Markley, J. L., Putter, I., & Jardetzky, O. (1968) *Science 161*, 1249.
Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321.
Paluh, J. L., & Yanofsky, C. (1986) *Nucleic Acids Res. 14*, 7851.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310.
Schevitz, R., Otwinowski, Z., Joachimiak, A., Lawson, C. L., & Sigler, P. B. (1985) *Nature 317*, 782.
Szilágyi, L., & Jardetzky, O. (1989) *J. Magn. Reson. 83*, 441.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *Mol. Biol. 169*, 949.
Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., & Sigler, P. B. (1987) *Nature 327*, 591.

Articles

Assignment of the Backbone ^1H and ^{15}N NMR Resonances of Bacteriophage T4 Lysozyme†

Lawrence P. McIntosh,*,‡ A. Joshua Wand,§ David F. Lowry,∥ Alfred G. Redfield,∥ and Frederick W. Dahlquist*,‡

Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, Oregon 97403, Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, and Departments of Biochemistry and Physics, Brandeis University, Waltham, Massachusetts 02254

Received January 22, 1990; Revised Manuscript Received March 23, 1990

ABSTRACT: The proton and nitrogen (^{15}NH–H^{α}–H^{β}) resonances of bacteriophage T4 lysozyme were assigned by ^{15}N-aided ^1H NMR. The assignments were directed from the backbone amide ^1H–^{15}N nuclei, with the heteronuclear single-multiple-quantum coherence (HSMQC) spectrum of uniformly ^{15}N enriched protein serving as the master template for this work. The main-chain amide ^1H–^{15}N resonances and H^{α} resonances were resolved and classified into 18 amino acid types by using HMQC and ^{15}N-edited COSY measurements, respectively, of T4 lysozymes selectively enriched with one or more of α-^{15}N-labeled Ala, Arg, Asn, Asp, Gly, Gln, Glu, Ile, Leu, Lys, Met, Phe, Ser, Thr, Trp, Tyr, or Val. The heteronuclear spectra were complemented by proton DQF-COSY and TOCSY spectra of unlabeled protein in H_2O and D_2O buffers, from which the H^{β} resonances of many residues were identified. The NOE cross peaks to almost every amide proton were resolved in ^{15}N-edited NOESY spectra of the selectively ^{15}N enriched protein samples. Residue specific assignments were determined by using NOE connectivities between protons in the ^{15}NH–H^{α}–H^{β} spin systems of known amino acid type. Additional assignments of the aromatic proton resonances were obtained from ^1H NMR spectra of unlabeled and selectively deuterated protein samples. The secondary structure of T4 lysozyme indicated from a qualitative analysis of the NOESY data is consistent with the crystallographic model of the protein.

T4 lysozyme is an endomuramidase required for the lytic growth of bacteriophage T4 (Tsugita & Ikeya-Ocada, 1972). From the time of the pioneering work of Streisinger and colleagues, this enzyme has been the subject of extensive genetic, thermodynamic, spectroscopic, and crystallographic studies and stands today as one of the best characterized systems for investigating the basis of protein structure and stability (Streisinger et al., 1961; Alber et al., 1987; Becktel & Baase, 1987; Hudson et al., 1987; Matthews, 1987; Weaver & Matthews, 1987). In earlier work, we have used NMR[1] to study selected regions of wild-type and variant T4 lysozymes, focusing on questions including the kinetics of hydrogen exchange and the pH dependence of the electrostatic stability of these proteins (Griffey et al., 1985a; Weaver et al., 1989; Anderson et al., 1990). However, more extensive investigations

† This work was supported by NSF Grant DMB8905322 and a grant from the Lucille B. Markey Chartible Trust to F.W.D., by NIH Grant GM20168 to A.G.R., by NIH Grant GM35940 to A.J.W., and by NIH Grants CA06927 and RR05539, an award from the Fanny Ripple Foundation, a grant from the Pew Trust, and an appropriation from the Commonwealth of Pennsylvania awarded to the Institute for Cancer Research. L.P.M. is a recipient of a Natural Sciences and Engineering Council of Canada 1967 Scholarship and an Alberta Heritage Foundation Studentship.

* Authors to whom requests for reprints should be addressed.
‡ University of Oregon.
§ Fox Chance Cancer Center.
∥ Brandeis University.

[1] Abbreviations: The amino acids are denoted by standard one- and three-letter codes, Asx is both aspartate and asparagine, and Glx is both glutamate and glutamine. Atoms in amino acids are named according to the IUPAC–IUB convention [(1970) *J. Mol. Biol. 52*, 1–17]; COSY, two-dimensional *J*-correlated spectroscopy; DQF, double-quantum-filtered; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulfonate; HMBC, heteronuclear multiple-bond coherence spectroscopy; HMQC, heteronuclear multiple-quantum coherence spectroscopy; HSMQC, heteronuclear single-multiple-quantum coherence spectroscopy; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE-correlated spectroscopy; ppm, parts per million; pH*, observed pH meter reading without corrections for the deuterium isotope effect; TPPI, time proportional phase incrementation; TOCSY, two-dimensional total correlation spectroscopy.

0006-2960/90/0429-6341$02.50/0

of the solution properties of T4 lysozyme required that the major task of deciphering the complex NMR spectra of the protein be accomplished. In this communication, we describe the complete assignment of the backbone ^1H and ^{15}N resonances of T4 lysozyme.

The difficulty of interpreting the NMR spectra of a protein increases dramatically with its size, as the increased number of protons and the broader widths of their signals lead to highly complex two-dimensional spectra with many degenerate and weak or missing cross peaks (Oppenheimer & James, 1989). Although the NMR spectra of proteins as large as hen egg white lysozyme (Redfield & Dobson, 1988) and ribonuclease A (Robertson et al., 1989) have been extensively assigned by use of two-dimensional ^1H NMR experiments, T4 lysozyme (164 residues, 18.7 kDa) is above the current practical size limit for such methodology. To interpret the NMR spectra of this protein, we followed an approach that embodied two features: (i) The assignment protocol focused on the H^N, H^α, and H^β proton resonances. The amide cross-peak regions of the two-dimensional ^1H NMR spectra of T4 lysozyme are simpler and more disperse than the complex side-chain cross-peak regions, and the NOEs between the H^N, H^α, and H^β protons can be used to specifically assign neighboring residues (Wüthrich, 1983, 1986; Englander & Wand, 1987; Chazin & Wright, 1987). (ii) The NMR experiments used for this approach resolved degeneracies due to overlapping resonances and permitted the unambiguous assignments of cross peaks in two-dimensional spectra to specific amide protons of known amino acid type.

The strategy that we adopted was to uniformly or selectively incorporate ^{15}N (and ^{13}C) nuclei into the backbone of T4 lysozyme. The resonances from labeled amide protons were resolved by ^1H and ^{15}N shifts in HMQC or HSMQC spectra, and the proton–proton scalar and dipolar cross peaks to the labeled protons were observed in ^{15}N-edited COSY and NOESY spectra and, to a lesser extent, in HMBC spectra. The H^N and H^α resonances were classified by amino acid type through extensive selective isotopic labeling, facilitating further analysis of the residue spin systems by COSY-type measurements and the assignment of resonances to specific residues by NOESY experiments. The backbone resonances of staphylococcal nuclease were recently assigned by a similar combination of isotope labeling and heteronuclear NMR experiments (Torchia et al., 1988a,b, 1989; Wang et al., 1990a,b).

The protocol leading to these assignments is summarized in seven steps: (i) We recorded the HSMQC spectrum of uniformly ^{15}N labeled T4 lysozyme. This was the master ^1H–^{15}N template from which the assignments were extended. (ii) The ^1H–^{15}N cross peaks were identified by amino acid type from HMQC spectra of T4 lysozyme selectively enriched with α-^{15}N-labeled Ala, Arg and Glx, Asp, Asp and Asn (Asx), Gly, Glu and Gln (Glx), Ile, Leu, Met, Phe, Ser and Gly, Thr, Trp, Tyr, or Val. (iii) The amide resonances of several residues were assigned by a variety of approaches, including multiple ^{13}C and ^{15}N labeling of T4 lysozyme (Griffey et al., 1986), amino acid substitutions in the protein, and saturation transfer HMQC experiments (Lowry et al., 1988). (iv) The H^α proton resonances of most residues were identified from ^{15}N-edited COSY spectra recorded for the selectively labeled lysozyme samples. (v) DQF-COSY and TOCSY spectra of unlabeled T4 lysozyme samples in H_2O and D_2O and HMBC spectra of uniformly ^{15}N enriched protein in D_2O were acquired to supplement the data obtained from the isotopically labeled proteins. The H^β proton resonances of many residues were identified from these spectra. (vi) The NOE interactions to each amide were resolved by ^{15}N-edited NOESY measurements of selectively labeled lysozyme samples. (vii) The residues of known amino acid type were specifically assigned on the basis of nearest-neighbor NOE connectivities.

Following this strategy, we have completely assigned the 160 (non-proline) backbone amide ^1H–^{15}N resonances of T4 lysozyme as well as the H^α and at least one H^β resonance of 157 and 121 residues, respectively, in this protein. A limited number of additional side-chain resonances of phage T4 lysozyme, particularly those from the alanine, valine, threonine, and tryptophan residues, were also identified by two-dimensional ^1H NMR techniques. The resonances from the phenylalanine and tyrosine side chains were delineated by using samples of protein with selectively deuterated aromatic rings. We have presented aspects of our work on T4 lysozyme previously (Griffey et al., 1985a, 1986; Dahlquist et al., 1985; McIntosh et al., 1987a,b; Lowry et al., 1988; Redfield et al., 1989; Muchmore et al., 1989; McIntosh & Dahlquist, 1990). In this publication, the full assignments of the backbone proton and nitrogen-15 NMR resonances of wild-type T4 lysozyme as well as partial assignments of the side-chain proton resonances are reported.

Materials and Methods

T4 Lysozyme Preparation. The genes encoding for wild-type and numerous variant T4 lysozymes are cloned into an inducible plasmid system for high levels of expression in a variety of *Escherichia coli* strains (Muchmore et al., 1989). With this expression system, 10–100-mg quantities of purified protein per liter of medium can be rapidly prepared (Griffey et al., 1985a; Muchmore et al., 1989). T4 lysozyme without any isotopic enrichment was obtained from *E. coli* grown in LBH media. T4 lysozyme with uniform ^{15}N enrichment was produced from a prototrophic *E. coli* strain grown in M9T media containing 0.75–1.0 g/L of $^{15}NH_4Cl$ or $(^{15}NH_4)_2SO_4$ as the sole nitrogen source (McIntosh et al., 1987; Muchmore et al., 1989). The ^{15}N isotopic enrichment was approximately 85% as judged by the relative intensities of the resonances from tryptophan indole $H^{\epsilon 1}$ protons directly coupled to ^{15}N and ^{14}N nuclei. T4 lysozymes selectively ^{15}N or ^{13}C enriched at one or more types of amino acids were prepared by using the appropriate auxotrophic *E. coli* strains, such as DL39 (LeMaster & Richards, 1988), grown in synthetic rich media containing one or more isotopically labeled amino acids (Griffey et al., 1985a; Muchmore et al., 1989). The strategies and procedures for producing ^{15}N isotopically enriched T4 lysozyme have been discussed elsewhere (Muchmore et al., 1989; McIntosh & Dahlquist, 1990) and a summary of this work is given in Table I. Similarly, T4 lysozymes containing aromatic amino acids with deuterated rings were prepared from *E. coli* DL39 grown in the synthetic rich media containing $[\delta_{1,2},\epsilon_{1,2},\zeta\text{-}^2H_5]$phenylalanine (130 mg/L) and $[\delta_1,\epsilon_2,\zeta_{2,3},\eta_2\text{-}^2H_5]$tryptophan (100 mg/L), or $[\epsilon_{1,2}\text{-}^2H_2]$tyrosine (200 mg/L) and $[\delta_1,\epsilon_2,\zeta_{2,3},\eta_2\text{-}^2H_5]$tryptophan (100 mg/L).

$^{15}NH_4Cl$ was obtained from Isotec and MSD Isotopes and $(^{15}NH_4)_2SO_4$ was bought from Monsanto Research Corp. ^{15}N- and ^{13}C-labeled amino acids were purchased from Cambridge Isotopes, ICON, MSD Isotopes and Tracer Technology. $[\delta_{1,2},\epsilon_{1,2},\zeta\text{-}^2H_5]$Phenylalanine was purchased form MSD isotopes. $[\epsilon_{1,2}\text{-}^2H_2]$Tyrosine (>95% isotopic purity) and $[\delta_1,\epsilon_2,\zeta_{2,3},\eta_2\text{-}^2H_5]$tryptophan (>85% isotopic purity) were synthesized as described by Kinsey et al. (1981).

NMR Samples. T4 lysozyme samples were approximately 1–4 mM protein in 95% H_2O/5% D_2O or in 99% D_2O solutions containing 100 mM potassium chloride, 30 mM potas-

Table I: ^{15}N Labeling of T4 Lysozyme

residue(s) α-^{15}N labeled	[α-^{15}N]amino acid[a]	amino acid/liter[a]	host genotype[b]
uniform	$^{15}NH_4Cl$, $(^{15}NH_4)_2SO_4$	0.75–1.0 g	prototrophic
Ala	D,L-alanine	600 mg	*asp*C,*avt*A,*ilv*E,*tyr*B
Arg, Gln, Glu[c]	L-glutamate	380 mg	*asp*C,*avt*A,*ilv*E,*tyr*B
Asp[d]	L-aspartate	100 mg	*asn*A,*asn*B,*asp*C,*tyr*B
Asn, Asp[e]	L-aspartate	200 mg	*asp*C,*avt*A,*ilv*E,*tyr*B
Gln, Glu[f]	L-glutamate	550 mg	*asp*C,*avt*A,*ilv*E,*tyr*B
Gly	glycine	250–375 mg	*gly*A
Ile	L-isoleucine	50 mg	*asp*C,*avt*A,*ilv*E,*tyr*B
Leu	L-leucine	60 mg	*asp*C,*ilv*E,*tyr*B
Lys	L-lysine	160 mg	*lys*A
Met	L-methionine	150 mg	*met*C
Phe	L-phenylalanine	50–80 mg	*asp*C,*avt*A,*ilv*E,*tyr*B *ilv*C,*phe*A[g,k] *phe*A[h,k]
Ser, Gly[i]	glycine	500 mg	*ser*A
Thr	L-threonine	150 mg	*thr*
Trp	L-tryptophan	35 mg	*trp*
Tyr	L-tyrosine	150 mg	*phe*A,*ilv*E[j,k]
Val	L-valine	60–100 mg	*ilv*C[k]

[a] The quantity of labeled acid (<98% enrichment) per liter of medium given in Muchmore et al., (1989). The media contained all other amino acids in unlabeled form, except for those to be derived from the supplied labeled material by the bacterial metabolic pathways. [b] The genotype of the bacterial strains related to amino acid metabolism (Bachmann, 1983). The strain DL39 (*asp*C,*ilv*E,*tyr*B) was a gift from D. LeMaster (LeMaster & Richards, 1988) and was modified by P1 transduction to introduce the *avt*A::Tn5 lesion (Muchmore et al., 1989). [c] No glutamine or arginine was supplied in the medium. The arginine residues were labeled in the α and ϵ positions. Also, the alanine residues were approximately 20% enriched, the lysine residues ca. 5%. [d] The conversion of aspartate to asparagine was prevented by the lesions in the asparagine synthetase genes. The threonine residues were enriched ca. 20% even though the media contained unlabeled threonine. [e] Addition of unenriched asparagine to the medium did not prevent the metabolic conversion of the [α-^{15}N]aspartate to [α-^{15}N]-asparagine. The threonine residues were enriched ca. 20% and the lysine residues ca. 5%. By omitting arginine from the medium, the arginine residues were ^{15}N enriched in the side-chain η positions. [f] No glutamine was present in the medium. The alanine residues were enriched ca. 20% and the lysine residues ca. 5%. [g] The media also contained 75 mg/L [1-^{13}C]valine or 150 mg/L [1,2-$^{13}C_2$]glycine to produce [^{15}N]Phe- and [^{13}C]Val- or -Gly-labeled T4 lysozyme, respectively. [h] The medium contained 80 mg/L [1-^{13}C]leucine to produce [^{15}N]Phe- and [^{13}C]Leu-enriched protein. [i] No serine was present in the media. The cysteine residues were enriched ca. 5%. [j] The media also contained [1-^{13}C]valine, [1-^{13}C]alanine, or [1,2-$^{13}C_2$]glycine to produce [^{15}N]Tyr- and [^{13}C]Val-, -Ala-, or -Gly-labeled T4 lysozyme, respectively. [k] The appropriate *E. coli* host would have been DL39, DL39 *avt*A::Tn5, or derivatives thereof, but these strains were unavailable when the particular proteins were produced.

sium phosphate, 1 mM reduced dithiothreitol, and 0.01% sodium azide at an observed pH* of 5.3–5.7. T4 lysozyme was transferred to D_2O solutions by passing approximately 0.5 mL of protein solution through a 2.5-mL Quik-sep spin column (Isolab Inc.) containing Sephadex G-25 preequilibrated in deuterated buffer. The protein was stored in D_2O buffer for at least 10 h to allow the most labile protons to exchange before recording NMR spectra. To completely exchange the amide protons of T4 lysozyme for deuterons, the protein was first transferred to a D_2O solution of 25 mM KCl at approximately pH* 2.0, then reversibly unfolded by heating to 40–50 °C, refolded by cooling, and finally transferred to the above specified D_2O buffer by using a spin column.

NMR Experiments. The NMR spectra were usually recorded with NMR spectrometers operating at 11.9 T, using 1–4 mM protein samples at pH 5.2–5.7 and at 20 °C. Under these conditions typical amide proton T_1 and T_2 relaxation times are 300–500 ms and 15–20 ms, respectively, giving proton line widths of 15–20 Hz (Griffey et al., 1986). T4 lysozyme is monomeric in solution to at least 3 mM as judged by equilibrium centrifugation and light scattering (L. P. McIntosh and J. H. Geiselmann, not shown). However, at concentrations above 4 mM, the line widths of T4 lysozyme begin to increase. The recycle times, including acquisition, were usually 650 ms for the two-dimensional experiments. The chemical shifts were set relative to an external proton reference of DSS and an external nitrogen standard of $^{15}NH_4Cl$ [2.9 M in 1 M HCl, 20 °C, 24.93 ppm relative to NH_3; Levy and Lichter (1979)].

The NMR experiments were recorded on a custom-built instrument operating at 500 MHz (A.G.R.), a General Electric GN500 spectrometer (F.W.D.), and on Bruker AM500 and AM600 spectrometers (A.J.W.). All the spectra discussed in this report were measured at 11.9 T (500 MHz for ^1H) unless stated otherwise. The NMR data were processed with software written by A.G.R. and the FTNMR package obtained from Hare Research.

(A) ^1H–^{15}N Correlation Experiments. Single-bond ^1H–^{15}N correlation spectra of ^{15}N-labeled proteins in H_2O and D_2O solutions were measured primarily by the heteronuclear multiple-quantum coherence experiment [HMQC; Bax et al. (1983)], as described previously for T4 lysozyme (Griffey et al., 1985a). A heteronuclear single-multiple-quantum coherence experiment [HSMQC; Zuiderweg (1990)], which combines the HMQC and Hahndor (Redfield, 1983) experiments, was also used to obtain ^1H–^{15}N correlation spectra with narrower line widths in the ^{15}N dimension. A delay of 4.5 ms instead of 5.4 ms was used as nominal $(2J_{NH})^{-1}$ time period to reduce the loss of signal due to relaxation. The signal from H_2O or residual HDO was suppressed by using a modified jump–return pulse sequence (Griffey et al., 1985a) or by selective presaturation. The ^1H observe sweep width was 6666 Hz when the transmitter was set to the resonance frequency of water for selective saturation (Zuiderweg et al., 1986) and 4000 Hz when the transmitter was set in the amide spectral region and selective excitation was employed. The ^{15}N sweep width was 4000–5000 Hz. Typically, spectra were collected with 1024 complex data points in the t_2 domain and 256–300 increments in t_1. Quadrature detection in the nitrogen dimension was by the method of States et al. (1983). WALTZ-16 or MLEV-64 broad-band ^{15}N decoupling (Levitt et al., 1982; Shaka et al., 1983) was applied during the acquisition period. The experiments usually required 8–16 h, although useable HMQC spectra can be recorded in as little as 8 min (A.G.R., unpublished). The data were processed with Gaussian resolution enhancement and zero filling to 2048 and 1024 data points in the ^1H and ^{15}N dimensions, respectively.

The single-bond ^1H-^{15}N coupling constants ($^1J_{NH}$) were measured from ^{15}N-coupled HMQC spectra of T4 lysozyme uniformly enriched with ^{15}N and selectively labeled with [α-^{15}N]Glu, -Gln, and -Arg, with [^{15}N]Leu, with [^{15}N]Lys, or with [^{15}N]Val. Delays of 5.4 ms [$=(2J_{NH})^{-1}$] were used to obtain pure absorptive proton doublets, assuming an average coupling constant of 92 Hz. The spectra were collected as 1024 complex points in t_2, apodized with a shifted sine-bell function, and zero filled to 8192 data points in the proton dimension for a final digital resolution of 1.22 Hz per point.

^1H–^{15}N multiple-bond correlation spectra (HMBC) of uniformly ^{15}N labeled T4 lysozyme in 99% D_2O were measured as described by Bax and Summers (1986) and Bax et al. (1988) with delays of 50–70 ms. The spectra were processed in a mixed-mode format (Bax & Marion, 1988).

(B) Two-Dimensional ^1H NMR. Phase-sensitive DQF-COSY (Shaka & Freeman, 1983; Rance et al., 1984),

NOESY (Macura & Ernst, 1980; Kumar et al., 1980), and TOCSY (Braunschwiler & Ernst, 1983; Bax & Davis, 1985) spectra were recorded of T4 lysozyme samples in H_2O and D_2O solutions. Spectral widths of 6410 and 7353 Hz in both dimensions were used for measurements at 500 MHz and a spectral width of 8333 Hz was used for experiments at 600 MHz. The H_2O signal was suppressed by selective saturation (Zuiderweg et al., 1986), and the SCUBA pulse sequence (Brown et al., 1988) was used to recover intensity from protein proton resonances also saturated by the decoupler. Mixing times of 25–70 ms using the MLEV-17 pulse train without trim pulses were employed for the TOCSY experiments (Bax & Davis, 1985). NOESY spectra were recorded with mixing times ranging from 30 to 110 ms.

The data were collected with 1024 complex points in t_2 and 400–800 increments in t_1 using TPPI (Marion & Wüthrich, 1983) and were transformed after zero filling to yield real two-dimensional spectra of 2048 × 2048 points. All spectra were processed with Guassian enhancement in the t_2 domain, followed by base-line flattening after Fourier transformation by using a fourth-order polynomial function. The DQF-COSY spectra were processed by using multiplication by an unshifted sine-bell in the t_1 domain and the NOESY and TOCSY spectra were apodized by using a 30° shifted sine-bell. To reduce t_1 ridges in the latter spectra, the first data point in the t_1 interferogram was multiplied by 0.3 prior to real Fourier transformation (Otting et al., 1986a).

Two-dimensional ^1H NMR spectra of uniformly ^{15}N labeled T4 lysozyme were also recorded with MLEV-64 decoupling of the nitrogen-15 nuclei during both the t_1 and t_2 time periods.

(C) ^{15}N-Edited ^1H NMR Experiments. ^{15}N-ω_2-Edited COSY and NOESY spectra of uniformly and selectively ^{15}N labeled T4 lysozyme were recorded using the conventional phase-sensitive COSY and NOESY pulse sequences with the final 90° read pulse replaced by a difference echo sequence (Emschwiller et al., 1960; Bendall et al., 1981; Freeman et al., 1981):

^1H	(90_x)	----Δ----	(180_α)		----Δ----	$acq_{(\beta+\alpha+\alpha)}$
^{15}N			(90_x)	(90_β)		decouple

A delay of 4.5 ms was used for Δ [ca. $(2J_{NH})^{-1}$] and β was cycled $(x, -x)$ to select for resonances from ^{15}N-labeled protons. To improve the suppression of the resonances from unlabeled protons, α was incremented in successive 90° steps after every second scan [EXORCYCLE; Bodenhausen et al. (1977)]. MLEV-64 broad-band ^{15}N decoupling was applied during t_1 and t_2. The resonance from H_2O was suppressed by selective presaturation. The sweep width was 6666 Hz in both proton dimensions and the data were collected as 180–300 complex t_1 increments (States et al., 1983) with 1024 complex points in the t_2 domain. The edited COSY spectra were processed by using a squared skewed unshifted sine-bell for apodization in the t_2 domain and an unshifted sine-bell for apodization of the t_1 interferogram. The edited NOESY spectra were recorded with a mixing period of 30–100 ms for the uniformly ^{15}N labeled protein and 100 ms for selectively labeled proteins. These were processed in a fashion similar to the conventional ^1H NOESY spectra. The data was zero filled to 2048 points in both dimensions but the spectral region upfield of water was discarded after the t_2 transform to yield final data matrices of 2048 points in ω_1 and 1024 points in ω_2. These and similar isotope-edited experiments have been discussed by several research group (Otting et al., 1986b; Bax & Weiss, 1987; Griffey & Redfield, 1987; McIntosh et al., 1987b; Senn et al., 1987a,b; Fesik et al., 1987; Wörgötter et al., 1988) and recently reviewed by Otting and Wüthrich (1990).

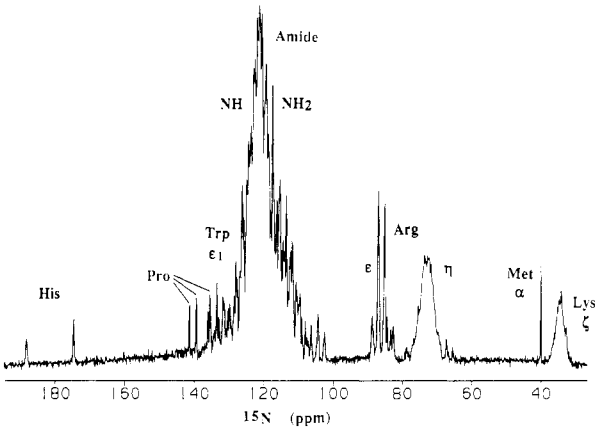

FIGURE 1: ^{15}N NMR spectrum of uniformly ^{15}N enriched T4 lysozyme in H_2O buffer, pH 5.6, at 20 °C. Proton decoupling was applied only during the aquisition period and NOE enhancement of the ^{15}N resonances was not utilized. The recycle time was 12 s to allow relaxation of the water resonance and prevent saturation of protons on nitrogen nuclei due to chemical exchange. The ^{15}N sweep width was 10 kHz and the spectrum was accumulated with 1284 scans.

The HMQC–COSY spectrum of uniformly ^{15}N labeled T4 lysozyme in 99% D_2O was acquired as described by Clore et al. (1988) and Gronenborn et al. (1989).

RESULTS

^1H–^{15}N Spectra of T4 Lysozyme—Amide Assignments by Amino Acid Type

(A) ^{15}N NMR Spectrum of ^{15}N Labeled T4 Lysozyme. The one-dimensional ^{15}N NMR spectrum of uniformly ^{15}N labeled T4 lysozyme is presented in Figure 1. Many of these resonances are classified by functional group on the basis of chemical shift and the presence or absence of attached protons (Gust et al., 1975). The two singlets at 187.8 and 174.1 ppm arise from the imidazole ring nitrogens of the sole histidine residue in T4 lysozyme. The three singlets at 140.9, 138.9, and 135.0 ppm are easily assigned to the imino nitrogens of the three proline residues in this protein since they are absent in a DEPT experiment and hence arise from nitrogen nuclei without any directly bonded protons (Doddrell et al., 1982). The resonances of the tryptophan $N^{\epsilon 1}$ nitrogens, the backbone secondary amide nitrogens, and the side-chain primary amide nitrogens fall in the range of 135–103 ppm. The peaks near 85 ppm are from the N^{ϵ} and those near 72 ppm correspond to the $N^{\eta 1}$ and $N^{\eta 2}$ of the 13 arginine guanidinium side chains. The single peak at 39.6 ppm is assigned to the α-nitrogen of the N-terminal methionine residue from the ^{15}N NMR spectrum of T4 lysozyme selectively labeled with [^{15}N]-methionine (not shown). Finally, the peaks near 34 ppm arise from the charged N^{ζ} nitrogens of the 13 lysine side chains.

(B) ^1H–^{15}N Correlation Spectra of Uniformly ^{15}N Labeled T4 Lysozyme. The ^{15}N NMR spectrum of T4 lysozyme is useful to broadly classify the peaks by nitrogen type, and in the few cases stated above, specific assignments can be made. In a similar fashion, the ^1H NMR spectrum of this protein can be subdivided into chemical shift regions characteristic of protons in various chemical environments (Bundi & Wüthrich, 1979; Gross & Kalbitzer, 1988). However, neither one-dimensional experiment offers the resolution required to assign all the resonances in each spectral region.

Our basic strategy was to focus our attention on a subset of the protons in T4 lysozyme: namely, those that are directly bonded to nitrogen atoms. In Figure 2, a portion of the ^1H–

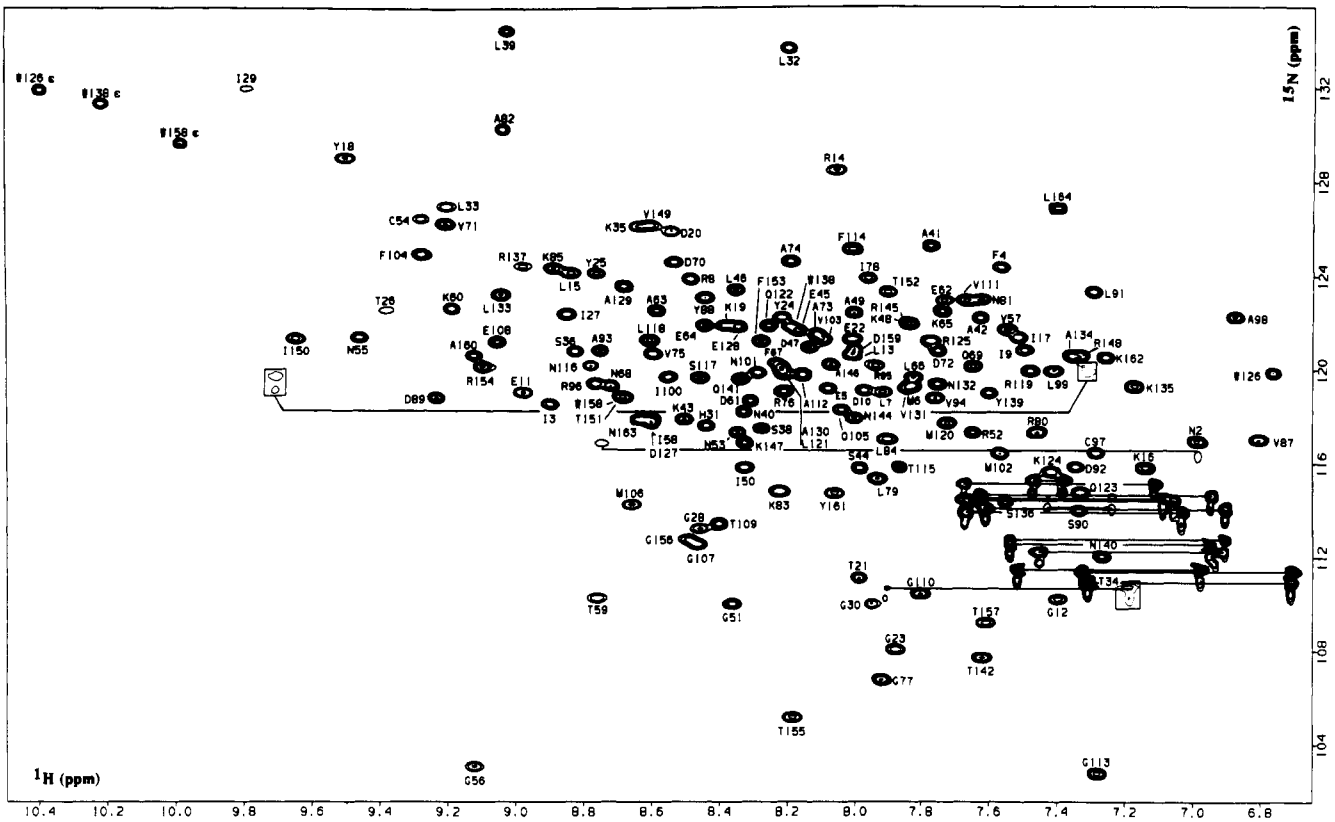

FIGURE 2: HSMQC spectrum of uniformly ^{15}N enriched T4 lysozyme in H_2O buffer. Only the amide region of the spectrum is shown. The assignments of all 160 backbone amide and three tryptophan indole $^{15}N^{\epsilon 1}H$ cross peaks are indicated. The asparagine and glutamine side-chain $^{15}NH_2$ groups appear as two peaks at a common ^{15}N chemical shift, as indicated by the horizontal lines. These peaks have not been specifically assigned. The peaks enclosed in boxes are drawn at a contour level, which is approximately one-third lower than used for the remainder of the spectrum.

^{15}N HSMQC spectrum of T4 lysozyme uniformly enriched in ^{15}N is presented. Each cross peak in the spectrum corresponds to a ^1H–^{15}N pair of the amide ^{15}NH and $^{15}NH_2$ and tryptophan $^{15}N^{\epsilon 1}H$ groups in T4 lysozyme. The HSMQC (or equivalently, the HMQC) experiment edits the proton NMR spectrum of T4 lysozyme, yielding only resonances from ^{15}N-labeled protons, and adds the dispersion of the indirectly detected ^{15}N dimension. We have also recorded the H(S)MQC spectra of labeled T4 lysozyme in D_2O buffer. Only the amides that remain protonated are detected and hence the ^1H–^{15}N groups in lysozyme can be further differentiated on the basis of hydrogen exchange kinetics [see McIntosh et al. (1987a) for representative spectra].

The HSMQC spectrum of T4 lysozyme is a signature of the protein backbone. The assignments of the ^1H–^{15}N cross peaks are indicated in Figure 2 and listed in Table II (for clarity, we will refer to the specific assignments of resonances throughout this section, although these assignments were obtained at a later stage of this procedure). In contrast to the highly overlapping amide regions of the one-dimensional ^1H and ^{15}N NMR spectra, the peaks in the HSMQC spectra are remarkably well resolved. Each of the expected 160 secondary amide ^{15}NH groups (164 residues less the N terminus and 3 prolines) is detected as a single peak. In only seven cases, two or more backbone amides have overlapping resonances in both ^1H and ^{15}N dimensions, and these peaks can be separated by recording spectra at different temperatures or pH (not shown) or distinguished by selective isotopic labeling or by their differential hydrogen exchange rates. We have found that the resolution in the ^{15}N dimension is slightly higher in the HSMQC experiment than in the HMQC experiment (Zuiderweg, 1990), thus helping to identify peaks in the crowded regions of Figure 2.

The HSMQC spectrum also contains cross peaks from the side-chain ^1H–^{15}N pairs that are not significantly broadened by proton exchange with the solvent. The peaks corresponding to the indole $^{15}N^{\epsilon 1}H$ of tryptophan residues 126, 138, and 158 are indicated in Figure 2. These resonances were assigned by selective tryptophan to tyrosine substitutions (L.P.M., unpublished). The asparagine γ and glutamine δ $^{15}NH_2$ groups yield two ^1H–^{15}N peaks in the HSMQC spectra with a common ^{15}N shift and a distinctive line shape due to the passive coupling of the geminal protons. All of peaks from the 17 $^{15}NH_2$ groups in T4 lysozyme are identified in Figure 2. The 13 arginine $^{15}N^{\epsilon}H$ peaks are also detected in the HSMQC spectrum [not shown; see Muchmore et al. (1989)]. These primary amide and guanidinium resonances are not presently assigned to specific residues as the spin systems of the asparagine, glutamine, and arginine residues have not yet been fully identified.

(C) ^1H–^{15}N Correlation Spectra of Selectively ^{15}N Labeled T4 Lysozymes. The ^1H–^{15}N HSMQC spectrum of uniformly ^{15}N labeled T4 lysozyme served as the master spectrum for assigning the backbone resonances of this protein. The first step toward these assignments involved classifying each HMQC cross peak by amino acid type. We prepared samples of T4 lysozyme selectively α-^{15}N enriched at one or more of every amino acid except histidine and cysteine (Table I). In Figure 3, the HMQC spectra of T4 lysozyme selectively α-^{15}N enriched at the alanine, the aspartate and asparagine, the arginine and glutamate and glutamine, and the serine and glycine residues are illustrated. By direct comparison, the corresponding peaks in the HSMQC spectrum of uniformly ^{15}N labeled protein were identified by amino acid type. We found that the spectra of many different protein samples could be accurately compared provided that similar experimental

Table II: Chemical Shifts of Assigned ^1H and ^{15}N Resonances of T4 Lysozyme

residue	chemical shift (ppm)[a] α-^{15}N	H^N	H^α	H^β	others
Met1	39.9				
Asn2	116.9	6.99	4.49		
Ile3	118.6	8.89	4.12		
Phe4	124.3	7.56	3.83	3.17, 3.05	H^δ 6.77, H^ϵ 6.93, H^ζ 6.99
Glu5	119.2	8.06	3.70	1.80	
Met6	119.3	7.84	3.30		
Leu7	119.0	7.91	3.97		
Arg8	123.9	8.48	3.41	1.73	
Ile9	120.8	7.52	3.54	2.08	
Asp10	119.1	7.97	4.49	2.60	
Glu11	119.0	8.98	4.59	2.11, 1.93	H^γ 2.30
Gly12	110.3	7.40	4.13, 3.77		
Leu13	120.1	7.98	4.40	1.69	
Arg14	128.5	8.05	4.69	1.90	
Leu15	124.1	8.83	4.36	1.60	
Lys16	115.7	7.15	5.43	1.93, 1.78	H^γ 1.53
Ile17	121.4	7.53			
Tyr18	129.0	9.48	5.13	3.11, 3.01	H^δ 7.19, H^ϵ 6.57
Lys19	121.9	8.35	4.58	1.59	
Asp20	125.9	8.54	4.60		
Thr21	111.2	8.00	4.00	4.32	H^γ 1.38
Glu22	121.3	7.99	4.17	1.70	
Gly23	108.1	7.88	3.98, 3.34		
Tyr24	122.2	8.16	4.97	2.97, 2.91	H^δ 7.12, H^ϵ 6.81
Tyr25	124.1	8.73	4.90	2.81	H^δ 7.01, H^ϵ 6.71
Thr26	122.6	9.37	5.06	2.80	H^γ 1.02
Ile27	122.3	8.85	4.61		
Gly28	113.3	8.47	4.28, 3.58		
Ile29	132.0	9.79	4.75	1.97	
Gly30	110.0	7.94			
His31	117.6	8.44	4.61	3.68	$H^{\delta 2}$ 8.65, $H^{\epsilon 1}$ 6.38
					$N^{\delta 1}$ 188.0, $N^{\epsilon 2}$ 174.4[b]
Leu32	133.7	8.17	4.06	1.41	
Leu33	127.0	9.18	4.48	1.73	
Thr34	110.9	7.31	4.48		
Lys35	126.1	8.61	4.26	1.42	
Ser36	120.8	8.81	4.66	4.03, 3.88	
Pro37	[c]				
Ser38	117.4	8.26	4.74	3.88	
Leu39	134.4	9.00	3.37	1.71	
Asn40	118.2	8.31	4.29	2.73, 2.62	
Ala41	125.2	7.77	4.24	1.63	
Ala42	122.2	7.62	3.74	1.23	
Lys43	117.9	8.49	3.73	1.88	
Ser44	115.7	7.98	4.30	4.02	
Glu45	121.6	8.16	4.14	2.02	
Leu46	123.4	8.33	3.97		
Asp47	120.9	8.12	4.85	2.72	
Lys48	122.0	7.84	4.02	1.92	$H^{\beta,\gamma}$ 1.61, 1.42
Ala49	122.4	8.00	4.18	1.50	
Ile50	115.8	8.32	4.07	1.96	H^γ 0.93
Gly51	110.0	8.36	4.17, 3.77		
Arg52	117.3	7.65	4.56	1.74	
Asn53	117.3	8.32	4.72	2.85, 2.70	
Cys54	126.5	9.28	4.50	3.05	
Asn55	121.4	9.45	4.32	3.14, 2.57	
Gly56	103.1	9.10	4.09, 3.31		
Val57	121.7	7.54	5.18	1.96	H^γ 0.99
Ile58	117.8	8.60	4.90		
Thr59	110.2	8.76	4.64		
Lys60	122.6	9.17	3.88		
Asp61	118.6	8.29	4.38	2.60	
Glu62	122.9	7.73	3.90	2.58	
Ala63	122.5	8.57	3.91	1.51	
Glu64	121.8	8.43	3.88		
Lys65	122.5	7.76	4.12	1.98	
Leu66	119.6	7.83	3.91	1.77	
Phe67	120.1	8.22	5.07	3.39	3.39 H^δ 7.57, H^ϵ 7.04
Asn68	119.3	8.71	4.29	2.93	
Gln69	120.0	7.65	4.23	2.44	
Asp70	124.6	8.52	4.64	3.22	
Val71	126.2	9.20	3.64	1.94	H^γ 0.84, 0.47
Asp72	120.8	7.75	4.38	2.87	
Ala73	121.2	8.10	4.04	1.58	
Ala74	124.6	8.18	4.19	1.50	
Val75	120.7	8.59	3.26	2.16	H^γ 1.05, 0.78
Arg76	119.0	8.22	3.82	1.90	
Gly77	106.8	7.92	3.86, 3.70		
Ile78	123.9	7.96	3.35	1.79	
Leu79	115.3	7.93	3.80		
Arg80	117.3	7.46	4.36	1.86	
Asn81	123.0	7.63	4.76		
Ala82	130.3	9.01	4.10	1.48	
Lys83	114.8	8.22	4.34	1.91	
Leu84	117.0	7.90	4.41	2.16	
Lys85	124.3	8.87	3.88		
Pro86	[c]				
Val87	117.0	6.80	3.51	2.28	H^γ 1.11, 0.95
Tyr88	123.0	8.43	3.72	3.13	H^δ 7.03, H^ϵ 6.73
Asp89	118.8	9.21	4.31	2.68	
Ser90	114.0	7.34	4.44	4.12	
Leu91	123.3	7.30	4.47	2.34	
Asp92	115.7	7.36	4.64		
Ala93	120.8	8.72	3.83	1.55	
Val94	118.8	7.77	3.18	1.32	H^γ 0.56, −0.78
Arg95	120.1	7.94	3.39	1.64	
Arg96	119.4	8.77	3.70		
Cys97	116.4	7.30	4.08		
Ala98	122.2	6.89	3.68	0.41	
Leu99	119.9	7.41	4.22	2.29	
Ile100	119.7	8.55	3.42	1.87	
Asn101	119.9	8.28	4.06		
Met102	116.4	7.58	4.36	2.19	
Val103	121.3	8.10	3.32	2.22	H^γ 0.90
Phe104	125.0	9.26	4.13	3.39, 3.24	H^δ 7.31[d]
Gln105	118.2	8.03	4.42	2.21	
Met106	114.2	8.65	4.78	2.15	
Gly107	112.6	8.47	4.49, 3.97		
Glu108	121.2	9.05	3.68	2.00, 1.89	H^γ 2.24
Thr109	113.4	8.38	3.84	4.13	H^γ 1.23
Gly110	110.4	7.77	3.82, 3.59		
Val111	122.9	7.66	3.65	1.85	H^γ 0.93, 0.84
Ala112	119.8	8.13	3.88	1.43	
Gly113	102.8	7.28	4.15, 3.80		
Phe114	125.1	7.99	5.03	3.35, 3.13	H^δ 7.50, H^ϵ 7.27, H^ζ 7.21
Thr115	115.6	7.87	3.69	4.07	H^γ 1.27
Asn116	120.2	8.76	4.54	2.98, 2.71	
Ser117	119.7	8.45	4.19	4.00	
Leu118	121.2	8.59	3.87	1.91	
Arg119	119.9	7.50	4.08	1.96	
Met120	117.7	7.72	3.97		
Leu121	119.8	8.20	3.87		
Gln122	121.9	8.23	3.98	2.28	
Gln123	114.8	7.33	4.10	1.74	
Lys124	115.6	7.42	1.40		
Arg125	121.2	7.77	4.24	1.42	
Trp126	119.8	6.76	4.35	3.87, 3.27	$H^{\epsilon 1}$ 10.38, $^{15}N^{\epsilon 1}$ 132.0
					$H^{\delta 1}$ 7.42, $H^{\zeta 2}$ 7.58, $H^{\eta 2}$ 7.35, $H^{\zeta 3}$ 6.92, $H^{\epsilon 3}$ 7.26
Asp127	117.8	8.59	4.43	2.72	
Glu128	121.8	8.33	3.97	2.11	
Ala129	123.6	8.66	4.00	1.32	
Ala130	120.0	8.20	4.07	1.59	
Val131	119.3	7.82	3.60	2.13	H^γ 1.09, 0.92
Asn132	119.4	7.74	4.33	2.77	
Leu133	123.2	9.01	4.20	2.09	
Ala134	120.6	7.39	3.67	0.90	
Lys135	119.3	7.18	4.42	1.84	
Ser136	114.4	7.56	4.66	4.79	
Arg137	124.2	8.96	4.26	2.02	
Trp138	121.7	8.17	4.55	3.68, 3.59	$H^{\epsilon 1}$ 10.21, $^{15}N^{\epsilon 1}$ 131.4
					$H^{\delta 1}$ 7.30, $H^{\zeta 2}$ 7.66, $H^{\eta 2}$ 7.16, $H^{\zeta 3}$ 6.95, $H^{\epsilon 3}$ 7.78
Tyr139	119.0	7.60			H^δ 6.71, H^ϵ 6.63
Asn140	112.0	7.27	4.37	2.89, 2.84	
Gln141	119.6	8.32	4.22	2.37, 2.22	
Thr142	107.7	7.62	4.74	4.51	H^γ 1.17
Pro143	[c]				
Asn144	117.9	8.00	4.31	2.62	
Arg145	122.0	7.84	3.91		
Ala146	120.2	8.07	3.71	0.02	
Lys147	116.9	8.31	3.96	1.93	
Arg148	120.6	7.35	3.80		
Val149	126.1	8.59	3.36	2.58	H^γ 1.30, 1.10
Ile150	121.4	9.63	3.90	2.11	
Thr151	118.8	8.67	3.95	4.24	H^γ 1.31
Thr152	123.3	7.90	3.95	4.31	H^γ 0.56, $H^{\gamma 1}$ 5.54
Phe153	121.2	8.27	4.14	3.33	H^δ 6.95, H^ϵ 7.40, H^ζ 6.72
Arg154	120.1	9.06	4.10	2.09	
Thr155	105.2	8.19	4.26		

Table II (Continued)

residue	chemical shift (ppm)[a] α-^{15}N	H^N	H^α	H^β	others	residue	chemical shift (ppm)[a] α-^{15}N	H^N	H^α	H^β	others
Gly156	112.7	8.49	3.68, 3.15			Ala160	120.6	9.12	4.21	1.24	
Thr157	109.2	7.61	4.67	4.40	H^γ 1.09	Tyr161	114.7	8.06	4.28	2.79	H^δ 7.14, H^ϵ 6.61, H^η 11.29
Trp158	118.8	8.66	5.11	3.79, 2.99	$H^{\epsilon 1}$ 9.97, $^{15}N^{\epsilon 1}$ 129.7 $H^{\delta 1}$ 7.23, $H^{\zeta 2}$ 7.42, $H^{\eta 2}$ 6.93, $H^{\zeta 3}$ 6.38, $H^{\epsilon 3}$ 7.27	Lys162	120.5	7.26	4.13	1.81	
						Asn163	117.8	8.60	4.59	2.92, 2.82	
Asp159	120.8	7.99	4.31			Leu164	126.8	7.40	4.21	1.65, 1.56	

[a] The assignments are reported for T4 lysozyme at pH 5.6 and 20 °C. The proton chemical shifts are referenced to external DSS with an error of ±0.02 ppm, and the nitrogen chemical shifts are referenced to NH_3 with an error of ±0.2 ppm. [b] The imidazole nitrogen resonances are tentatively assigned from HMBC spectra. [c] The imino nitrogen resonances of the three proline residues are at 140.9, 138.9, and 135.0 ppm. [d] Tentatively assigned as a H^δ resonance of Phe104.

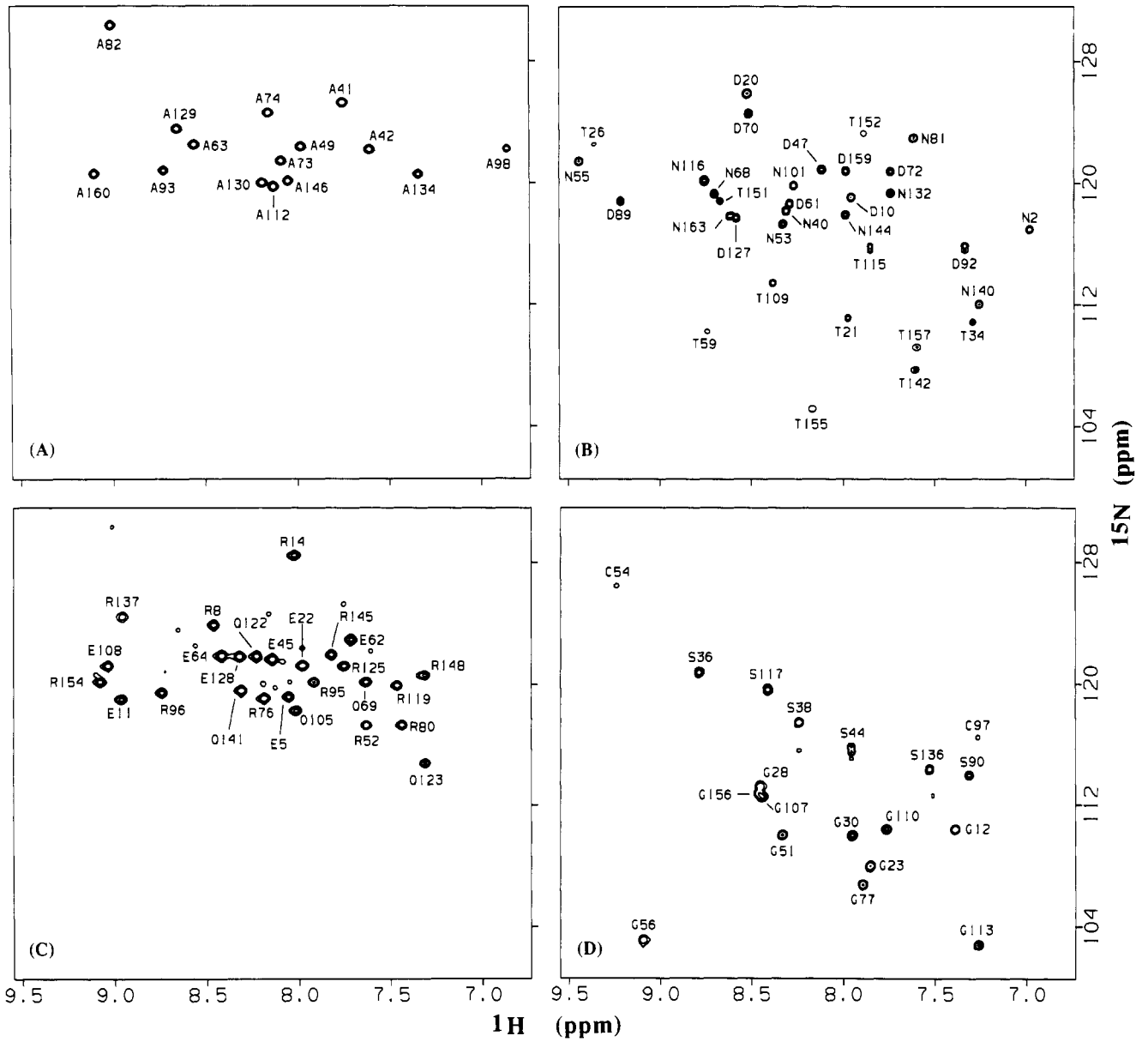

FIGURE 3: HMQC spectra of selectively labeled T4 lysozyme. (A) [^{15}N]alanine-labeled protein. (B) [^{15}N]aspartate- and -asparagine-labeled protein. The threonine residues are ^{15}N enriched to approximately 20%. (C) [^{15}N]glutamate-, -glutamine-, and -arginine-labeled protein. The alanine residues are ^{15}N enriched to approximately 20% and can be identified by comparison to panel A. (D) [^{15}N]serine- and -glycine-labeled T4 lysozyme. The cysteine residues are approximately 5% ^{15}N enriched.

conditions were carefully maintained.

Following this approach, we have analyzed the HMQC spectra of 16 selectively labeled T4 lysozyme samples (Ala, Arg, Asp, Asx, Gly, Glx, Ile, Leu, Lys, Met, Phe, Ser, Thr, Trp, Tyr, and Val; Table I). The peaks corresponding to asparagine were distinguished from those of aspartate by comparison of the spectra of proteins that were α-^{15}N enriched in both the aspartate and asparagine residues and only in the aspartate residues. In contrast, the resonances from glutamate and glutamine were not distinguished. The T4 lysozyme sample containing isotopically α-^{15}N (and ϵ-^{15}N) enriched arginine was prepared with [^{15}N]glutamate as the source of the isotope and thus was also labeled at the glutamate and glutamine backbone amides (Table I). The amide and side-chain resonances from the arginines were easily identified by comparison to the spectrum of lysozyme labeled with only the

Table III: Single Bond $^1J_{NH}$ Coupling Constants of T4 Lysozyme

group	$^1J_{NH}$ (Hz)[a] mean[b]	range
backbone amide ^1H–^{15}N	92 (1)	90–98
side-chain amide ^1H–^{15}N	90 (1)[c]	89–91
	87 (1)[d]	87–89
tryptophan ^1H–$^{15}N^{\epsilon 1}$	99[e]	
arginine ^1H–$^{15}N^{\epsilon}$	92 (1)	90–93

[a] The magnitude of the single-bond ^1H–^{15}N coupling constant measured from ^{15}N-coupled HMQC spectra zero filled to a digital resolution of 1.22 Hz/point. [b] The mean coupling constant determined from resolved resonances. The standard deviations are given in parentheses. [c] The mean measured coupling constant from the downfield proton resoance of each $^{15}NH_2$ pair, tentatively ascribed to the trans proton (Redfield & Waelder, 1979). [d] The mean measured coupling constant from the upfield proton resonance of each $^{15}NH_2$ pair, tentatively ascribed to the cis protons. [e] The measured coupling constant for free $^{15}N^{\epsilon 1}$ tryptophan in buffer conditions simlar to those used for these studies is 98.3 Hz.

latter two amino acids. In a similar fashion, samples of protein labeled with only glycine and with both glycine and serine were analyzed (Table I). The ^1H–^{15}N peaks from the two cysteine residues were identified as minor peaks in the HMQC spectra of T4 lysozyme labeled with [^{15}N]glycine. The peak from the sole histidine was identified as a remaining unassigned ^1H–^{15}N signal in the HSMQC spectrum of uniformly enriched protein. All 160 secondary amide peaks were accounted for and classified as 1 of 18 types of amino acids by this procedure.

(D) Single-Bond $^1J_{NH}$ Coupling Constants. The single-bond $^1J_{NH}$ coupling constants of the ^1H–^{15}N pairs in the backbone and side chains of T4 lysozyme were measured from the ^{15}N-coupled HMQC spectra of uniformly or selectively ^{15}N enriched protein samples. These data are presented in Table III. The magnitudes of the backbone amide $^1J_{NH}$ coupling constants range from 90 to 98 Hz, but cluster very closely to a mean value of 92 Hz (standard deviation 1 Hz). The largest observed $^1J_{NH}$ constant is 98 Hz for Lys124. The coupling constants of the glycine residues are the most variable, ranging from 93 to 97 Hz in T4 lysozyme. Aside from this, we did not observe any obvious dependence of the backbone amide $^1J_{NH}$ on amino acid type or on structural parameters such as hydrogen bonding or the deviations of the amide peptide bonds from planarity (the dihedral angle ω) as determined from the crystallographic model of T4 lysozyme (Weaver & Matthews, 1987). LeMaster and Richards (1985) described similar results for *E. coli* thioredoxin. Although studies of smaller molecules have suggested that the backbone $^1J_{NH}$ coupling constants may depend on such parameters (Llinás et al., 1976), this does not appear fruitful for investigating the structures of proteins.

The mean $^1J_{NH}$ coupling constant for the downfield proton resonance of each side-chain primary amide in T4 lysozyme is 90 Hz, whereas that for the upfield proton is 87 Hz (Table III). On the basis of ^1H chemical shift, the downfield proton resonances are tentatively assigned to the trans (E) protons and the upfield resonance to the cis (Z) protons (Redfield & Waelder, 1979). This is consistant with studies of model primary and constrained secondary amides, which have demonstrated that the $^1J_{NH}$ coupling is slightly greater to a trans proton than to a cis proton (Sogn et al., 1973; Witanoski et al., 1986; Bystrov, 1976).

Identification of the ^{15}NH–H^{α}–H^{β} Spin Systems

We dissected the ^1H–^{15}N HMQC spectrum of T4 lysozyme by amino acid class using selective isotopic labeling. Although several peaks were assigned to specific residues by a variety of approaches discussed below, a general method was required to specifically identify all the amide resonances. As demonstrated by Wüthrich and co-workers using the sequential assignment method, the spin systems detected for the amino acids of a protein by J-correlated experiments can be placed within the polypeptide sequence by using NOE connectivities (Wüthrich, 1983, 1986). The sequential connections are established primarily by the NOEs between one or more of the H^N, H^{α}, and H^{β} protons of neighboring residues (Billeter et al., 1982; Wüthrich et al., 1984; Englander & Wand, 1987; Chazin & Wright, 1987). We followed this general method with the exceptions that emphasis was placed on assigning only the H^N, H^{α}, and H^{β} proton resonances of each residue and that these partial spin systems were identified by amino acid class using selective labeling.

The next stage of the isotope-aided assignment strategy was 2-fold: we identified at least the H^{α} and H^{β} resonances of most residues in the DQF-COSY and TOCSY spectra of unlabeled T4 lysozyme and classified these resonances by amino acid type using the HMQC and ^{15}N-edited COSY spectra of the selectively labeled proteins.

(A) N^N-H^{α} Fingerprint of T4 Lysozyme. The H^N–H^{α} region of a two-dimensional J-correlated spectrum of a protein is often referred to as the fingerprint region, with each residue yielding one cross peak, except for the glycine residues which may display two and the N-terminal and the proline residues which produce none (Wüthrich, 1986). In Figure 4, this section of the ^{15}N-edited COSY spectrum of uniformly ^{15}N enriched T4 lysozyme is presented.

T4 lysozyme is a 164-residue protein and contains 3 proline and 11 glycine residues. Maximally, we expect 160 H^N–H^{α} cross peaks plus 11 additional H^N–$H^{\alpha'}$ cross peaks from the glycine residues. From the first DQF-COSY spectra of unlabeled T4 lysozyme in H_2O and D_2O solutions recorded at 500 MHz, we identified approximately 120 and 65 candidate H^N–H^{α} cross peaks, respectively. At least 40 cross peaks were initially unaccounted for, and no obvious cross peaks from the glycine residues were apparent in these spectra. There are several reasons why these cross peaks were not detected, including being degenerate with other cross peaks, being saturated by the preirradiation at the resonance of water, or being too weak. Since T4 lysozyme has a high content of α-helical secondary structure (Weaver & Matthews, 1987), many of the $^3J_{HN\text{-}H\alpha}$ coupling constants are expected to be small, leading to weak antiphase cross peaks between the relatively broad H^N and H^{α} resonances (Neuhaus et al., 1985).

Subsequently, four additional H^N–H^{α} cross peaks close to the resonance of water and six glycine H^N–H^{α} and H^N–$H^{\alpha'}$ cross peaks were observed in the TOCSY spectra of T4 lysozyme recorded with the SCUBA presaturation recovery sequence. TOCSY experiments were useful for detecting weak or overlapping cross peaks due to the in-phase absorptive line shape. Later we recorded the (SCUBA) DQF-COSY spectrum of T4 lysozyme at 600 MHz. This spectrum had improved water suppression, higher signal-to-noise, and greater spectral dispersion than those recorded at lower field strength, and six more weak H^N–H^{α} cross peaks and four near-degenerate pairs of H^N–H^{α} cross peaks were identified. Finally, from isotope-edited COSY spectra of uniformly and selectively ^{15}N labeled T4 lysozyme samples, we detected additional weak cross peaks and identified several apparently single cross peaks as being the overlapping signals from degenerate H^N and H^{α} resonances.

All total, we have identified H^N–H^{α} COSY cross peaks for 157 residues in T4 lysozyme, including eight pairs and two triplets of cross peaks overlapping in both dimensions (Table

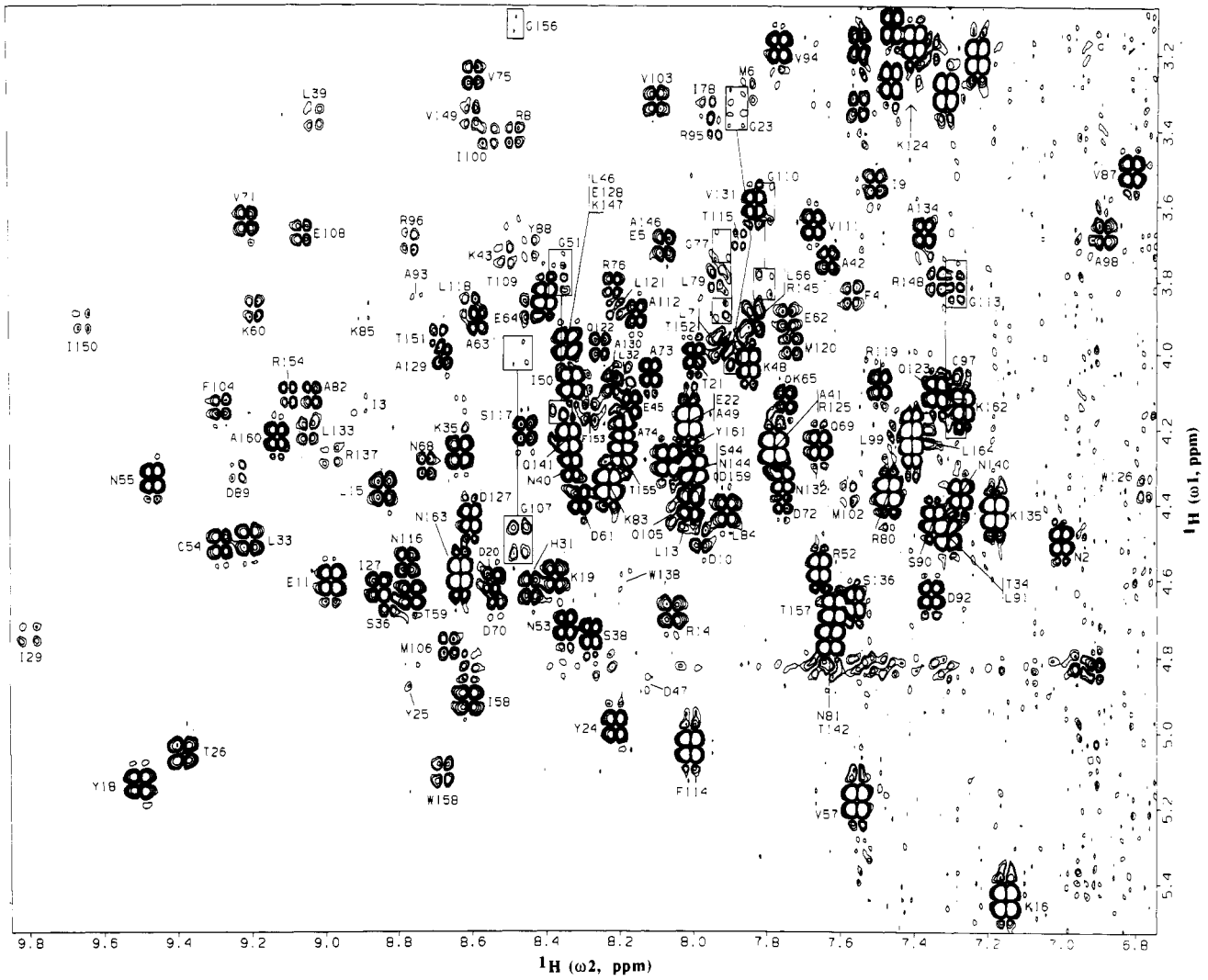

FIGURE 4: H^N–H^α region of an ^{15}N-ω_2-edited COSY spectrum of uniformly ^{15}N labeled T4 lysozyme. This figure is essentially the same as the DQF-COSY fingerprint of the unlabeled protein. The H^N–H^α cross peaks of 151 residues are identified in the figure. The cross peak of K124 at 7.42 and 1.40 ppm is not shown. The cross peaks of G12, G28, G56, F67, and N101 are not seen in this spectrum but have been identified in other TOCSY or COSY spectra. The cross peaks of I17, G30, and Y139 have not been identified. The strong cross peaks in the upper right-hand corner of the spectrum arise from the side chains of the arginine residues ($^{15}NH^\epsilon$–H^δ).

II; Figure 4). The H^α resonances of Ile17, Gly30, and Tyr139 remain undetected. The H^α protons were associated with the corresponding amide ^{1}H–^{15}N nuclei and thereby identified by amino acid type by three approaches: alignment of the HMQC ^{1}H–^{15}N cross peaks with the COSY H^N–H^α cross peaks; use of HMQC–COSY and HMBC experiments, which yield cross peaks between the ^{15}N and $^1H^\alpha$ resonances; and from ^{15}N-edited COSY spectra of the selectively labeled T4 lysozymes.

(B) Alignment of HMQC and DQF-COSY Spectra. The simplest approach toward identifying the H^α resonance corresponding to each ^{1}H–^{15}N pair is to align the HMQC ^{1}H–^{15}N cross peaks with the proton COSY H^N–H^α cross peaks. We sequentially recorded the HMQC and DQF-COSY spectra (with ^{15}N decoupling during t_1 and t_2) of the same sample of T4 lysozyme uniformly enriched in ^{15}N. These measurements were made with protein in H_2O buffer, with protein freshly transferred into D_2O buffer, and with lysozyme after storage for 4 months in deuterated buffer. By carefully maintaining identifical experimental conditions, the two species could be exactly aligned and the HMQC ^{1}H–^{15}N and the DQF-COSY H^N–H^α cross peaks associated on the basis of common amide proton chemical shift and hydrogen exchange rate. Following this approach, the ^{15}N–H^α spin systems of approximately 90 residues were confidently identified and those of a further 50 residues were tentatively assigned to one of a few possibilities.

(C) HMQC–COSY and HMBC Spectra of Uniformly ^{15}N Enriched T4 Lysozyme. A method for associating H^α resonances with the HMQC ^{1}H–^{15}N cross peaks is the two-dimensional HMQC–COSY experiment, which displays the H^N–H^α cross peak at the amide ^{15}N chemical shift (Clore et al., 1988; Weber & Mueller, 1989). We have recorded the HMQC–COSY spectrum of uniformly ^{15}N labeled T4 lysozyme, which was stored in D_2O buffer for 4 months, and have identified approximately 55 ^{15}NH–H^α spin systems corresponding to residues with amide proton that are protected from rapid hydrogen exchange (Figure 10). The H^α resonance of N101, which was never detected in the DQF-COSY or ^{15}N-edited COSY spectra of T4 lysozyme, was observed as a weak cross peak in the HMQC–COSY spectrum. This may be due to removal of the ^{15}N–^{1}H heteronuclear dipolar broadening in the former experiment, as discussed by Bax et al. (1989). We did not record the HMQC–COSY spectrum of T4 lysozyme in H_2O due to the interference from the strong residual signal of water.

An alternative method for identifying the ^{15}N and H^α resonances of amino acids is the HMBC experiment, which exploits the small intraresidue two-bond scalar coupling between the two nuclei (Bystrov, 1976; Bax et al., 1988; Clore et al., 1988). The assignments of about 25 $^{15}N_i$–H^α_i cross

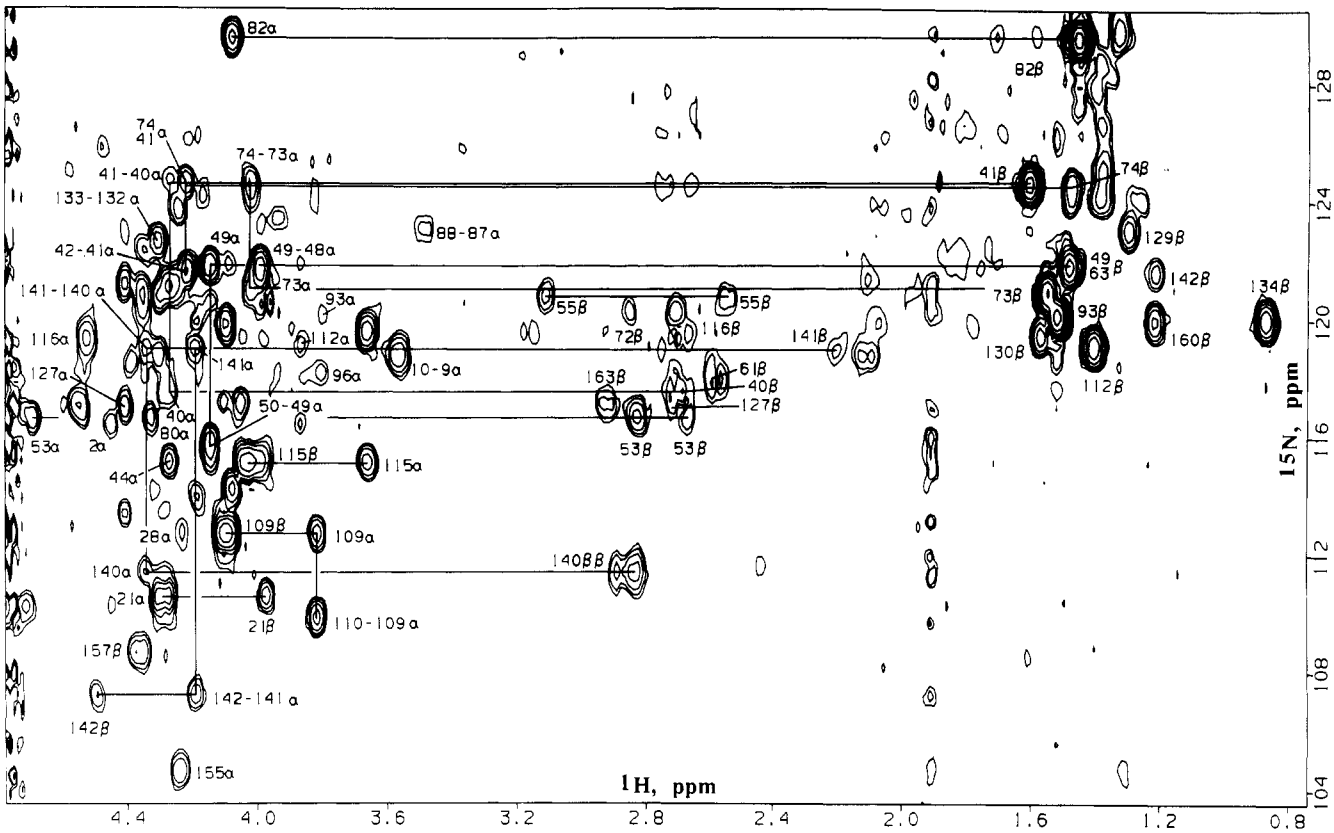

FIGURE 5: A portion of the ^1H–^{15}N HMBC spectrum of uniformly ^{15}N enriched T4 lysozyme. The protein was stored in D_2O buffer for approximately 1 year at 5 °C before the spectrum was accumulated and all but the most protected amides were deuterated. The amide ^{15}N chemical shifts are approximately 0.7 ppm upfield relative to the data given in Table II due to the deuterium isotope shift on ^{15}N (Hansen, 1983). The delay time was 70 ms in order to detect the weak multiple-bond ^1H–^{15}N couplings. A selection of the cross peaks are annotated. The intraresidue $^{15}N_i$–H^{α}_i and $^{15}N_i$–H^{β}_i cross peaks are labeled by the residue number. The interresidue $^{15}N_i$–H^{α}_{i-1} cross peaks are labeled by the numbers of both residues. The horizontal lines connect cross peaks common to one ^{15}N nuclei, and the vertical lines indicate sequential assignments based on the interresidue correlations.

peaks were confirmed from the HMBC spectrum of uniformly ^{15}N labeled T4 lysozyme in D_2O buffer (Figure 5). Alone, this approach was not very useful for correlating ^{15}N and H^{α} resonances, as only a fraction of the possible $^{15}N_i$–H^{α}_i cross peaks were observed. Presumably this was due to the short transverse relaxation times of the nuclei. Also, the interpretation of HMBC spectra is complicated by the occurrence of interresidue $^{15}N_i$–H^{α}_{i-1} and intraresidue $^{15}N_i$–H^{β}_i cross peaks and by a deuterium isotope shift of approximately −0.7 ppm on the ^{15}N resonances upon deuteration of the amides (Hansen, 1983; Bax et al., 1988). Nevertheless, the information obtained from Figure 5 was useful as support for the conclusions drawn through other approaches.

(D) ^{15}N-ω_2-Edited COSY Spectra of Selectively ^{15}N Labeled T4 Lysozyme. Isotopic labels in proteins can also be exploited to edit two-dimensional ^1H NMR spectra. Proton–proton scalar and dipolar interactions from or to protons directly bonded to ^{15}N or ^{13}C nuclei can be selectively observed (or suppressed) by replacing the first (ω_1-edit) or last (ω_2-edit) pulse of a conventional proton two-dimensional experiment with a difference echo pulse sequence (Otting et al., 1986b; Bax & Weiss, 1987; Griffey & Redfield, 1987; Otting & Wüthrich, 1990). By use of these approaches with selectively labeled protein samples, complex NMR spectra are dissected and cross peaks unambiguously assigned to residues of known amino acid class.

We measured the ^{15}N-edited COSY spectra of T4 lysozyme uniformly ^{15}N labeled or selectively enriched with ^{15}N-labeled Ala, Asx, Arg and Glx, Gly and Ser, Ile, Leu, Lys, Met, Phe, Thr, Tyr, or Val. In addition, we recorded this spectrum of [^{15}N]tryptophan-labeled T4 lysozyme containing the substitution of Trp126 to Tyr (the yield of [^{15}N]Trp-labeled wild-type protein was low and the sample was too dilute for these experiments). In each of these spectra, only the H^N–H^{α} COSY cross peaks of the particular ^{15}N-enriched amino acids are detected. This is shown in Figure 6 for T4 lysozyme isotopically enriched with [^{15}N]alanine and -valine, and in Figure 4 for the uniformly labeled protein. From these measurements, we have identified the H^{α} resonances of 144 out of a total of 156 selectively labeled residues, assigning these unambiguously by amino acid class. The glutamine and glutamate residues were not distinguished specifically. Fortunately, the isotope effect of ^{15}N on the chemical shift of protons is very small (Hansen, 1983), and the edited COSY spectra from labeled protein samples could be directly compared with the DQF-COSY spectra recorded for unlabeled T4 lysozyme.

The H^N–H^{α} cross peaks of the 12 remaining labeled residues (Ile17, Tyr25, Gly28, Gly30, Asp47, Gly56, Phe67, Gly77, Tyr88, Asn101, Tyr139, and Gly156) were not observed in the ω_2-edited COSY spectra of selectively labeled T4 lysozymes. With the exception of Ile17, Gly30, and Tyr139, these were detected in either the edited COSY or HMQC–COSY spectrum of the uniformly labeled protein or in DQF-COSY or TOSCY spectra of unlabeled protein samples. These cross peaks are all very weak, and in addition, the [^{15}N]tyrosine and [^{15}N]serine/-glycine samples were relatively dilute in isotopically labeled protein and hence gave spectra of comparatively lower signal-to-noise.

The ^{15}N-edited COSY experiment was very useful for assigning the H^{α} resonances to T4 lysozymes for several reasons:

(i) With selectively labeled proteins, most H^{α} proton signals were unambiguously identified and classified by amino acid

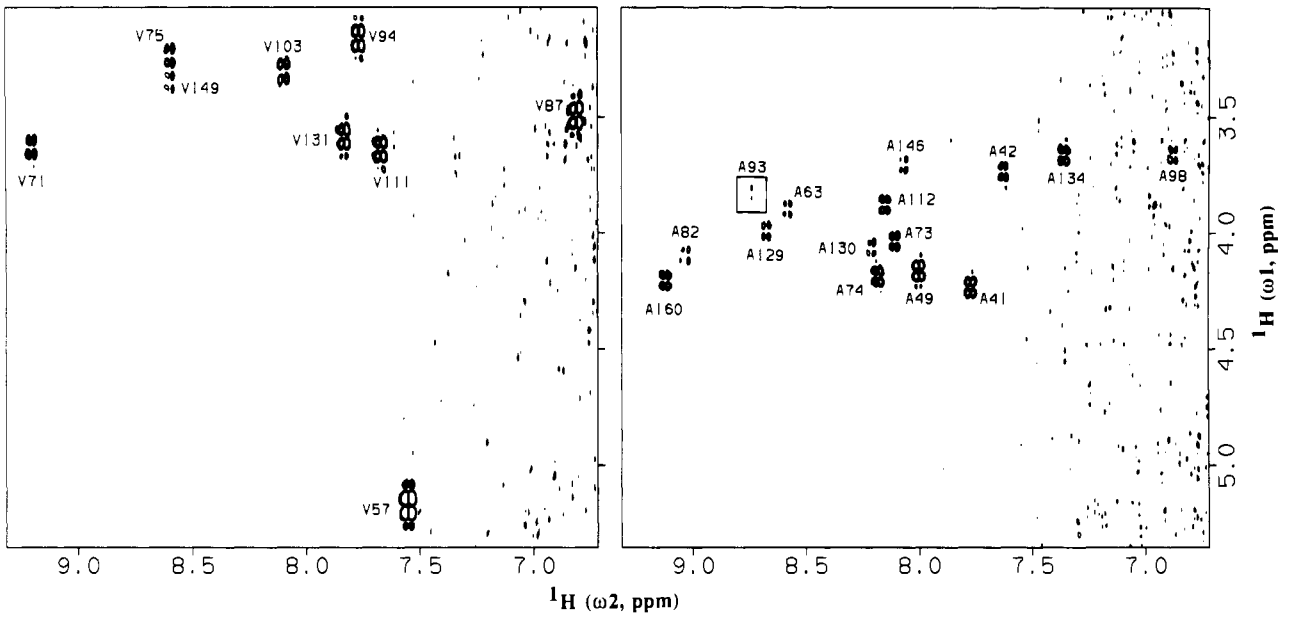

FIGURE 6: H^N–H^α correlation regions of the ^{15}N-ω_2-edited COSY spectra of T4 lysozyme selectively labeled with [^{15}N]valine (left) and -alanine (right). The cross peak of alanine 93 in the box is shown at a contour level five times lower than the full spectrum.

type. Of particular note, the edited COSY spectrum of [^{15}N]glycine-labeled lysozyme revealed at least one H^N–H^α cross peaks for 6 out of 11 glycine residues and, in three cases, both H^N–H^α and H^N–$H^{\alpha'}$ cross peaks.

(ii) This approach proved useful for sorting out crowded regions of the COSY fingerprint of T4 lysozyme and for identifying weak cross peaks and degenerate cross peaks. There are eight occurrences of two residues with overlapping H^N and H^α resonances and two cases of three residues with degenerate H^N–H^α cross peaks in the spectrum of T4 lysozyme (Figure 4). These degenerate spin systems are difficult to identify when only ^{1}H NMR experiments with unlabeled protein are used.

(iii) This experiment also has excellent elimination of the residual signal from water after presaturation, and H^α resonances near the decoupler frequency were easily detected.

(iv) The approach of recording the ^{15}N-edited COSY spectra of selectively labeled proteins was also helpful in distinguishing H^N–H^α cross peaks from arginine side-chain H^α–H^γ cross peaks. For example, an anaomalous COSY peak at 7.42 and 1.40 ppm was clearly assigned to the H^N–H^α of a lysine residue from the ^{15}N-edited COSY spectrum of [α-^{15}N]lysine-labeled protein; this was subsequently identified as Lys124. In the crystallographic model, the H^α of Lys124 is directly above the plane of the aromatic ring of Trp126, suggesting that its upfield-shifted resonance is due to a ring current effect. The sample of T4 lysozyme prepared with a medium containing [^{15}N]glutamate but lacking arginine (Table I) was also isotopically enriched at the arginine side-chain N^ϵ positions, and in this case, the H^ϵ–H^γ cross peaks were identified in the ^{15}N-edited COSY spectrum of the protein (not shown; see Figure 4).

A difficulty in this method arose when two or more labeled amino acids had similar amide proton chemical shifts. For example, two cross peaks corresponding to Asp72 and Asn132 were identified in the edited COSY spectrum of [^{15}N]Asx-labeled protein but could not be distinguished as the two residues have identical amide proton chemical shifts. Also, the cross peaks from Leu99 was obscured by the strong cross peak of Leu164 in the spectrum of [^{15}N]leucine-labeled lysozyme. In both cases, the peaks were distinguished by different amide hydrogen exchange rates, as determined by using both unlabeled and uniformly labeled protein samples. Similarly, the cross peaks from Gly28 and Gly107 were also difficult to resolve due to similar amide proton chemical shifts. The assignments were made by recording HMQC and edited COSY spectra of [^{15}N]Ser/[^{15}N]Gly-labeled protein at several temperatures.

(E) Extended Assignments of Spin Systems Using DQF-COSY and TOCSY Experiments. At this stage, we had identified all 160 backbone amide ^{1}H–^{15}N HMQC resonances by amino acid class and the corresponding H^α resonances of 157 residues. The next step toward assigning the NMR spectra of this protein involved identifying the possible H^β resonances of each residue. This process was constrained by our knowledge of the amino acid type of each ^{15}NH–H^α spin system. In principle, the H^β resonances can be identified in the DQF-COSY spectrum of a protein, but we were unable to unambiguously identify many H^α–H^β cross peaks due to extensive overlap in both the H^α and H^β spectral regions and many weak or missing cross peaks due to broad resonances. The H^β resonances of only approximately 30 residues were confidently identified solely from the DQF-COSY spectra of T4 lysozyme in H_2O and D_2O buffers.

The problems of overlapping side-chain resonances encountered with COSY experiments can often be circumvented by using RELAY and TOCSY experiments, which yield multistep scalar correlations from amide protons (Wagner, 1983). Unfortunately, due to short proton transverse relaxation times (approximately 15–20 ms for most amide protons) and many small $^3J_{NH-H\alpha}$ couplings, a RELAY spectrum of T4 lysozyme was not useful. We recorded TOCSY spectra of T4 lysozyme in H_2O and D_2O solutions using an MLEV-17 mixing sequence ranging from 25 to 70 ms in duration. The most informative spectra were recorded with a mixing period of 35 ms, and the cross peaks involving the amide protons were generally restricted to the H^α and H^β protons (Figure 7). From these spectra we confidently identified at least one H^β resonance for approximately 90 additional residues (Table II). Together with the assignments based on the DQF-COSY spectra alone, H^β resonances for approximately 120 residues in T4 lysozyme were identified.

The upfield aliphatic cross-peak region of the COSY spectra of T4 lysozyme is very complex and hence difficult to interpret.

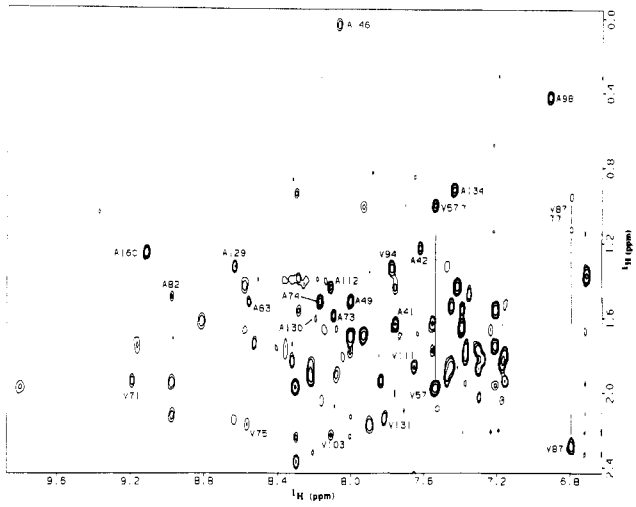

FIGURE 7: A section of a TOCSY spectrum of unlabeled T4 lysozyme showing the cross peaks between amide and side-chain protons. The observed H^N–H^β and H^N–H^γ cross peaks of the alanine and valine residues are indicated. The spectrum was recorded with a 35-ms MLEV-17 mixing period.

Also the TOCSY spectra with the longer mixing times were of poor signal-to-noise and yielded relatively few relayed cross peaks from the amide protons. As a result, we have not been able to confidently assign the complete side-chain spin systems of many residues. Of the 114 residues with two H^β protons, we have identified at least one H^β resonance for 84 residues and both H^β resonances for only 25 residues (Table II). For a few residue types, the spin systems were relatively easy to delineate from COSY and TOCSY spectra. Most notably, we have defined the complete spin systems of all 15 alanine and 9 valine residues (Figure 7) and 8 of the 11 threonine residues using the DQF-COSY and TOCSY spectra.

The spin systems of the 11 glycines in T4 lysozyme were particularly difficult to identify. The glycine H^N–H^α and H^N–$H^{\alpha'}$ cross peaks are weak due to the additional splitting caused by the passive coupling of the two H^α protons, and the cross-peak fine structure depends on the magnitudes of the $^3J_{NH\text{-}H\alpha}$ and $^3J_{NH\text{-}H\alpha'}$ coupling constants. Also, three glycine residues have similar amide proton chemical shifts under the conditions used for these experiments (Figure 1). At least one H^N–H^α cross peak from six glycine residues was observed in the edited COSY spectrum of T4 lysozyme labeled with [^{15}N]glycine. The H^N–H^α cross peaks of four additional glycines were detected in 500-MHz TOCSY and 600-MHz DQF-COSY spectra of unlabeled protein. The $H^{\alpha'}$ resonances of these 10 residues were identified by H^N–$H^{\alpha'}$ cross peaks in TOSCY and DQF-COSY spectra and by the strong H^α–$H^{\alpha'}$ cross peaks in DQF-COSY spectra. The identifications of these spin systems were also supported by edited NOESY spectra of [^{15}N]glycine-labeled T4 lysozyme.

We have recorded the ^{15}N-edited TOCSY spectra of uniformly and selectively labeled T4 lysozyme (not shown) but, overall, this did not prove very useful. One early edited TOCSY measurement of [^{15}N]Lys-labeled T4 lysozyme identified the H^α resonance of Lys16, from which several additional amide assignments were deduced using saturation transfer HMQC experiments (Lowry et al., 1987). The edited TOCSY experiment has lower signal-to-noise than the unedited experiment due to the additional time requirement of the difference echo pulse sequence. Also, unedited TOCSY spectra of the unlabeled protein were relatively straightforward to interpret due to the limited number of cross peaks involving amide protons, and the simplification offered by isotope labeling was not required.

(F) Extended Assignments of Spin Systems Using ^1H–^{15}N HMBC Experiments. Intraresidue $^{15}N_i$–H^β_i correlations can be observed in the HMBC spectra of labeled proteins, although the intensities of the cross peaks depend strongly upon the relaxation rates of the nuclei and the magnitudes of the χ_1-dependent three-bond coupling constants (Bystrov, 1976; Bax et al., 1988). We observed only a very limited number of $^{15}N_i$–H^β_i cross peaks in the HMBC spectra of uniformly ^{15}N enriched T4 lysozyme in D_2O buffer (Figure 5). Most notably, cross peaks between the amide ^{15}N and the methyl groups of all 15 alanine residues were observed. In addition, the $^{15}N_i$–H^β_i cross peaks of several threonine, asparagine, and aspartate residues were identified. These assignments are summarized in Figure 10.

We had hoped to identify backbone and side-chain resonances of residues such as histidine ($^{15}N^\delta$–H^β), proline (^{15}N–H^α, $-H^\beta$, or $-H^\delta$), asparagine ($^{15}N^{\delta 2}$–H^β), glutamine ($^{15}N^{\epsilon 2}$–H^γ), and the N-terminal methionine (^{15}N–H^α or $-H^\beta$) using the HMBC experiment. Unfortunately, cross peaks to the ^{15}N resonances of these groups were not observed in HMBC spectra of ^{15}N-labeled T4 lysozyme recorded with 50–70-ms delays. This is likely due to weak heteronuclear J couplings and the broad resonances in T4 lysozyme.

Poorly resolved ^{15}N–^{1}H cross peaks from the side chains of the arginine and lysine residues were observed in the HMBC spectra of ^{15}N-enriched T4 lysozyme (not shown). The arginine $^{15}N^\epsilon$–H^δ and $^{15}N^\epsilon$–H^γ cross peaks lay near 85 (^{15}N) and 3.2 and 1.5–1.9 ppm (^{1}H), respectively. The lysine $^{15}N^\zeta$–H^ϵ and $^{15}N^\zeta$–H^δ cross peaks lay near 33 (^{15}N) and 3.0 and 1.7 ppm (^{1}H), respectively. Due to the large number of arginine and lysine residues in T4 lysozyme, these data are of little help for specific residue assignments.

Assignment of Selected Residues

During the early stages of this research, we applied several approaches to unambiguously assign a limited number of selected ^{1}H–^{15}N resonances from T4 lysozyme. These assignments served as valuable reference points for the complete analysis of the NMR spectra of this protein.

(A) Comparison of the HMQC Spectra of Wild-Type and Variant Proteins. An extensive library of T4 lysozyme variants with defined amino acid substitutions has been developed at the University of Oregon. The amide ^{1}H and ^{15}N resonances of Tyr25 were identified from a missing cross peak in the HMQC spectrum of [^{15}N]tyrosine-labeled T4 lysozyme with the substitution of Tyr25 to Gly. Similarly, those of Trp126 were assigned from the spectrum of [^{15}N]tryptophan-labeled protein with the replacement of tyrosine for this residue. The amide resonances of Cys54 and Cys97 were also assigned from spectra of uniformly ^{15}N labeled T4 lysozyme with threonine and alanine, respectively, at these positions. We have observed extensive changes in the HMQC spectra of uniformly ^{15}N enriched T4 lysozyme variants (J.-R. Lu, L.P.M., and F.W.D., unpublished), and therefore this approach is best suited to experiments with selectively labeled samples where confusion due to small changes in the chemical shifts of the few labeled amide groups is less likely.

(B) Multiple ^{15}N and ^{13}C Isotopic Labeling. By incorporating an α-^{15}N-labeled amino acid A and an α-^{13}C-labeled amino acid B into T4 lysozyme, any dipeptides of the sequence B-A will simultaneously contain both isotopic labels (Kainoshi & Tsuji, 1982). The ^{1}H–^{15}N HMQC peak of residue A will be split or broadened due to the $^1J_{CN}$ scalar coupling. Using this approach, we assigned the amide resonances of Phe67, Phe104, Phe114, Tyr24, and Tyr88 (Dahlquist et al., 1985; Griffey et al., 1986). The amide resonances of Phe4 and

Phe153 were tentatively assigned from hydrogen exchange experiments (Griffey et al., 1986). Although this is a very powerful method to identify selected amide resonances, it has the disadvantages that many labeled protein samples are required and that the $^1J_{CN}$ coupling constant is comparable to the ^{15}N line widths in the ω_1 dimension of HMQC spectra. This latter problem may be alleviated by several approaches, including using HSMQC experiments, direct ^{15}N detection (Leighton & Lu, 1987), or ^{13}C–^{15}N HMQC experiments (Westler et al., 1988), or by comparing spectra acquired with and without ^{13}C decoupling (Campbell Burk et al., 1989).

(*C*) *One-Dimensional Isotope-Directed NOE Measurements.* We have measured the NOEs from specific amide protons in selectively ^{15}N labeled T4 lysozyme samples using the isotope-directed NOE experiment (Griffey et al., 1985b; McIntosh et al., 1987a,b). As described previously, the amide proton resonances from Asn101, Met102, Val103, and Phe104 were assigned by using this approach (Muchmore et al., 1989). We generated an extensive library of isotope-directed NOE spectra of labeled amide protons in T4 lysozyme and assigned or tentatively assigned over 30 resonances using these data. The isotope-directed NOE measurement has the advantage of correlating three or four resonance frequencies and is useful for clarifying specific identifications (McIntosh et al., 1987a). However, this approach met limited success due to the difficulty in comparing one-dimensional spectra from many labeled protein samples and we found that the two-dimensional edited NOESY experiment was more useful.

(*D*) *Saturation Transfer HMQC Experiments.* The saturation transfer HMQC experiment detects the NOE from protons at any given frequency to ^{15}N-labeled protons using the HMQC experiment as the read sequence of a standard NOE measurement. This experiment yields two-dimensional spectra that are equivalent to slices of a three-dimensional heteronuclear NOESY spectrum and requires about 2 h of recording time per spectrum. This experiment was developed to help assign the resonances of residues in extended structures by using NOEs involving the H^α protons. The amide ^1H and ^{15}N resonances of Lys16, Ile17, Val57, and Ile58 in a β-sheet structure in T4 lysozyme were assigned by using this approach to observe the NOEs from the H^α of Lys16 (Lowry et al., 1988).

(*E*) *Hydrogen Exchange Experiments.* Amide protons in proteins have different hydrogen exchange kinetics, depending on factors including hydrogen bonding, solvent accessibility, and local electrostatic environment (Englander & Kallenbach, 1984). Several amides in T4 lysozyme have very slow exchange rates at pH* 5.6–6.5 and remain protonated after a year in D_2O buffer at 5 °C. These were assigned to residues Ala98 through Phe104 and Thr152 through Arg154 by two-dimensional ^1H–^1H- and ^1H–^{15}N-edited NOESY measurements of uniformly ^{15}N labeled T4 lysozyme (McIntosh et al., 1987a; Muchmore et al., 1989). These residues comprise two α-helices and form the hydrophobic core of the C-terminal lobe of the protein (Weaver & Matthews, 1987).

Specific Residue Assignments

(*A*) *Specific Assignments from ^{15}N-Edited NOESY Experiments.* At this stage we had identified the ^{15}NH–H^α–H^β spin systems of most residues in T4 lysozyme by amino acid class and had assigned a limited number of these to specific positions along the polypeptide backbone. The next step in this task involved the assignment of these spin systems to specific residues based on the nearest-neighbor NOE connectivities (Wüthrich, 1986). T4 lysozyme contains almost 1200 protons and its full NOESY spectrum is a formidable challenge to interpret due to the myriad of overlapping cross peaks. We focused only on the NOEs involving amide protons (H^N_i–H^N_{i+1}, H^α_i–H^N_{i+1}, H^β_i–H^N_{i+1}, H^α_i–H^N_{i+3}), using ^{15}N-edited NOESY measurements of selectively labeled samples to identify the cross peaks to the H^N protons of individual residues of known amino acid type.

Similar to the ^{15}N-ω_2-edited COSY experiment, the ^{15}N-ω_2-edited NOESY measurement yields an asymmetric two-dimensional spectrum with cross peaks arising due to dipolar interactions from all possible protons to only ^{15}N-bonded protons. This experiment proved invaluable for completely assigning the backbone resonances of T4 lysozyme for many reasons:

(i) The NOEs to virtually every amide proton were measured from the ^{15}N-edited NOESY spectra of T4 lysozyme selectively α-^{15}N labeled with Ala, Arg, Asx, Gly, Glx, Ile, Leu, Met, Phe, Thr, Tyr, Phe, or Val and of T4 lysozyme with the mutation Trp126 to Tyr labeled with [^{15}N]Trp. In Figure 8, the spectra of the proteins labeled with [^{15}N]alanine and -valine are presented. Of the 156 amide protons from these 14 amino acid types, the NOESY cross peaks to 145 were clearly identified. Only 11 amide protons were degenerate in amide proton chemical shift with a second labeled amide proton and hence the NOESY cross peaks were not unambiguously identified. Of the remaining four residues in T4 lysozyme that were not isotopically labeled, the amide proton resonances of Cys54, Cys97, and Trp126 were clearly resolved in the edited NOESY spectrum of uniformly ^{15}N labeled protein, and accordingly, NOEs to these protons were readily identified. The amide proton of His31 was not resolved and potential NOEs to this proton could not be unambiguously identified.

(ii) A major difficulty in analyzing the NOESY spectra of larger proteins is to resolve and unambiguously assign cross peaks between two interacting protons. With the ^{15}N-edited NOESY experiment we have resolved the NOEs to amide protons of known amino acid type by observing only the NOEs to a limited number of labeled amides. In most cases, the resonances of labeled protons in the edited spectra of the selectively enriched lysozymes are sufficiently disperse that each was located at a distinct column along the diagonal of the two-dimensional map and all cross peaks along that column were confidently ascribed to that amide proton (Figure 8). This permitted us to dissect crowded regions of the complete NOESY spectrum of T4 lysozyme and to avoid possible difficulties in aligning cross peaks. This was particularly useful when a weak, yet important, cross peak was otherwise obscured by strong, nearby cross peaks, as we frequently observed in the crowded H^N–H^α region of the NOESY spectra of unlabeled T4 lysozyme. Unfortunately, using this approach, we are generally not able to unambiguously assign the side-chain resonances, which give NOESY cross peaks to the labeled amides, to specific protons due to incomplete resonance assignments and severe overlap in the upfield region of the ^1H NMR spectrum of T4 lysozyme. This will require further side-chain assignments and an analogous approach such as ^1H–^{13}C heteronuclear NMR (Fesik et al., 1990; Zuiderweg et al., 1990).

(iii) The ^{15}N-edited NOESY spectrum of uniformly ^{15}N labeled T4 lysozyme is similar to the downfield region (along ω_2) of a protein NOESY spectrum without the confusing overlap of NOE cross peaks to aromatic protons. Additionally, the NOEs between two labeled protons are observed as cross peaks that are symmetrically disposed about the diagonal, whereas NOEs from unlabeled protons to ^{15}N-labeled protons

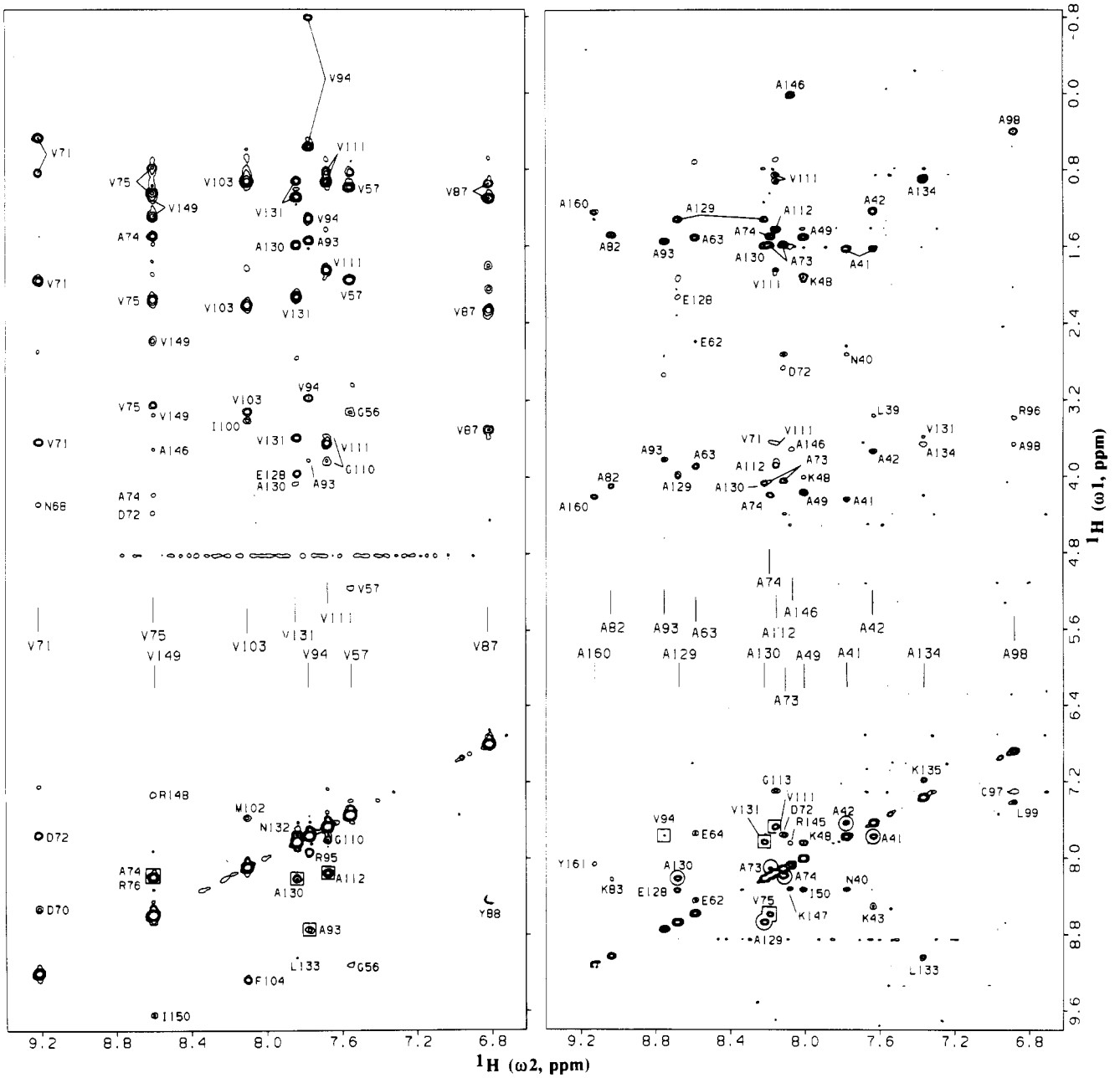

FIGURE 8: ^{15}N-ω_2-Edited NOESY spectra of T4 lysozyme selectively labeled with [^{15}N]valine (left) and -alanine (right). The amide NH shifts of the valine and alanine residues are indicated by the vertical lines near the middle of the spectra and the assignments of the resonances that give NOEs to these amides are presented beside the cross peaks. The three H^N_i–H^N_{i+1} NOE cross peaks between sequential alanine residues are enclosed by circles in the spectrum on the right. The four complementary H^N_i–H^N_{i+1} NOE cross peaks between adjacent valine and alanine residues are enclosed by boxes. By comparison to the edited COSY spectra shown in Figure 4, the intraresidue H^N_i–H^α_i cross peaks can be identified.

are seen as single cross peaks that lack any symmetric mate (McIntosh et al., 1987a). This proved very useful in distinguishing amide–amide proton NOEs from amide–aromatic NOEs.

The basic strategy that we employed to assign the backbone resonances of T4 lysozyme is summarized as follows:

(i) We recorded the ^{15}N-ω_2-edited NOESY spectra of uniformly and selectively labeled T4 lysozyme. The NOESY spectra were recorded with a mixing time of 100 ms, for which the H^N_i–H^N_{i+1} cross peaks of residues in helical regions (ca. 2.8 Å separation) and the intraresidue H^α_i–H^N_i cross peaks (maximum distance of 2.9 Å) of most residues are observed as strong peaks, while many interresidue H^α_i–H^N_{i+1} cross peaks [maximum distance of 3.6 Å; Wüthrich et al., (1984)] are observed as weak peaks. In Figure 8, the ^{15}N-edited NOESY spectra of T4 lysozyme labeled with [^{15}N]alanine and -valine are presented, and in Figure 9, the downfield portion of the spectrum of uniformly labeled protein is shown.

(ii) The ^{15}N-edited NOESY spectra of the selectively labeled T4 lysozymes were compared to the corresponding HMQC and ^{15}N-edited COSY spectra to determine the ^{15}N chemical shift and to identify the intraresidue H^α_i–H^N_i cross peaks of each labeled residue (compare Figures 3, 6, and 8 for the alanine residues). The NOESY data were also used to help identify or confirm the identification of the resonances of the H^β and any additional protons in the spin system of each residue.

(iii) The ^{15}N-edited NOESY spectra from each selectively labeled sample were mapped on the ^{15}N-edited NOESY spectrum of T4 lysozyme uniformly enriched in ^{15}N (Figure 9) by simply overlaying plots of the various spectra. This assimilated the data from many different protein samples into

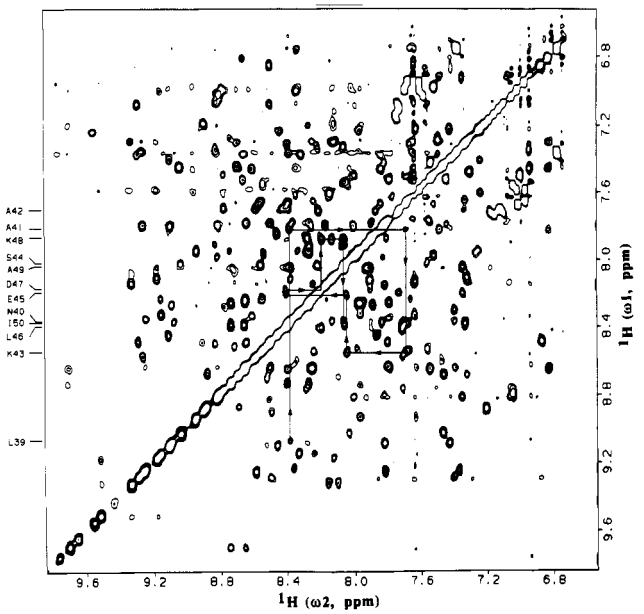

FIGURE 9: Amide region of the ^{15}N-ω_2-edited NOESY spectra of uniformly ^{15}N labeled T4 lysozyme, recorded with a mixing time of 100 ms. The H^N_i–H^N_{i+1} connectivities joining residues 39–50 are indicated.

a single data set, thereby allowing careful alignment of the NOESY cross peaks.

(iv) Using this master ^{15}N-edited NOESY spectrum, as well as additional NOESY spectra of T4 lysozyme in H_2O and D_2O solutions recorded at 500 and 600 MHz, we assigned the backbone proton resonances to specific residues on the basis of NOESY connectivities to amide protons of known amino acid type. Figure 10 summarizes the observed H^N_i–H^N_{i+1}, H^α_i–H^N_{i+1}, H^β_i–H^N_{i+1}, and H^α_i–H^N_{i+3} NOEs that led to these conclusions. The initial assignments of several residues, discussed previously, were useful as starting points for interpreting the NOESY data.

We focused our attention on primarily the amide H^N_i–H^N_{i+1} cross peaks, as most of these were unambiguously identified in *both* dimensions by amino acid type. That is, when a particular H^N_i–H^N_{i+1} cross peak was assigned as an NOE to a single proton of amino acid type A from the ^{15}N-edited NOESY spectrum of T4 lysozyme ^{15}N labeled at each A residue and the corresponding symmetry-related peak was likewise assigned as an NOE to a single amide proton of residue type B, this led us to identify the peak as being between the amide protons of an A- and a B-type residue. This is illustrated in Figure 8 for [^{15}N]alanine- and -valine-enriched T4 lysozymes. The H^N_i–H^N_{i+1} NOEs from each of the three alanine–alanine dipeptides are observed as symmetry-related cross peaks about the diagonal of the ^{15}N-edited NOESY spectrum of [^{15}N]Ala lysozyme because all of the alanine amide protons are isotopically labeled (these assignments were distinguished from coincidental degeneracies by using a dou-

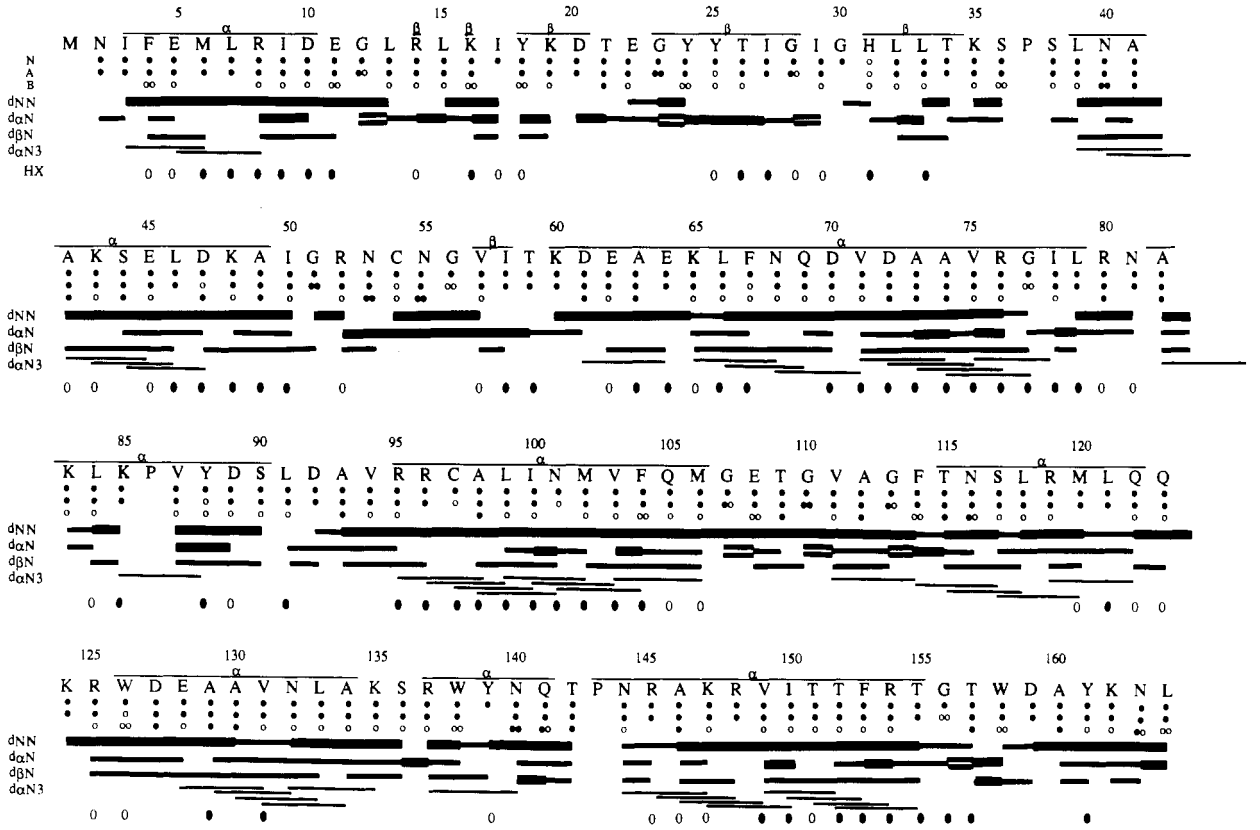

FIGURE 10: Summary of the assignment of the backbone resonances of T4 lysozyme. The sequence and secondary structure as defined by the crystallographic model are indicated on the top of each section. The three rows of dots labeled NAB indicate the extent of the resonance assignment with the first row referring to ^{15}N–^{1}H, the second to H^α, and the third to H^β. Solid dots show that the assignment was deduced from HMQC (^{15}N–^{1}H), ω_2-edited COSY (H^α), or HMBC (H^β) spectra of a uniformly or selectively labeled protein sample and open dots indicated that the assignment was from unedited COSY or TOCSY data. The next four rows summarize the NOE connectivities used to assign the backbone resonances. These connectivities are the H^N_i–H^N_{i+1} (d_{NN}), H^α_i–H^N_{i+1}, ($d_{\alpha N}$), H^β_i–H^N_{i+1} ($d_{\beta N}$), and H^α_i–H^N_{i+3} ($d_{\alpha N3}$) interactions. For the d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$ connectivities, thick bars indicate that the assignment was unambiguously assigned to both protons whereas thin bars indicate that the assignment was consistant but not unambiguously assigned due to possible degeneracies. For $d_{\alpha N3}$, this was not distinguished. The absence of a bar connecting protons of neighboring residues does not necessarily imply that there is no observed NOE. The last row of ovals summarized the hydrogen exchange kinetics of the amide protons at pH 5.6. An open oval indicates that the amide was still protonated after 10 h in D_2O at room temperature and a solid oval indicates that the amide was still protonated after 3 months in D_2O at 5 °C. Residues without an oval are completely deuterated within 10 h after transfer to the D_2O buffer at room temperature.

ble-edited NOESY experiment in which both the first and last ^1H 90° pulses are replaced by difference echo pulse sequences, and hence, only cross peaks between labeled protons are detected). As a more general example, the H^N_i–H^N_{i+1} NOEs from each of the four alanine–valine dipeptides are detected as cross peaks on only one side of the diagonal in each of the spectra shown in Figure 8 and were identified as complementary NOEs only after the two data sets were compared. In several cases, amide–amide NOESY cross peaks were nearly overlapping or degenerate with other peaks when mapped onto the spectrum of the uniformly labeled protein, and assignments were made by considering both additional NOE interactions and the constraint of the primary structure of T4 lysozyme.

In T4 lysozyme, there are 156 sequential pairs of amino acids, excluding proline residues. Of these, 83 pairs are unique in the sequence of T4 lysozyme, and thus approximately half of the residues in T4 lysozyme could potentially be assigned by pairwise comparison of the ^{15}N-edited NOESY spectra of selectively labeled proteins (we consider glutamine and glutamate as a single residue type and assume that the orientation of the amino acid pair could also be determined, for example, from H^α_i–H^N_{i+1} or H^β_i–H^N_{i+1} NOEs). The amino acid pairs or triplets that are not unique in the primary sequence of the protein were specifically assigned by using NOE connectivities to distinct flanking amino acids. In total, almost 130 amino acids were unambiguously linked on the basis of nearest-neighbor H^N_i–H^N_{i+1} NOEs (Figure 10). T4 lysozyme has a high content of helical secondary structure (Weaver & Matthews, 1987) and many diagnostic stretches of strong H^N_i–H^N_{i+1} NOEs are observed in its NOESY spectra (Wüthrich et al., 1984). In Figure 9, the assignments of the amide proton resonances of residues Leu39 through Ile50 are outlined.

Specific assignments were also derived from NOESY cross peaks between amide protons and H^α and H^β protons. In a limited number of cases, the H^α proton resonances did not overlap with any other resonances and thus H^α_i–H^N_{i+1} NOEs could be unambiguously identified between two residues. Several sequences in extended secondary structures were assigned on the basis of these observed NOEs (Figure 10). In general, however, the aliphatic protons were highly degenerate and incompletely assigned, and thus H^α_i–H^N_{i+1}, H^α_i–H^N_{i+3}, or H^β_i–H^N_{i+1} NOEs involving an amide proton could only be tentatively ascribed to a specific H^α or H^β proton. These NOESY connectivities were used to support assignments based on amide–amide interactions.

(*B*) *Sequential Assignments from* ^1H-^{15}N *HMBC Experiments.* Sequential assignment can also be obtained by exploiting the ψ_i-dependent interresidue $^{15}N_i$–H^α_{i-1} scalar coupling (Bystrov, 1976; Bax et al., 1988a). A limited number of $^{15}N_i$–H^α_{i-1} cross peaks were observed in the HMBC spectra of uniformly ^{15}N enriched T4 lysozyme (Figure 5). These have been tentatively assigned to I9-N10, N40-A41, A41-A42, S44-E45, K48-A49, A49-I50, Q69-D70, V87-Y88, D89-S90, T109-G110, G113-F114, N116-T117, D127-E128, N132-L133, N140-Q141, Q141-T142, and A160-Y161. From the crystallographic model of T4 lysozyme (Weaver & Matthews, 1987), all of these pairs have ψ_i ranging between −18 and −48° and are expected to have $|^3J_{NH}|$ greater than 3.5 Hz (Bystrov, 1976). These HMBC assignments were obtained by using the assignments deduced from the NOESY spectra of T4 lysozyme and serve to confirm the results presented in Table II and Figure 10. In general the cross peaks in the HMBC spectra of T4 lysozyme were not umambiguously assigned due to the limited number of observed cross peaks, degeneracies of the ^{15}N or ^1H chemical shifts, and the overlap of interresidue and intraresidue cross peaks.

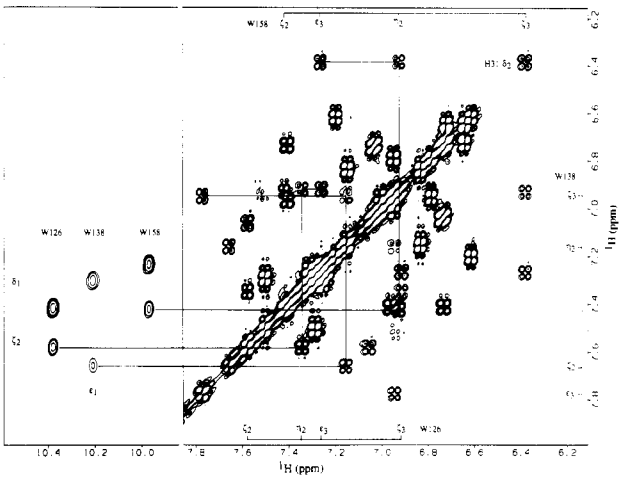

FIGURE 11: Assignment of the aromatic proton resonances of the three tryptophan residues in T4 lysozyme using NOESY (left) and DQF-COSY (right) spectra recorded at 600 MHz. Additional cross peaks from the phenylalanine and tyrosine aromatic and side-chain amide protons are in the spectrum but are not specifically labeled.

Assignment of the Proton Resonances from the Aromatic Side Chains

(*A*) *Histidine.* The $H^{\delta 2}$ and $H^{\epsilon 1}$ proton resonances of the single histidine residue in T4 lysozyme were identified at 6.43 and 8.70 ppm, respectively, from a weak cross peak in the DQF-COSY spectrum of the protein. The $H^{\delta 2}$ proton also displays NOEs to the H^α and H^β protons assigned to His31. The imidazole ring nitrogens of this residue have ^{15}N chemical shifts at 188.0 and 174.4 ppm, as shown in Figure 1; these have been tentatively assigned to $N^{\delta 1}$ and $N^{\epsilon 2}$, respectively, using HMQC and HMBC spectra (Stockman et al., 1989). The cross peak between $H^{\epsilon 1}$ and $N^{\delta 1}$ is significantly weaker than between this proton and $N^{\epsilon 2}$ (Blomberg et al., 1977). The pK_a of His31 is approximately 9.1 (Anderson et al., 1990), and both the proton and nitrogen chemical shifts of the imidazole ring indicated that the histidine is in its acidic form at the pH of these measurements. The 14 ppm downfield shift of one imidazole nitrogen resonance ($N^{\delta 1}$) relative to the other suggests that the nitrogen may be involved in a hydrogen bond (Bachovchin, 1986). His31 and Asp70 form a strong salt bridge, as shown by crystallographic, thermodynamic, and NMR studies (Weaver & Matthews, 1987; Anderson et al., 1990). From the crystallographic data, the salt bridge has been modeled to involve the $N^{\delta 1}$ of the histidine ring.

(*B*) *Tryptophan.* T4 lysozyme has three tryptophan residues at positions 126, 138, and 158. Previously, we assigned the indole $H^{\epsilon 1}$ proton resonances of these three side chains to resonances at 10.38, 10.21, and 9.97 ppm, respectively, by selective tryptophan to tyrosine substitutions. The $H^{\delta 1}$ protons of the indole rings were assigned from $H^{\epsilon 1}$–$H^{\delta 1}$ cross peaks in COSY spectra and the $H^{\zeta 2}$ protons were identified from the $H^{\epsilon 1}$–$H^{\zeta 2}$ NOESY cross peaks. Finally, by use of DQF-COSY and TOCSY experiments, the resonances from the four adjacent protons of each tryptophan ring were identified. These assignments are illustrated in Figure 11.

The identification of the indole proton and nitrogen resonances by mutational substitutions were completely independent of the backbone resonances assignments by isotope-aided experiments. In all three cases, NOEs between the $H^{\epsilon 3}$ and $H^{\delta 1}$ ring protons with the H^α and H^β protons were ob-

served, confirming the assignments of these spin systems (not shown).

(*C*) *Tyrosine.* In addition to three tryptophans, T4 lysozyme has five phenylalanine and six tyrosine residues. The DQF-COSY cross peaks of all these aromatic spin systems are present in Figure 11. To confidently assign the proton resonances of the phenylalanine and tyrosine side chains, we prepared samples of T4 lysozyme with selectively deuterated aromatic rings (LeMaster, 1989). The amide protons of these protein samples were also partially or completely exchanged for deuterons to avoid confusing overlap of backbone and side-chain amide cross peaks in the COSY and NOESY spectra.

The resonances of the H^{δ} and H^{ϵ} ring protons of the tyrosine residues were identified from H^{δ}–H^{ϵ} cross peaks in the DQF-COSY spectrum of T4 lysozyme in which the tryptophan and phenylalanine aromatic rings were deuterated (Figure 12). The H^{δ} and H^{ϵ} resonances were distinguished by using spectra recorded from T4 lysozyme labeled with [$\epsilon_{1,2}$-$^2H_2$]tyrosine. Somewhat surprisingly, there is only one H^{δ}–H^{ϵ} cross peak observed for each tyrosine residue, indicating that the $H^{\delta 1}$ and $H^{\delta 2}$ and the $H^{\epsilon 1}$ and $H^{\epsilon 2}$ proton chemical shifts are each equivalent, possibly due to ring flipping, which is rapid on the experimental time scale.

The six tyrosine ring spin systems were specifically identified by NOE connectivities to the backbone H^{α} and H^{β} protons. The backbone resonance of tyrosine residues 18, 24, 25, 88, and 161 were identified previously by selective ^{15}N labeling of T4 lysozyme and the NOE connectivities between the H^{α} and H^{β} protons and the ring H^{δ} (and H^{ϵ}) protons were used to specifically assign the ring spin systems of these five residues (Figure 12). The remaining tyrosine ring was assigned to Y139 by default. We have not confidently identified the backbone proton resonances of this residue. The most downfield signal in the ^{1}H NMR spectrum of T4 lysozyme at 11.3 ppm is assigned as the phenolic H^{η} proton resonance of Tyr161 by NOE connectivities.

(*D*) *Phenylalanine.* Complementary to the assignment of the tyrosine ring proton resonances, we assigned the phenylalanine ring spin systems of T4 lysozyme from spectra recorded with protein labeled with deuterated tyrosine and tryptophan. The protein was in D_2O buffer, allowing partial exchange of the amide protons. The H^{δ}, H^{ϵ}, and H^{ζ} proton resonances of four phenylalanine rings were identified by DQF-COSY and TOCSY experiments. The $H^{\delta 1}$ and $H^{\delta 2}$ and the $H^{\epsilon 1}$ and $H^{\epsilon 2}$ proton resonances of these four aromatic rings were equivalent. By use of NOESY connectivities, these aromatic rings were assigned to phenylalanine residues 4, 67, 114, and 153 (Figure 12). We were unable to identify the resonances from the aromatic ring protons of Phe104 in DQF-COSY experiments. However, a phenylalanine ring proton resonating at 7.31 ppm was identified in this deuterated protein sample and was assigned to Phe104 from NOESY experiments. The side chain of this residue is buried in the interior of the protin (Weaver & Matthews, 1987) and may have broad aromatic proton resonances due to its motional or environmental properties.

Amide Hydrogen Exchange

A major goal of our studies of T4 lysozyme is to probe the structure and dynamics of this protein using NMR to measure the kinetics of amide hydrogen exchange. In previous studies and over the course of this work, we have recorded spectra of T4 lysozyme stored in D_2O buffers for times ranging from hours to a year (McIntosh et al., 1987a; Muchmore et al., 1989). Amide protons were distinguished by their differential exchange rates and this was very useful for assigning overlapping and crowded cross peaks. In Figure 10, a qualitative summary of the exchange rates of the backbone amide protons in T4 lysozyme is presented. Further details of these studies will be reported elsewhere.

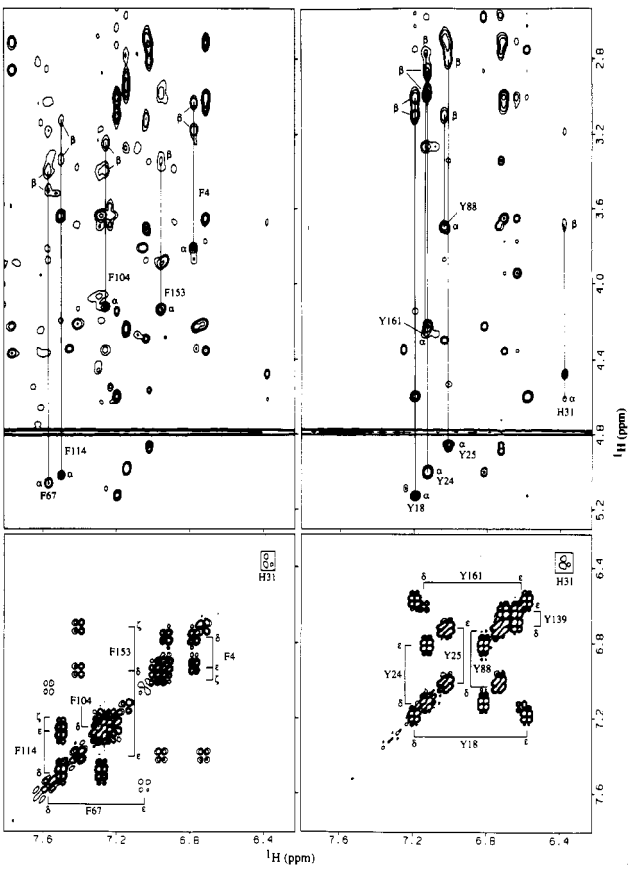

FIGURE 12: Assignment of the phenylalanine (left) and tyrosine (right) aromatic proton resonances of T4 lysozyme using DQF-COSY (bottom) and NOESY (top) spectra recorded at 500 MHz. To assign the phenylalanine resonances, T4 lysozyme was prepared with [δ_1,ϵ_2,$\zeta_{2,3}$,η_2-$^2H_5$]tryptophan and [$\epsilon_{1,2}$-$^2H_2$]tyrosine and was transferred to D_2O buffer. The protein was not completely unfolded and many protected amide groups are protonated. The aromatic spin systems are identified in the DQF-COSY spectrum, and the NOE interactions between the ring H^{δ} protons and the backbone H^{α} and H^{β} protons used to specifically assign the residues are shown in the NOESY spectra as indicated. To assign the tyrosine residues, T4 lysozyme was prepared with [δ_1,ϵ_2,$\zeta_{2,3}$,η_2-$^2H_5$]tryptophan and [$\delta_{1,2}$,$\epsilon_{1,2}$,ζ-$^2H_5$]-phenylalanine and the protein was unfolded in D_2O buffer to completely deuterate the exchangeable sites. The spin systems are identified in the DQF-COSY spectra. The H^{δ} and H^{ϵ} protons were distinguished by comparison with the spectra of the protio-phenylalanine sample, which contained [$\epsilon_{1,2}$-$^2H_2$]tyrosine. The tyrosine rings were specifically assigned by the NOE interactions between the ring H^{δ} protons and the backbone H^{α} and H^{β} protons as indicated. The His $H^{\delta 2}$ diagonal peak is shown in a box at a 5 times lower contour level in the DQF-COSY spectra, and the backbone H^{α} and H^{β} proton is labeled in the NOESY spectra.

DISCUSSION

We have assigned the proton and nitrogen (^{15}NH–H^{α}–H^{β}) resonances of bacteriophage T4 lysozyme using ^{15}N-aided ^{1}H NMR. These assignments were directed from the ^{1}H–^{15}N nuclei of each amide in T4 lysozyme, with the HSMQC spectrum of the uniformly enriched protein serving as the master template for this work. By extensive selective labeling of T4 lysozyme, the amide ^{15}N–^{1}H resonances were classified as 1 of 18 amino acid types. Protons coupled to the isotopically tagged protons by scalar or dipolar interactions were identified by ^{15}N-edited COSY and NOESY experiments and, to a lesser extent, HMQC–COSY, HMBC, and saturation transfer

HMQC experiments. The isotope-aided data were complemented by proton DQF-COSY, TOCSY, and NOESY data of unlabeled lysozyme. Spectra were also recorded with protein samples in H_2O and D_2O buffers to differentiate amide protons by hydrogen exchange. We confidently assigned the backbone resonances of specific residues in T4 lysozyme using NOE connectivities between protons in $^{15}NH-H^{\alpha}-H^{\beta}$ units (Figure 10) and by considering the constraints imposed by selective isotopic labeling. The resonances from several residues were also unambiguously assigned by a variety of techniques such as multiple $^{15}N/^{13}C$ labeling and mutational substitution.

In summary, we identified the $^1H-^{15}N$ resonances of all 160 backbone amides, the H^{α} resonances of 157 residues, and at least one H^{β} resonance of approximately 120 residues in T4 lysozyme. A limited number of additional side-chain assignments, particularly of the valine and threonine residues, were derived from *J*-correlated spectra of unlabeled lysozyme. Assignments of the aromatic side chains of the histidine, tryptophan, tyrosine, and phenylalanine amino acids were obtained by two-dimensional ^1H NMR studies of unlabeled and selectively deuterated protein samples. These data are compiled in Table II.

Assignment Strategy

The sequential assignment method pioneered by Wüthrich and co-workers (Wagner & Wüthrich, 1982; Wüthrich, 1983, 1986) has become well established for interpreting the ^1H NMR spectra of small proteins. Unfortunately, the assignment of the NMR spectra of macromolecules becomes dramatically more difficult with increasing molecular weight. Nevertheless, the NMR spectra of a growing number of proteins in the size range of 10–20 kDa are being interpreted by a variety of elegant approaches. These include assignment algorithms such as the main-chain-directed strategy (Englander & Wand, 1987; Wand et al., 1989; Feng et al., 1989); NMR methods such as ^1H and ^{13}C multiple-quantum NMR experiments (Dalvit & Wright, 1987; Oh et al., 1988) and two- and three-dimensional heteronuclear NMR experiments (Marion et al., 1989a,b; Stockman et al., 1989; Zuiderweg & Fesik, 1990); and preparation of isotopically labeled samples using hydrogen exchange (Redfield & Dobson, 1988), random fractional deuteration (LeMaster & Richards, 1988; Wang et al., 1990a), and uniform or selective ^2H, ^{13}C, or ^{15}N enrichment (Torchia et al., 1988a,b, 1989; Campbell Burk, et al., 1989; Wang et al., 1990b). Many of these approaches have been discussed in detail recently (Oppenheimer & James, 1989).

The strategy we applied to assign the NMR spectra of T4 lysozyme utilized uniform and selective ^{15}N labeling to identify $^{15}NH-H^{\alpha}$ nitrogen and proton resonances by amino acid class and to resolve and assign proton COSY and NOESY cross peaks involving these labeled amide protons. We have previously discussed the methodology for isotopic labeling of T4 lysozyme using bacterial expression systems (Muchmore et al., 1989; McIntosh & Dahlquist, 1990). There are several advantages to using this approach to assign the NMR spectra of larger proteins [also see Senn et al. (1987b) and Torchia et al. (1989)]. These include the following.

(i) In order to assign the NMR spectrum of a protein, it is very useful to have a reference spectrum from which at least one resonance of each residue can be identified and from which extended assignments can be drawn. The H(S)MQC spectrum of uniformly ^{15}N labeled T4 lysozyme is a complete fingerprint of the protein, yielding an $^1H-^{15}N$ cross peak for every (non-proline) backbone amide, as well as for many side chains [Figure 2; McIntosh et al. (1987) and Muchmore et al. (1989)]. Large, uniform $^1J_{NH}$ couplings, the lack of overlap from the signal of water, good spectral dispersion, and high sensitivity all contribute to the usefulness of this experiment. Even for a protein the size of T4 lysozyme, there are very few degeneracies in the H(S)MQC spectrum. In contrast, the conventional COSY H^N-H^{α} fingerprint of T4 lysozyme (Figure 4) has several degenerate cross peaks, and many weak or missing cross peaks due to cancellation of the antiphase quartets of broad proton resonances and due to preirradiation at the resonance of water. This COSY spectrum is particularly difficult to analyze as T4 lysozyme is primarily an α-helical protein (Weaver & Matthews, 1987) and has limited H^{α} resonance dispersion (Pardi et al., 1984; Szilágyi & Jardetzky, 1989) and many small $^3J_{NH-H\alpha}$ couplings (Pardi et al., 1983).

(ii) The assignment protocol is "amide directed", focusing on the $H^N-H^{\alpha}-H^{\beta}$ resonances. As has been discussed previously (Englander & Wand, 1987; Chazin & Wright, 1987), the amide region of the ^1H NMR spectrum of a protein is simpler and more disperse than the diverse upfield region and hence it is relatively straightforward to identify the linear $H^N-H^{\alpha}-H^{\beta}$ spin systems by COSY and relayed COSY experiments (Wagner, 1983). We have identified the $^1H-^{15}N$ and H^{α} resonances of virtually all the residues and at least one H^{β} resonance of approximately 75% of the residues in T4 lysozyme. Residue-specific assignments and secondary structure analysis are made using NOE connectivities between H^N, H^{α}, and H^{β} protons (Billeter et al., 1982; Wüthrich et al., 1984; Englander & Wand, 1987).

(iii) By use of selectively labeled T4 lysozyme samples, the $^1H-^{15}N$ and H^{α} resonances in HMQC and ^{15}N-edited COSY spectra were resolved and unambiguously identified by 1 of 18 amino acid classes. In the sequential assignment method, the complex side-chain regions of *J*-correlated spectra must be extensively analyzed in order to identify the residue spin systems of a protein and to classify these into categories of amino acid types depending upon the observed connectivity patterns (Wüthrich, 1983, 1986). In the case of larger proteins such as T4 lysozyme, which has almost 1200 protons in 164 residues, this analysis is exceedingly difficult. With selective isotope labeling, this problem is circumvented, as ^{15}NH and H^{α} proton resonances are unambiguously identified by amino acid class. Also, further analysis of the side-chain resonances of each residue is facilitated by the knowledge of its amino acid type. The amino acid composition of T4 lysozyme was confirmed by selective labeling.

(iv) It is increasingly difficult to analyze the two-dimensional NOESY spectra of larger proteins for many reasons. With a growing number of protons having broader resonances, fewer NOESY cross peaks are resolved. Even when they are resolved, it is difficult to unambiguously assign a cross peak to a pair of interacting protons because there are many possible degeneracies in one or both dimensions. One solution is to record NOESY spectra of protein samples under various conditions, such as pH or temperature, to differentially move proton resonances. This is unlikely to help assign cross peaks in dense regions of the spectra. As demonstrated in this study, using the ^{15}N-edited NOESY experiment with selectively labeled T4 lysozyme samples, cross peak to amide protons are resolved and assigned to amides of known amino type (Figure 8). The problem of resolution and assignment is overcome in the amide dimension by selective labeling (except, for example, when the labeled residues have degenerate amide proton resonances). In addition, by identifying the amide protons by amino acid class, the possible assignments of resonances using NOESY data become highly constrained. Many specific

identifications can be deduced simply from pairwise NOE connectivities, given knowledge of the sequence of the protein. Residues may also be assigned by "default" if all other members of an amino acid class are confidently account for.

The ^{15}N-edited NOESY experiment is ideally suited for assigning residues, such as those in helical and turn secondary structures, which produce strong H^N_i–H^N_{i+1} NOEs. Generally, we could not unambiguously identify the side-chain protons that interact with the labeled amide protons using this approach, as we have not completely assigned the side-chain resonances and because these resonances are highly degenerate. Accordingly, the ^{15}N-edited NOESY experiment is less suited for assigning residues, such as those in extended secondary structures, on the basis of H^α_i–H^N_{i+1} or H^α_i–H^α_j NOEs. In principle these latter problems can be overcome by analogous ^{13}C isotopic labeling and ^1H–^{13}C heteronuclear experiments (Bax & Weiss, 1987; Torchia et al., 1989; Zuiderweg et al., 1990; Fesik et al., 1990) or by deuteration of the protein (LeMaster, 1989).

(vi) With ^{15}N labeling, the ^{15}N and ^1H resonances of nitrogen-containing side chains can be identified. Although we have not completely assigned these resonances, this information is useful for studies of proteins, such as for the investigation of salt bridges and hydrogen bonding.

This assignment strategy has disadvantages such as the need to prepare and record spectra of many isotopically labeled proteins, as well as the requirement to carefully maintain identical experimental conditions in order to reliably compare data. It was not overly expensive or time consuming to obtain the data presented in this paper due to the efficiency of the T4 lysozyme expression system and the ease of purification of the protein. However, this is unlikely to hold true for many biologically interesting molecules. Also, we have not assigned any resonances from the N-terminal methionine or the three proline residues in T4 lysozyme as these do not have protonated amide nitrogens. Although the ^{15}N resonances of these groups are directly observed in the ^{15}N NMR spectrum of labeled T4 lysozyme (Figure 1), we did not identify proton resonances from the N-terminal methionine or proline residues using the HMBC experiment.

An alternative approach to selective isotopic labeling involves studying a single uniformly ^{15}N labeled protein sample with recently developed heteronuclear three-dimensional NMR experiments (Fesik & Zuiderweg, 1988, 1989; Marion et al., 1989a,b; Kay et al., 1989; Zuiderweg & Fesik, 1990). These experiments resolve proton–proton scalar and dipolar interactions by the ^{15}N shift of directly bonded nitrogen-15 nuclei. We have recently measured the three-dimensional [^1H–^1H–^{15}N] TOCSY–HMQC and NOESY–HMQC spectra of uniformly ^{15}N labeled T4 lysozyme (Fesik & Zuiderweg, 1990). Similar to the edited two-dimensional experiments used for this work, the three-dimensional experiments resolve TOCSY and NOSEY cross peaks to the labeled protons, albeit at lower resolution, and can be interpreted to assign the resonances of many protons. However, these experiments generally do not identify the amide protons by specific amino acid type. The constraints imposed by selective isotopic labeling lend great confidence to the assignments reported herein. It seems likely that a combination of judiciously chosen selectively labeled protein samples along with a single uniformly labeled protein will be very useful for the assignment of the NMR spectra of increasingly larger proteins.

The next major step in analyzing the NMR spectra of T4 lysozyme is to extend these backbone assignments to the complete side chains. From the DQF-COSY and TOCSY spectra of T4 lysozyme we have thus far identified the full spin systems of many residues (Table II) including the alanines, valines, threonines, tryptophans, histidine, and several with AMX spin systems (Wüthrich, 1986). Using selectively deuterated aromatic amino acids, we also assigned the aromatic rings of the phenylalanines and tyrosines in T4 lysozyme. Further analysis of the COSY spectra is very difficult as the upfield region is very complex, with many broad, overlapping cross peaks. We have prepared a sample of T4 lysozyme uniformly enriched (>97%) in both ^{15}N and ^{13}C and are applying promising three-dimensional ^1H–^{13}C experiments to assign the ^1H and ^{13}C resonances of the protein (Fesik et al., 1990; Ikura et al., 1990) and to resolve the NOEs between these side-chain protons (Zuiderweg et al., 1990; Kay et al., 1990).

Secondary and Global Structure of T4 Lysozyme

The high-resolution crystal structure of numerous wild-type and mutant T4 lysozymes have been determined by Matthews and co-workers (Weaver & Matthews, 1987). In the early stages of this work, we liberally consulted the crystallographic model of this protein to interpret our NMR data (Griffey et al., 1986; McIntosh et al., 1987b). However, as we developed an extensive library of selectively labeled proteins, the spectra could be interpreted unambiguously and the assignments reported in Table II were made on the basis of only the NMR data and the known sequence of T4 lysozyme.

Many features of the secondary structure of a protein can be determined coincidentally with the assignment of its NMR spectrum (Wüthrich, 1986). In Figure 10, a summary of selected NOE connectivities between backbone protons of T4 lysozyme is presented. As explained previously, the NOE cross peaks between amide protons and side-chain protons were generally not unambiguously assigned to specific side-chain protons due to extensive degeneracies, so this represents only a preliminary analysis of the NOESY spectra of T4 lysozyme. With this in mind, we find that the secondary structural features indicated by the NMR data are qualitatively consistent with the crystallographic model of this protein. Previously we had identified Lys16, Ile17, Val57, and Ile58 as components of an antiparallel β sheet in T4 lysozyme. This protein also has nine regular α helices. On the basis of diagnostic NOE connectivities such as H^N_i–H^N_{i+1}, H^β_i–H^N_{i+1}, and H^α_i–H^N_{i+3}, these helices are observed by NMR measurements, although their exact bounds are difficult to define. In the crystallographic model of T4 lysozyme, residues 108–113 form an "irregular 3_{10}" helix and we also observe continuous H^N_i–H^N_{i+1} NOE connectivities between these amino acids. Residues 159–164 have been modeled as a distorted α helix, although the C-terminal three residues have very high thermal factors (Weaver & Matthews, 1987). We do observe H^N_i–H^N_{i+1} NOE connectivities between residues 158–164, indicative of a helical structure. However, the amide proton of Leu164 has an exceptionally long transverse relaxation time, as indicated by its sharp resonances (Figure 2), suggesting that the C terminus of T4 lysozyme is relatively mobile. This will influence the apparent secondary structure of this region as determined from the NOESY data (Zuiderweg et al., 1989).

At this stage, we have assigned very few side-chain NOEs required to establish the tertiary structure of T4 lysozyme. However, several unambiguous long-range NOEs, such as those between the side chains of V94 and W158 and between the aromatic rings of F4 and F67, as well as a demonstrated salt bridge between H31 and D70 (Anderson et al., 1990) and the above mentioned antiparallel β-sheet (Lowry et al., 1987), indicate that the overall global fold of T4 lysozyme in solution

as measured by NMR is similar to that determined for the crystalline state.

With knowledge of the backbone assignments of a protein, hydrogen exchange studies can also be cautiously interpreted in terms of amide hydrogen bonding and solvent accessibility (Englander & Kallenbach, 1987). In Figure 10, the hydrogen exchange kinetics of the amide groups of T4 lysozyme are qualitatively summarized. Five contiguous sequences of amides show significantly retarded exchange kinetics and these amides are in helical secondary structures, as indicated by the NOESY data and demonstrated in the crystal structure. Several of these helices form the hydrophobic cores of lysozyme. However, amides in several helical regions on the molecule's surface display relatively fast hydrogen exchange kinetics, demonstrating that these are readily accessible to the solvent. Further exchange studies to probe the structure and dynamics of wild-type and mutant lysozymes are in progress.

In conclusion, this report has summarized the assignment of the ^{1}H and ^{15}N NMR spectra of wild-type T4 lysozyme. Using these spectra as a reference, we can rapidly interpret the NMR spectra of T4 lysozymes with defined amino acid substitutions. We are now in the position to investigate the solution properties of wild-type and variant T4 lysozymes. Experiments focusing on the dynamics and stability of the native proteins and on the kinetics of the transition between the unfolded and folded conformations of T4 lysozyme are currently underway. In combination with thermodynamic, crystallographic, and additional spectroscopic studies, we hope to develop an integrated view of how the amino acid sequence of T4 lysozyme dictates the properties of the folded conformation of this protein.

Acknowledgments

We thank Tom Alber, Eric Anderson, Jim Davis, David LeMaster, Brian Matthews, David Muchmore, Christian Nielson, Michael Strain, Terrence Oas, and Joan Wozniak for invaluable help and advice, David Wemmer for running preliminary spectra on T4 lysozyme, Stephen Fesik and Erik Zuiderweg for measuring 3D spectra of labeled T4 lysozyme, Dennis Torchia for sending preprints of their studies on staphyloccocal nuclease, and Dennis Hare and Hare Research for supplying the FTNMR software.

References

Alber, T. A., Dao-pin, S., Nye, J. A., Muchmore, D. C., & Matthews, B. W. (1987) *Biochemistry 26*, 3754–3758.

Anderson, D. E., Becktel, W. J., & Dahlquist, F. W. (1990) *Biochemistry 29*, 2403–2408.

Bachmann, B. J. (1983) *Microbiol. Rev. 47*, 180–230.

Bachovchin, W. W. (1986) *Biochemistry 25*, 7751–7759.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Bax, A., & Summers, M. F. (1986) *J. Am. Chem. Soc. 108*, 2093–2094.

Bax, A., & Weiss, M. (1987) *J. Magn. Reson. 71*, 571–575.

Bax, A., & Marion, D. (1988) *J. Magn. Reson. 78*, 186–191.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Reson. 55*, 301–315.

Bax, A., Sparks, S. W., & Torchia, D. A. (1988) *J. Am. Chem. Soc. 110*, 7926–7927.

Bax, A., Kay, L. E., Sparks, S. W., & Torchia, D. A. (1989) *J. Am. Chem. Soc. 111*, 408–409.

Becktel, W. J., & Baase, W. A. (1987) *Biopolymers 26*, 619–623.

Bendell, M. R., Pegg, D. T., Doddrell, D. M., & Field, J. (1981) *J. Am. Chem. Soc. 103*, 934–936.

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

Blomberg, F., Maurer, W., & Rüterjans, H. (1977) *J. Am. Chem. Soc. 99*, 8149–8159.

Bodenhausen, G., Freeman, R., & Turner, D. L. (1977) *J. Magn. Reson. 27*, 511–514.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 77*, 166–169.

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

Bystrov, V. F. (1976) *Prog. Nucl. Magn. Reson. Spectrosc. 10*, 41–81.

Campbell Burk, S., Papastavros, M. Z., McCormick, F., & Redfield, A. G. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 817–820.

Chazin, W. J., & Wright, P. E. (1987) *Biopolymers 26*, 973–977.

Clore, G. M., Bax, A., Wingfield, P., & Gronenborn, A. M. (1988) *FEBS Lett. 238*, 17–21.

Dahlquist, F. W., Griffey, R. H., McIntosh, L. P., Muchmore, D. C., Oas, T. G., & Redfield, A. G. (1985) in *Synthesis and Applications of Isotopically Labeled Compounds* (Muccino, R. R., Ed.) pp 533–538, Elsevier Scientific, New York.

Dalvit, C., & Wright, P. E. (1987) *J. Mol. Biol. 194*, 313–327.

Doddrell, D. M., Pegg, D. T., & Bendall, M. R. (1982) *J. Magn. Reson. 48*, 323–327.

Emschwiller, M., Hahn, E. L., & Kaplan, D. (1960) *Phys. Rev. 118*, 414–424.

Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521–655.

Englander, S. W., & Wand, A. J. (1987) *Biochemistry, 26*, 5953–5958.

Feng, Y. Roder, H., Englander, S. W., Wand, A. J., & Di Stefano, D. L. (1989) *Biochemistry 28*, 195–203.

Fesik, S. W., & Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 78*, 588–593.

Fesik, S. W., & Zuiderweg, E. R. P. (1990) *Q. Rev. Biophys. 23*, 97–131.

Fesik, S. W., Gampe, R. T., Jr., & Rockway, T. W. (1987) *J. Magn. Reson. 74*, 366–371.

Fesik, S. W., Eaton, H. L., Olejniczak, E. T., Zuiderweg, E. R. P., McIntosh, L. P., & Dahlquist, F. W. (1990) *J. Am. Chem. Soc. 112*, 886–888.

Freeman, R., Mareci, T. H., & Morris, G. A. (1981) *J. Magn. Reson. 42*, 341–345.

Griffey, R. H., & Redfield, A. G. (1987) *Q. Rev. Biophys. 19*, 51–82.

Griffey, R. H., Redfield, A. G., Loomis, R. E., & Dahlquist, F. W. (1985a) *Biochemistry 24*, 817–822.

Griffey, R. H., Jarema, M. A., Kunz, S., Rosevar, P. R., & Redfield, A. G. (1985b) *J. Am. Chem. Soc. 107*, 711–712.

Griffey, R. H., Redfield, A. G., McIntosh, L. P., Oas, T. G., & Dahlquist, F. W. (1986) *J. Am. Chem. Soc. 108*, 6816–6817.

Gronenborn, A., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.

Gross, K.-H., & Kalbitzer, H. R. (1988) *J. Magn. Reson. 76*, 87–99.

Gust, D., Moon, R. B., & Roberts, J. D. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4696–4700.

Hansen, P. E. (1983) *Annu. Rep. NMR Spectrosc. 15*, 103–234.

Hudson, B., Harris, D., Gray, T., & McIntosh, L. (1987) *Biophys. J. 51*, 419a.

Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 86*, 204–209.

Kainosho, M., & Tsuji, T. (1982) *Biochemistry 21*, 6273–6279.

Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72–84.

Kay, L. E., Ikura, M., & Bax, A. (1990) *J. Am. Chem. Soc. 112*, 888–889.

Kinsey, R. A., Kintanan, A., & Oldfield, E. (1981) *J. Biol. Chem. 256*, 9028–9036.

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

LeMaster, D. M. (1989) *Methods Enzymol. 177*, 23–43

LeMaster, D. M., & Richards, F. M. (1985) *Biochemistry 24*, 7263–7268.

LeMaster, D. M., & Richards, F. M. (1988) *Biochemistry 27*, 142–150.

Levitt, M. H., Freeman, R., & Frenkiel, T. (1982) *J. Magn. Reson. 47*, 328–330.

Levy, G. C., & Lichter, R. L. (1979) *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, J. Wiley & Sons, New York.

Llinâs, M., Horsley, W. J., & Klein, M. P. (1976) *J. Am. Chem. Soc. 98*, 7554–7558.

Lowry, D. F., Redfield, A. G., McIntosh, L. P., & Dahlquist, F. W. (1988) *J. Am. Chem. Soc. 110*, 6885–6886.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 92–117.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989a) *Biochemistry 28*, 6150–6156.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., & Bax, A. (1989b) *J. Am. Chem. Soc. 111*, 1515–1517.

Matthews, B. W. (1987) *Biochemistry, 26*, 6885–6888.

McIntosh, L. P., & Dahlquist, F. W. (1990) *Q. Rev. Biophys. 23*, 1–38.

McIntosh, L. P., Dahlquist, F. W., & Redfield, A. G. (1987a) *J. Biomol. Struct. Dyn. 5*, 21–34.

McIntosh, L. P., Griffey, R. H., Muchmore, D. C., Nielson, C. P., Redfield, A. G., & Dahlquist, F. W. (1987b) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1244–1248.

Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., & Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44–73.

Neuhaus, D., Wagner, G., Vasâk, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257–273.

Oh, B. Y., Westler, W. M., Darba, P., & Markley, J. L. (1988) *Science 240*, 908–911.

Oppenheimer, N. J., & James, T. L., Eds. (1989) *Methods Enzymol. 173 and 177.*

Otting, G., & Wüthrich, K. (1990) *Q. Rev. Biophys. 23*, 39–96.

Otting, G., Widmer, H., Wagner, G., & Wüthrich, K. (1986a) *J. Magn. Reson. 66*, 187–193.

Otting, G., Senn, H., Wagner, G., & Wüthrich, K. (1986b) *J. Magn. Reson. 70*, 500–505.

Pardi, A., Wagner, G., & Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 445–454.

Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1984) *Biochem. Biophys. Res. Commun. 117*, 479.

Redfield, A. G. (1983) *Chem. Phys. Lett. 96*, 537–540.

Redfield, A. G., & Waelder, S. (1979) *J. Am. Chem. Soc. 101*, 6151–6162.

Redfield, C., & Dobson, C. M. (1988) *Biochemistry 27*, 122–136.

Redfield, A. G., McIntosh, L. P., Dahlquist, F. W. (1989) *Arch. Biol. Med. Exp. 22*, 129–138.

Robertson, A. D., Purisima, E. O., Eastman, M. A., & Scherga, H. A. (1989) *Biochemistry 28*, 5930–5938.

Senn, H., Euguster, A., Otting, G., Suter, F., & Wüthrich, K. (1987a) *Eur. Biophys. J. 14*, 301–306.

Senn, H., Otting, G., & Wüthrich, K. (1987b) *J. Am. Chem. Soc. 109*, 1090–1092.

Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169–173.

Shaka, A. J., Keeler, J., Frenkiel, T., & Freeman, R. (1983) *J. Magn. Reson. 52*, 335–338.

Sogn, J. A., Gibbons, W. A., & Randall, E. W. (1973) *Biochemistry 12*, 2100–2105.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Stockman, B. J., Reily, M. D., Westler, W. M., Ulrich, E. L., & Markley, J. L. (1989) *Biochemistry 28*, 230–236.

Streinsinger, G., Mukai, F., Dryer, W. J., Miller, B., & Horiruchi, S. (1961) *Cold Spring Harbor Symp. Quant. Biol. 26*, 25–30.

Szilâgyi, L., & Jardetzky, O. (1989) *J. Magn. Reson. 83*, 441–449.

Torchia, D. A., Sparks, S. W., & Bax, A. (1988a) *J. Am. Chem. Soc. 110*, 2320–2321.

Torchia, D. A., Sparks, S. W., & Bax, A. (1988b) *Biochemistry 27*, 5135–5141.

Torchia, D. A., Sparks, S. W., & Bax, A. (1989) *Biochemistry 28*, 5509–5524.

Tsugita, A., & Ikeya-Ocada, Y. (1972) in *Proteins: Structure and Function* (Funatsu, A., Hiromi, H., Imahori, K., Murachi, T., & Narita, K., Eds.) Vol. 1, pp 225–284, Wiley, New York.

Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.

Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–366.

Wand, A. J., Di Stefano, D. L., Feng, Y., Roder, H., & Englander, S. W. (1989) *Biochemistry 28*, 186–194.

Wang, J., LeMaster, D. L., & Markley, J. L. (1990a) *Biochemistry 29*, 88–101.

Wang, J., Hinck, A. P., Loh, S. N., & Markley, J. L. (1990b) *Biochemistry 29*, 102–113.

Weaver, L. H., & Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189–199.

Weaver, L. H., Gray, T. M., Grütter, M. G., Anderson, D. E., Wozniak, J. A., Dahlquist, F. W., & Matthews, B. W. (1989) *Biochemistry 28*, 3793–3797.

Weber, P. L., & Müller, L. (1989) *J. Magn. Reson. 81*, 430–434.

Westler, W. M., Stockman, B. J., Markley, J. L., Hosoya, Y., Mikake, Y., & Kainosho, M. (1988) *J. Am. Chem. Soc. 110*, 6256–6258.

Witanoski, M., Stefaniak, L., & Webb, G. A. (1986) *Annu. Rep. NMR Spectrosc. 18*, 1–761.

Wörgötter, E., Wagner, G., Vasâk, M., Kägi, J. H. R., & Wüthrich, K. (1988) *J. Am. Chem. Soc. 110*, 2388–2393.

Wüthrich, K. (1983) *Biopolymers 22*, 131-138.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, J. Wiley and Sons, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715-740.
Zuiderweg, E. R. P. (1990) *J. Magn. Reson. 86*, 346-357.
Zuiderweg, E. R. P., & Fesik, S. W. (1988) *J. Magn. Reson. 78*, 588-593.
Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387-2391.
Zuiderweg, E. R. P., Hallenga, K., & Olejniczak, E. T. (1986) *J. Magn. Reson. 70*, 336-343.
Zuiderweg, E. R. P., Nettesheim, D. G., Mollison, K. W., & Carter, G. W. (1989) *Biochemistry 28*, 172-185.
Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., & Fesik, S. W. (1990) *J. Magn. Reson. 86*, 210-216.

A 21S Enzyme Complex from HeLa Cells That Functions in Simian Virus 40 DNA Replication in Vitro†

Linda H. Malkas,‡ Robert J. Hickey,‡ Congjun Li, Nina Pedersen,§ and Earl F. Baril*
Worcester Foundation for Experimental Biology, 222 Maple Avenue, Shrewsbury, Massachusetts 01545
Received October 30, 1989; Revised Manuscript Received March 9, 1990

ABSTRACT: A sedimentable complex of enzymes for DNA synthesis was partially purified from the combined low-salt nuclear extract–postmicrosomal supernatant solution of HeLa cell homogenates by poly(ethylene glycol) precipitation in the presence of 2 M KCl, discontinuous gradient centrifugation, Q-Sepharose chromatography, and velocity gradient centrifugation. In addition to the previously described 640-kDa multiprotein DNA polymerase α–primase complex [Vishwanatha et al. (1986) *J. Biol. Chem. 261*, 6619-6628], the enzyme complex also has associated topoisomerase I, DNA-dependent ATPase, RNase H, DNA ligase, a simian virus 40 origin recognition, dA/dT sequence binding protein [Malkas & Baril (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 70-74], and proliferating cell nuclear antigen. Essentially all of the T antigen dependent simian virus 40 in vitro replication activity in the combined nuclear extract–postmicrosomal supernatant solution resides with the sedimentable complex of enzymes for DNA synthesis. Sedimentation analysis on a 10–35% glycerol gradient in the presence of 0.5 M KCl indicates that the enzyme complex is 21S. The associated enzymes for DNA synthesis and in vitro simian virus 40 replication activity cofractionate throughout the purification of the 21S complex. The DNA polymerase and in vitro simian virus 40 replication activities are both inhibited by monoclonal antibody (SJK 132-20) to human DNA polymerase α and by 5–10 μM butylphenyl-dGTP, indicating that the association of DNA polymerase α with the 21S enzyme complex is essential for the initiation of SV40 DNA replication in vitro.

Chromosomal DNA replication in animal cells is a complex process, the mechanism and regulatory control of which are poorly understood. An understanding of this process will require information on the organization and control of the DNA synthesizing machinery, as well as replicons, replication origins (ori),[1] and their cis-acting control elements. Through the development of improved purification procedures, it is now apparent that DNA polymerase α, and the analogous DNA polymerase I from yeast (Campbell, 1986), as isolated from a broad variety of eukaryotes exists as a multiprotein complex that includes primase and other proteins in addition to the polymerase α catalytic subunit [reviewed in Kaguni and Lehman (1988) and Lehman and Kaguni (1989)]. The results of studies of the DNA synthesizing machinery in prokaryotes, by the use of in vitro systems for replication of bacteriophage DNAs (Kornberg, 1980, 1982, 1988; Alberts, 1985; Richardson, 1983), however, have demonstrated that DNA replication requires the concerted action of several enzymes and nonenzymic proteins in addition to the DNA polymerase holoenzymes. This was further established by recent studies on the initiation of replication of *Escherichia coli* (Funnell et al., 1986; Kornberg, 1988) and λ bacteriophage (Mensa-Wilmot et al., 1989) chromosomes in vitro using purified proteins.

Simian virus 40 (SV40)[1] replication uses the host cell DNA synthesizing apparatus, and the only viral-encoded protein that

† This work was supported by U.S. Public Health Service Grants P-3012708 and CA-15187 and by National Research Service Award CA-08173 to L.H.M.
* Author to whom correspondence should be addressed.
‡ Present address: Department of Pharmacology and Experimental Therapeutics, University of Maryland School of Medicine, Baltimore, MD.
§ Present address: Mikrobiolgisk Institut, Oster Farimagsgade, Copenhagen, Denmark.

[1] Abbreviations: AAN, aminoacetonitrile hemisulfate; bp, base pair(s); BSA, bovine serum albumin; BuPdGTP, N^2-(*p-n*-butylphenyl)-dGTP; dA/dT, runs of deoxyadenines or deoxythymines; DEAE, diethylaminoethyl; DTT, dithiothreitol; EDTA-Na_3, ethylenediaminetetraacetic acid trisodum salt, pH 7.5; EGTA-Na_3, [ethylenebis(oxyethylenenitrilo)]tetraacetic acid trisodium salt, pH 7.5; ori, replication origin; PCNA, proliferating cell nuclear antigen; PEG, poly(ethylene glycol); NE, 0.15 N KCl nuclear extract; S-3, postmicrosomal supernatant solution; P-4, sedimentable subfraction of the combined nuclear extract–postmicrosomal supernatant solution; S-4, nonsedimentable subfraction of the combined nuclear extract–postmicrosomal supernatant solution; PMSF, phenylmethanesulfonyl fluoride; SDS, sodium dodecyl sulfate; ss, single stranded; SV40, simian virus 40; T-ag, simian virus 40 large, tumor (T) antigen; Tris, tris(hydroxymethyl)aminomethane; TBS, 50 mM Tris-HCl, pH 7.5, and 0.05 M NaCl; TCA, trichloroacetic acid.

0006-2960/90/0429-6362$02.50/0