Biochemistry

 Volume 19, Number 2 January 22, 1980

Comparisons of Ring-Current Shifts Calculated from the Crystal Structure of Egg White Lysozyme of Hen with the Proton Nuclear Magnetic Resonance Spectrum of Lysozyme in Solution†

Stephen J. Perkins* and Raymond A. Dwek

ABSTRACT: Detailed ring-current calculations with some consideration for protein flexibility are described for egg white lysozyme of hen. These are based on 15 sets of crystal coordinates of lysozyme and over 60 assigned resonances in the ^{1}H nuclear magnetic resonance (NMR) spectrum of lysozyme. These analyses evaluate two basic assumptions: that the refined protein crystal structure is similar to the protein structure in solution and that the present ring-current theories in application to the aromatic amino acids offer a good description of the observed shifts in the protein NMR spectrum. The dipolar, Johnson–Bovey, and Haigh–Mallion equations are tested. The conventional ring-current intensity factors for the tyrosine and phenylalanine rings of 0.94 and 1.00 are found to be satisfactory, but those for the tryptophan ring need to be increased. With this, ring-current effects offer good explanations for the chemical shifts of the aliphatic and aromatic protons in lysozyme for which the NMR signals have to date been resolved and assigned. Ring-current effects do not, however, explain directly the shifts of the six tryptophan N–H protons nor those of the α protons. Applications of these calculations are described to compare the tetragonal and triclinic forms of native lysozyme, to propose several assignments of ^{1}H NMR signals in the spectrum of lysozyme, and to analyze some of the conformational changes which occur on the binding of Gd(III) and GlcNAc sugars to lysozyme. The uses and limitations of the calculations for protein NMR are briefly discussed.

The chemical shifts of signals in the nuclear magnetic resonance (NMR) spectrum of a protein change on passing from the disordered "random-coil" state to the folded tertiary structure. The shift changes can be as large as several parts per million (ppm)[1] in the total range of about 10 ppm in protein ^{1}H NMR spectra. If the signal has been assigned to a proton of an amino acid residue in the protein, the shift change is easily measured by taking the difference between the observed shifts in the protein spectrum and the shifts of the random-coil amino acid in water, which are available in standard compilations (McDonald & Phillips, 1969; Roberts & Jardetsky, 1970; Gelan & Anteunis, 1972; Bundi & Wüthrich, 1979). It is usual to analyze the shift differences on the assumption that they arise from ring-current interactions of the proton with the local magnetic fields of the aromatic rings of histidine, phenylalanine, tyrosine, and tryptophan residues in the protein (Sternlicht & Wilson, 1967; Cowburn et al., 1970; Perkins et al., 1977a). This procedure utilizes the protein crystallographic coordinates together with one of the ring-current equations in common usage. The applications of these calculations have included those of protein NMR signal assignment, comparison of homologous proteins, the characterization of ligand binding sites on a protein, the conformational changes attendant on ligand binding, and various aspects of protein dynamics, besides the classical application of establishing the identity of the crystal and solution conformations of the protein (Sternlicht & Wilson, 1967; Cowburn et al., 1970; Campbell et al., 1975; Browne et al., 1976; Dower et al., 1977; Perkins et al., 1977a; Perkins & Wüthrich, 1979).

Many crystallographic and NMR studies of egg white lysozyme of hen (hereafter termed "lysozyme") have been reported, and these have been reviewed (Blake et al., 1978). The determination of the crystal structure of tetragonal lysozyme (Blake et al., 1965) and studies of the binding of *N*-acetylglucosamine (GlcNAc) and *N,N,N''*-triacetylchitotriose ($GlcNAc_3$) to lysozyme led to the identification of the three

† From the Laboratory of Molecular Biophysics, Zoology Department, Oxford OX1 3PS, United Kingdom (S.J.P.), and the Biochemistry Department, Oxford OX1 3QU, United Kingdom (R.A.D.). *Received January 18, 1979.* Supported in part by the Medical Research Council and the European Molecular Biology Organisation.

* Correspondence should be addressed to this author at the European Molecular Biology Laboratory, c/o C.E.N.G., L.M.A., 85X, 38041 Grenoble, France.

[1] Abbreviations used: DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; βMeGlcNAc, β-methyl-*N*-acetylglucosaminide; lysozyme, egg white lysozyme of hen (EC 3.2.1.17) (the lysozyme coordinates are abbreviated as listed in Table I); ^{1}H NMR, proton nuclear magnetic resonance; ppm, parts per million; rms, root mean square.

0006-2960/80/0419-0245$01.00/0

subsites A, B, and C in the active-site cleft of lysozyme. Further model-building studies suggested that there were three additional subsites D, E, and F (Blake et al., 1967a,b). Subsite D was subsequently confirmed by crystallography (Beddell et al., 1970; Ford et al., 1974). These studies thus led to the basic clues to the enzymatic activity of lysozyme (Johnson et al., 1968). Since that time, many other crystallographic studies of lysozyme have been reported, which have investigated the binding of metals, the triclinic form of crystalline lysozyme, chemical modifications, and protein denaturation. Currently the tetragonal and triclinic forms of native lysozyme are being extensively refined.

Many studies of the protein ^1H NMR spectrum of lysozyme have been described, and these generally confirm the findings from protein crystallography. The work of Williams and co-workers has been reviewed recently (Williams, 1976; Dobson, 1977; Blake et al., 1978), and other studies include those of Kaptein (1978), Lenkinski et al. (1978, 1979), and Chapman et al. (1978). Over 60 assignments have now been reported for protein NMR signals arising from more than 20 of the 129 amino acids of lysozyme. Other studies have explored the ^{13}C NMR spectrum of lysozyme and the ^1H, ^{13}C, and ^{19}F NMR of the binding of the GlcNAc sugars to lysozyme.

The ability to explain well the conformation-dependent chemical shifts in the NMR spectrum of a globular protein implies that the biological activity of the protein in solution may be appreciated in molecular terms. Ring-current interactions give rise to significant shift effects in lysozyme. As lysozyme is well-characterized by protein crystallography and protein NMR, it serves as a model enzyme to investigate ring-current calculations in detail. This study therefore includes the following: (1) the advances made in ring-current theories (Mallion, 1978) and their applications to aromatic amino acids (Giessner-Prettre & Pullman, 1969; Perkins et al., 1977b), using all three of the ring-current equations in common usage (Pople, 1956; Johnson & Bovey, 1958; Haigh & Mallion, 1971); (2) a thorough computerized analysis of ring currents, taking into its scope a detailed printout, the calculation of the mean square planes of the aromatic rings, a realistic representation of the methyl protons, and several considerations of protein flexibility; (3) the highly refined crystal structures of tetragonal (D. E. P. Grace & D. C. Phillips, unpublished results) and triclinic (L. H. Jensen, unpublished results) lysozyme; (4) the resolution and assignment of many resonances in the ^1H NMR spectrum of lysozyme [Blake et al. (1978), Dobson (1977), and references cited therein]. These calculations are used to establish the form of the equations for use with proteins and to compare the triclinic and tetragonal crystal forms of lysozyme with its solution structure. The applicability of the ring-current equations is tested for several groups of protons in lysozyme. Several assignments of ^1H NMR signals are reported, and some of the conformational changes on the binding of lanthanide and sugar inhibitors to lysozyme are discussed. Preliminary accounts are given in Perkins et al. (1977a) and in Dobson (1977).

Theory

The three ring-current equations considered in this study are the classical dipolar equation of Pople (1956), the semiclassical Johnson & Bovey (1958) equation, and the quantum-mechanical Haigh & Mallion (1971) equation. The explicit forms of these equations with symbols are defined in companion studies (Perkins & Wüthrich, 1979; Perkins, 1979). The ring-current shifts δ_R are expressed in parts per million, and all distances are given in nanometers. Note that negative

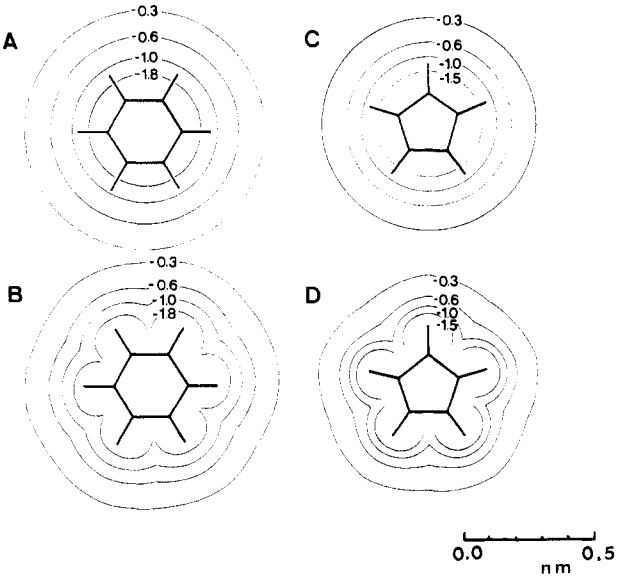

FIGURE 1: Contour diagrams to show the magnitude of the downfield shift in ppm in the plane of the aromatic ring by using the values for B which were used in this study and a unit value for i (see Theory). (A) and (C) show the cylindrical symmetry of the ring-current shift predicted by the Johnson–Bovey equation for the six- and five-membered rings, where there is no dependence of shift on the angle ϕ (eq 1). (B) and (D) show the hexagonal and pentagonal symmetry of the shift which is predicted from the ϕ dependence of the shift in the Haigh–Mallion equation.

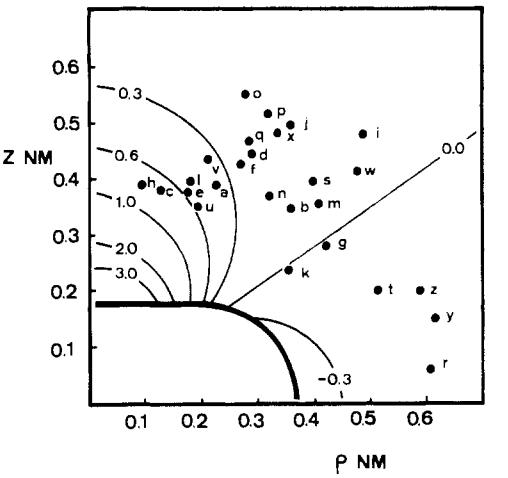

FIGURE 2: Contour diagram of the cross section of one-quarter of an aromatic ring to show the Johnson–Bovey ring-current shifts as a function of ρ and z. The heavy line indicates the van der Waals thickness of the aromatic ring. The filled circles indicate 26 methyl groups in lysozyme which are within 0.7 nm of the ring center of the six-membered rings of phenylalanine, tyrosine, and tryptophan. The letters a–z are identified in Table II.

numbers indicate downfield shifts which are experienced for protons lying in the ring plane (Figure 1). Figure 1 and Figure 2 thus summarize the variation of the ring-current shifts relative to the aromatic ring [see also Figure 4 of Perkins & Wüthrich (1979)].

In what is a strictly empirical application to protein NMR, the general form of the ring-current equation can be defined by the expression

$$\delta_R \times 10^{-6} = iBG(\rho,z,\phi) \qquad (1)$$

where i is a ring-current intensity factor which is specific for the aromatic ring type, B is a constant of proportionality, and $G(\rho,z,\phi)$ is a spatial term relating the cylindrical coordinates

ρ, z, and ϕ of the proton to each ring plane and center of the protein aromatic rings.

The geometrical term $G(\rho,z,\phi)$, which is different for each of the dipolar, Johnson–Bovey, and Haigh–Mallion equations (Figure 1), is defined explicitly in Perkins & Wüthrich (1979) and Perkins (1979). This takes on particularly large values in the short-range vicinity of each aromatic ring, which is exemplified in the case of the Johnson–Bovey equation by Figure 2. For comparison of the three ring-current equations in this region of space in this study, the three ring-current equations are normalized by the appropriate choice of the constant of proportionality B (below), such that the calculated shifts at large distances converge to similar values. At large distances, the individual forms of $G(\rho,z,\phi)$ for the Johnson–Bovey and Haigh–Mallion equations reduce to that of the dipolar equation (Haigh & Mallion, 1972; Mallion, 1978).

Previously, well-refined values for the constant B have been obtained for aromatic protons directly bonded to the rings in a variety of mono- and polycyclic molecules (Dailey, 1964; Haigh et al., 1970; Mallion, 1971) by means of regression analyses based on 66 experimental shifts. For the dipolar equation, where B is partly defined by the susceptibility χ_L (Perkins & Wüthrich, 1979), then $\chi_L = 29.6 \times 10^{-6}$ by this procedure. For the 1958 Johnson–Bovey equation, the calculated shifts are reduced by the factor of 0.597. For the Haigh–Mallion equation, the regression leads to the factor of 1.56 originally given (Haigh & Mallion, 1972). During the course of this study, it became clear that all the calculated shifts uniformly underestimated the observed shifts with these values of B. Hence, a value of $\chi_L = 49.5 \times 10^{-6}$ is used for the dipolar equation, which is based on Pauling's (1936) classical value for the susceptibility of the aromatic ring. The Johnson–Bovey equation is used in its 1958 form. The value of B of 1.56 in the Haigh–Mallion equation has to be scaled up by the factor of 2.6 to give 4.06 [see Haigh & Mallion (1972) and Mallion (1978)].

It is also necessary to allow for the specific nature of the aromatic ring in the protein by use of the ring-current intensity factor i, which is referenced to the above values of B by setting it to unity for phenylalanine. From the theoretical values of Giessner-Prettre & Pullman (1969), the values of i for phenylalanine, tyrosine, histidine, and the five- and six-membered rings of tryptophan are 1.00, 0.94, 0.53, 0.56, and 1.04 in that order. In the text below, and in another study (Perkins & Wüthrich, 1979), it is established that these values of i for tyrosine and phenylalanine rings are satisfactory. However, from this study and two others (Perkins, 1979; S. J. Perkins, unpublished experiments), it is clear that the two values of i for tryptophan require recalibration, and the description of this is given below.

Methods

A Fortran program RCCAL was written to compute the ring-current shifts for any proton position in the protein crystal structure by the three ring-current equations. Previous descriptions or accounts of its use have been given in Perkins (1977), Perkins et al. (1977a,b), Perkins & Wüthrich (1979), Perkins (1979), S. J. Perkins (unpublished experiments), and Perkins, Johnson, Phillips, and Dwek (unpublished experiments). The proton coordinates are computed from the protein crystal structure by using standard bond lengths and angles (Weast, 1971). Methyl protons are represented by 12 proton positions equally spaced in 30° steps starting from the eclipsed conformation with the penultimate-but-one bond; the calculated ring-current shifts could thus be weighted by a Boltzmann function corresponding to a 12.6 kJ mol^{-1} rotamer energy barrier at 27 °C. One provision in the program allowed for the movement of the aromatic ring plane about χ_2, the C^β–C^γ bond of the aromatic residue. This facility was used either to introduce local conformational changes of the aromatic ring in the protein coordinates or to represent the thermal vibrations of the aromatic ring in terms of a Boltzmann potential energy function corresponding to complete rotation of the aromatic ring about χ_2 in the protein molecule. Random-coil spectra of the protein (Bundi & Wüthrich, 1979) or these spectra corrected for the calculated ring-current shifts may be plotted also, together with an analysis of the predicted protein spectrum in terms of the subspectra of the individual amino acids (see Figure 5 below).

Results and Discussion

Source of Crystallographic and NMR Data. The protein crystal coordinates for lysozyme which are used in the calculation are listed in Table I, together with a description of the source of each set of coordinates. All X-ray diffraction data were collected at room temperature (~20 °C). The chemical shifts in the ^{1}H NMR spectrum of lysozyme in ^2H_2O are given in Tables II and III at ~54 °C and p^{2}H 4.7 (the p^{2}H is uncorrected for isotope effects), relative to acetone at 2.215 ppm downfield of sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at 0 ppm. Assignments to the amino acid type and to the individual residue in the NMR spectrum of lysozyme were taken from Dobson (1975, 1977) and Blake et al. (1978) and references cited therein.

The observed shifts for lysozyme are corrected for the chemical shifts of the amino acids in the random-coil state by reference to the most recent compilation by Bundi & Wüthrich (1979). Comparison of the set of about 80 of these shifts, other than those of the N–H protons and the α protons, shows rms differences of 0.11 ppm with the random-coil data of McDonald & Phillips (1969), 0.10 ppm with those of Gelan & Anteunis (1972), and 0.13 ppm with those of Roberts & Jardetsky (1970). Futher comparisons showed, for example, a rms difference of 0.08 ppm between the 1970 and 1972 data sets and 0.12 ppm between the 1969 and 1972 data sets. Limiting the comparisons to the shifts used for Tables II and III does not influence significantly these estimates of the reliability of the random-coil shifts.

Several assignments are based on calculated ring-current shifts, and these are identified in Tables II and III. For the shifts used in the regression analyses below, which are marked in Tables II and III, this affects signals from five residues, i.e., the choice between Met-105 and Met-12, the choice between Leu-17 and Leu-8 and the group of Leu-25, -56, -75, -83, -84, and -129, and finally the choice between Trp-28 and the group of Trp-62, -63, -108, -111, and -123. Below, it will be seen that the observed shifts from these assignments agree to within 0.3 ppm of the predicted shifts, or 1.5 times the rms difference of the regressions. It is also found that all large conformation-dependent shifts that are observed in the aromatic and aliphatic regions of the lysozyme spectrum readily receive explanation in terms of ring-current effects, with the exception of N–H protons and α protons (see below). Furthermore, on normalization of the slopes m of eq 2, the regression analyses are seen to converge. These three considerations show that these particular assignments are well-founded at the present time.

Observed Conformation-Dependent Shifts. In order to understand more fully the ^{1}H NMR spectrum of lysozyme, we performed several experiments in order (1) to study the effects of temperature, concentration, and ionic strength on the ^{1}H NMR spectrum, (2) to confirm the chemical shifts

Table I: Summary of the Lysozyme Coordinates Used in This Study and Some Related Sets

coordinate set	description
W2	tetragonal lysozyme coordinates from the hand-built wire model which was obtained from the electron density map calculated from the multiple isomorphous replacement phases (Blake et al., 1967a)
M2 (not used)	derived from the W2 set by the application of a mathematical model-building procedure to optimize the molecular stereochemistry of the amino acids to that seen in the corresponding small molecules [Diamond (1966); listed in Imoto et al. (1972)]
RS5D (also RS6A, RS9A, RS12A, and RS16 which were not used)	the RS5D set is the optimal result of a series of real-space refinements of the tetragonal lysozyme coordinates starting from the M2 set and seeking to find the best fit between the protein coordinates and the electron density map calculated from the multiple isomorphous phases; five sets were produced (Diamond, 1974); the R^a for the RS5D set is 0.35
tetragonal[b]	this is obtained from a crystallographic refinement of tetragonal lysozyme at 0.18-nm resolution, starting from the RS5D set and minimizing the differences between the observed and calculated structure factors (D. E. P. Grace and D. C. Phillips, unpublished results); the R is 0.22
lysozyme-sugar; lysozyme-Gd^{III}	the following coordinates were obtained by real-space refinement based on the RS5D set as starting point and the multiple isomorphous replacement phases; the crystallographic data to 0.2–0.25-nm resolution were obtained for lysozyme in acetate-free media at pH 5.1, lysozyme–GlcNAc, lysozyme–βMeGlcNAc, lysozyme–$GlcNAc_3$, and lysozyme–Gd^{III}–GlcNAc; all were obtained from tetragonal lysozyme in acetate buffer at pH 4.7 except for the first one and the third one which is the same as for the first one (Blake et al., 1967b; Perkins, 1977; Perkins et al., 1978, 1979)
triclinic[b]	this is obtained in the course of refinement of triclinic lysozyme at 0.15-nm resolution, where the R is 0.173 (L. H. Jensen, unpublished results)

[a] $R = \Sigma|F_{obsd} - F_{calcd}|/\Sigma F_{obsd}$ where F_{obsd} and F_{calcd} are the observed and calculated structure factors. [b] The unit cell volume is 29.6 nm^3/molecule in tetragonal lysozyme and 26.2 nm^3/molecule in triclinic lysozyme.

given in Tables II and III, and (3) to extend the known signal assignments to include those of the high field shifted signals of Ile-98 and Met-105.

The temperature dependence of the shifts in the NMR spectrum between 10 and 66 °C for a 5 mM solution of lysozyme at p^2H 5.3 showed that most chemical shifts were unchanged and that the ring current shifted signals decrease in shift within a maximum range of 0.05 ppm. Exceptions to this were the high field shifted methyl signals of Leu-17, which increased and decreased their shifts by 0.07 and 0.16 ppm, respectively, and three single-proton resonances upfield of −0.2 ppm from DSS, which decreased their conformation-dependent shifts by 0.25, 0.07, and 0.37 ppm between 10 and 66 °C. There was little evidence from a plot of chemical shift against temperature of a temperature transition in lysozyme at 25 °C, which has been proposed by Jardetsky and others [Cozzone et al. (1975), but see Dobson (1975) and Shindo et al. (1977)] to account for small shift changes seen in the ^{13}C NMR spectrum of lysozyme. The effects of varying the concentration of lysozyme between 2.5 and 20 mM and the ionic strength between 0 and 0.83 by using sodium chloride in a solution of 5 mM lysozyme, both at p^2H 5.0 and 54 °C, were negligible. Exceptions were the downfield-shift changes of about 0.3 ppm which were seen for the C(2)H and C(4)H protons of His-15 when the concentration or ionic strength was increased. This can be accounted for by the participation of His-15 in a dimerization equilibrium with Glu-35 (Shindo et al., 1977). Accordingly, the effects of solution conditions as appropriate for the crystallization of lysozyme in 0.7–0.9 M sodium chloride and the change of temperature were taken to be negligible for the present study.

Nuclear Overhauser experiments were made in order to obtain further information on the high field shifted single-proton signals in the ^1H spectrum of lysozyme (S. J. Perkins, unpublished results). The saturation of the Ile-98 γ-CH signal at −2.03 ppm from DSS identified another γ-CH signal at 0.72 ppm from DSS and the α-CH signal at 2.82 ppm from DSS, thus completely identifying the Ile-98 spin system. Decoupling experiments and saturation of the signal at −0.91 ppm from DSS led to its assignment to a β-CH proton of Met-105. The evidence included the observation of selective effects at positions corresponding to the shifts of the C(2,6)H and C(3,5)H signals of Tyr-23 and a third signal at 7.37 ppm from DSS and the location of the α-CH signal and two other coupled protons at 3.82 and 0.38 ppm from DSS, respectively. Calculation with the triclinic and tetragonal lysozyme proton coordinates showed that a β-CH proton of Met-105 experiences a very large upfield ring-current shift and that this proton is very close to the Tyr-23 aromatic protons and the Trp-108 and Trp-111 C(4)H protons (Figure 4C). Some of these results were also found by Chapman et al. (1978) independently and are consistent with the above results.

Dipolar, Johnson–Bovey, and Haigh–Mallion Equations. Initial ring-current calculations were based on the dipolar, Johnson–Bovey, and Haigh–Mallion equations, the Giessner-Prettre–Pullman (1969) ring-current intensity factors i [previously employed also for lysozyme in Perkins et al. (1977a)], and the refined tetragonal and triclinic coordinates. The 45 experimental shifts in Tables II and III which were used for the regression analyses were selected on the basis of criteria given in the footnote to Table II. The ability of the regressions in exploring the spatial term $G(\rho,z,\phi)$ of eq 1 is demonstrated in Figure 2. This shows an analysis in terms of ρ and z of the arrangements of 26 methyl group–aromatic ring interactions in the refined tetragonal coordinates, where a methyl group is within 0.7 nm of the ring center of any six-membered ring of tryptophan, phenylalanine, and tyrosine. The methyl groups are seen to be well distributed in the cylindrical solid space surrounding the aromatic rings.

Figure 3 demonstrates that with tetragonal lysozyme, the calculated and observed shifts show good linear relationships for all three ring-current equations for both the aliphatic and aromatic C–H protons of lysozyme. The intercepts are close to zero. This shows that ring-current effects provide the dominant explanation for the experimental shifts and that shift effects from carbonyl groups, α helices, and so on may be

Table II: Experimental and Calculated Chemical-Shift Changes between the Random-Coil NMR Spectrum of Lysozyme and the Observed NMR Spectrum of Lysozyme, Using Residues Which Contain Methyl Groups and Which Are Resolved by ^{1}H NMR

resonance assignment	evidence for assignment	chemical shift from DSS	shift change from the random-coil value[a]	shift calcd from the triclinic coordinates (Table I): Johnson–Bovey	shift calcd from the refined tetragonal coordinates (Table I)			methyl group (Figure 2)
					Johnson–Bovey[c]	Haigh–Mallion	dipole	
Leu-17: δ-CH_3	ring current	−0.63	1.53 or 1.57§	2.08	1.80 (1.56, Trp-28)	1.64	2.09	a, b
δ-CH_3		−0.10	1.04 or 1.00§	1.02	1.21 (0.99, Trp-28)	1.15	1.41	c
γ-CH		0.68	0.97§	0.79	0.77 (0.65, Trp-28)	0.70	0.91	
Ile-98: δ-CH_3	Glu-35 ionization	−0.01	0.90§	1.33	0.99 (1.10, Trp-63)	0.88	1.28	f, g
γ-CH		−2.03	3.22 or 3.51	5.00	4.33 (4.17, Trp-63; 0.21, Trp-108)	3.64	5.12	
γ-CH		0.72	0.76 or 0.47	1.17	2.24 (2.26, Trp-63)	1.98	2.59	
γ-CH_3		−0.25	1.19§	1.04	1.34 (0.88, Trp-108; 0.47, Trp-63)	1.28	1.48	d, e
β-CH		1.52	0.37§	0.36	0.49 (0.66, Trp-63; −0.11, Trp-108)	0.47	0.57	
α-CH		2.82	1.40	0.58	0.99 (0.90, Trp-63)	0.99	1.00	
Leu-8: δ-CH_3	ring current	−0.01	0.91 or 0.95§	1.03	0.89 (0.82, Phe-3)	0.82	0.93	h, i
δ-CH_3		0.57	0.37 or 0.33§	0.10	0.22 (0.12, Phe-3)	0.27	0.21	j, k
γ-CH		1.47	0.18§	0.27	0.29 (0.18, Phe-3)	0.29	0.28	
Met-105: ε-CH_3	ring current	0.00	2.13§	1.88	2.21 (0.90, Trp-28; 0.80, Trp-111; 0.30, Trp-108; 0.15, Tyr-23)	2.10	2.63	l, m, n, o
γ-CH		n.a.[d]		1.72	1.67 (0.73, Trp-111; 0.68, Tyr-23; 0.58, Trp-28; −0.37, Trp-108)	1.53	1.85	
γ-CH		n.a.		3.31	4.02 (2.87, Tyr-23; 0.87, Trp-111; 0.43, Trp-28; −0.16, Trp-108)	3.40	4.91	
β-CH		−0.91	2.91 or 3.07	4.52	3.01 (2.67, Trp-111; 0.36, Trp-28; 0.28, Tyr-23; −0.28, Trp-108)	2.69	3.48	
β-CH		n.a.		3.09	3.43 (2.92, Trp-111; 0.39, Tyr-23; 0.28, Trp-28; −0.17, Trp-108)	3.07	3.98	
α-CH		3.82	0.69	0.53	0.20 (0.48, Trp-111; 0.25, Tyr-23; 0.15, Trp-28; −0.73, Trp-108)	0.12	0.21	
Ile-88: δ-CH_3	His-15 ionization	0.27	0.62§	0.32	0.29 (0.16, Phe-3)	0.29	0.32	p
Leu-56: δ-CH_3	Ln(III), ring current	0.28	0.62 or 0.66§	0.61	0.60 (0.72, Trp-108; −0.26, Trp-28)	0.52	0.77	q, r
δ-CH_3		0.53	0.41 or 0.37§	0.16	0.12 (0.15, Trp-108; −0.21, Trp-28)	0.13	0.14	s, t
γ-CH		1.22	0.43§	0.14	0.13 (0.13, Trp-108; −0.12, Trp-28)	0.12	0.17	
Thr-51: γ-CH_3	Ln(III)	0.32	0.91§	0.87	0.61 (0.65, Tyr-53)	0.56	0.61	u
β-CH		3.76	0.46§	0.09	−0.21 (−0.19, Tyr-53)	−0.10	−0.20	
Val-92: γ-CH_3	His-15 ionization	0.48	0.46 or 0.49§	0.37	0.37 (0.16, Trp-28)	0.35	0.43	
γ-CH_3		0.60	0.37 or 0.34§	0.29	0.39 (0.25, Trp-28; 0.11, Tyr-20)	0.38	0.43	
β-CH		1.92	0.21§	0.26	0.33 (0.14, His-15; 0.11, Trp-28)	0.31	0.40	
Ala-107: β-CH_3	Ln(III)	0.64	0.76§	1.01	0.76 (0.68, Trp-108)	0.73	0.87	v, w, x
α-CH		3.87	0.48	0.34	0.22 (0.23, Trp-108)	0.21	0.26	
Val-109: γ-CH_3	Ln(III)	1.03	−0.09 or −0.06§	−0.05	−0.09	−0.09	−0.11	
γ-CH_3		1.14	−0.18 or −0.21§	−0.12	−0.17	−0.16	−0.21	
β-CH		2.19	0.06§	−0.08	−0.13	−0.12	−0.16	
Ala-110: β-CH_3	Ln(III)	1.38	0.02§	−0.13	−0.15	−0.14	−0.17	y
α-CH		4.31	0.04	−0.02	−0.04	−0.04	−0.06	
Met-12: ε-CH_3	ring current	1.66	0.47§	0.12	0.10 (0.16, Trp-108; −0.20, Trp-28)	0.11	0.11	z

[a] The observed shifts marked § here and in Table III are the 45 shifts used in the regression analyses. Excluded from these are all tentative assignments, single protons with large ring-current shifts, α-protons, and N–H protons for reasons stated in the text. In this column, two shifts are given where a 1:1 correspondence between the random-coil and observed signal is not established, and their mean is used in the regression analyses. [b] Ala-31 is a less likely assignment. [c] Here and in Table III, ring-current shifts larger than |0.10| ppm are shown in parentheses. [d] n.a., not assignable further with certainty. Another coupled proton (probably the β-CH proton from its spatial proximity to the first β-CH proton) resonates at 0.40 ppm and two other coupled protons very probably resonate at 0.55 and 0.08 ppm from DSS.

Table III: Experimental and Calculated Chemical-Shift Changes between the Random-Coil NMR Spectrum of Lysozyme and the Observed NMR Spectrum of Lysozyme, Using Residues Which Contain Aromatic Rings

resonance assignment	evidence for assignment	chemical shift from DSS	shift change from the random-coil value	shift calcd from the triclinic coordinates (Table I): Johnson–Bovey	shift calcd from the refined tetragonal coordinates (Table I)		
					Johnson–Bovey	Haigh–Mallion	dipole
Tyr-20: C(2,6)H	by elimination	6.98	−0.12§	−0.25	−0.12	−0.11	−0.13
C(3,5)H		7.24	−0.09§	−0.09	−0.04	−0.04	−0.04
Tyr-23: C(2,6)H	chem. modif.	6.71	0.10§	0.15	0.14 (0.13, Trp-111)	0.15	0.17
C(3,5)H		7.05	0.15§	0.27	0.15 (0.19, Trp-111)	0.21	0.15
Tyr-53: C(2,6)H	Ln(III)	6.83	0.03§	−0.09	−0.06	−0.07	−0.10
C(3,5)H		7.09	0.06§	−0.04	−0.03	−0.03	−0.04
His-15: C(2)H	His-15 ioniz-	8.34	0.05§	0.03	0.03	0.03	0.04
C(4)H	ation	7.23	0.03§	0.10	0.11	0.10	0.12
Trp-28: C(2)H	by elimination	7.30	−0.06§	−0.01	0.02	0.01	0.04
N(1)H	and ring current	9.36	0.86	−0.12	−0.09 (−0.12, Tyr-23)	−0.10	−0.07
C(4)H		6.76	0.89§	0.73	0.74 (0.51, Trp-108)	0.70	0.86
C(5)H	ring current	6.28	0.89§	0.64	0.99 (0.66, Trp-108; 0.13, Trp-111)	1.02	1.01
C(6)H		6.85	0.39§	0.04	0.05 (0.17, Trp-111; −0.24, Trp-108)	0.07	0.08
C(7)H		7.77	−0.27§	−0.19	0.00 (0.16, Trp-111; −0.16, Tyr-20)	0.00	0.04
Trp-62: β-CH[a]	CIDNP[b]	1.89	1.31 or 1.43	4.34	4.29 (4.22, Trp-63)	3.32	5.87
β-CH[a]	CIDNP	1.89	1.43 or 1.31	1.06	1.86 (1.81, Trp-63)	1.61	2.26
C(2)H	chem.	7.03	0.21§	0.03	0.03	0.09	0.08
N(1)H	modif.	10.00	0.22	0.08	0.04	0.04	0.05
C(4)H	CIDNP	7.20	0.64§	1.03	0.46 (0.43, Trp-63)	0.42	0.55
C(5)H		n.o.[c]		0.31	0.15 (0.15, Trp-63)	0.14	0.18
C(6)H	CIDNP	7.05	0.19§	0.13	0.07	0.07	0.09
C(7)H		n.o.		0.12	0.08	0.08	0.10
Trp-63: β-CH[a]	CIDNP	3.53	−0.33 or −0.21	−0.15	−0.12 (−0.10, Trp-62)	−0.12	−0.14
β-CH[a]	CIDNP	3.42	−0.10 or −0.22	−0.19	−0.15 (−0.13, Trp-62)	−0.15	−0.18
C(2)H	chem. modif.,	7.63	−0.39§	−0.26	−0.25 (0.10, Trp-108; −0.30, Trp-62)	−0.25	−0.32
N(1)H	spin label, and GlcNAc	10.21	0.01	0.29	0.05 (0.21, Trp-108; −0.15, Trp-62)	0.07	0.07
C(4)H	CIDNP	7.77	−0.12	−0.13	−0.09	−0.10	−0.11
C(5)H		7.08	0.09	−0.10	−0.03	−0.04	−0.03
C(6)H	CIDNP	7.46	−0.22	−0.02	0.12	0.11	0.16
C(7)H		n.o.		0.26	0.37 (0.24, Trp-62; 0.16, Trp-108)	0.35	0.47
Trp-108: C(2)H	Ln(III), pH, and	7.08	0.16§	0.01	−0.02	−0.02	−0.02
N(1)H	chem. modif.	9.98	0.24	0.06	0.00 (0.11, Trp-63)	−0.01	0.01
C(4)H		7.38 or 7.19	0.27 or 0.46	0.50	0.56 (0.25, Trp-111; 0.14, Tyr-23; 0.10, Trp-28)	0.56	0.64
C(5)H	ring current	6.49 or 6.49	0.68 or 0.68	0.39	0.33 (0.23, Trp-111; 0.15, Tyr-23; 0.13, Trp-63; −0.16, Trp-28)	0.34	0.38

	C(6)H		7.19 or 7.38	0.05 or 0.14	0.16	0.01 (0.24, Trp-63; 0.10, Trp-111; −0.37, Trp-28)	−0.03	0.04
Trp-111:	C(7)H		n.o.		0.18	0.11 (0.25, Trp-63; −0.19, Trp-28)	0.08	0.16
	C(2)H	chem.	7.03	0.21[δ]	0.32	0.32 (0.22, Trp-28; 0.10, Trp-123)	0.30	0.41
	N(1)H	modif. and ring current	10.35	−0.13	0.21	0.20 (0.28, Trp-28)	0.17	0.27
	C(4)H		n.o.		0.00	−0.01 (0.12, Trp-28; −0.16, Trp-108)	−0.01	−0.02
	C(5)H		n.o.		−0.03	−0.05	−0.05	−0.05
	C(6)H		n.o.		−0.07	−0.13 (0.11, Trp-28; −0.17, Tyr-23)	−0.15	−0.12
	C(7)H		n.o.		−0.04	−0.09 (0.18, Trp-28; −0.21, Tyr-23)	−0.11	−0.05
Trp-123:	C(2)H	N–H exchange and spin label	7.55	−0.31[δ]	−0.10	−0.10	−0.10	−0.10
	N(1)H		10.68	−0.46	−0.11	−0.15 (−0.14, Phe-34)	−0.15	−0.14
	C(4)H		n.o.		−0.18	−0.15 (−0.15, Phe-38)	−0.15	−0.16
	C(5)H		n.o.		−0.01	−0.03 (0.10, Phe-34; −0.11, Phe-38)	−0.04	−0.04
	C(6)H		n.o.		0.08	0.06	0.10	0.06
	C(7)H		n.o.		−0.26	−0.33 (−0.34, Phe-34)	−0.32	−0.32

[a] The calculated shifts for these protons make no allowance for relative conformational changes between the β protons and the indole ring of the same residue on passing from the random-coil state to the folded protein. [b] CIDNP, chemically induced dynamic nuclear polarization. [c] n.o., not observable.

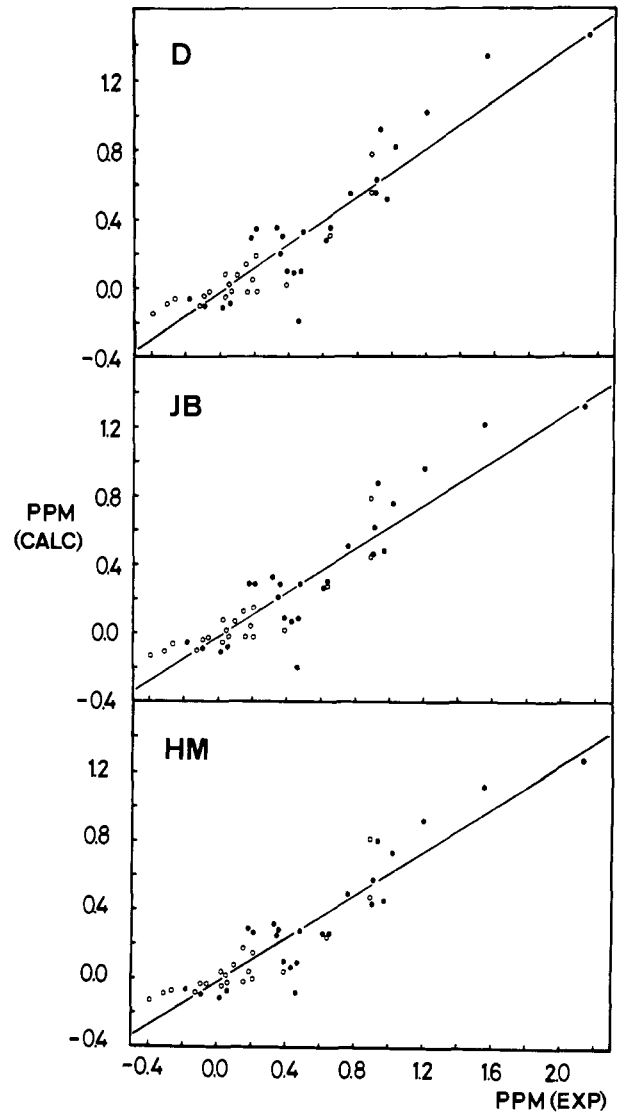

FIGURE 3: Plots of 45 experimental shifts from the NMR spectrum of lysozyme for the methyl and coupled protons (Table II) (●) and the aromatic protons (Table III) (○) vs. the shifts calculated on the basis of the dipolar equation (D), the Johnson–Bovey equation (JB), and the Haigh–Mallion equation (HM) with the refined tetragonal coordinates of lysozyme. The lines represent the least-squares regression line based on 45 data points for protons in Tables II and III. The calculated shifts correspond to the initial ring-current calculations.

neglected. Standard regression analyses were performed to give equations of the form

$$\delta_{\text{calculated}} = m\delta_{\text{observed}} + c \qquad (2)$$

For example, for the Johnson–Bovey equation

$$\delta_{\text{calculated}} = 0.663\delta_{\text{observed}} - 0.025 \qquad (3)$$

The corresponding results for the three equations and both refined coordinate sets are shown in Table IV, together with the rms differences in parts per million and the standard regression coefficients r^2.

Table IV shows that the slopes m of these regressions are not equal to unity, where the calculated shifts substantially underestimate the observed shifts. Experience gained in the course of this study, such as with Leu-8 and Thr-51 (Table II), shows that the ring-current shifts calculated for tyrosine and phenylalanine rings with the values of B and i defined under Theory are satisfactory. Supporting evidence for this comes from a study based on the pancreatic trypsin inhibitor (Perkins & Wüthrich, 1979), where tyrosines and phenyl-

Table IV: Results of the Regression Analyses Based on 45 Observed and Calculated Shifts Which Are Selected on the Basis of Criteria Given in Footnote *a* of Table II

regression analysis		slope m	intercept c	rms difference[a]	regression coefficient r^2
	Tetragonal Coordinates				
dipolar	initial	0.715	−0.025	0.204	0.857
	final	1.145	−0.053	0.193	0.870
Johnson–Bovey	initial	0.663	−0.025	0.214	0.845
	final	0.989	−0.046	0.190	0.874
Haigh–Mallion	initial	0.628	−0.012	0.210	0.851
	final	0.928	−0.035	0.183	0.882
	Triclinic Coordinates				
dipolar	initial	0.725	−0.028	0.240	0.813
	final	1.150	−0.036	0.213	0.847
Johnson–Bovey	initial	0.674	−0.025	0.231	0.824
	final	0.991	−0.033	0.201	0.861
Haigh–Mallion	initial	0.644	−0.024	0.207	0.855
	final	0.930	−0.029	0.183	0.882

[a] The rms differences are the normalized differences between the calculated and the observed shifts in ppm.

alanines are the only aromatic rings. Tables II and III, however, show that the majority of the ring-current shifts arise from tryptophan rings. This is the justification for recalculations which were made by using the Johnson–Bovey equation and the tetragonal and triclinic coordinates in which the ring-current intensity factors of $i = 0.56$ and 1.04 for tryptophan were increased. The values of i for the two rings were varied in such a way that the resulting regression slopes were always very close to unity. The best agreement between observed and calculated shifts is found with $i = 1.7$ and 1.2 for the five- and six-membered rings, respectively. The magnitudes of the calculated shifts were found to be more sensitive to the six-membered ring current factor, but the rms difference was more sensitive to that of the five-membered ring. This showed that both factors require empirical readjustment, but it should be noted that the values of $i = 1.7$ and 1.2 corresponded to a broad minimum of values for the rms differences and regression coefficients. Comparison of the initial and final regression lines in Table IV shows that the slopes m for each equation have become very similar in the tetragonal and triclinic coordinates and that the rms difference has improved by up to 15%. These convergences, while small, justify further the premise that the ring-current factors for tryptophan are underestimates of those corresponding to the experimental observations. The significance of these improvements is emphasized from comparisons of the ring-current shifts calculated from the crystal structures of cytochrome *c* and bovine trypsin with the observed NMR shifts (Perkins, 1979; S. J. Perkins, unpublished experiments). The factors of 1.7 and 1.2 lead to much more satisfactory agreements between experiment and calculation than the Giessner-Prettre–Pullman (1969) factors.

The value of $i = 1.7$ for the five-membered ring by this analysis may appear to be physically unreasonable by comparison with the value for the six-membered ring. It should, however, be noted that a ring radius of 0.1182 nm was employed in the calculations for the five-membered ring, which is smaller than the value of 0.139 nm used with the six-membered ring [Table I of Perkins et al. (1977a)]. The factor of 1.7 would be substantially less (~30%) if the ring radius for the five-membered ring had been taken as 0.139 nm, which was previously employed by Sternlicht & Wilson (1967). For this reason, it is emphasized that the ring-current intensity factors i here are empirically derived on the basis of experimental observation as one way to improve the agreement between experiment and the Johnson–Bovey calculation, having ruled out other sources which contribute to the conformation-dependent shift, quod vide.

Table IV shows that the overall agreement between observation and calculation is slightly improved in the case of the tetragonal coordinates. For each coordinate set, comparisons between the dipolar, Johnson–Bovey, and Haigh–Mallion equations show that there are minor improvements in that order. Overall, within the error of the analyses, the three ring-current equations lead to essentially similar results for a protein containing tyrosine, tryptophan, and phenylalanine residues. Below, the Johnson–Bovey equation is used for reasons of its popularity and its cost effectiveness (where it is simpler and cheaper than the Haigh–Mallion equation) and to avoid the point-dipole approximation [which is shown to be important in the case of cytochrome *c* (Perkins, 1979)]. The calculated shifts given in Tables II and III correspond to these final values of i, and the following text is based on these calibrated shifts.

The final rms differences of 0.18–0.21 ppm in Table IV can be explained in terms of contributions from three sources. (1) The shifts of the random-coil data sets used as the reference shifts are one limiting factor of the analyses, where rms differences of 0.10–0.13 ppm are seen on comparing different compilations of these shifts (see above). The protein chemical shifts are measured to a precision of 0.05 ppm (see above). (2) Limitations exist in the ring-current theories themselves. Regression analyses based on 66 nonhindered aromatic protons in a series of polycyclic aromatic compounds (but based on only one reference shift, that of the benzene protons, rather than about 80 reference shifts for the amino acid protons) lead to rms deviations of 0.078 ppm for the Johnson–Bovey equation (Mallion, 1971) and 0.085 ppm for the Haigh–Mallion equation (Haigh et al., 1970). (3) Differences in the atomic positions in the crystallographic coordinates increase the rms differences (see below). Through-space secondary shift mechanisms therefore do not appear to play a systematic and significant role in determining the observed shifts of the aliphatic and aromatic C–H proton signals in lysozyme. These secondary shifts might have arisen, for example, by van der Waals effects, electric field and diamagnetic susceptibility effects of the peptide carbonyl groups and charged groups, and so on (Jackman & Sternhell, 1969), but they are known to be short ranged in their effects.

Wire-Model, Real Space Refined, and Calculated-Phase Coordinates of Lysozyme. Regression analyses are now presented for the wire-model tetragonal coordinates of lysozyme W2 (Blake et al., 1967a) and the real space refined tetragonal coordinates of RS5D (Diamond, 1974). These corresponded respectively to an initial set and a preliminary refined set of protein coordinates (Table I) obtained in the course of a crystallographic study. Deviations for individual aromatic atoms from the mean square plane of their own ring ranged up to 0.016, 0.003, and 0.002 nm for the W2, RS5D, and refined sets, respectively. The crystallographic R factors are given in Table I. The regression analyses with the Johnson–Bovey equation for the W2 set lead to

$$\delta_{\text{calculated}} = 1.041\delta_{\text{observed}} - 0.077 \qquad (4)$$

$$\text{rms difference} = 0.247 \text{ ppm}$$

$$r^2 = 0.804$$

and for the RS5D set lead to

$$\delta_{calculated} = 1.067\delta_{observed} - 0.046 \quad (5)$$

rms difference = 0.201 ppm

$$r^2 = 0.861$$

Comparison with the corresponding values of $m = 0.989$ and $c = -0.046$ in Table IV for the phase-refined tetragonal coordinates shows that the intercepts and rms differences decrease and the r^2 values increase with the increased quality of the crystallographic coordinates. This is to be expected if the crystal and solution structures of lysozyme are similar.

Closer inspection of the individual predictions shows that qualitatively similar shifts are calculated with the wire-model, real space refined, and phase-refined sets of coordinates. For the 19 methyl groups of Table II, Figure 2 shows that none approach the regions of very large shift, and therefore the calculated shifts are not sensitive to small variations in atomic positions. Further, it is expected that undue variations in the predicted shifts are reduced by the use of the Boltzmann averaging of 12 rotamer positions for the methyl protons. Accordingly, it was found that the rescaled calculated shifts for the largest shifted methyl protons, such as those of Leu-17, Ile-98, and Leu-8 (Table II), are variable within a range of 0.5 ppm. For the remaining 42 methyl groups of lysozyme, intense NMR signals which were observed in the ^{1}H NMR spectrum of lysozyme between 1.6 and 0.7 ppm from DSS correspond well with the small shift changes from the random-coil values predicted from the calculations. The signals between 1.6 and 1.2 ppm from DSS will correspond primarily to methyl groups from the group of 12 alanine and 17 threonine residues, and those between 1.3 and 0.7 ppm from DSS likewise will correspond to the group of six isoleucine, eight leucine, and six valine residues.

The calculated shifts for single protons where they are close to an aromatic ring are more sensitive to the coordinates used. For example, in the three tetragonal sets, that of the β-CH proton of Met-105 varies in a range of 1.9 ppm. Likewise, the γ-CH proton of Leu-56 is downfield shifted, upfield shifted, and almost unshifted in the three sets of coordinates. When a single proton is in hydrophobic contact with an aromatic ring, small changes in position easily lead to substantial shift changes, which is readily appreciated on inspection of Figure 2 for the regions where ρ is close to zero. With the three tetragonal and the triclinic coordinates, the high field shifted one-proton signals upfield of 0 ppm from DSS include a number of single-proton lines from Leu-17, Lys-33, Arg-61, Trp-62, Ile-98, Met-105, and Lys-116 (Figure 5). The highest field line of all is consistently predicted to be the γ-CH proton of Ile-98, and the next one is predicted to be the β-CH proton of Met-105, except in the refined tetragonal coordinates where it takes third place after a γ-CH proton of Arg-61 (Figure 5). With the Giessner-Prettre–Pullman (1969) ring-current factors, the second most high field shifted proton would have been a Met-105 γ-CH proton which has been shifted by its proximity to Tyr-23 (Figure 4C), but this clearly conflicts with the spectroscopic evidence given above. The agreement between the rescaled calculated shifts and the observed shifts is, however, qualitative. The Ile-98 and Met-105 signals are observed at −2.03 and −0.91 ppm from DSS, respectively, while the refined triclinic coordinates lead to predicted shift positions at −3.81 and −2.35 ppm from DSS, thus overestimating the upfield shift by well over 1 ppm. The uncertainity in the protein coordinates in regions of high ring current shift gradient thus has a pronounced effect on the outcome of the calculations, and accordingly these shifts were not included in the regression analyses.

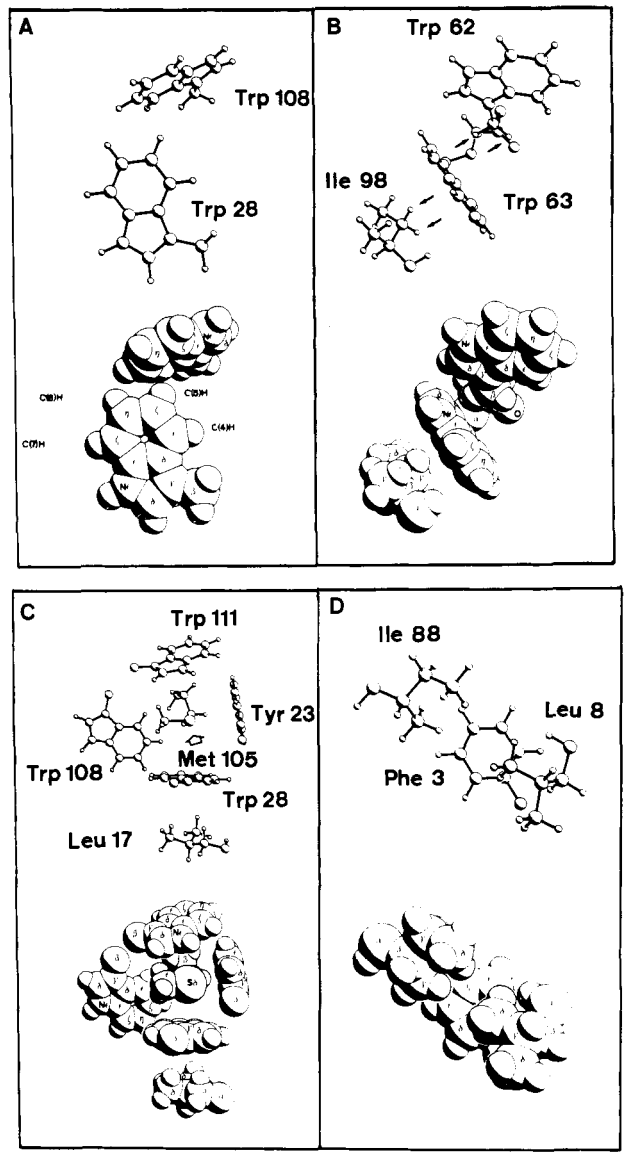

FIGURE 4: Computer drawings of local structures in the refined tetragonal crystal coordinates of lysozyme to show the most important ring-current interactions in the protein. The upper figures are skeletal diagrams of the atomic coordinates while the lower figures are drawn by using atomic radii slightly less than the van der Waals radii of the individual atoms. Oxygen, nitrogen, and sulfur atoms are denoted by capital letters, and carbon atoms are denoted by Greek letters where visible. (A) The ring-current interaction between the ring of Trp-108 and the four benzenoid protons of Trp-28. (B) The interaction between the ring of Trp-63 and the methylene protons of Ile-98 and Trp-62 (arrowed). As Trp-63 moves on the binding of the GlcNAc sugar to lysozyme in subsite C, the chemical shifts of these four protons will act as sensitive probes at the molecular level of the protein enzymatic activity. (C) The interactions in the aromatic box, where Met-105 is surrounded by the rings of Tyr-23, Trp-28, Trp-108, and Trp-111. (D) The interaction between Phe-3, Leu-8, and Leu-88, where the ring of Phe-3 is depicted beneath the side chains of Leu-8 and Ile-88.

C–H Aromatic Protons of Lysozyme. Previously, it has been shown that the relatively small shifts calculated for the six tyrosine and six tryptophan C(2)H signals are in accord with experimental observations (Perkins et al., 1977a). Subsequently, the individual assignments for the tyrosine signals have been reported (Dobson et al., 1978). Table III shows that the relative sequence in the observed spectrum of lysozyme of each of the three C(2,6)H and C(3,5)H proton signals of Tyr-20, Tyr-23, and Tyr-53 is correctly predicted by all three ring-current equations. Likewise, in the light of the individual assignments of the Trp C(2)H signals (Cassels et al., 1978),

their relative positions in the NMR spectrum are correctly predicted for all the six C(2)H protons in tetragonal lysozyme and triclinic lysozyme (Table III). The trivial exceptions are Trp-28 and Trp-108 in tetragonal lysozyme where the shift difference is insignificant at 0.04 ppm. The individual assignments for the pair of C(2)H signals belonging to Trp-28 and Trp-111 are presently tentative (Cassels et al., 1978). The additional evidence from the recalibrated ring-current shifts now calculated with the refined lysozyme coordinates shows that these assignments are good ones on the ground that the shift difference is larger than the rms differences of the above regressions.

The C(2)H shift of Trp-62 is underestimated by about 0.2 ppm. While this is at the limit of the rms differences of the regression analyses, it is interesting that a rotation of Trp-62 toward Trp-63 by 30° leads to an additional upfield shift of 0.16 ppm on the calculated shift. This is principally the result of the ring-current field of Trp-63 acting on the Trp-62 C(2)H proton (Figure 4B). It improves the agreement between observed and calculated shifts. Both Trp-62 and Trp-63 are on the surface of the protein. There is thus some evidence for a minor conformational difference between the solution and crystal states of lysozyme.

Four of the six aromatic four-spin systems of tryptophan have been wholly or partly identified to date. Three signals of one four-spin system are identified (Table III) by a combination of evidence from spin-decoupling (Dobson, 1975) and chemically induced dynamic nuclear polarization (Kaptein, 1978) experiments. This spin system was tentatively assigned to Trp-63 (Kaptein, 1978) and is consistent with the ring-current calculations which predict that the Trp C(4)H proton is one of the two downfield-shifted C(4)H protons together with that of Trp-123. The C(4)H and C(6)H proton signals of a second four-spin system have been assigned to Trp-62 from similar experiments by Kaptein (1978). This assignment is supported by the large shifts which are seen experimentally and are predicted (Figure 4B and Table III) for the methylene protons of Trp-62. The large upfield shift of the C(4)H proton of Trp-62 is in accord with what is predicted in Table III.

At this point, it is concluded that all the resolved and assigned aromatic C–H signals in the spectrum of lysozyme are consistent with the ring-current calculations. This detailed justification is required because the aromatic shifts in the pancreatic trypsin inhibitor are not explained by ring-current shifts, with nonsystematic deviations of as much as 0.4–0.6 ppm (Wagner & Wüthrich, 1975; Snyder et al., 1975; S. J. Perkins, unpublished results). The present calculations can therefore be used to propose assignments. A third four-spin system of tryptophan has been fully identified, and this was tentatively assigned to Trp-63 from a qualitative consideration of ring-current effects (Campbell et al., 1975; Dobson, 1975). Subsequently, the assignment was revised to that of Trp-28 on the basis of a prediction using the RS5D coordinates (Perkins et al., 1977a). This reassignment is supported by the use of the refined coordinates here. Not only is the relative order of the four observed shifts rationalized by this assignment to within 0.3 ppm of their calculated shifts but also the predicted upfield shift of the C(5)H triplet signal (Figure 4A) is evidently the largest of all the C(5)H or C(6)H tryptophan protons in lysozyme (Table III). This is in complete agreement with the experimentally observed spectrum of lysozyme where this proton resonates at 6.28 ppm from DSS. [Note that the C(7)H signal of Trp-28 is chemical-shift coincident with the C(4)H signal of Trp-63 assigned by Kaptein (1978), appearing in a two-proton doublet at 7.77 ppm from DSS. The statement by Kaptein (1978) that the above assignment of the Trp-63 C(4)H signal conflicts with the assignment of the Trp-28 C(7)H signal is therefore not fully justified at present.]

A fourth spin system of a tryptophan with an upfield-shifted triplet signal at 6.49 ppm from DSS has been partly identified by spin-decoupling experiments (Dobson, 1975). Both sets of refined coordinates predict (Table III) that only Trp-108 of the remaining four tryptophans will have an upfield-shifted triplet signal, from the C(5)H proton. This suggests a tentative assignment to Trp-108.

Shifts Not Explained Directly by Ring-Current Effects. Five of the aromatic N–H protons of the six indole rings of tryptophan in lysozyme were first observed and identified as such by Glickson et al. (1969, 1971). The sixth N–H proton has been recently identified and the assignments for all six have been updated (Cassels et al., 1978). After this paper was submitted for publication, Glickson has published revisions of his five original assignments (Lenkinski et al., 1979), and these new assignments agree with those updated by Williams and also the choice made above between Trp-28 and Trp-111. Comparison of the observed and calculated shifts (Table III) shows that the N–H shift of Trp-28 differs by 1 ppm and that while there are moderate differences of 0.2 to 0.3 ppm for the N–H protons of Trp-62, -108, -111, and -123, the relative positions of the calculated shifts are quite unrelated to the observed order in the NMR spectrum. Ring-current effects do not directly explain the N–H proton shifts, and this has been ascribed to differences in the molecular environments of the tryptophan N–H protons in the folded protein (Perkins et al., 1977a).

Ring-current effects can offer an indirect explanation of the indole N–H chemical shifts. It is clear from consideration of Figures 1 and 2 that one major contribution to the shifts will be from the ring-current field of its own ring. Variations in the N–H bond length will change the ring-current shift experienced by the indole N–H proton. By use of a correction factor of 0.597 for the constant B in eq 1 in the case where protons are directly bonded to the aromatic ring (Theory), model calculations for an indole ring with the rescaled values of i for tryptophan show that the observed N–H shifts can be explained by repositioning the N–H proton relative to its location in the so-called random-coil state in water. The N–H protons of Trp-111 and Trp-123 would have to be about 0.02 nm closer to the indole ring. Similarily, the N–H protons of Trp-62 and Trp-108 would have to be about 0.02 nm further away from the indole ring, while the N–H proton of Trp-28 would have to be further away by about 0.1 nm. Thus, quite subtle changes in atomic positions are sufficient in order to explain the observed N–H shifts for Trp-62, Trp-108, Trp-111, and Trp-123 by ring-current effects.

Inspection of the individual environments of the indole N–H protons now offers an explanation for the shifts of the N–H protons of Trp-28, Trp-108, and Trp-111. The solvent exposures of the N–H protons of Trp-62, Trp-63, and Trp-123 are 50–80% while those of Trp-28, Trp-108, and Trp-111 are completely buried with solvent exposures of 0% (Shrake & Rupley, 1973). A survey of the atomic coordinates for electronegative atoms which are close enough to the three buried indole N–H protons to act as hydrogen-bond acceptors, using both the refined tetragonal and triclinic coordinates, shows that the nearest ones to Trp-28 are the main-chain oxygen atoms of Tyr-23 (0.25 and 0.21 nm from the N–H proton in the tetragonal and triclinic coordinates, respectively) and Leu-17 (0.29 and 0.27 nm, respectively). Similarily, the nearest one to Trp-108 is the main-chain oxygen atom of

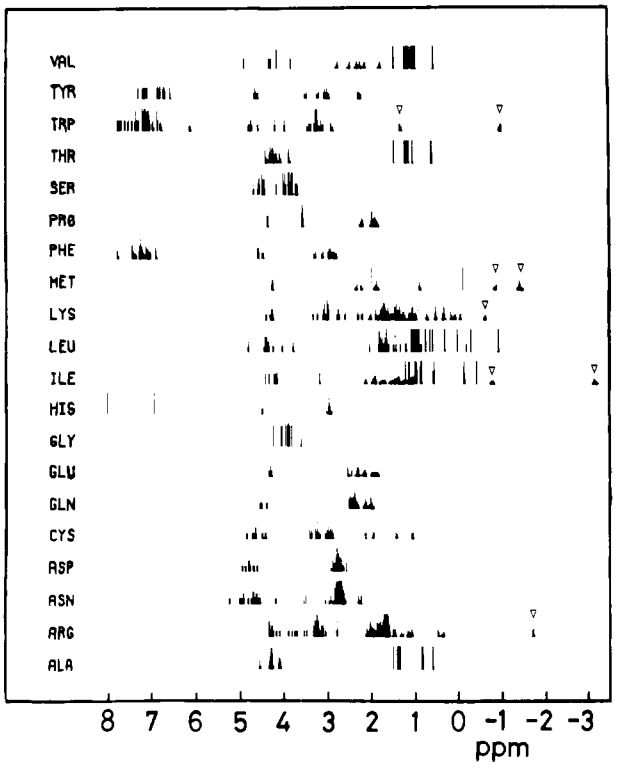

FIGURE 5: Schematic computer simulation of the ^{1}H NMR spectrum of lysozyme in terms of the subspectra of the 20 amino acid types present in lysozyme, using the refined tetragonal coordinates of lysozyme and the Johnson–Bovey equation defined under Theory. Large upfield-shifted single-proton signals for Lys-33, Trp-62, Met-105, Ile-98, and Arg-61 are denoted by inverted triangles.

Leu-56 (0.27 and 0.21 nm, respectively), while the nearest one to Trp-111 is the side-chain oxygen or nitrogen atom of Asn-27 (0.18 and 0.18 nm, respectively). All other possible proton acceptors are more than 0.35 nm away. These distances therefore suggest the hypothesis that the 1-ppm upfield shift of the N–H signal of Trp-28 is the consequence of two hydrogen-bond acceptors situated relatively far away from this N–H proton. Likewise, the 0.3-ppm upfield shift of the Trp-108 N–H proton is the consequence of one acceptor positioned relatively far away, and the 0.3-ppm downfield shift of the Trp-111 N–H proton is the result of the closely positioned Asn-27 side chain. However, in general, we note that there are several other possible sources of chemical-shift variation for hydrogen-bonded N–H protons [see Jackman & Sternhell (1969)], so this is by no means the only explanation for the lysozyme N–H indole shifts.

A second class of shifted resonances which are not explained by ring-current effects are the α-proton resonances. Four of these are seen downfield of the water signal at shifts between 5.8 and 5.4 ppm from DSS, while Figure 5 shows that no resonances from any of the amino acid residues occur in this spectral region through ring-current effects. Comparison of the α-proton shifts in the random-coil tetrapeptides with those in the free amino acids shows rms differences of 0.67 and 0.62 ppm, respectively, with the 1970 and 1972 data sets. This shows that the shift of the α proton is sensitive to its immediate chemical environment; it will be affected in particular by the orientations of adjacent main-chain carbonyl groups, from which the induced shifts become significant at distances less than 0.4 nm (Sternlicht & Wilson, 1967; Jackman & Sternhell, 1969).

Another instance of shifts which are seen as a result of the protein tertiary structure and which are not explained by ring-current effects is those following ionization. Large through-space shifts of this type of 0.2 to 0.3 ppm are seen only on the C(2)H and N(1)H protons of Trp-108 when the active-site Glu-35 carboxyl group becomes ionized at a pH of about 6. From the NMR evidence, a specific interaction between Glu-35 and Trp-108 has been proposed which is characteristic of the family of lysozymes (Dobson, 1977).

Protein Flexibility. Ring current shifted signals may be affected by two mechanisms based on considerations of protein flexibility. The protein can coexist in several different conformations of similar energy content. Alternatively or additionally, the protein has a large degree of thermal vibrational motion about the time-averaged structure of the crystal conformation.

The effect of the first mechanism can be evaluated by a comparison of each pair of the individual predictions from the refined triclinic and tetragonal coordinates. Differences will be the consequence of conformational changes caused by the different lattice contacts in the two crystal forms. Above, it has been shown from the regression analyses that the two structures are generally similar. However, it is known that the segment of the β-pleated sheet between residues 70 and 74 to the left of the active-site cleft at the protein surface (when the protein is viewed conventionally) lies in different positions in tetragonal and triclinic lysozymes (L. H. Jensen, private communication). The ring-current predictions in this study are unaffected because no appropriately situated protons are in the close vicinity of residues 70–74. There are, however, noticeable differences which are greater than 0.3 ppm in the individual predictions for a δ-CH_3 group of Leu-17 (0.6 ppm), the C(4)H proton of Trp-62 (0.4 ppm), and the δ-CH_3 proton of Ile-98 (0.5 ppm) in triclinic lysozyme and for the β-CH proton of Thr-51 (0.6 ppm) in tetragonal lysozyme. That these differences are seen with different protons in the two coordinate sets shows that their source lies in strictly localized differences in the protein coordinates themselves. That for Trp-62 and Ile-98 in triclinic lysozyme is ascribable to the positions of these residues relative to Trp-63 (see Figure 4B). For example, the distance between the centers of the six-membered rings of Trp-62 and Trp-63 is 0.68 nm in triclinic lysozyme and 0.75 nm in tetragonal lysozyme. In the case of Thr-51 (Table II), the cylindrical coordinates ρ,z of the β-CH and the mean γ-CH_3 protons of Thr-51 relative to those of Tyr-53 are, respectively, 0.27,0.22 and 0.11,0.36 in triclinic lysozyme and 0.32,0.17 and 0.19,0.35 in tetragonal lysozyme. The difference between the coordinate sets in the atomic positions of Thr-51 relative to those of Tyr-53 is thus only of the order of 0.1 nm. This is sufficient to cause a 0.6-ppm difference between the observed and calculated shifts for the β-CH proton of Thr-51 in tetragonal lysozyme but not in triclinic lysozyme. For Leu-17, the difference is likewise small between the two coordinate sets and now favors tetragonal lysozyme on the basis of the observed NMR shift. The cylindrical coordinates ρ,z for the mean δ-CH_3 and γ-CH protons of Leu-17 relative to those of the six-membered ring of Trp-28 (Figure 4C) are 0.08,0.38, 0.28,0.38, and 0.15,0.56 for triclinic lysozyme and 0.13,0.38, 0.23,0.39, and 0.13,0.57 for tetragonal lysozyme. This shows that the relative differences between the coordinates here are of the order of 0.05 nm. Thus, while there are cases of large shift differences between the two forms, one of the two coordinate sets always gives differences between observation and calculation of less than 0.3 ppm. It is not likely that these differences in relative differences are significant.

The second mechanism in which protein flexibility may influence ring current shift interactions is through the existence

of large thermal motions in the atomic positions of the protein side chains. This can be detected crystallographically and by NMR. A plot of the crystallographic temperature factors for individual residues against the squares of their distance from the molecular centroid shows that the protein surface is more mobile than the protein core in refined tetragonal lysozyme (Blake et al., 1978; Sternberg et al., 1979). Comparison of the temperature coefficients of chemical shift of the ring-current signals of lysozyme supports this conclusion. The ϵ-CH_3 and β-CH signals of Met-105, which is buried in a hydrophobic box in the protein core (Figure 4C), at 0.00 and −0.91 ppm from DSS change by 0.01 and 0.07 ppm from 10 to 66 °C, while the γ-CH signal of Ile-98, which is near the protein surface, at −2.03 ppm from DSS shifts by 0.37 ppm over the same temperature interval.

Several relatively simple ring-current models may be examined in order to account for these shift dependences with temperature. An expansion of the whole protein, which can either include the aromatic rings or not, leads to decreased ring-current shifts because of the increase in the ring radius in the Johnson–Bovey equation and the increase in ρ,z. The absence of a temperature coefficient for the shifts of the Met-105 signals implies that this model is inadequate; in fact, ring-current shifts are generally considered to be temperature independent (Dwek, 1973). A second model presumes that the source of the temperature dependence lies in thermal vibrations of the aromatic rings about χ_2, the aromatic C^β–C^γ bond. Such "thermal" ring-current shifts were calculated on the basis of rotating the plane of the aromatic ring in five steps at 5° intervals about χ_2 on each side of the aromatic ring which is seen in the refined tetragonal crystal structure. The ring-current shifts were calculated for each of the eleven plane positions, and these shifts were weighted by a Boltzmann function to correspond to the population of the aromatic ring at each value of χ_2 [see Perkins & Wüthrich (1979)]. The Boltzmann curve as a function of χ_2 is modeled on the basis of energy functions calculated for the aromatic rings in the pancreatic trypsin inhibitor, in particular, that for Phe-45 (Gelin & Karplus, 1975; Hetzel et al., 1976). Increase of temperature can thus be represented by redistribution of the Boltzmann populations and by a loosening of the protein structure (i.e., the energy barrier for 180 or 360° aromatic ring rotations about χ_2 has been decreased). Calculations were based on the temperature difference between 27 and 77 °C for energy curves with barriers to rotation of 30–130 kJ mol^{-1}. Both factors lead to very similar relative shift changes. Considerably larger effects are found from the calculations based on the loosening of the protein structure than from those based on the temperature difference itself. Only these larger shifts are comparable in magnitude with the observable NMR shift changes with temperature such as that for Ile-98. Many of the shifts are steady to within 0.02 ppm or are reduced with increased temperature by amounts ranging up to 0.2 ppm for the C(5)H proton of Trp-28. Interestingly, there are rare occasions where the shift increases with temperature. This corresponds to the dominance of the r^{-3} distance term [where $r = (\rho^2 + z^2)^{1/2}$] in the "thermal" calculation and not the angular term if analogy is made with a dipolar ring-current equation for this comparison. In conclusion, the shift changes with temperature seen by NMR can therefore be assigned to regions of the protein which have become looser with rise of temperature, and this receives support from the model calculations just outlined. However, the use of ring-current calculations to study positions and fluctuations in proteins is still only in the most preliminary stages.

Comparison of the Calculations for Lysozyme, Lysozyme–Sugar, and Lysozyme–Lanthanide Coordinates. Protein coordinates for crystallographic data for six lysozyme, lysozyme–GlcNAc sugar, and lysozyme–lanthanide complexes listed in Table I have been obtained by real-space refinement methods, using the same methods and parameterizations that were used in the calculation of the tetragonal RS5D native coordinates (Perkins, 1977; Perkins et al., 1978, 1979). Because these coordinates are preliminary in nature, only some of the main results with the ring-current calculations are described.

Acetate-free lysozyme has been found to be almost identical in conformation with the native enzyme (Perkins et al., 1978). A comparison of the predicted shifts for the native RS5D coordinates with those from the acetate-free lysozyme coordinates acts as a control in the following calculations. The rescaled shifts for the Trp-28 C(5)H triplet proton using the RS5D and acetate-free coordinates are 0.90 and 0.96 ppm, respectively. In the lysozyme–GlcNAc, lysozyme–βMeGlcNAc, and lysozyme–$GlcNAc_3$ coordinates, the shifts are 1.32, 1.14, and 1.07 ppm, respectively. The difference is in good agreement with the observed upfield shift of this triplet signal at 6.28 ppm from DSS by up to 0.2 ppm on the addition of GlcNAc and $GlcNAc_3$ to lysozyme (Dobson & Williams, 1975). The triplet signal at 6.49 ppm from DSS which is probably that of the Trp-108 C(5)H proton shows no shift change in the presence of sugar, and this is in agreement with the rescaled calculated shifts for this proton which are steady to within 0.05 ppm. While Trp-108 has moved closer to Trp-28 by 0.03 nm, a shift change is only seen for Trp-28 and not for Trp-108. The computer analysis shows that this results from the location of the upfield-shifted Trp-108 proton close to the isoshielding line of zero ring-current shift of Trp-28.

In the lysozyme–Gd^{III} and lysozyme–Gd^{III}–GlcNAc coordinates (Perkins et al., 1979), a conformational change is seen in which the ring of Tyr-53 has moved by 5.4 and 6.4° relative to its position in the native RS5D coordinates. It might therefore be expected that a shift change would be seen for the nearby Thr-51 protons which are upfield shifted by Tyr 53 (Table II). However, the predicted shifts are steady to within 0.04 ppm, while the observed downfield-shift change of the Thr-51 methyl signal in the presence of diamagnetic lanthanide is 0.04–0.06 ppm (Campbell et al., 1975). On closer inspection of the coordinates, the cylindrical coordinates ρ,z of the mean γ-CH_3 proton coordinates of Thr-51 relative to Tyr-53 are found to be 0.19,0.31, 0.15,0.33, and 0.16,0.33. The conformational change thus corresponds to a 0.04-nm movement of the Thr-51 methyl group along one of the isoshielding lines of chemical shift of Tyr-53 and, hence, rationalizes the lack of observable shift change in the Thr-51 signal.

Application and Limitation of Ring-Current Calculations in Protein NMR. In this study, the crystallographic structures have been shown to be consistent with the NMR data for aliphatic and aromatic C–H signals with rms differences of 0.18–0.21 ppm between 45 observed and calculated shifts from a total of over 60 assigned signals. Among the 45 shifts, there are few individual differences greater than 0.3 ppm. Those greater than 0.3 ppm are fully accounted for by differences in atomic positions between the two crystal forms, and these are unlikely to be of significance with respect to the solution state. The study of ring current shifted aliphatic and aromatic protons by NMR can therefore be quantitatively seen in terms of proton–aromatic ring interactions for at least the case of lysozyme, but not in the cases of hydrogen-bonded N–H

protons or α protons, or when the shifts are particularly large. In several instances, exemplified by the indole N–H protons or the single-proton shifts of Ile-98 and Met-105, qualitative interpretations are possible. Because no extrinsic probes are added to the NMR sample, there are no constraints on the interpretations, i.e., the biological significance of observed shift changes in terms of molecular rearrangements. For example, the otherwise highly useful lanthanide paramagnetic reagents are known to inhibit lysozyme and prevent the binding of βMeGlcNAc in one of the six subsites of lysozyme (Perkins, Johnson, Phillips, and Dwek, unpublished experiments).

One chief limitation is that only localized structures within the protein may be investigated. Of the 61 methyl groups and 13 aromatic rings in lysozyme, which leads to a total of 793 possible interactions, only 14 methyl signals are sizably affected by large ring-current shifts. High-quality structural information is thus less readily available from the shifts of the remaining 47 methyl signals. Additionally, it is possible that a conformational change is observed crystallographically and no corresponding shift change is detected for a favorably shifted proton signal. Examples of this were cited in the preceding section. The general difference between crystallography and NMR spectroscopy is that the information from the former is vectorial and global, while that from the latter probes mainly pairwise interactions.

The study of ring current shifted proton signals by NMR is more reproducible by its nature than an equivalent study by X-ray crystallography and is also a more sensitive means of establishing the similarity or dissimilarity of the microenvironments of certain amino acids in the protein under a range of experimental conditions. The precision inherent in each atomic position determined by protein crystallography is of the order of several hundredths of a nanometer. Because it is the relative distance between the aromatic ring(s) and the proton which is of relevance, these errors are of the order of 0.05–0.1 nm, which was shown by the comparison of localized structures in triclinic and tetragonal lysozymes above. This leads to calculated shift differences of as much as 0.6 ppm from the experimental shifts, which can be appreciated from the rule of thumb that the shifts are approximately proportional to the inverse cube of the distance. However, by protein NMR, these large conformation-dependent chemical shifts can be reproducibly measured within precisions of 0.01–0.05 ppm, precisions which are 1 order of magnitude higher than that from protein crystallography.

Acknowledgments

Professor L. H. Jensen is thanked for generously making the coordinates of triclinic lysozyme available, D. E. P. Grace and Sir David Phillips are thanked for those of tetragonal lysozyme, and Dr. C. M. Dobson and Professor R. J. P. Williams are thanked for the tyrosine and tryptophan aromatic assignments. We thank Sir David Phillips and Dr. L. N. Johnson for useful discussions and criticisms.

References

Beddell, C. R., Moult, J., & Phillips, D. C. (1970) *Mol. Prop. Drug Recept. Ciba Found. Symp., 1970*, 85–112.

Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1967a) *Proc. R. Soc. London, Ser. B 167*, 365–377.

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1967b) *Proc. R. Soc. London, Ser. B 167*, 378–388.

Blake, C. C. F., Grace, D. E. P., Johnson, L. N., Perkins, S. J., Phillips, D. C., Cassels, R., Dobson, C. M., Poulsen, F. M., & Williams, R. J. P. (1978) *Mol. Interact. Act. Proteins, Ciba Found. Symp., 1978*, 137–185.

Browne, C. A., Campbell, I. D., Kiener, P. A., Phillips, D. C., Waley, S. G., & Wilson, I. A. (1976) *J. Mol. Biol. 100*, 319–343.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285–297.

Campbell, I. D., Dobson, C. M., & Williams, R. J. P. (1975) *Proc. R. Soc. London, Ser. A 345*, 41–59.

Cassels, R., Dobson, C. M., Poulsen, F. M., & Williams, R. J. P. (1978) *Eur. J. Biochem. 92*, 81–97.

Chapman, G. E., Abercrombie, B. D., Cary, P. D., & Bradbury, E. M. (1978) *J. Magn. Reson. 31*, 459–469.

Cowburn, D. A., Bradbury, E. M., Crane-Robinson, C., & Gratzer, W. B. (1970) *Eur. J. Biochem. 14*, 83–93.

Cozzone, P. J., Opella, S. J., Jardetsky, O., Bethou, J., & Jolles, P. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2095–2098.

Dailey, B. P. (1964) *J. Chem. Phys. 41*, 2304–2310.

Diamond, R. (1966) *Acta Crystallogr. 21*, 253–266.

Diamond, R. (1974) *J. Mol. Biol. 82*, 371–391.

Dobson, C. M. (1975) D. Phil. Thesis, University of Oxford.

Dobson, C. M. (1977) in *NMR in Biology* (Dwek, R. A., Campbell, I. D., Richards, R. E., & Williams, R. J. P., Eds.) Academic Press, London.

Dobson, C. M., & Williams, R. J. P. (1975) *FEBS Lett. 56*, 362–365.

Dobson, C. M., Ferguson, S. J., Poulsen, F. M., & Williams, R. J. P. (1978) *Eur. J. Biochem. 92*, 99–103.

Dower, S. K., Wain-Hobson, S., Gettins, P., Givol, D., Jackson, W. R. C., Perkins, S. J., Sunderland, C. A., Sutton, B. J., Wright, C. E., & Dwek, R. A. (1977) *Biochem. J. 165*, 207–223.

Dwek, R. A. (1973) *NMR in Biochemistry*, Oxford University Press, Oxford.

Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C., & Tjian, R. (1974) *J. Mol. Biol. 88*, 349–371.

Gelan, J., & Anteunis, M. (1972) *^{1}H 300 MHz NMR Spectra of the Natural Amino Acids*, Rijksuniversiteit Gent, Belgium.

Gelin, B. R., & Karplus, M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2002–2006.

Giessner-Prettre, C., & Pullman, B. (1969) *C. R. Hebd. Seances Acad. Sci., Ser. D 268*, 1115–1117.

Glickson, J. D., McDonald, C. C., & Phillips, W. D. (1969) *Biochem. Biophys. Res. Commun. 35*, 492–498.

Glickson, J. D., Phillips, W. D., & Rupley, J. A. (1971) *J. Am. Chem. Soc. 93*, 4031–4038.

Haigh, C. W., & Mallion, R. B. (1971) *Mol. Phys. 22*, 955–970.

Haigh, C. W., & Mallion, R. B. (1972) *Org. Magn. Reson. 4*, 203–228.

Haigh, C. W., Mallion, R. B., & Armour, E. A. G. (1970) *Mol. Phys. 18*, 751–766.

Hetzel, R., Wüthrich, K., Deisenhofer, J., & Huber, R. (1976) *Biophys. Struct. Mech. 2*, 159–180.

Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., & Rupley, J. A. (1972) *Enzymes, 3rd Ed. 7*, 665–868.

Jackman, L. M., & Sternhell, S. (1969) *Applications of NMR Spectroscopy in Organic Chemistry*, Pergamon Press, Oxford.

Johnson, C. E., & Bovey, F. A. (1958) *J. Chem. Phys. 29*, 1012–1014.

Johnson, L. N., Phillips, D. C., & Rupley, J. A. (1968) *Brookhaven Symp. Biol. 21*, 120–138.

Kaptein, R. (1978) *Nuclear Magnetic Resonance Spectroscopy in Molecular Biology* (Pullman, B., Ed.) pp 211-219, Reidel, Dordrecht, Holland.
Lenkinski, R. E., Agresti, D. G., Chen, D. M., & Glickson, J. D. (1978) *Biochemistry 17*, 1463-1468.
Lenkinski, R. E., Dallas, J. L., & Glickson, J. D. (1979) *J. Am. Chem. Soc. 101*, 3071-3077.
Mallion, R. B. (1971) *J. Chem. Soc. B*, 681-686.
Mallion, R. B. (1978) *Nuclear Magnetic Resonance Spectroscopy in Molecular Biology* (Pullman, B., Ed.) pp 183-191, Reidel, Dordrecht, Holland.
McDonald, C. C., & Phillips, W. D. (1969) *J. Am. Chem. Soc. 91*, 1513-1521.
Pauling, L. (1936) *J. Chem. Phys. 4*, 673-677.
Perkins, S. J. (1977) D. Phil. Thesis, University of Oxford.
Perkins, S. J. (1979) *J. Magn. Reson.* (in press).
Perkins, S. J., & Wüthrich, K. (1979) *Biochim. Biophys. Acta 576*, 409-423.
Perkins, S. J., Johnson, L. N., Phillips, D. C., & Dwek, R. A. (1977a) *FEBS Lett. 82*, 17-22.
Perkins, S. J., Dower, S. K., Gettins, P., Wain-Hobson, S., & Dwek, R. A. (1977b) *Biochem. J. 165*, 223-225.
Perkins, S. J., Johnson, L. N., Machin, P. A., & Phillips, D. C. (1978) *Biochem. J. 173*, 607-616.
Perkins, S. J., Johnson, L. N., Machin, P. A., & Phillips, D. C. (1979) *Biochem. J. 181*, 21-36.
Pople, J. A. (1956) *J. Chem. Phys. 24*, 1111-1111.
Roberts, G. C. K., & Jardetsky, O. (1970) *Adv. Protein Chem. 24*, 447-545.
Shindo, H., Cohen, J. S., & Rupley, J. A. (1977) *Biochemistry 16*, 3879-3882.
Shrake, A., & Rupley, J. A. (1973) *J. Mol. Biol. 79*, 351-371.
Snyder, G. H., Rowan, R., Karplus, S., & Sykes, B. D. (1975) *Biochemistry 14*, 3765-3777.
Sternberg, M. J. E., Grace, D. E. P., & Phillips, D. C. (1979) *J. Mol. Biol. 130*, 231-253.
Sternlicht, H., & Wilson, D. (1967) *Biochemistry 6*, 2881-2892.
Wagner, G., & Wüthrich, K. (1975) *J. Magn. Reson. 20*, 435-445.
Weast, R. C., Ed. (1971) *Handbook of Chemistry and Physics*, 51st ed., p F-154, Chemical Rubber Co., Cleveland, OH.
Williams, R. J. P. (1976) The Merck Sharp & Dohme Research Laboratories Scientific Lecture, Merck Sharp & Dohme, Rahway, NJ.

Proton Magnetic Resonance Study of *p*-Mercuribenzoate Binding and Structural Changes in Methemoglobin†

Saburo Neya and Isao Morishima*

ABSTRACT: Interaction of human adult methemoglobin with *p*-mercuribenzoate (*p*MB) was examined at 21 °C by monitoring the hyperfine-shifted proton nuclear magnetic resonance (NMR) spectra of several high- and low-spin derivatives. The NMR spectra show that the heme methyl proton resonances from the β subunits in methemoglobin were selectively affected by the binding of *p*MB regardless of whether the heme iron was saturated with high-spin or low-spin ligand. This observation suggests that the binding of *p*MB to methemoglobin induces a localized tertiary structural change around the β heme, leaving the α heme unaffected. The structural change of the β subunit was correlated with an increase in the high-spin character of the β heme iron. A model study of the azide-methemoglobin complex suggested that the increase of the high-spin character of the β heme iron is due to a conformational change of the proximal histidine which weakens the interaction between the heme iron and the proximal base. A similar and more pronounced spectral change due to binding of *p*MB was observed for the isolated β subunit. The NMR spectral change in the isolated β subunit also suggests that the binding of *p*MB to methemoglobin induces a localized conformational change within the β subunit.

The binding of *p*-mercuribenzoate, *p*MB,[1] to hemoglobin has been examined by many investigators either to determine sulfhydryl reactivity as a measure of quaternary conformation (Gibson, 1973) or to study the functional property of the *p*MB-bound hemoglobin and subunits (Antonini & Brunori, 1971). The spectral properties of the native and *p*MB-bound hemoglobin have been compared in several ways using, for example, visible absorption (Banerjee et al., 1969), circular dichroism (Geraci & Li, 1969), and ^{13}C NMR (Dill et al., 1978). The function of the *p*MB-bound hemoglobin was discussed on the basis of the spectroscopic results. However, the *p*MB-induced spectral changes were not intensively interpreted in terms of structure.

Recently Olson (1976a) reported that the visible absorption change induced by *p*MB is similar to that induced by an allosteric effector, inositol hexaphosphate. Since the change of visible absorption due to the binding of inositol hexaphosphate was proposed to induce a localized tertiary structural change in the β subunits of methemoglobin (Hensley et al., 1978; Edelstein & Gibson, 1975; Olson, 1976b), he speculated that the *p*MB-induced absorption change may reflect a localized structural alteration in the β subunits (Olson, 1976a). However, others (Perutz et al., 1974) have suggested that inositol hexaphosphate induces the quaternary conformational transition of methemoglobin from the relaxed to the tensed state. Thus, the structural basis for the visible absorption

† From the Department of Hydrocarbon Chemistry, Faculty of Engineering, Kyoto University, Kyoto 606, Japan. *Received May 17, 1979.* This work was supported by research grants from the Ministry of Education, Japan, and from the Toray Research Foundation for Science and Technology.

[1] Abbreviations used: *p*MB, *p*-mercuribenzoate; NMR, nuclear magnetic resonance; ppm, parts per million; Me_4Si, tetramethylsilane; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; β(SH), the β subunit with free cysteines; β(*p*MB), the β subunit with *p*MB-bound cysteines.

0006-2960/80/0419-0258$01.00/0