J|A|C|S
ARTICLES
Published on Web 10/09/2002

Probing the Exposure of Tyrosine and Tryptophan Residues in Partially Folded Proteins and Folding Intermediates by CIDNP Pulse-Labeling

Charles E. Lyon,[†] Eui-Sik Suh,[†] Christopher M. Dobson,[‡] and P. J. Hore*,[†]

Contribution from the Oxford Centre for Molecular Sciences and the Physical and Theoretical Chemistry Laboratory, Oxford University, United Kingdom and the Department of Chemistry, University of Cambridge, Cambridge, United Kingdom

Received January 28, 2002

Abstract: A nuclear magnetic resonance (NMR) technique has been devised to probe the structures of disordered, partially folded states of proteins at the level of individual amino acid residues. Chemically induced dynamic nuclear polarization (CIDNP) is first generated in exposed aromatic side-chains of the denatured state and then transferred to the high-resolution NMR spectrum of the native state by stimulating rapid refolding of the protein. Crucial improvements in sensitivity were achieved by carrying out the polarization-producing photochemistry in a deoxygenated sample of the disordered state of the protein in a magnetic field of 4.0 T and recording the ^{1}H NMR spectrum of the refolded native state at 9.4 T (400 MHz). Application of this method to the low pH molten-globule state of α-lactalbumin reveals remarkably nativelike environments for the aromatic residues in the primary hydrophobic core of the protein. This result provides compelling evidence that the detailed fold of a protein can be established prior to the formation of the cooperative close-packed native structure.

Introduction

Partially folded states of proteins are frequently populated in the process of protein folding and are implicated in many other aspects of cell biology, including translocation across membranes and disease states resulting from protein aggregation.[1,2] Such species are often substantially disordered and exceptionally difficult to characterize by structural methods such as X-ray diffraction and NMR spectroscopy because they resist crystallization and frequently exhibit very poorly resolved NMR spectra.[3,4] One approach to overcoming the low spectral resolution is to transfer information from the non-native state to the highly resolved NMR spectrum of the ordered native state. This tactic has been used with great success in hydrogen exchange experiments, to probe the existence of secondary structure in folding intermediates.[5,6] Of particular value, however, would be an NMR technique capable of providing direct evidence on the environments of the individual side-chains that determine the fold of globular proteins. Here, we show that this can be realized using chemically induced dynamic nuclear polarization (CIDNP), opening the door to detailed structural studies of disordered states of proteins.

The CIDNP technique probes the accessibility of aromatic side-chains in proteins by means of cyclic photochemical reactions.[7−9] Conventionally, a solution of the protein together with a small quantity of a flavin photosensitizer is irradiated inside the NMR probe using a laser. The photoexcited dye reacts with *exposed* amino acid residues to give transient radical pairs whose magnetic interactions lead to nuclear polarization in the intact protein.[7,8] In addition to information on side-chain accessibilities, the attractive features of the method are its selectivity (only histidine, tyrosine and tryptophan are polarized), and sensitivity (the ^{1}H NMR signals are typically 2−5 times stronger than normally expected).[10] Applications of CIDNP include studies of protein interactions that modify the exposure of surface residues, for example binding to cofactors, inhibitors, nucleic acids, micelles, and other proteins.[8,11,12] The method has recently been extended to monitor protein folding in real

* To whom correspondence should be addressed. Physical and Theoretical Chemistry, Laboratory, Oxford University, South Parks Road, Oxford OX1 3QZ, United Kingdom. E-mail: peter.hore@chem.ox.ac.uk. Tel: +44 1865 275415. Fax: +44 1865 275410.

[†] Oxford Centre for Molecular Sciences and the Physical and Theoretical Chemistry Laboratory, Oxford University.

[‡] Department of Chemistry, University of Cambridge.

(1) Ptitsyn, O. B. *Adv. Protein Chem.* **1995**, *47*, 83−229.
(2) Radford, S. E.; Dobson, C. M. *Cell* **1999**, *97*, 291−298.
(3) Baum, J.; Dobson, C. M.; Evans, P.; Hanley, C. *Biochemistry* **1989**, *28*, 7−13.
(4) Schulman, B. A.; Kim, P. S.; Dobson, C. M.; Redfield, C. *Nature Struct. Biol.* **1997**, *4*, 630−634.
(5) Clarke, J.; Itzhaki, L. S. *Curr. Opin. Struct. Biol.* **1998**, *8*, 112−118.
(6) Englander, S. W. *Ann. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 213−238.
(7) Kaptein, R.; Dijkstra, K.; Nicolay, K. *Nature* **1978**, *274*, 293−294.
(8) Hore, P. J.; Broadhurst, R. W. *Prog. Nucl. Magn. Reson. Spectrosc.* **1993**, *25*, 345−402.
(9) Dobson, C. M.; Hore, P. J. *Nature Struct. Biol.* **1998**, *5*, 504−507.
(10) Kaptein, R. In *NMR Spectroscopy in Molecular Biology*; Pullman, B., Ed.; D. Reidel: Dordrecht, 1978; pp 211−229.
(11) Siebert, H. C.; von der Lieth, C. W.; Kaptein, R.; Beintema, J. J.; Dijkstra, K.; van Nuland, N.; Soedjanaatmadja, U. M. S.; Rice, A.; Vliegenthart, J. F. G.; Wright, C. S.; Gabius, H. J. *Proteins: Struct., Funct., Genet.* **1997**, *28*, 268−284.
(12) Siebert, H. C.; Adar, R.; Arango, R.; Burchert, M.; Kaltner, H.; Kayser, G.; Tajkhorshid, E.; von der Lieth, C. W.; Kaptein, R.; Sharon, N.; Vliegenthart, J. F. G.; Gabius, H. J. *Eur. J. Biochem.* **1997**, *249*, 27−38.

10.1021/ja020141w CCC: $22.00

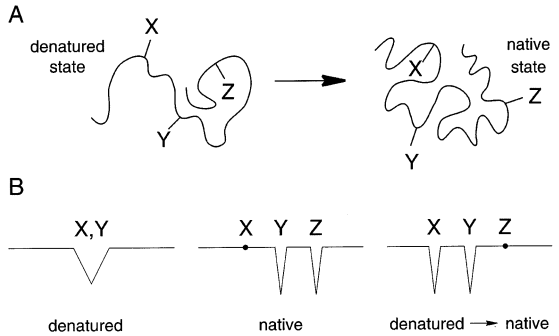

Figure 1. Basis of the CIDNP pulse-labeling method. (A) Schematic structures of denatured and native states of a protein containing three potentially CIDNP-active amino acid side-chains, X, Y, and Z. (B) Schematic CIDNP spectra of the denatured state (left), the native state (center), and of the native state with polarizations appropriate to the denatured form of the protein (right). The poor resolution of the spectrum of the denatured state is indicated by the broad, overlapping resonances from X and Y. Dots on the baselines of spectra indicate the chemical shifts of unpolarized resonances.

time using stopped flow techniques.[13,14] For example, rapid injection of a solution of denatured protein into refolding buffer in the NMR tube can be used to produce a change in solvent conditions that triggers refolding.[15] By recording CIDNP spectra at short intervals thereafter one can monitor changes in the accessibility of aromatic residues as the protein refolds to its native state.[13,14]

The principle of the experiment described in this paper is illustrated schematically in Figure 1. The idea is to polarize the partially folded denatured protein and then to refold it rapidly before the CIDNP relaxes. In this way, enhancements are transferred to the native state spectrum whose chemical shifts allow the exposed residues in the denatured state to be identified. Figure 1A shows schematic structures of denatured and native states of a protein containing three potentially polarizable amino acid side-chains, X, Y, and Z. X and Y are exposed to solvent in the partially folded denatured state, whereas Y and Z are exposed in the native state. The conventional CIDNP spectra (Figure 1B) comprise broad, overlapping lines for the denatured state (left), and relatively narrow, well-resolved resonances for the native state (center). Only one (emissively) enhanced NMR signal is shown for each exposed side-chain. Also shown in Figure 1B (right) is the spectrum immediately after rapid refolding of the polarized denatured state. The observation of resonances from residues X and Y in this spectrum reveals their accessibility in the denatured state. Efficient magnetization transfer requires that the refolding is at least as fast as nuclear spin relaxation.

Although the CIDNP spectra must be recorded in an intense magnetic field, to maximize chemical shift dispersion and hence resolution, strong fields do not necessarily result in the optimum polarization.[16] Experiments with amino acids have established that a field of ~4 T gives the strongest enhancements for all three CIDNP-active residues—histidine, tyrosine, and tryptophan.[17] As we describe below, this field-dependence allows a crucial 500% gain in sensitivity to be achieved by generating polarization at 4.0 T and observing it at 9.4 T.

This "CIDNP pulse-labeling" technique is demonstrated using bovine α-lactalbumin, a protein that exists at low pH in a compact, partially folded molten globule state (the A-state), characterized by extensive secondary structure without rigid side-chain packing.[18−20] The structural heterogeneity and millisecond conformational fluctuations of the A-state result in an NMR spectrum so poorly resolved that very few resonances can be assigned.[4] This form of α-lactalbumin is of particular interest because of its resemblance, in both overall structure and dynamics, to the transient intermediate state that is formed during folding of this protein.[15,21,22] Previous studies have shown that the A-state can be rapidly converted into the native state by a sudden jump in pH.[14,23,24] The four tyrosine and four tryptophan residues in α-lactalbumin are distributed throughout the structure of the native state and therefore report on local structure in different regions of the protein. CIDNP pulse-labeling reveals remarkably nativelike environments for the aromatic residues in the hydrophobic core of the A-state.

Methods

^{1}H CIDNP experiments were performed either at 400 MHz (9.4 T) on a Varian XL-400 spectrometer or at 600 MHz (14.1 T) on a home-built instrument. The light source was a Spectra-Physics Stabilite 2016−05 argon ion laser, operating in multi-line mode at 4 W output power, principally at wavelengths 488 and 514 nm. A mechanical shutter, controlled by the spectrometer, was used to produce light pulses of 100−500 ms duration. Flavin mononucleotide (FMN) was employed throughout as the photosensitizer, at a concentration of 0.2 mM for the experiments that involved high field irradiation (at 9.4 T) and 0.8 mM for those at low field (4.0 T). All CIDNP spectra are the result of subtracting free induction decays measured without irradiation ("dark") from the same number recorded, under identical conditions, *with* irradiation ("light"). All of the NMR experiments were carried out at 25 °C. All of the NMR samples were made up in D_2O. Dissolved oxygen was removed from protein samples by three cycles of a freeze, pump, thaw procedure.[25] The samples were frozen in a Schlenk tube in an ethanol/dry ice mixture at −77 °C, evacuated to 0.2 mbar, and thawed in warm water. Refolding kinetics were recorded with an Applied Photophysics SX17-MV stopped-flow fluorimeter; tryptophan fluorescence, excited at 280 nm, was monitored at 320 nm.

High Field Irradiation. For experiments involving irradiation of the NMR sample in the probe of the 400 MHz spectrometer, light was conducted into a 5 mm NMR tube from the side (i.e., through the coil) by way of mirrors and a cylindrical quartz rod surmounted by a right-angle prism inside the NMR probe.[26] In the experiments at 600 MHz, the light was brought into the sample tube from above by means of a 1 mm diameter optical fiber, the end of which was positioned at the bottom of a Wilmad coaxial Pyrex insert (WGS 5BL), 4 mm above

(13) Hore, P. J.; Winder, S. L.; Roberts, C. H.; Dobson, C. M. *J. Am. Chem. Soc.* **1997**, *119*, 5049−5050.
(14) Maeda, K.; Lyon, C. E.; Lopez, J. J.; Cemazar, M.; Dobson, C. M.; Hore, P. J. *J. Biomol. NMR* **2000**, *16*, 235−244.
(15) Balbach, J.; Forge, V.; van Nuland, N. A. J.; Winder, S. L.; Hore, P. J.; Dobson, C. M. *Nature Struct. Biol.* **1995**, *2*, 865−870.
(16) Closs, G. L. In *Advances in Magnetic Resonance*;Waugh, J. S., Ed.; Academic Press: San Diego, 1974; Vol. 7, pp 157−229.
(17) Lyon, C. E.; Lopez, J. J.; Cho, B.-M.; Hore, P. J. *Molec. Phys.* **2002**, *100*, 1261−1269.
(18) Ptitsyn, O. B. In *Protein Folding*; Creighton, T. E., Ed.; W. H. Freeman: New York, 1992; pp 243−300.
(19) Dobson, C. M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 636−640.
(20) Arai, M.; Kuwajima, K. *Adv. Protein Chem.* **2000**, *53*, 209−282.
(21) Kuwajima, K. *Proteins: Struct., Funct., Genet.* **1989**, *6*, 87−103.
(22) Ptitsyn, O. B.; Pain, R. H.; Semisotnov, G. V.; Zerovnik, E.; Razgulyaev, O. I. *FEBS Lett.* **1990**, *262*, 20−24.
(23) Balbach, J.; Forge, V.; Lau, W. S.; Jones, J. A.; van Nuland, N. A. J.; Dobson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7182−7185.
(24) Wirmer, J.; Kuhn, J.; Schwalbe, H. *Angew. Chem., Int. Ed.* **2001**, *40*, 4248−4251.
(25) Oppenheimer, N. J. In *Methods in Enzymology*; Oppenheimer, N. J., James, T. L., Eds.; Academic Press: San Diego, 1989; Vol. 176, pp 78−89.
(26) Broadhurst, R. W.; Dobson, C. M.; Hore, P. J.; Radford, S. E.; Rees, M. L. *Biochemistry* **1991**, *30*, 405−412.

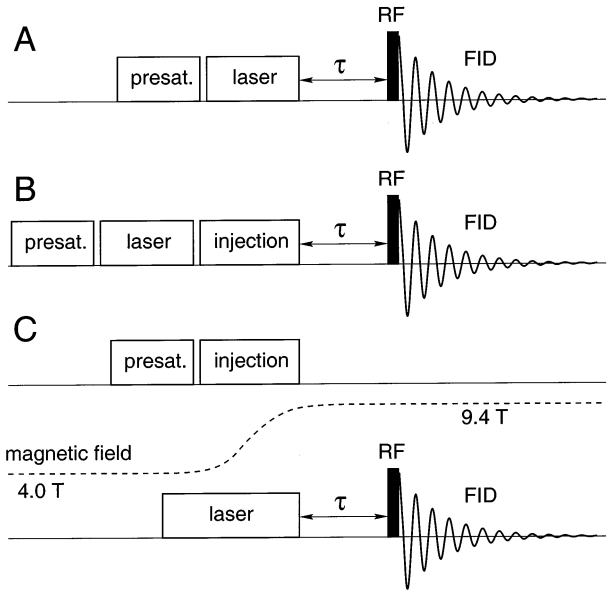

Figure 2. Pulse sequences for ^{1}H CIDNP measurements. (A) High field (9.4 T) laser irradiation, without injection. (B) High field (9.4 T) irradiation, with injection. (C) Low field (4.0 T) irradiation, with injection. In all three cases, the 400 MHz radio frequency (RF) pulse used to excite the NMR free induction decay (FID) had a flip angle of 90°. Prior to each experiment, the proton spins in the NMR tube were presaturated using a train of randomly spaced, randomly phased 90° pulses. In the experiments involving rapid sample mixing, the interval (τ) between injection and detection allowed (when appropriate) refolding of the denatured protein to its native state. The overlap of the injection with the end of the laser pulse in (C) ensured irradiation of the whole of the solution of the denatured state as it passed through the irradiated region at the lower end of the capillary (at 4.0 T) on its way to the probe (9.4 T).

the top of the NMR coil.[14,27] The latter arrangement required no modifications to the NMR probe. The pulse sequence for this experiment is shown in Figure 2A.

Real-time refolding experiments at 400 MHz were performed by injecting (in 200 ms) 300 μL of buffer solution from a pneumatically driven (10 bar) gastight syringe (SGE, model 2.5MDR-GT) placed outside the magnet, via 0.5 mm internal diameter PTFE tubing, into 400 μL of a solution of the denatured protein, together with 0.2 mM FMN, in the NMR tube (see Figure 2B).[13−15]

Low Field Irradiation. The sensitivity of the pulse-labeling experiment was optimized by generating CIDNP at low field (4.0 T) and observing it at high field (9.4 T). As described in detail elsewhere,[17] fields in the range from 0.1 to 7 T are conveniently available 50 to 13 cm above the center of the superconducting solenoid of the 9.4 T (400 MHz) magnet. The beam from the argon ion laser was directed via a 1 mm optical fiber and right angle prism onto a glass capillary (length, 3 cm; internal diameter, 2 mm), connected at both ends to 0.6 mm internal diameter PTFE tubing. This irradiation cell was placed within the magnet bore at a position where the field strength, as measured with a Hall probe, was 4.0 T (17 cm above the center of the magnet). The syringe described above, connected to the upper PTFE tubing, enabled rapid transfer of the irradiated solution into a 5 mm NMR sample tube. The NMR tube was fitted with a Wilmad coaxial insert with its stem cut down so that the end of the lower PTFE tube could be placed 4 mm above the top of the NMR coil, just below the surface of the solution inside the NMR tube. Further details can be found in ref 17.

Experiments were performed using a 500 ms laser flash overlapped by a 200 ms injection pulse (see Figure 2C). Typically, a 300 μL aliquot of a solution containing denatured protein was transferred into 400 μL of refolding buffer in the NMR tube. The NMR signal was detected at a variable time (100−200 ms) after the end of the injection, to allow complete mixing, recovery from the mechanical shock of the transfer, and refolding of the denatured protein.

Theoretical Background

The simple diffusion model of CIDNP,[28−31] for a radical pair with a single proton, predicts that the nuclear polarization in the diamagnetic recombination product is proportional to[17]

$$|\Delta g\mu_B B_0/\hbar + a/2|^{1/2} - |\Delta g\mu_B B_0/\hbar - a/2|^{1/2}$$

where Δg is the difference in the g-values of the two radicals (one derived from the amino acid side-chain, the other from the flavin photosensitizer), B_0 is the strength of the magnetic field in which the photochemistry takes place, a is the proton hyperfine coupling constant, and other symbols have their usual meanings. The square roots arise from the $t^{-3/2}$ time dependence of the radical re-encounter probability.[28]

The polarization predicted by this expression has a sharp maximum at the field where the difference in the electron Zeeman interactions of the two radicals ($\Delta g\mu_B B_0/\hbar$) equals half the hyperfine interaction ($a/2$), i.e., when one of the two terms in the equation vanishes.[32] At higher fields, the polarization diminishes rapidly. For radicals containing several nuclei with significant hyperfine couplings, the theory can easily be extended to predict the field dependence of the CIDNP effect. The result is that the maximum is broadened by the additional magnetic interactions and shifted to a field strength somewhat larger than $\hbar a/2\Delta g\mu_B$. For the tyrosine/flavin radical pair, $\hbar a/2\Delta g\mu_B = 0.6$ T; when all of the significant hyperfine couplings are included, the maximum for the tyrosine H3,5 protons occurs at ~3 T.

Experimental measurements of the CIDNP field dependence for tyrosine agree well with this simple model, using the g-values and hyperfine couplings of the tyrosyl and flavin radicals.[17] Tryptophan and histidine, whose radicals are not well characterized, show similar behavior, with optimum fields of ~4 T.[17]

Results

Preliminary Experiments. Figure 3 compares the 600 MHz ^{1}H NMR and CIDNP spectra of the native state and the A-state of α-lactalbumin: of the three histidine, four tryptophan, and four tyrosine residues in this molecule, only His 68, Trp 118, and Tyr 18 are sufficiently exposed on the surface of the native protein to react efficiently with the photosensitizer.[33−38] For the A-state, the reduced signal intensity from the tryptophan residues relative to the tyrosines, indicates that the hydrophobic tryptophan side-chains are significantly protected from the solvent in the partially folded form and/or the tyrosine residues have

(27) Scheffler, J. E.; Cottrell, C. E.; Berliner, L. J. *J. Magn. Reson*. **1985**, *63*, 199−201.
(28) Noyes, R. M. *J. Chem. Phys.* **1954**, *22*, 1349−1359.
(29) Adrian, F. J. *J. Chem. Phys.* **1970**, *53*, 3374−3375.
(30) Adrian, F. J. *J. Chem. Phys.* **1971**, *54*, 3912−3917.
(31) Kaptein, R. *J. Am. Chem. Soc.* **1972**, *94*, 6251−6262.
(32) Brocklehurst, B. *Molec. Phys*. **1999**, *96*, 283−291.
(33) Berliner, L. J.; Kaptein, R. *J. Biol. Chem*. **1980**, *255*, 3261−3262.
(34) Berliner, L. J.; Kaptein, R. *Biochemistry* **1981**, *20*, 799−807.
(35) Berliner, L. J.; Koga, K.; Nishikawa, H.; Scheffler, J. E. *Biochemistry* **1987**, *26*, 5769−5774.
(36) Improta, S.; Molinari, H.; Pastore, A.; Consonni, R.; Zetta, L. *Eur. J. Biochem.* **1995**, *227*, 78−86.
(37) Improta, S.; Molinari, H.; Pastore, A.; Consonni, R.; Zetta, L. *Eur. J. Biochem*. **1995**, *227*, 87−96.
(38) Alexandrescu, A. T.; Broadhurst, R. W.; Wormald, C.; Chyan, C. L.; Baum, J.; Dobson, C. M. *Eur. J. Biochem*. **1992**, *210*, 699−709.

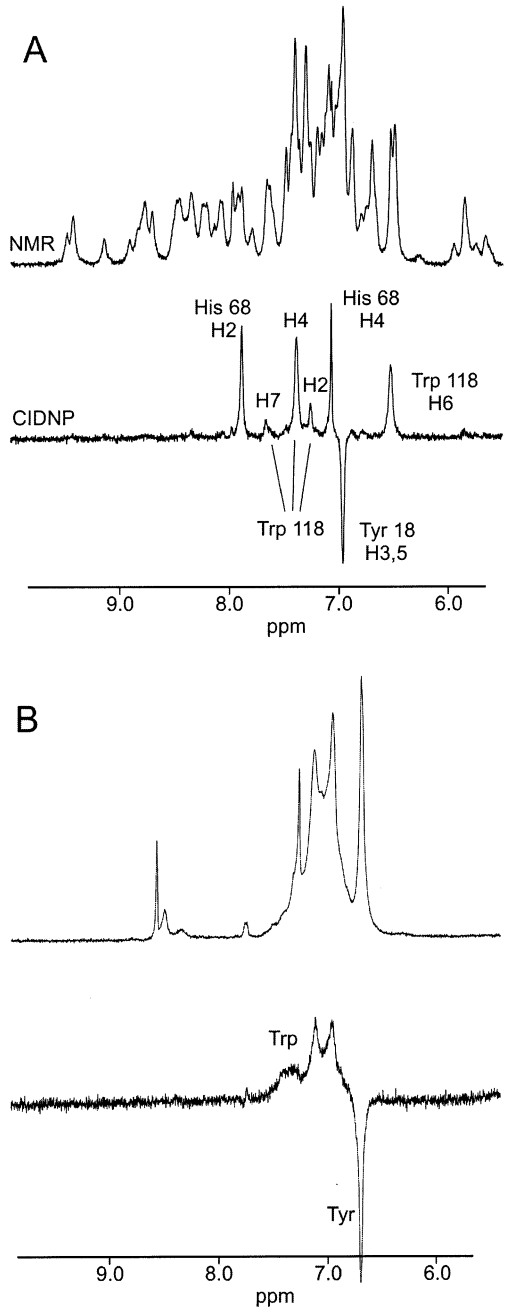

Figure 3. Aromatic regions of the 600 MHz ^{1}H NMR (above) and CIDNP (below) spectra of (A) the native state (1.5 mM, pH 7.0) and (B) the molten globule A-state (1.5 mM, pH 2.0) of bovine α-lactalbumin in D_2O. The NMR spectra are the result of 64 scans. The CIDNP spectra are the difference between 16 "light" scans (100 ms light flashes at 4 W power) and 16 "dark" scans recorded with the pulse sequence of Figure 2A, using a delay τ of 10 ms to allow the photochemically formed radicals to recombine before detection of the NMR signals.

an increased accessibility to the photosensitizer.[14] However, the poor resolution in this spectrum prevents the direct identification of the residues that are exposed to solvent in the A-state.

Figure 4 shows the time dependence of the integrated 400 MHz CIDNP intensities of the molten globule A-state of α-lactalbumin obtained in three different ways, none of which involved refolding the protein. In the first experiment, a solution of the A-state (1.5 mM, pH 2) was polarized at 9.4 T using a 500 ms light pulse. The signal intensities were measured with various delays after the end of the irradiation period (Figure 4,

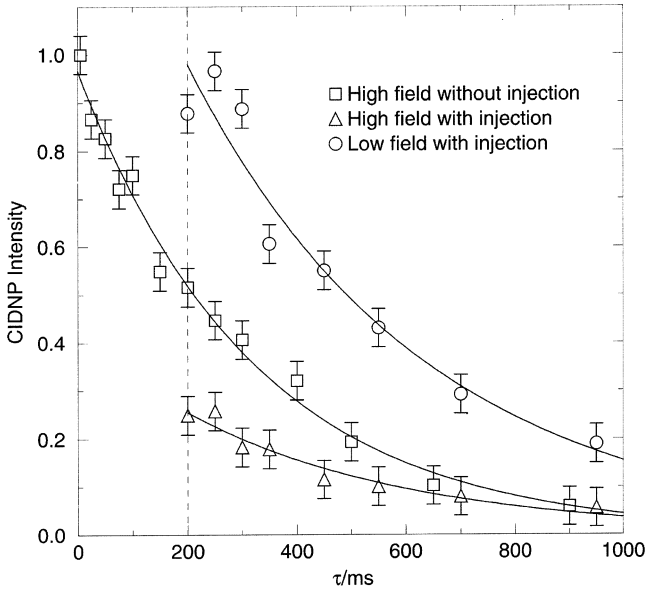

Figure 4. ^{1}H CIDNP intensities (in arbitrary units) of the A-state of bovine α-lactalbumin as a function of time τ after the end of the 500 ms laser light pulse. *Squares*: high field irradiation, without injection; pulse sequence as in Figure 2A. *Triangles*: high field irradiation, with injection; pulse sequence as in Figure 2B. *Circles*: low field irradiation, with injection; pulse sequence as in Figure 2C. The solid lines are exponential fits, to guide the eye. The dashed vertical line marks the end of the injection pulse in the second and third experiments.

squares). By 200 ms, about half of the signal has been lost as a result of spin−lattice relaxation. In the second experiment, a 400 μL sample of the A-state in the NMR tube was polarized exactly as before, except that the laser pulse was immediately followed by the injection of 300 μL of the same solution of the (unpolarized) A-state (1.5 mM, pH 2). CIDNP intensities were then recorded at intervals after the end of the 200 ms injection period (Figure 4, triangles). Comparing the measured intensities at 200 ms with and without the injection, it is clear that the mixing produces a further ∼30% attenuation of the enhancement, over and above the 50% loss by relaxation. This extra reduction in signal strength comes about because the polarized molecules generated within the 300 μL sensitive region of the NMR coil are spread by the injection into a final volume of 700 μL, leading to a ∼2.3-fold "dilution" of the signal. In principle, the use of smaller injectant volumes should improve matters both by shortening the time required for the injection and by reducing the dilution factor; however, no significant improvements could be found experimentally. As expected from these exploratory measurements, attempts to implement the pulse-labeling technique by injecting refolding buffer (pH 7) into a solution of the molten globule (pH 2) gave very little signal immediately after the 200 ms injection (not shown).

In the third experiment, the solution of the A-state (300 μL, 1.5 mM, pH 2) was polarized at 4.0 T in the irradiation cell in the magnet bore (see Methods) and then injected (200 ms) into 400 μL of the same solution in an NMR tube in the probe at 9.4 T. As before, the CIDNP signals were measured at different times after the end of the injection (Figure 4, circles). The improvement in signal strength at 200 ms, immediately after the injection, is dramatic. Compared to the injection experiment with high field irradiation (Figure 4, triangles), the low field approach gives a *five*-fold boost in sensitivity that compensates for the losses that arise from sample dilution and spin relaxation.

In an attempt to increase the signal strength still further, the effect of dissolved O_2 was assessed. Oxygen plays two roles in the photochemical reactions that generate CIDNP: it quenches excited flavin triplets,[39] and it re-oxidizes any photoreduced flavin.[40] The latter reaction is welcome in experiments that require several laser flashes because it prolongs the lifetime of the sample, but is largely irrelevant in the "single-shot" experiments discussed here.[14] However, triplet quenching, which competes with radical pair formation, can result in a serious loss of polarization. For example, deoxygenation of a sample of histidine (0.5 mM, pH 7) by bubbling with N_2 for 5 min produced an ~3-fold increase in the CIDNP. Bubbling with O_2, in contrast, removed the signal completely. All protein samples were therefore subjected to three cycles of freeze−pump−thaw degassing, resulting in typically a *doubling* of the CIDNP intensity. The 10-fold increase in sensitivity obtained by the combination of low field irradiation and deoxygenation proved crucial to the success of the technique in the present experiments.

TFE State of Lysozyme. From a variety of experiments, including real-time NMR[15] and CIDNP,[14] it is known that the A-state of α-lactalbumin can be refolded to its native state within a second by a sudden pH jump in the presence of Ca^{2+}. Before applying CIDNP pulse-labeling to α-lactalbumin, the technique was first tested on a protein that can be refolded much more rapidly. The system chosen was hen lysozyme, a small monomeric protein with a backbone fold similar to that of α-lactalbumin. CIDNP of the native state shows that two of the six tryptophan residues are exposed, in agreement with static accessibility calculations.[41] Under most conditions, lysozyme folds to its native state on roughly the same time scale as α-lactalbumin. In the presence of 2,2,2-trifluoroethanol (TFE), however, the folding process is accelerated by up to one order of magnitude.[42] Lysozyme was therefore denatured in 30% TFE; under these conditions, it is present in a partially folded state with a greater α-helical content than the native protein, but few specific tertiary interactions.[43] Stopped-flow fluorescence measurements, with the same final conditions as used in the CIDNP experiments below, (6.4% TFE, pH 5.2) showed that conversion of the TFE state of lysozyme to the native state is essentially complete within 100 ms, i.e., within the dead-time of the CIDNP injection.

Figure 5A shows the CIDNP spectrum of the TFE state of lysozyme recorded *without refolding* by polarizing 150 μL of 1.5 mM protein in 30% (v/v) TFE, pH 2 at 4.0 T and injecting it into a 550 μL sample of the same solution in the NMR probe at 9.4 T. Figure 5B shows the corresponding spectrum *with refolding*, obtained in exactly the same way, with the exception that the NMR tube prior to injection contained 550 μL of sodium acetate buffer (20 mM, pH 5.2). Spectra were recorded 100 ms after the end of the 200 ms injection. The differences between the two spectra are striking and confirm that successful transfer of polarization to the native state has taken place. The single emissive tyrosine peak in the TFE state has been replaced by three distinct tyrosine resonances; several resolved tryptophan signals are also apparent. The absence of polarization from histidine residues is likely to be a simple consequence of their inability to compete with tyrosine and tryptophan residues for the excited flavin.[44]

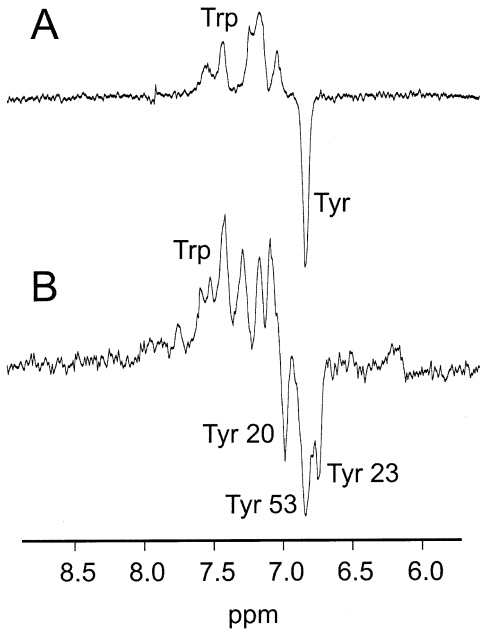

Figure 5. ^{1}H CIDNP spectra of the TFA state of hen lysozyme, with polarization generated at low field (4.0 T) and detected at high field (9.4 T), *without* (A) and *with* (B) refolding to the native state. The CIDNP difference spectra are obtained by subtracting N "light" and N "dark" spectra: $N = 1$ for (A), $N = 4$ for (B).

Table 1. ^{1}H NMR Assignments at Native Chemical Shifts of H3,5 Resonances of Tyrosine Residues Exposed in the TFE State of Hen Lysozyme

residue	Tyr 20	Tyr 23	Tyr 53
this work[a]	7.00	6.76	6.82
literature[b]	6.99	6.71	6.83

[a] Chemical shifts (ppm) from Figure 5B. [b] Chemical shifts from ref 45.

The chemical shifts of native hen lysozyme[45] allow the three emissive peaks to be assigned to the three tyrosines in the protein (Table 1), which evidently have essentially identical chemical shifts and similar exposures to solvent in the TFE state. By contrast, only a very weak signal from one tyrosine residue, Tyr 23 is observed in the spectrum of the native state,[41] reflecting the lack of significant exposure of any of the three tyrosines in this state. The situation is less easily interpreted for the tryptophan residues because of the overlap of the large number of polarized resonances. There can be up to five peaks in the aromatic region from each residue: protons H2, H4, and H6 are directly polarized, whereas H5 and H7 receive polarization by cross relaxation.[41] Although it is impossible to assign these peaks unambiguously, it is evident that there are contributions from at least three and possibly all six tryptophan residues, in comparison to the two residues (Trp 62 and Trp 123) observed in the native state. The greater exposure of aromatic residues in the TFE state of lysozyme is consistent with its lack of globular structure[43] and the ability of this medium to solvate both polar and hydrophobic residues.

(39) Connolly, P. J.; Hoch, J. C. *J. Magn. Reson.* **1991**, *95*, 165−173.
(40) Müller, F. In *Chemistry and Biochemistry of Flavoenzymes*; Müller, F., Ed.; CRC Press: Boca Raton, Florida, 1991; Vol. 1, pp 1−71.
(41) Hore, P. J.; Kaptein, R. *Biochemistry* **1983**, *22*, 1906−1911.
(42) Lu, H.; Buck, M.; Radford, S. E.; Dobson, C. M. *J. Mol. Biol.* **1997**, *265*, 112−117.
(43) Buck, M.; Radford, S. E.; Dobson, C. M. *Biochemistry* **1993**, *32*, 669−678.
(44) Winder, S. L.; Broadhurst, R. W.; Hore, P. J. *Spectrochim. Acta* A **1995**, *51*, 1753−1761.
(45) Dobson, C. M.; Ferguson, S. J.; Poulsen, F. M.; Williams, R. J. P. *Eur. J. Biochem.* **1978**, *92*, 99−107.

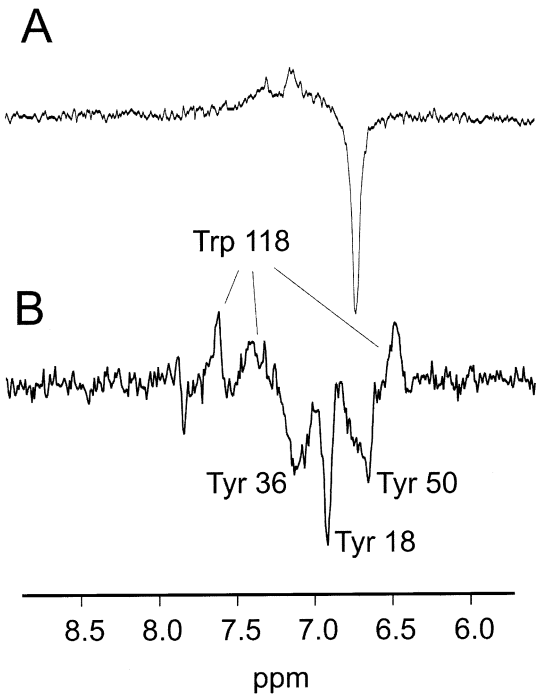

Figure 6. ^{1}H CIDNP spectra of the molten globule A-state of bovine α-lactalbumin, with polarization generated at low field (4.0 T) and detected at high field (9.4 T), *without* (A) and *with* (B) refolding to the native state. The CIDNP difference spectra are obtained by subtracting *N* "light" and *N* "dark" spectra: $N = 1$ for (A), $N = 8$ for (B).

Table 2. ^{1}H NMR Assignments at Native Chemical Shifts of Tyrosine and Tryptophan Residues Exposed in the Molten Globule A-State of Bovine α-Lactalbumin

residue	Tyr 18	Tyr 36	Tyr 50	Trp 118	Trp 118	Trp 118
proton(s)	H3,5	H3,5	H3,5	H4	H6	H7
this work[a]	6.94	7.12	6.68	7.42	6.52	7.54
literature[b]	6.94	7.17	6.72	7.42	6.50	7.63

[a] Chemical shifts (ppm) from Figure 6B. [b] Chemical shifts from ref 38.

A State of α-Lactalbumin. Figure 6 shows CIDNP spectra of α-lactalbumin obtained using the same procedure as for lysozyme, but with a 200 ms delay after the injection to give time for the slower refolding. Under the conditions of these experiments, stopped-flow fluorescence showed that the refolding was ∼90% complete after 200 ms. The A-state spectrum, without refolding (Figure 6A), was recorded by injecting 300 μL of protein solution (1.5 mM, pH 2.0) into 400 μL of sodium cacodylate buffer (200 mM, pH 2.0, with no Ca^{2+}). The pulse-labeled spectrum, with refolding (Figure 6B), was measured in the same way except that the NMR tube contained 400 μL of sodium cacodylate buffer (200 mM, pH 7.0, with 65 mM Ca^{2+}). The minor differences between Figure 6A and the CIDNP spectrum of the A-state in Figure 3B arise from cross relaxation effects during the 200 ms interval between the end of the injection and the radio frequency pulse.

The pulse-labeled spectrum (Figure 6B) is very different from the spectra of the native state and of the A-state. Three emissive tyrosine resonances and three absorptive tryptophan resonances are visible, whose polarization can only have arisen from photoreactions of accessible side-chains in the A-state. Comparison of the chemical shifts of these resonances with the ^{1}H assignments for native α-lactalbumin[38] (Table 2) shows that these peaks arise from Tyr 18, Tyr 36, Tyr 50, Trp 118 and, tentatively, Trp 60. These assignments were made on the basis that significant polarization is expected for protons H2, H4, and H6 of a tryptophan side-chain (absorptive polarization), and H3,5 of a tyrosine (emissive polarization).[10] The only ambiguity is Trp 60 whose resonances fall close to those of Trp 118 and Tyr 18. The shoulder at 6.75 ppm on side of the Tyr-50 peak in Figure 6B is all that remains of the strong tyrosine polarization of the A-state.

Table 3. Exposed Tyrosine and Tryptophan Residues in the Native State and the Molten Globule A-state of α-lactalbumin, as Determined by CIDNP Spectroscopy[a]

	α-domain					interface	β-domain	
	Y18	W26	Y103	W104	W118	Y36	Y50	W60
N-state	√	×	×	×	√	×	×	×
A-state	√	×	×	×	√	√	√	

[a] Ticks and crosses indicate residues identified as accessible or inaccessible, respectively, to FMN.

Discussion

Bovine α-lactalbumin has two subdomains: the α-domain comprises four α-helices and one small region of 3_{10}-helix near the C-terminus; the remainder of the protein (residues 40−81) constitutes the β-domain, containing three short β-strands. The side-chain accessibilities determined from the CIDNP data for the different states of α-lactalbumin are summarized in Table 3. Remarkably, all five probes in the core of the α-domain of the A-state have nativelike accessibilities, including the exposed hydrophobic side-chain of Trp 118 and the buried polar residue Tyr 103. This observation strongly implies that the native fold is substantially preserved in the α-domain and that the side-chains are well organized within the structure, even though they are not yet locked-in to the unique closely packed structure of the native state. It is evident that the interactions between side-chains that determine the fold of the α-domain in the native state of the protein are already established in the molten globule state and that rather little structural rearrangement is required to achieve the well ordered tertiary structure in the native state.[18−20] This finding complements NMR studies of the main chain resonances that reveal that this core retains a highly compact structure even in the presence of high concentrations of denaturants,[4] and that the amide hydrogens in the regions of the polypeptide chain that contribute to the core structure are substantially protected from hydrogen exchange.[46]

In contrast to this finding, two tyrosine residues, one at the heart of the β-domain (Tyr 50), the other close to its interface with the α-domain (Tyr 36), are exposed in the molten globule state but not in the native state. Both of these residues are located in regions of the polypeptide chain that have been found to be among the least resistant in the molten globule state to unfolding in the presence of denaturants[4] and whose amides show no significant protection against hydrogen exchange.[46] These findings suggest that the protein structure in these regions has sufficient flexibility to allow the two tyrosine side-chains access to the flavin probe and hence to become polarized in the molten globule state. This result is particularly significant in the light of the conclusions from CIDNP studies of the collapsed state formed early in the folding of the homologous protein,

(46) Schulman, B. A.; Redfield, C.; Peng, Z.−y.; Dobson, C. M.; Kim, P. S. *J. Mol. Biol.* **1995**, *253*, 651−657.

lysozyme.[26] In this transient state, it appears that the hydrophobic tryptophan residues are largely buried, and the polar tyrosines exposed, consistent with a relatively disorganized hydrophobic collapse. The present results indicate that the aromatic side-chain environments in the molten globule state of α-lactalbumin in the β-domain are similar to those in the initially formed collapsed lysozyme species, but are substantially more nativelike in the core region of the α-domain.

Conclusion

In this paper, we have shown that it is possible to induce selective nuclear polarization in aromatic residues of a partially folded protein and then to overcome the poor NMR spectral resolution characteristic of such states by transferring the polarization to the well-resolved spectrum of the native protein. An important aspect of the experimental procedure is the boost in signal strength achieved by producing nuclear polarization in a magnetic field substantially smaller than that required to record the NMR data. For success, the experiment requires that appropriate conditions can be found in which refolding to the native state is at least as rapid as nuclear spin relaxation. The latter is typically of the order of 100−1000 ms, whereas folding for proteins of the size appropriate for NMR studies is often complete in times that are substantially shorter.[47] The method described here should therefore have applicability to a wide range of systems.

The technique presented here could be developed in several directions. First, a more intense light source would increase the NMR enhancements to the extent that one should be able to detect polarization transfer, through nuclear Overhauser effects, to neighboring groups in the partially folded state. Observable as additional, identifiable peaks in the spectrum of the refolded native state, such transfers would give extra structural insights by revealing residue contacts, as demonstrated using a related *radio frequency* pulse-labeling approach,[23] but with greater residue-selectivity than could be achieved by nonselective saturation of the aromatic proton signals. There would also, undoubtedly, be a sensitivity and resolution benefit from using a higher field spectrometer (≥600 MHz) than the 400 MHz instrument employed here, particularly in conjunction with cryoprobes. More rapid and more efficient injection methods would also boost sensitivity by reducing the signal attenuation caused by relaxation and dilution effects. Improved deoxygenation of the protein samples and the refolding buffer solution, perhaps by enzymatic oxygen-scavenging,[48] would also increase signal strengths. There is also scope for the use of CIDNP photosensitizers other than flavins: particularly attractive is 2,2′-dipyridyl which polarizes tyrosine and tryptophan side-chains efficiently and allows improved time resolution by the use of pulsed UV lasers.[49] Second, the experiment could be performed in H_2O to allow observation of the tryptophan indole proton resonances, which have comparable enhancements to the aromatic protons.[50] With only one NMR line per residue in the NH region of the spectrum, as opposed to five aromatic peaks, this would greatly ease the problem of assigning tryptophan resonances. Third, it should also be feasible to observe the CIDNP of the tryptophan indole *nitrogens* in ^{15}N labeled proteins, with the dual advantage of reduced relaxation losses during the injection and refolding interval (^{15}N has slower spin−lattice relaxation) and, if preliminary experiments on ^{15}N labeled lysozyme are a guide,[50] even stronger enhancements than for the indole and aromatic protons. Such measurements might even be used to determine heteronuclear relaxation rates in molten globules, potentially providing direct information on side-chain mobility. These extensions of the CIDNP pulse-labeling technique, which achieves the key step of selectively exciting particular residues in a disordered protein, therefore have the potential to describe the nature of at least the most highly structured regions of partially folded states in atomic detail.

Acknowledgment. We are pleased to acknowledge support from the EPSRC, BBSRC, and MRC through the Oxford Centre for Molecular Sciences. Further funding was provided by the EU (RTD project HPRI-1999-CT-50006), and by INTAS (project 01-2126). The research of C.M.D. is supported in part by a Program Grant from the Wellcome Trust.

JA020141W

(47) Pain, R. H., Ed. *Mechanisms of Protein Folding*; Oxford University Press: Oxford, 2000.

(48) Englander, S. W.; Calhoun, D. B.; Englander, J. J. *Anal. Biochem.* **1987**, *161*, 300−306.

(49) Morozova, O. B.; Yurkovskaya, A. V.; Tsentalovich, Yu. P.; Forbes, M. D. E.; Hore, P. J.; Sagdeev, R. Z. *Molec. Phys.* **2002**, *100*, 1187−1195.

(50) Lyon, C. E.; Jones, J. A.; Redfield, C.; Dobson, C. M.; Hore, P. J. *J. Am. Chem. Soc.* **1999**, *121*, 6505−6506.