Protein Science (1997), 6:1101–1109. Cambridge University Press. Printed in the USA.

SYMPOSIUM REPORT

Hydrogen exchange: The modern legacy of Linderstrøm-Lang

S.W. ENGLANDER, L. MAYNE, Y. BAI,[1] AND T.R. SOSNICK[2]

Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania, 19104-6059

(RECEIVED October 22, 1996; ACCEPTED January 27, 1997)

Abstract

This discussion, prepared for the Protein Society's symposium honoring the 100th anniversary of Kaj Linderstrøm-Lang, shows how hydrogen exchange approaches initially conceived and implemented by Lang and his colleagues some 50 years ago are contributing to current progress in structural biology. Examples are chosen from the active protein folding field. Hydrogen exchange methods now make it possible to define the structure of protein folding intermediates in various contexts: as tenuous molten globule forms at equilibrium under destabilizing conditions, in kinetic intermediates that exist for less than one second, and as infinitesimally populated excited state forms under native conditions. More generally, similar methods now find broad application in studies of protein structure, energetics, and interactions. This article considers the rise of these capabilities from their inception at the Carlsberg Labs to their contemporary role as a significant tool of modern structural biology.

Keywords: cooperativity; folding intermediates; hydrogen exchange; Linderstrøm-Lang; molten globule; protein folding

The hydrogen exchange approach was conceived by Kaj Linderstrøm-Lang and implemented by him and his collaborators at the Carlsberg laboratories in the early 1950s. Pauling had just discovered the α-helix and β-sheet and postulated that they were stabilized by hydrogen bonds. In those exciting days at the dawn of modern protein science, Lang realized that peptide group NH hydrogens participate in continual exchange with the hydrogens of solvent, just as in the already understood polar side chains. He set out to test Pauling's ideas by measuring the exchange behavior of these hydrogens. Lang created entirely novel methods to measure H-D exchange, and together with his colleagues at the Carlsberg Labs, he studied hydrogen exchange in a number of proteins and peptides under various solution conditions (Hvidt & Linderstrøm-Lang, 1954, 1955a, 1955b; Krause & Linderstrøm-Lang, 1955; Linderstrøm-Lang, 1955a, 1955b; Berger & Linderstrøm-Lang, 1957; Benson & Linderstrøm-Lang, 1959; Hvidt et al., 1960). In comparison with modern capabilities the information available to Lang was severely limited in resolution and even in accuracy. Remarkably, he saw past these limitations and quickly moved past his initial goal of merely testing for hydrogen bonding. He inferred that exchange is controlled by the dynamic behavior of protein molecules, as we believe today, and he saw the relationships that connect the hydrogen exchange rate with protein dynamics and energetics (Linderstrøm-Lang, 1958; Hvidt & Nielsen, 1966). The equations he formulated continue to provide the theoretical basis for protein and nucleic acid hydrogen exchange that all workers still use today (Hvidt et al., 1960; Hvidt & Nielsen, 1966; Barksdale & Rosenberg, 1982; Englander & Kallenbach, 1984; Woodward, 1994; Gueron & Leroy, 1995; Scholtz & Robertson, 1995). In this leap of imagination and intuition, Lang and John Schellman (Linderstrøm-Lang & Schellman, 1959) foreshadowed the whole field of protein dynamics, years before the first protein structures were solved. This, together with his experimental ingenuity, opened the door to hydrogen exchange studies.

The recent literature documents the expanding use of hydrogen exchange methods to measure the fundamental parameters and behaviors of protein molecules. Hydrogen exchange (HX) methods are being used to determine the structure of protein folding intermediates with lifetimes less than one second (Englander & Mayne, 1992; Baldwin, 1993; Woodward, 1994) and to characterize partially unfolded forms that exist at infinitesimal levels under fully native conditions (Bai et al., 1995; Chamberlain et al., 1996). HX is being used to define protein structures under extreme conditions as in loosely formed equilibrium molten globules at acid pH (Hughson et al., 1990; Kuroda et al., 1992; Chyan et al., 1993; Ptitsyn, 1995b), in organic solvents (Wu & Gorenstein, 1993), in

Reprint requests to: S.W. Englander, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania, 19104-6059; e-mail: walter@hx2.med.upenn.edu.

[1]Present address: Department of Molecular Biology, Scripps Research Institute, LaJolla, California 92037-9701

[2]Present address: Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637

the crystalline state (Pedersen et al., 1991; Gallagher et al., 1992) and even in the dry state (Desai et al., 1994). HX is being used to measure structure changes in complexes of proteins with massive chaperonins (Okazaki et al., 1994; Robinson et al., 1994; Zahn et al., 1994), with antibodies (Paterson et al., 1990; Mayne et al., 1992; Orban et al., 1994) and enzymes (Werner & Wemmer, 1992; Jones et al., 1993), in micelles (Thornton & Gorenstein, 1994), and even in whole active muscle (Rodgers et al., 1996).

This article briefly reviews the development of these capabilities, their physical bases, and illustrates their utility for problems in contemporary structural biology.

Hydrogen exchange methodology

The peptide group NH hydrogens and also the polar side chain NHs and OHs that hydrogen exchange methods measure are uniformly distributed throughout all protein molecules. Access to their exchange behavior can, in principle, provide detailed information at an amino acid-resolved level on protein structure, structure change, dynamics, and energetics. This information is encoded in the language of hydrogen exchange rates, within the immense dynamic range of rates over which the typical protein exchanges its various hydrogens. We need only record the stream of hydrogen exchange information that proteins continually and non-perturbingly emit and interpret these signals in terms of protein parameters.

For many years, however, this level of resolution was beyond reach and it was necessary to rely on methodologies that read out HX rate behavior in a summed, structurally unresolved way (Hvidt & Linderstrøm-Lang, 1955a; Linderstrøm-Lang, 1958; Englander & Englander, 1978). Attempts to increase the structural resolution available included the combination of tritium exchange labeling with protein fragmentation and HPLC separation methods (Rosa & Richards, 1979; Englander et al., 1985) and the use of H-D exchange together with crystallization and neutron diffraction analysis (Kossiakoff, 1982; Mason et al., 1983). The development of multidimensional nuclear magnetic resonance (NMR) methods (Wuthrich, 1986; Ernst et al., 1988; Bax, 1994) dramatically solved the problem of structural resolution (Wagner & Wuthrich, 1982) and made accessible the full power of site-resolved hydrogen exchange. The analytical capability of NMR also made it possible to exploit the inherently powerful capability of hydrogen exchange labeling (Rogero et al., 1986). In the labeling mode one can perform direct hydrogen isotope labeling experiments under conditions most suitable for the question being asked, for example under any of the extreme conditions noted above. The structure-sensitive hydrogen isotope labeling pattern imposed during the experiment can then be trapped and read out under conditions most favorable for the NMR analysis.

Methodological developments are continuing. The great potential of mass spectrometry is being adapted to hydrogen exchange analysis (Anderegg & Wagner, 1995; Miranker et al., 1996; Zhang et al., 1996). A range of useful techniques employing various forms of spectroscopy or physical separation together with H-H, H-D, and H-T exchange are now available (listed in Englander et al., 1996). Also important progress is being made in the interpretation of HX data, as suggested below.

The physical bases of protein hydrogen exchange

Linderstrøm-Lang and his colleagues considered various possible mechanisms that might underlie protein hydrogen exchange, including alternative penetration models (Bryan, 1970; Lumry & Rosenberg, 1975; Richards, 1979; Tsuboi & Nakanishi, 1979; Woodward et al., 1982) and unfolding models (Wagner & Wuthrich, 1979; Kossiakoff, 1982; Englander & Kallenbach, 1984) that have dominated thinking in this area in subsequent decades. He decided in favor of an unfolding mode (Hvidt et al., 1960; Hvidt & Nielsen, 1966), as pictured in Figure 1. In this view, hydrogens involved in hydrogen bonds can exchange only during the small fraction of time when the H-bond donor and acceptor are transiently separated in a dynamic opening-closing reaction. There is now strong support for this view. H-exchange occurs through a distinct chemical reaction in which the peptide group is attacked by either a OH^- or H_3O^+ ion. A hydrogen-bonded proton is sterically inaccessible to such a mechanism. For most purposes, the equations shown in Figure 1 connect the measured exchange rate (k_{ex}) with the free peptide chemical rate (k_{ch}) and the equilibrium constant for the dominant opening, which may be any one of the kinds of openings pictured, since $K_{ap} = K_{local} + K_{subglobal} + K_{global}$ when the structure is stable (i.e., $K_{ap} \ll 1$). Lang favored the view that H-bonds might break individually, as in the local pathway in Figure 1, but he also later considered larger, more cooperative opening modes, following the work of John Schellman who had at that time launched his early investigations into protein cooperativity at the Carlsberg Labs.

How is one to decide which of the opening modes pictured in Figure 1 is at work in any given case? In a stable protein, these

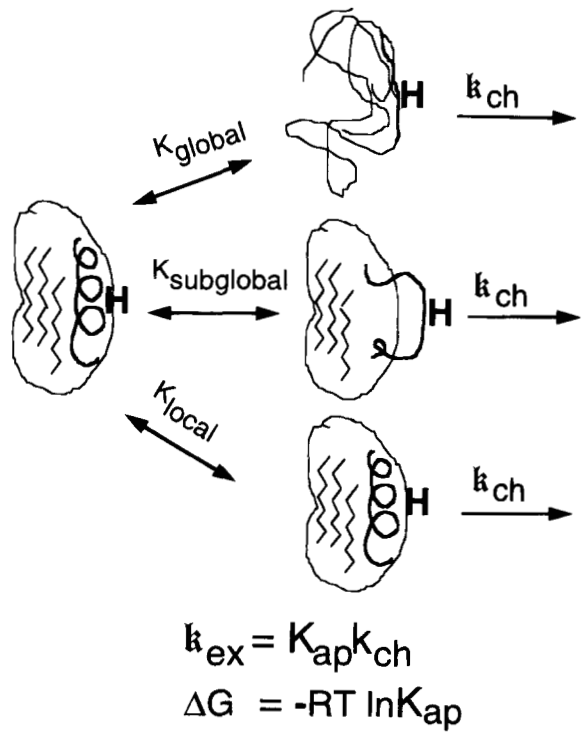

Fig. 1. The unfolding model for protein hydrogen exchange. H-bonded hydrogens in a native protein (left) can exchange only when the H-bond donor and acceptor groups are separated in some transient, high energy unfolding reaction, which may represent a local fluctuation (bottom), a whole molecule unfolding (top), or some intermediate-sized opening (middle). Exchange is governed by the equations shown in the commonly occurring EX2 case, where the structure is stable and reclosing is fast relative to the chemical exchange rate (k_{ch}).

dynamic modes occur at immeasurably low levels. A useful approach to this problem is suggested by the classical melting experiment diagrammed in Figure 2 (left). To evalute the stability of a protein against unfolding, one commonly finds it necessary to drive the protein through its global unfolding transition at high temperature or denaturant concentration, as in Figure 2. Through the transition zone the unfolding equilibrium constant (K_{unf}) can be measured and from this the free energy for unfolding ($\Delta G = -RT \ln K_{unf}$) and the dependence of ΔG on denaturant concentration can be obtained. A graph of these data extrapolated to zero denaturant provides the free energy of stabilization (Pace, 1986). The slope of the extrapolated curve, the so-called m value given by $-(\partial \Delta G/\partial C)_{T,P}$, relates to the denaturant-sensitive surface exposed in the unfolding reaction (Tanford, 1970; Schellman, 1987).

This approach can be extended to the unfolding reactions that determine protein hydrogen exchange (Kim & Woodward, 1993; Mayo & Baldwin, 1993; Bai et al., 1994a; Bai et al., 1994b; Qian et al., 1994). The data shown in Figure 2(right) for equine cytochrome *c* (cyt *c*) include results from classical melting experiments at high GdmCl. At lower GdmCl, Figure 2 also shows results for some of the slowest exchanging hydrogens in cyt *c*, measured by NMR and processed through the equations in Figure 1 to yield ΔG values (Bai et al., 1994b). The HX data produce a curve that smoothly joins the classical melting data. These slowest hydrogens are so well protected in native cyt *c* that they can exchange only during the small fraction of time when the protein transiently visits its globally unfolded state. Thus hydrogen exchange can measure the global unfolding reaction at conditions far below the melting transition, in the region inaccessible to classical methods. A similar conclusion, obtained from HX experiments that differ somewhat in detail, has been obtained for a number of other proteins (Loh et al., 1993; Orban et al., 1994; Perrett et al., 1995; Swint-Kruse & Robertson, 1996).

It is interesting that the curve derived from HX data in the region invisible to classical melting experiments diverges somewhat from the linear extrapolation usually used in melting studies (Pace, 1986). Possible reasons include the presence of residual protecting structure in the unfolded protein at low denaturant or effects imposed by site binding of denaturant (Tanford, 1970) or solvent displacement (Schellman, 1990). The results do make it clear that HX provides a means for evaluating protein stability and changes therein.

This capability is promising—global unfolding can be measured under native conditions, using simple 1D NMR, perhaps even with larger proteins, without having to risk the uncertainties and difficulties often encountered at extreme denaturing conditions. Even more interesting is the fact that the method can be extended to study the free energy, surface exposure, and identity of smaller subglobal protein unfolding reactions.

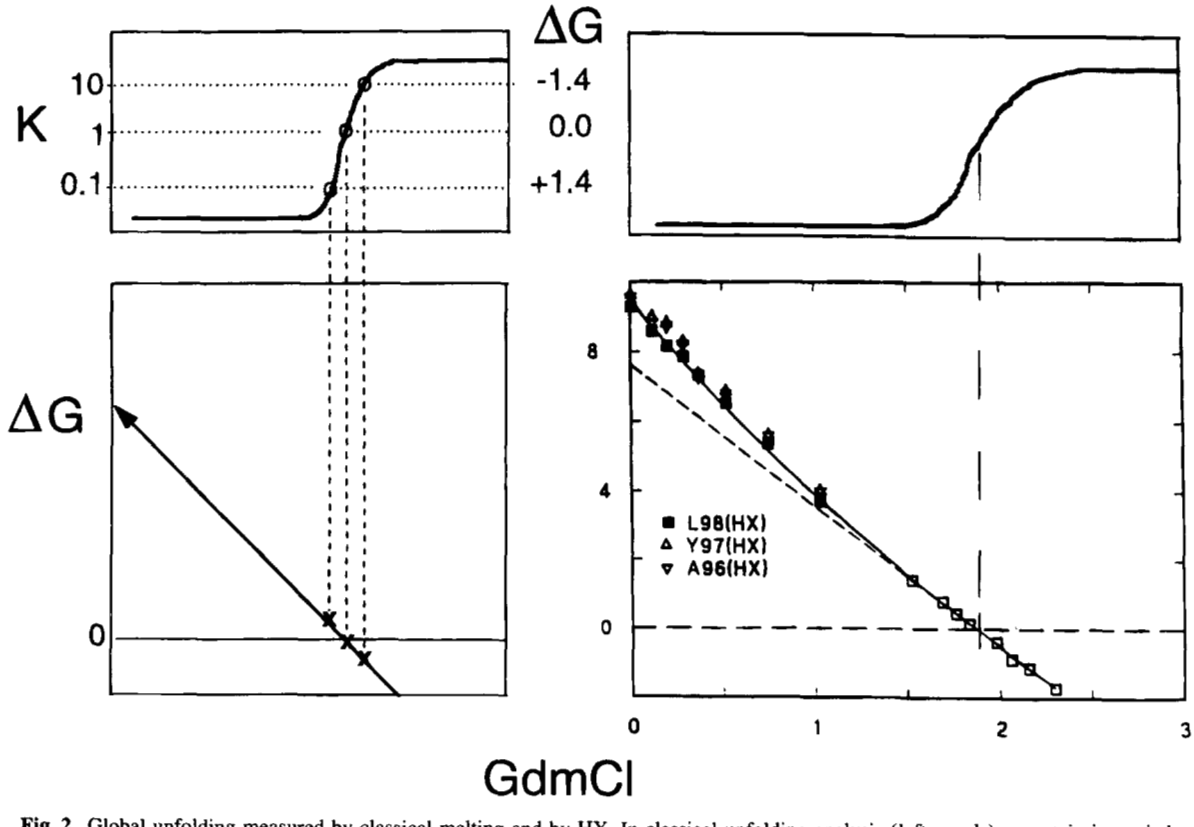

Fig. 2. Global unfolding measured by classical melting and by HX. In classical unfolding analysis (left panels), a protein is carried through its melting transition where the unfolding equilibrium constant and the ΔG for unfolding can be measured. The slope of a plot of ΔG against denaturant concentration yields a measure of surface exposure, the m value, and the curve can be extrapolated to estimate the stabilization free energy at zero denaturant. Data for cyt *c* (right) include classical melting results (open squares) and results obtained from HX measurements of the slowest NHs, which exchange by way of transient global unfolding even under native conditions (data at pH 7, 50 °C).

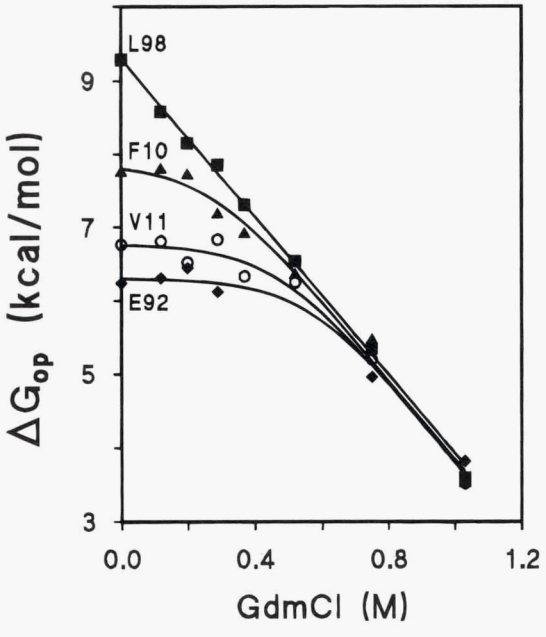

Fig. 3. HX results for some hydrogens in cyt *c*. Leu 98 serves as a marker for the global unfolding isotherm. At low denaturant some other hydrogens exchange through local fluctuations with small m value. Their exchange becomes dominated by the global unfolding when it is promoted by increasing denaturant concentration.

Local and subglobal unfolding units

Most hydrogens exchange by way of openings that are much smaller and more probable than global unfolding. Figure 3 shows HX results, computed as above for a few of these faster exchanging NHs. The ΔG values obtained are essentially independent of GdmCl concentration, showing that these hydrogens exchange through small molecular fluctuations that have very small m values (surface exposure). The global unfolding, indicated in Figure 3 by the Leu 98 NH, makes a negligible contribution to their exchange at low GdmCl concentration. However when GdmCl is increased, the global unfolding reaction, with large surface exposure and correspondingly large m, is sharply enhanced and overtakes the small fluctuations. The curves for the individual NHs merge into the global unfolding isotherm as the global unfolding pathway rises to dominate their exchange. This behavior is exhibited by all the H-bonded residues in the amino and carboxyl-terminal helices of cyt *c* (Fig. 4).

One can expect that all the exchanging NHs eventually join the global isotherm in this way. In fact they do. But before that occurs, something quite surprising is seen, as summarized in Figure 5. Figure 5 shows Leu 98 as a marker for the global isotherm and one of the hydrogens, Phe 10, that join it. Another marker proton, the Leu 68 NH, traces out a curve with m and ΔG values that are large, but not as large as for global unfolding. This hydrogen must exchange through a sizeable but still subglobal unfolding reaction. The HX data measure the free energy of this unfolding and its physical size in terms of the m parameter. The identity of the

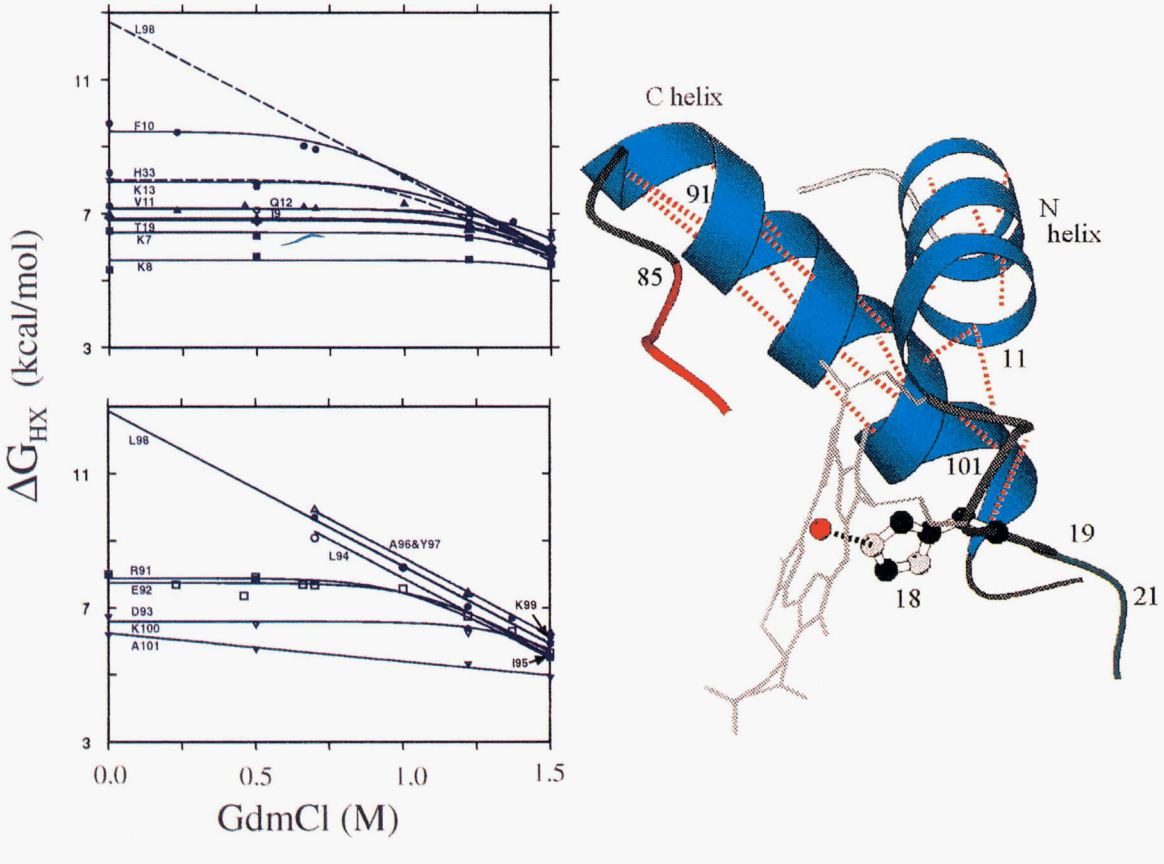

Fig. 4. Hydrogens that join the highest energy global unfolding isotherm represent the amino and carboxyl-terminal helices of cyt *c*.

unfolding segment is revealed by the set of amino acids that join the unfolding isotherm; Tyr 67 is shown as an example. Two additional and successively smaller but still sizeable unfoldings are indicated by the lower lying isotherms in Figure 5.

The segments represented in these HX isotherms account for the entire cyt *c* molecule (Fig. 5). The highest energy unfolding isotherm (blue) reflects the unfolding of the amino and carboxyl helices of cyt *c*. The product of this reaction is the globally unfolded state (Fig. 2). The next lower lying isotherm (green) is joined by all the residues in the 60s helix and several others that appear to identify an entire omega loop on the right side of the protein. The yellow isotherm is joined by five residues that span the large omega loop at the bottom of the protein (when viewed in the standard orientation of Fig. 5). Finally three residues point to the red omega loop on the left side of the molecule.

This behavior is not due to some artifact of denaturant action; similar though less complete results were obtained in cyt *c* experiments that used temperature rather than GdmCl as a perturbant (Bai et al., 1995). Further, the same approach applied to a different protein, bacterial ribonuclease H, reveals fully analogous results (Raschke & Marqusee, 1997).

Equilibrium substructure and kinetic folding

These results indicate that native protein molecules are composed of cooperative units that are smaller than the whole protein. The cooperative units in cyt *c* involve entire secondary structural segments, individually or in pairs. It seems likely that the ability to take protein molecules apart into their constituent cooperative units in this way will have significance for such issues as protein cooperativity and globular protein design, and perhaps also for the biological evolution of protein molecules since one supposes that contemporary protein molecules have evolved from smaller units.

Another implication has to do with protein unfolding and refolding behavior. Protein molecules are required by thermodynamic principles to occupy all possible higher energy states according to their Boltzmann factors and, over time, to cycle through all of these states. That is, protein molecules in solution must unfold and refold all the time, even under fully native conditions. This behavior is invisible to almost any measurement one can think of since most techniques are swamped by signals from the predominant native state. Hydrogen exchange is an exception because the native state contributes nothing to the HX rates that one measures. Measurable hydrogen exchange behavior is produced by the cycling of protein molecules through their normally invisible higher energy states. Available evidence now suggests that some of these states, the sizeable subglobally unfolded forms identified by the native state HX experiment, are the intermediate structures that define the kinetic folding pathway of cyt *c*.

In considering this possibility, one wants to determine the identity of the putative intermediate structures. The interpretation of the HX data in these terms is not wholly straightforward. The problem is formalized in Figure 6, which places the higher energy, partially unfolded states found for cyt *c* on a free energy scale. In native cyt *c*, all the cooperative units (blue, green, yellow, red) are folded (N = BGYR). The up-pointing arrows in Figure 6 represent possible reactions by which the particular unit indicated can unfold. For example, I_1 is defined by the unfolding-dependent exchange of the hydrogens in the red segment, I_2 by the exchange of the yellow set of NHs, and so on. The down arrows represent the reverse, refolding reactions. It should be appreciated that each unfolding reaction must be matched by an equivalent refolding reaction, according to the principle of microscopic reversibility, since these HX measurements were done with cyt *c* at equilibrium, under fully native conditions (pH 7, 30 °C, low GdmCl).

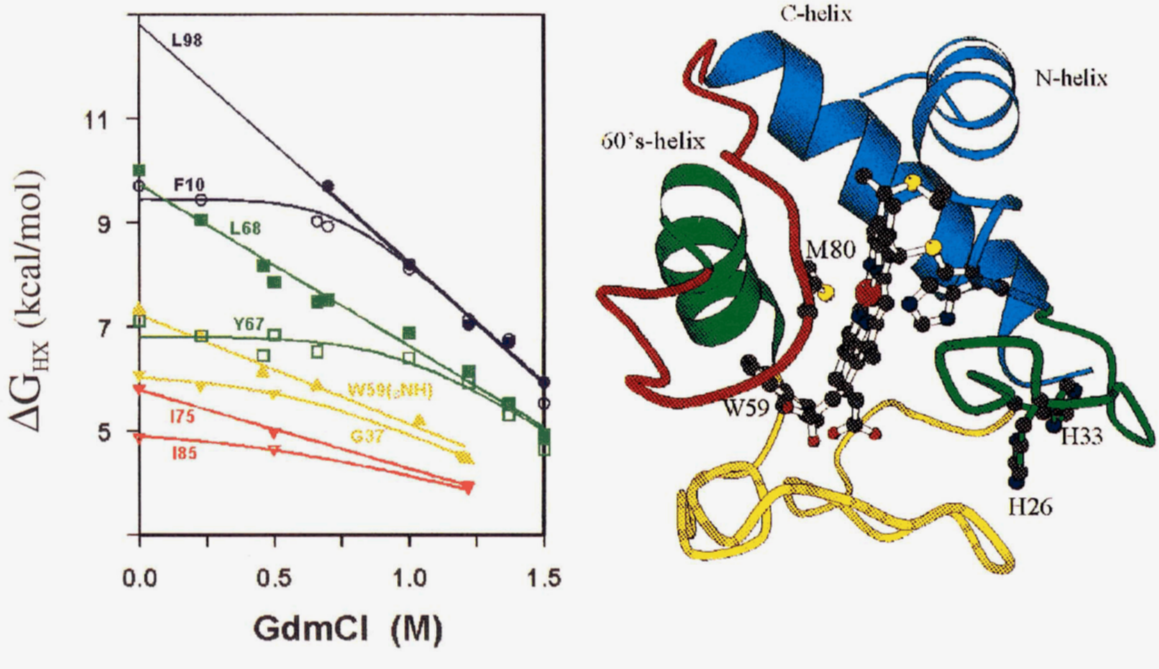

Fig. 5. The four unfolding isotherms portrayed by HX data for cyt *c*, including the global unfolding isotherm (blue, see Fig. 4) and three lower lying subglobal unfoldings (data obtained at pH 7, 30 °C). In each case, one isothermal marker proton and one NH that joins it are shown (from Bai et al., 1995). The groupings are color-coded to indicate their relative free energy levels (red to blue). The protein segments defined by the four HX isotherms account for the whole cyt *c* protein, as shown by the color coding.

Figure 6 pictures alternative views of the unfolding and refolding reactions that may underlie the HX data. The cooperative units may be considered to unfold independently while all the other units remain closed, as suggested in the right set of reaction arrows in Figure 6. For example, the I_3 state would represent green alone open. An alternative possibility is that the cooperative units unfold sequentially, as in the left column of up arrows, so that first the red unit is unfolded to yield I_1 (=YGB), then the yellow unit is unfolded to yield I_2 (=GB), then the green to yield I_3 (=B), and finally the blue is lost to yield the fully unfolded state. If so, then the matching down arrows must define the major refolding pathway of cyt *c*, as suggested at the bottom of Figure 6. The measured exchange does not directly distinguish between these alternative interpretations. For example, in the high energy state I_3 that allows the green NHs to exchange, it is obvious that the blue segments remain closed because the blue NHs do not yet exchange. However one cannot tell whether the red and yellow units are open or closed in the I_3 intermediate. The HX data identify the protein segment that newly opens to produce each intermediate state, but the data do not reveal the condition of the segments that have already exchanged in lower energy openings.

For both extreme views in Figure 6, the lowest level unfolding, opening of the red unit, is the same. Both models produce the same 6 kcal state (I_1) with only the red unit unfolded. Thus the identity of I_1 is not ambiguous. Unfolding of the next higher cooperative unit (−yellow) produces a two-fold ambiguity. The HX data might reflect an intermediate with only the yellow unit open (independent model; I_2 = BGR) or with both red and yellow open (sequential model; I_2 = BG). The exchange of the hydrogens in the green unit presents a four-fold ambiguity—opening of green alone (independent model), or green together with yellow, or with red, or with both yellow and red (sequential model). The opening of the blue unit has the potential to be eight-fold ambiguous. Fortunately, the identity of this last opening is well-defined by direct measurement; unfolding of the two terminal helices (blue) produces the globally unfolded state (Fig. 2B). Thus we have uncertainty only in the identity of two intermediate states, the 7.4 kcal state represented by $I_2(-Y)$ in Fig. 6 and the 10 kcal state, $I_3(-G)$.

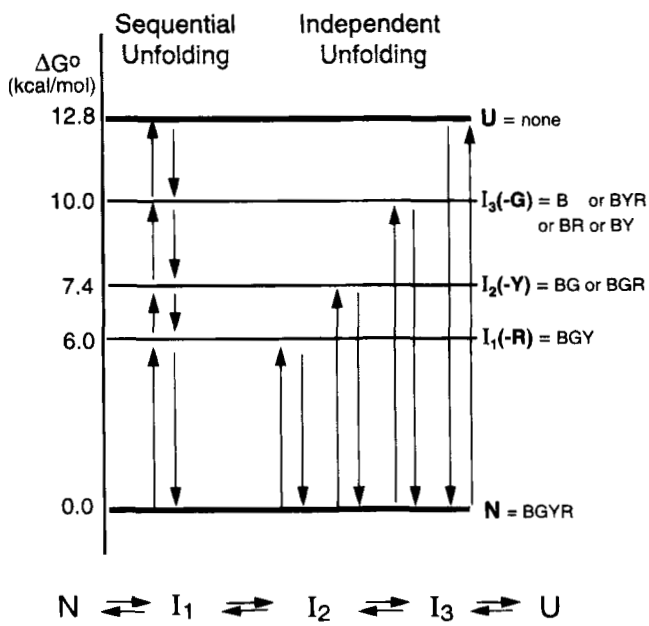

Fig. 6. An energy level diagram indicating the unfolding reactions and the intermediates that might underlie the HX results. A sequential unfolding model (left column of arrows) would correspond to the N to U sequence shown at the bottom, and must then also trace the refolding sequence. In the independent unfolding model, single cooperative segments unfold separately while the others remain closed. The possible identities of each intermediate state detected in the HX experiment are shown at the right in terms of the cooperative unit measured to be newly opened in forming each unfolding intermediate (boldface) and the cooperative units that may still remain folded in each case. B, G, Y, and R represent the blue, green, yellow and red segments, respectively.

The HX pulse labeling experiment

A hydrogen exchange pulse labeling experiment performed in a kinetic folding mode helps to resolve some of the ambiguity by independently identifying one of these forms. This experiment was originally designed to determine the structure of folding intermediates that are only transiently populated, on a sub-second time scale, in kinetic folding experiments (Kim & Baldwin, 1982; Roder et al., 1988; Udgaonkar & Baldwin, 1988; Englander & Mayne, 1992; Baldwin, 1993; Woodward, 1994). The protein is initially unfolded in high GdmCl concentration in D_2O so that all the NHs exchange to ND. At time zero of a kinetic folding experiment, the denaturant solution is diluted into H_2O in a rapid mix apparatus. The protein starts to refold. Conditions in this phase are set to avoid D to H exchange, e.g., pH 6 and 10 °C where the exchange lifetime of even freely exposed NHs is several seconds. After some time of folding, D to H exchange is initiated by pulsing the refolding solution to a higher pH in a second rapid mix, e.g., to pH 9.5 and 10 °C where freely exposed NHs exchange in ~1 msec. Residues in still unfolded segments quickly become H-labeled; those already involved in H-bonded structure are protected from exchange and remain as ND. The solution is neutralized in a third mixing. The protein folds to its native state, trapping the H-D exchange pattern imposed during the pulse. A 2D NMR spectrum of the refolded native protein identifies the residues that were already protected (D-labeled) and those still unprotected (H-labeled) at the time of the labeling pulse. A series of such experiments with the labeling pulse imposed after different folding times can in favorable cases trace out the time course for structure formation at many H-bonded sites in the protein and thus identify the structures of some partially folded kinetic intermediates.

In cyt *c* the pulse labeling experiment identifies an early folding intermediate in which the amino and carboxyl helices are formed before any other H-bonded structure (Roder et al., 1988). This is the same as the initial refolding intermediate (blue) visualized in the sequential unfolding model (Fig. 6). This result resolves the four-fold ambiguity of the $I_3(-G)$ state in favor of the form suggested by the sequential model. It is also pertinent that the amino and carboxyl terminal segments, when excised from the protein and placed into solution together, show some independent folding stability (Kuroda, 1993; Wu et al., 1993). This independent stability is consistent with the possible role of the N- and C-helices as an initial folding intermediate, as suggested by the sequential model.

An analogous result has been found for bacterial ribonuclease H. The highest energy partially unfolded form identified by the native state HX experiment in ribonuclease H, with helices A , B, and D more or less formed and the rest of the protein not yet folded (Chamberlain et al., 1996), was found to be similar to the first intermediate populated in kinetic refolding experiments studied by

HX pulse labeling (Raschke & Marqusee, 1997). Further, the Marqusee group has found yet another confirmatory coincidence. The same most stable form seen in the native state and pulse labeling experiments just described appears to match the low pH molten globule intermediate defined by direct hydrogen exchange labeling (Dabora et al., 1996).

Molten globules, crossover curves, and protein cooperativity

As a final example, we consider the ability of hydrogen exchange to define the structure of proteins in their equilibrium molten globule form. A number of proteins, when placed in mildly destabilizing conditions, usually in acid pH, have been found to assume a somewhat expanded but still partially structured form known as a molten globule (Kuwajima, 1989; Fink, 1995; Ptitsyn, 1995a). The degree of secondary structure present, indicated by circular dichroism measurements, varies greatly among the large number of known protein molten globules but all share the loss of tertiary side chain packing. Molten globules in general are loose, dynamically disordered forms that cannot be crystallized for diffraction analysis. This is unfortunate since the suggestion is strong that some if not all of these represent equilibrium analogs of kinetic folding intermediates (Ptitsyn et al., 1990; Matthews, 1993; Ptitsyn, 1994, 1995b).

The hydrogen exchange labeling capability makes it possible to specify the secondary structural elements present in a protein under molten globule conditions by identifying many of the NHs that are protected from exchange. The protein is placed into D_2O under molten globule conditions and allowed to exchange for increasing periods of time. Samples are returned to native conditions and subjected to NMR analysis. The results measure the exchange rate of the various NHs that can be detected in the native state and thus show which NHs were protected in the disordered form. This analysis has now been done for four proteins, with similar results. Elements of secondary structure do exist and these strongly resemble the native structure (Baum et al., 1989; Hughson et al., 1990; Jeng et al., 1990; Chyan et al., 1993; Dabora et al., 1996). The conclusion that molten globule forms represent kinetic folding intermediates is especially firm in the case of apomyoglobin and ribonuclease H where the same partially structured forms have also been demonstrated as kinetic intermediates by use of the HX pulse labeling experiment (Jennings & Wright, 1993; Raschke & Marqusee, 1997).

These results raise some interesting questions. Why is it that some proteins exhibit molten globule forms and others do not? How do the intermediate forms seen at equilibrium by hydrogen exchange in both native and molten globule conditions fit with our long-standing notions of two-state protein cooperativity? These issues are illuminated in Figure 7 which expands the results in Figure 5. Figure 7 emphasizes that at low denaturant concentration the partially unfolded forms are more stable than the fully unfolded state. The intermediates are normally invisible however due to their minimal population. When denaturant is increased in melting experiments, the free energy of the U state decreases and crosses N at the transition midpoint. In the transition region (arrows in Fig. 7), near the U–N crossover point, the intermediate forms are still only minimally populated, because the U state crosses them before it crosses N. This behavior explains the common observation that intermediate forms are not detected in melting experiments. It also illuminates the equilibrium molten globule issue. When some intermediate form exists at sufficiently low free energy, like the lower dashed line in Figure 7, it may be induced to cross N before U and thus become visible as an incompletely native equilibrium form.

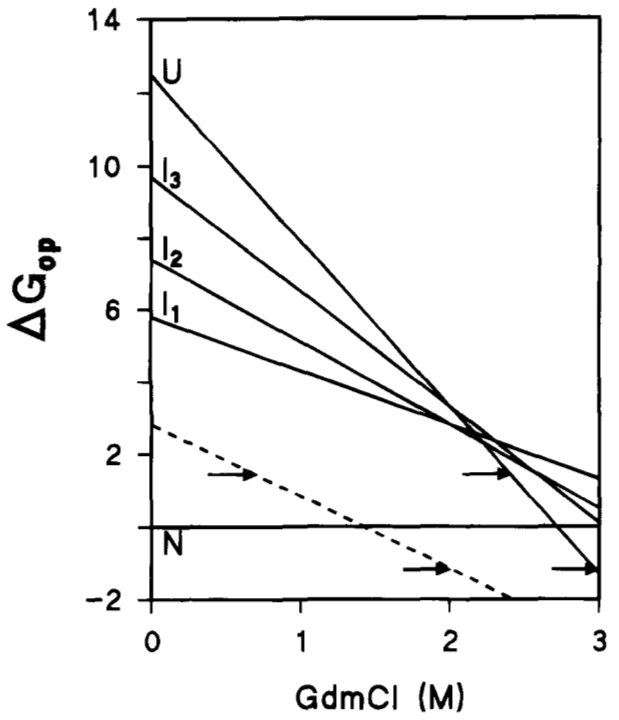

Fig. 7. A crossover diagram derived by extrapolating the unfolding isotherms in Figure 5 (from Bai & Englander, 1996). The dependence on denaturant of the free energy of the different unfolded states is shown relative to the native state as a reference. First-order transitions measurable at equilibrium occur when U crosses N or when a sufficiently low lying intermediate state (dashed line) crosses N before U does. The latter produces a detectable equilibrium molten globule. Arrows show the measurable transition region (95% N to 5% N). The failure of most intermediates to cross N before U does so is responsible for the usually observed two-state nature of protein melting.

Coda

The results described here illuminate some of the unusual capabilities of hydrogen exchange approaches. The HX pulse labeling experiment, performed in a kinetic folding mode, can detect and characterize kinetic folding intermediates that exist for less than one second. The native-state HX experiment defines a new and previously unsuspected level of cooperative protein substructure, and appears to reveal intermediate forms that constitute the entire protein folding pathway. The HX labeling capability can be used to define secondary structural elements in molten globule intermediates at equilibrium under moderately extreme conditions. These capabilities by no means exhaust the hydrogen exchange repertoire; the recent literature documents a wide range of other striking applications (see Englander et al., 1996).

The hydrogen exchange approach was created through the genius of Kaj Linderstrøm-Lang over 50 years ago. It remains one of his enduring contributions. Hydrogen exchange experiments transcend issues of three-dimensional structure—how proteins look—and delve into the substructural foundations of protein molecular design, the dynamic behavior of protein molecules, i.e., their time dimension and also their energy dimension, and thus can probe the fundamental secrets of how proteins work. These are capabilities that we are only now beginning to appreciate.

Acknowledgments

This work was supported by NIH research grant GM31847.

References

Anderegg RJ, Wagner DS. 1995. Mass spectrometric characterization of a protein ligand interaction. *J Amer Chem Sci 117*:1374–1377.

Bai Y, Englander SW. 1996. Future directions in folding: The multi-state nature of protein structure. *Proteins: Struct Funct Genet 24*:145–151.

Bai Y, Englander JJ, Mayne L, Milne JS, Englander SW. 1994a. Thermodynamic parameters from hydrogen exchange measurements. In: Ackers GK, Johnson ML, eds. *Energetics of Biological Macromolecules* vol. 259 of *Methods Enzymology*. pp 344–356.

Bai Y, Milne JS, Mayne L, Englander SW. 1994b. Protein stability parameters measured by hydrogen exchange. *Proteins: Struct Funct Genet 20*:4–14.

Bai Y, Sosnick TR, Mayne L, Englander SW. 1995. Protein folding intermediates studied by native state hydrogen exchange. *Science 269*:192–197.

Baldwin RL. 1993. Pulsed H/D-exchange studies of folding intermediates. *Curr Opin Struct Biol 3*:84–91.

Barksdale AD, Rosenberg A. 1982. Acquisition and interpretation of hydrogen exchange data from peptides, polymers, and proteins. *Methods Biochem Anal 28*:1–113.

Baum J, Dobson CM, Evans PA, Hanley C. 1989. Characterization of a partly folded protein by NMR methods: Studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry 28*:7–13.

Bax A. 1994. Multidimensional nuclear magnetic resonance methods for protein studies. *Curr Opin Struct Biol 4*:738–744.

Benson EE, Linderstrøm-Lang K. 1959. Deuterium exchange between myoglobin and water. *Biochim Biophys Acta 32*:579–581.

Berger A, Linderstrøm-Lang K. 1957. Deuterium exchange of poly-D,L-alanine in aqueous solution. *Arch Biochem Biophys 69*:106–118.

Bryan WD. 1970. The mechanism of hydrogen exchange in proteins. *Recent Prog Surf Sci 3*:101–120.

Chamberlain AK, Handel TM, Marqusee S. 1996. Detection of rare partially folded molecules in equilibrium with the native conformation of RNase H. *Nature Struct Biol 3*:782–787.

Chyan CL, Wormald C, Dobson CM, Evans PA, Baum J. 1993. Structure and stability of the molten globule state of guinea pig α-lactalbumin: A hydrogen exchange study. *Biochemistry 32*:5681–5691.

Dabora JM, Pelton JG, Marqusee S. 1996. Structure of the acid state of E coli ribonuclease H. *Biochemistry 35*:11951–11958.

Desai UR, Osterhout JJ, Klibanov AM. 1994. Protein structure in the lyophilized state: A hydrogen isotope exchange study with bovine pancreatic trypsin inhibitor. *J Amer Chem Soc 116*:9420–9422.

Englander JJ, Rogero JR, Englander SW. 1985. Protein hydrogen exchange studied by the fragment separation method. *Anal Biochem 147*:234–244.

Englander SW, Englander JJ. 1978. Hydrogen-tritium exchange. *Methods Enzymol 49*:24–39.

Englander SW, Kallenbach NR. 1984. Hydrogen exchange and structural dynamics of proteins and nucleic-acids. *Q Rev Biophys 16*:521–655.

Englander SW, Mayne L. 1992. Protein folding studied by hydrogen-exchange labeling and two-dimensional NMR. *Annu Rev Biophys Biomol Struct 21*:243–265.

Englander SW, Sosnick TR, Englander JJ, Mayne L. 1996. Mechanisms and uses of hydrogen exchange. *Curr Opin Struct Biol 6*:18–23.

Ernst RR, Bodenhausen G, Wokaun A. 1988. *Principles of nuclear magnetic resonance in one and two dimensions*. Oxford: Clarendon Press.

Fink AL. 1995. Compact intermediate states in protein folding. *Annu Rev Biophys Biomol Struct 24*:495–522.

Gallagher W, Tao F, Woodward C. 1992. Comparison of hydrogen exchange rates for bovine pancreatic trypsin inhibitor in crystals and in solution. *Biochemistry 31*:4673–4680.

Gueron M, Leroy JL. 1995. Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol 261*:383–413.

Hughson FM, Wright PE, Baldwin RL. 1990. Structural characterization of a partly folded apomyoglobin intermediate. *Science 249*:1544–1548.

Hvidt A, Johansen G, Linderstrøm-Lang K. 1960. Deuterium and ^{18}O exchange. In: Alexander P, Block RJ, eds. *The composition, structure, and reactivity of proteins*. New York: Pergamon Press. pp 101–130.

Hvidt A, Linderstrøm-Lang K. 1954. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim Biophys Acta 14*:574–575.

Hvidt A, Linderstrøm-Lang K. 1955a. Exchange of deuterium and ^{18}O between water and other substances. 3. Deuterium exchange of short peptides. *CR Trav Lab Carlsberg (Sér Chim) 29*:385–402.

Hvidt A, Linderstrøm-Lang K. 1955b. The pH dependence of the deuterium exchange of insulin. *Biochim Biophys Acta 18*:308–312.

Hvidt A, Nielsen SO. 1966. Hydrogen exchange in proteins. *Adv Protein Chem 21*:287–386.

Jeng M-F, Englander SW, Elöve GA, Wand AJ, Roder H. 1990. Structural description of acid-denatured cytochrome *c* by hydrogen exchange and 2D NMR. *Biochemistry 29*:10433–10437.

Jennings PA, Wright PE. 1993. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science 262*:892–896.

Jones DN, Bycroft M, Lubienski MJ, Fersht AR. 1993. Identification of the barstar binding site of barnase by NMR spectroscopy and hydrogen–deuterium exchange. *FEBS Lett 331*:165–172.

Kim KS, Woodward C. 1993. Protein internal flexibility and global stability: effect of urea on hydrogen exchange rates of bovine pancreatic trypsin inhibitor. *Biochemistry 32*:9609–9613.

Kim PS, Baldwin RL. 1982. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu Rev Biochem 51*: 459–489.

Kossiakoff AA. 1982. Protein dynamics investigated by the neutron diffraction-hydrogen exchange technique. *Nature 296*:713–721.

Krause IM, Linderstrøm-Lang K. 1955. Exchange of water and ^{18}O between water and other substances. 2. Alternative methods. *CR Lab Carlsberg (Sér Chim) 29*:367–384.

Kuroda Y. 1993. Residual helical structure in the C-terminal fragment of cytochrome *c*. *Biochemistry 32*:1219–1224.

Kuroda Y, Kidokoro S, Wada A. 1992. Thermodynamic characterization of cytochrome *c* at low pH: Observation of the molten globule state and of the cold denaturation process. *J Mol Biol 223*:1139–1153.

Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Struct Funct Genet 6*:87–103.

Linderstrøm-Lang K. 1955a. Deuterium exchange between beta-lactoglobulin and water. *Society of Biolog Chemists (Souvenir)*:191–194.

Linderstrøm-Lang K. 1955b. Deuterium exchange between peptides and water. *The Chemical Society's Special Publication No 2*.

Linderstrøm-Lang K. 1958. Deuterium exchange and protein structure. In: Neuberger A, ed. *Symposium on protein structure*. London: Methuen.

Linderstrøm-Lang K, Schellman JA. 1959. Protein structure and enzyme activity. In: Boyer PD, Lardy H, Myrback K, eds. *The Enzymes, 2nd ed*. New York: Academic Press. pp 443–510.

Loh SN, Prehoda KE, Wang J, Markley JL. 1993. Hydrogen exchange in unligated and ligated staphylococcal nuclease. *Biochemistry 32*:11022–11028.

Lumry R, Rosenberg A. 1975. The mobile defect hypothesis of protein function. *Col Int C N R S L'Eau Syst Biol 246*:55–63.

Mason SA, Bentley GA, McIntyre GJ. 1983. Deuterium exchange in lysozyme at 1.4 Å resolution. In: Schoenborn BP, ed. *Neutrons in biology: Neutron scattering analysis for biological structures*. New York: Plenum.

Matthews CR. 1993. Pathways of protein folding. [Review]. *Annu Rev Biochem 62*:653–683.

Mayne L, Paterson Y, Cerasoli D, Englander SW. 1992. Effect of antibody binding on protein motions studied by hydrogen exchange labeling and two-dimensional NMR. *Biochemistry 31*:10678–10685.

Mayo SL, Baldwin RL. 1993. Guanidinium chloride induction of partial unfolding in amide proton exchange in RNase A. *Science 262*:873–876.

Miranker A, Robinson CV, Radford AE, Dobson CM. 1996. Investigation of protein folding by mass spectrometry. *FASEB J 10*:93–101.

Okazaki A, Ikura T, Nikaido K, Kuwajima K. 1994. The chaperonin GroEL does not recognize apo-alpha-lactalbumin in the molten globule state. *Nature Struct Biol 7*:439–445.

Orban J, Alexander P, Bryan P. 1994. Hydrogen-deuterium exchange in the free and immunoglobin G-bound protein GB1-domain. *Biochemistry 33*:5702–5710.

Pace CN. 1986. Determination and anlysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol 131*:266–280.

Paterson Y, Englander SW, Roder H. 1990. An antibody binding site on a protein antigen defined by hydrogen exchange and two-dimensional NMR. *Science 249*:755–759.

Pedersen TG, Sigurskjold BW, Andersen KV, Kjaer M, Poulsen FM, Dobson CM, Redfield C. 1991. A NMR study of the hydrogen exchange behaviour of lysozyme in crystals and solution. *J Mol Biol 218*:413–426.

Perrett S, Clarke J, Hounslow AM, Fersht AR. 1995. Relationship between equilibrium amide hydrogen exchange behavior and the folding pathway of barnase. *Biochemistry 34*:9288–9298.

Ptitsyn OB. 1994. Kinetic and equilibrium intermediates in protein folding. *Protein Engineering 7*:593–596.

Ptitsyn OB. 1995a. Molten globule and protein folding. [Review]. *Adv Protein Chem 47*:83–229.

Ptitsyn OB. 1995b. Structures of folding intermediates. *Curr Opin Struct Biol 5*:74–78.

Ptitsyn OB, Pain RH, Semisotnov GV, Zerovnik E, Razgulyaev OI. 1990. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett 262*:20–24.

Qian H, Mayo SL, Morton A. 1994. Protein hydrogen exchange in denaturant: quantitative analysis by a two-process model. *Biochemistry 33*:8167–8171.

Raschke TM, Marqusee S. 1997. The kinetic folding intermediate of ribonuclease H resembles partially unfolded molecules detected under native conditions and in the acid molten globule. *Nature Struct Biol 4*:298–304.

Richards FM. 1979. Packing defects, cavities, volume fluctuations, and access to the interior of proteins, including some general comments on surface area and protein structure. *Carlsberg Res Commun 44*:47–63.

Robinson CV, Gross M, Eyles SJ, Ewbank JJ, Mayhew M, Hartl FU, Dobson CM, Radford SE. 1994. Conformation of GroEL-bound alpha-lactalbumin probed by mass spectrometry. *Nature 372*:646–651.

Roder H, Elöve GA, Englander SW. 1988. Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR. *Nature 335*:700–704.

Rodgers ME, Englander JJ, Englander SW, Harrington WF. 1996. Measurement of protein structure change in active muscle by hydrogen–tritium exchange. *Biophys Chem 59*:221–230.

Rogero JR, Englander JJ, Englander SW. 1986. Individual breathing reactions measured by functional labeling and H-exchange methods. *Enzyme Structure* vol. 131L of *Methods Enzymology*. pp 508–517.

Rosa JJ, Richards FM. 1979. An experimental procedure for increasing the structural resolution of chemical hydrogen exchange measurements on proteins: Application to ribonuclease S peptide. *J Mol Biol 133*:399–416.

Schellman JA. 1987. Selective binding and solvent denaturation. *Biopolymers 26*:549–559.

Schellman JA. 1990. A simple model for solvation in mixed solvents. Applications to the stabilization and destabilization of macromolecular structures. *Biophys Chem 37*:121–140.

Scholtz JM, Robertson AD. 1995. Hydrogen exchange techniques. *Methods Molec Biol 40*:291–311.

Swint-Kruse L, Robertson AD. 1996. Temperature and pH dependences of hydrogen exchange and global stability for ovomucoid third domain. *Biochemistry 35*:171–180.

Tanford C. 1970. Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv Protein Chem 24*:1–95.

Thornton K, Gorenstein DG. 1994. Structure of glucagon-like peptide (7-36) amide in a dodecylphosphocholine micelle as determined by 2D NMR. *Biochemistry 33*:3532–3539.

Tsuboi M, Nakanishi M. 1979. Overall and localized fluctuation in the structure of a protein molecule. *Adv Biophys 12*:101–130.

Udgaonkar JB, Baldwin RL. 1988. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature 335*:694–699.

Wagner G, Wuthrich K. 1979. Structural interpretation of amide proton exchange in BPTI and related proteins. *J Mol Biol 134*:75–94.

Wagner G, Wuthrich K. 1982. Amide proton exchange and surface conformation of BPTI in solution: studies with 2D NMR. *J Mol Biol 160*:343–361.

Werner MH, Wemmer DE. 1992. Identification of a protein binding surface by differential amide hydrogen exchange measurements. *J Mol Biol 225*:873–889.

Woodward CK. 1994. Hydrogen exchange rates and protein folding. *Curr Opin Struct Biol 4*:112–116.

Woodward CK, Simon I, Tuchsen E. 1982. Hydrogen exchange and the dynamic structure of proteins. *Mol & Cell Biochem 48*:135–160.

Wu J, Gorenstein DG. 1993. Structure and stability of cytochrome *c* in nonaqueous solvents by 2D NH exchange NMR. *J Amer Chem Soc 115*:6843–6850.

Wu LC, Laub PB, Elöve GA, Carey J, Roder H. 1993. A noncovalent peptide complex as a model for an early folding intermediate of cytochrome c. *Biochemistry* 32:10271-10276.

Wuthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.

Zahn R, Spitzfaden C, Ottiger M, Wuthrich K, Pluckthun A. 1994. Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. *Nature 368*:261–265.

Zhang Z, Post CB, Smith DL. 1996. Amide hydrogen exchange determined by mass spectrometry: Application to rabbit muscle aldolase. *Biochemistry 35*:779–791.