Maeda, M., Saneyoshi, M., & Kawazoe, Y. (1971) *Chem. Pharm. Bull. 19*, 1641-1649.
Pardi, A., Morden, K. M., Patel, D. J., & Tinoco, I., Jr. (1983) *Biochemistry 22*, 1107-1113.
Patel, D. J., Kozlowski, S. A., Rice, J. A., Broka, C., & Itakura, K. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7281-84.
Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., & Jones, R. A. (1986) *Biochemistry 25*, 1027-1036.
Pichat, L., Godbillon, J., & Herbert, M. (1973) *Bull. Chim. Soc. Fr.*, 2715-2719.
Rabi, J. A., & Fox, J. J. (1973) *J. Am. Chem. Soc. 95*, 1628-1632.
Richter, E., Loeffler, J. E., & Taylor, E. C. (1960) *J. Am. Chem. Soc. 82*, 3144-3146.
Roy, S., Papaztavros, M. Z., & Redfield, A. G. (1982) *Nucleic Acids Res. 10*, 8341-8349.
Sanchez, V., Redfield, A., Johnston, P. D., & Tropp, J. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5659-5662.
Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Nucleic Acids. Res. 12*, 4305-4047.
Wemmer, D. E., Chou, S.-H., Hare, D. R., & Reid, B. R. (1984) *Biochemistry 23*, 2262-2268.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Sequential ^{1}H NMR Assignments and Secondary Structure of Hen Egg White Lysozyme in Solution†

Christina Redfield* and Christopher M. Dobson
Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, England
Received July 14, 1987; Revised Manuscript Received October 13, 1987

ABSTRACT: Assignments for ^{1}H NMR resonances of 121 of the 129 residues of hen egg white lysozyme have been obtained by sequence-specific methods. Spin systems were identified with phase-sensitive two-dimensional (2-D) correlated spectroscopy and single and double relayed coherence transfer spectroscopy. For key types of amino acid residues, particularly alanine, threonine, valine, and glycine, complete spin systems were identified. For other residues a less complete definition of the spin system was found to be adequate for the purpose of sequential assignment. Sequence-specific assignments were achieved by phase-sensitive 2-D nuclear Overhauser enhancement spectroscopy (NOESY). Exploitation of the wide range of hydrogen exchange rates found in lysozyme was a useful approach to overcoming the problem of spectral overlap. The sequential assignment was built up from 21 peptide segments ranging in length from 2 to 13 residues. The NOESY spectra were also used to provide information about the secondary structure of the protein in solution. Three helical regions and two regions of β-sheet were identified from the NOESY data; these regions are identical with those found in the X-ray structure of hen lysozyme. Slowly exchanging amides are generally correlated with hydrogen bonding identified in the X-ray structure; a number of exceptions to this general trend were, however, found. The results presented in this paper indicate that highly detailed information can be obtained from 2-D NMR spectra of a protein that is significantly larger than those studied previously.

Hen egg white lysozyme, an enzyme containing 129 amino acid residues, was one of the first proteins to be studied by NMR[1] (Meadows et al., 1967; McDonald & Phillips, 1967; Sternlicht & Wilson, 1967; Cohen & Jardetsky, 1968). At an early stage it was found that many resonances in the complex ^{1}H NMR spectrum are strongly perturbed by magnetic interactions within the folded protein (McDonald & Phillips, 1967; Sternlicht & Wilson, 1967). This results in a number of individual resonances being sufficiently well resolved to be observed separately from the mass of overlapping signals. Assignments for these resolved resonances were proposed, initially on the basis of calculations of ring-current shifts associated with aromatic residues, with the assumption that the structure of the protein in solution was identical with that defined in the crystalline state by X-ray diffraction studies (Sternlicht & Wilson, 1967; McDonald & Phillips, 1970).

The number of assignments in the spectrum of lysozyme has increased steadily over the last 15 years as higher field spectrometers have become available and as new methods for resolving and identifying resonances have been developed. Assignments have been reported for protons of some 50 of the

† This work was supported by the U.K. Science and Engineering Research Council. This work is a contribution from the Oxford Enzyme Group which is supported by the SERC.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2-D, two dimensional; lysozyme, hen egg white lysozyme; COSY, two-dimensional *J*-correlated spectroscopy; RELAY, two-dimensional relayed coherence transfer spectroscopy; TM, trapezoidal multiplication; DM, double exponential multiplication; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional NOE spectroscopy; type J, amino acid residue belonging to the group consisting of Trp, Tyr, Phe, His, Asp, Asn, Cys, and Ser; type U, amino acid residue belonging to the group consisting of Lys, Arg, Met, Gln, Glu, Leu, and Ile; type X, any of the 20 common amino acid residues; $d_{\alpha N}(i,j)$, NOE connectivity between the αCH proton on residue i and the NH proton on residue j; $d_{NN}(i,j)$, NOE connectivity between the NH proton on residue i and the NH proton on residue j; $d_{\beta N}(i,j)$, NOE connectivity between the βCH proton on residue i and the NH proton on residue j; $d_{\alpha N}$, $d_{\alpha N}(i,i+1)$; d_{NN}, $d_{NN}(i,i+1)$; $d_{\beta N}$, $d_{\beta N}(i,i+1)$; $d_{\alpha P\delta}$, NOE connectivity between the αCH proton on residue i and the δCH proton on proline residue $i+1$.

0006-2960/88/0427-0122$01.50/0

129 amino acid residues of lysozyme (Meadows et al., 1967; Campbell et al., 1975a; Chapman et al., 1978; Dobson et al., 1978; Cassels et al., 1978; Lenkinski et al., 1979; Poulsen et al., 1980; Delepierre et al., 1982, 1984; Redfield et al., 1982; Hore & Kaptein, 1983). Most of these were achieved with one-dimensional techniques, and the majority of the resonances were of side-chain protons. All have relied to at least some extent on the broad similarity of the crystal and solution structure of the protein. Two-dimensional techniques, however, have been shown to be of considerable value in the spectroscopy of hen lysozyme (Boyd et al., 1983, 1985a) and have been used extensively in the assignment of the spectrum of the closely related human lysozyme (Boyd et al., 1985b).

The availability of the assignments has permitted detailed studies of individual residues of the protein to be carried out. As well as investigation of the overall structure of the molecule in solution, considerable emphasis has been placed on dynamic properties of the protein (Campbell et al., 1975b; Poulsen et al., 1980; Blake et al., 1981). These range from local fluctuations of side chains (Olejnizcak et al., 1981) to the events associated with hydrogen exchange (Wedin et al., 1982; Delepierre et al., 1983) and the cooperative folding and unfolding of the structure (Dobson & Evans, 1984). In addition, studies have recently been extended to the unfolded state of the protein, where assignments have been made by correlating the spectra of the different states with magnetization transfer techniques (Dobson et al., 1984).

In this paper we report a detailed analysis of 2-D NMR spectra directed toward assignment of the ^{1}H spectrum of hen lysozyme by sequential methods. This approach does not require any assumption about the structure of the protein but does require very extensive analysis of cross-peaks in COSY and NOESY experiments. The approach has recently been used to assign in detail a number of proteins containing up to about 100 residues (Arseniev et al., 1982; Wagner & Wüthrich, 1982; Keller et al., 1983; Štrop et al., 1983; Zuiderweg et al., 1983; Holak & Prestegard, 1986; Klevit et al., 1986; Kline & Wüthrich, 1986; Wagner et al., 1986; Clore et al., 1987; Cooke et al., 1987; Driscoll et al., 1987; Sukumaran et al., 1987) although the task becomes increasingly difficult as the number of residues increases. Nevertheless, we have been able to achieve almost total assignment of the main-chain resonances in lysozyme, a protein of 129 residues. In addition, assignments have enabled independent determination of the secondary structure of lysozyme in solution and permitted a comparison with the structure in the crystalline state. They have also enabled assignments made on the basis of nonsequential methods to be checked. The existence of resonance assignments for virtually every residue (125 out of 129) of lysozyme offers opportunities for increasingly detailed study of the structure and dynamics of an enzyme in solution.

Materials and Methods

Hen lysozyme was obtained from Sigma Chemical Co. and was dialyzed extensively at pH 3 before use. All NMR samples were 7 mM in lysozyme at pH 3.8, and experiments were run at 35 °C. Lysozyme samples were prepared in four ways in order to exploit the wide range of hydrogen exchange rates exhibited by the protein. Nonexchanged lysozyme was prepared by dissolving 50 mg of lyophilized protein in 0.5 mL of 90% H_2O/10% D_2O at 35 °C; the spectrum of a sample prepared in this way contains resonances from all amide hydrogens. Fully exchanged lysozyme was prepared by dissolving the protein in D_2O and heating at 80 °C for 10 min; the spectrum of a sample prepared in this way does not contain resonances from any labile hydrogens. Partly exchanged lysozyme was prepared by dissolving the protein in 99.98% D_2O at 35 °C; the spectrum of such a sample contains resonances from slowly exchanging amide hydrogens. Reverse-exchanged lysozyme was prepared by dissolving lyophilized, fully exchanged lysozyme in 80% H_2O/20% D_2O at 35 °C; the spectrum of such a sample contains resonances from rapidly exchanging amide hydrogens (Keller et al., 1983). Back-exchange of slowly exchanging amides in the reverse-exchanged samples was avoided by first adding 0.1 mL of D_2O to the lyophilized protein and then bringing the volume up to 0.5 mL with H_2O.

All the NMR experiments were performed on the home-built 500-MHz spectrometer of the Oxford Enzyme Group. The spectrometer is equipped with an Oxford Instruments Co. magnet, a GE/Nicolet 1280 computer and 293B pulse programmer, and a Bruker probe. Phase-sensitive J-correlated spectroscopy (COSY) (Aue et al., 1976; Bax & Freeman, 1981) and single and double relayed coherence transfer spectroscopy (RELAY) (Eich et al., 1982; Bax & Drobny, 1985) experiments were performed according to the method of States et al. (1982) and with standard phase cycling schemes. Data sets consisting of 512 t_1 increments of 64, 96, and 128 transients were collected for the COSY, RELAY, and double-RELAY experiments, respectively. Each experiment took between 9.5 and 24 h to complete. A sweep width of 7042 Hz was used in both dimensions. In RELAY and double-RELAY experiments mixing periods (2τ) of 36 ms were used. Spectra were resolution enhanced in t_2 by trapezoidal multiplication (TM) and double exponential multiplication (DM) and in t_1 by TM only. After zero filling twice in the t_1 dimension and once in t_2, the digital resolution was 3.5 Hz/point.

Phase-sensitive NOESY spectra (Jeener et al., 1979; Anil Kumar et al., 1980) were obtained by a different procedure. Sign discrimination in F_1 was achieved by placing the transmitter offset at the downfield edge of the spectrum. This procedure eliminates the "anti-diagonal", which often runs through 2-D spectra obtained with the transmitter offset placed in the center of the spectrum. A sweep width of 14084 Hz was used in both dimensions. Prior to complex Fourier transformation in t_1, the imaginary part of the data set was zeroed. A mixing time of 150 ms which was randomly varied by 10% was used in all NOESY experiments (Macura et al., 1981). Data sets consisting of 512 t_1 increments of 4K data points were collected. Spectra were resolution enhanced in t_2 by TM and DM. No resolution enhancement was used in the t_1 dimension. After zero filling, the digital resolution was 7.0 Hz/point.

All spectra are presented as contour plots. In COSY, RELAY, and double-RELAY spectra both positive and negative levels are shown; in NOESY spectra only positive levels are shown. The contours are spaced logarithmically, so that the nth level corresponds to an intensity of 0.7^n times the chosen scaling factor. In all spectra presented here the horizontal and vertical axes represent F_1 and F_2, respectively.

Results and Discussion

The fingerprint region of the COSY spectrum of hen lysozyme is shown in Figure 1. The assignment of the NH–αCH cross-peaks in this spectrum was carried out by the standard two-stage procedure (Wüthrich, 1986). The first stage involved the assignment of each cross-peak to a type of amino acid. Information about coupled spin systems was obtained from experiments, such as COSY and RELAY, which exploit through-bond scalar couplings. At this stage some cross-peaks could be assigned to a unique amino acid type whereas others could be assigned only to more general

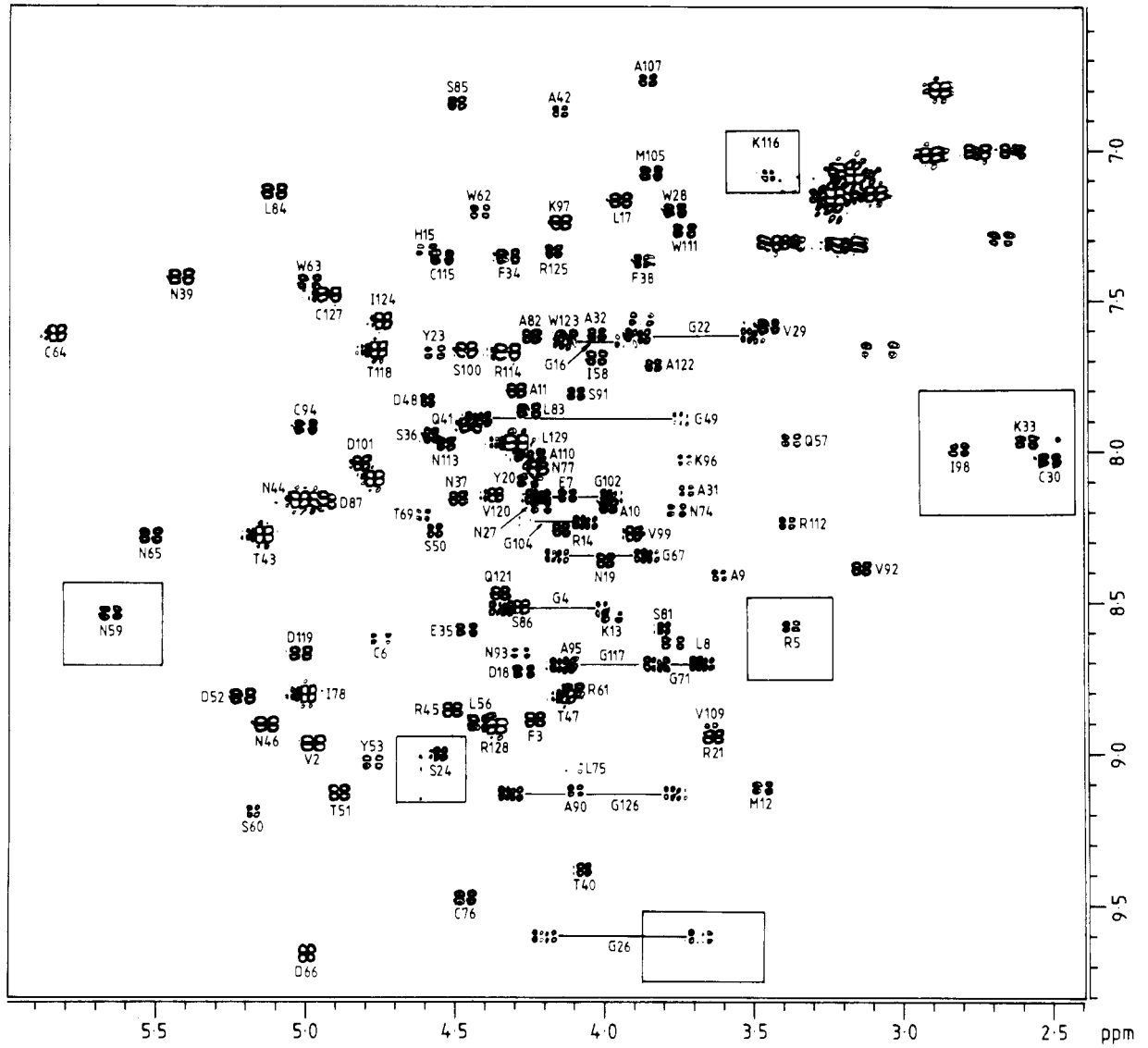

FIGURE 1: Fingerprint region of the phase-sensitive 500-MHz COSY spectrum of nonexchanged hen lysozyme. Both positive and negative contour levels are plotted with a scaling factor of 150; cross-peaks in boxes are plotted at a lower level with a scaling factor of 50. Cross-peaks from W108 and I88 were lost because of saturation of the water peak.

classes of amino acid types. The specificity of the assignment at this first stage depended on whether the complete spin system could be identified and to what extent the topology of the spin system was unique. The second stage involved the assignment of the cross-peaks to specific amino acid residues in the protein sequence. The NOESY spectrum was searched for sequential NOE effects involving neighboring NH, αCH, and βCH protons. Short peptide segments were built up from these sequential NOE effects. The sequences of these peptide segments were compared with the known lysozyme sequence. Sequence-specific assignments were obtained when there was a unique occurrence of the peptide segment in the protein sequence. This method of assignment has been discussed in detail by Wüthrich and co-workers (Billeter et al., 1982; Wüthrich, 1986).

Analysis of Spin Systems. Hen lysozyme is composed of 129 amino acid residues including 2 proline residues. The fingerprint region of the COSY spectrum should, therefore, contain 127 cross-peaks (or pairs of cross-peaks in the case of glycine). At the contour level chosen, cross-peaks from 110 amino acid residues can be seen in the COSY spectrum in Figure 1. At a lower contour level cross-peaks from an additional 10 amino acid residues are visible as illustrated in the boxed regions of Figures 1 and 2. There are several possible reasons for the absence of some cross-peaks. First, intensity may be very low because of either a small NH–αCH coupling constant or a large line width; overlap of antiphase components in the COSY spectrum leads to cancellation of most of the cross-peak intensity (Neuhaus et al., 1985). Second, irradiation of the H_2O resonance can lead to loss of cross-peak intensity for two reasons (Wüthrich, 1986). If the αCH resonance overlaps with the strong H_2O signal at 4.65 ppm, then the NH–αCH cross-peak will disappear when the αCH resonance is saturated along with the solvent. This problem was overcome by collecting spectra at different temperatures since the H_2O resonance shifts significantly with temperature. At 35 °C two cross-peaks are obscured by the solvent line, but both are seen at 55 °C. One of these peaks has a slowly exchanging amide and can be seen in the partly exchanged spectrum at 35 °C. In addition, cross-peaks can be lost if saturation is transferred from the suppressed H_2O signal to rapidly exchanging amide protons. The experiments shown here were carried out at pH 3.8, near to the pH minimum for hydrogen exchange, and none of the backbone amide resonances were significantly affected by saturation of the water resonance. With the additional two peaks identified in the high-temperature spectrum, NH–αCH cross-peaks from a total of 122 residues have been identified, indicating that

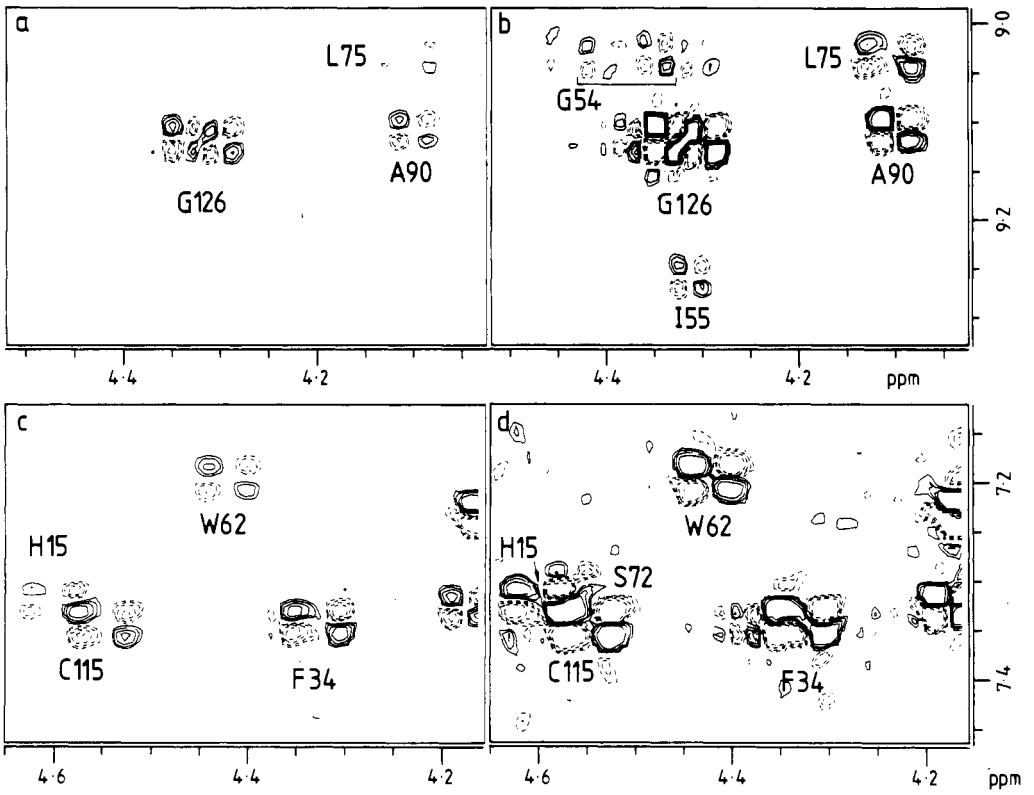

FIGURE 2: Expansions of the fingerprint region of the phase-sensitive 500-MHz COSY spectrum of nonexchanged hen lysozyme plotted at two different contour levels. Positive and negative contours are plotted with solid and broken lines, respectively. Spectra in a and c are plotted with a scaling factor of 150. Spectra in b and d are plotted with a scaling factor of 50; cross-peaks of G54, I55, and S72 are visible in these spectra.

cross-peaks from 5 amino acid residues have not been detected.

The majority of cross-peaks in the fingerprint region of the spectrum have the same antiphase fine structure and cannot be assigned to an amino acid type on the basis of their appearance in the COSY spectrum alone. The cross-peaks of glycine, however, have a unique fine structure pattern that results from the coupling of the NH to two αCH protons. Pairs of cross-peaks from all 12 glycine residues can be seen in Figures 1 and 2; the overlap of cross-peaks from two glycine residues, G71 and G117, is overcome at 55 °C.

Fingerprint-region cross-peaks of alanine, threonine, valine, isoleucine, and leucine can also be identified uniquely if a connectivity between the NH and the methyl peaks can be established. These connectivities can be identified in COSY, RELAY, and double-RELAY experiments. Complete spin systems can be traced through COSY spectra unambiguously when resonances have unique chemical shift values. However, resonance overlap becomes an increasing problem as the number of residues increases; in this study additional data from RELAY and double-RELAY spectra were required to allow unambiguous analysis of COSY spectra and the identification of spin systems. All 12 alanine spin systems of lysozyme were identified from NH–βCH$_3$ cross-peaks in RELAY spectra as shown in Figure 3. The connectivity between the NH and γCH$_3$ groups of threonine, valine, and isoleucine could be most clearly established from NH–γCH$_3$ cross-peaks in double-RELAY spectra; cross-peaks for several threonine and valine residues of lysozyme are shown in Figure 3. In the absence of double-RELAY peaks, spin systems of these residues were identified from NH–βCH and αCH–γCH$_3$ RELAY peaks and from NH–γCH$_3$, NH–βCH, and αCH–γCH$_3$ cross-peaks in NOESY spectra; αCH–γCH$_3$ RELAY peaks for several valine and threonine residues are shown in Figure 4. By use of a combination of these experiments, fingerprint-region COSY peaks for six of seven threonine residues, five of six

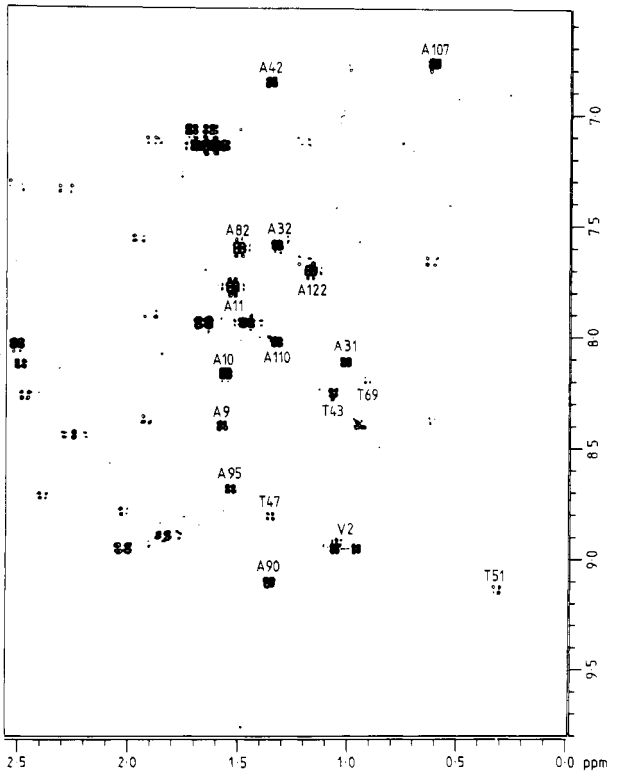

FIGURE 3: Phase-sensitive 500-MHz double-RELAY spectrum of nonexchanged hen lysozyme. NH–βCH$_3$ peaks of all 12 alanine residues and NH–γCH$_3$ peaks of one valine and four threonine residues are labeled. The alanine peaks were assigned unambiguously because they also appear in the RELAY spectrum. The valine and threonine peaks appear in the double-RELAY spectrum only.

valines, and three of six isoleucines were identified. The αCH resonance of the seventh threonine is at a position where there are no NH–αCH cross-peaks in the COSY spectrum; this

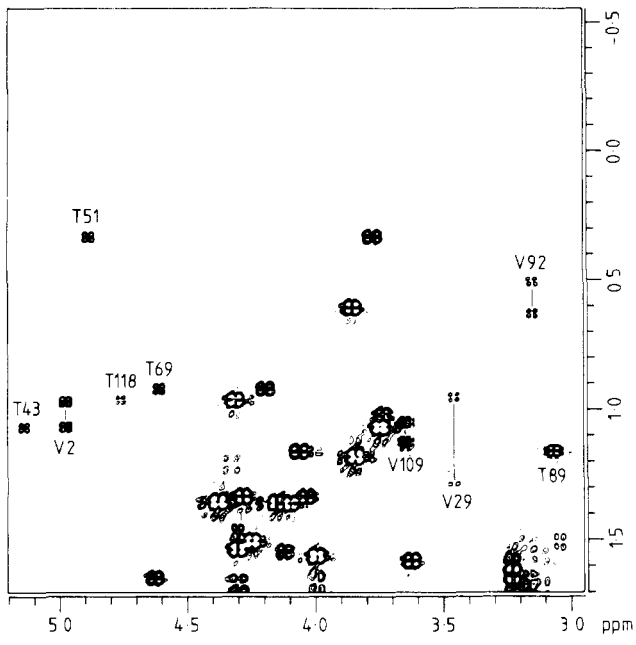

FIGURE 4: The 500-MHz phase-sensitive RELAY spectrum of fully exchanged hen lysozyme. αCH–γCH_3 RELAY peaks of five threonine and four valine residues are labeled.

threonine accounts for one of the five missing COSY cross-peaks. No data are, at this stage, available for the NH and the αCH positions of the sixth valine spin system although the βCH and γCH_3 resonances were identified. The βCH and γCH_3 positions of the three remaining isoleucine residues were identified, but the NH and αCH positions for these residues were not found. Identification of the γCH_2–δCH_3 resonances of the isoleucine spin systems was carried out on the basis of NOE effects between the βCH, γCH_2, γCH_3, and δCH_3 resonances. In principle, the αCH position of a leucine residue can be identified from pairs of αCH–δCH_3 cross-peaks in double-RELAY spectra; none of the eight leucine residues in lysozyme, however, gives rise to detectable peaks. RELAY peaks between βCH and δCH_3 resonances were, nevertheless, observed for some of the leucine residues. The αCH position of a leucine can be identified by an alternative approach using NOE effects. The αCH proton is always in close proximity ($\leq$2.5 Å) to one of the δCH_3 groups, and a strong NOE should be observed between these resonances. In general, this is the closest αCH to the δCH_3 group, although the NOE effect may be difficult to assign unambiguously when several methyl resonances overlap or when the methyl group gives effects to more than one αCH resonance position. The NH position of the leucine can then be identified from the NH–αCH COSY peak and from NH–βCH RELAY peaks. Fingerprint-region cross-peaks for two of the eight leucine residues were identified unambiguously in this way.

The glycines and the five methyl-containing residues account for 51 of the 129 amino acid residues of lysozyme. At this stage, cross-peaks in the fingerprint region arising from 40 residues have been assigned uniquely to one of these six residue types, and in addition, one of the five missing cross-peaks has been accounted for by a threonine residue. The 78 remaining residues in lysozyme fall into three categories. First, the proline residues do not have an amide proton and do not give a peak in the fingerprint region. The two remaining categories can be distinguished, in principle, on the basis of βCH chemical shift values and cross-peak fine structure patterns. The βCH chemical shifts were obtained from COSY spectra when the αCH resonance was resolved or otherwise from the NH–βCH peaks in RELAY spectra. Eight types of amino acid residues (Asp, Asn, Tyr, His, Phe, Trp, Ser, Cys) have two βCH protons and no γCH protons. In fully exchanged COSY spectra the αCH and βCH cross-peaks have easily recognizable fine structure patterns characteristic of AMX spin systems. In random coil structures these residues have βCH chemical shifts in the range 2.75–3.9 ppm (Bundi & Wüthrich, 1979). In this study, spin systems with βCH resonances downfield of 2.5 ppm and αCH–βCH cross-peak patterns characteristic of AMX spin systems were assigned to this group of eight residues and are classified as type J.[2] The observation of NOE effects between certain aromatic ring protons and the βCH protons has been used by other workers to distinguish tryptophan, tyrosine, phenylalanine, and histidine residues from the other four type J spin systems (Wüthrich, 1986). In view of the large number of type J residues and the overlap in the βCH region of the spectrum, this distinction was not made at this point. Instead, these NOE effects were used at a later stage to reinforce assignments of NH–αCH cross-peaks arising from aromatic residues. Hen lysozyme contains 52 type J residues, and 29 of these could be unambiguously identified on the basis of the COSY and RELAY spectra. The five remaining types of residues (Met, Lys, Arg, Gln, Glu) have βCH protons that are coupled to two γCH protons. Resonances beyond the β-position are usually difficult to identify because the large number of coupled protons gives rise to complex coupling patterns in which overlap of antiphase components leads to cancellation of cross-peak intensity and because the similarity of βCH and γCH chemical shifts can give rise to cross-peaks close to the diagonal. In random coil structures these five residues have βCH chemical shifts in the range 1.7–2.2 ppm (Bundi & Wüthrich, 1979). In this study spin systems with βCH resonances upfield of 2.3 ppm were assigned to this group of residue types and are classified as type U. The eight leucine and isoleucine residues that were not identified above also fall into the type U category. Hen lysozyme contains 32 type U residues; 14 of these were unambiguously identified on the basis of COSY and RELAY spectra. There are 39 fingerprint-region cross-peaks in the COSY spectrum which, at this stage of the analysis, could not be assigned to a particular group of amino acid residues; these are classified as type X.

Sequential Assignments. The second stage of assignment involves the correlation of a cross-peak in the fingerprint region with a specific amino acid residue in the protein sequence. The NOESY spectrum of hen lysozyme was searched for sequential d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$ connectivities (Wüthrich et al., 1984; Wüthrich, 1986); the d_{NN} and $d_{\alpha N}$ connectivities identified in the NOESY spectrum are shown in Figures 5 and 6 and are summarized in Figure 7. The length of a sequentially assigned peptide segment is limited by several factors. Overlap of the NH resonances of two neighboring residues prevents the identification of a d_{NN} connectivity and leads to a break in the sequential assignment. Similarly, overlap of an αCH resonance with the solvent resonance can result in a missing $d_{\alpha N}$ connectivity because the αCH resonance is saturated along with the solvent. Proline residues in the sequence can also lead to breaks; the $d_{\alpha P\delta}$ connectivities observed for proline are more difficult to assign unambiguously than the $d_{\alpha N}$ connectivities seen for the other amino acids. Some of these problems were overcome, for example, by collecting spectra at different temperatures or by searching NOESY spectra for $d_{\beta N}$ connectivities. Overlap of NH or αCH chemical shifts can also lead to ambiguities in the sequential assignment procedure;

[2] The classifications type J and type U were chosen because the letters J and U are the only ones not used in the one-letter amino acid code.

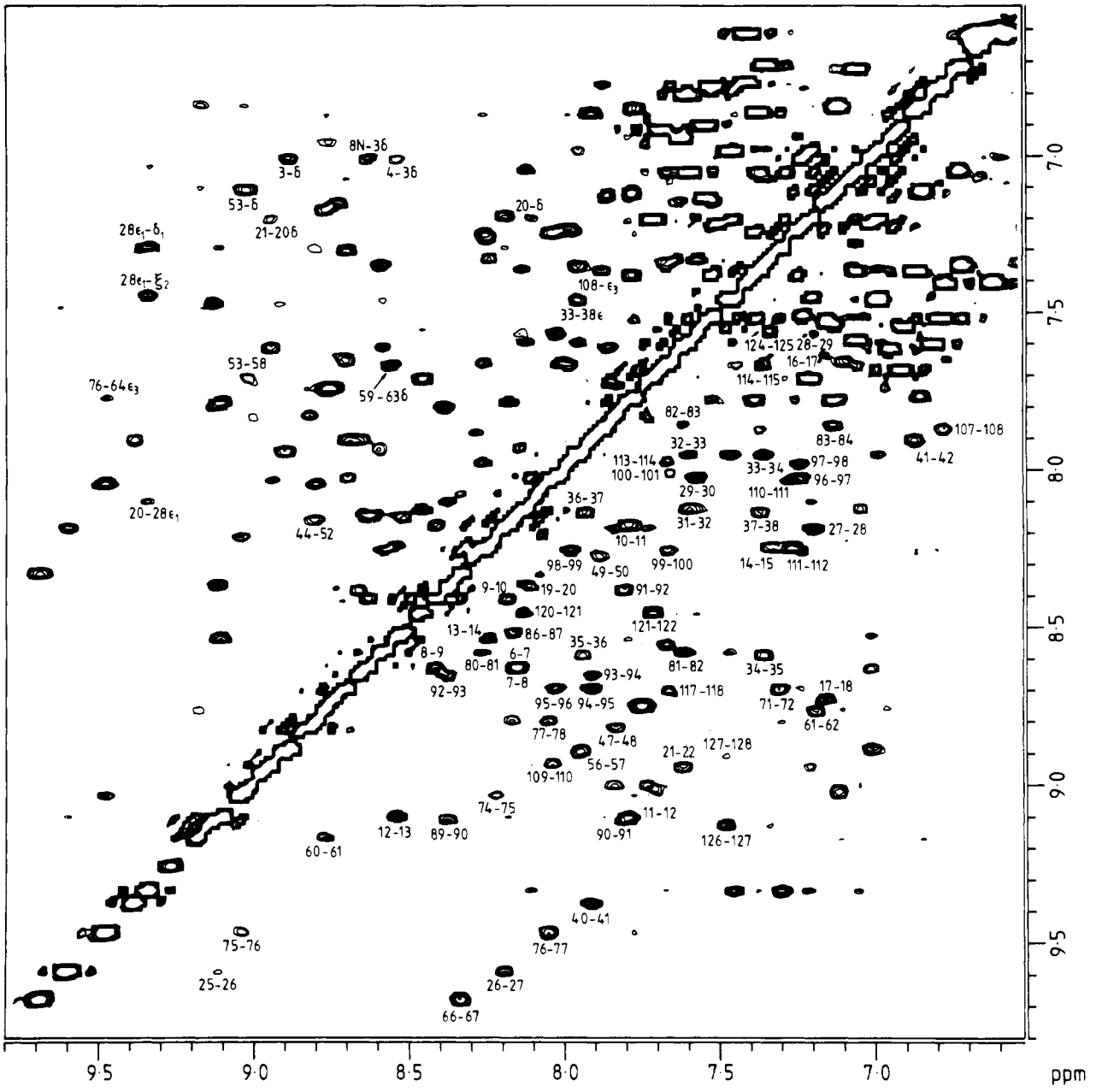

FIGURE 5: NH and aromatic region of the phase-sensitive 500-MHz NOESY spectrum of nonexchanged hen lysozyme, τ_m = 150 ms. Connectivities involving pairs of neighboring NH protons are labeled in the region below the diagonal. Long-range d_{NN} connectivities and NH–aromatic connectivities are labeled in the region above the diagonal.

these problems were overcome, in some cases, by exploiting differences in hydrogen exchange rates. The segments of sequentially assigned residues and the order in which these segments were assigned are summarized in Figure 7.

The probability of identifying a unique peptide segment within the lysozyme sequence generally increases as the length of the sequentially assigned peptide segment increases. In this study, however, it was found to be particularly important to identify, initially, sequences that contain several residues, such as glycine and alanine, which had been assigned uniquely in the fingerprint region in the first stage of assignment. All 12 of the alanine spin systems have been identified in the spectrum of hen lysozyme, and these residues are scattered throughout the protein sequence shown in Figure 7. A study of the lysozyme sequence shows that all the alanine residues are found in close sequential proximity to other easily identifiable residues such as threonine and valine. Therefore, the alanine residues were chosen as the starting point for the sequential assignment task.

The longest stretch of sequential connectivities identified in the spectrum of hen lysozyme is illustrated in Figure 8. Peptide segment 1, consisting of 13 residues, was assigned to T89 through D101 on the basis of the A-X-V and A-X-U-I-V sequences. The NH chemical shift of T89 does not correspond to any of the six threonine spin systems identified earlier; therefore, T89 must correspond to the threonine residue with a missing cross-peak in the fingerprint region of the COSY spectrum. Problems of overlap in the NH and αCH chemical shifts were overcome by use of the additional $d_{\alpha N}$, $d_{\beta N}$, and $d_{\alpha N}(i,i+3)$ connectivities observed and by exploitation of the range of hydrogen exchange rates observed for the amide hydrogens of lysozyme. COSY spectra collected under three different sets of hydrogen exchange conditions are shown in Figure 8. The amide hydrogens of residues 92–99 exchange slowly with D_2O and are observed in the partly exchanged spectrum, Figure 8b. The amide hydrogens of residues 89–91 and 100–101, on the other hand, exchange rapidly with D_2O and are observed in the reverse-exchanged spectrum, Figure

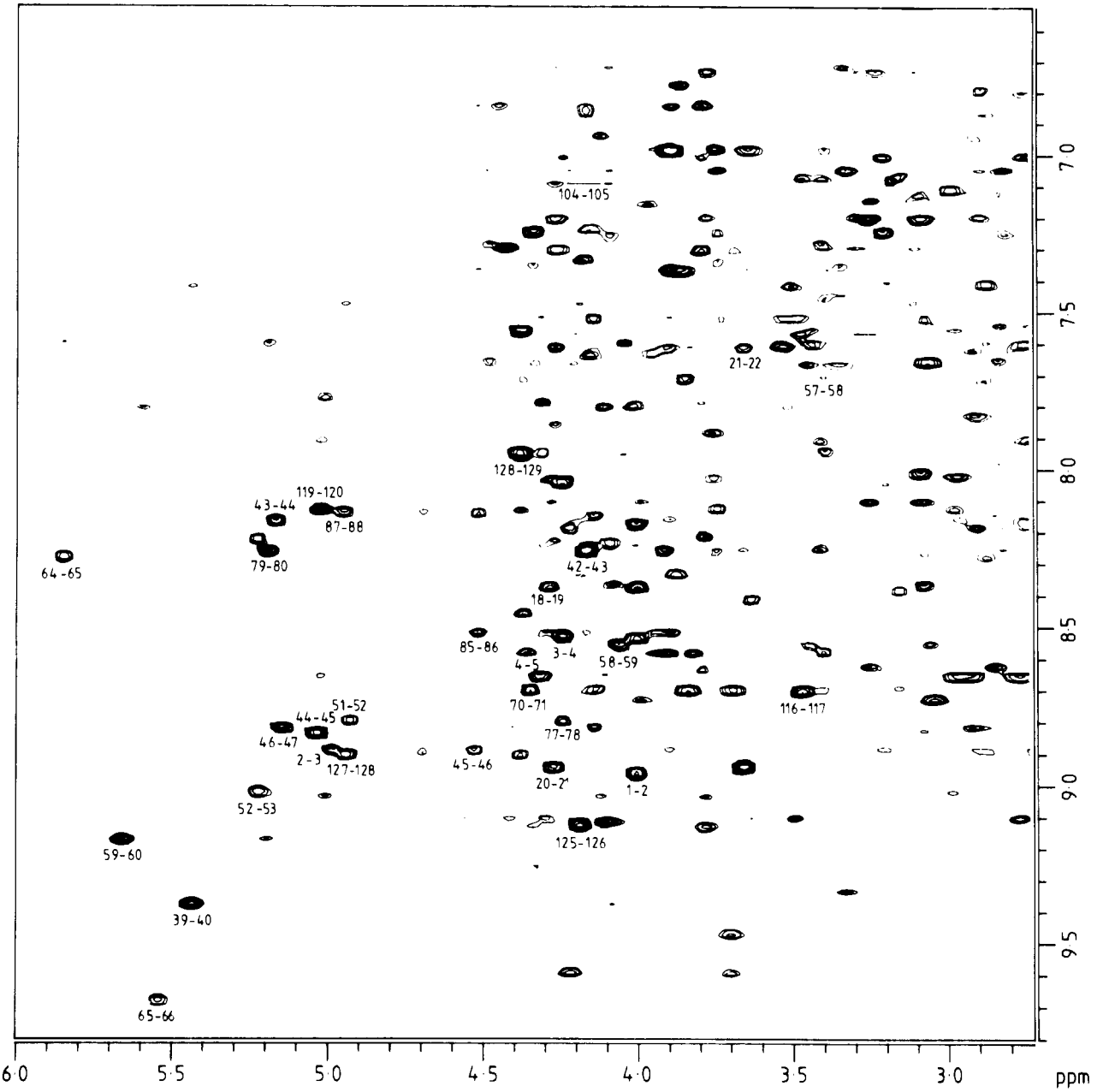

FIGURE 6: Phase-sensitive 500-MHz NOESY spectrum of nonexchanged hen lysozyme, τ_m = 150 ms. Sequential $d_{\alpha N}$ connectivities identified in the assignment process are labeled.

8c. The simplification observed in the fingerprint region of the COSY spectra was very useful for resolving cross-peaks in crowded regions of the spectrum. For example, the cross-peak of A95 is obscured in the nonexchanged spectrum by overlap with one of the G117 cross-peaks. A95 has a slowly exchanging amide and gives a clear cross-peak in the partly exchanged spectrum whereas G117 has a rapidly exchanging amide and gives a cross-peak in the reverse-exchanged spectrum. Similar spectral simplification is observed in the NH–αCH region of the NOESY spectrum. Additional simplification occurs in the NH–NH region of the NOESY spectrum as a cross-peak between a pair of amide hydrogens is observed in the partly exchanged or the reverse-exchanged spectrum only if the two amides have similar hydrogen exchange properties. An NOE cross-peak between one slowly exchanging amide and one rapidly exchanging amide will only appear in the nonexchanged spectrum. Thus, d_{NN} connectivities for residues 92–99 are seen in the partly exchanged NOESY spectrum, Figure 8b, whereas the connectivities between residues 89–91 and residues 100–101 are seen in the reverse-exchanged NOESY spectrum, Figure 8c. The d_{NN} connectivities involving residues 91–92 and 99–100 appear in the nonexchanged spectrum only, Figure 8a. The approach described above is very useful when two or more peaks in the fingerprint region share the same NH chemical shift but have different hydrogen exchange characteristics.

A second sequence of 10 residues, peptide segment 2, was identified from d_{NN} connectivities shown in Figures 5 and 9. The segment was assigned to residues C6 through H15 on the basis of the unique L-A-A-A sequence. In cases where overlap in the NH chemical shifts occurred, these resonances could be resolved by means of the large number of $d_{\alpha N}$ and $d_{\alpha N}$-(i,i+3) connectivities shown in Figure 9. For example, the observed $d_{\alpha N}$ connectivity from the NH of A11 to the αCH of A10 allowed the cross-peaks of A10 and N27 to be distinguished even though they share the same NH chemical shift and hydrogen exchange characteristics.

A peptide segment of eight residues, segment 3, was assigned to A31 through F38 on the basis of the A-A sequence. Resonances of N39 through D48 were assigned on the basis

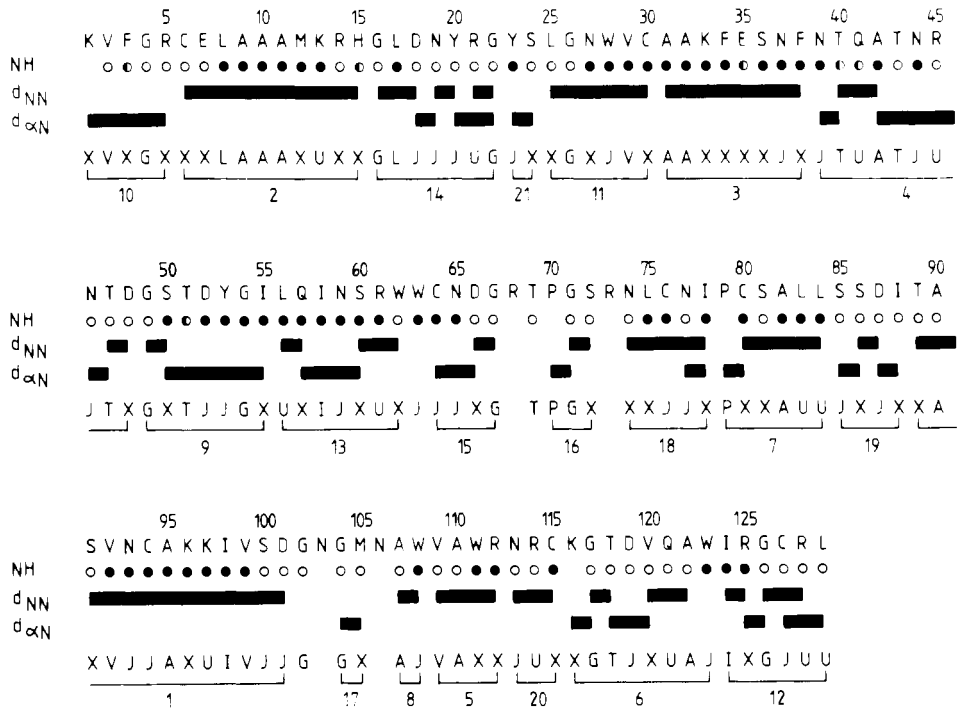

FIGURE 7: Amino acid sequence of hen lysozyme, summary of hydrogen exchange rates, and summary of the NOE connectivities used in the sequential assignment procedure. Filled, half-filled, and empty circles indicate amides with slow, intermediate, and rapid exchange rates, respectively (see text). Peptide segments identified in the assignment procedure and the amino acid types of the residues in these segments are shown. The numbers beneath these segments indicate the order in which they were assigned.

of the unique T-U-A-T sequence occurring in peptide segment 4. The $d_{\alpha N}$ connectivities for residues 42–47 are illustrated in detail in Figure 10. The four remaining alanine residues were assigned on the basis of the sequentially assigned peptide segments 5–8 summarized in Figure 7. The NH–αCH fingerprint-region cross-peak of V120, in peptide segment 6, had not been identified as a valine in the analysis of spin systems, and therefore, this residue accounts for the missing valine spin system. The spin system of W123 was assigned from a $d_{\beta N}$ connectivity to the βCH_3 resonance of A122 in peptide segment 6. There was no evidence for a $d_{\alpha N}$ connectivity between this pair of residues, and the near overlap of the two NH resonances prevented the identification of a d_{NN} connectivity. No cross-peak is observed in the COSY fingerprint region for C80 in peptide segment 7; this residue accounts for the second missing cross-peak in the fingerprint region. The αCH position of this residue has not been identified in this study. Assignments for 61 of the 129 residues of hen lysozyme were obtained with only the 12 alanine spin systems as a starting point.

Five of the seven threonine residues have at this stage been assigned by virtue of their proximity to alanine residues in the sequence. T51 can be assigned from a series of NOE connectivities, peptide segment 9, involving residues G49 through I55. Six of these residues, S50 through I55, have slowly exchanging amides, and their $d_{\alpha N}$ connectivities can be clearly seen in the partly exchanged NOESY spectrum shown in Figure 11. The connectivities involving the αCH resonances of S50 and Y53 were not observed in the nonexchanged NOESY spectrum because of their close proximity to the saturated water peak. The last threonine cross-peak in the fingerprint region was assigned to T69 by elimination. No NOE effects from T69 to R68 or P70 could, however, be unambiguously identified in the 2-D spectra. Four of the six valine residues have already been identified as a result of their proximity to alanine residues. The two remaining valine residues, V2 and V29, were assigned from peptide segments 10 and 11 as summarized in Figure 7. Residues K1, in peptide segment 10, and L25, in segment 11, do not give cross-peaks in the fingerprint region of the COSY spectrum; these two residues account for the third and fourth residues of the five missing fingerprint-region cross-peaks. One of the three complete isoleucine spin systems was assigned as a result of its proximity to an alanine residue in the sequence. The two remaining isoleucine residues, I124 and I58, were identified from the sequential NOE connectivities of peptide segments 12 and 13. Six of the eight leucine residues of hen lysozyme were assigned because of their proximity in the sequence to alanine, threonine, valine, or isoleucine residues. NH and αCH resonances for one of the remaining leucine residues were identified in the first stage of assignment. This spin system is contained in the seven-residue peptide segment 14 assigned to residues G16 through G22.

Eight of the 12 glycine residues have at this stage been assigned on the basis of their proximity to other identified residue types. Two of the four remaining glycines can now be assigned on the basis of their positions relative to type J and type U residues in the sequence. Peptide segments 15 and 16 were assigned to residues C64 through G67 and P70 through S72, respectively. The αCH resonances of one of the two remaining glycine residues are involved in $d_{\alpha N}$ connectivities to a type X spin system, segment 17. At this point it is not possible to distinguish between G102-N103 and G104-M105.

Assignments for 107 of the 129 residues of lysozyme have now been made with alanine, threonine, valine, isoleucine, leucine, and glycine residues as reference points. At this stage of the assignment, gaps of two or more residues are found only at five positions in the sequence, 23–24, 73–78, 85–88, 102–106, and 113–115. The restrictions on the length and the sequence of these unassigned groups of residues now make feasible assignment on the basis of the type J and type U classifications alone. A peptide segment of five residues, segment 18, must correspond to R73–N77, N74–I78, or G102–N106, the only remaining gaps of five residues. The

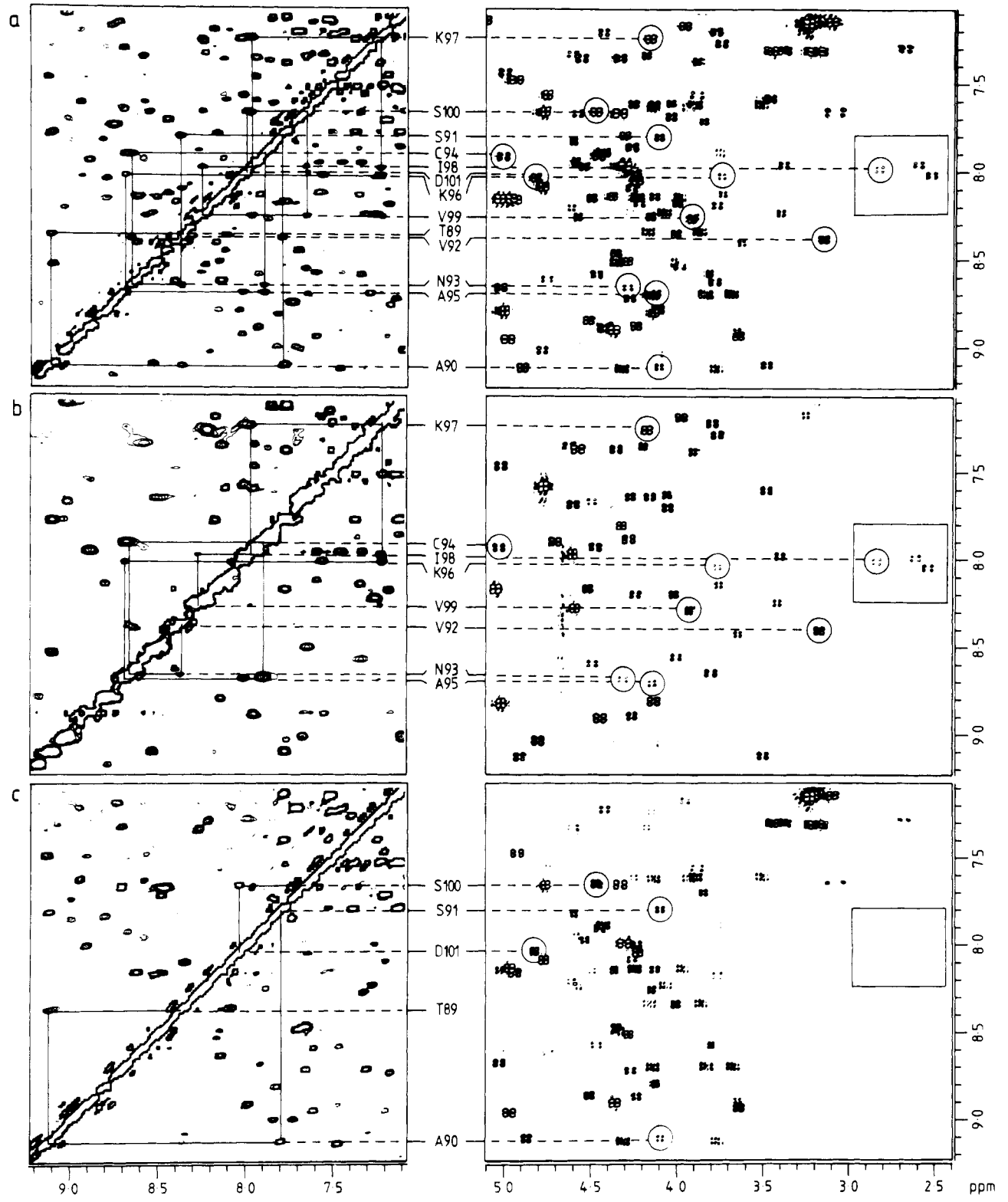

FIGURE 8: Phase-sensitive 500-MHz NOESY (left) and COSY (right) spectra of nonexchanged (a), partly exchanged (b), and reverse-exchanged (c) hen lysozyme. Sequential d_{NN} NOE connectivities for residues 89–101 are shown on the left, and the corresponding fingerprint-region COSY peaks are circled on the right.

fragment was assigned to N74 through I78 on the basis of the J–J pair. Segment 19, which contains four residues, was assigned to S85 through I88 on the basis of the J-X-J-X sequence. The three residues of segment 20 with sequence J-U-X were assigned to N113 through C115. The last peptide segment, segment 21, was assigned to Y23-S24 on the basis of the J-X sequence.

Resonances corresponding to 121 of the 129 residues of hen lysozyme have been assigned with the sequential NOE data presented here. The assignments of the three tyrosine and three phenylalanine residues were reinforced by NOE effects observed between the H^{δ} ring protons and one or both of the NH and αCH protons. In addition, NOE effects between tryptophan ring protons and backbone protons reinforced the assignments of W28, W108, and W123. Resonances of W63, R68, R73, G102, N103, G104, M105, and N106 have not been assigned at this stage. Four of the five missing fingerprint-region COSY peaks have been accounted for by K1, L25,

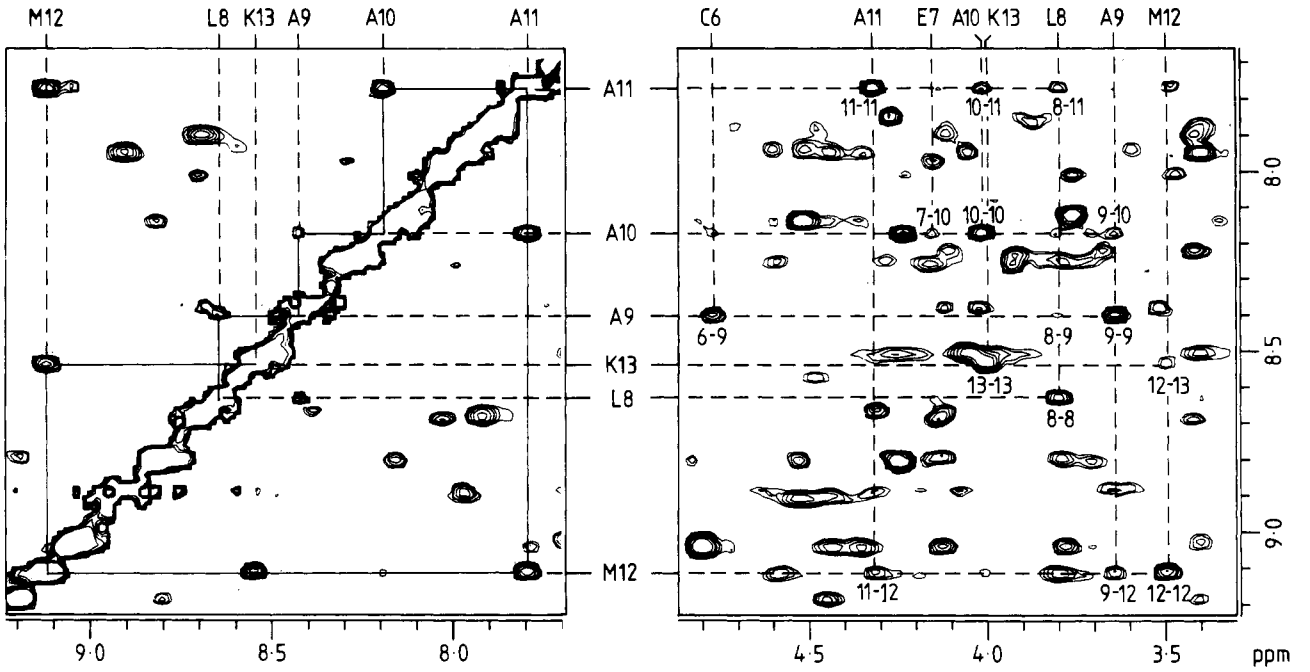

FIGURE 9: Phase-sensitive 500-MHz NOESY spectra of partly exchanged hen lysozyme. Sequential d_{NN} connectivities for residues 8–13 are shown on the left; $d_{\alpha N}(i,i)$, $d_{\alpha N}(i,i+1)$, and $d_{\alpha N}(i,i+3)$ connectivities for residues 6–13 are shown on the right.

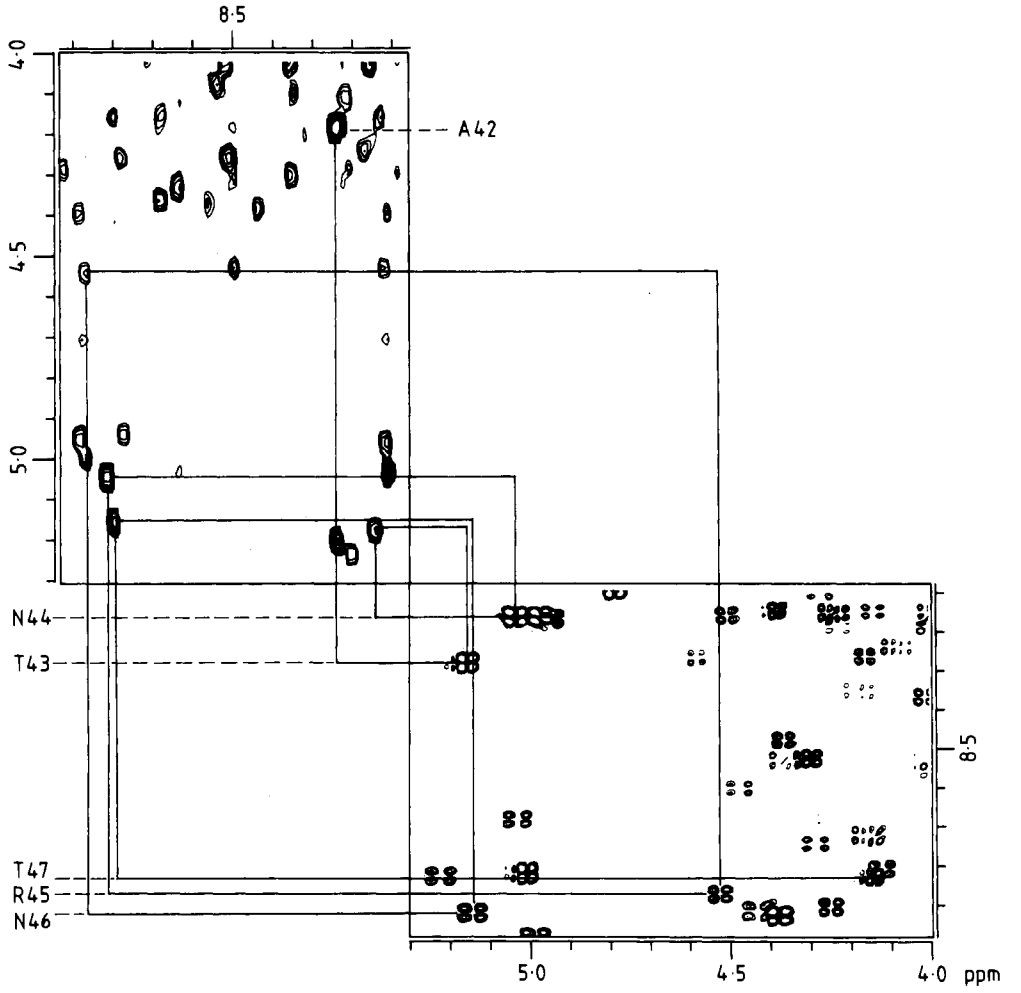

FIGURE 10: Phase-sensitive 500-MHz NOESY (upper left) and COSY (lower right) spectra of nonexchanged hen lysozyme. The sequential assignment of residues 42–47 via $d_{\alpha N}$ connectivities is shown.

C80, and T89. Unassigned cross-peaks from two glycines and five other residues appear in the fingerprint region of the COSY spectrum.

Structural Correlations. In addition to sequential assignment information, NOE effects involving NH, αCH, and βCH protons also provide information about the secondary structure within the protein. NOE connectivities involving pairs of neighboring NH protons are characteristic of helical and tight

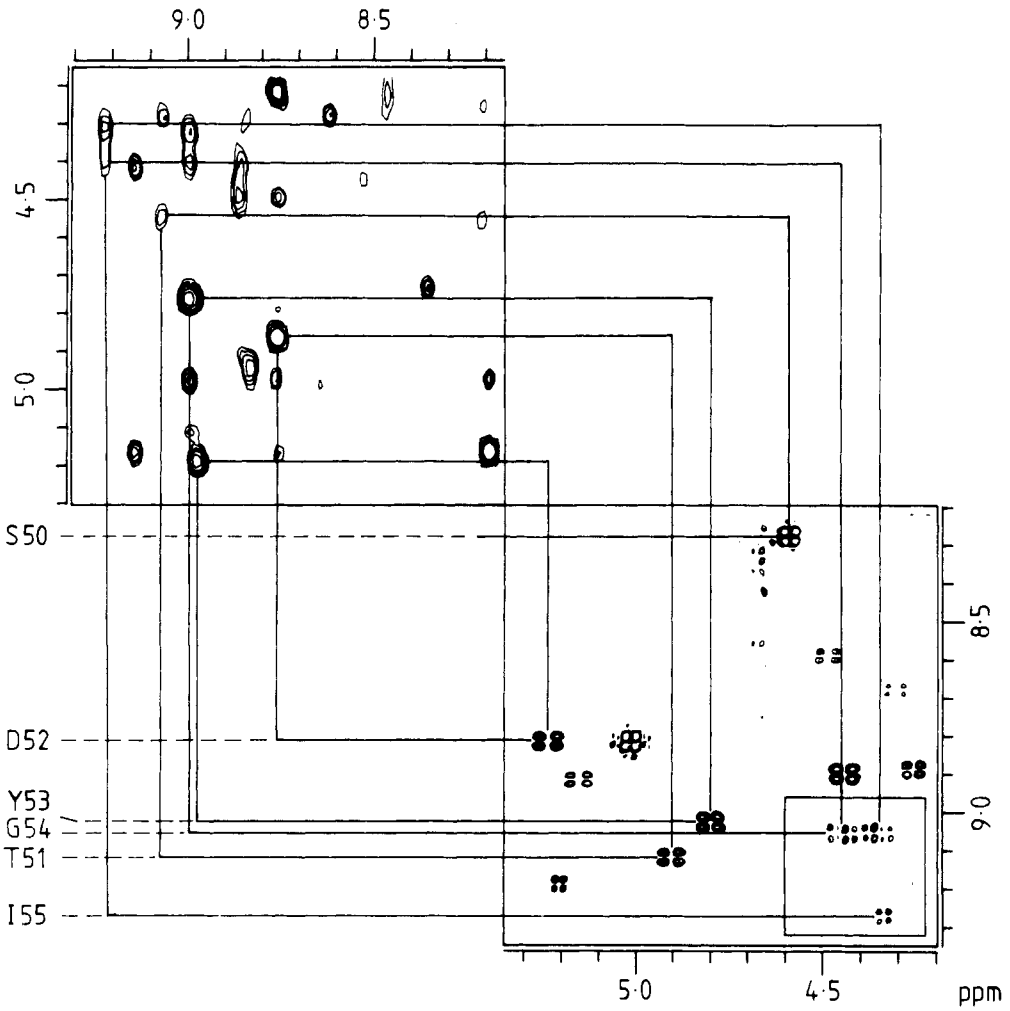

FIGURE 11: Phase-sensitive 500-MHz NOESY (upper left) and COSY (lower right) spectra of partly exchanged hen lysozyme. The sequential assignment of residues 50–55 via $d_{\alpha N}$ connectivities is shown.

turn structures whereas connectivities of the $d_{\alpha N}$ type are characteristic of β-strand structures (Wüthrich, 1986). Long-range NOE effects can supplement the short-range information and can give information about the interaction of β-strands in β-sheet structures.

Extended regions of d_{NN} sequential connectivities were observed for residues 6–15, 25–38, and 89–101 as shown in Figure 7. These regions were identified as helices with a high degree of confidence on the basis of the number of consecutive d_{NN} connectivities. Assignment of helical secondary structure was confirmed by the observation of $d_{\alpha N}(i,i+3)$ connectivities. These connectivities have been illustrated for residues 6–15 in Figure 9. Virtually all of the expected $d_{\alpha N}(i,i+3)$ connectivities were also found for residues 89–101. A smaller number of these connectivities were observed for residues 25–38; many of the NH resonances associated with these residues overlap with aromatic proton resonances, making unambiguous assignment of NOESY cross-peaks somewhat more difficult. Shorter stretches of d_{NN} connectivities were observed for residues 74–78, 80–84, and 107–115. However, no long-range $d_{\alpha N}(i,i+3)$ connectivities have yet been identified for these regions, and therefore, assignment of these regions as helical on this basis alone would be premature.

Extended regions of $d_{\alpha N}$ sequential connectivities were observed for residues 1–5, 42–47, 50–55, and 57–60. The last three of these regions are connected by short segments of d_{NN} connectivities characteristic of tight turns. This pattern of NOE connectivities indicates the presence of a triple-stranded antiparallel β-sheet. This structure was confirmed by several

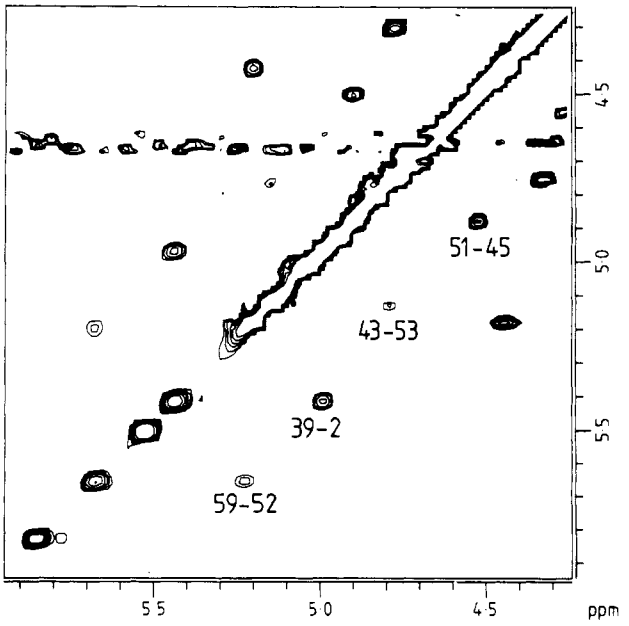

FIGURE 12: Phase-sensitive 500-MHz NOESY spectrum of fully exchanged hen lysozyme. Long-range $d_{\alpha\alpha}$ connectivities characteristic of β-sheet structures are labeled.

long-range interstrand d_{NN}, $d_{\alpha N}$, and $d_{\alpha\alpha}$ connectivities observed in the NOE spectra shown in Figures 5, 6, and 12. Four interstrand NOE effects, $d_{NN}(44,52)$, $d_{\alpha\alpha}(45,51)$, $d_{\alpha\alpha}$-(43,53), and $d_{\alpha N}(45,52)$, confirm the interaction between

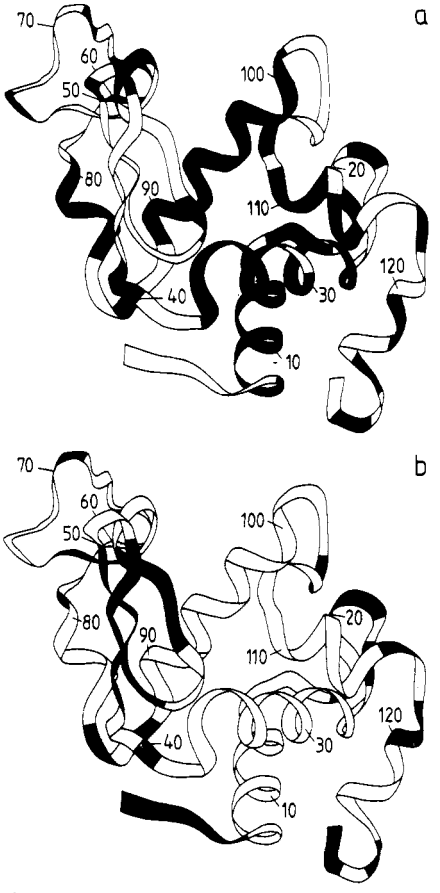

FIGURE 13: Schematic ribbon representation of the backbone of hen lysozyme derived from the X-ray structure. Regions characterized by d_{NN} NOE connectivities are shaded in (a) whereas regions characteristized by $d_{\alpha N}$ connectivities are shaded in (b).

strands 42–47 and 50–55. Three interstrand NOE effects, $d_{NN}(53,58)$, $d_{\alpha\alpha}(52,59)$, and $d_{\alpha N}(57,53)$, confirm the interaction between strands 50–55 and 57–60. Two long-range interactions between residue 2 and residues 39 and 40, $d_{\alpha N}$-(2,40) and $d_{\alpha\alpha}(2,39)$, have been identified from NOESY spectra. These connectivities indicate that residue 2 and residues 39–40 form two strands of a short antiparallel β-sheet. No further long-range connectivities involving residues 2 or 39-40 have been identified, thus giving no evidence for a third strand in this β-sheet.

The NOE data described above provide a good basis for the identification of the major components of secondary structure in lysozyme. Further analysis of the NOESY spectrum will yield additional distance constraints, which can be used as the input for structure determination algorithms such as distance geometry and restrained molecular dynamics (Braun et al., 1983; Kaptein et al., 1985). The secondary structure determined from NOE connectivities can, however, be compared immediately with the secondary structure determined from X-ray diffraction studies (Blake et al., 1967). The backbone structure of hen lysozyme in the crystalline state is illustrated schematically in Figure 13. In Figure 13a regions of the sequence which give d_{NN} NOE connectivities are shaded; in Figure 13b regions which give $d_{\alpha N}$ connectivities are shaded. It can be seen that regions which were identified earlier as helical from d_{NN} connectivities correspond exactly with the major helices in the X-ray structure. The triple-stranded β-sheet identified earlier agrees with the β-sheet identified in the X-ray structure; the β-strands are characterized by $d_{\alpha N}$ connectivities whereas the tight turns are characterized by d_{NN} connectivities. The connectivities identified between residues 2 and 39–40 are in agreement with the small antiparallel β-sheet involving residues 1–3 and 38–40.

The NMR experiments carried out with partly exchanged and reverse-exchanged lysozyme allow a qualitative comparison of hydrogen exchange behavior at pH 3.8 and 35 °C to be made. Amides that give cross-peaks of greater intensity in the partly exchanged COSY than in the reverse-exchanged COSY have slow hydrogen rates, $t_{1/2} \geq 1$–5 h. Amides that give cross-peaks of greater intensity in the reverse-exchanged COSY have a rapid exchange rate, $t_{1/2} \leq 1$–5 h, whereas amides that give peaks of equal intensity in the two spectra have intermediate exchange rates, $t_{1/2} \sim 1$–5 h. The observed hydrogen exchange behavior is summarized in Figure 7. These exchange rates correlate well with the secondary structure of the protein; the majority of slowly exchanging amides are found in the helices and the β-sheet region. The qualitative exchange rates observed here can be compared with hydrogen bonds found in the hen lysozyme crystal structure (Handoll, 1985). Forty-nine of the 58 slowly exchanging amides are hydrogen bonded to main-chain carbonyl oxygens. Three of the nine remaining amides, S50, R61, and I78, are hydrogen bonded to side-chain oxygens. Of the rest, the amides of I55, L56, and N59 are hydrogen bonded to buried water molecules whereas those of F38, A42, and C64 are not specified as being involved in any hydrogen bonds. Thirty-seven of the 57 rapidly exchanging amides are either not internally hydrogen bonded or are hydrogen bonded to water molecules. Five of the 20 remaining amides are hydrogen bonded to side-chain oxygen atoms. The 15 remaining amides are hydrogen bonded to main-chain carbonyl oxygens. The amides of R14, N113, and R114 are located in helical structures whereas the amides of N46 and G49 are located in the β-sheet region. The amides of the other 10 residues, G16, N19, Y20, G22, S72, N77, D101, T118, V120, and A122, are located in irregular structures (Handoll, 1985). A slow amide exchange rate is often taken as an indication of a hydrogen bond and used as a distance constraint in structure determination techniques. If the hydrogen bonds found in the X-ray structure are maintained in solution, then the preliminary results presented here indicate that some care must be taken when interpreting slow hydrogen exchange rates in terms of hydrogen bonds.

Assignments for 121 of the 129 residues of lysozyme are listed in Table I. In previous studies assignments for resonances of some 50 amino acid residues had been made with information derived from the crystal structure; this information included ring current shift data and interproton distances. The sequential assignments presented here are in complete agreement with previously reported assignments (Poulsen et al., 1980; Redfield et al., 1982; Delepierre et al., 1984). This provides further evidence for the close similarity of the X-ray and solution tertiary structures. In view of the agreement between the assignments made by these two procedures and of the similarity of the secondary structure derived from these NMR data with that determined by X-ray diffraction, it is possible to propose assignments for some of the eight remaining residues. The αCH resonances of both W63 and M105 have been assigned with reference to the crystal structure (Poulsen et al., 1980; Delepierre et al., 1984). There are two unassigned cross-peaks in the fingerprint region which have αCH positions that correspond to those of W63 and M105. A $d_{\alpha N}$ connectivity between a glycine residue and a type X residue was described earlier. The assignment of the type X residue as M105 results in the assignment of the glycine as G104; the last glycine, G102, is assigned by elimination. No connectivities between G102, G104, M105, or A107 and the unas-

Table I: Sequence-Specific Assignments of Hen Egg White Lysozyme[a]

residue	NH	αCH	βCH	γCH and others	residue	NH	αCH	βCH	γCH and others
Lys-1		3.99			Gly-67	8.34	4.16, 3.86		
Val-2	8.96	4.97	2.00	γCH_3 0.95, 1.05	Arg-68				
Phe-3	8.87	4.23	2.73, 3.21	δCH 7.03; εCH 7.22; ζCH 7.51	Thr-69	8.20	4.60	4.22	γCH_3 0.92
Gly-4	8.51	4.34, 4.01			Pro-70		4.35		
Arg-5	8.56	3.38	2.15		Gly-71	8.69	3.83, 3.68		
Cys-6	8.60	4.74	2.84, 3.23		Ser-72	7.30	4.56	3.75, 4.24	
Glu-7	8.14	4.12	2.20, 2.30		Arg-73				
Leu-8	8.63	3.77	0.97, 1.58	γCH 1.49; δCH_3 −0.05, 0.55	Asn-74	8.18	3.75	2.01, 3.12	
Ala-9	8.40	3.61	1.59		Leu-75	9.03	4.10	1.52, 2.15	γCH 1.68; δCH_3 0.55, 0.88
Ala-10	8.17	3.99	1.58		Cys-76	9.46	4.46	3.69	
Ala-11	7.79	4.30	1.56		Asn-77	8.04	4.22	2.54, 3.21	
Met-12	9.10	3.47		ϵCH_3 1.66	Ile-78	8.79	4.99	1.72	γCH 1.08, 1.51; γCH_3 0.91; δCH_3 0.89
Lys-13	8.54	3.97	1.97, 2.24						
Arg-14	8.25	4.14	1.79, 1.87		Pro-79		5.21	2.06, 2.39	
His-15	7.32	4.60	2.61, 3.77	δCH 8.66; εCH 7.53	Cys-80	8.24			
Gly-16	7.62	4.13, 3.92			Ser-81	8.58	3.80		
Leu-17	7.15	3.94	0.33	γCH 0.74; δCH_3 −0.66, −0.09	Ala-82	7.61	4.25	1.53	
Asp-18	8.71	4.27	2.42, 3.04		Leu-83	7.85	4.25	2.11	γCH 1.71; δCH_3 0.75, 1.05
Asn-19	8.35	3.99	2.84, 3.07		Leu-84	7.13	5.11	1.90	γCH 1.74; δCH_3 1.01, 1.04
Tyr-20	8.08	4.25	3.06, 3.25	δCH 7.22; εCH 6.99	Ser-85	6.83	4.49	3.89, 4.17	
Arg-21	8.93	3.63	1.82		Ser-86	8.50	4.27	3.89, 3.95	
Gly-22	7.60	3.89, 3.50			Asp-87	8.16	4.92	2.68, 2.96	
Tyr-23	7.66	4.57	2.53, 3.32	δCH 7.05; εCH 6.71	Ile-88	8.11	4.67	1.87	γCH 0.40, 1.21; γCH_3 0.80; δCH_3 0.23
Ser-24	8.98	4.54							
Leu-25	9.09	4.42		γCH 1.61; δCH_3 0.87, 1.00	Thr-89	8.38	3.07	4.07	γCH_3 1.18
Gly-26	9.59	4.20, 3.68			Ala-90	9.11	4.09	1.38	
Asn-27	8.19	4.22	2.35, 2.87		Ser-91	7.79	4.09	3.51, 4.02	
Trp-28	7.19	3.76	3.24, 3.30	N1H 9.36; C2H 7.31; C4H 6.77; C5H 6.29; C6H 6.82; C7H 7.46	Val-92	8.38	3.14	1.94	γCH_3 0.52, 0.65
					Asn-93	8.66	4.28	2.80, 2.92	
					Cys-94	7.91	5.00	2.72, 3.39	
Val-29	7.57	3.44	1.96	γCH_3 0.96, 1.28	Ala-95	8.70	4.14	1.56	
Cys-30	8.01	2.50	2.64, 2.98		Lys-96	8.01	3.72	1.71	
Ala-31	8.12	3.72	1.02		Lys-97	7.23	4.14		
Ala-32	7.60	4.02	1.33		Ile-98	7.98	2.81	1.51	γCH −2.04, 0.73; γCH_3 −0.24; δCH_3 0.01
Lys-33	7.95	2.58							
Phe-34	7.34	4.32	2.30, 3.20	δCH 7.26; εCH 7.37, ζCH 7.51	Val-99	8.26	3.90	2.43	γCH_3 1.21, 1.36
Glu-35	8.58	4.46	2.07		Ser-100	7.65	4.46	4.11, 4.20	
Ser-36	7.94	4.58	3.59, 4.51		Asp-101	8.03	4.81	3.06, 3.12	
Asn-37	8.14	4.49	2.52, 3.35		Gly-102	8.14	4.22, 3.97		
Phe-38	7.35	3.87	3.63, 3.76	δCH 6.99; εCH 7.45; ζCH 6.97	Asn-103				
Asn-39	7.41	5.41	2.85, 3.46		Gly-104	8.23	4.25, 4.06		
Thr-40	9.38	4.07	4.61	γCH_3 1.65	Met-105	7.06	3.84	−0.97, 0.48	γCH 0.55; ϵCH_3 0.00
Gln-41	7.90	4.45			Asn-106				
Ala-42	6.86	4.14	1.36		Ala-107	6.75	3.85	0.64	
Thr-43	8.27	5.14	3.79	γCH_3 1.06	Trp-108	7.88	4.71	3.24, 3.37	N1H 9.98; C2H 7.10; C4H 7.36; C5H 6.50; C6H 7.17; C7H 6.93
Asn-44	8.15	5.02	2.74						
Arg-45	8.85	4.50	1.72, 1.84						
Asn-46	8.89	5.13	2.86, 2.91		Val-109	8.91	3.65	2.20	γCH_3 1.04, 1.12
Thr-47	8.80	4.12	4.37	γCH_3 1.37	Ala-110	8.00	4.27	1.35	
Asp-48	7.81	4.58	2.67, 3.10		Trp-111	7.26	3.72		N1H 10.35; C2H 7.04; C4H 7.28; C5H 7.04; C6H 7.35; C7H 7.51
Gly-49	7.88	4.42, 3.74							
Ser-50	8.25	4.57	3.78						
Thr-51	9.12	4.88	3.76	γCH_3 0.34	Arg-112	8.23	3.39	2.05	
Asp-52	8.80	5.21	2.07, 2.71		Asn-113	7.97	4.54	2.69	
Tyr-53	9.02	4.76	2.69, 2.99	δCH 7.12; εCH 6.83	Arg-114	7.66	4.32	1.24	
Gly-54	9.04	4.43, 4.32			Cys-115	7.35	4.54	2.50, 2.64	
Ile-55	9.27	4.31	1.66	γCH 1.06, 1.54; γCH_3 0.91; δCH_3 0.77	Lys-116	7.06	3.46	−0.16, 1.28	
					Gly-117	8.70	4.14, 3.83		
Leu-56	8.89	4.41	1.48, 1.76	γCH 1.23; δCH_3 0.28, 0.52	Thr-118	7.65	4.76	4.33	γCH_3 0.98
Gln-57	7.95	3.37	2.03, 2.15		Asp-119	8.66	5.01	2.77, 2.97	
Ile-58	7.68	4.02	1.87	γCH 1.80; γCH_3 1.08; δCH_3 0.98	Val-120	8.13	4.36	2.20	γCH_3 1.09, 1.14
					Gln-121	8.46	4.34	2.22, 2.30	
Asn-59	8.52	5.65	3.10, 3.40		Ala-122	7.70	3.83	1.21	
Ser-60	9.18	5.17	4.44		Trp-123	7.61	4.12	3.45, 3.55	N1H 10.68; C2H 7.55; C4H 7.53; C5H 7.14; C6H 7.10; C7H 7.76
Arg-61	8.78	4.10	1.50, 1.73						
Trp-62	7.19	4.42		N1H 10.05; C2H 7.11; C4H 7.11; C5H 7.03; C6H 7.19; C7H 7.45	Ile-124	7.55	4.74	2.23	γCH 1.28, 1.49; γCH_3 0.87; δCH_3 0.98
Trp-63	7.43	4.98	3.37, 3.48	N1H 10.21; C2H 7.65; C4H 7.76; C5H 6.85; C6H 7.08; C7H 7.22	Arg-125	7.33	4.17	1.83, 1.94	
					Gly-126	9.12	4.32, 3.76		
					Cys-127	7.47	4.92	2.67, 3.10	
Cys-64	7.60	5.83	2.58, 3.09		Arg-128	8.90	4.36	1.80, 1.88	
Asn-65	8.27	5.51	2.47, 2.86		Leu-129	7.96	4.30	1.68	γCH 1.44; δCH_3 0.76, 0.88
Asp-66	9.65	4.99	2.25, 3.28						

[a] Chemical shifts are in ppm referenced to DSS and are accurate to ±0.02 ppm. Values are for lysozyme at 35 °C, pH 3.8.

signed residues N103 and N106 have been identified, and at this stage these two residues must remain unassigned. In addition, no sequential-based or crystal structure based assignments for R68 and R73 have been possible. There are three unassigned cross-peaks in the fingerprint region of the spectrum; the peak at 8.15–4.96 ppm is type J, the peak at 8.08–4.76 ppm is type U, and the peak at 7.55–3.87 ppm is type X.

Resonances arising from 121 of the 129 residues of lysozyme have been assigned by sequential methods. Breaks occur in the sequential assignment process at 20 points in the sequence as illustrated in Figure 7. Using the X-ray structure, it is possible to determine the types of connectivities that should be observed at these points and then to explain why these breaks occur. Two of the breaks occur between a proline and its preceding residue. The δCH resonances of the two proline residues have not been identified in COSY or RELAY spectra, and therefore, $d_{\alpha P\delta}$ NOE connectivities could not be identified unambiguously. Two of the breaks occur because of the absence of $d_{\alpha N}$ connectivities. Both of the αCH resonances involved are close to the water resonance, and the connectivities are absent because the αCH resonances are saturated when the water is irradiated. The 16 remaining breaks occur because of the absence of d_{NN} connectivities. Seven of these connectivities are missing because the two NH resonances involved have very similar ($\Delta\delta \leq 0.1$ ppm) chemical shift values. Five of these connectivities are missing because the two NH resonances involved have chemical shift values in the 7.0–7.8 ppm range. The NH–NH cross-peaks overlap with the intense aromatic and asparagine and glutamine side-chain NH cross-peaks and cannot be identified unambiguously.

The assignments presented in this paper provide the starting point for a variety of important studies. Further analysis of the NOESY spectrum will yield the required input for structure determination procedures. Assignments for all the active site residues, including the catalytically important E35 and D52, which were not previously available, will allow studies of substrate and inhibitor interactions to be carried out. The detailed knowledge of the spectrum of the folded protein can be used to characterize the structure and dynamics of the protein in its unfolded state. Finally, the results presented here indicate that detailed information can be obtained for proteins that are relatively large in NMR terms without using information obtained from the X-ray structure.

ACKNOWLEDGMENTS

We thank N. Soffe and J. Boyd for their assistance. C.R. acknowledges Lady Margaret Hall, University of Oxford, for an E. P. A. Cephalosporin Junior Research Fellowship.

REFERENCES

Anil Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Arseniev, A. S., Wider, G., Joubert, F. J., & Wüthrich, K. (1982) *J. Mol. Biol. 159*, 323–351.

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542–561.

Bax, A., & Drobny, G. P. (1985) *J. Magn. Reson. 61*, 306–320.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

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

Blake, C. C. F., Cassels, R., Dobson, C. M., Poulsen, F. M., Williams, R. J. P., & Wilson, K. S. (1981) *J. Mol. Biol. 147*, 73–95.

Boyd, J., Dobson, C. M., & Redfield, C. (1983) *J. Magn. Reson. 55*, 170–176.

Boyd, J., Dobson, C. M., & Redfield, C. (1985a) *J. Magn. Reson. 62*, 543–550.

Boyd, J., Dobson, C. M., & Redfield, C. (1985b) *Eur. J. Biochem. 153*, 383–396.

Braun, W., Wider, G., Lee, K. H., & Wüthrich, K. (1983) *J. Mol. Biol. 169*, 921–948.

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

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

Campbell, I. D., Dobson, C. M., & Williams, R. J. P. (1975b) *Proc. R. Soc. London, B 189*, 503–509.

Cassels, R., Dobson, C. M., Poulsen, F. M., & Williams, R. J. P. (1978) *Eur. J. Biochem. 92*, 81–98.

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

Clore, G. M., Sukumaran, D. K., Nilges, M., & Gronenborn, A. M. (1987) *Biochemistry 26*, 1732–1745.

Cohen, J. S., & Jardetsky, O. (1968) *Proc. Natl. Acad. Sci. U.S.A. 60*, 92–99.

Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H., & Sheard, B. (1987) *Nature (London) 327*, 339–341.

Delepierre, M., Dobson, C. M., & Poulsen, F. M. (1982) *Biochemistry 21*, 4756–4761.

Delepierre, M., Dobson, C. M., Selverajah, S., Wedin, R. E., & Poulsen, F. M. (1983) *J. Mol. Biol. 168*, 687–692.

Delepierre, M., Dobson, C. M., Howarth, M. A., & Poulsen, F. M. (1984) *Eur. J. Biochem. 145*, 389–395.

Dobson, C. M., & Evans, P. E. (1984) *Biochemistry 23*, 4267–4270.

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

Dobson, C. M., Evans, P. A., & Williamson, K. L. (1984) *FEBS Lett. 168*, 331–334.

Driscoll, P. C., Hill, H. A. O., & Redfield, C. (1987) *Eur. J. Biochem.* (in press).

Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731–3732.

Handoll, H. H. G. (1985) D. Phil. Thesis, Oxford University.

Holak, T. A., & Prestegard, J. H. (1986) *Biochemistry 25*, 5766–5774.

Hore, P. J., & Kaptein, R. (1983) *Biochemistry 22*, 1906–1911.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179–182.

Keller, R. M., Baumann, R., Hunziker-Kwik, E. H., Joubert, F. J., & Wüthrich, K. (1983) *J. Mol. Biol. 163*, 623–646.

Klevit, R. E., Drobny, G. P., & Waygood, E. B. (1986) *Biochemistry 25*, 7760–7769.

Kline, A. D., & Wüthrich, K. (1986) *J. Mol. Biol. 192*, 869–880.

Lenkinski, R. E., Dallas, J. L., & Glickson, J. D. (1979) *J. Am. Chem. Soc. 101*, 3071–3077.

Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.

McDonald, C. C., & Phillips, W. D. (1967) *J. Am. Chem. Soc. 89*, 6332–6341.

McDonald, C. C., & Phillips, W. D. (1970) in *Fine Structure of Proteins and Nucleic Acids* (Fasman, G. D., & Timasheff, S. N., Eds.) Dekker, New York.
Meadows, D. H., Markley, J. L., Cohen, J. S., & Jardetsky, O. (1967) *Proc. Natl. Acad. Sci. U.S.A. 58*, 1307-1313.
Neuhaus, D., Wagner, G., Vašák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257-273.
Olejniczak, E. T., Poulsen, F. M., & Dobson, C. M. (1981) *J. Am. Chem. Soc. 103*, 6574-6580.
Poulsen, F. M., Hoch, J. C., & Dobson, C. M. (1980) *Biochemistry 19*, 2597-2607.
Redfield, C., Poulsen, F. M., & Dobson, C. M. (1982) *Eur. J. Biochem. 128*, 527-531.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.
Sternlicht, H., & Wilson, D. (1967) *Biochemistry 6*, 2881-2892.
Štrop, P., Wider, G., & Wüthrich, K. (1983) *J. Mol. Biol. 166*, 641-667.
Sukumaran, D. K., Clore, G. M., Preuss, A., Zarbock, J., & Gronenborn, A. M. (1987) *Biochemistry 26*, 333-338.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347-366.
Wagner, G., Neuhaus, D., Wörgötter, E., Vašák, M., Kägi, J. H. R., & Wüthrich, K. (1986) *Eur. J. Biochem. 157*, 275-289.
Wedin, R. E., Delepierre, M., Dobson, C. M., & Poulsen, F. M. (1982) *Biochemistry 21*, 1098-1103.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715-740.
Zuiderweg, E. R. P., Kaptein, R., & Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 279-292.

Flavodoxin from *Anabaena* 7120: Uniform Nitrogen-15 Enrichment and Hydrogen-1, Nitrogen-15, and Phosphorus-31 NMR Investigations of the Flavin Mononucleotide Binding Site in the Reduced and Oxidized States†

Brian J. Stockman, William M. Westler, Eddie S. Mooberry, and John L. Markley*
Department of Biochemistry, 420 Henry Mall, University of Wisconsin—Madison, Madison, Wisconsin 53706
Received July 7, 1987; Revised Manuscript Received September 4, 1987

ABSTRACT: Interactions between flavin mononucleotide (FMN) and apoprotein have been investigated in the reduced and oxidized states of the flavodoxin isolated from *Anabaena* 7120 (M_r ~21 000). ^1H, ^{15}N, and ^{31}P NMR have been used to characterize the FMN–protein interactions in both redox states. These are compared with those seen in other flavodoxins. Uniformly enriched [^{15}N]flavodoxin (>95% isotopic purity) was isolated from *Anabaena* 7120 grown on $K^{15}NO_3$ as the sole nitrogen source. ^{15}N insensitive nucleus enhanced by polarization transfer (INEPT) and nuclear Overhauser effect (NOE) studies of this sample provided information regarding protein structure and dynamics. A ^1H-detected ^{15}N experiment allowed the correlation of nitrogen resonances to those of their attached protons. Over 90% of the expected N–H cross peaks could be resolved in this experiment.

Flavodoxins constitute a group of low molecular weight (M_r 14 000–23 000), FMN[1]-containing flavoproteins that mediate electron transfer at low redox potential between the prosthetic groups of other proteins (Mayhew & Ludwig, 1975). In some organisms, flavodoxin is produced constitutively, while in others it is produced only under conditions of limiting iron. Flavodoxins serve as a replacement for the iron-containing protein, ferredoxin, in electron-transfer reactions (Tollin & Edmondson, 1980).

The FMN cofactor serves as the redox carrier in flavodoxins. The coenzyme can exist in three oxidation states, two of which have variable protonation states: oxidized (FMN), one electron reduced or semiquinone ($FMNH^\bullet$ or $FMN^{\bullet-}$), and two electron reduced or hydroquinone ($FMNH_2$ or $FMNH^-$). The ability for flavin-containing proteins to conduct one- or two-electron transfers permits them to mediate between one-electron- and two-electron-transfer pathways. All three redox states of flavodoxin exist in vitro, but redox reactions probably only occur between the reduced and semiquinone states in vivo (Simondson & Tollin, 1980).

The redox potentials of both transitions in FMN are altered by its association with the apoprotein. The transition from reduced to semiquinone states in flavodoxin has a midpoint potential of −400 to −500 mV, as compared to −124 mV in free FMN (Simondson & Tollin, 1980; Sykes & Rogers, 1984). Coenzyme in different redox states appears to interact differently with the apoprotein. The redox potential for the

† Supported by USDA Competitive Research Grant 85-CRCR-1-1589. This study made use of the National Magnetic Resonance Facility at Madison, which is supported in part by NIH Grant RR023021 from the Biomedical Research Technology Program, Division of Research Resources. Equipment in the facility was purchased with funds from the University of Wisconsin, the NSF Biological Biomedical Research Technology Program (Grant RR023021), NIH Shared Instrumentation Program (Grant RR02781), and the U.S. Department of Agriculture. A preliminary account of this research has been presented (Stockman & Markley, 1987).
* Author to whom correspondence should be addressed.

[1] Abbreviations: BIRD, bilinear rotation decoupling; COSY, correlated spectroscopy; FMN, flavin mononucleotide; INEPT, insensitive nucleus enhanced by polarization transfer; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser spectroscopy; SDS, sodium dodecyl sulfate; Tris, tris(hydroxymethyl)aminomethane; TSP, (trimethylsilyl)propionic acid.

0006-2960/88/0427-0136$01.50/0