} cm^2.

A Corning Model 12 pH meter was used with a sodium-insensitive electrode. All pH measurements are direct meter readings; we have used pH to represent H_2O solutions and pH* for high percentage ^2H_2O solutions.

Coat protein samples were prepared by adding H_2O or ^2H_2O to the lyophilized powder of coat protein and $NaDodSO_4$. Solutions of 0.05–0.5 N KOH and a 0.2 N solution of HCl were used to adjust the pH of the samples.

NMR Spectroscopy. ^1H NMR spectra were obtained at 360 MHz on Bruker WH-360 spectrometer. The coat-protein samples were typically 4 mM in protein in 0.5 mL of ^2H_2O in a 5-mm NMR tube. The spectrometer was used in the pulsed Fourier transform mode for the decoupling experiments. Otherwise, most of the spectra were obtained with rapid scan correlation spectroscopy. ^1H chemical shifts are referenced to internal DSS.

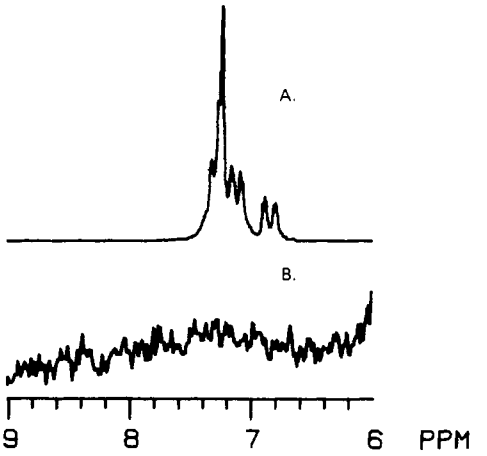

FIGURE 2: Aromatic regions of ^1H NMR spectra at 360 MHz. (A) $NaDodSO_4$-solubilized fd coat protein in D_2O at 45 °C, pH* 9, 5 mM protein, and 10 mM borate. There was a total of 256 scans in the correlation mode at a 1-s recycle delay. (B) Intact fd, same spectral region as (A), 25 °C, pH* 7, 50 mg/mL. There was a total of 8192 accumulations in the FT mode at a 1-s recycle delay.

The ^{13}C spectra were obtained on a Nicolet NT-150 spectrometer at 37.7 MHz with square wave modulated proton decoupling at 150.0 MHz in 20-mm sample tubes having a sample volume of 10 mL. The ^{13}C spectra of the coat protein resulted from signal-averaging 10^3 free induction decays and "block-averaging" the Fourier-transformed spectra to reduce dynamic range problems from the large detergent resonances. The inversion–recovery sequence (180–τ–90–τ) was used to measure T_1 and gated decoupling for NOE measurements. ^{13}C chemical shifts are referenced to external Me_4Si.

Results

^1H NMR Studies of fd and Its Coat Protein in Solution. The aromatic spectral regions of the ^1H NMR spectra of the fd coat protein in $NaDodSO_4$ micelles and of the fd virion are shown in Figure 2. The spectra are very different despite similar protein concentrations in the two samples. Figure 2A is a well-resolved ^1H NMR spectrum of a protein in solution, while Figure 2B shows essentially no evidence of intensity in the aromatic region even after very long periods of signal averaging.

The extremely different appearances of these ^1H NMR spectra reflect the differences in rates of reorientation of the aromatic groups of the coat protein. fd is a highly organized 16×10^6 dalton particle with corresponding slow rotational diffusion; as a result, the proton resonances are severely broadened by ^1H–^1H dipolar interactions that are not motionally averaged in the virus particle. The aromatic protons are strongly coupled to nearly all of the other protons in the virus by spin diffusion; therefore the aromatic resonances cannot be considered as independent spins reflecting the properties of their sites (Cross et al., 1981).

The detergent-solubilized coat protein has an overall rotational correlation time at 50 °C of about 15 ns, and the aromatic rings have substantial internal mobility (Cross & Opella, 1980). The presence of these relatively rapid motions enables the protein to give a well-resolved ^1H NMR spectrum with intensities, line widths, and chemical shift dispersion typical of a native low molecular weight protein. The aromatic region of the solubilized coat protein spectrum can be analyzed with the standard protein NMR approach, using selective decoupling, resolution enhancement, isotopic substitution, and pH titrations as well as various continuous-wave and pulsed NMR procedures.

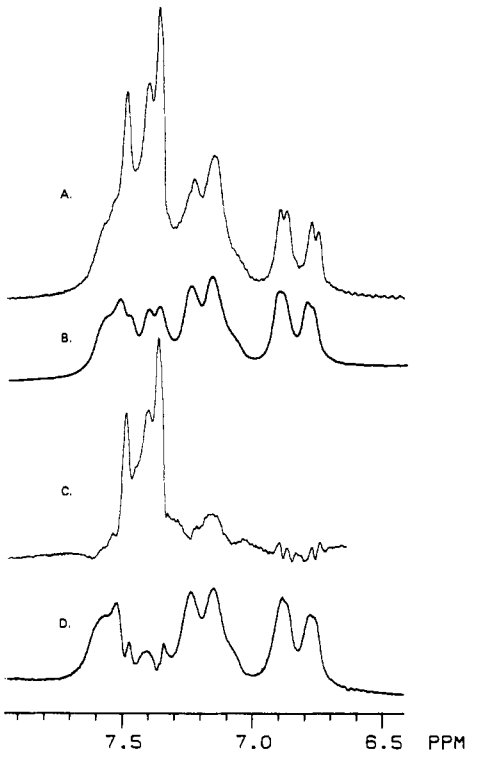

FIGURE 3: [$^2H_5$]Phenylalanine-labeled fd coat protein, $NaDodSO_4$ solubilized in D_2O, pH*9 at 45 °C, ^1H NMR spectra at 360 MHz, 256 scans in the correlation mode at a 1-s recycle delay. (A) Unlabeled protein (4 mM) for comparison. (B) Labeled protein (4 mM) showing ~70% incorporation of [$^2H_5$]Phe. (C) Spectrum A – spectrum B. (D) Spectrum B – one-third of spectrum C.

The spectrum of Figure 2A contains resonances from only the protons bonded to aromatic carbons in the protein since this sample was heated in ^2H_2O solution to completely exchange the oxygen and nitrogen bonded protons for deuterons. There are a total of 28 aromatic protons from the three Phe, two Tyr, and one Trp of the protein.

Several lines of evidence assign the four prominent two-proton lines located at 6.8, 6.9, 7.1, and 7.2 ppm in Figure 2A to the two tyrosines. First of all, the two furthest upfield peaks are near the expected position of Tyr 3,5 protons and the more downfield resonances are located appropriately for the Tyr 2,6 protons. When the NMR data have the effective resolution enhanced by mathematical manipulations, all of these peaks appear as doublets as in Figure 4A and particularly clearly in Figure 5A,B. These doublet patterns for the two-proton resonances are characteristic of tyrosine rings reorienting rapidly about C_β–C_γ bonds. All four resonances are well resolved. The separation between the centers of the3,5 proton lines is about 0.12 ppm, a value for structure-induced chemical shift dispersion typical of a native protein (Wüthrich, 1976). These four peaks, and no others, change chemical shift position when the sample pH is altered in the pH range greater than 10. This helps assign them to the only groups that can titrate in this range, the phenolic rings of tyrosine. The intensities of the lines assigned to tyrosine are not diminished when [$^2H_5$]Phe is substituted for Phe in the protein. fd-infected *E. coli* cells grown on a defined medium where the phenylalanine is replaced by [$^2H_5$]phenylalanine yield virus with about 70% of the coat protein phenylalanines replaced by [$^2H_5$]Phe. The ^1H NMR spectrum of this selectively labeled coat protein is shown in Figure 3B and in resolution-enhanced form in Figure 5. Much of the aromatic proton intensity observed in the normal coat protein is missing in these spectra. The downfield tyrosine doublets from the 2,6 protons are much more obvious

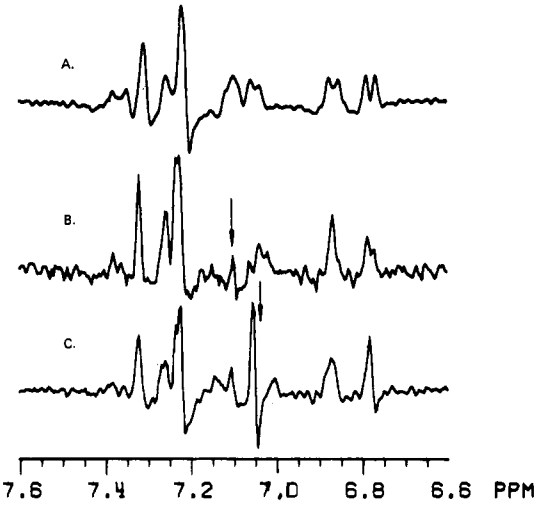

FIGURE 4: ^1H homonuclear decoupling at 360 MHz of tyrosine $C_{2,6}$ protons from $C_{3,5}$ protons in the $NaDodSO_4$ solubilized fd coat protein in D_2O at 45 °C, pH* 9. Each spectrum represents 256 accumulations in the FT mode at a 1-s recycle delay and resolution enhancement by multiplication with a trapezoidal function. (A) Decoupler off. (B) Decoupler at 7.11 ppm from DSS collapsing doublet at 6.87. (C) Decoupler at 7.06 ppm collapsing doublet at 6.78.

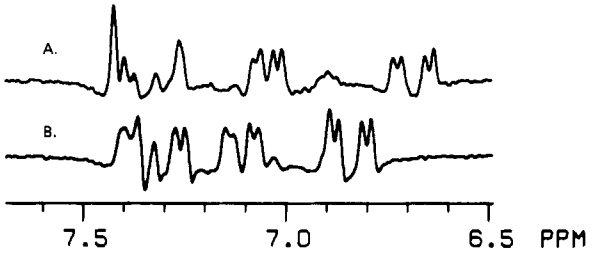

FIGURE 5: pH titration of tyrosines in 4 mM $NaDodSO_4$ solubilized [$^2H_5$]Phe fd coat protein in D_2O, 45 °C. ^1H NMR spectra at 360 MHz and 256 scans in the correlation mode at a 1-s recycle delay. Resolution is enhanced by multiplication by a trapezoidal function. (A) pH* 12.5. (B) pH* 9.0.

in the spectra of the deuterium-labeled protein. Subtraction of the [$^2H_5$]Phe-labeled coat protein spectrum from the unlabeled protein spectrum generates Figure 3C. Figure 3A has resonances from three Phe, two Tyr, and one Trp, Figure 3B has resonances from two Tyr and one Trp and ~30% of the Phe intensity, Figure 3C has resonances from the three Phe, and Figure 3D has resonances from only the two Tyr and one Trp. Most of the Phe and Trp lines can be assigned by inspection. The majority of the phenylalanine protons give resolved narrow resonances near 7.4 ppm, but there is some broad Phe intensity at 7.15 ppm. There is also significant chemical shift dispersion among the resonances from the three Phe residues.

Single-frequency ^1H decoupling experiments shown in Figure 4 assign the 3,5 and 2,6 pairs of proton resonances to their respective rings. The arrows in Figure 4B,C show the frequency of the decoupler irradiation. The results are clear-cut and show that the 3,5 protons at 6.87 ppm are coupled to the 2,6 protons at 7.11 ppm and that the 3,5 protons at 6.78 ppm are from the same tyrosine as the 2,6 protons at 7.06 ppm. These decoupling results reinforce the assignment of these resonances to the two tyrosines.

Spectral shifts as a function of pH are useful for making resonance assignments and determining p*K*as of individual groups. The effects of changing the pH on coat-protein ^1H NMR spectra are shown in Figure 5 with the [$^2H_5$]Phe-labeled protein for clarity. The dominant effect is that the four peaks assigned to the two tyrosines change chemical shift position.

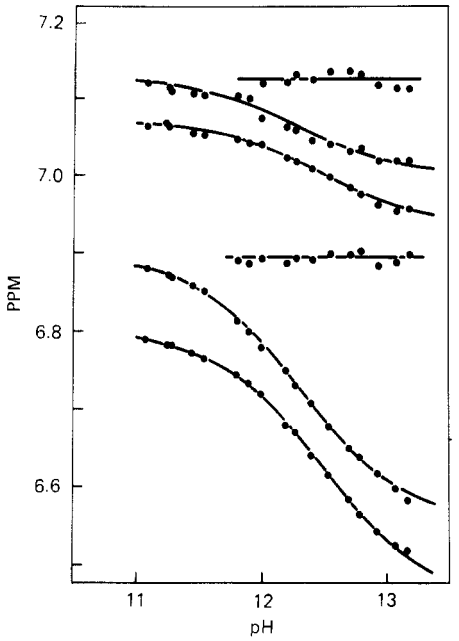

FIGURE 6: pH titration of ^1H resonances of fd coat protein at 360 MHz between 6.50 and 7.15 ppm from DSS. The data have been collected from spectra such as those in Figure 5 and recorded under the conditions given in Figure 5. All titration curves are theoretical Henderson–Hasselbach titration curves superimposed on the data.

Small or no chemical shift changes occur among the Phe or Trp resonances when the pH is altered. Similar behavior is seen in the ^{13}C NMR spectra.

Figure 6 graphically summarizes the influence of pH on many of the aromatic proton chemical shifts. There are two distinct classes of resonances, titrating and nontitrating, which are from tyrosines and from phenylalanines and tryptophan, respectively. At high pH, two resonances become visible which do not titrate. Comparison of spectra from the [$^2H_5$]Phe-labeled protein and the normal protein shows no reduction of the intensity of the nonitrating resonance at 6.88 ppm, indicating that this peak, seen as a broad triplet in Figure 5A, is from Trp-26 and is probably the C_5 proton. The other nontitrating resonance that is resolved at high pH is at 7.12 ppm and is clearly reduced in size in the [$^2H_5$]Phe-labeled protein spectra; therefore it can be assigned to one of the phenylalanines.

The pK_as of Tyr-21 and Tyr-24 of the coat protein have been determined from the titration curves in Figures 6–8. These pK_as are around 12.5 which is extraordinarly high for tyrosines which usually have pK_as near 10.5. The errors in the pK_a determinations came from two sources since it is difficult to accurately reference the chemical shifts internally while using correlation spectroscopy at very high pH values and because the errors inherent in measuring high pH values even by using a sodium-intensitive electrode are very significant.

Figures 7 and 8 serve to increase the accuracy of the pK_a determinations and demonstrate that there is a significant difference in pK_a between Tyr-21 and Tyr-24. A plot of the separation between the 3,5 and 2,6 proton resonances from the same tyrosine (Figure 7) removes the chemical shift reference as a source of error because only the difference in parts per million is being plotted. The more upfield Tyr resonances come from a tyrosine with a pK_a = 12.5 ± 0.1. By plotting the chemical shift separation between the two pairs of tyrosine 2,6 protons (Figure 8), the difference in pK_a between these two residues can be determined with high accuracy (<0.05 pH unit) even though the absolute values of the pK_as are known less accurately. There is a 0.20 pH unit difference

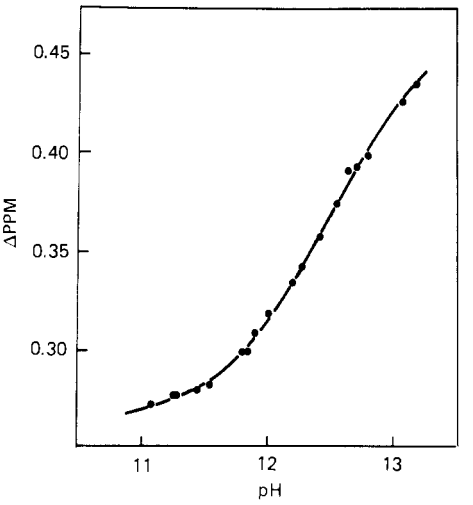

FIGURE 7: Separation of the $C_{3,5}$ protons from the $C_{2,6}$ protons of the single tyrosine with low pH resonances at 6.78 and 7.06 ppm. The resulting titration fits a theoretical curve with a pK_a of 12.5.

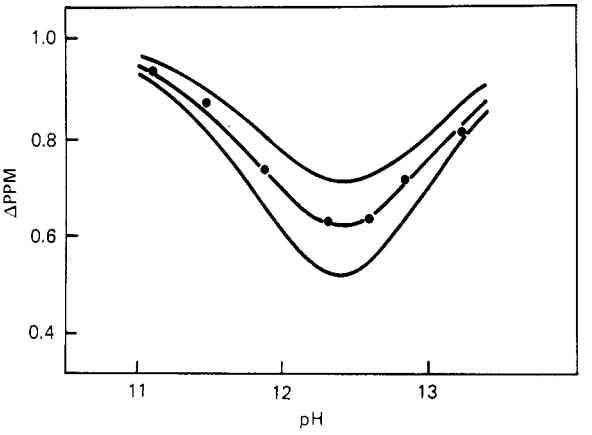

FIGURE 8: Separation of the two pairs of tyrosine $C_{2,6}$ protons. The three curves represent the expected separation of the tyrosine resonances over the titratable pH range with pK_a differences of 0.15, 0.20, and 0.25 pH unit.

between the pK_as of Tyr-21 and Tyr-24 of fd coat protein in $NaDodSO_4$ micelles. The tyrosine with the upfield protons has pK_a = 12.5 ± 0.1 and the tyrosine with the downfield protons has pK_a = 12.3 ± 0.1; however, the difference between the pK_a is accurately known to be 0.20 pH unit.

The points in Figures 6–8 are experimental, while the lines are theoretical Henderson–Hasselbalch titration curves. The tyrosine pK_as have also been determined spectrophotometrically on these samples of fd coat protein; the measured pK_as of both tyrosines is about 12.4 because of the lack of resolution between the two tyrosines. Spectrophotometric titration of fd virus in solution does not follow a typical Henderson–Hasselbalch pattern; there is essentially no titration of the tyrosines until pH 12, indicating that the tyrosine pK_as in the virus are equal to or higher than the average value of 12.4 determined for the coat protein alone.

The low-field region of teh ^1H NMR spectrum of coat protein in H_2O solution is shown in Figure 9. This spectrum contains resonances from many of the oxygen-and nitrogen-bound "exchangeable" protons in addition to the carbon-bound aromatic protons seen in Figure 2A. A single-proton resonance at 9.8 ppm disappears when the sample is transferred to ^2H_2O solution. This resonance can be attributed to the indole N–H of Trp-26 based on its characteristic chemical shift position (McDonald et al., 1971). This resonance has a temperature dependence of 0.006 ppm/°C which is indicative of a non-

Table I: Chemical Shifts and Relaxation Parameters of the Aromatic Residues of fd Coat Protein

amino acid	carbon site	^{13}C shift[a] (ppm)	^{13}C $\Delta\nu_{1/2}$ (Hz)	^{13}C T_1[a] (ms)	^{13}C NOE[a]	^{1}H shift[b]
Trp	γ 3	109.5	9			
	δ_1 2	125.6	20			
	δ_2 9	128.3	8			
	ϵ_2 8	136.7	8			7.52
						7.57
	ϵ_3 4	119.2	50[c]			
	ζ_2 7	112.7	50			
	ζ_3 5	122.0	40			6.88
	η_2 6	119.2	50[c]			
Tyr	γ 1	129.0				
	δ 2,6	131.0		180	1.8	7.06, 7.11
	ϵ 3,5	116.8	18[c]	170	1.6	6.78, 6.87
	ζ 4	156.2, 156.4	6			
Phe	γ 1	137.5, 137.8, 138.0	5			
	δ 2,6	130.0	13[c]	210	2.1	7.12, 7.35,
	ϵ 3,5	129.5	12[c]	190	2.2	7.42, 7.49
	ζ 4	127.8	20[c]			

[a] pH 9.0, 50 °C. [b] pH 8.4, 45 °C. [c] These line widths are from overlapping resonances.

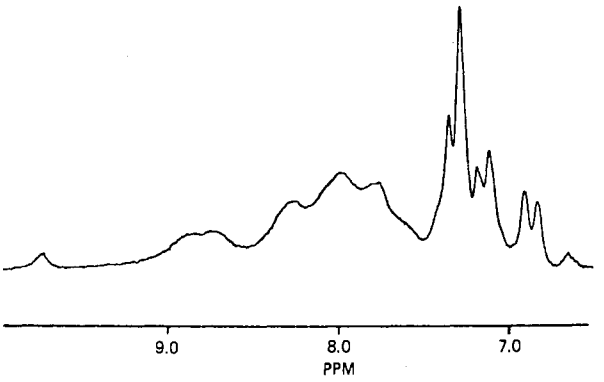

FIGURE 9: Complete low-field ^{1}H NMR spectrum of $NaDodSO_4$ solubilized fd coat protein at 360 MHz in H_2O; 4 mM protein and 10 mM borate, pH 8.4, at 30 °C. The spectrum represents 64 scans in the correlation mode at a 1-s recycle delay. Nitrogen- and oxygen-bound protons are observed as well as the aromatics. The indole N–H has a resonance at 9.8 ppm from DSS.

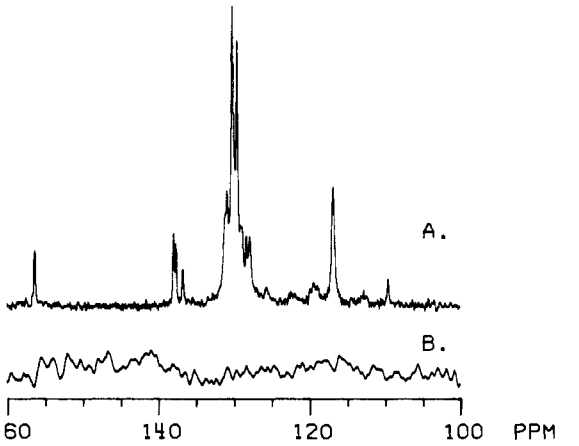

FIGURE 10: Aromatic regions of the natural abundance ^{13}C NMR spectra at 38 MHz (A) $NaDodSO_4$ solubilized fd coat protein in D_2O at 50 °C, pH* 9, 5 mM protein, and 10 mM borate; 81 blocks of 1024 scans at a 2-s recycle delay were summed. (B) Intact fd, same spectral region as A at 25 °C, pH* 7, 50 mg/mL. Thirty-six hours of signal averaging with a 2-s recycle delay.

hydrogen-bonded proton (Opella et al., 1980). At pH* 7.4 and 30 °C in $^{2}H_2O$ solution, this resonance has a half-life for exchange with solvent deuterons of 10 min (Cross & Opella, 1980).

Natural-Abundance ^{13}C NMR of fd and Its Solubilized Coat Protein. Figure 10 compares the aromatic region of the natural-abundance ^{13}C NMR spectra of fd in solution to that of the coat protein in $NaDodSO_4$ micelles. Both spectra were obtained under similar high-resolution conditions with low-power square wave modulated proton decoupling which is sufficiently intense to collapse carbon–proton spin–spin couplings but not the dipole–dipole couplings. The severe line broadening and low intensity of Figure 10B are a consequence of substantial ^{13}C–^{1}H dipolar couplings and ^{13}C chemical shift anisotropy that are not averaged by molecular motion. As in the ^{1}H NMR spectra, the solubilized coat protein reorients rapidly enough to give a high-resolution ^{13}C NMR spectrum. There are a total of 38 carbons contributing to the aromatic spectral region of Figure 10A. At lease one resonance from each of the six aromatic residues can be resolved, and, in the case of Trp-26, essentially all eight lines are distinguishable.

Table I summarizes the assignments of lines in the ^{13}C NMR spectrum of Figure 10A to carbon types; most of these are made by comparing chemical shifts of free amino acids of peptides to those of the protein. The most downfield resonances at 156.25 ppm are from the two Tyr C_ζ because of their unique chemical shift position and sensitivity to pH. In some spectra, these two C_ζ resonances are resolved to the base line. Three partially resolved lines are observed near 138 ppm from the Phe C_γ; the dispersion among these γ carbons is less than that typically seen for globular proteins in solution; however, it is significant compared to the single line observed for the Phe C_γ in the spectra of denatured proteins. Nearly all of the carbons from Trp-26 are readily assigned to resonances in Figure 10 by inspection on the basis of model compound chemical shifts. The four C_ϵ from Tyr-21 and Tyr-24 are not resolvable in the line at 117 ppm. There are some difficulties associated with sorting out the resonances in the reigon 128–132 ppm because some of the protonated carbons have narrow lines.

In ^{13}C NMR spectra of native proteins at moderate magnetic field strengths, the nonprotonated carbon resonances are typically much narrower than the protonated carbon resonances. This is because the dominant relaxation mechanism for carbons of proteins is dipole–dipole interactions with nearby protons. Since the efficiency of relaxation in this regime goes as $1/r^6$ where r is the internuclear distance, those carbons bonded to protons have much broader lines (shorter T_2) than those carbons farther away from protons. Most aromatic groups have limited internal mobility, and the presence of motions more rapid than the overall tumbling of the protein

is generally manifested through changes in T_1 or NOE rather than substantial line narrowing (Opella et al., 1974). Trp-26 of the coat protein in the micelles shows typical ^{13}C relaxation behavior of a rigidly held aromatic group in a protein with sharp nonprotonated resonances $C_\epsilon 2$ at 137 ppm and C_γ at 109 ppm, with the protonated carbons having broad lines at 113, 119–120, 122, and 126 ppm.

One of the most striking features of the ^{13}C spectrum of the coat protein in micelles is the very narrow lines from the phenylalanine C_δ and C_ϵ protonated ring carbons at 129.5 and 130.0 ppm. The line widths of these Phe carbon resonances is substantially less than those typically observed for protonated ring carbons in native proteins. The nonprotonated carbons can generally be disitnguished as a class from protonated carbons by their line width and their response to weak noise-modulated proton decoupling. This procedure greatly broadens the resonances of proton-bearing carbons with their large $^{13}C-^1H$ couplings but can be arranged to have a negligible effect on the line widths of nonprotonated carbons with their small $^{13}C-^1H$ couplings. Neither relying on qualitative line width differences nor the application of off-resonance noise-modulated proton decoupling noise-modulated proton decoupling provided assignments of carbon types to the overlapping resonances in the region 128–132 ppm. This spectral region contains resonances from six Phe C_ϵ, six Phe C_δ, three Phe C_ζ, two Tyr C_γ, four Tyr C_δ, and one Trp_{δ_2}. By combination of off-resonance noise modulated decoupling with ^{13}C spin–echoes, the two classes of aromatic carbons could be distinguished (Opella & Cross, 1980). The narrow lines at 129.0 and 128.3 ppm are from nonprotonated carbons. Since the line at 129.0 ppm is about twice the intensity of the other, it is probably from the two Tyr C_γ while the 128.3 line is from the Trp_{δ_2}.

All of the nonprotonated aromatic carbons of coat protein in $NaDodSO_4$ are narrow, and some of the phenylalanine protonated ring carbons are narrow. Tyr-21 and Tyr-24 display intermediate protonated carbon line widths as seen by comparing the four overlapping C_ϵ resonance at 117 ppm to the Trp carbons near 120 ppm and the Phe carbons at 130 ppm.

The differences in line width among the protonated carbons of the aromatic residues of fd coat protein are a reflection of the different rates of internal motion of these groups. More information can be obtained on the dynamical properties of the coat protein with additional relaxation measurements; the results of T_1 and NOE determinations are given in Table I. Notice that the phenylalanine ring carbons have larger NOE than the others; therefore the intensities of the Phe resonances in Figure 10 are relatively larger than other lines.

Discussion

The aromatic amino acids of fd coat protein have NMR properties that are different from those of previously studied native proteins. Some of the aromatic rings have relatively rapid motion compared to the α-carbon backbone. The two tyrosines have extremely high pK_as. The limited data presented here on the intact virus indicate that the coat proteins are part of a highly constrained structure.

fd coat protein in the presence of excess detergent has a stable native conformation. The evidence for this comes from the chemical shift dispersion among resonances in the ^1H and ^{13}C spectra, broad line widths of some ^{13}C resonances, and the slow rates of solvent exchange of many amide protons. General features of the structural integrity have been described previously by analyzing the protein α-carbon backbone relaxation parameters and the amide exchange rates (Cross & Opella, 1980); most of the properties are similar to those observed for globular proteins in solution, although a few of the α-carbon resonances have narrow line widths and other relaxation properties consistent with part of the protein backbone being mobile.

Trp-26 is near the center of the hydrophobic sequence in the coat protein and one residue away from a tyrosine. Essentially all of the carbons of Trp-26 give resolved resonances with its nonprotonated carbons giving narrow lines and the proton bearing carbons giving broad lines as expected for an immobilized bulky ring system in a protein that has an isotropic rotational correlation time of 15 ns. Only one of the carbon bound proton resonances of Trp-26 is well resolved and that is observed as a triplet at 6.88 ppm at high pH; this is probably from the C_5 proton of the indole ring. The rest of the Trp proton resonances are broad and unresolved, which is in accord with a lack of internal motion of the Trp side chain. The indole N–H proton can be clearly seen in the protein spectrum obtained in H_2O. The exchange rate of this proton is significantly slower than that for a random coil polypeptide and indicates that the residue is shielded from the solvent. Hydrogen bonding of this N–H appears to be ruled out by the temperature dependence of its chemical shift which is 0.006 ppm/°C. These results are all consistent with the Trp-26 rings being in a hydrophobic environment of a highly structured region of the protein.

Tyr-21 and Tyr-24 are near the end of the hydrophobic stretch of residues adjacent to the acidic hydrophilic portion of the sequence. Their proton resonances are completely resolved, although the ^{13}C spectra show only slight chemical shift inequivalence for the ζ and γ carbons. The relaxation parameters of the ϵ carbon resonances and the characteristic doublet structure of the ^1H spectra show that these resonances for both tyrosines are moving more rapidly than the protein backbone. These results are fully compatible with those of Sykes and co-workers on 3-fluorotyrosine labeled coat protein (Hagen et al., 1978).

The pK_as of the tyrosines of the detergent-solubilized coat protein have been determined by ^1H NMR chemical shift titration and found to be 12.5 and 12.3. Spectrophotometric titration of the protein gives approximately the mean of these values. Tyrosines in peptides and proteins typically have pK_as between 10 and 11, with a few reports of "nontitrating" residues (Snyder et al., 1979). The tyrosines of fd coat protein have extraordinarily high pK_as which are fully reversible. While the high pK_as are compatible with the tyrosines existing in an extremely hydrophobic environment like the nearby Trp-26, they are also accessible to solvent because of the reversibility of the titration as well as the reasonable exchange rate for the Trp-26 indole N–H and the amide protons throughout the backbone. The relatively rapid internal motions of the tyrosine rings are different from the immobilized Trp ring separated by only one residue in the sequence.

These coat-protein tyrosines have been shown to have very high pK_as in the virus by others (Dunker et al., 1979; Day & Wiseman, 1978). The finding of this extraordinary feature for the membrane-bound and DNA-bound forms of the proteins suggests that the Tyr pK_as are a consequence of the protein folding rather than intermolecular interactions. It also argues against large conformational changes in the hydrophobic portion of the protein upon virus assembly.

The three phenylalanines are among the few hydrophobic amino acids in the hydrophilic end regions of the protein structure. Chemical shift resolution among the ring γ carbons and the ring protons reflects their being in different environ-

ments. The aromatic regions of the ^{13}C spectra are dominated by very narrow phenylalanine ring resonances. The large NOE's, long T_1's, and small line widths of these resonances indicate that there must be phenylalanine residues undergoing rapid motion in the protein. The remaining question is whether all three Phe rings are rapidly moving. The use of [$^2H_5$]-Phe-labeled proteins and difference spectroscopy shows that one of the Phe's is less mobile and gives rise to broad peaks dispersed in the spectrum, one of which is observed as a nontitrating resonance only at high pH with the tyrosine protons shifted out of the way. Broad intensity with small NOE from a single phenyl ring could easily be lost in the 128–132-ppm region of overlapping lines on its ^{13}C spectrum. It is not possible to say which of the phenylalanines this is at the present time, with all three phenylalanines in α-helical regions according to secondary structure predictions based on the amino acid sequence.

References

Cavalieri, S. J., Goldthwait, D. A., & Neet, K. E. (1976) *J. Mol. Biol. 102*, 713–722.

Chang, C. N., Blobel, G., & Model, P. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 361–365.

Cross, T. A., & Opella, S. J. (1980) *Biochem. Biophys. Res. Commun. 92*, 478–484.

Cross, T. A., Gall, C. M., & Opella, S. J. (1981) in *Proceedings of the 7th Biennial Bacteriophage Assembly Meeting*, Liss, NY (in press).

Day, L. A. (1969) *J. Mol. Biol. 39*, 265–277.

Day, L. A., & Wiseman, R. L. (1978) in *The Single-Stranded DNA Phages*, pp 605–626, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Denhart, D. T., Dressler, D., & Ray, D. S., Eds. (1978) in *The Single-Stranded DNA Phages*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Dunker, A. K., Williams, R. W., & Peticolas, W. L. (1979) *J. Biol. Chem. 254*, 6444–6448.

Green, N. M., & Flanagan, M. T. (1976) *Biochem. J. 153*, 729–732.

Hagen, D. S., Weiner, J. H., & Sykes, B. D. (1978) *Biochemistry 17*, 3860–3866.

Hagen, D. S., Weiner, J. H., & Sykes, B. D. (1979) *Biochemistry 18*, 2007–2012.

Knippers, R., & Hoffmann-Berling, H. (1966) *J. Mol. Biol. 21*, 281–292.

Kornberg, A. (1974) *DNA Synthesis*, W. H. Freeman, San Francisco, CA.

Lim, V. I. (1974) *J. Mol. Biol. 88*, 873–894.

Makino, S., Woolford, J. L., Tanford, C., & Webster, R. (1975) *J. Biol. Chem. 250*, 4327–4337.

Marvin, D. A., Pigram, W. J., Wiseman, R. L., Wachtel, E. J., & Marvin, F. J. (1974) *J. Mol. Biol. 88*, 581–600.

McDonald, C. C., Phillips, W. D., & Glickson, J. D. (1971) *J. Am. Chem. Soc. 93*, 235–246.

Miller, J. H. (1972) *Experiments in Molecular Genetics*, p 431, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Nakashima, Y., & Konigsberg, W. (1974) *J. Mol. Biol. 88*, 598–600.

Newman, J., Swinney, H. L., & Day, L. A. (1977) *J. Mol. Biol. 116*, 593–606.

Nozaki, Y., Reynolds, J. A., & Tanford, C. (1978) *Biochemistry 17*, 1239–1246.

Opella, S. J., & Cross, T. A. (1980) *J. Magn. Reson. 37*, 171–175.

Opella, S. J., Nelson, D. J., & Jardetzky, O. (1974) *J. Am. Chem. Soc. 96*, 7157–7159.

Opella, S. J., Cross, T. A., DiVerdi, J. A., & Sturm, C. F. (1980) *Biophys. J. 32*, 531–548.

Smilowitz, H., Carson, J., & Robbins, P. W. (1972) *J. Supramol. Struct. 1*, 8–18.

Snyder, G. H., Rowan, R., III, Karplus, S., & Sykes, B. D. (1975) *Biochemistry 14*, 3765–3777.

Williams, R. W., & Dunker, A. K. (1977) *J. Biol. Chem. 252*, 6253–6255.

Wuthrich, K. (1976) in *NMR in Biological Research: Peptides and Proteins*, pp 65–122, American Elsevier, New York.