Residual Structure in Unfolded Proteins Revealed by Raman Optical Activity

Gary Wilson, Lutz Hecht, and Laurence D. Barron*

Chemistry Department, The University, Glasgow G12 8QQ, U.K.

Received June 4, 1996[⊗]

ABSTRACT: Because of its ability to probe directly the chiral elements of the peptide backbone, together with the very short time scale of the scattering process, vibrational Raman optical activity (ROA) can provide new information on structure in non-native states of proteins. Here we report ROA studies of hen egg white lysozyme and bovine ribonuclease A in unfolded denatured states, prepared by reducing all the disulfide bonds. ROA spectra of unfolded lysozyme at 45, 20, and 2 °C, and of unfolded ribonuclease A at 35 and 20 °C, are presented and discussed. At 45 and 20 °C, unfolded lysozyme appears to contain very little extended secondary structure, but at 2 °C there could be roughly 20% of the native amount of α-helix present but little β-sheet. Unfolded ribonuclease A, on the other hand, appears to contain roughly 50% of its native-like secondary structure, including both α-helix and β-sheet, at 20 °C; similar secondary structure persists at 35 °C, but the amount is reduced. The most striking result is the observation of three sharp ROA bands in the extended amide III region, originating in coupled $C_\alpha-H$ and N−H deformations, which might monitor directly the dominant intrinsic propensities for residues to adopt particular ϕ,ψ angles, averaged over the different amino acids in the mobile heteropolypeptide. Specifically, positive bands at ∼1300 and 1314 cm^{-1} appear to monitor propensities for α-helix and β-structure, respectively, and a negative band at ∼1237 cm^{-1} appears to monitor that for the poly(L-proline) II helix. These signals are generated by individual residues clustering in the most favorable regions of the Ramachandran plot and are present even in the absence of signals from the corresponding extended secondary structures. At 45 °C, the 1300 and 1314 cm^{-1} ROA bands of unfolded lysozyme coalesce into a single sharp band from which an analysis similar to that used for exchange effects in NMR suggests a rate of $\sim 2.6 \times 10^{12}$ s^{-1} for interconversion between the individual residue conformations at this temperature.

Interest in non-native states of proteins is increasing rapidly on account of their importance in studies of protein folding, stability, and function (Kim & Baldwin, 1990; Dill & Shortle, 1991; Creighton, 1992; Shortle, 1993; Pain, 1994; Dill *et al.*, 1995). It is now recognized that the term "non-native", or "denatured", state embraces a plethora of structures ranging from ideal random coils at one extreme to collapsed molten globules with native-like tertiary folds at the other (Dobson, 1994). Unfortunately, the heterogeneity of such states has made their detailed characterization extremely difficult, with NMR[1] and far-UV CD providing most information to date. Recently, however, Raman optical activity (ROA), which measures vibrational optical activity by means of a small difference in the intensity of Raman scattering from chiral molecules in right and left circularly polarized incident light (Barron, 1982; Barron & Hecht, 1994; Nafie & Che, 1994), has emerged as an incisive new probe of biopolymer conformation in aqueous solution (Barron *et al.*, 1996) and has already provided valuable new insight into the complexity of order within the molten globule state (Wilson *et al.*, 1995, 1996a). Here we report ROA measurements on unfolded states of hen egg white lysozyme at 45, 20, and 2 °C and ribonuclease A at 35 and 20 °C which give new information on the ϕ,ψ propensities of the individual residues and reveal the existence of elements of secondary structure.

The reason ROA is proving so useful in this area is that it combines the very short time scale of conventional Raman scattering ($\sim 3.3 \times 10^{-14}$ s for a vibration with wavenumber 1000 cm^{-1}), which is much shorter than that of conformational changes, with the ability to cut through the complexity of conventional vibrational spectra and probe directly the chiral elements of the backbone and side groups. Thus ROA provides a superposition of "snapshot" spectra from all the distinct chiral conformers present in the sample with secondary structure, tertiary loop structure, and certain side groups giving rise to distinct recognizable bands which can be monitored independently. Furthermore, ROA signals from biopolymers appear to be generated by short-range mechanisms involving relatively local vibrational coupling, which suggests that they are closely related to the ϕ,ψ angles of the individual residues and so might monitor directly the ϕ,ψ propensities of amino acids in denatured states that are currently the focus of much interest (Swindells *et al.*, 1995; Serrano, 1995; Arcus *et al.*, 1995; Fiebig *et al.*, 1996). The importance of short-range mechanisms is reinforced by the realization that secondary structure irregularity and force constant variation, which is a *sine qua non* of heteropolypeptide structure, can suppress collective extended modes of vibration leaving localized motions to dominate the vibrational spectrum (Krimm & Reisdorf, 1994).

We have published previously and discussed briefly a room temperature ROA spectrum of unfolded hen egg white lysozyme (Ford *et al.*, 1995) acquired with an earlier generation of instrument. However, the spectra shown here

[†] This work was supported by the Biotechnology and Biological Sciences Research Council. L.D.B. is an Engineering and Physical Sciences Research Council Senior Fellow.

[⊗] Abstract published in *Advance ACS Abstracts,* September 1, 1996.

[1] Abbreviations: ROA, Raman optical activity; NMR, nuclear magnetic resonance; UV CD, ultraviolet circular dichroism; ESR, electron spin resonance; PPII, poly(L-proline) II.

S0006-2960(96)01314-1 CCC: $12.00

are of significantly higher quality, and the 45 and 2 °C spectra not measured before have turned out to be highly significant. The ROA of unfolded ribonuclease A has not been reported previously.

MATERIALS AND METHODS

We purchased hen egg white lysozyme from Sigma (grade I) and ribonuclease A from Fluka and used them without further purification. The buffer solutions were prepared using Analar grade chemicals and distilled and deionized water. Native lysozyme and ribonuclease A were studied in sodium acetate buffer at pH 5.4 and the reduced proteins in sodium citrate buffer in the pH range 2.0−3.0.

Unfolded lysozyme was prepared by reducing all four disulfide bonds using the method of Creighton (1989), modified at the purification stage to dialysis six times against 3 L of 0.01 M HCl. We used this method of denaturation because the other commonly used methods, namely, chemical denaturation and thermal denaturation, generate samples which are unsuitable for ROA measurements (chemically denatured proteins have intense Raman bands from the denaturants, and thermally denatured proteins often show considerable light scattering due to the formation of aggregates). Disulfide reduction is reversible so the free sulfhydryl groups are usually carboxymethylated to prevent reoxidation; however, we found that the resulting carboxymethylated protein was insufficiently soluble for ROA measurements, so we left the sulfhydryl groups unblocked and maintained a low pH to prevent reoxidation as was done in an NMR study of denatured lysozyme (Evans *et al.*, 1991). The solution was lyophilized and finally prepared for ROA measurements by dissolving in citrate buffer at pH 2.0. A similar procedure was followed to prepare reduced ribonuclease A except that citrate buffer at pH 2.6 was used because this sample tends to form a gel at lower pH. The number of free sulfhydryl groups was checked for both proteins using the Ellman assay (Creighton, 1989) and found to be between seven and eight per molecule before and after ROA data acquisition.

Protein solutions at ∼70 mg/mL, prepared in small glass sample tubes, were mixed with a little activated charcoal (pharmaceutical grade) to remove traces of fluorescing impurities and centrifuged. Solutions were subsequently filtered through Millipore GV4 (0.22 μm) filters directly into quartz microfluorescence cells, which were again centrifuged gently (to clear any remaining dust particles) prior to mounting in the ROA instrument. Residual visible fluorescence from remaining traces of impurities, which can give large background levels in Raman spectra, was allowed to "burn down" by leaving the sample to equilibrate for several hours in the laser beam before acquiring ROA data. This procedure, which is standard in the Raman spectroscopy of proteins (Cary, 1982), did not appear to harm our samples. For example, their conventional Raman spectra remained unchanged, and there was no slow buildup with time of the background fluorescence which usually accompanies photodegradation of proteins. Also, both native lysozyme and native ribonuclease A were found to maintain their full enzymatic activity after several days of exposure to the laser beam (as measured for lysozyme by the standard procedure based on cell wall hydrolysis of *Micrococcus lysodeikticus* and for ribonuclease A by that based on digestion of yeast nucleic acid).

The instrument used for the ROA measurements utilizes backscattering, which is essential for aqueous solution samples of biopolymers and employs a single-grating stigmatic spectrograph based on a novel transmission diffraction grating fitted with a backthinned CCD camera as detector and a holographic notch filter to block the Rayleigh line (Hecht & Barron, 1994; Barron *et al.*, 1996). ROA is measured by synchronizing the spectral acquisition with an electrooptic modulator used to switch the polarization of the incident argon ion laser beam between right and left circular at a suitable rate. The ROA spectra are presented in the form of a circular intensity difference $I^R - I^L$, where I^R and I^L are the Raman-scattered intensities in right and left circularly polarized incident light. The conventional Raman spectra are presented as the corresponding circular intensity sum $I^R + I^L$. The experimental conditions were as follows: laser wavelength 514.5 nm; laser power at the sample ∼700 mW; spectral bandwidth ∼12 cm^{-1}; ROA recording time ∼10 h.

The temperatures were established by blowing dry temperature-controlled air over the sample cell using an FTS Systems Model TC-84 air jet crystal cooler. The quoted temperatures will be underestimates by several degrees of the true sample temperatures along the laser beam since measurements were taken at the outside wall of the cell.

RESULTS AND DISCUSSION

Native Lysozyme and Ribonuclease A. The backscattered Raman and ROA spectra of native hen egg white lysozyme and bovine ribonuclease A measured at room temperature (∼20 °C) are shown in Figure 1. The main spectral regions are marked on the lysozyme ROA spectrum, while on that of ribonuclease A the wavenumbers of several important ROA bands are indicated. For a detailed discussion of the assignments of the various ROA bands we refer to Ford *et al.* (1995) and Barron *et al.* (1996). Small discrepancies (up to ∼5 cm^{-1}) in the quoted wavenumbers for some of the ROA bands between this and earlier protein ROA papers originate in different calibration protocols used with different spectrographs and also in uncertainties in some of the ROA band peak positions due to noise fluctuations.

ROA in the backbone C_α−C and C_α−N skeletal stretch region appears to be dominated by secondary structure. Thus α-helix generates positive intensity in the range ∼880−960 cm^{-1}, together with a couplet centered at ∼1100 cm^{-1}, negative at low wavenumber and positive at high (Wen *et al.*, 1994a; Wilson *et al.*, 1996b). Positive ROA intensity in the range ∼1000−1060 cm^{-1} has previously been associated with β-sheet (Wen *et al.*, 1994b), but a difficulty arose with this assignment when concanavalin A, which contains a large amount of multistranded antiparallel β-sheet, was found to have little ROA intensity in this region (Barron *et al.*, 1996). However, this difficulty could be resolved by our recent suggestion (Wilson *et al.*, 1996b) that this positive ROA band, together with negative intensity on either side with the complete pattern covering the range ∼950−1070 cm^{-1}, originates in the right-twisted antiparallel β-strand ribbon (equivalent to a left-handed helical conformation of the polypeptide backbone; Salemme, 1983; Maccallum *et al.*, 1995) that occurs in the double-stranded β-sheet as in ribonuclease A which shows an archetypal example of this ROA band pattern (Figure 1). In general, the average twist

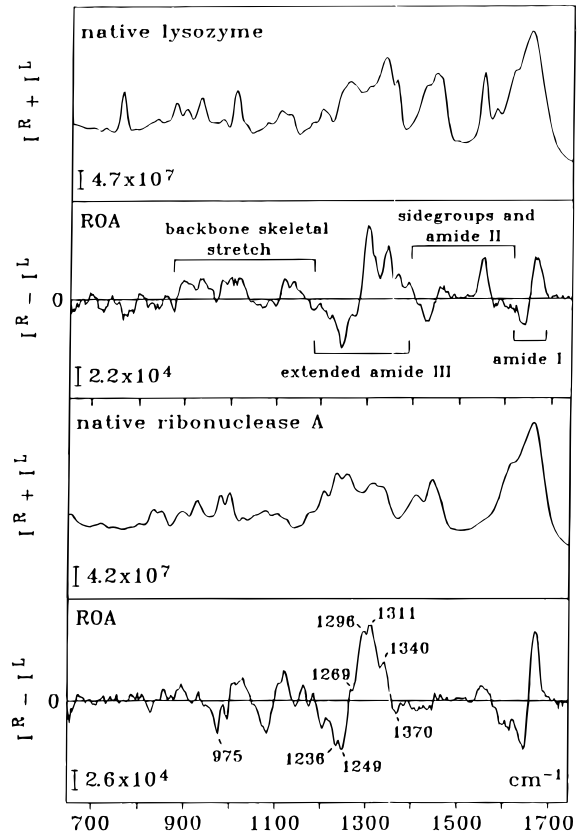

FIGURE 1: Backscattered Raman ($I^R + I^L$) and ROA ($I^R - I^L$) spectra of native hen egg white lysozyme (top pair) and of native bovine ribonuclease A (bottom pair), both in acetate buffer at pH 5.4. An approximate division into distinct wavenumber regions is indicated on the ROA spectrum of lysozyme, with the wavenumbers of some important bands indicated on that of ribonuclease A.

is largest for double-stranded ribbons and becomes increasingly suppressed the more strands are present and the longer they are (Richardson, 1982; Salemme, 1983), which could explain the absence of this band pattern in the ROA spectrum of concanavalin A. Support for this assignment comes from the normal mode calculations of Sengupta and Krimm (1987) on a left-handed extended helix conformation, which is very similar to a right-twisted β-strand, which gives a characteristic $C_\alpha-C$ stretch frequency slightly higher than that for the α-helix and so could well correspond to the negative ROA band peaking at ∼975 cm^{-1} in the ROA spectrum of ribonuclease A.

Both secondary and loop structure generate ROA signals in the extended amide III region. A positive peak at ∼1296 cm^{-1} with a smaller negative peak at ∼1265 cm^{-1} has been assigned to the α-helix (Wen *et al.*, 1994a), and a positive peak at ∼1313 cm^{-1} has been assigned to the β-sheet (Wen *et al.*, 1994b). There are two clear loop signatures in this region (Barron *et al.*, 1996): one, a positive band at ∼1340 cm^{-1}, has been assigned recently to a rigid loop structure with local order corresponding to that of a 3_{10}-helix (Wilson *et al.*, 1995, 1996b); the other, a negative band at ∼1237 cm^{-1}, was previously unassigned. However, as discussed below, from a suggested correlation of the dominant extended amide III peaks of unfolded lysozyme with the major ϕ,ψ propensities for individual residues in mobile heteropolypeptide structure, we now tentatively propose that the local order responsible for the negative ROA loop band at ∼1237 cm^{-1} is that of a left-handed poly(L-proline) II (PPII) helix. Also, the negative ROA bands in the range ∼1340−1380 cm^{-1}

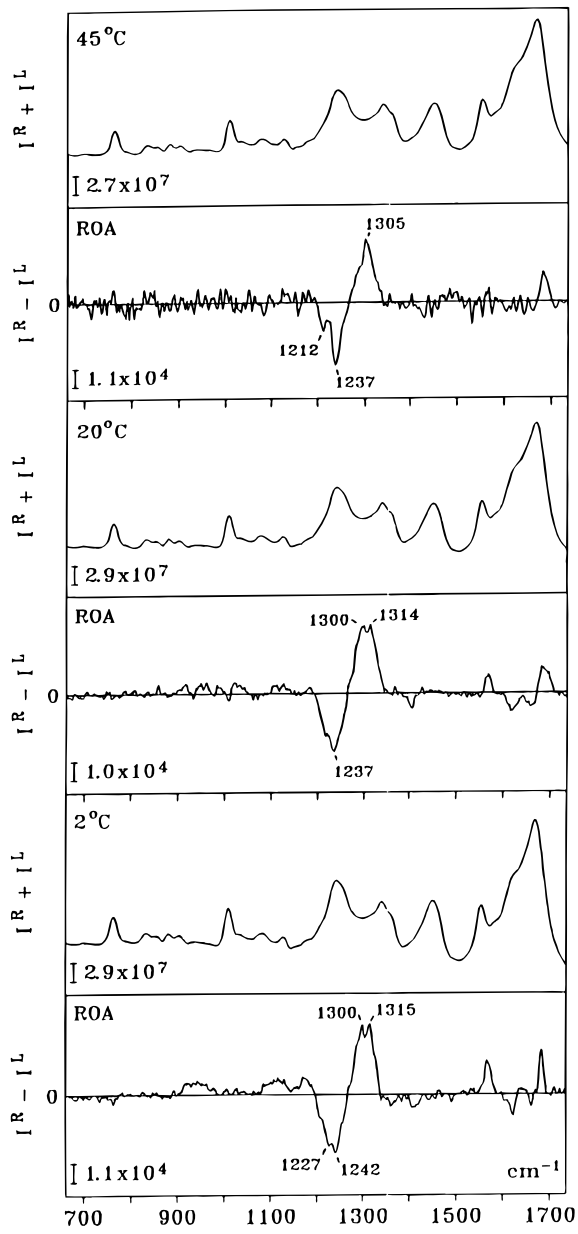

FIGURE 2: Backscattered Raman and ROA spectra of reduced hen egg white lysozyme in citrate buffer, pH 2.0, at 45 °C (top pair), 20 °C (middle pair), and 2 °C (bottom pair).

that are prominent in β-sheet proteins have been assigned to β-turns (Wen *et al.*, 1994b).

Unfolded Lysozyme. Figure 2 shows the backscattered Raman and ROA spectra of reduced lysozyme at 45 °C (top pair), 20 °C (middle pair) and 2 °C (bottom pair). Looking first at the 20 °C spectrum, it is striking that most of the ROA bands shown by native lysozyme have disappeared with the spectrum now dominated by a relatively unstructured extended amide III couplet. However, the small amount of structure remaining in this couplet could be of great significance because it appears to monitor the ϕ,ψ propensities of the individual residues averaged over the different amino acids in the mobile heteropolypeptide structure. Thus the two positive ROA bands at ∼1300 and 1314 cm^{-1} [which were not clearly resolved and went unnoticed in the earlier study of Ford *et al.* (1995)] are close to the extended amide III bands mentioned above which have been assigned to the α-helix and β-sheet, respectively, in native proteins. The negative band at ∼1237 cm^{-1} coincides with the second

extended amide III band assigned to the loop structure mentioned above; so if we identify the associated local order with that of the PPII helix, which is now recognized as a major conformational element in protein architecture (Adzhubei & Sternberg, 1993), we obtain a remarkable correlation of the three dominant extended amide III ROA peaks of unfolded lysozyme with the three main ϕ,ψ distribution peaks associated with α-helix, β-structure, and M conformation, the last corresponding to the PPII helix, determined from an analysis of all residues in a set of 68 proteins from the crystallographic structural database (Adzhubei *et al.*, 1987). In addition to the three main bands, a number of small bumps can be seen on the sides of the extended amide III ROA bands which might originate in smaller clusters of residues populating less favorable regions of ϕ,ψ space.

On reducing the temperature to 2 °C, the two positive extended amide III ROA bands have become sharper, while the negative band at ∼1237 cm^{-1} has split into two components, one at ∼1227 cm^{-1} and the other at ∼1242 cm^{-1}, suggesting that different types of local order are becoming differentiated in the putative PPII helix loop structure. Conversely, on raising the temperature to 45 °C, the positive α-helix and β-structure ROA bands at ∼1300 and 1314 cm^{-1} have coalesced into a single sharp band at ∼1305 cm^{-1}, which suggests that at this elevated temperature the distinction between the two local residue conformations is lost. Also the negative PPII helix peak at ∼1237 cm^{-1} has become sharper, and a negative ROA band at ∼1212 cm^{-1} (of unknown provenance at the time of writing) that appears as a shoulder in the lower temperature spectra now stands out clearly.

The coalescence and further narrowing of amide III ROA bands with increasing temperature in unfolded lysozyme are reminiscent of chemical exchange effects in NMR where closely spaced lines collapse into a single line at the temperature at which the lifetime broadening is greater than the line separation. A similar mechanism might be operative in the present ROA case, with the broadening produced by the finite lifetimes of the interconverting conformers. In fact, the effect of exchange in vibrational spectroscopy has attracted considerable interest and controversy (Wood & Strauss, 1990; Turner *et al.*, 1991; Strauss, 1992) and is still being hotly debated (Turner *et al.*, 1995). The basic problem appears to be that the time spent in the transition state during the exchange process, and which is very short on the NMR time scale, is on the vibrational time scale, which means that the "reactive" motion involving transitions between the potential energy wells associated with the separate conformers is mixed up with "nonreactive" motion involving large amplitude vibrations of a particular conformer within its own well. This suggests that the simple analysis used in NMR is not directly applicable to the vibrational case. But despite this difficulty, some authors nonetheless use an analysis analogous to that for exchange in NMR to deduce approximate values for rate constants at the temperature at which vibrational band coalescence is observed [*e.g.*, Cavagnat *et al.* (1992) and Turner *et al.* (1995)], and we shall do the same here. Thus from the uncertainty relation, the lifetime broadening is given by $\delta E \approx h/2\pi\tau$, where τ is the lifetime, so that two spectral bands characteristic of two interconverting states and separated in frequency by $\delta\nu$ will coalesce when τ is of the order of $1/2\pi\delta\nu$. The rate constant $k = 1/\tau$ is therefore of the order of $2\pi\delta\nu$, so we deduce from the ∼14 cm^{-1} separation of the α-helix and β-structure ROA bands that the rate constant at 45 °C for interconversion between the corresponding conformers is $\sim 2.6 \times 10^{12}$ s^{-1}, which happens to be of the same order as the rate of rotations involving changes in ϕ and ψ about $C_\alpha-N$ and $C_\alpha-C$ bonds, respectively, in single residues in the polypeptide backbone (Creighton, 1993).

Recently, Swindells *et al.* (1995) found that there were three main regions of ϕ,ψ space for individual residues in "coil" regions of native proteins which they deduced by including all residues not in the α-helix or β-strand. However, these regions, called a, b, and p, seem too diffuse to correspond with the sharp appearance of the three extended amide III ROA bands of unfolded lysozyme discussed above, which suggests that the actual propensities are more tightly focused in a mobile heteropolypeptide in aqueous solution than in coil regions of native proteins where constraints from the structured environment can distort the intrinsic propensities.

One feature conspicuous by its absence from the ROA spectrum of unfolded lysozyme is any hint of the sharp positive band at ∼1340 cm^{-1} in the native protein assigned to the 3_{10}-helix loop structure. We have previously noticed that 3_{10}-helix loops appear to be labile and are often associated with mobile regions in both native and molten globule protein states, the nascent rigid structure being encouraged to form by stabilizing perturbations such as ligand binding or temperature reduction (Wilson *et al.*, 1995, 1996a). It therefore seems that, unlike the PPII helix loop structure, there is little propensity for individual residues to take up this conformation in mobile unfolded protein states despite appearing in coil regions in native proteins in the structural database.

The backbone skeletal stretch region of unfolded lysozyme at 20 °C shows only slight hints of ROA bands, indicating that there is very little extended secondary structure present. However, the ROA spectrum at 2 °C shows small but clear signatures of α-helix, namely, the broad band peaking at ∼950 cm^{-1} and the positive component on the high wavenumber side of the couplet at ∼1100 cm^{-1} shown by native lysozyme (Figure 1). The intensities relative to that of the native protein suggest that there might be, very roughly, ∼20% of the native amount of α-helix present. However, there is no hint of any twisted β-strand structure. The tryptophan ROA couplet, positive at ∼1555 cm^{-1} and negative at ∼1580 cm^{-1}, which appears weakly in the 20 °C spectrum, has increased in step with the α-helix bands in the 2 °C spectrum, suggesting that the conformational freedom of the tryptophans associated with the sections of the α-helix has been reduced. These α-helix sections might include residues corresponding to the C (88−99) and D (108−115) helices in the native protein which, from UV CD studies on a peptide comprising residues 84−129 of hen egg white lysozyme in water at pH 2.0 and 25 °C, have been shown to have a high intrinsic propensity for α-helix formation (Yang *et al.*, 1995). In another study, this time of a peptide comprising residues 36−108, it was again found that residues corresponding to helix C had a high propensity for α-helix formation (Ueda *et al.*, 1994). At 45 °C virtually all traces of the secondary structure and tryptophan ROA bands have disappeared into the noise background.

A small amide I couplet remains in the 20 °C (and also in the 45 °C) spectrum of unfolded lysozyme but is shifted

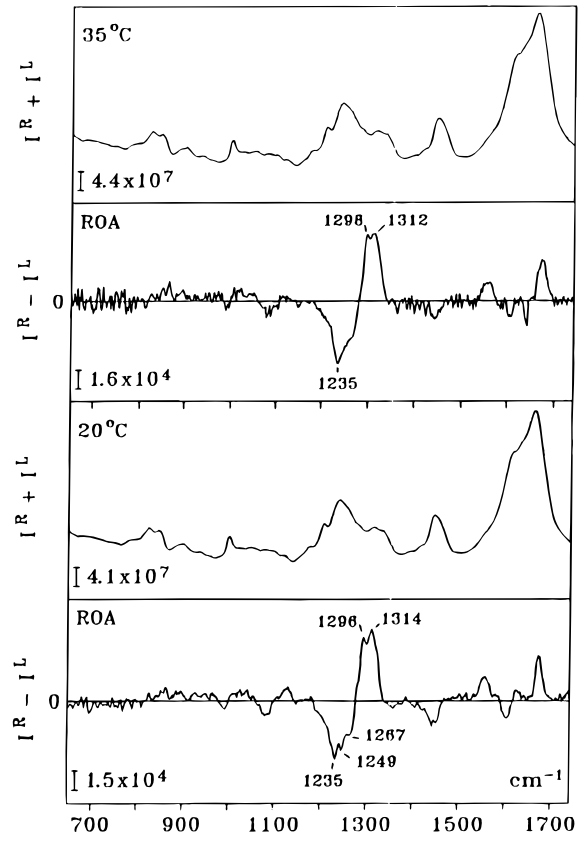

FIGURE 3: Backscattered Raman and ROA spectra of reduced bovine ribonuclease A in citrate buffer, pH 2.6, at 35 °C (top pair) and 20 °C (bottom pair).

along with its parent Raman band to higher wavenumber relative to that in the native protein, in accord with the amide I assignment of "random coil" in the conventional Raman spectroscopy of proteins (Tu, 1986). This residue amide I couplet presumably arises from the local chiral environments of the carbonyl oscillators. The negative ROA band at ~1620 cm^{-1} originating in aromatic side groups is now clearly resolved from the negative part of the shifted amide I couplet.

Unfolded Ribonuclease A. Figure 3 shows the Raman and ROA spectra of reduced ribonuclease A at 35 °C (top pair) and 20 °C (bottom pair). Unfortunately, it was not possible to measure the ROA at temperatures much lower or much higher than these due to excessive light scatter from gel formation and aggregation, respectively. It is immediately apparent that unfolded ribonuclease A has more ordered structure at 20 °C than unfolded lysozyme at this temperature. In the extended amide III region the three bands assigned above to individual residues occupying regions of ϕ,ψ space associated with α-helix, β-structure, and PPII helix are quite clear. However, there are small but potentially significant differences from the corresponding ROA bands in unfolded lysozyme at 20 °C. Thus in ribonuclease A the positive α-helix band, as well as being slightly weaker than the positive β-structure band, is shifted by ~4 cm^{-1} to lower wavenumber relative to its position in lysozyme, which could be due to the presence of a significant amount of α-helical secondary structure (the amide III signature of α-helical secondary structure in native proteins is a couplet with a positive band at ~1296 cm^{-1} and a negative band at ~1265 cm^{-1}, which is clearly perceptible in the ROA spectrum of unfolded ribonuclease A at 20 °C). This positive α-helix band shows a small but definite shift by ~2 cm^{-1} to higher wavenumber in the 35 °C ROA spectrum of ribonuclease A, where it also now matches the height of the adjacent β-structure band, concomitant with a general decrease in other α-helix ROA bands (*vide infra*), which suggests that the band seen at ~1300 cm^{-1} in unfolded lysozyme at 20 °C is characteristic of individual residues taking up the ϕ,ψ angles of α-helix but without secondary structure. Also there is a small negative band at ~1360 cm^{-1}, suggesting the presence of a small amount of β-turn structure. As in unfolded lysozyme, unfolded ribonuclease A shows no hint of the positive ROA band at ~1340 cm^{-1}, assigned to the 3_{10}-helix loop structure, that is present in the native protein. In fact, it is striking just how narrow and smooth-sided most of the extended amide III ROA couplet has become at 35 °C, with no sign of the many small bumps perceptible here in unfolded lysozyme at all temperatures, which suggests there are very few residues taking up ϕ,ψ angles other than those corresponding to the three main bands.

In the backbone skeletal stretch region, unfolded ribonuclease A at 20 °C shows ROA band structure in the range ~950−1150 cm^{-1}, similar to that of the native protein but reduced in intensity by ~50%. This suggests that this sample of the unfolded protein contains ~50% of secondary structure similar, but not identical, to that in the native protein, including both right-twisted β-strand and α-helix. Some of this α-helix might be associated with the amino-terminal segment since residual structure has been detected here in reduced ribonuclease A by nonradiative energy transfer (Haas *et al.*, 1988), the C-peptide of ribonuclease A (residues 1−13) was found using UV CD to populate helical conformations to ~30% at 1.7 °C (Brown & Klee, 1971), and helical conformations have been detected using two-dimensional NMR in an analogue of this C-peptide (Osterhout *et al.*, 1989). On raising the temperature to 35 °C, hints of the ROA signatures of secondary structure in the backbone skeletal stretch region are just perceptible, suggesting that a small amount of secondary structure is still present.

The Two-Group Model and Mobile Heteropolypeptide ROA. A possible qualitative explanation for the general appearance of the ROA spectrum of an unfolded protein, in particular, the loss of ROA in the backbone skeletal stretch region and the sharp peaks in the extended amide III region reflecting local conformational propensities of individual residues, emerges from a consideration of the dependence of ROA on torsion angles. Many prominent protein ROA signals appear to originate in short-range vibrational coupling for which a simple two-group model provides useful insight (Barron & Buckingham, 1974). In this model, two achiral axially symmetric bonds are held together in a twisted chiral structure with ROA being generated by interference at the detector between two light waves scattered independently from the two groups. In the special case where the axes of the two bonds are perpendicular to the connecting bond, the ROA intensities $I^R - I^L$ generated by idealized normal modes corresponding to the symmetric (+) and antisymmetric (−) combinations of the local *stretch* coordinates of the two bonds are proportional to $\pm\sin 2\theta$, where θ is the torsion angle, whereas for idealized normal modes corresponding to the *deformations* of the angles between the two bonds and the connecting bond the ROA intensities are proportional to $\mp\sin\theta$ (Barron, 1982). For a more general geometry, the ROA in stretch and deformation modes contains terms

depending on both sin 2θ and sin θ. A number of chiral two-group structures can be identified in the polypeptide backbone. In particular, the arrangement H−C_α−N−H is likely to dominate the ROA in the extended amide III region since this is associated largely with coupled C_α−H and N−H deformations (Diem, 1993; Ford *et al.*, 1994), while the arrangements C_α−C(O)−N−C_α, C(O)−N−C_α−C(O), N−C_α−C(O)−N, C_β−C_α−N−C(O), and C_β−C_α−C(CO)−N will all contribute to ROA in the backbone skeletal stretch region. ROA in the extended amide III region is therefore likely to depend primarily on the Ramachandran ϕ angle which is related to the two-group torsion angle through $\theta = \phi - 60°$. ROA in the backbone skeletal stretch region, on the other hand, is expected to have a complicated dependence on both the ϕ and ψ angles. Hence much less ROA intensity in the backbone skeletal stretch region is expected to survive than in the extended amide III region as the residues explore the range of ϕ and ψ angles characteristic of mobile heteropolypeptide structure.

Essentially, then, the exquisite sensitivity of ROA to conformational mobility is a direct consequence of its dependence on absolute chirality which leads to cancellation of contributions from enantiomeric structures such as two-group units with equal and opposite torsion angles, whereas contributions to observables which are "blind" to chirality, such as the conventional Raman band intensities, are additive. This explains why the conventional Raman spectra of reduced lysozyme at the three different temperatures in Figure 2, and those of reduced ribonuclease A at the two temperatures in Figure 3, are virtually identical despite significant changes appearing in the corresponding ROA spectra. UV CD is also sensitive to mobility, but less dramatically so than ROA due to its primary dependence on electronic rather than nuclear motion, because it too depends on absolute chirality.

We emphasize that, although the two-group model has conceptual value, because of the complexity of biopolymer normal modes of vibration (even if relatively localized), it has little practical value in predicting observed ROA spectral features. *Ab initio* methods appear to be mandatory for realistic ROA computations (Polavarapu, 1990; Helgaker *et al.*, 1994), but the application to peptides and proteins (Polavarapu & Deng, 1994) is not yet developed sufficiently to give reliable interpretations.

CONCLUDING REMARKS

The first ROA spectra of unfolded proteins presented here suggest that ROA provides a unique perspective on residual structure, sensing both the ϕ,ψ propensities of individual residues in mobile unordered heteropolypeptide structure and the presence of sections of extended secondary structure. However, because the time scale of the Raman scattering process is very short, it is not possible to tell from the ROA spectra whether the sections of secondary structure are transient or persist for a significant length of time.

The unfolded states of lysozyme and ribonuclease A at room temperature show rather different characteristics. Unfolded lysozyme is highly mobile with individual residues showing high propensities for the ϕ,ψ angles associated with the α-helix, β-structure, and PPII helix, whereas superimposed on these propensities in unfolded ribonuclease A is a significant amount of native-like secondary structure. However, since secondary structure starts to become perceptible in unfolded lysozyme at 2 °C, it would appear that stretches of extended ordered structure are nascent and might be encouraged to form with changes in the physical or chemical conditions (through suppression of mobility from loss of solvation following hydrophobic collapse, for example). These ROA observations reinforce recent NMR data which fit models of protein folding in which local nucleation sites are latent in the denatured state (Arcus *et al.*, 1995). It would be of great interest to characterize the nature of the transition by following the changes in the ROA spectra of selected unfolded proteins over a range of temperatures in conjunction with differential scanning calorimetry measurements to discover whether it is gradual or, if cooperative, whether the associated phase transition is first order or continuous. Such studies would complement those on the development of the native-like tertiary fold in the molten globule state which we have recently shown is controlled by a continuous phase transition (Wilson *et al.*, 1996a). Also, studies of unfolded states in nonaqueous solvents might reveal whether or not solvation by water molecules is an essential requirement for the maintenance of "dynamic disorder" in heteropolypeptides.

It is remarkable that, by observing the coalescence and narrowing with increasing temperature of the ϕ,ψ propensity ROA bands in the extended amide III region, it is possible to monitor the interconversion of conformers at rates $\sim 10^{12}$ s^{-1}. Since the simplest version of absolute reaction rate theory predicts a maximum rate for a chemical process of 6.2×10^{12} s^{-1} at 25 °C corresponding to a zero free energy barrier (Creighton, 1993), or a little less than this when collisional damping is included (Brooks *et al.*, 1988), our observations suggest that, at physiological temperatures, the individual residues in mobile heteropolypeptide structure in unfolded lysozyme "flicker" between distinct conformational states at rates close to the maximum theoretically possible. Interestingly, our ROA studies also indicate that unordered homopolypeptides such as poly(L-lysine) at low pH behave quite differently from unordered heteropolypeptides since, as well as containing sections of well-defined conformational elements such as an α-helix and, possibly, a left-handed helix corresponding to a right-twisted β-strand, there is no hint of the extremely rapid interconversion of conformers observed in unfolded lysozyme (Wilson *et al.*, 1996b). These ROA exchange effects clearly warrant more detailed studies in the future since they appear to be a source of valuable new information about very fast processes in proteins.

The absence of any sign of a positive band at $\sim$1340 cm^{-1} in the ROA spectra of the two unfolded proteins which appear in the ROA spectra of the native proteins where it is characteristic of 3_{10}-helix loop structure reinforces our earlier conclusions that the 3_{10}-helix is finely tuned and can only exist in a relatively structured environment (Wilson *et al.*, 1995, 1996a,b). This result is consistent with recent ESR studies (Smythe *et al.*, 1995a) which concluded that short alanine-based peptides exist primarily in the α-helical conformation and not in the 3_{10} form as previously suggested (Miick *et al.*, 1992). It would therefore appear that the propensity for individual residues to take up the 3_{10}-helix ϕ,ψ angles in the mobile heteropolypeptide structure is virtually zero, which seems to militate against the suggestion that the 3_{10}-helix is a precursor to the α-helix in the folding pathway (Millhauser, 1995). A related suggestion is that local α- to 3_{10}-helical transitions could be important for a

number of biological processes (Smythe *et al.*, 1995b). In fact, we have recently used ROA to monitor changes in the structure of lysozyme on ligand binding (Ford *et al.*, 1995) and do indeed detect a possible increase in 3_{10}-helix (*via* an increase in the positive ~1340 cm^{-1} ROA band). However, there appears to be no interconversion with α-helix since there is no concomitant decrease in the α-helix ROA bands; rather, the 3_{10}-helix seems to coalesce out of the mobile conformationally heterogeneous structure. Nor does there appear to be any interconversion of α- and 3_{10}-helix between the native and acid molten globule states of α-lactalbumin (Wilson *et al.*, 1995, 1996a) despite the plausible suggestion (Smythe *et al.*, 1995b) that it would facilitate the transition between the associated compact and expanded structures.

On a related theme, recent work by Minor and Kim (1996) has shown that a particular 11 amino acid sequence folds as an α-helix when in one position but as a β-sheet when in another position of the primary sequence of protein GB1, which suggests that even these two core elements of secondary structure are quite finely tuned and susceptible to perturbations from the environment. However, the present ROA study has indicated that, unlike the 3_{10}-helix, both the α-helix and β-sheet can exist as an extended structure in the unfolded state which is probably due to the high propensities for individual residues to take up the corresponding regions of ϕ,ψ space. It is intriguing that the intrinsic α-helix and β-structure (and indeed PPII helix) propensities averaged over the different individual residues in unfolded lysozyme at 20 °C and ribonuclease A at 35 °C are very similar despite the quite different sequences in these two proteins. This could well be connected with Minor and Kim's discovery since it suggests that factors other than intrinsic residue propensities ultimately determine the actual conformation adopted by a particular stretch of the peptide backbone in a native protein. Similarly, Parthasarathy *et al.* (1995) have stressed recently that it is the sequence context of residues rather than their individual fixed propensities that determines helix formation. All this reinforces Minor and Kim's conclusion that their results underscore the importance of context-dependent effects in protein folding; however, their more specific suggestion that their results support views of protein folding that favor tertiary interactions as dominant determinants of structure (Dill *et al.*, 1995) is not consistent with the results of ROA studies of non-native states of α-lactalbumin (Wilson *et al.*, 1996a) which reveal that an almost native-like complement of secondary structure can exist in the collapsed globule in the absence of any tertiary loop structure (*e.g.*, in the apo state at neutral pH at 60 °C and in the acid molten globule state at 35 °C). Indeed, from ROA observations of the development of a highly native-like tertiary fold in acid molten globule α-lactalbumin around a core of native-like secondary structure with decreasing temperature, the opposite would appear to obtain; namely that once native-like secondary structure has become established in a collapsed globule, it can predetermine the formation of the correct native-like tertiary fold (Wilson *et al.*, 1996a). Similarly, recent studies of helix formation using time-resolved infrared spectroscopy (Williams *et al.*, 1996) support the notion that secondary structure forms before long-range tertiary contacts are made and that it is not necessary for a peptide sequence to have a strong propensity for helix formation in aqueous solution in order to serve as a nucleation site.

The first ROA measurements on unfolded proteins presented here have shown that ROA fills a crucial gap in the armory of existing experimental techniques in that it can monitor separately the small populations of conformers with distinct secondary structure, together with the intrinsic ϕ,ψ propensities averaged over the different individual residues, within a conformational ensemble of a heteropolypeptide and which are of crucial importance in the early stages of protein folding. More secure assignments of some of the ROA spectral features of unfolded proteins and a better understanding of how they are generated will evolve in the course of our ongoing ROA studies of non-native protein states.

ACKNOWLEDGMENT

We thank Drs. A. Cooper, D. D. MacNicol, and J. K. Tyler for helpful discussions.

REFERENCES

Adzhubei, A. A., & Sternberg, M. J. E. (1993) *J. Mol. Biol. 229*, 472−493.
Adzhubei, A. A., Eisenmenger, F., Tumanyan, V. G., Zinke, M., Brodzinski, S., & Esipova, N. G. (1987) *J. Biomol. Struct. Dyn. 5*, 689−704.
Arcus, V. L., Vuilleumier, S., Freund, S. M. V., Bycroft, M., & Fersht, A. R. (1995) *J. Mol. Biol. 254*, 305−321.
Barron, L. D. (1982) *Molecular Light Scattering and Optical Activity*, Cambridge University Press, Cambridge.
Barron, L. D., & Buckingham, A. D. (1974) *J. Am. Chem. Soc. 96*, 4769−4773.
Barron, L. D., & Hecht, L. (1994) in *Circular Dichroism Principles and Applications* (Nakanishi, K., Berova, N., & Woody, R., Eds.) pp 179−215, VCH Publishers, New York.
Barron, L. D., Hecht, L., & Bell, A. F. (1996) *Appl. Spectrosc. 50*, 619−629.
Brooks, C. L., Karplus, M., & Pettitt, B. M. (1988) *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics* (*Adv. Chem. Phys. 71*), Wiley, New York.
Brown, J. E., & Klee, W. A. (1971) *Biochemistry 10*, 470−476.
Carey, P. R. (1982) *Biochemical Applications of Raman and Resonance Raman Spectroscopies*, Academic Press, New York.
Cavagnat, D., Banisaeid-Vahedie, S., Lespade, L., & Rodin, S. (1992) *J. Chem. Soc., Faraday Trans. 88*, 1845−1861.
Creighton, T. E. (1989) in *Protein Structure: a Practical Approach* (Creighton, T. E., Ed.) pp 155−167, IRL Press, Oxford.
Creighton, T. E., Ed. (1992) *Protein Folding*, W. H. Freeman and Co., New York.
Creighton, T. E. (1993) *Proteins*, W. H. Freeman and Co., New York.
Diem, M. (1993) *Modern Vibrational Spectroscopy*, Wiley, New York.
Dill, K. A., & Shortle, D. (1991) *Annu. Rev. Biochem. 60*, 795−825.
Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D., & Chan, H. S. (1995) *Protein Sci. 4*, 561−602.
Dobson, C. M. (1994) *Curr. Biol. 4*, 636−640.
Evans, P. A., Topping, K. D., Woolfson, D. N., & Dobson, C. M. (1991) *Proteins: Struct., Funct., Genet. 9*, 248−266.
Fiebig, K. M., Schwalbe, H., Buck, M., Smith, L. J., & Dobson, C. M. (1996) *J. Phys. Chem. 100*, 2661−2666.
Ford, S. J., Wen, Z. Q., Hecht, L., & Barron, L. D. (1994) *Biopolymers 34*, 303−313.
Ford, S. J., Cooper, A., Hecht, L., Wilson, G., & Barron, L. D. (1995) *J. Chem. Soc., Faraday Trans. 91*, 2087−2093.
Haas, E., McWherter, C. A., & Scheraga, H. A. (1988) *Biopolymers 27*, 1−21.
Hecht, L., & Barron, L. D. (1994) *Faraday Discuss. 99*, 35−47.
Helgaker, T., Ruud, K., Bak, K. L., Jorgensen, P., & Olsen, J. (1994) *Faraday Discuss. 99*, 165−180.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.

Krimm, S., & Reisdorf, W. C., Jr. (1994) *Faraday Discuss. 99*, 181−197.
Maccallum, P. H., Poet, R., & Milner-White, E. J. (1995) *J. Mol. Biol. 248*, 374−384.
Miik, S. M., Martinez, G. V., Fiori, W. R., Todd, A. P., & Millhauser, G. L. (1992) *Nature 359*, 653−655.
Millhauser, G. L. (1995) *Biochemistry 34*, 3873−3877.
Minor, D. L., & Kim, P. S. (1996) *Nature 380*, 730−734.
Nafie, L. A., & Che, D. (1994) *Adv. Chem. Phys. 85* (3), 105−149.
Osterhout, J. J., Jr., Baldwin, R. L., York, E. J., Stewart, J. M., Dyson, H. J., & Wright, P. E. (1989) *Biochemistry 28*, 7059−7064.
Pain, R. H., Ed. (1994) *Mechanisms of Protein Folding*, IRL Press, Oxford.
Parthasarathy, R., Chaturvedi, S., & Go, K. (1995) *Prog. Biophys. Mol. Biol. 64*, 1−54.
Polavarapu, P. L. (1990) *J. Phys. Chem. 94*, 8106.
Polavarapu, P. L., & Deng, Z. (1994) *Faraday Discuss. 99*, 151−163.
Richardson, J. S. (1982) *Adv. Protein Chem. 34*, 167−339.
Salemme, F. R. (1983) *Prog. Biophys. Mol. Biol. 42*, 95−133.
Sengupta, P. K., & Krimm, S. (1987) *Biopolymers 26*, S99−S107.
Serrano, L. (1995) *J. Mol. Biol. 254*, 322−333.
Shortle, D. (1993) *Curr. Opin. Struct. Biol. 3*, 66−74.
Smythe, M. L., Nakaie, C. R., & Marshall, G. R. (1995a) J. *Am. Chem. Soc. 117*, 10555−10562.
Smythe, M. L., Huston, S. E., & Marshall, G. R. (1995b) *J. Am. Chem. Soc. 117*, 5445-5452.
Strauss, H. L. (1992) *J. Am. Chem. Soc. 114*, 905−907.
Swindells, M. B., MacArthur, M. W., & Thornton, J. M. (1995) *Nat. Struct. Biol. 2*, 596−603.
Tu, A. T. (1986) Adv. Spectrosc. 13, 47−112.
Turner, J. J., Grevels, F.-W., Howdle, S. M., Jacke, J., Haward, M. T., & Klotzbücher, W. E. (1991) *J. Am. Chem. Soc. 113*, 8347−8353.
Turner, J. J., Gordon, C. M., & Howdle, S. M. (1995) *J. Phys. Chem. 99*, 17532−17538.
Ueda, T., Nakashima, A., Hashimoto, Y., Miki, T., Yamada, H., & Imoto, T. (1994) *J. Mol. Biol. 235*, 1312−1317.
Wen, Z. Q., Hecht, L., & Barron, L. D. (1994a) *J. Am. Chem. Soc. 116*, 443−445.
Wen, Z. Q., Hecht, L., & Barron, L. D. (1994b) *Protein Sci. 3*, 435−439.
Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H., & Dyer, R. B. (1996) *Biochemistry 35*, 691−697.
Wilson, G., Ford, S. J., Cooper, A., Hecht, L., Wen, Z. Q., & Barron, L. D. (1995) *J. Mol. Biol. 254*, 747−760.
Wilson, G., Hecht, L., & Barron, L. D. (1996a) *J. Mol. Biol. 261*, 341−347.
Wilson, G., Hecht, L., & Barron, L. D. (1996b) *J. Chem. Soc., Faraday Trans. 92*, 1503−1510.
Wood, K. A., & Strauss, H. L. (1990) *J. Phys. Chem. 94*, 5677−5684.
Yang, J. J., Buck, M., Pitkeathly, M., Kotik, M., Haynie, D. T., Dobson, C. M., & Radford, S. E. (1995) *J. Mol. Biol. 252*, 483−491.

BI961314V