

## Protein Folding, Structure and Function

**Heinrich Roder, Ph.D.**, *Senior Member; Adjunct Professor of Biochemistry and Biophysics, and Member of the Graduate Group in Biochemistry and Molecular Biophysics, University of Pennsylvania*

**Hong Cheng, Ph.D.**, *Staff Scientist*

**Harvey H. Hensley, Ph.D.**, *Staff Scientist*

**Dharmaraj Samuel, Ph.D.**, *Postdoctoral Fellow*

**Pavel Elistratov,\*** *Graduate Student, Moscow State University, Moscow, Russia*

**Paul W. Riley, B.S.**, *Graduate Student, Temple University, Philadelphia, PA*

**Colin M. Hayden,\*** *Student Assistant, La Salle University, Philadelphia, PA*

**Lisa Wang,\*** *Student Assistant, Germantown Academy, Fort Washington, PA*



Much of our research is aimed at understanding the mechanism of protein folding, focusing primarily on early structural events in the transition from the unfolded to the native (folded) state of a protein. Many proteins encounter partially organized structures during early stages of folding, and these are thought to be important for directing a protein toward its unique native conformation. However, important questions remain to be answered concerning the nature and origin of the kinetic barriers encountered during early stages of folding and the structural properties of the intermediate states. Are compact partially folded states the result of a non-specific chain collapse or more specific folding events? What is the relative importance of local vs. long-range interactions? Do intermediates contain native-like tertiary interactions? For some proteins, intermediates are populated even at equilibrium, and this raises further questions concerning the origin of structural cooperativity, the relationship between kinetic and thermodynamic intermediates, and any residual interactions remaining in the denatured state.

We have addressed these fundamental questions by coupling advanced mixing techniques for rapid initiation of folding reactions with a variety of detection methods, including intrinsic and extrinsic fluorescence probes and H/D exchange labeling experiments with NMR detection. We use these approaches, often in combination with protein engineering, to gain detailed insight into the folding mechanism of a diverse set of proteins (for a recent comprehensive review, see (1)).

Studies of protein folding *in vitro* provide the necessary framework for investigating the folding of proteins in their cellular environments, protein trafficking and degradation. Protein stability and folding also play a central role in our understanding of the biological consequences of mutations and in *de novo* design of proteins. Moreover, insight into the properties of protein-folding intermediates is critical for a mechanistic understanding and treatment of a wide range of diseases that can be linked to the aggregation of partially denatured or misfolded forms of proteins.

We are especially interested in elucidating the structural basis for formation of amyloid fibrils, using fragments of the human prion as a test case. By coupling NMR with quenched hydrogen exchange techniques, we can observe the degree of solvent protection in the fibrillar state of an amyloidogenic peptide or protein. Together with molecular dynamics simulations (with R.L. Dunbrack<sup>§</sup>), the data provide unique insight into the structural basis of prion protein fibril formation. We also rely on rapid mixing methods to explore the folding mechanism of the human prion protein in order to gain insight into the role of partially folded intermediates in the conversion of the normal cellular form of the protein into its cytotoxic and infectious aggregates (in collaboration with A.C. Apetri<sup>a</sup> and W.K. Surewicz<sup>a</sup>).

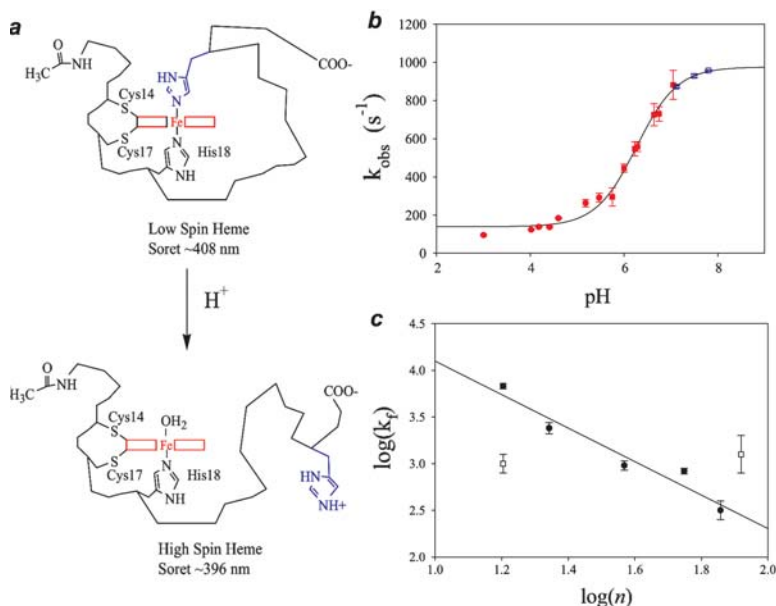
In other ongoing projects we make extensive use of multidimensional NMR methods to investigate the structure, dynamics and molecular interactions of various proteins of biological and/or clinical interest, including blood coagulation factor XI (with P.N. Walsh<sup>b</sup>), domain 5 of high molecular weight kininogen and its role in angiogenesis (with R.W. Colman<sup>b</sup>) and protein modules involved in signal transduction, such as CARD domains (with Y.C. Park<sup>§</sup>) and PDZ domains (with Z. Bu<sup>§</sup>).

**Kinetics of loop formation and breakage in unfolded cytochrome *c*.** Roder, in collaboration with Kurchan,<sup>d</sup> Bowler<sup>d</sup>

The formation of loops is among the earliest events in protein folding. By measuring the rate of formation for different loops in unfolded proteins, we can gain insight into the dynamic properties of the denatured state and detect any residual tertiary interactions. Conversely, the rate of loop breakage is related to the persistence of nascent structures during initial stages of folding. Cytochrome *c* offers a unique opportunity to study these dynamic events because of the strong affinity of the heme iron for deprotonated histidines, (see Figure 1a). His18 remains bound to the heme under both native and denaturing conditions while additional histidines elsewhere in the sequence (in blue in Figure 1a) have a strong tendency to form non-native ligand interactions in the denatured state.

Rates of loop formation and breakage were measured by monitoring the heme absorbance changes associated with pH jumps in either direction across the histidine titration range (2). The use of continuous- and stopped-flow

mixing methods allowed us to measure the pH-dependent rates for a series of variants of yeast cytochrome *c* containing engineered His residues at several positions designed to form loops ranging in size from 16 to 83 residues. Figure 1b shows representative kinetic results for the formation (red symbols) and breakage (black symbols) of a 56-residue loop involving a His at position 73. A simple two-step mechanism allowed us to extract the elementary rates of loop formation,  $k_f$ , and breakage,  $k_b$ , as a function of loop size. The magnitude and range of  $k_f$  values (300–7000  $s^{-1}$ ) measured for different variants (Figure 1c) are consistent with loop formation being reaction-limited rather than diffusion-limited. The observed decrease in  $k_f$  with increasing loop size yields a scaling factor of  $-1.8$ . This is consistent with the values expected for a random coil with excluded volume. The rate of loop breakage,  $k_b$ , is lower than the rate of dissociation of free histidine from the metal ion and goes through a minimum at an intermediate loop size of 37 residues. Our conclusion that a substantial amount of residual tertiary structure persists in the denatured state is



**Figure 1.** a) Schematic illustration of loop formation in unfolded cytochrome *c* mediated by histidine-heme ligation. Since only the neutral form of His can ligate the positively charged heme iron, loop formation is coupled with deprotonation of the His side chain, and vice versa. b) Plot of the observed rate of His73 linkage to the heme in iso-1 cytochrome *c* vs. pH. Rates of loop formation ( $k_f$ , squares) and breakage ( $k_b$ , circles) were measured in upward and downward pH jump experiments, respectively, using stopped-flow (red) or continuous-flow (black) methods. c) Logarithmic plot of the rate of loop formation,  $k_f$  vs. loop size,  $n$ . The observed slope ( $-1.8$ ) is consistent with the loop closure dynamics of a Gaussian (random) chain with excluded volume.

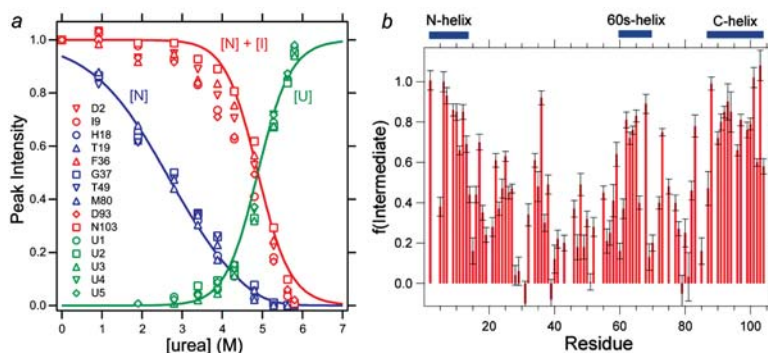
further supported by the observed variation in loop breakage rates for different histidines.

**Structural analysis of partially unfolded states of cytochrome *c* by NMR.** *Latypov, Cheng, Wang, Roder*

In contrast to other small proteins, cytochrome *c* is found to populate partially structured intermediate states not only in kinetic experiments, but also under certain equilibrium conditions. To gain further insight into the structure and stability of such equilibrium intermediates, we monitored the reversible denaturant-induced unfolding transition of oxidized horse cytochrome *c* using a variety of spectroscopic methods (3), such as, fluorescence, circular dichroism (CD) and heme absorbance bands. The normalized transition curves for different probes diverge between 1 and 2.5 M GuHCl, indicating that a partially unfolded intermediate state is populated under mildly denaturing conditions. The use of global fitting techniques allowed us to simultaneously fit a three-state equilibrium model to the combined data, and thus revealed the spectroscopic properties of the equilibrium intermediate. In particular, a global three-state fit of the heme absorbance data in the Soret region (350–450 nm) made it possible to deconvolute the heme absorbance spectrum of the intermediate. To rule out the possibility that heme misligation plays a role in stabilizing intermediates, we showed that a cyt *c* variant in which the predominant non-native ligand, His33, is replaced by Asn shows identi-

cal three-state equilibrium behavior as the wild-type protein.

The fact that the cytochrome *c* intermediate is well populated within the equilibrium unfolding transition (e.g., it reaches nearly 50% of the total population in 4.3 M urea) makes it possible to study its structural properties in more detail by the use of NMR (3). By recording a series of 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation NMR spectra vs. urea concentration, we were able to follow the changes in peak intensity for individual backbone amide groups (Figure 2). Some residues show greatly diminished intensity already at low urea concentrations (blue in Figure 2), while others, especially those in  $\alpha$ -helical regions (red), remain at their native level prior to the main unfolding transition. The main unfolding transition gives rise to a new set of sharp peaks at random-coil positions (green). The intensities of the blue peaks track the population of the native state (N) predicted by the optical data (solid line) while the red peaks follow the urea dependence expected for the combined population of native and intermediate states (I). The NMR data thus provide striking evidence for the formation of a partially structured state at intermediate urea concentrations. While many peaks approach the limiting cases indicated by the blue and red lines in Figure 2a, others fall in between, indicating partial disruption or dynamic contributions. In Figure 2b the data are parametrized in terms of a fractional contribution of the intermediate to the intensity of each assigned HSQC peak.



**Figure 2.** NMR analysis of an equilibrium-folding intermediate in cytochrome *c*. **a)**  $^{15}\text{N}$ - $^1\text{H}$  cross peak intensities vs. urea concentrations, including residues in locally unfolded (blue) and structured (red) regions of the intermediate state. Representative peaks assigned to the fully unfolded form are shown in green. **b)** Fractional contribution of intermediate population to N-state peak intensity. Values near 1 indicated regions that retain native-like structure and values near 0 indicate locally unfolded segments.

Values near 1 indicating ordered regions are seen mainly for the relatively hydrophobic segments of the cytochrome *c* sequence, including the three  $\alpha$ -helices and a cluster of apolar residues in the 30s. Together with the optical studies (3), the results indicate that the partially unfolded form of cytochrome *c* popu-

lated at low denaturant concentrations contains many native-like features, including a cluster of stable  $\alpha$ -helices and a partially buried heme group. On the other hand, the native Met80-iron linkage and other specific tertiary interactions are disrupted in the intermediate, and most of the loop regions are disordered.

## Publications

1. Roder, H., Maki, K., Cheng, H. Early events in protein folding explored by rapid mixing methods. *Chem. Rev.* (in press).
2. Kurchan, E., Roder, H., Bowler, B.E. Kinetics of loop formation and breakage in the denatured state of iso-1-cytochrome *c*. *J. Mol. Biol.* **353**:730-743, 2005.
3. Latypov, R.F., Cheng, H., Roder, N.A., Zhang, J., Roder, H. Structural characterization of an equilibrium unfolding intermediate in cytochrome *c*. *J. Mol. Biol.* **357**:1009-1025, 2006.

<sup>§</sup> Fox Chase researcher

\* Personnel left Fox Chase

<sup>a</sup> A.C. Apetri, W.K. Surewicz: Case Western Reserve University, Cleveland, OH

<sup>b</sup> P.N. Walsh, R.W. Colman: Temple University, Philadelphia, PA

<sup>c</sup> E.W. Kurchan, B.E. Bowler: University of Denver, Denver, CO