Structural and Dynamical Properties of a Denatured Protein. Heteronuclear 3D NMR Experiments and Theoretical Simulations of Lysozyme in 8 M Urea†

Harald Schwalbe,‡ Klaus M. Fiebig, Matthias Buck,§ Jonathan A. Jones, Shaun B. Grimshaw, Andrew Spencer,|| Steffen J. Glaser,‡ Lorna J. Smith, and Christopher M. Dobson*

Oxford Centre for Molecular Sciences, University of Oxford, New Chemistry Laboratory, South Parks Road, Oxford OX1 3QT, England, and Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, England

Received January 9, 1997; Revised Manuscript Received April 14, 1997⊗

ABSTRACT: Oxidized and reduced hen lysozyme denatured in 8 M urea at low pH have been studied in detail by NMR methods. ^{15}N correlated NOESY and TOCSY experiments have provided near complete sequential assignment for both ^{1}H and ^{15}N resonances. Over 900 NOEs, including 130 (i, $i + 2$) and 23 (i, $i + 3$) NOEs, could be identified by analysis of the NOESY spectra of the denatured states, and ^{3}J(HN, Hα) coupling constants and ^{15}N relaxation rates have been measured. The coupling constant and NOE data were analyzed by comparisons with theoretical predictions from a random coil polypeptide model based on amino acid specific ϕ,ψ distributions extracted from the protein data bank. There is significant agreement between predicted and experimental NMR parameters suggesting that local conformations of the denatured states are largely determined by short-range interactions within the polypeptide chain. This result is supported by the observation that the chemical shift, coupling constant, and NOE data are little affected by whether or not the four disulfide bridge cross-links are formed in the denatured protein. The relaxation data, however, show significant differences between the oxidized and reduced protein. Analysis of the relaxation data in terms of simple dynamics models provides evidence for weak clustering of hydrophobic groups near tryptophan residues and increased barriers to motion in the more compact conformers formed when the polypeptide chain is cross-linked by the disulfide bridges. Using this information, a structural description of these denatured states is given in terms of an ensemble of conformers, which have a complex relationship between their local and global characteristics.

It is being increasingly recognized that the denatured states of proteins can provide significant insight into fundamental issues such as the relationship between the sequence of a protein and its three-dimensional structure, protein-folding pathways, the stability of proteins and their turnover in the cell, and the transport of proteins across membranes (Dill & Shortle, 1991; Shortle, 1996; Smith et al., 1996a). In contrast to native folded protein structures, an extensive variety of which have been determined by X-ray diffraction or NMR techniques (Thornton et al., 1995; Hendrickson & Wüthrich, 1995), only a limited number of denatured proteins have been examined in detail experimentally, and procedures for the study of these non-native states are still being developed (Lattman, 1994; Wüthrich, 1994; Kataoka & Goto, 1996; Shortle, 1996). A number of challenging features are associated with the structural characterization of denatured proteins. In the case of NMR techniques, one of the major problems, which hampers studies by homonuclear 2D NMR techniques, is the limited chemical shift dispersion which leads to significant spectral overlap. Despite this problem, some successful studies using magnetization transfer techniques (Dobson et al., 1984; Baum et al., 1989; Evans et al., 1991; Harding et al., 1991; Alexandrescu et al., 1993) and hydrogen exchange procedures (Roder et al., 1985; Wand et al., 1986; Radford et al., 1992; Buck et al., 1994) have been performed and insight has also come from the use of smaller peptide fragments where there is less spectral overlap (Dyson & Wright, 1993).

Major advances have come, however, with the application of heteronuclear 3D NMR techniques to the study of denatured proteins (Neri et al., 1992; Alexandrescu et al., 1994; Arcus et al., 1994; Logan et al., 1994; Buck et al., 1995a; Frank et al., 1995, Pan et al., 1995; Zhang & Forman-Kay, 1995; Wong et al., 1996). These enable in favorable cases resonance assignment and the measurement of structural data such as coupling constants, NOEs, and relaxation rates to be carried out with a detail comparable to studies of native folded proteins, although interpretation has proved difficult. Here, we report a system where such a detailed investigation has been possible and describe the results of heteronuclear 3D NMR studies of the 129 residue protein

† Contribution from the Oxford Centre for Molecular Sciences, which is supported by the U.K. EPSRC, BBSRC, and MRC.

* To whom correspondence should be addressed at the New Chemistry Laboratory, University of Oxford.

‡ Institut für Organische Chemie, Universität Frankfurt/Main, Marie-Curie-Strasse 11, 60449 Frankfurt/Main, Germany.

§ Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138.

|| Institute of Food Research.

⊗ Abstract published in *Advance ACS Abstracts,* June 15, 1997.

[1] Abbreviations: 2D and 3D, two and three dimensional; CIDNP, chemically induced dynamic nuclear polarisation; HMQC, heteronuclear multiple-quantum correlation; HSQC, heteronuclear single-quantum correlation; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement and exchange spectroscopy; sc, side chain; ppm, parts per million; $R_1 = (T_1)^{-1}$, $R_2 = (T_2)^{-1}$, $R_1\rho = (T_1\rho)^{-1}$, longitudinal, transverse, and rotating frame relaxation rates, in inverse seconds; TOCSY, total correlation spectroscopy; ^{3}J(HN, Hα), scalar coupling constant between HN and Hα in Hertz; αN(i, $i + 1$), NOE cross-peak between Hα of residue i and HN of residue $i + 1$; $\Delta\delta$, chemical shift difference between experimental values and those found in short unstructured peptides.

S0006-2960(97)00049-4 CCC: $14.00

hen lysozyme, denatured in 8 M urea. This work extends and complements previous studies of hen lysozyme under a range of denaturing conditions by NMR techniques (Evans et al., 1991; Radford et al., 1992; Buck et al., 1993, 1994, 1995a,b, 1996) and also a recent X-ray scattering study of the urea denatured protein (Chen et al., 1996).

With the increase in experimental data now available for denatured proteins, another complex feature of unfolded states becomes significant, namely their conformational heterogeneity. In contrast to native-state proteins, which can in general be adequately represented by a single conformation about which only limited fluctuations occur, a denatured state must be described in terms of an ensemble of interconverting conformers (Smith et al., 1996a). A range of rates of interconversion between conformers may be observed for different parts of the sequence, and where there is fast exchange the structural parameters will be averaged over the contributing conformers. Under such circumstances, the methods for the interpretation of, for example, NMR data developed for native proteins will not be appropriate and several different approaches have been considered. One strategy is to interpret parameters such as chemical shifts and coupling constants in terms of their deviations from empirical values measured in small unstructured peptides (Bundi et al., 1979; Merutka et al., 1995; Wishart et al., 1995). Another procedure, which we have developed and also apply in the current study of denatured lysozyme, is to use NMR parameters predicted from a model for a random coil as a framework for interpreting the experimental data (Fiebig et al., 1996; Smith et al., 1996a,b). In this model, the random coil is assumed to be a state in which there are no nonlocal interactions along the polypeptide chain and descriptions of the ϕ,ψ populations of each residue in the protein are taken from the distribution of ϕ,ψ torsion angles in the protein data base. From these distributions, coupling constants have been predicted for the different types of amino acid in a random coil state and ensembles of random coil conformers were generated, using a Monte Carlo procedure, from which NOE intensities have been calculated.

Through the detailed experimental characterization of denatured lysozyme by NMR techniques and the interpretation of the experimental data using the random coil model as a baseline for identifying regions of nonrandom structure, we have been able to build up a description of the conformational properties of the denatured protein which we present here. One interesting aspect of the current work is that the denatured protein has been characterized both with the four native disulfide bridges present (referred to as the oxidized form) and in the absence of the disulfide bridges with the cysteine groups reduced and methylated (referred to as the reduced form). The dynamical properties of the protein are found to differ considerably between the oxidized and reduced protein, and we describe a model to interpret the relaxation data of both an unbranched polypeptide chain and one which is cross-linked by disulfide bridges.

MATERIALS AND METHODS

Hen egg white lysozyme was expressed in *Aspergillus niger* using $^{15}NH_4Cl$ as the sole nitrogen source and purified from filtered culture medium (MacKenzie et al., 1996). Preparation of ^{15}N-labeled, reduced S-methylated hen egg white lysozyme (HEWL-S^{Me}) from the oxidized ^{15}N-labeled protein was carried out using procedures described previously (Ching-Li & Atassi, 1973) and verified by electrospray mass spectrometry. Resonances of the S-methyl group between 2.00−2.08 ppm in the 3D NOESY spectra of HEWL-S^{Me} and the characteristic upfield shift of Hα in cysteine compared to oxidized cystine were observed.

NMR samples were prepared to contain approximately 3 mM protein in 8 M urea. The pH was adjusted to 2.0 with hydrochloric acid. Experiments were performed at 20 °C using NMR spectrometers belonging to the Oxford Centre for Molecular Sciences with ^{1}H operating frequencies of 500.2 and 600.2 MHz. 3D NOESY-HSQC, HSQC-NOESY-HSQC, and TOCSY-HSQC experiments (Driscoll et al., 1990; Frenkiel et al., 1990; Ikura et al., 1990, Kay et al., 1990) using gradient sensitivity enhanced back transfer (Schleucher et al., 1994) were recorded with mixing times of 200 ms in the NOESY and 62 ms in the TOCSY experiments. The recycle delay was 1.3 s in each case. The data set typically comprised 128, 64, and 512 complex points in t_1, t_2, and t_3 for the NOESY-HSQC and TOCSY-HSQC experiments, with a digital resolution of 42.2, 17.7, and 6.0 Hz per complex point, respectively, at 600 MHz. The data set for the HSQC-NOESY-HSQC experiment comprised 64, 32, and 512 complex points in t_1, t_2, and t_3 with a digital resolution of 13.0, 25.9, and 13.2 Hz per complex point, respectively, at 600 MHz. Forward linear prediction, as implemented in Felix 2.3 (Hare Research & Biosym Inc.), was used to extend the data in the ^{15}N dimension from 64 to 96 complex points, and all 3D data sets were zero-filled to yield a final matrix size of 256, 128, and 2048 real data points.

^3J(HN, Hα) coupling constants were measured from HMQCJ experiments using line shape simulations (Kay & Bax, 1990). All the HMQCJ experiments were performed on a 600 MHz spectrometer. The data sets comprised 512 and 2048 complex points in t_1 and t_2, with resolutions of 2.2 and 3.0 Hz per complex point. Shifted (55°) squared sine-bell apodization was used in t_1, and the data matrices were zero filled to 1024 real data points. Coupling constants were fitted to doublets in ω_1, after summing over the cross-peak in ω_2 and four times zero-filling, by optimization of the coupling constant and line widths for each of the doublet components using home-written software (Redfield et al., 1991). Approximate average ^3J(HN, Hα) coupling constants for glycine residues were determined as the distance between the two outer lines of the pseudo triplet along ω_1 in the HMQCJ. No correction factors have been applied to acount for differential relaxation (Harbison, 1993).

Longitudinal ($R_1 = 1/T_1$), transverse ($R_2 = 1/T_2$), and rotating frame ($R_{1\rho} = 1/T_{1\rho}$) relaxation rates and heteronuclear NOE effects of ^{15}N nuclei were measured using procedures described in the literature (Kay et al., 1989; Boyd et al., 1990; Palmer et al., 1992; Akke & Palmer, 1996) in conjunction with gradient sensitivity enhanced back transfer (Schleucher et al., 1994; Farrow et al., 1995). The R_1, R_2, and $R_{1\rho}$ relaxation measurements used a series of 6−10 experiments with mixing times ranging from 20 to 1400 ms and 17.6 to 351 ms, respectively. A relaxation delay of 1.5 s was used for all relaxation experiments except for the NOE experiments in which the relaxation delay was 4 s. A CPMG duty cycle delay of 0.5 ms was used in the R_2 experiments. For the $R_{1\rho}$ experiment, a power of 1.5 kHz was used for ^{15}N spin locking and a power of 2 kHz for proton decoupling. Calculations have been performed that show that coherence transfer due to scalar coupling between ^{1}H and ^{15}N spins

can be neglected at these spin locking field strengths. $R_{1\rho}$ rates were corrected for offset effects using the measured R_1 rates. The heteronuclear NOE effect was calculated as the ratio of peak heights in spectra recorded with and without ^{1}H saturation. Relaxation rates were fitted as single exponential decays to peak height data. The errors in the primary intensity data were taken as the standard deviation of spectral noise.

The prediction of ^3J(HN, Hα) coupling constants and NOEs for a random coil used ϕ,ψ distributions taken from a data base of 85 high-resolution protein structures (Swindells et al., 1995; Fiebig et al., 1996; Smith et al., 1996b). ^3J(HN, Hα) coupling constants were calculated from these distributions as population weighted averages using the ϕ distribution for each amino acid and the Karplus equation with the parameters of Pardi et al. (1984):

$$^3J(\mathrm{HN, H}\alpha) = 6.4\cos^2(\phi - 60°) - 1.4\cos(\phi - 60°) + 1.9 \quad (1)$$

This assumes that there is rapid interconversion between conformers in the random coil ensemble so NMR parameters will be averaged.

NOE intensities were predicted using a simplified pentapeptide model (Fiebig et al, 1996). The backbone atoms N, HN, Cα, Hα, C, and O were modeled explicitly, and Cβ and Hβ atoms were added for all amino acids but glycine for which both Hα atoms were modeled explicitly. All H$_\beta$ atoms were treated as united atoms positioned collinearly with the Cα−Cβ bond and were used to estimate the intensity of NOEs involving side chains. Steric overlap between N, HN, Cα, Hα, C, O, Cβ, and Hβ atoms was disallowed, but interactions involving other side chain atoms (Cγ and C$_\delta$) were neglected. Ensembles of conformations were generated by a Monte Carlo procedure which for each conformation selects at random ϕ,ψ angles from the ϕ,ψ distributions in the protein data base.

Distributions of interproton distances were extracted from the conformational ensembles, and NOE intensities calculated assuming a two spin approximation and r^{-6} averaging (Abragam, 1961). Predicted values of NOE intensities converged with ensemble sizes of 10^5 conformers. For comparison with the experimental data, NOEs with predicted intensities greater than those expected for fixed interproton distances of 2.5, 3.5, and 4.3 Å were classified as strong, medium, and weak, respectively. These values were chosen by comparing the experimental intensities of intraresidue NOEs between glycine Hα and HN protons in the 3D NOESY-HSQC spectra with predicted intensities assuming complete conformational averaging around the ϕ torsion angle. For the NOEs between side chain protons and amide protons (scN), the intensities predicted for the united atom Hβ in the model were used with standard intensity categorization unless the amino acid side chain concerned contained a methyl group. In these cases, the weak intensity category was increased to include NOEs with an intensity greater than that expected for a fixed interproton separation of 4.7 Å. This was to account for the increased NOE intensity arising from multiple protons and the specific relaxation properties of methyl groups (Wüthrich, 1986). A similar increase in the weak intensity category was used for NOEs involving the Hα of glycine due to the increased intensity expected with two degenerate protons.

The relaxation data for reduced denatured lysozyme were analyzed using a simple model. This model assumes that the influence of neighboring residues decays exponentially as the distance (in number of peptide bonds) from a given residue increases. Experimental $R_{1\rho}$ data were fitted to the following expression:

$$R_{1\rho}(i) = R_{1\rho}^{\mathrm{intrin}} \sum_{j=1}^{N} \exp\left(-\frac{|(i-j)|}{\lambda}\right) \quad (2)$$

where $R_{1\rho}^{\mathrm{intrin}}$ is a measure of the intrinsic relaxation rate and will thus depend on the temperature and viscosity of the solution, λ is the persistence length of the polypeptide chain (in number of residues), and N is the total chain length of the denatured protein (number of residues).

For oxidized denatured lysozyme, a relaxation model similar to the above was used, but two modifications were introduced to allow for the presence of the disulfide bridges. Firstly, disulfide bonds were modeled using a topological distance matrix, dm_{ij}, which counts the number of covalent bonds along the shortest path from residue i to residue j. Here, peptide bonds and disulfide bridges are treated as being equivalent, each one being counted as an "effective" covalent bond. In addition to using the distance matrix as a measure of how far two residues are separated along the polypeptide chain, a chemical exchange contribution was introduced into the model. This exchange contribution is assumed to arise from conformational averaging associated with the disulfide bridges. In the model, we approximate the chemical exchange contribution by assuming that the chemical shift differences of amide nitrogen resonances in the different conformers associated with the disulfide bridges decrease exponentially with distance (in residues) from the corresponding cysteine residue while the exchange rate is kept constant. The following equation was used to fit R_2 rates of oxidized denatured lysozyme:

$$R_2(i) = R_2^{\mathrm{intrin}} \sum_{j=1}^{N} \exp\left(-\frac{\mathrm{dm}_{ij}}{\lambda_1}\right) + R_2^{\mathrm{exch}} \sum_{k=1}^{N_{\mathrm{Cys}}} \exp\left(-\frac{|(i - \mathrm{Cys}_k)|}{\lambda_2}\right) \quad (3)$$

where R_2^{intrin} is a measure of the intrinsic relaxation rate as defined in eq 2, R_2^{exch} is the amplitude of the exchange contribution from the disulfides, Cys_k is the kth cysteine residue, and N_{cys} is the total number of cysteine residues that are involved in disulfide bridges. λ_1 and λ_2 are persistence lengths of the polypeptide chain with regard to the effects of the intrinsic stiffness of the polypeptide chain and exchange contribution from the disulfides; in fitting the equation to the experimental data, it was assumed that the value of λ_2 was equal to that of λ_1.

RESULTS AND DISCUSSION

Experimental Measurements

Resonance Assignments. Assignment of the ^{1}H and ^{15}N resonances of both reduced and oxidized lysozyme denatured in 8 M urea at pH 2 and 20 °C was carried out using a combination of 3D TOCSY-HSQC, HSQC-NOESY-HSQC, and NOESY-HSQC experiments. Despite the restricted resolution of the spectra compared to those of the native state,

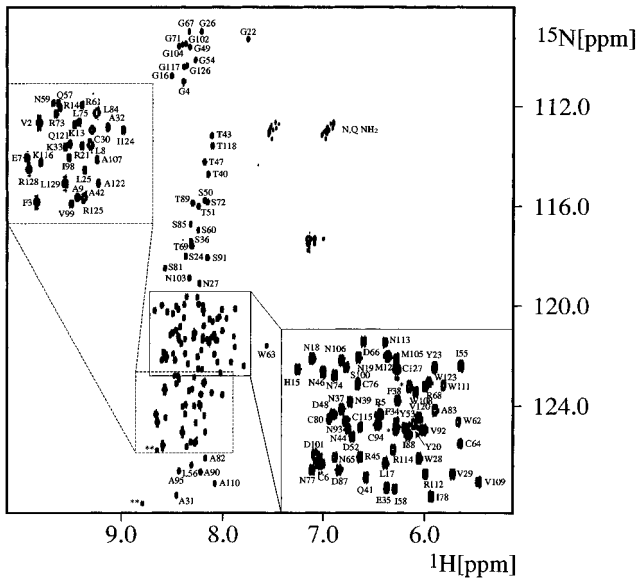

FIGURE 1: ^{15}N-^{1}H HSQC spectrum of the reduced denatured state of lysozyme in 8 M urea at 20 °C and a ^{1}H frequency of 500 MHz showing the backbone amide cross-peaks. Minor peaks from trace impurities are marked with two asterisks (**).

a number of features facilitated their assignment. First, although the ^{1}H chemical shift dispersion is small (~1 ppm for NH), the dispersion of the ^{15}N chemical shifts (~20 ppm) is large enough to yield spectra with well-resolved resonances for the majority of residues in the protein; this can be seen in Figure 1, which shows the ^{1}H-^{15}N HSQC spectrum of reduced denatured lysozyme. The resonances cluster in three regions in the spectra, largely reflecting the characteristic ^{15}N chemical shifts of the different types of amino acids when in unfolded conformations (Wishart et al., 1995). Second, the chemical shifts of side chain protons (e.g., Hβ and Hγ) resonate at the chemical shifts found for short unstructured peptides (Merutka et al., 1995), allowing direct assignment of peaks in the 3D TOCSY-HSQC to a specific type of residue. This leads to particularly straightforward identification of glycine, alanine, valine, leucine, serine, threonine, tryptophan, histidine, and arginine residues. Spectra of native proteins do not usually display this feature because of the chemical shift perturbations induced by longer range specific contacts. Third, cross-peaks in the 3D TOCSY-HSQC are in general intense due to ^{3}J(HN, Hα) coupling constant values of 5.0−8.0 Hz which are large compared to the line widths typical of unfolded conformations. A particularly important result in this present study is that sequential NN(i, $i+1$), αN(i, $i+1$), and scN(i, $i+1$) NOEs have been found between all residues and αN(i, $i+2$) and scN(i, $i+2$) NOEs between the majority of residues of the protein.

The ^{1}H-^{15}N HSQC spectrum of the oxidized state of the protein, however, contains several signals with line broadening, indicating the presence of motions on an intermediate NMR time scale. For some residues the line broadening is sufficiently severe that no resonances could be observed. Such residues are all located near the disulfide bridges Cys 30−Cys 115 (residues 29 and 112−115), Cys 64−Cys 80 (residues 64−66 and 80−83), and Cys 76−Cys 94 (residues 76−79 and 94−100). Residues close to the fourth disulfide bridge Cys 6−Cys 127 do not show any increased line width. Overall, a total of 120 residues could be assigned in the reduced state and 103 in the oxidized state of denatured lysozyme.

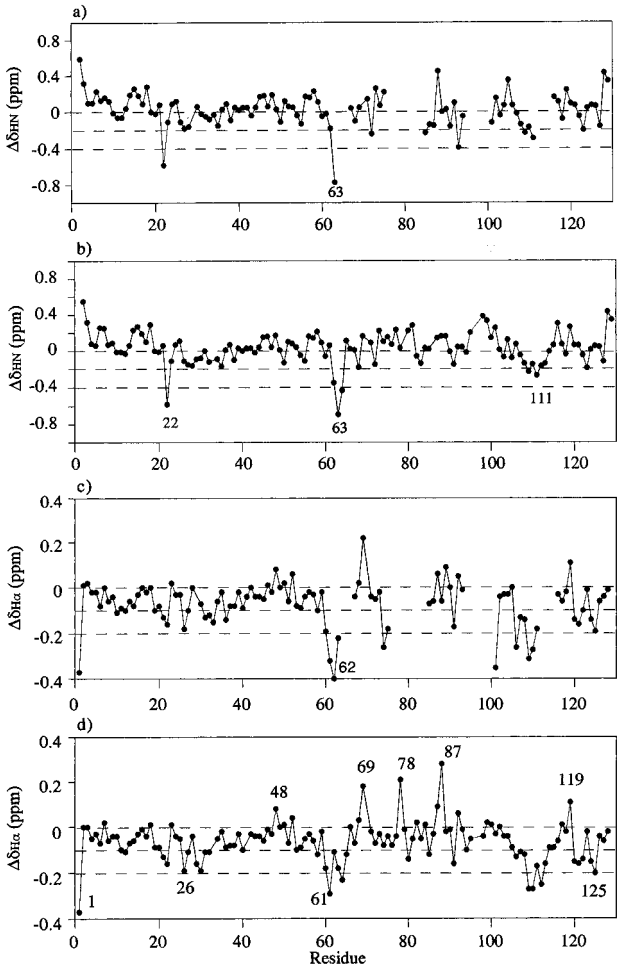

FIGURE 2: (a,b) HN and (c,d) Hα chemical shift perturbations from chemical shifts measured in short unstructured peptides ($\Delta\delta = \delta^{expt} - \delta^{empirical}$) (Merutka et al., 1995) plotted as a function of protein sequence for oxidized (a, c) and reduced (b, d) denatured lysozyme in 8 M urea at pH 2. Residues of interest are labeled by sequence number.

Chemical Shifts and ^{3}J(HN, Hα) Coupling Constants. The chemical shifts measured for the oxidized and the reduced states of denatured lysozyme are very close to empirically measured values for unstructured peptides (Merutka et al., 1995). Also, the pattern of chemical shift perturbation is extremely similar in the reduced and oxidized proteins (Figure 2): $\langle\Delta\delta\rangle = 0.19 \pm 0.18$ ppm for the HN chemical shift and 0.12 ± 0.10 ppm for the Hα chemical shift in both the oxidized and reduced states. Large upfield shifts of the HN resonances are observed for the two N- and C-terminal residues, which can be attributed to electrostatic effects arising from the termini, the same effects giving a large downfield shift for the Hα resonance of Lys 1. In addition to these effects at the termini, large downfield shifts of the HN resonances are found for Gly 22, Trp 62, Trp 63, and Cys 64 and marked deviations of Hα chemical shifts are seen in the regions of residues 19−32 ($\langle\Delta\delta_{H\alpha}\rangle = -0.10/-0.11$ ppm for the reduced and oxidized form, respectively), residues 58−63 (−0.22/−0.17 ppm), residues 106−112 (−0.23/−0.19 ppm) and residues 120−125 (−0.13/−0.15 ppm).

^{3}J(HN, Hα) coupling constants for 57 residues in the oxidized and 104 residues in the reduced form of the protein were measured using the HMQCJ experiment (Kay & Bax, 1990). For those residues with line widths larger than the coupling constant, fitting is not reliable; the coupling constant

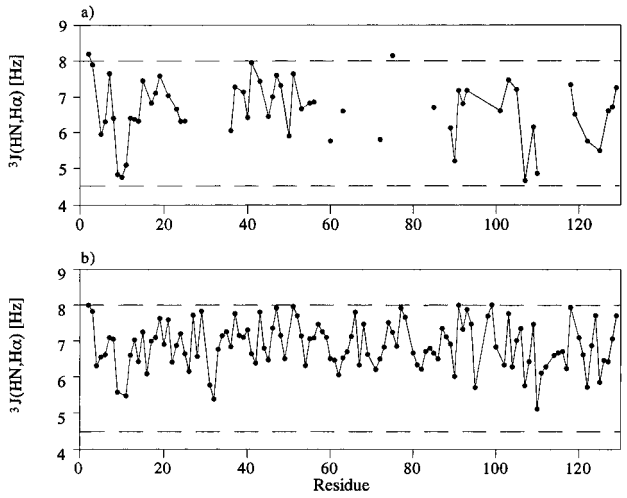

FIGURE 3: ^3J(HN, Hα) coupling constants measured in HMQCJ (Kay & Bax, 1990) experiments for oxidized (a) and reduced (b) denatured lysozyme. Values of ^3J(HN, Hα) coupling constants indicative for fully populated α helical or β strand conformations are shown by horizontal broken lines (Smith et al., 1996b).

data are therefore incomplete, especially for the oxidized protein. For both states, the coupling constants can be seen to vary substantially along the sequence ranging between 4.8 and 8.2 Hz (Figure 3), and at first sight the values appear to be randomly distributed without detectable clustering.

NOE Data. Figure 4a shows ω_1, ω_3 strips at the respective ^{15}N frequency, ω_2, from the 3D NOESY-HSQC spectrum for residues 107−116 of reduced denatured lysozyme. The full data set of identified NOEs for both the reduced and oxidized states is summarized in Figure 5. In this figure, all NOEs arising from side chain protons have been grouped together (scN NOEs). From these two figures, it is immediately apparent that many NOEs are observed in the spectra of denatured lysozyme. Virtually all sequential NN(i, $i + 1$), αH(i, $i + 1$), and scN(i, $i + 1$) NOEs are observed together with a large number of αN(i, $i + 2$) and scN(i, $i + 2$) NOEs (shown with dashed arrows in Figure 4a) and some NN(i, $i + 2$), NN(i, $i + 3$), αN(i, $i + 3$), and scN(i, $i + 3$) NOEs. Many of the medium range (i, $i + 2$) and (i, $i + 3$) NOEs are, however, close to the noise level. This is demonstrated in Figure 4b where the ω_1, ω_3 strip for Trp 111 is plotted at different contour levels. At the highest contour level (L_1), only the NH,N autocorrelation peak and the sequential αN(i, $i + 1$) and intraresidue βN cross-peaks are observed. At a lower level (L_2), the αN intraresidue cross-peak and the sequential NN(i, $i + 1$) and βN(i, $i + 1$) cross peaks from Ala 110 are also observed. At the next level (L_3), one of the two intraresidue cross peaks from the aromatic protons of Trp 111 is visible, and at the lowest level (L_4), a second cross-peak from aromatic protons as well as sequential αN(i, $i + 2$) and βN(i, $i + 2$) cross-peaks from Val 109 are also seen. The data summarized in Figure 5 were extracted from peaks visible at the lowest contour level, L_4.

Despite limitations in the experimental signal-to-noise ratios and severe cross-peak overlap, even in the 3D NOESY-HSQC experiments, which hampered the analysis of medium-range NOEs, a total of 21 NN(i, $i + 2$), 59 αN(i, $i + 2$), 50 scN(i, $i + 2$), 1 NN(i, $i + 3$), 9 αN(i, $i + 3$), and 13 scN(i, $i + 3$) NOEs have been identified in the spectra of the oxidized and/or reduced denatured protein. The (i, $i + 2$) cross-peak density is uniformly high except in the region toward the middle of the polypeptide sequence, between residues 68 and 97, where a much smaller total number of medium range cross peaks has been identified. The distribution of (i, $i + 3$) cross peaks is, in contrast, highly nonuniform; (i, $i + 3$) cross peaks are only observed in the regions 6−13, 16−22, 28−38, 51−61, 95−101, 104−111, and 119−123. It is interesting to compare the pattern of NOEs seen in the reduced and oxidized protein. Most of the NOEs are observed in the oxidized as well as in the reduced protein (indicated by black boxes in Figure 5). However, major differences are apparent for some residues located in the vicinity of three of the four disulfide bonds, particularly residues in the regions 28−30, 63−65, 75−84, 94−100, and 101−115. A larger number of NOE cross peaks is also observed in the oxidized state in the N and C terminal regions (residues 1−6 and 123−129). The latter finding is consistent with the heteronuclear relaxation studies discussed below, which show increased flexibility and smaller effective correlation times for residues at the N- and C-termini of the reduced protein.

Heteronuclear Relaxation Data. Relaxation data, R_1, R_2, and $R_{1\rho}$, and the heteronuclear NOE could be extracted for 80 and 97 of the 129 main chain nitrogen nuclei for oxidized and reduced denatured lysozyme, respectively (Figure 6). The R_1 and heteronuclear NOE data are very similar for the two states, but there are considerable differences in the transverse relaxation rates R_2 and $R_{1\rho}$. In the following, we concentrate on these transverse relaxation rates and their differences in the oxidized and reduced protein. In the reduced state, the R_2 and $R_{1\rho}$ relaxation rates are very similar; the relaxation rates follow a bell-shaped curve along the sequence with an average relaxation rate of 4.2 s^{-1} for residues in the middle of the chain. Smaller rates are observed at both ends of the sequence, and there are a few clusters with larger R_2 and $R_{1\rho}$ relaxation rates, located in the vicinity of five of the six tryptophan residues (Trp 62, Trp 63, Trp 108, Trp 111, and Trp 123). For oxidized denatured lysozyme, however, the transverse relaxation rates vary considerably along the sequence. The R_2 values range from 5 s^{-1} (near Ser 50) to 17 s^{-1} (near Cys 64 and Arg 115), with even larger values being probable for residues which could not be observed or assigned in the HSQC spectrum of the oxidized state because of considerable broadening of the resonances.

Structural Analysis

Coupling Constant and NOE Data. The nonuniform distribution of ^3J(HN, Hα) coupling constants and the large number of NOEs observed for the urea denatured states of lysozyme prompted us to use a model for a random coil as a framework for interpreting the experimental data. In this model, the random coil state of a polypeptide chain is presumed to be one in which there are no specific nonlocal interactions between residues (Fiebig et al., 1996; Smith et al., 1996a,b). In such a state, individual residues will however have local conformational preferences and will sample the low energy ϕ,ψ conformations in Ramachandran space (Ramachandran, 1963). In the model, the ϕ,ψ populations for residues in a random coil are derived from the protein data bank. The distribution of ϕ,ψ torsion angles in the protein data bank has been found to resemble the experimental ϕ,ψ energy surface (Swindells et al., 1995; Serrano et al., 1995; Smith et al., 1996a), the effects of specific nonlocal interactions present in individual proteins

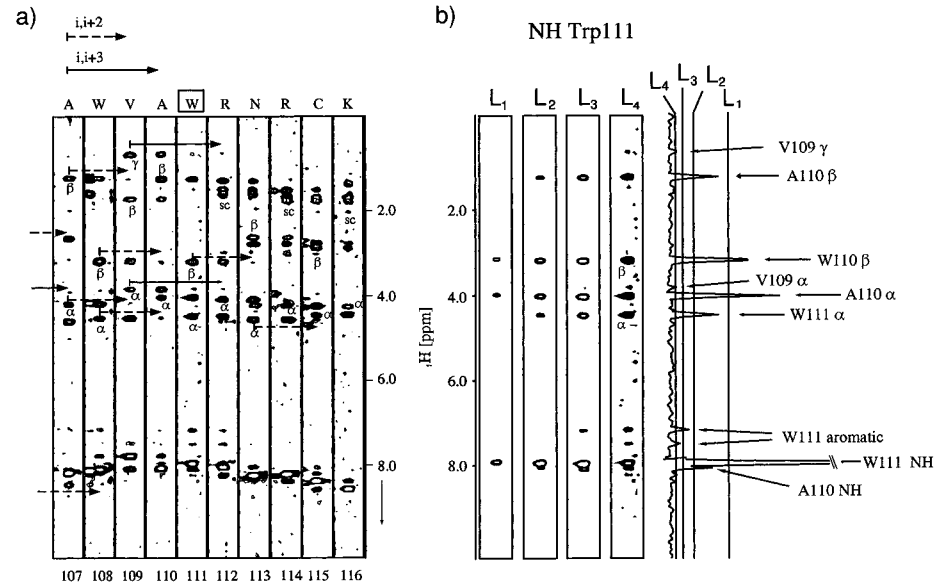

FIGURE 4: (a) ω_1, ω_3 strips from 3D NOESY-HSQC spectrum with a mixing time of 200 ms showing residues 107−116 in reduced denatured lysozyme. (b) The strip corresponding to Trp 111 is plotted at four different signal-to-noise levels which are indicated by vertical lines in the ω_1 1D trace. Analyzed cross-peaks in the 1D trace are marked with arrows and annotated according to the resonance frequency in ω_1.

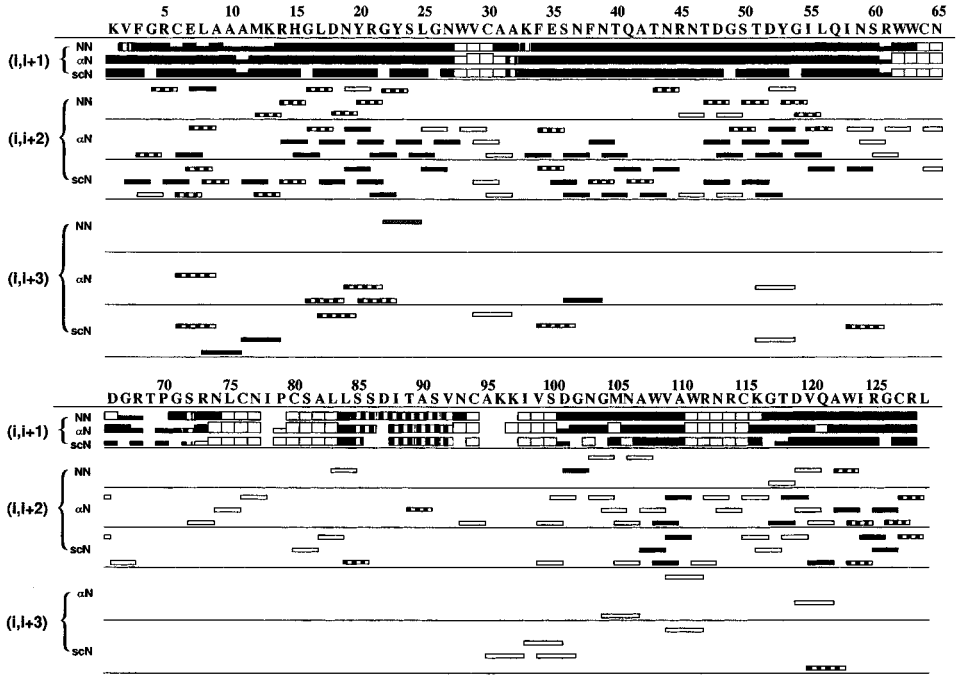

FIGURE 5: Summary of short- and medium-range NOEs for the 8 M urea denatured states of lysozyme at pH 2 observed in 3D NOESY-HSQC spectra recorded with a mixing time of 200 ms. The data are the sum of the NOEs seen for the reduced and the oxidized protein. Black bars represent NOEs observed in both states, whereas NOEs only observed for the oxidized protein are represented by dashed bars and NOEs only observed in the reduced form by open bars. Relative intensities as determined by counting contour lines in the 2D plots are indicated as strong, medium, and weak by the height of the bars. All NOEs identified involving side chain protons are grouped together as scN NOEs.

being averaged out over the set of structures. From the protein data base ϕ,ψ distributions, coupling constants and NOEs have been predicted for a random coil using the procedures described in the Materials and Methods section.

The mean experimental ^3J(HN, Hα) value for residues in denatured lysozyme where measurements could be made is 6.9 Hz. This is close to the value predicted from the ϕ,ψ distribution of all residues in the protein data base (6.6 Hz), where there is overall 45 and 40% population of the α and β regions of ϕ,ψ space, respectively (Smith et al., 1996b). If, however, coupling constants are predicted for different residue types in a random coil from the amino acid specific ϕ,ψ distributions in the protein data base, the values are found to vary significantly among the different amino acids. These differences reflect the variations in the torsion angle preferences of the different amino acids (Serrano, 1995; Smith et al., 1996b), small coupling constants being predicted for residues with high populations of α space, e.g., alanine (5.8 Hz) and glutamate (6.2 Hz) and large values for residues which favor β or extended conformations such as amino acids with β-branched side chains, e.g., valine (7.3 Hz) and threonine (7.5 Hz) (Figure 7). A good correlation is seen between these predicted values and the experimental data for both denatured states of lysozyme (Figure 8 shows data

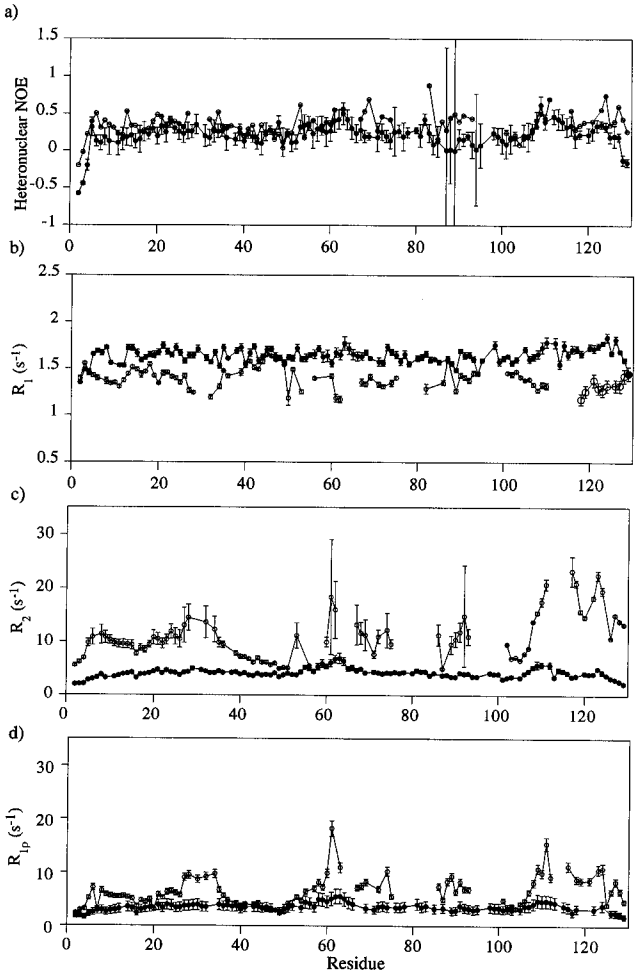

FIGURE 6: Relaxation data of main chain ^{15}N sites from reduced and oxidized denatured lysozyme (a) heteronuclear NOE at ^{15}N spectrometer frequency of 60.8 MHz; (b) longitudinal relaxation rates $R_1 = 1/T_1$; (c) transverse $R_2 = 1/T_2$ relaxation rates; (d) relaxation rates in the rotating frame $R_{1\rho} = 1/T_{1\rho}$ measured at 60.8 MHz. Values are plotted as a function of sequence using filled symbols for the reduced protein and open symbols for the oxidized protein.

for the reduced protein; correlation coefficient = 0.89, slope = 1.03, $\chi^2 = 0.85$). Much of the observed variation in the experimental coupling constant data (Figure 3) therefore reflects the different ϕ,ψ preferences of the various amino acids, rather than the presence of any persistent nonlocal interactions along the polypeptide chain.

The experimental ^3J(HN, Hα) coupling constants are, however, on average 0.3 Hz larger than the predicted random coil values. This suggests that the population of α space in denatured lysozyme is slightly lower overall than in the ϕ,ψ distributions used in the predictions. In these calculations, all residues in the set of proteins were used to predict the ϕ,ψ populations (ALL parameter set). An alternative approach is to use in the random coil predictions only residues that are not in recognized regions of secondary structure within the structures in the protein data base (COIL parameter set) (Smith et al., 1996b). This gives rise to ^3J(HN, Hα) values that are in general larger by 0.4−0.5 Hz than those predicted with all residues; as more experimental data for denatured proteins become available, it should become apparent which ϕ,ψ distribution is most appropriate to use. Some larger deviations in experimental ^3J(HN, Hα) coupling constants from the predicted values are seen, however, in specific parts of the sequence, particularly for the oxidized protein. For example, there is a clustering of coupling constant values for residues 5−12 and 107−110 in the oxidized protein that deviate by up to 1.1 Hz. In these regions, it seems likely that the intrinsic ϕ,ψ preferences are being modulated by interactions with residues that are either neighboring or more distant in the amino acid sequence. In particular, we note that, in the first region, there are three adjacent alanine residues (9−11) and that in the second region, increased heteronuclear relaxation rates are also observed (see below).

Using a Monte Carlo procedure and the simplified pentapeptide model shown in Figure 7c, ensembles of random coil conformers were calculated using the protein data base ϕ,ψ distributions as described in the Materials and Methods section (Fiebig et al., 1996). We first consider the ensemble generated using the protein data base ϕ,ψ distributions of all the amino acids. From this, interproton distance distributions have been extracted, the statistics of which are summarized in Table 1. Comparing the effective interproton distances for a random coil, $\langle r^{-6}\rangle^{-1/6}$, with the interproton distance in regions of regular secondary structure, two features are immediately apparent. Firstly, in contrast with α helical or β strand secondary structure, both the αN(i, i + 1) and the NN(i, i + 1) effective interproton distances are short in a random coil, reflecting the population of both α and β conformers. Secondly, the effective αN(i, i + 2) interproton distance (4.2 Å) is shorter than that in either α helical (4.5 Å) or β strand (>5.5 Å) secondary structure. This reflects the significant population of mixed $\beta\alpha$ conformations in a random coil; in these conformers the αN(i, i + 2) interproton distance is approximately 3.7 Å. Comparison of these two features of the effective interproton distances in the ensemble of random coil conformers with the experimental data for denatured lysozyme provides an explanation for the observation of both strong sequential αN(i, i + 1) and NN(i, i + 1) NOEs for almost all residues in the protein and for the significant number of αN(i, i + 2) NOEs identified throughout the protein sequence.

A more detailed NOE prediction takes into account the amino acid sequence and uses the residue specific ϕ,ψ distributions from the protein data base discussed above. The resulting predicted relative NOE intensities for the lysozyme sequence, calculated from interproton distance distributions in the random coil ensembles using the two spin approximation, are shown in Figure 9. Variations are predicted in the intensities of all the types of NOE, the ratio of sequential αN to NN NOE intensity, for example, being predicted to vary by a factor of 3 (Fiebig et al., 1996). Many of the greatest variations observed occur when glycine residues are involved; this arises from the different conformational properties of glycine reflecting the reduced steric constraints in the absence of a side chain. For example, the NN(i, i + 2) NOEs are predicted to be intense when glycine is in position (i + 1) as this interproton separation can be short (~3.8 Å) when residue (i + 1) adopts a conformation in the left-hand region of ϕ,ψ space. A similar situation occurs when asparagine is in position (i, i + 1) as this residue can also populate significantly α_L conformations (Swindells et al., 1995).

In order to compare in detail the NOE predictions, taking into account the amino acid sequence, with the experimental data various factors must be considered. Firstly, the approximate intensity below which NOEs will not be observable in the spectrum must be chosen (the "cutoff"). This is

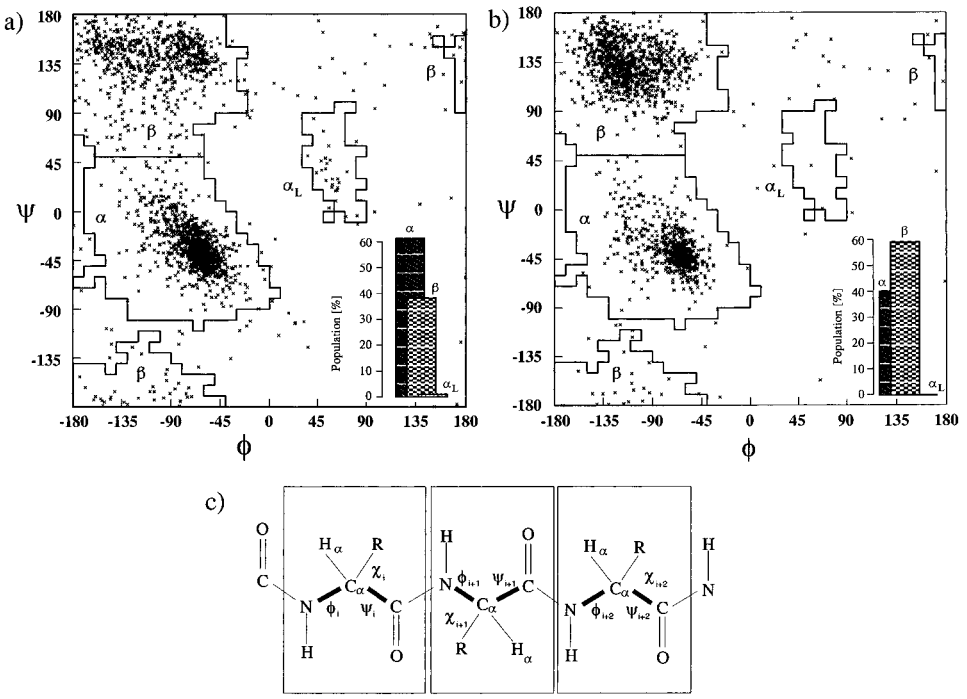

FIGURE 7: Distribution of ϕ,ψ torsion angles for alanine (a) and valine (b) residues in a data base of 85 high-resolution protein structures. Histograms show the percentage populations of the α, β, and α_L regions of ϕ,ψ space, the definitions of these regions being taken from Morris et al. (1992). (c) The simplified pentapeptide model used to generate an ensemble of random coil conformers from the ϕ,ψ populations in the protein data base using a Monte Carlo procedure.

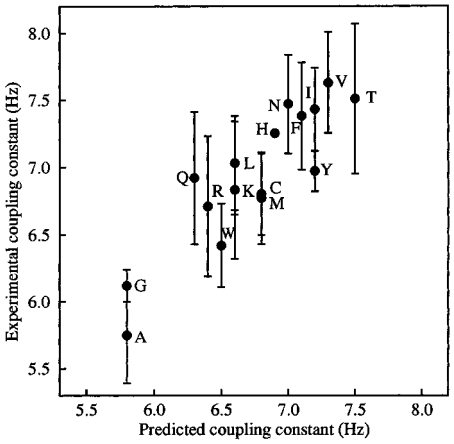

FIGURE 8: Predicted coupling constants from the random coil model versus mean experimental coupling constants for reduced denatured lysozyme. The points are labeled with the one letter amino acid codes and error bars show ± 1 standard deviation. The values for aspartic acid and glutamic acid are not included in this comparison; the side chains of these residues will be protonated at the low pH (2.0) used in the experimental study, and predicted coupling constants are not available for these conditions.

dependent both on the signal-to-noise ratio in the spectrum (which may not be uniform due to line broadening caused by intramolecular exchange or exchange with the solvent) and the effective correlation time. The latter is likely to vary between different parts of the sequence and for side chain compared to main chain protons. We discuss this in more detail later. Here, a single set of intensity cutoffs has been used as an approximation for the whole comparison, chosen by comparing in an overall manner the predictions with the experimental data (see Materials and Methods). Secondly, there is considerable overlap in the experimental spectra; in these comparisons, NOEs that are predicted but could not be observed due to chemical shift degeneracy are excluded. In addition, as the resonances of many methyl groups in the amino acid side chains are well resolved compared to the β

Table 1: Comparison of the Effective Interproton Distances in the Ensemble of Random Coil Conformers with Interproton Distances in Regions of Regular Secondary Structure

distance	random coil[a] (Å)	α helix (Å)	β strand (Å)
αN ($i, i+1$)	2.5	3.5	2.2
NN ($i, i+1$)	2.7	2.8	4.3
αN ($i, i+2$)	4.2	4.4	6.1
NN ($i, i+2$)	4.6	4.2	6.9
αN ($i, i+3$)	4.6	3.4	8.9
NN ($i, i+3$)	5.5	4.8	10.7

[a] Effective interproton distances, $\langle r^{-6}\rangle^{-1/6}$, are given for an ensemble of random coil conformers generated using the ϕ,ψ torsion angle populations of all residues in the protein data base.

protons, we consider NOEs between all side chain protons and the amide protons (scN) rather than just βN NOEs. The scN NOE intensities were estimated from intensities predicted for the Hβ pseudoatoms of the simplified peptide model as described in the Materials and Methods. For residues where the side chain involves a methyl group, the lower limit for the NOE intensity cutoff of these scN NOEs was increased from an effective distance of 4.3 Å to one of 4.7 Å to take into account the multiple protons and specific relaxation properties of methyl groups. This increase for NOEs from methyl groups is analogous to that used when experimental NOE intensities are converted into distance restraints in protein structure determination procedures for native proteins (Wüthrich, 1986).

Figure 10d shows a comparison between NOEs predicted from the random coil model using the chosen cutoffs and the combined experimental data from both of the urea denatured states. We have combined the two data sets, despite some differences in the details of their experimental coupling constants and chemical shifts, since the overall patterns are similar. Variations between the NOEs identified for the two states result in general from differences in line widths rather than reflecting any differences in conformational preferences. This is particularly apparent in the

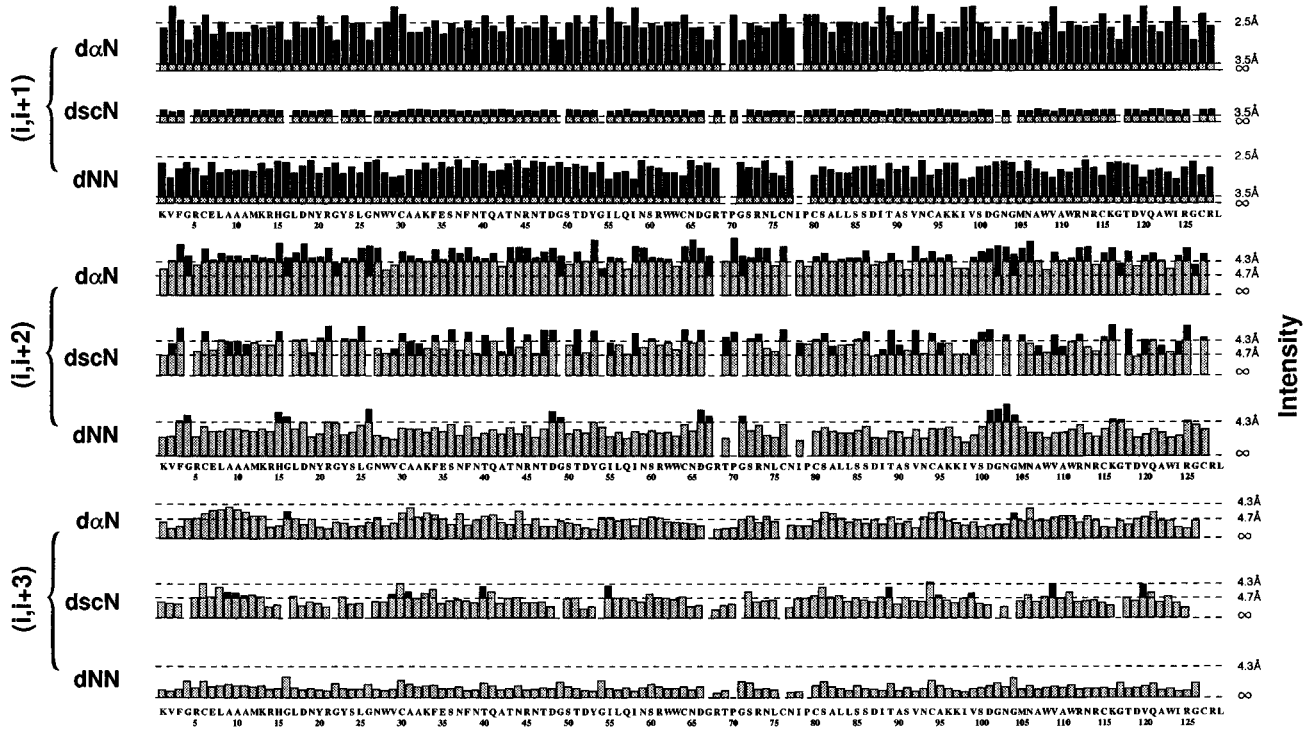

FIGURE 9: Variations in NOE intensities predicted, using the amino acid specific ϕ,ψ distributions in the protein data base, for the lysozyme sequence behaving as a random coil. The intensities expected for NOEs arising from various fixed interproton distances are shown for comparison (infinite interproton distances for zero intensity). NOE intensities above chosen cutoff values (see text) are drawn in black. For (i, $i + 1$) NOEs, the medium intensity cutoff is used (intensity greater or equal to that expected for 3.5 Å separation). For (i, $i + 2$) and (i, $i + 3$) NOEs, the weak cutoff is used (intensity greater or equal to that expected for 4.3 Å fixed separation; for scN NOEs where the residue involved contains a methyl group and for NOEs involving glycine Hα the cutoff is increased to 4.7 Å).

vicinity of the disulfide bridges in the oxidized protein where the NMR resonances are significantly broadened. In Figure 10d, the observed and predicted NOE cross peaks are shown in black. In white, cross peaks that are observed but not predicted are indicated, while dashed boxes show those NOEs that are predicted but not observed. Overall, the agreement between the prediction and the experimental data is remarkably good. Both αN(i, $i + 1$) and NN(i, $i + 1$) NOEs are predicted and observed throughout the sequence. αN(i, $i + 1$) cross peaks are found to be uniformly stronger than intraresidue αN cross peaks, consistent with the prediction of the model. Throughout the sequence, αN(i, $i + 2$) and scN(i, $i + 2$) NOEs are also predicted and observed; 56 out of 82 predicted and potentially observable αN(i, $i + 2$) and 45 out of 72 predicted and potentially observable scN(i, $i + 2$) NOEs are seen. Twelve (44%) of the predicted but not observed αN(i, $i + 2$) NOEs and 16 (22%) of the predicted but not observed scN(i, $i + 2$) NOEs are within the region of residues 60−100, which displays only a sparse number of NOEs in total; this region is considered in more detail later. Few of the αN or scN(i, $i + 2$) NOEs [2 αN(i, $i + 2$) and 6 scN(i, $i + 2$)] are observed but not predicted. For the NN(i, $i + 2$) NOEs, 11 of 27 predicted and potentially observable cross peaks are observed, but another 10 NN(i, $i + 2$) cross peaks are observed but not predicted by the model. In summary, therefore, the occurrence of αN(i, $i + 2$) and scN(i, $i + 2$) cross peaks are broadly consistent with the proposed random coil model, although the prediction of NN(i, $i + 2$) NOEs is less successful.

In the spectra, only 23 (i, $i + 3$) NOEs are observed at the signal-to-noise ratios of our experiments. All the medium range (i, $i + 3$) NOEs are close to the cutoff of the experimental signal-to-noise (and similarly close to the cutoff used in the predictions), and their observation in the spectra will therefore be affected by variations in correlation time through the protein sequence. While the agreement between prediction and experiment is poor for the αN(i, $i + 3$) NOEs (2 out of 8 predicted NOEs are observed, and 7 NOEs are observed but not predicted), the agreement with predictions is remarkable for the scN(i, $i + 3$) NOEs; 10 out of 13 predicted and potentially observable scN(i, $i + 3$) NOEs are seen in the spectra and only three experimental scN(i, $i + 3$) NOEs are not predicted. Many, but not all, of these NOEs involve the side chains of residues with methyl groups and so indicate the importance of the intensity enhancement of these NOEs due to the multiple protons and relaxation properties of methyl groups which, as discussed above, were included in the predictions by decreasing the intensity cutoff for residues with methyl groups. The successful prediction of scN(i, $i + 3$) on the basis of this random coil model shows the power of such an approach in providing a framework for the interpretation of experimental data for denatured proteins. It is important to note, however, that these (i, $i + 3$) NOEs are weak in the spectra of urea denatured lysozyme and would not have been observed if the signal-to-noise levels in the experimental spectra were not so high.

Analysis of NMR Relaxation Data. Heteronuclear relaxation measurements are sensitive both to motions on a subnanosecond time scale and to slow conformational exchange in the millisecond time scale (Wagner, 1993). The relaxation data for reduced denatured lysozyme have, however, a very simple pattern. In the central region of the protein sequence, the R_2 relaxation rates of most residues approach plateau values (3.5 s^{-1}) while in the terminal parts of the sequence the relaxation rates are significantly smaller. This implies that the relaxation properties of a given amide are in general not markedly influenced by the specific nature of its neighbors but just reflect that it is part of a polypeptide

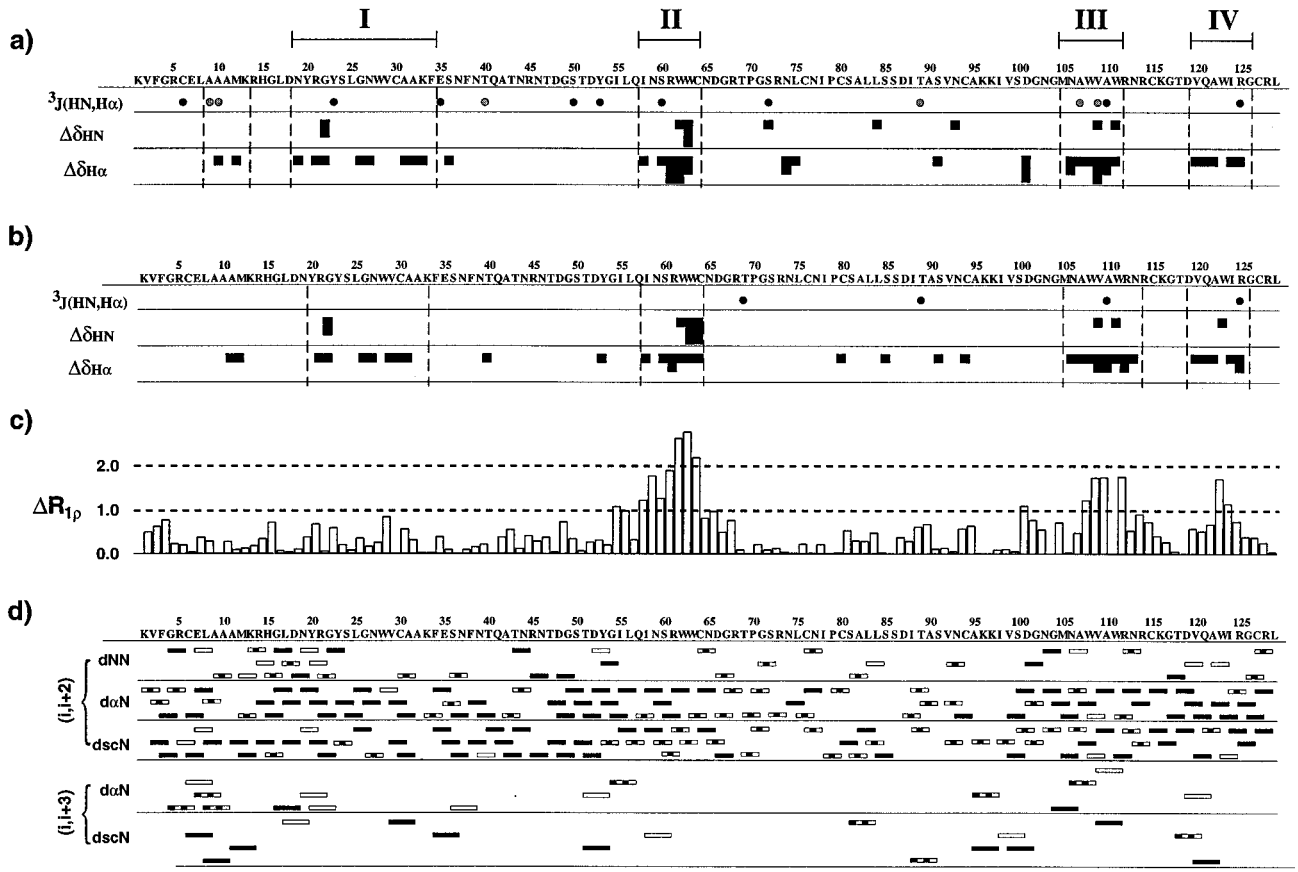

FIGURE 10: Summary of the NMR parameters defining the conformational properties of the urea denatured states of lysozyme. Regions of significant deviations from random coil behavior are labeled I−IV. Panels a and b show coupling constant and chemical shift data for oxidized and reduced lysozyme, respectively. ^3J(HN, Hα) coupling constants with marked deviations (larger than 1 or 2 standard deviations) from random coil predictions are shown with circles, significant HN ($\Delta\delta > -0.2$ ppm by 1 unit, $\Delta\delta > -0.4$ ppm by 2 units, and $\Delta\delta > -0.6$ ppm by 3 units) and Hα ($\Delta\delta > -0.1$ ppm by 1 unit, $\Delta\delta > -0.2$ ppm by 2 units, and $\Delta\delta > -0.3$ ppm by 3 units) chemical shift perturbations from empirical values for unstructured peptides are indicated by vertical bars. In panel c, deviations of heteronuclear $R_{1\rho} = 1/T_{1\rho}$ relaxation rates for the reduced protein from random coil predictions are shown. (d) Comparison between the combined experimental NOE data set and the predictions for a random coil are shown. NOEs observed and predicted by the random coil model are indicated by black horizontal bars, predicted but not observed NOEs by dashed horizontal bars, and observed but not predicted NOEs by open horizontal bars.

chain. Further experimental data have been reported recently for reduced FK506 binding protein in 8 M urea and for cysteine free protein G in 7.4 M urea (Logan et al., 1994, Frank et al., 1995). For both proteins, relaxation measurements follow the general pattern observed here with R_2 relaxation rates, which vary between 3.0 and 4.0−5.0 s^{-1} in protein G and FK506 binding protein, respectively; again, lower rates are found at the terminal parts of the protein. All these results suggest that a simple model of segmental motion can to a first approximation explain the experimental data. This conclusion is qualitatively similar to that from heteronuclear ^{13}C R_1 measurements on atactic polystyrene (Allerhand et al., 1972). Here the relaxation rates were found to increase by 40% on increasing the molecular mass from 2 to 10 kDa and then to remain constant for homopolymers as large as 860 kDa. This implies that for longer polymer chains, individual segments of the polymer move independently in solution, i.e., the correlation time is largely independent of the overall tumbling rate and affected only by segmental motions.

In this work, we introduce two simple models to characterize the relaxation behavior of reduced and oxidized denatured lysozyme (see Materials and Methods). As their central hypothesis, these models both assume that relaxation contributions due to neighboring residues in the peptide chain decay exponentially as the distance from a given residue increases. Figure 11a shows the transverse relaxation rates, $R_{1\rho}$, of reduced denatured lysozyme. Since the $R_{1\rho}$ experiment reduces chemical exchange contributions to the relaxation rate (Szyperski et al., 1993), we have fitted these data to a minimalistic model. We note, however, that the R_2 and $R_{1\rho}$ values are very similar for reduced denatured lysozyme (Figure 6c,d), indicating the absence of significant chemical exchange contributions to R_2. Experimental $R_{1\rho}$ data for the ith residue were fitted to a sum of exponential functions over all residues j of the chain with two variable parameters, namely a measure of the intrinsic relaxation rate, $R_{1\rho}^{\text{intrin}}$, and the persistence length λ (eq 2 in Materials and Methods). The best two-parameter fit of the $R_{1\rho}$ data for the reduced denatured protein was found for $\lambda = 7$ residues and $R_{1\rho}^{\text{intrin}} = 0.25$ s^{-1}. Figure 11a demonstrates that for most parts of the sequence this simple model reproduces well the pattern of experimental transverse relaxation rates. There are, however, three regions where significantly larger relaxation rates are observed experimentally. Their location correlates with the positions of several of the tryptophan residues in the sequence. The largest relaxation rates are observed around Trp 62 and Trp 63, but elevated $R_{1\rho}$ values are also observed near Trp 108 and Trp 111 as well as Trp 123 (and its neighbor Ile 124). Clearly, in these regions the simple approximation of uniform interactions between neighboring

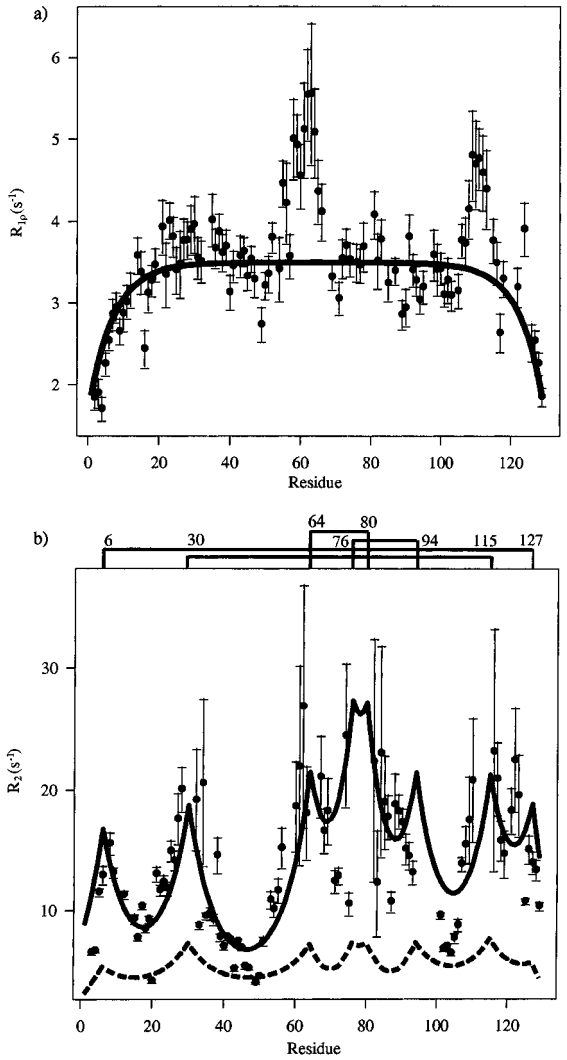

FIGURE 11: (a) Experimental transverse $R_{1\rho} = 1/T_{1\rho}$ relaxation rates for reduced denatured lysozyme fitted to the model for the dynamics of an unbranched polymer chain (eq 2 in the Materials and Methods). (b) Experimental transverse $R_2 = 1/T_2$ relaxation rates for oxidized denatured lysozyme fitted to the three parameter dynamics model (eq 3), which takes into account both the cross-linking of the chain by the disulfide bridges and the chemical exchange contributions (solid line). The predictions from the model if the chemical exchange contributions are removed is included for comparison (dashed line).

residues breaks down. We therefore conclude that interactions among bulky hydrophobic side chains introduces an additional degree of stiffness into the chain. The local hydrophobic clusters affect the relaxation behavior of a number of neighboring residues, the effects decaying rapidly with distance but in an asymmetric manner. It is interesting that not all the tryptophan residues in the sequence are involved in hydrophobic clusters that significantly alter the relaxation behavior of the polypeptide chain; two closely space tryptophan side chains or a tryptophan with an adjacent hydrophobic residue seem to be required.

Experimental transverse relaxation rates, R_2, for the oxidized form are shown in Figure 11b. These rates differ significantly from those of the reduced denatured protein. From a comparison of the $R_{1\rho}$ and R_2 rates of the oxidized protein (see Figure 6c,d), it is clear that, in contrast to the reduced protein, chemical exchange contributes significantly to the R_2 rates of amide groups near the disulfide bridges of the oxidized protein. In particular, the smallest values of R_2 for the oxidized protein (Figure 6d) are close to those of the reduced form, but R_2 values near the disulfide bridges are elevated by 25 s^{-1}. A similar observation of significant chemical exchange contributions to R_2 rates has been made previously for residues in the vicinity of disulfide bridges in native interleukin 4 (Redfield et al., 1992) and BPTI (Szyperski et al., 1993). In both these studies, increased line widths suggest that interconversion rates are in the millisecond time range. Thus, one contribution to the elevated R_2 rates observed here for oxidized denatured lysozyme could be slow conformational averaging around the χ_1 torsion angle of cysteine residues associated with changes in chirality of the disulfide bridges. Another contribution could come from the increased chain stiffness due to the additional conformational constraints arising from disulfide cross-linking and the complicated network of loops present in the oxidized form of the protein. We note that correlations have been observed previously between relaxation data for lysozyme denatured in trifluoroethanol and the positions of the disulfide bridges in the protein sequence (Buck et al., 1996).

To analyze the above mentioned chemical exchange and intrinsic relaxation contributions, the three parameter model introduced in the Material and Methods was used (eq 3). The model is similar to the one used to analyze the transverse relaxation rates for reduced lysozyme but has been modified to include extra terms which account for the disulfide cross-links and chemical exchange contributions. The intrinsic cross-linking effects of the disulfide bridges are modeled by using the topological distance between residues instead of simply the number of intervening residues along the primary sequence. Consequently, the chain is considered as a network of interconnecting loops rather than a linear array of residues. Chemical exchange is modeled as exponentially decreasing contributions arising from each of the cysteine residues involved in a disulfide bridge.

Experimental and predicted R_2 data using the approach are compared in Figure 11b. The figure shows that there is good agreement for much of the polypeptide chain suggesting that the presence of the disulfide bridges is indeed a major contributor to the experimentally observed R_2 pattern. We concentrate first on the residues far from the disulfide bridges, particularly in the long loops between residues 30 and 64 and between residues 94 and 115. Here, similar dynamical behavior is predicted to that for an unbranched chain in agreement with the experimentally observed R_2 and $R_{1\rho}$ rates in these regions which are about 5 s^{-1}, similar in magnitude to those of the reduced protein. In the vicinity of the disulfide bridges, however, the dynamic behavior is very different and exchange contributions need to be included in the model before the experimental data can be fitted successfully (solid line Figure 11b). The best fit is found with $R_2^{exch} = 10.9$ s^{-1}, $R_2^{intrin} = 0.29$ s^{-1} and $\lambda = 7$ residues. Chemical exchange effects therefore contribute significantly to the R_2 relaxation data of oxidized denatured lysozyme. We note that, if the model is fitted to the $R_{1\rho}$ data for the oxidized form, rather than the R_2 data, the R^{exch} values are found to be reduced by a factor of 3. In the central part of the sequence, where there is a complex network of disulfide cross-links, the fit of the model to the experimental R_2 data is less good, even with the inclusion of the exchange contributions, than at the termini of the polypeptide chain. This may in part be because the experimental line widths are so broad that R_2 values cannot be measured accurately; indeed in many cases they could not be measured at all.

Moreover, the dynamics around the branched disulfide bridges 64−80 and 76−94 will be very complicated and the simple models used here may not be sufficient to describe fully the relaxation behavior. One interesting observation from the current analysis is that both the predicted and the experimental R_2 values are larger at the C- than at the N-terminus. This reflects both the increased disulfide bridge density (Cys 115 and Cys 127) at the C-terminus and also the closer proximity of the C-terminus (Leu 129) to a disulfide bridge (Cys 127) than the N-terminus (Cys 6).

Overall Description of the Denatured State

The analysis of the experimental data for reduced and oxidized lysozyme in 8 M urea, presented in the previous sections, has used various models to describe a denatured protein. If we concentrate first on the global properties of the polypeptide chain, the analysis of the experimental relaxation data for the reduced protein demonstrates that in the absence of disulfide bridges the chain can be approximately described, from a dynamic point of view, as a homopolymer with few residue specific effects. An exception to this is in regions of the sequence where there are hydrophobic clusters involving some tryptophan residues which we consider in more detail later. The persistence length, a measure of the chain stiffness, which gives the best fit of the experimental data to the model is about 7 residues, a value comparable to the persistence length of 10 residues determined by X-ray scattering techniques for various denatured proteins, including apocytochrome *c* and phosphoglycerate kinase (Gast et al., 1995). It is particularly encouraging that two completely different approaches give rather similar values for the persistence lengths of denatured polypeptide chains.

For the oxidized protein, the experimental R_2 relaxation data can be explained if the presence of disulfide bridges cross-linking the chain and undergoing conformational exchange processes is taken into account. Such disulfide bridges will clearly restrict the conformational space accessible to a denatured polypeptide chain and reduce the effective radius of gyration. In order to visualize this, the distribution of the values of the radius of gyration within a conformational ensemble has been modeled for a random coil polypeptide chain with and without the disulfide cross-links present in oxidized lysozyme (Figure 12). The predicted distributions are scaled to be consistent with the experimental radius of gyration for oxidized denatured lysozyme determined by X-ray scattering (22 Å) (Chen et al., 1996). With this scaling, the average radius of gyration predicted for the reduced protein (∼30 Å) is considerably larger than that reported from the experimental X-ray scattering studies (24 Å). The reduced radius of gyration relative to the random coil prediction is fully consistent with the picture of the reduced denatured state of lysozyme provided by this work in which the hydrophobic clusters identified involving several of the tryptophan residues would be expected to lead to a higher proportion of compact conformers within the ensemble than expected for a complete random coil.

The broadening of some NMR resonances due to conformational exchange, as seen in oxidized denatured lysozyme, is a common feature of the NMR spectra of many non-native protein conformations whether or not disulfide bridges are present (Shortle, 1996). One example where the phenomenon has been characterized in detail, including the changes in line width with increasing concentrations of the denaturant urea, is the Δ131Δ fragment of staphylococcal nuclease where the resonances of many of the amide groups that are in the β sheet in the native protein are broadened to such an extent that they cannot be observed (Wang et al., 1995; Wang & Shortle, 1995). Similar effects, but involving much more of the polypeptide sequence, are seen in molten globules such as those of α-lactalbumins (Baum et al., 1989; Dobson, 1992). In these states, the conformational ensemble is biased toward more compact conformers and the presence of regions of secondary structure will increase the persistence length compared to a denatured polypeptide chain. This is likely to give rise to large energy barriers and hence slower rates for interconversion between structures within the conformational ensemble, resulting in extensive broadening of many of the NMR resonances (Dobson 1992; Ptitsyn, 1995). A comparable situation has been observed in NMR studies of a number of designed proteins where the interior side chains are less specifically packed than in native proteins, again resulting in slow conformational interconversion within a compact ensemble of states (Betz et al., 1993, 1995). Such states as these therefore lie in between the two extremes of fully denatured proteins, where fast interconversion between conformers gives rise to sharp NMR resonances, and native folded proteins, where there is in general one unique conformation, with only rapid vibrations and librations about the average coordinates, and sharp NMR resonances are once again observed.

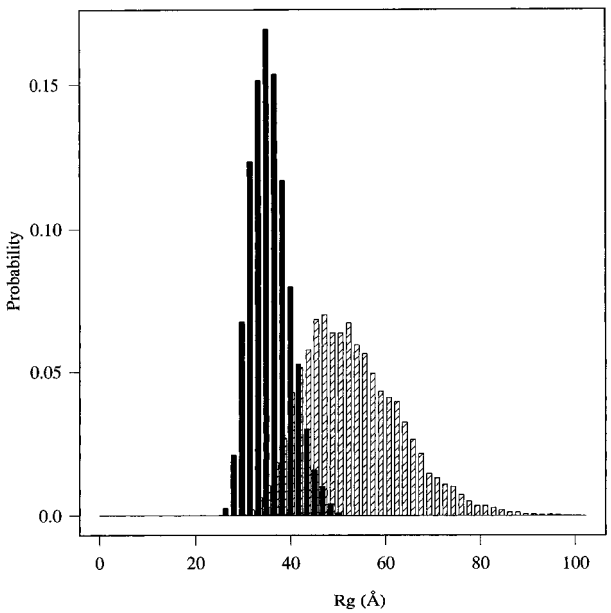

FIGURE 12: Model derived distribution of the radius of gyration (R_g) in the ensemble of random coil conformers for oxidized (filled bars) and reduced (dashed bars) lysozyme. Distributions were calculated from molecular dynamics simulations using a simple ball-and-stick model where each residue is represented by a 5 Å sphere. Residue−residue interactions were modeled as repulsive van der Waals forces with a cutoff at 7 Å. Peptide and disulfide bond constraints were simulated using the RATTLE algorithm, as described by Allen and Tildesley (1987), in conjunction with velocity Verlet molecular dynamics and coupling to a stochastic heat bath. Initial conformations of the oxidized form were generated by distance geometry projection. Distributions of R_g were extracted from ensembles of random coil conformations, which were generated at high temperature ($k_BT \gg E_{vdW}$). Subsequently, the distribution of R_g for the oxidized form was scaled such that the average R_g matched the experimental value of 22 Å reported by Chen et al. (1996). The same scaling factor was used for the reduced form.

Considering now the local conformational properties of denatured lysozyme, comparisons of predictions of coupling constants and NOEs using the ϕ,ψ population of all residues in the protein data base with the experimental data shows that to a first approximation these can again be described without considering the details of the amino acid sequence. The considerable number of NOEs identified experimentally, particularly both sequential $\alpha N(i, i + 1)$ and $NN(i, i + 1)$ NOEs and $\alpha N(i, i + 2)$ NOEs, and the mean experimental coupling constant value are in agreement with this simple model. However, extension to take into account the amino acid sequence of the polypeptide chain leads to a significantly improved correlation and can explain at least to a good approximation many experimentally observed features, such as the variation of $^3J(HN, H\alpha)$ coupling constants along the sequence and the nonuniform distribution of $scN(i, i + 3)$ NOEs. Overall, therefore, we have a picture of denatured lysozyme in which each amino acid in the polypeptide chain appears to sample conformational space largely independently of its neighbors, and the main chain torsion angle populations reflect predominantly the ϕ,ψ preferences of the individual amino acids involved. This picture assumes that interactions in denatured states of proteins are entirely local and their properties can therefore be predicted simply from the primary sequence of the protein.

Although the majority of data for denatured lysozyme agree rather closely with the predictions from the protein data base model for a random coil, some significant deviations are observed in specific regions of the sequence. Here, we use the predictions from the model as a framework for investigating these. We note that the presence of a considerable number of NOEs, particularly medium-range $(i, i + 2)$ and $(i, i + 3)$ NOEs, in NMR spectra of denatured proteins has frequently been interpreted as indicating the presence of persistent nonrandom structure. The predictions of the random coil model indicate that interpretation of such data is complex but that the data base predictions can be used as a baseline from which deviations from random coil behavior can be identified.

Figure 10 provides an overview of the NMR parameters found in the urea denatured states of lysozyme. In four regions of the protein, namely residues Asn 19−Phe 34, Ile 58−Cys 64, Met 105−Arg 112, and Asp 119−Arg 125, marked deviations from random coil predictions can be ascertained for a variety of parameters in both the oxidized and reduced protein. Three of these regions include the tryptophan residues Trp 62 and Trp 63, Trp 108 and Trp 111, and Trp 123 discussed above. The chemical shift perturbations for amide and Hα resonances and the increased $R_{1\rho}$ relaxation rates for residues in these regions are consistent with the presence of hydrophobic clusters involving adjacent aromatic and/or hydrophobic side chains. An aromatic cluster involving Trp 62 and Trp 63 has been proposed previously in thermally denatured oxidized lysozyme to explain large deviations in chemical shift values from those seen in short unstructured peptides (Evans et al., 1991); a similar hydrophobic cluster has also been identified in urea denatured 434-repressor (Neri et al., 1992). The results for denatured lyszoyme are also in close agreement with photo-CIDNP studies of the denatured protein, which showed that on average only one of the six tryptophan side chains in the sequence is available for reaction with photoexcited flavin (Broadhurst et al., 1991).

There is also some evidence from the chemical shift and coupling constant perturbations for the denatured states of lysozyme (Figure 10) for the presence of specific interactions involving residues in the region of Asn 19−Phe 34. We postulate an interaction between the aromatic ring of Tyr 20 and the amide group of Gly 22 at the start of this region. This type of interaction has been identified and shown to be preserved in urea, in short peptide fragments corresponding to this region of the lysozyme sequence (Yang et al., 1996; Smith, D. K., Pitkeathly, M., Smith, L. J., & Dobson, C. M., unpublished results). Such Tyr-X-Gly interactions have also been characterized in detail in peptides based on part of the BPTI sequence (Kemmink & Creighton, 1993, 1995; Lumb & Kim, 1994). It is interesting that this interaction, and whatever others are present in this region, do not give rise to considerable deviations in the relaxation rates of the reduced protein from the predictions of the simple dynamic model in contrast to the hydrophobic clusters involving the tryptophan residues. With regard to all these deviations from random behavior, we note that they are most readily identified from chemical shift perturbations and the heteronuclear relaxation rates rather than from deviations in the number or location of NOEs from the protein data base predictions for a random coil. Indeed, the regions of local structure present in denatured lysozyme do not appear to give rise to significant numbers of medium range NOEs, additional to those expected for a random coil, that can be identified in the current experimental spectra.

Although the NOEs observed for the urea denatured states of lysozyme in general resemble those expected for a random coil, the analysis of the NOE comparison shows that a smaller number of medium-range NOEs are observed overall even with high signal-to-noise ratios for residues 65−100 compared to the rest of the protein sequence in both proteins. This may partly reflect increased line widths for some of these residues. However, another possible explanation takes into account the fact that the random coil ensemble contains conformers with a range of global characteristics, as illustrated by the radius of gyration distributions shown in Figure 12. The intensities of any NOEs may not arise uniformly from the different conformers within this ensemble; in particular, the greatest intensity contribution is likely to come from the most compact states where fast internal motions will be quenched to some extent, giving longer effective correlation times. In a denatured protein, these compact states may not be randomly collapsed but instead certain regions of the sequence could be involved preferentially in a more compact core. If this was the case, the NOE intensity increase arising from the contribution of these compact states, via a mechanism analogous to the transferred NOE (Albrand et al., 1979; Clore & Gronenborn, 1982), would not be uniform along the sequence but would be concentrated around the residues involved in the core even if the population of such compact states was to be a small population of the complete ensemble of conformers. It is particularly interesting in this regard that for urea denatured lysozyme the regions with the highest density of observed NOEs, namely those corresponding to the A, B, D, and C-terminal 3_{10} helices in the native protein, are those regions that have been found to be involved in persistent structure in the equine lysozyme and homologous α-lactalbumin molten globules formed under milder denaturing conditions (Schulman et al., 1997; Morozova et al., 1995). This correlation suggests that it is possible that the more compact

conformers within the denatured lysozyme ensemble could resemble to some extent these molten globules, with a core consisting of residues corresponding to the native A, B, D, and 3_{10} helices while residues corresponding to the native C helix and β domain retain a higher degree of mobility.

Although further work is clearly required to understand the characteristics of the compact states in the conformational ensemble of denatured lysozyme in more detail, the proposed coupling of the collapse of the polypeptide chain to give compact conformers with the formation of some native-like characteristics even in the denatured protein is of particular interest. The fact that the same variation in NOE intensity with sequence is seen in both the oxidized and reduced denatured protein suggests that any nonrandom collapse is an intrinsic property of the sequence and is not directed by the disulfide bridges, although the cross-links are likely to stabilize the resulting compact structure. This interpretation of the experimental results does not require the presence of any specific local secondary structure in the compact states; indeed, the close resemblance of the observed NOEs to those predicted for a random coil suggest that these states are highly disordered at a local level. Moreover, the fact that the global relaxation data are not significantly affected by the nonrandom nature of the compact states indicates that the population of such conformers is low. The ability of these few states to influence significantly the NOE data of the ensemble, as discussed above, suggests that there are exciting possibilities for using NOEs to characterize certain compact species present within a conformational ensemble.

CONCLUSIONS

In this paper, through the analysis of experimental data from heteronuclear NMR techniques, we have been able to provide a detailed description of the structures of two denatured states of a protein considering both their local and global characteristics. One of the most important aspects of the analysis presented here has been the use of random coil models to provide a framework for interpreting the experimental data. For the two denatured states of lysozyme close similarities have been seen between the predictions from the random coil models and much of the experimental data, although differences have been observed for some specific regions of the sequence. At present, the extent to which such comparisons of experimental data with the models can be made is limited for the NOEs by resonance overlap and the signal-to-noise ratios present in the NMR spectra. Major advances, however, are being made as new heteronuclear NMR procedures are being applied to denatured proteins, such as the recently developed experiments for the unambiguous identification of side chain to side chain NOEs through indirect detection on well-resolved ^{15}N or ^{13}C resonances (Zhang et al., 1996). As the wealth of experimental data for denatured proteins increases, random coil models will not only be crucial for aiding the analysis of the data but will also be able to be developed further on the basis of the experimental NMR data.

Now that NMR procedures can be used to provide such detailed structural and dynamical information about denatured proteins comparisons with data from other experimental techniques, which are used to study these non-native states, also become important. Techniques such as X-ray scattering, through probing the global characteristics, provide information complementary to that coming from NMR data. Taken together, these techniques are therefore able to provide much insight into these unfolded states which have considerable significance for our understanding of protein stability, folding and structure. Of particular significance from the results reported here, is the possible coupling of the global and local characteristics of the conformational ensemble. Further insight into this may be provided through the analysis of the complex dependence of the NOEs observed for non-native protein conformations on both interproton distances and effective correlation times. It is possible that these effects may enable the characteristics of the most compact states in the conformational ensemble to be detected preferentially. These states are likely to be of very considerable value in our quest to understand the molecular basis of the formation of the native structures of globular proteins.

ACKNOWLEDGMENT

We thank Donald A. MacKenzie, David J. Jeenes, and David B. Archer of the Institute of Food Research, Norwich, U.K., for the supply of ^{15}N-labeled hen lysozyme. This is a contribution from the Oxford Centre for Molecular Sciences which is supported by the U. K. Engineering and Physical Sciences Research Council, the Biotechnology and Biological Sciences Research Council and the Medical Research Council. H.S. was supported by a Human Capital and Mobility grant from the European Commission. K.M.F acknowledges support from the Burrows Wellcome Fund by a Hitchings-Elion Fellowship. S.J.G. acknowledges a Heisenberg Stipend from the Deutsche Forschungsgemeinschaft. L.J.S. is a Royal Society University Research Fellow. The research of C.M.D. is supported in part by an International Research Scholars award from the Howard Hughes Medical Foundation.

SUPPORTING INFORMATION AVAILABLE

Table of ^{1}H and ^{15}N resonance assignments for oxidized and reduced hen lysozyme denatured in 8 M urea and table of heteronuclear relaxation rates [$R_1 = (T_1)^{-1}$, $R_2 = (T_2)^{-1}$, $R_{1\rho} = (T_{1\rho})^{-1}$] and heteronuclear NOEs determined for oxidized and reduced hen lysozyme denatured in 8 M urea (11 pages). Ordering information is given on any current masthead page.

REFERENCES

Abragam, A. (1961) *Principles of Nuclear Magnetism*, Clarendon Press, Oxford.
Akke, M., & Palmer, A. G., III (1996) *J. Am. Chem. Soc. 118*, 911−912.
Albrand, J. P., Birdsall, B., Feeney, J., Roberts, G. C. K., & Burgen, A. S. V. (1979) *Int. J. Biol. Macromol. 1*, 37−41.
Alexandrescu, A. T., Evans, P. A., Pitkeathly, M., Baum, J., & Dobson, C. M. (1993) *Biochemistry 32*, 1707−1718.
Alexandrescu, A. T., Abeygunawardana, C., & Shortle, D. (1994) *Biochemistry 33*, 1063−1072.
Allen, M. P., & Tildesley, D. J. (1987) *Computer simulations of liquids*, Clarendon Press, Oxford.
Allerhand, A., & Hailstone, R. K. (1972) *J. Chem. Phys. 56*, 3718−3720.
Arcus, V. L., Vuilleumier, S., Freund, S. M. V., Bycroft, M., & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9412−9416.
Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7−13.
Betz, S. F., Raleigh, D. P., & DeGrado, W. F. (1993) *Curr. Opin. Struct. Biol. 3*, 601−610.
Betz, S. F., Bryson, J. W., & DeGrado, W. F. (1995) *Curr. Opin. Struct. Biol. 5*, 457−463.

Boyd, J., Hommel, U., & Campbell, I. D. (1990) *Chem. Phys. Lett. 175*, 477−482.
Broadhurst, R. W., Dobson, C. M., Hore, P. J., Radford, S. E., & Rees, M. L. (1991) *Biochemistry 30*, 405−412.
Buck, M., Radford, S. E., & Dobson, C. M. (1993) *Biochemistry 32*, 669−678.
Buck, M., Radford, S. E., & Dobson, C. M. (1994) *J. Mol. Biol. 237*, 247−254.
Buck, M., Schwalbe, H., & Dobson, C. M. (1995a) *Biochemistry 34*, 13219−13232.
Buck, M., Boyd. J., Redfield, C., MacKenzie, D. A., Jeenes, D. J., Archer, D. B., & Dobson, C. M. (1995b) *Biochemistry 34*, 4041−55.
Buck, M., Schwalbe, H., & Dobson, C. M. (1996) *J. Mol. Biol. 257*, 669−683.
Bundi, A., & Wüthrich, K (1979) *Biopolymers 18*, 285−297.
Chen, L. L., Hodgson, K. O., & Doniach, S. (1996) *J. Mol. Biol. 261*, 658−671.
Ching-Li Lee, & Atassi, M. Z. (1973) *Biochemistry 12*, 2690−2695.
Clore, G. M., & Gronenborn, A. M. (1982) *J. Magn. Reson. 48*, 402−417.
Dill, K. A., & Shortle, D. (1991) *Ann. Rev. Biochem. 60*, 795−825.
Dobson, C. M. (1992) *Curr. Opin. Struct. Biol. 2*, 6−12.
Dobson, C. M., Evans, P. A., & Williamson, K. L. (1984) *FEBS Lett. 168*, 331−334.
Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 3542−3556.
Dyson, H. J., & Wright, P. E. (1993) *Curr. Opin. Struct. Biol. 3*, 60−65.
Evans, P. A., Topping, K. D., Woolfson, D. N., & Dobson, C. M. (1991) *Proteins: Struct., Funct., Genet. 9*, 248−266.
Farrow, N. A., Zhang, O., Forman-Kay, J. D., & Kay, L. E. (1995) *Biochemistry 34*, 868−878.
Fiebig, K. M., Schwalbe, H., Buck, M., Smith, L. J., & Dobson, C. M. (1996) *J. Phys. Chem 100*, 2661−2666.
Frank, M. K., Clore, G. M., & Gronenborn, A. M. (1995) *Protein Sci. 4*, 2605−2615.
Frenkiel, T., Bauer, C., Carr, M. D., Birdsall, B., & Feeney, J. (1990) *J. Magn. Reson. 90*, 420−425.
Gast, K., Damaschun, H., Eckert, K., Schulze-Forster, K., Maurer, H. R., Müller-Frohne, M., Zirwer, D., Czarnecki, J., & Damaschun, G. (1995) *Biochemistry 34*, 13211−13218.
Harbison, G. S. (1993) *J. Am. Chem. Soc. 115*, 3026−3027.
Harding, M. M., Williams, D. H., & Woolfson, D. N. (1991) *Biochemistry 30*, 3120−3128.
Hendrickson, W. A., & Wüthrich, K. (1996) *Macromolecular Structures 1995*, Current Biology Ltd., London.
Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 86*, 204−209.
Kataoka, M., & Goto, Y. (1996) *Folding Des. 1*, R107−R114.
Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110−126.
Kay, L. E., Torchia, D. A., & Bax, A. (1989) *Biochemistry 28*, 8972−8979.
Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 89*, 496−514.
Kemmink, J., & Creighton, T. E. (1993) *J. Mol. Biol. 234*, 861−878.
Kemmink, J., & Creighton, T. E. (1995) *J. Mol. Biol. 245*, 251−260.
Lattman, E. E. (1994) *Curr. Opin. Struc. Biol. 4*, 87−92.
Logan, T. M., Theriault, Y., & Fesik, S. W. (1994) *J. Mol. Biol. 236*, 637−648.
Lumb, K. J., & Kim, P. S. (1994) *J. Mol. Biol. 236*, 412−420.
MacKenzie, D. A., Spencer, J. A., Le Gal-Coëffet, M. F., & Archer, D. B. (1996) *J. Biotech. 46*, 85−93.
Merutka, G., Dyson, H. J., & Wright, P. E. (1995) *J. Biomol. NMR 5*, 14−24.
Morozova, L. A., Haynie, D. T., Arico-Muendel, C., van Dael, H., & Dobson, C. M. (1995) *Nat. Struct. Biol. 2*, 871−875.
Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992) *Proteins: Struct., Funct., Genet. 12*, 345−364.
Neri, D., Billeter, M., Wider, G., & Wüthrich, K. (1992) *Science 257*, 1559−1563.
Palmer, A. G., & Case, D. A. (1992) *J. Am. Chem. Soc. 114*, 9059−9067.
Pan, H., Barbar, E., Barany, G., & Woodward, C. (1995) *Biochemistry 34*, 13974−13981.
Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741−751.
Ptitsyn, O. B. (1995) *Adv. Protein Chem. 47*, 83−229.
Radford, S. E., Dobson, C. M., & Evans, P. A. (1992) *Nature 358*, 302−307.
Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963) *J. Mol. Biol. 7*, 95−99.
Redfield, C., Smith, L. J., Boyd, J., Lawrence, G. M. P., Edwards, R. G., Smith, R. A. G., & Dobson, C. M. (1991) *Biochemistry 30*, 11029−11033.
Redfield, C., Boyd, J., Smith, L. J., Smith, R. A. G., & Dobson, C. M. (1992) *Biochemistry 31*, 10431−10437.
Roder, H., Wagner, G., & Wüthrich, K. (1985) *Biochemistry 24*, 7407−7411.
Schleucher, J, Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S. J., Sørensen, O. W., & Griesinger, C. J. (1994) *J. Biomol. NMR 4*, 301−306.
Schulman, B. A., Kim, P. S., Dobson, C. M., & Redfield, C. (1997) *Nat. Struct. Biol.* (in press).
Serrano, L. (1995) *J. Mol. Biol. 254*, 322−333.
Shortle, D. R. (1996) *Curr. Opin. Struct. Biol. 6*, 24−30.
Smith, L. J., Fiebig, K. M., Schwalbe, H., & Dobson, C. M. (1996a) *Folding Des. 1*, R95−R106.
Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., & Dobson, C. M. (1996b) *J. Mol. Biol. 255*, 496−506.
Swindells, M. B., MacArthur, M. W., & Thornton, J. M. (1995) *Nat. Struct. Biol. 2*, 596−603.
Szyperski, T., Luginbuehl, P., Otting, G., Güntert, P., & Wüthrich, K. (1993) *J. Biomol. NMR 3*, 151−164.
Thornton, J. M., Jones, D. T., MacArthur, M. W., Orengo, C. M., & Swindells, M. B. (1995) *Philos. Trans. R. Soc. London B 348*, 71−79.
Wagner, G. (1993) *Curr. Opin. Struct. Biol. 3*, 748−754.
Wand, A. J., Roder, H., & Englander, S. W. (1986) *Biochemistry 25*, 1107−1114.
Wang, Y., & Shortle, D. (1995) *Biochemistry 34*, 15895−15905.
Wang, Y., Alexandrescu, A. T., & Shortle, D. (1995) *Philos. Trans. R. Soc. London B 348*, 27−34.
Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., & Sykes, B. D. (1995) *J. Biomol. NMR 5*, 67−81.
Wong, K. B., Freund, S. M. V., & Fersht, A. R. (1996) *J. Mol. Biol. 259*, 805−818.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K. (1994) *Curr. Opin. Struct. Biol. 4*, 93−99.
Yang, J. J., van der Berg, B., Pitkeathly, M., Smith, L. J., Bolin, K. A., Keiderling, T. A., Redfield, C., Dobson, C. M., & Radford, S. E. (1996) *Folding Des. 1*, 473−484.
Zhang, O., & Forman-Kay, J. D. (1995) *Biochemistry 34*, 6784−6794.
Zhang, O., Forman-Kay, J. D., Shortle, D., & Kay, L. E. (1996) *J. Biomol. NMR* (in press).

BI970049Q