Townend, R., and Timasheff, S. N. (1960), *J. Am. Chem. Soc. 82*, 3168.

Williams, M. A., Gregory, D. W., and Harrison, P. M. (1966), *Biochem. J. 99*, 10p.

Magnetic Resonance Studies of Macromolecules. I. Aromatic–Methyl Interactions and Helical Structure Effects in Lysozyme*

H. Sternlicht and D. Wilson

ABSTRACT: Lysozyme nuclear magnetic resonance spectra display a number of upfield resonances (some occur above tetramethylsilane) which change in a striking manner as the enzyme denatures. These resonances arise primarily from methyl–aromatic magnetic interactions. A comparison of optical rotatory dispersion measurements with the changes in these upfield resonances during thermal denaturation indicates that helical region unwinding occurs simultaneously with the opening up of the protein. Furthermore, a preliminary reconstruction of the upfield resonances based on the X-ray crystallographic coordinates and aromatic ring current data gives results which are in good agreement with the observed spectra. Barring the possibility of a fortuitous agreement, this is highly suggestive that the enzyme structure in solution and in the crystal state are very similar, if not identical. Some efforts at direct experimental identification of the upfield resonances are also briefly reported. The lysozyme spectra show changes in the α-hydrogen region upon folding. Electric field effects, which arise from the permanent dipoles of the peptide units in a helix configuration, and, probably, magnetic anisotropy contributions from the carbonyls are important. Resonances from α-hydrogens in infinite right-handed helices are predicted to occur $\sim$0.5 ppm downfield of the corresponding resonance for infinite left polypeptides. The P_E resonance (text) in the α-hydrogen region grows as the enzyme folds. This resonance may be associated with shifts arising from helical ordering.

Magnetic resonance as a technique to study protein conformation has been explored by a number of investigators over the past decade (Kowalsky and Cohn, 1964; Kowalsky, 1962; Mandel, 1965). Recently high-resolution proton resonance studies at 220 Mcycles of a number of enzymes have been reported by Phillips and collaborators (McDonald and Phillips, 1967). These studies clearly establish the magnetic resonance technique as a powerful tool for protein investigations. A few of the areas where significant contributions from magnetic resonance are imminent follow. (1) (protein structure) The first X-ray crystal structure of an enzyme (lysozyme) has recently been determined by Phillips and collaborators (Phillips, 1966; Blake *et al.*, 1965). Questions have been raised concerning the nature of the enzyme structure when in solution. M. A. Raftery (private communication) has indirect evidence of the similarity of structures from spectrophotometric and magnetic resonance studies of substrate and inhibitor binding to lysozyme in solution. It is our intention in this paper to present more direct experimental verification of the similarity of crystal and solution structures; (2) protein folding and hydrophobic interactions (Schachman, 1963); and (3) practical calculations of protein conformation. Within the past few years, considerable progress has been made in understanding the stereochemical principles of polypeptide chain formation (Ramakishnan and Ramachandran, 1965; Brant and Flory, 1965; Leach *et al.*, 1966a,b; Vanderkooi *et al.*, 1966; Liquori, 1966). Efforts are now being made to "predict" protein conformation (Leach *et al.*, 1966a,b; Vanderkooi *et al.*, 1966; Levinthal, 1966). The results can be checked against nuclear magnetic resonance observations, where practical.

In this laboratory, over the past year, we have been concerned with nuclear magnetic resonance studies in proteins. In particular we have been interested in spectral changes that occur as a consequence of the tertiary protein structure. We have limited ourselves primarily to lysozyme since any conclusions reached by magnetic resonnace could ultimately, in principle, be checked against Phillip's (1966) X-ray determination. Our own study at 100 Mcycles has benefited considerably from the parallel studies done at 220 Mcycles (McDonald and Phillips, 1967). In this paper we discuss two aspects of tertiary protein structure: hydrophobic interactions and helical structure. By

* From the Department of Chemistry, University of California, Berkeley, California 94720. *Received April 3, 1967.* This study was aided by Grant No. GM-14313-01 from the National Institutes of Health.

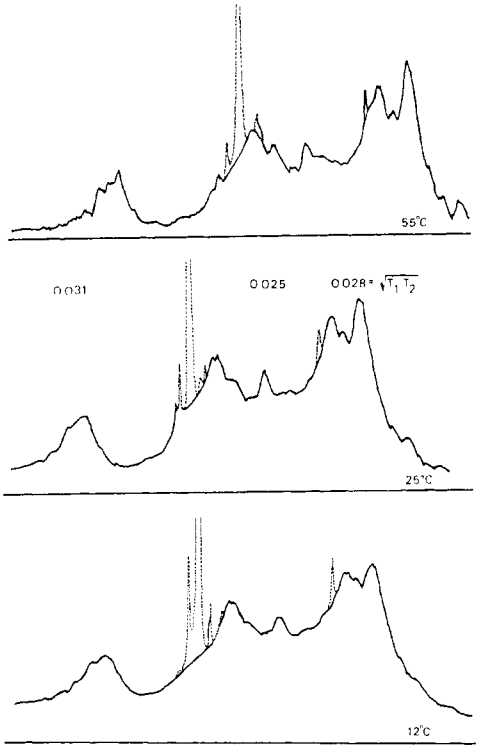


FIGURE 1: Lysozyme, 10% by weight, pD 7.4.

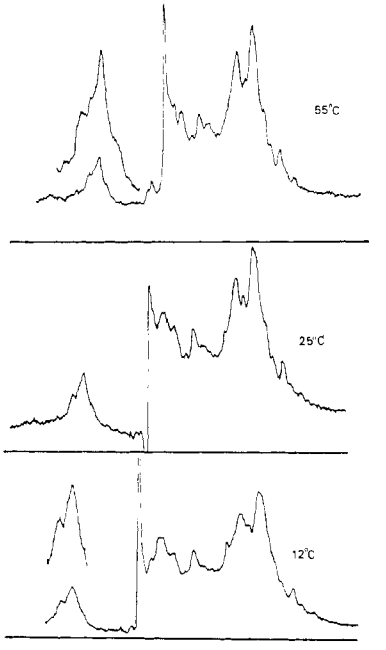


FIGURE 2: Lysozyme, 10% by weight, pD 4.2. The aromatic regions are enlarged.

hydrophobic interactions we mean, more specifically, interactions primarily between the methyls of valine, leucine, and isoleucine with aromatic residues as evidence by upfield shifts of the methyls in the folded configuration.

Results

The lysozyme used was supplied by Sigma Chemical Co. Deuterium replacement of the exchangeable protons was accomplished by dialyzing the H_2O solution against considerable excess of D_2O for approximately 24 hr, and then lyophilizing once or twice. This procedure was found to give good exchange results with virtually no denaturation precipitate. Lysozyme is very soluble, and causes no obvious change in viscosity even at high enzyme concentrations of 10%. The nuclear magnetic resonance spectra (Figures 1 and 2) done at 100 Mcycles are a function of pH and temperature as has already been noted by others (McDonald and Phillips, 1967). The increased line width with pH arises from the dimerization of lysozyme with increasing pH. One concludes from a centrifuge study (Sophianopoulos and Van Holde, 1964) that at pH 7, a 10% by weight lysozyme solution in H_2O is approximately 55% dimer, whereas at pH 5 it is approximately 10% dimer, while at pH 4.2 it is essentially all monomer. By increasing temperature one decreases the amount of dimerization. This is qualitatively confirmed by nuclear magnetic resonance. The spectra at pD 4.2 are considerably sharper and less temperature sensitive than the pD 7.4 spectra. The pD 7.4 spectra sharpen dramatically with increasing temperature as a result of dimer breakup. It is estimated that dimerization broadens the resonance by a factor of about 1.5–2. This broadening suggests increased rigidity in the lysozyme complex, and may be a result of the increased mass or may possibly be related to some definite structural feature of lysozyme. A "wing" of one monomeric unit may, for example, fit into the cleft formed by the two wings of the second, restricting considerably the "breathing" motion of the monometic unit.

Lysozyme has 700 nonexchanging protons (approximately five-sevenths of the total number of protons are nonexchanging). This agrees within 10% with the experimental integrated intensity. A gross decomposition of the spectra into seven main groups based on earlier work (Mandel, 1965; Bovey and Tiers, 1959) is shown in Figure 3. The assignment in the aromatic region follows Mandel. A "saturation" study was done in an effort to distinguish broad line resonances from narrow line resonances. A number of regions were selected (Figure 1b, pD 7.4, temperature 25°) and progressively saturated. The resonance intensity as a function of radiofrequency field was fitted to the Bloch equations and an effective $\sqrt{T_1T_2}$ was calculated (T_1 and T_2 denote the spin-lattice and transverse relaxation times, respectively). The $\sqrt{T_1T_2}$ values varied by approximately 20% about an average value of 0.025 sec suggesting that if T_1 is approximately T_2, the spectra consists of a large number of unresolvable resonances with an average intrinsic line width of $\Delta\nu = (\pi T_2)^{-1}$ approximately 15 cycles/sec. This would

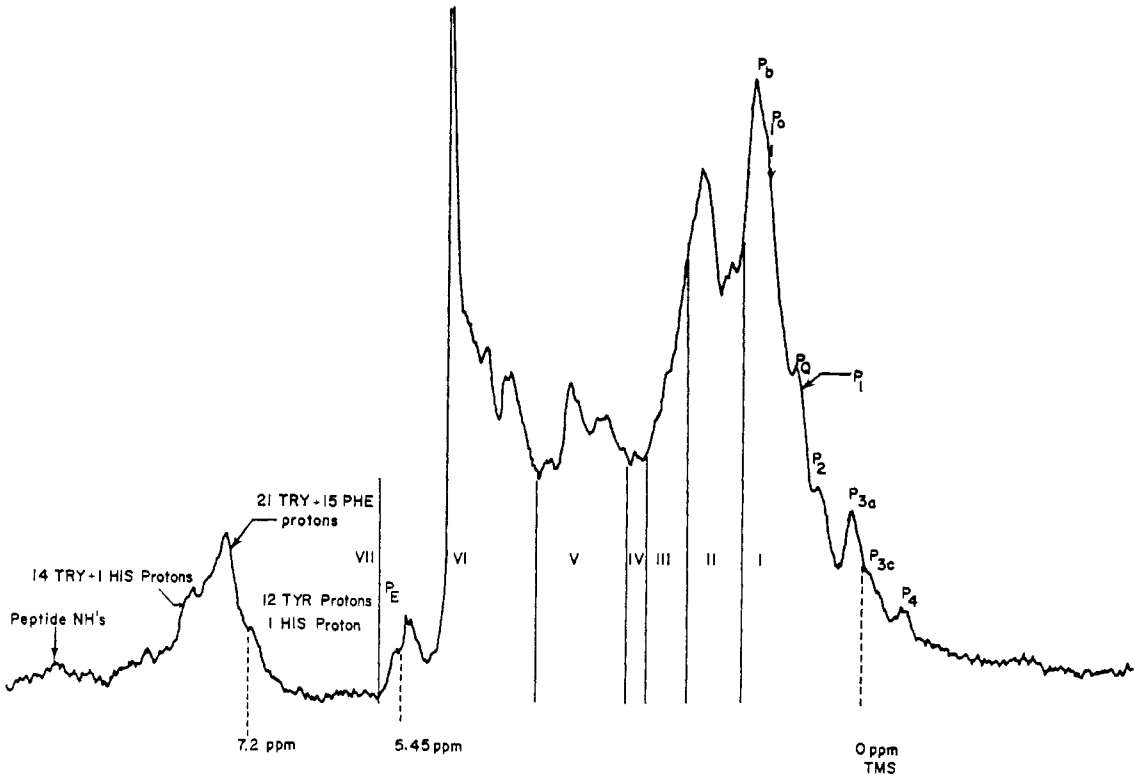


FIGURE 3: Lysozyme, pD 4.2, 55°. A decomposition of the spectrum into seven main groups. The upfield P_i resonances are discussed in the text and in Table I. Group I: CH_3 in valine, leucine, and isoleucine; group II: $CHCH_2$ in leucine, $(CH_2)_3$ in lysine, CH_2 in isoleucine, CH_3 in alanine and threonine; group III: CH in valine and isoleucine, $(CH_2)_2$ in arginine, proline, glutamine, and glutamic acid; group IV: $(CH_2)_2$ in glutamine and glutamic acid; group IV: $(CH_2)_2$ in glutamine and glutamic acid; group V: β-hydrogens except for those appearing in the above groups; group VI: α-hydrogens in peptide chain; group VII: aromatic hydrogens.

imply that the intrinsic line width of the pD 4.2 spectra was smaller than 15 cycles/sec at room temperature, and still smaller at the elevated temperatures. Preliminary attempts at a line-fitting reconstruction of the resonance spectra and a comparison with the 220-Mcycle spectra suggest that the average width at approximately 55°, pD 4.2 is approximately 15–18 cycles/sec.

Although any conclusions drawn from a saturation study of this type must be regarded with caution, it would appear that the methyl region I is somewhat sharper than the average. This could be ascribed to the methyls retaining more rotational motion than the bulk of the protons. The central resonance of the aromatic region, for which a value of approximately 0.031 sec is estimated for $\sqrt{T_1T_2}$, also appears sharper than the average. This central resonance comes from the phenylalanine-tryptophan region. There are three or four tryptophans in the cleft regions, Trp 62, 63, 108, and 111, which may have more motional freedom than other aromatics. (After examining a Corey–Pauling–Koltum space-filling model built by M. A. Raftery from Phillip's X-ray coordinates we feel that this hypothesis is very reasonable.) The α-hydrogen region could not be studied by saturation because of overlap with the HDO resonance. The rather small intrinsic line widths found in lysozyme are interesting considering that water protons have line widths of $\sim$1 cycle/sec.

Studies of the α-hydrogen region are obscured by the HDO resonance. We could not establish the presence of a helix-induced broadening of the "intrinsic" line width. It would appear inferentially from a relative comparison of the α-hydrogen region with other regions of the spectra as to temperature and pH sensitivity, and from an examination of the 220-Mcycle spectra that the broadening is not significant. The α-hydrogen region shows changes during folding over essentially the whole region which we do not understand. One striking feature, for example, which can be studied even at 100 Mcycle is the appearance upon folding of a doublet resonance centered at $P_E = 4.6$ with an integrated intensity of $\sim$15 protons. The increase in intensity appears to follow Figure 4b below.

The valine (Val), leucine (Leu), isoleucine (Ile) methyl regions occur at $P_O = \tau$ 9.0 (on the τ scale

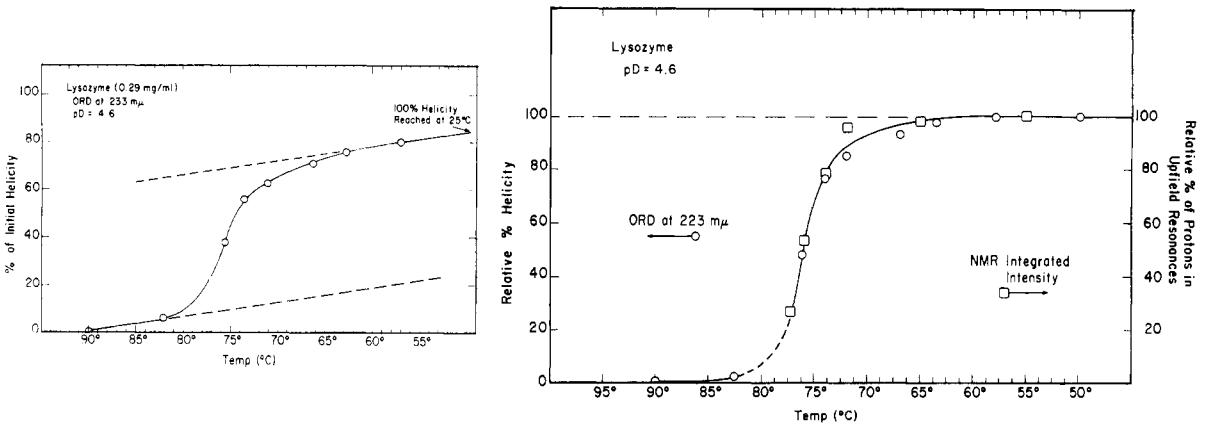


FIGURE 4: Optical rotatory dispersion and resonance measurements. (a) (left) Optical rotatory dispersion measurement (233 mμ) or lysozyme, 0.03 by weight, pD 4.8. We have arbitrarily assumed 100% relative helicity at 25° and 0% at 90°. (b) (right) A comparison of the upfield resonance intensity (P_Q plus P_1–P_4) as a function of temperature with optical rotatory dispersion after correcting for the linear tail of part a.

TABLE I: Upfield Resonance Positions[a] (τ scale) and Intensity.

Mcycles	P_B	P_O	P_Q	P_1	P_2	P_3	P_4
			Position				
100 (pD 4.2)	8.7	N.O.[b]	9.16	N.O.	9.47	(9.84) (N.O.) (10.0)	10.45
(pD 7.4)	8.7	N.O.	9.14	N.O.	9.49	(9.82) (N.O.) (10.0)	10.5
220	8.75	9.0	9.10	9.23	9.50	(9.85) (9.92) (10.0)	10.45
			Intensity[c] (number of protons)				
	85	13	9	15	8	13	3

[a] Positions given are estimated to be accurate to within τ ±0.07. [b] N.O. = not observed. The P_O resonance in the folded enzyme is obscured by P_B in the 100-Mcycle case. [c] Estimated error within 20% on the average. P_Q and P_1 may be in error by as much as 35%. P_B consists of ~85 protons, approximately 55 of which are methyl protons of valine, leucine, and isoleucine which shift downfield from P_O when the enzyme folds.

according to which the TMS[1] peak is τ +10.00 by definition) in TFA and water solutions for unfolded proteins. A number of resonances P_Q, P_2, P_{3A}, P_{3C}, and P_4 occur upfield of P_O with maxima at τ 9.16, 9.47, 9.84, 10, and 10.45 ± 0.07, respectively, in aqueous solution under folded conditions (Figure 3). The P_Q resonance shows a hint of a shoulder at τ 9.25. In Table I we compare the 100-Mcycle resonance results with the values we deduce from the 220-Mcycle spectra. The agreement as to position is quite good. At the higher frequency P_Q and P_3 are seen to split into P_Q and P_1, P_{3A}, and P_{3B}, respectively. A small resonance is now present at P_O. The estimate of the number of protons in each resonance is difficult and may be in error by as much as 20% on the average. Raising the temperature sharpens the resonances until ~71° is reached whereupon the resonances P_B, P_Q, and P_1–P_4 decrease with further temperature increase while resonance P_O increases. (The intensity decrease in the P_B, P_A, and P_1–P_4 upon unfolding essentially reappears in P_O so that the over-all integrated intensity remains constant.) The 220-Mcycle study of the unfolding is remarkable indeed for its clarity and potential for quantitative evaluation. In Figure 4 we present a comparison of the

[1] Abbreviations used: TMS, tetramethylsilane; TFA, trifluoroacetic acid.

integrated intensity of the upfield resonances P_Q plus P_1–P_4 taken from the 220-Mcycle study (10% lysozyme by weight) with the optical rotatory dispersion measurements done at 233 mμ (0.03% lysozyme by weight). The optical rotatory dispersion presumably measures helicity and was done on a Cary 60 spectrapolarimeter. The nuclear magnetic resonance changes shown are relative and temperature reversible. The temperature range over which unfolding occurs is essentially the same in the two types of measurements despite the factor of 300 difference in concentration. This would indicate that there are no concentration-dependent stability factors that influence the unfolding. (In this connection we have examined spectra of 1% lysozyme (pD 4.2) and have seen the same upfield resonances with the same intensity distribution.) There is some arbitrariness in the interpretation of the optical rotatory dispersion data. The optical rotatory dispersion shows a linear tail below 65° and above 80°, respectively. In Figure 4a we have arbitrarily taken the helicity to be 100% at 25° and 0% at 90°, and have assumed a two-state transition from a folded to an unfolded configuration. The nuclear magnetic resonance integrated intensity, on the other hand, is essentially constant below 70° although changes can be observed in other regions such as the aromatic at lower temperature. However, there are no large changes in the nuclear magnetic resonance between 70 and 25° that would suggest that a large change in helicity between 70 and 25° is reasonable. Consequently in Figure 4b we have redefined the end states by subtracting out the linear regions on the assumption that they represent factors other than tertiary structure; *i.e.*, we extrapolate the linear regions of Figure 4a and assume that the upper line represents the completely folded state while the lower line represents the completely unfolded state. The two extrapolated lines are parallel to each other as one would expect them to be if this end-state interpretation were correct.

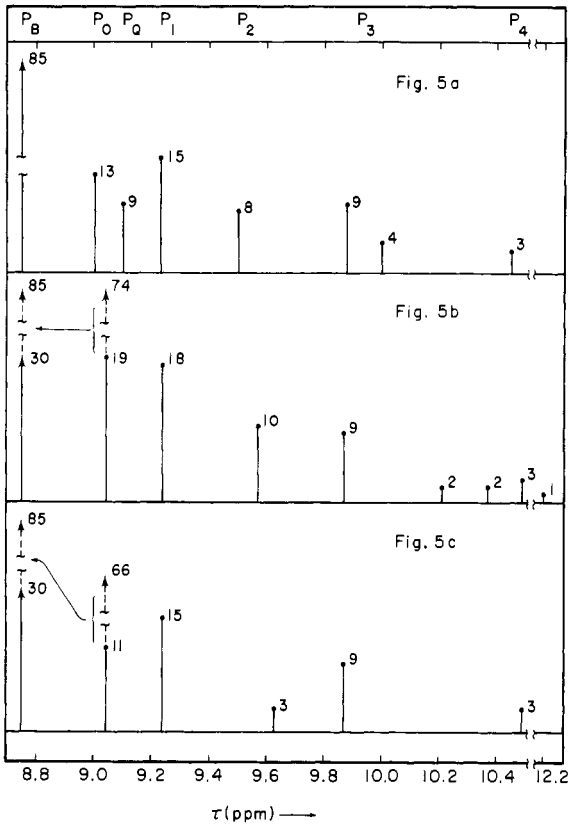


FIGURE 5: The number of protons in the upfield resonances. (a) The observed number. (b) The predicted number using X-ray coordinate data and ring-current shifts. Both group I and II protons contribute. (c) The group I (methyl) proton contribution only.

Discussion

Methyl–Aromatic Interactions. Below we show that the upfield resonances arise primarily from methyl–aromatic magnetic interactions as first suggested by the Du Pont group (McDonald and Phillips, 1967). Neither the methyl nor the aromatic residues are necessarily in the helical region. The methyl–aromatic interactions come from the three-dimensional folding of the enzyme. In Table I we estimate that ~48 protons occur in the upfield resonances. Assuming no group II (Figure 3) contributions to the upfield resonances, this means that a maximum of 8 residues of valine, leucine, and isoleucine out of a total of 20 are somehow perturbed. Large upfield shifts generally occur when protons are in the diamagnetic, secondary field regions produced by the large ring currents of aromatics (Johnson and Bovey, 1958). Electric fields from ionized groups or permanent dipoles can also cause appreciable upfield shifts (Buckingham *et al.*, 1960). The planar peptide has a permanent dipole moment. In the case of a helix, these dipoles can add and produce large fields. We have estimated the strength of these fields at the methyls and conclude that while they might make some contribution to P_Q, they certainly should not contribute significantly to the other upfield resonances.

In Figure 5 we present a comparison of the observed shifts and intensities with those calculated using the X-ray lysozyme coordinates and Johnson–Bovey tables (Johnson and Bovey, 1958) for ring-current-induced shifts. (The lysozyme coordinates were privately communicated to this laboratory by D. C. Phillips with the clear understanding that they are preliminary in nature and subject to further refinement.) The calculation of the shift positions was done by treating all heteronuclear aromatics as if they were phenyl rings. Tryptophan was replaced by two fused phenyl rings, each making a shift contribution of a lone phenyl group. The group II proton positions were determined by assuming tetrahedral bond angles, using the X-ray coordinates of the carbon positions. The methyl protons of group I, on the other hand, were assumed to be located at the center of the base of the cone formed by the freely rotating CH_3 group about the C–C bond.

TABLE II: Predicted Shifts into Region above τ 9.0.[a]

Residue	Proton(s)	Interacting Aromatics	Unfolded[a,b] Resonance Position	Predicted[c] Upfield Shift	No. of Protons	Spectrum Region
Val 92	Methyl	His 15	9.06	+0.20 (0.20)	3	P_1
Leu 8	Methyl	Phe 3	9.04	+0.18	3	P_1
	Methyl	Phe 3	9.04	+0.76	3	P_3
Leu 17	Methyl	Tyr 20, Trp 28	9.04	+1.45 (0.31)	3	P_4
	Methyl	Trp 28	9.04	+0.59 (0.26)	3	P_2
	Methylene (C_β)	Trp 28	8.29	+1.92 (1.71)	1	P_3–P_4
Leu 56	Methyl	Trp 108	9.04	+0.25 (0.16)	3	P_1
Leu 75	Methyl	Trp 63	9.04	+0.12	3	P_1
Ile	Methyl (C_δ)	Phe 3, His 15	9.00	+0.31 (0.21)	3	P_1
	Methylene (C_γ)	His 15	8.65	+0.75 (0.75)	1	P_1
Ile 98	Methyl (C_δ)	Trp 63, Trp 108	9.00	+0.92 (0.87)	3	P_3
	Methylene (C_{γ_1})	Trp 63, Trp 108	8.65	+1.70 (0.54)	1	P_4
	Methylene (C_{γ_1})	Trp 63, Trp 108	8.65	+3.55 (1.92)	1	
	Methyl (C_{γ_2})	Trp 63, Trp 108	9.00	+0.88 (0.26)	3	P_3
Thr 51	Methyl	Tyr 53	8.68	+0.89	3	P_2
Lys 1	Methylene (C_δ)	Phe 3	8.35	+1.15	1	P_2
Lys 33	Methylene (C_δ)	Phe 34, Phe 38, Trp 123	8.35	+1.29	1	P_2
	Methylene (C_γ)	Phe 34, Phe 38, Trp 123	8.35	+2.04 (0.05)	1	P_4
	Methylene (C_γ)	Phe 34, Phe 38, Trp 123	8.35	+0.88 (0.03)	1	P_1
	Methylene (C_β)	Phe 34, Phe 38, Trp 123	8.35	+0.88 (0.22)	1	P_1
Lys 96	Methylene (C_δ)	Tyr 20	8.35	+1.23	1	P_2
	Methylene (C_γ)	Tyr 20, Trp 28	8.35	+1.09	1	P_2
Lys 116	Methylene (C_β)	Trp 111	8.35	+1.87 (0.49)	1	P_3–P_4

[a] Relative to sodium 2,2-dimethyl-2-silapentane-5-sulfonate at τ 10.0. [b] Based on Mandel's 100-Mcycle data. [c] The contributions to the predicted shifts from five-membered rings in histidine and tryptophan are indicated in parentheses.

An IBM 7094 computer was employed in all calculations.

In Table II we present a detailed breakdown of the upfield shifts into the region of interest. Both regions I and II (Figure 3) contribute to the upfield resonances. The contributions from the group I protons are shown separately in Figure 5c. All calculated resonances which fall within $\lesssim$0.14 ppm of each other are lumped together and assigned a common resonance position. This is reasonably consistent with the experimental observation that resonances falling within ~15 cycles/sec of each other are unresolvable.

The agreement between the observed and calculated resonances as to position and intensity is quite good. For example, one can assign the highest upfield resonance involving three protons at τ 10.45 to the Tyr 20, Trp 28, and Leu 17 interaction. A resonance involving a single methylene proton from Ile 98 is calculated to occur still further upfield at τ 12.2, but was not observed in the single sweep experiments at 100 Mcycles. However, because of the approximate nature of the calculation, particularly with respect to five-membered ring contributions, the detailed shift assignments given in Table II must be regarded with caution. Apparently, however, the broad intrinsic line width in lysozyme permits one to use an approximate calculation procedure to reproduce the detailed pattern of upfield resonances.

Because of the approximate nature of the Johnson–Bovey tables as applied to heteroatomic aromatic systems, an independent experimental assignment of these upfield resonances would certainly be desirable. Tryptophan 62 can be modified to an oxidized indole by *N*-bromosuccinimide (Hayashi *et al.*, 1965). We have examined the nuclear magnetic resonance of the modified enzyme. According to our assignment (Table II) tryptophan 62 does not contribute to the upfield resonances. If converting tryptophan 62 to the oxidized indole does not alter the protein structure, we expect no change in the upfield resonances. This is indeed seen to be the case. The trimer–(*N*-acetylglucosamine)–lysozyme complex has been shown from X-ray studies (Blake *et al.*, 1967) to involve interactions of the trimer with the aromatic residues tryptophans 63 and 108. According to our assignment (Table II) we predict possible large nuclear magnetic resonance changes in P_3 and P_4 resonances of the complex since these regions apparently involve contributions from tryptophans 63 and 108.

In a still further effort to identify the upfield resonance H. Sternlicht and E. Wheeler (unpublished

results) spin-labeled lysozyme with a nitroxide radical, *N*-(1-oxyl-2,2,5,5-tetramethylpyrrolidinylmaleimide (Griffith and McConnell, 1966). It was our hope that the unpaired spin of the radical would magnetically perturb selective regions of the enzyme. Unfortunately, no magnetic perturbations could be seen in our 100-Mcycle studies.

Returning to the experimental spectra, it is observed that upon folding, approximately 55 protons shift from the P_O position downfield 0.25 ppm to the P_B position, ~13 remain at P_O, and ~48 appear in the upfield regions. Our calculations attribute 33 of these upfield protons to the original P_O group I protons, and 12 to group II protons. Our calculations further predict that only nine group I protons will be significantly shifted downfield, so that the bulk downfield shift (P_B) of 0.25 ppm does not arise from methyl–aromatic paramagnetic interactions. Indeed, we calculate that ~74 protons should appear in the P_O position. We hypothesize that the ~13 unshifted protons remaining in the P_O position correspond to methyls of valine, leucine, or isoleucine which are on the enzyme surface, and consequently are exposed to a water environment. The 55 protons shifted into the P_B regions, on the other hand, are probably in a hydrophobic environment in the protein interior but not near an aromatic. We have therefore indicated (Figure 5) a subtraction of 55 from 74, leaving 19 protons in the $P_O + P_Q$ region to depict a more realistic representation of the calculated spectrum.

The good agreement between the reconstructed upfield resonances and the experimental spectra as shown in Figure 5 must be regarded with caution because of the considerable experimental and theoretical uncertainties. However, in the absence of any evidence to the contrary, this agreement is highly suggestive that the enzyme structure in solution is similar, if not identical with that in the crystal. This is reasonable since the crystal has considerable mother liquid surrounding the protein.

It is currently hypothesized that folding *in vivo* may first involve residues near the amino acid end which act as nucleation regions for the rest of the protein. One sees from Table II that the upfield resonances involve contributions over the whole extent of the polypeptide. Within the experimental uncertainties, it would appear, at least during thermal unfolding, that there are no preferred regions which might act as nucleation regions for the subsequent folding of the rest of the protein upon cooling. The question of folding under *in vivo* conditions is still an open one.

An optical rotatory dispersion study at 233 mμ was undertaken to correlate the helix unwinding with the opening up of the enzyme as revealed by the disappearance of the upfield resonances in the transition temperature region. It was of some interest to us to establish whether during thermal unfolding (a) certain regions of the enzyme first move away from each other thus exposing helical regions to the solution with subsequent helix unwinding; (b) helical unwinding occurs first; (c) both helical unwinding and relative regional motions occur simultaneously. The agreement between the optical rotatory dispersion and the 220-Mcycle nuclear magnetic resonance measurements of McDonald and Phillips (Figure 4) was excellent. This indicates that helix unwinding occurs essentially simultaneously with the opening up of the enzyme during thermal unfolding. (Some caution must be exercised, however, since (I) Figure 4b is based on measurements in two different laboratories and there may well be experimental temperature discrepancies and (II) there is some uncertainty in defining the end states for the optical rotatory dispersion experiment.)

Helical Structure Effects. Lysozyme has ~42% helicity (right handed). There are six helical regions ranging from 6 to 13 peptides in length. One expects helical structure to produce changes in the nuclear magnetic resonance spectra. (I) It is known from polymer studies (Bovey *et al.*, 1959) that, barring polymer association, helix formation increases the α-hydrogen line width slightly relative to the unfolded state. (II) The permanent dipoles of the planar amide residues align in the helix. The electric field of the aligned dipoles may be sufficiently large to produce significant chemical shifts of the α-hydrogens and other near protons. (III) The planarity of the amide structure is ascribed to the partial double bond between the nitrogen and the carbonyl carbon. The magnetic anisotropy coming from the circulating currents in the nitrogen–carbonyl system can produce secondary fields which may add in the helix configuration to produce significant shifts. We have not been able to ascribe any resonance change in the α-hydrogen region with any degree of confidence to the helical structure. In this section we wish only to draw attention to the possibility of an appreciable downfield shift of certain α-hydrogens in right-handed helical regions as a result of electric field effects. (There may also be significant magnetic anisotropy contributions. We are unable to satisfactorily estimate its contribution. Some years ago Narasimham and Rogers (1959) studied the proton nonequivalence of the two nitrogen protons in formamide. To the best of our knowledge there is still no better treatment. They approximated the shielding field by a magnetic dipolar field coming from a dipole located on the C–O bond axis. The components of the induced magnetic dipole were determined by the principal magnetic susceptibilities of the C–O bond (X_{xx}, X_{yy}, and X_{zz}). The nitrogen protons are a distance r from the magnetic dipole.

Considerable criticism can be raised against the Narasimham–Rogers treatment. (1) It is a point dipole treatment, which is questionable since the C–O bond length is not small compared with r (Didoy and Gerry, 1961). (2) It is an empirical treatment which derives the X_{ii}'s from the observed shifts by varying the dipole location. If a mechanism other than the magnetic anisotropy were contributing significantly or even dominating the shift (*e.g.*, electrical field effect) this would not be known from the Narasimham–Rogers treatment, since it is an operational procedure concerned only with obtaining an effective X which would

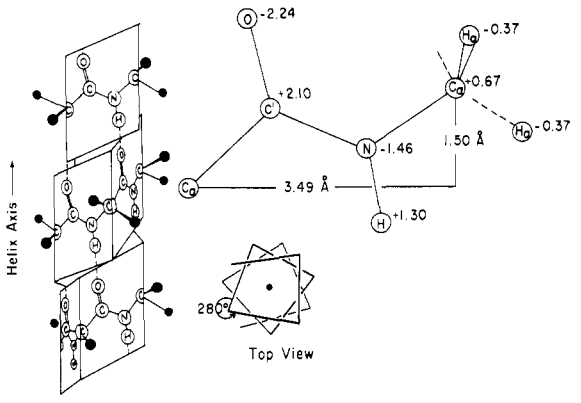


FIGURE 6: Monople assignment (in units of 10^{-10} esu) and α-helix parameters.

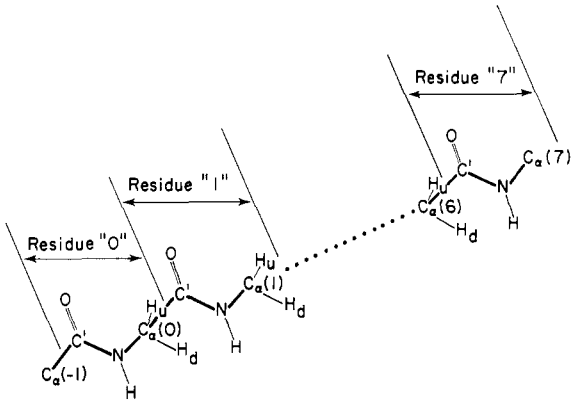


FIGURE 7: Atom- and residue-labeling scheme for an eight-residue helix for Tables III and IV.

reproduce the data. (3) Subsequent work by Richards (Hatton and Richards, 1960, 1962) on spin–spin interaction in formamide and dimethylformamide suggests that the *cis* proton (*cis* to the carbonyl) of formamide and the *cis*-methyl of dimethylformamide are upfield in the nonequivalent proton spectra. Nuclear Overhauser studies of proton saturation (Anet and Bourn, 1965) in dimethylformamide confirms Richards assignment. The Narasimham–Rogers best choice of X_{ii}'s, on the other hand, corresponds to the *cis* proton and *cis*-methyl being downfield in direct conflict with the experimental assignment.)

Electric field effects on protons shifts have been discussed by a number of authors (Buckingham *et al.*, 1960; Buckingham, 1960; Emsley *et al.*, 1965; Musher, 1962, Fraenkel *et al.*, 1960; Hruska *et al.*, 1963). Electric field effects in ions, in substituted halogen and nitro compounds, and in polar solvents have been clearly established. Buckingham (1960) estimates that the electric field ($\hat{E}$) contribution to the C–H proton chemical shift in solution is

$$\sigma_E = -2 \times 10^{-12} E_z \text{ (cgs)} \qquad (1a)$$

Subsequent estimates of the constant of proportionality vary from -2.5 to -3.5×10^{-12} (Musher, 1962). We shall use

$$\sigma_E = -2.5 \times 10^{-12} E_z \text{ (cgs)} \qquad (1b)$$

E_z is the component of $\hat{E}$ along the bond direction. The positive z is directed from carbon to hydrogen. A positive E_z draws electrons from the hydrogen and decreases the electronic shielding causing a downfield shift in the resonance (*i.e.*, σ_E is negative). According to the above estimate, $\sigma_E = -0.25$ ppm when $E_z = 10^5$ cgs (the field at the C–H hydrogen produced by a proton 7 A away along the C–H axis). The electric field clearly can be significant in protein systems, and should be examined.

Planar amides have permanent dipoles. Formamide (Kurland and Wilson, 1957), for example, has a dipole moment 3.8 D. directed from the oxygen to the nitrogen. If one subtracts out the N–H bond dipole in formamide, which is 1.3 D. directed from N to H (Smythe, 1938, 1955), one is left with the dipole contribution of the N═C═O bond system which we denote as $\mu_{amid.}$ [$|\mu_{amide}| = 2.9$ D.]. (The CH bond dipole in sp^2 hybridization in formamide is taken as zero; Gent, 1948.) In the case of a polypeptide a reasonably accurate dipole moment per residue can be calculated. One of the N–H bonds of formamide is replaced by a N–C_α bond with a dipole moment 0.45 D. directed from N to C_α (Smythe, 1938, 1955). The carbonyl (C′) hydrogen of formamide is similarly replaced by a C_α. The C′–C_α bond has a zero dipole moment. The C_α–H bond (sp^3) has a dipole moment 0.4 D. directed from the hydrogen to C_α which is opposite to what one expects on the basis of electronegativity (Gent, 1948). μ_{amide} is assumed to be the same as in formamide. Contributions to the total dipole moment from dipoles on R, the side chain bonded to C_α, can be included when necessary. For the purposes of simplicity consider polyglycine. We are interested in the electric fields at the α-hydrogens in an ordered α-helix configuration. The distance of closest approach of the α-hydrogens to some of the neighboring peptide atoms may be less than 2.5 A. The use of the dipole approximation to estimate electric fields would cause considerable error. Instead, we use a set of uniquely determinable monopoles; *i.e.*, charges at the various amide atoms, which at large distances reproduce the calculated dipole moment (Figure 6).

The bond angles and bond positions for the planar amide used are those given by Pauling (1960). The C_α sp^3 bond angle was taken as 109° 30 min. A right-handed α-helix was generated by requiring the translation along the helix axis to be 1.50 A/residue, and that there be 3.6 residues/turn (Figure 6). The "up" end of the helix is the carboxyl-terminal end while the "down" C_α–H of polyglycine is directed toward the NH_2 end. To the extent that electric field contributions from the R residues can be neglected, the electric field

TABLE III: The Contribution from Peptide m to the Electric Field Shifts[a] of $\alpha H(n)$.

	m								Total Shift
n	0	1	2	3	4	5	6	7	(ppm)
αH_u 0	~0.21	−0.87	−0.08	−0.01					−0.75
αH_u 6				0.07	0.17	0.03	0.21	−0.87	−0.39
αH_d 0	+0.58	−0.38	−0.05	−0.10					~0.05
αH_d 6					−0.08	0.05	0.58	−0.38	~0.15

[a] Shifts are relative to an abstracted $C_\alpha<^{H}_{H}$.

at the up α-H approximates the field at the α-H of *all* right-handed L-amino acid helices, whereas the field at the down α-H approximate the field at the α-H of *all* left-handed L-amino acid helices. In other words, the difference in the electric field induced shifts between the up and down α-hydrogen of polyglycine right-handed helix corresponds to the field-induced shift differences between the α-hydrogens of right and left helices of L-amino acids, respectively. The calculations are straightforward and the results are shown in Figure 7. The electric fields at the midpoint of the $C_\alpha H$ bonds were calculated; the projection along the bond was determined and substituted into eq 1b to give σ_E. This procedure was arbitrary. Had we used the field at the hydrogen nucleus instead, σ_E would change by $\lesssim 20\%$. The origin at 0 ppm in Figure 7 corresponds to the position of the abstracted $C_\alpha<^{H}_{H}$ group in its valence state in the peptide. (For our purposes the term "abstracted" means that (1) the field at the α-hydrogens coming from the monopole charge $+0.67 \times 10^{-10}$ esu on the C_α 1.1 A away is ignored. (2) The field at one of the hydrogens coming from the -0.37×10^{-10} esu charge on the other is also ignored. (3) The chemical shift of the $C_\alpha<^{H}_{H}$ hydrogens is set equal to zero.) The results shown in Figure 7 and Table III are not intended as "predictions" of numerical values to be meticulously compared with experiment. Rather they are intended to provide a framework for subsequent experiment in an area relatively devoid of any "theoretical" guidance. Many factors, undoubtedly, will cause experimental disagreement with the numerical values, *e.g.*, distortions from the model Pauling α helix (Phillips, 1966), inadequacies of the approximations, and neglected contributing mechanisms to the shift such as solvent effects. The charges on the planar amide can polarize the surrounding media, solvent, and other amide groups, which in turn can produce a "reaction field" that may affect the $C_\alpha H$ shifts. Reaction field effects have been identified as being significant in the case of CH_3CN in various solvents (Buckingham *et al.*, 1960) and in some aliphatic solvents (Hruska *et al.*, 1963) of high dielectric constants. In some polar solvents reaction field effects appear to be negligible (Sears and Hahne, 1966). The Onsager (1936) treatment of the reaction field which has met with some success is not applicable to protein systems because of the many oversimplifications inherent in the Onsager derivation (Laszlo and Musher, 1964; Kirkwood, 1939). The situation is exceedingly complicated in protein systems. We have ignored the reaction field effects and hope that their contributions, as well as those from the other factors above, will be sufficiently small or cancelling so that the general conclusions reached below will remain correct.

The electric field shifts are seen (Figure 7) to depend on the size of the helix and the position of the α-hydrogens. The up α-hydrogens shift downfield on the average relative to the origin, whereas the down hydrogens shift very little on the average. The numbering is such that in the case of an eight-peptide helix, for example, -1 refers to $C_\alpha H$ hydrogens at the beginning or bottom of the helix (NH_2 end) and 7 to the α-hydrogens at the end or top (carboxyl end).

The planar peptides are numbered as shown. The peptides joined to $C_\alpha(-1)$ from below and $C_\alpha(7)$ from above are not in the α-helical orientation for which the dihedral angles ϕ and ψ = 127.5 and 128.2°, respectively (Leach *et al.*, 1966b). Standard angle conventions are used. End-effect contributions from peptides below $C_\alpha(-1)$ and above $C_\alpha(7)$ have been neglected in Figure 7. In Table III we compile the contribution of near- and far-neighbor peptides to the fields at the 0 and 6 α-hydrogens. The fields at the α-hydrogens at intermediate positions on the helix can be readily determined from this table. A breakdown of contributions according to monopoles is given in Table IV. One sees that the $n + 1$ peptide group causes the n up and down α-hydrogens to shift 0.87 and 0.38 ppm, respectively, downfield relative to the origin. The monopole at the C′ of the $n + 1$ peptide makes the

TABLE IV: The Contributions from the Monopoles of Peptide m to the Electric Field Shifts[a] of αH_u ($n = 0$) and αH_d ($n = 0$).

	$m = 0$				$m = 1$						
	O	C′	N	H	O	C′	N	H	αC	αH_u	αH_d
αH_u	0.11	−0.29	**0.72**	−0.34	0.1	**−0.98**	0.39	−0.34	−0.11	0.03	0.03
αH_d	0.48	−0.51	0.72	−0.10	0.48	−0.98	0.17	−0.03	−0.05	0	0

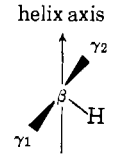
[a] Only the peptides which make the dominant contributions to the chemical shifts of αH are tabulated.

dominant contribution to this shift. The n-peptide group causes the n up and down α-hydrogens to shift 0.2 and 0.58 ppm upfield, respectively. The monopole on the nitrogen of the n peptide makes the dominant contribution. Contributions from next nearest neighbors are also significant. In general, peptides below an α-hydrogen tend to shift the α-hydrogen resonance upfield, while those above tend to shift its resonance downfield. In the case of the down hydrogen, however, the opposing shifts of the peptides above and below approximately cancel while in the case of the up hydrogen the contributions from the peptides above dominate.

Some effort was made to see how termination of the helix affects the shifts. Random chains consisting of three and two peptides were grafted on to the terminal 7 and -1 C_α's, respectively. The random sections in the literature sequence terminology (CD1, CD2, and CD3 (top), and CD7 and CD8 (bottom)) used were chosen from Kendrew's X-ray data of myoglobin (Liquori, 1966; Dickerson, 1964). The results are similar to Table III. The termination corrections are small, and do not affect our conclusions. In general, one expects similar effects from other random terminations.

The shift positions for an infinite right (R)- and left (L)-handed helices are shown in Figure 7. The (R) helix is predicted to shift downfield by ~0.5 ppm relative to the (L) helix. (The latter is only shifted slightly relative to the abstracted $C_\alpha<^{H}_{H}$ group taken as origin.) We further predict (Figure 7) that in the case of small (R) helices, some of the α-hydrogens near the beginning (NH end) of the helix may be shifted as much as 0.4 ppm further downfield than the α-hydrogens of the infinite helix.

Electric field effects at protons other than α-hydrogens are also possible. A valine residue (n) in a right-handed helix, for example, is in a "locked" configuration resulting from the valine β-H being confined to a pocket formed by the N–H of peptides n and $n + 1$. (The locked configuration can be schematically represented as



Choose an axis system x', y', z' such that (I) x' coincides with $C_\alpha - C_\beta$; (II) z' lies in the plane defined by the unit vectors along the C_α–H and C_α–C_β bonds. (z' makes an angle of 19.5° with C_α–H and is perpendicular to C_α–C_β.) In the initial configuration, $\theta = 0°$, C_β–H is "*cis*" to, and eclipses C_α–H. The locked configuration corresponds to $\theta \sim 200°$. (The positive sense is a clockwise rotation looking down the C_B–C_α bond toward C_α.) This approximate θ value holds for Pauling α helices. Helices in enzymes appear distorted (6) and the actual θ's will vary to some degree.) We

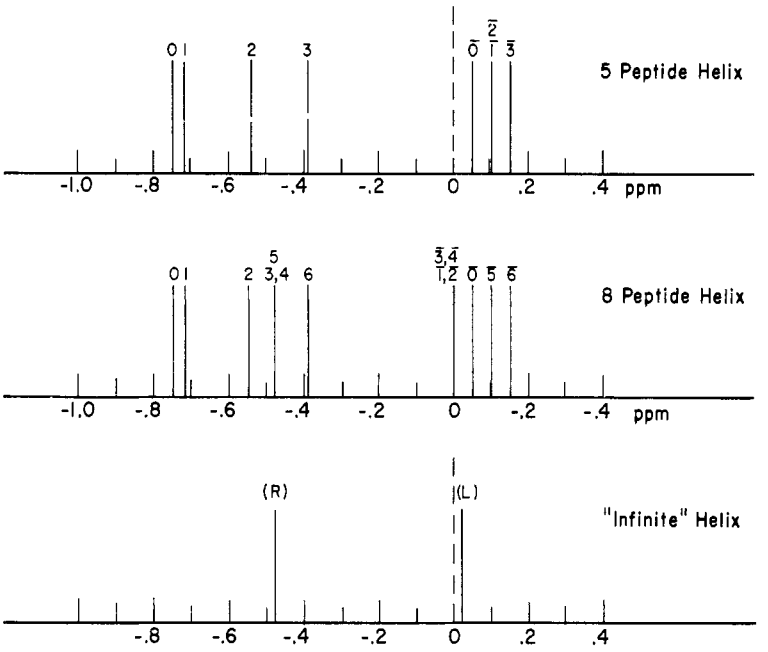


FIGURE 8: Electric field effects on the shifts of the α-hydrogens in polyglycine in a right-handed α-helix configuration. The down proton results are barred. The shifts are relative to an abstracted $C_\alpha<\frac{H}{H}$. We approximate the "infinite" helix by calculating the shift at αH(8) of a 16-peptide helix.

calculate a shift of 0.3 ppm upfield and 0.2 ppm downfield from the peptide monopoles for valines 0 and 6 of an eight-peptide right-handed helix, respectively, and (2) a 0.25-ppm shift downfield for an infinite right-handed helix.

J. A. Ferretti and A. M. Liquori (private communication) recently examined poly-(γ-L-benzylglutamate), poly-(β-L-benzylaspartate), and poly-(β-L-methylaspartate) in (1) $CDCl_3$ (helix-supporting solvent); (2) TFA, a nonsupporting solvent; and (3) mixtures of $CDCl_3$ and TFA. They found that random–helix (R) transitions, as represented by poly-(γ-L-benzylglutamate) in TFA and $CDCl_3$, respectively, are not accompanied by any appreciable change in α-hydrogen positions ($\tau \sim 5$ in TFA and $\tau \sim 5.0$ in $CDCl_3$), whereas the random–helix (L), as represented by poly-(β-L-benzylaspartate) and poly-(β-L-methylaspartate), is accompanied by a 0.6-ppm upfield shift ($\tau \sim 5.0$ in TFA and τ 5.6 in $CDCl_3$). No other protons in either the right- or left-handed helices show a shift change with solvent. In order to eliminate the possibility that TFA might shift resonances as a result of hydrogen bonding with the amide group, they prepared mixed TFA and $CDCl_3$ solvents and found a sharp helix (L)–random coil transition to occur at $\sim 3.5\%$ TFA for poly-(γ-L-benzylaspartate), poly-(β-L-benzylaspartate), and poly-(β-L-methylaspartate). The observed shifts of the random coil in the mixed solvent was the same as in pure TFA.

The observation of a 0.6-ppm upfield shift of the (L) helix relative to the (R) helix should be compared with our calculation of 0.5 ppm using the electric field perturbation. The absence of a shift in going from helix (R)–random would imply, as these authors have suggested, that the local environment in the random configuration is similar to the helix (R) although the long-range order obviously is not. This observation is supported by the fact that the majority of the residue in the nonhelical regions of sperm whale myoglobin (Liquori, 1966; Dickerson, 1964) have local conformations similar to a right-handed α helix. *If the Ferretti and Liquori conclusions can be extended to proteins in aqueous solutions, one expects that only a small, if at all detectable, chemical shift change would occur upon folding into regions of right-handed helices.* Caution must be exercised, however, since (1) experimentally only long helices were examined, whereas we predict that the chemical shifts depend on length. For the case of small (R) helices (Figure 7) some of the α-hydrogens (*i.e.*, those near the NH end) are predicted to be shifted ~ 0.3 ppm further downfield than the infinite helices; (2) disulfide linkages are common in proteins and persist even after unfolding. These may alter the local environment of neighboring residues causing them to have a conformation in the random state different from an α-right-handed helix; (3) although there may not be any solvent peptide-induced shift coming from possible hydrogen bonding of TFA to the amide group in the synthetic polypeptide study of Ferretti and Liquori there may be a solvent perturbation when a protein unfolds in water. Stewart *et al.* (1967) have recently done nuclear magnetic resonance studies of poly-L-

alanine and poly-L-leucine and the corresponding random DL forms in TFA and $CDCl_3$ and found very pronounced solvent-induced shifts as a result of TFA hydrogen bonding with the amide group. The situation is thus seen to be conflicting and requires more experimental work for clarification; (4) X-ray studies have shown that the α-helical regions in lysozyme are distorted from the classical Pauling helix. The CO group tends to point outward from the helix and the NH inward as opposed to being aligned along the helical axis. The extent of the distortion varies from slight to considerable. 3_{10} helices (three residues per turn, 2-A advance/residue) or approximations thereof, also occur in lysozyme. (In the case of the 3_{10} helix we have calculated essentially the same electric field effects as in the α helix.) The P_E resonance at τ 4.6 which appears upon folding may be an example of a helix induced shift for any of the reasons 1–4 above.

Acknowledgments

We wish to thank Professor Tinoco, Dr. W. Phillips, and Dr. J. Ferretti for a number of helpful discussions. Mr. Wheeler's assistance with the optical rotatory dispersion measurements is gratefully acknowledged. The authors are indebted to Dr. D. C. Phillips for making his X-ray coordinate data available to us.

References

Anet, F. A. L., and Bourn, A. (1965), *J. Am. Chem. Soc. 87*, 5250.

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C., and Sarma, V. R. (1967), *Proc. Roy. Soc. (London)* (in press).

Blake, C. C., Koenig, D. G., Mair, G. A., North, A. C., Phillips, D. C., and Sarma, V. R. (1965), *Nature 206*, 757.

Bovey, F. A., and Tiers, G. V. (1959), *J. Am. Chem. Soc. 81*, 2780.

Bovey, F. A., Tiers, G. V., and Filipovich, G. (1959), *J. Polymer Sci. 38*, 73.

Brant, D. A., and Flory, P. J. (1965), *J. Am. Chem. Soc. 87*, 2788, 2791.

Buckingham, A. D. (1960), *Can. J. Chem. 38*, 300.

Buckingham, A. D., Schaefer, T., and Schneider, W. G. (1960), *J. Chem. Phys. 32*, 1227.

Dickerson, R. E. (1964), *Proteins 2*, 634.

Didoy, J. R., and Gerry, F. (1961), *Compt. Rend. 253*, 422.

Emsley, J. W., Feeney, J., and Sutcliffe, L. H. (1965), High Resolution NMR Spectroscopy, Vol. 1, New York, N. Y., Pergamon, pp 85 ff.

Fraenkel, G., Carter, R. E., McLachlan, A. D., and Richards, J. H. (1960), *J. Am. Chem. Soc. 82*, 5846.

Gent, L. G. (1948), *Quart. Rev. (London) 2*, 383.

Griffith, O. H., and McConnell, H. M. (1966), *Proc. Natl. Acad. Sci. U. S. 55*, 8.

Hatton, J., and Richards, R. E. (1960), *Mol. Phys. 3*, 253.

Hatton, J., and Richards, R. E. (1962), *Mol. Phys. 5*, 137.

Hayashi, K., Imoto, T., Funatsu, G., and Funatsu, M. (1965), *J. Biochem. (Tokyo) 58*, 227.

Hruska, F., Bock, E., and Schaefer, T. (1963), *Can. J. Chem. 41*, 3034.

Johnson, C. E., Jr., and Bovey, F. A. (1958), *J. Chem. Phys. 29*, 1012.

Kirkwood, J. G. (1939), *J. Chem. Phys. 7*, 911.

Kowalsky, A. (1962), *J. Biol. Chem. 237*, 1807.

Kowalsky, A., and Cohn, M. (1964), *Ann. Rev. Biochem. 33*, 481.

Kurland, J., and Wilson, E. B. (1957), *J. Chem. Phys. 27*, 585.

Laszlo, P., and Musher, J. I. (1964), *J. Chem. Phys. 41*, 3906.

Leach, S. J., Némethy, G., Scheraga, H. A. (1966a), *Biopolymers 4*, 369.

Leach, S. J., Némethy, G., and Scheraga, H. A. (1966b), *Biopolymers 4*, 887.

Levinthal, C. (1966), *Sci. Am. 214*, 42.

Liquori, A. M. (1966), *J. Polymer Sci. 12C*, 209.

Mandel, M. (1965), *J. Biol. Chem. 240*, 1586.

McDonald, C. C., and Phillips, W. D. (1967), 2nd International Conference on Magnetic Resonance in Biological Systems, New York, N. Y., Pergamon (in press).

Musher, J. I. (1962), *J. Chem. Phys. 37*, 34.

Narasimham, P. T., and Rogers, M. T. (1959), *J. Phys. Chem. 63*, 1388.

Onsager, L. (1936), *J. Am. Chem. Soc. 58*, 1486.

Pauling, L. (1960), The Nature of the Chemical Bond, Ithaca, N. Y., Cornell University, p 496.

Phillips, D. C. (1966), *Sci. Am. 215*, 78.

Ramakishnan, C., and Ramachandran, G. W. (1965), *Biophys. J. 5*, 909.

Schachman, H. K. (1963), *Cold Spring Harbor Symp. Quant. Biol. 28*, 409.

Sears, R. E., and Hahn, E. (1966), *J. Chem. Phys. 45*, 2753.

Smythe, C. P. (1938), *J. Am. Chem. Soc. 60*, 183.

Smythe, C. P. (1955), Dielectric Behavior and Structure, New York, N. Y., McGraw-Hill, p 244.

Sophianopoulos, A. J., and Van Holde, D. E. (1964), *J. Biol. Chem. 239*, 2516.

Stewart, W. E., Mandelkern, L., and Glick, R. E. (1967), *Biochemistry 6*, 143.

Vanderkooi, B., Leach, S. J., Némethy, G., Scott, R. A., and Scheraga, H. A. (1966), *Biochemistry 5*, 2991.