

Structural and biophysical characterization of protein folding.

The protein folding problem addresses how amino acid sequences determine protein structures [1]. The problem remains of fundamental importance in molecular biology as the number of new sequences identified continues to vastly outpace known protein structures. Progress in this area has enormous potential dividends including streamlining structure determination, facilitating rational protein design, and understanding how proteins misfold in disease states.

My lab's approach to the protein folding problem has been to structurally characterize equilibrium intermediates that result when the native state is disrupted by denaturants or mutagenesis [2]. An improved understanding of the structural properties of initial states in protein folding is critical to the development of accurate theoretical models. Residual structure in denatured states has the potential to affect protein stability (by raising or lowering the free energy difference from the native state), as well as folding kinetics (depending on whether intermediates are on- or off-pathway). The role of partially folded structure in aggregative misfolding is becoming increasingly recognized [3].

While our studies use a multi-faceted experimental approach, the principle technique in our work is NMR spectroscopy. NMR has the advantage of being able to characterize structure in a solution environment relevant to protein folding, and is unique among spectroscopic techniques in its ability to provide information at atomic resolution

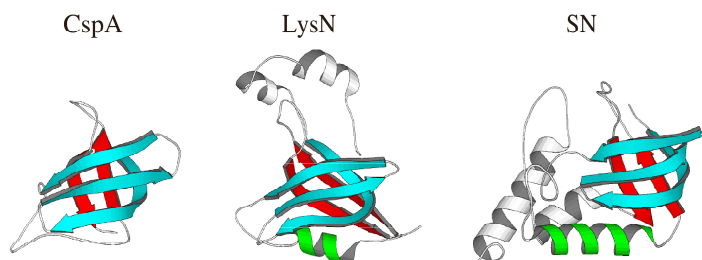
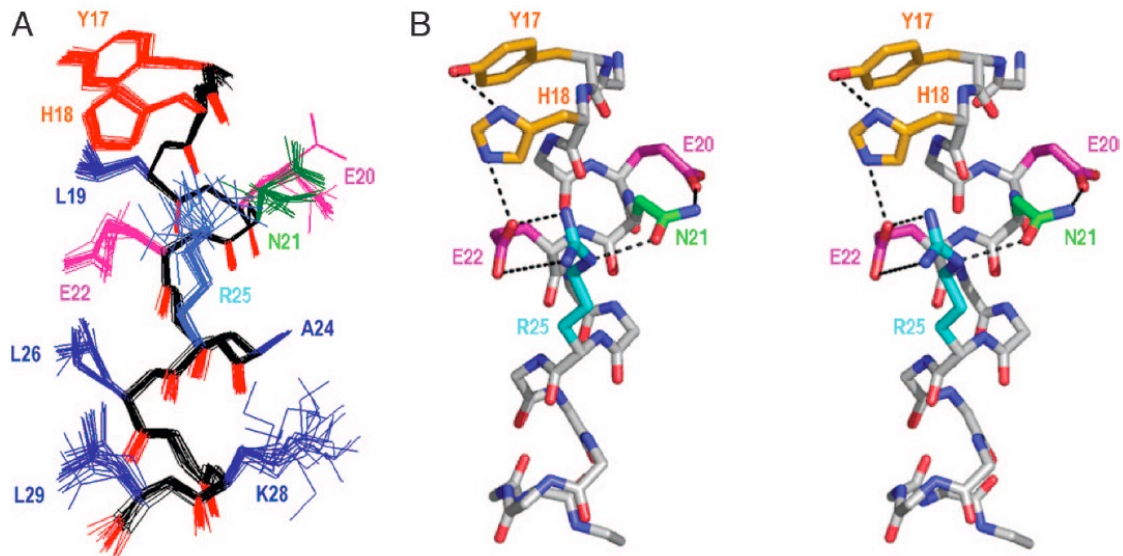


Fig. 1 : OB-fold proteins used for comparative protein folding studies. The structurally conserved β -barrel core of the OB-fold consists of a three-stranded β 1-3 menader (blue) and a two-stranded β 4-5 hairpin (red). Elements of structure that are unique to the individual proteins are left uncolored.

to unfolding. This is in fact what we found based on mutagenesis and hydrogen exchange experiments [6-8]. In addition to OB-fold proteins, we have looked at the folding of α -helical coiled coils [9-11] and protein fragments such as the ribonuclease S-peptide [12,13]. These systems have the advantage of a small size, and of a relatively simple α -helical structure that can be used to benchmark the NMR methods used to study non-native proteins (e.g. relaxation experiments [13], through-hydrogen bond coupling constants [12], and residual dipolar couplings [14]).

Our studies and those of others, have pointed to a large variety of folding intermediates [3,15-19]. These include sub-domains that retain part of the native tertiary structure independent of the rest of the protein, molten globules with native-like secondary structure but fluctuating side-chain structure, and highly disordered species that retain small amounts of native-like secondary structure. These results paint a picture of protein folding as a hierarchical assembly process, in which initially autonomous elements of structure become increasingly interdependent as the protein folds [20].

[4]. The model system we have used for most of our folding studies are three proteins that lack sequence homology but share an 'OB-fold' β -barrel structure [5]. The presence of a shared structural motif together with divergent elements of structure makes it possible to ask whether the conserved structures are the most stable



NMR structure of the ‘trigger site’, a folding initiation site for the α -helical GCN4 coiled-coil leucine zipper. (A) Best-fit superposition of the 20 lowest-energy NMR structures. The main chain is shown in black (carbons, nitrogens) and red (carbonyl oxygens). Selected side chains are shown in different colors and labeled according to their position in the GCN4 coiled-coil sequence (42). (B) Stereo side view of the p16-31 structure closest to the NMR ensemble average. Atoms of side chains forming the network of polar interactions are connected by dashed black lines.

Presumably, this reflects evolutionary constraints that favored unique functional native states over alternative conformations. Work on the OB-fold proteins in particular, suggests that the independent folding properties of a conserved β -barrel motif may be vestiges of evolutionary processes during which the autonomy of the motif was sacrificed to achieve a maximally cooperative integrated structure [6]. Partially formed structures are particularly prone to aggregation. This has been observed for all three of the OB-fold proteins under acid denaturing conditions. In all three proteins aggregation is initially promoted by the association of the conserved β 1-3 meander (Fig. 1), which corresponds to the most persistent structure under denaturing conditions [21-23].

Long-term goals:

- An unresolved issue in protein folding is whether a global folding topology persists under highly denaturing conditions [24]. We are using NMR residual dipolar couplings (RDCs) and distance restraints from engineered paramagnetic spin labels to test for the presence of conserved long-range structure in denatured states of OB-fold proteins.

- Insights from our protein folding studies will be used to develop strategies for modeling protein structures. A first effort in this direction has been the development of an approach for supplementing structure determination in cases where there are limited experimental data, with ‘hydrophobic distance restraints’ derived directly from the protein sequence. [25] Most work on structure prediction has focused on modeling new folds. If folds could be modeled reliably, there still remains the problem of how to go from a coarse fold-topology model to a usable high-resolution structure. Often, it is necessary to have multiple structures of a protein; for example, with and without an inhibitor. The problem

of determining closely related structures is easily tractable by X-ray crystallography (through Fourier difference maps), but current NMR approaches can require months of data collection per structure. We are exploring the application of RDCs to this problem. We anticipate that RDC-based refinement methods could be of general use in modeling closely related structures, for example obtaining snapshots of structural changes over a range of solution conditions.

Recent Publications

- Guardino, K.M., Sheftic, S.R., Slattery, R.E., & Alexandrescu, A.T. (2009) "Relative stabilities of conserved and non-conserved structures in the OB-fold superfamily" *Int. J. Mol. Sci.* 10, 2412-2430.
- Steinmetz, M.O., Jelesarov, I., Matousek, W.M., Honnappa, S., Jahnke, W., Missimer, J.H., Frank, S., Alexandrescu, A.T. & Kammerer, R.A. (2007) "Molecular basis of coiled coil formation". *Proc. Natl. Acad., USA.* 104, 7062-7067.
- Matousek, W.M., Ciani, B., Fitch, C.A., Garcia-Moreno E., B., Kammerer, R.A. & Alexandrescu, A.T. (2007) "Electrostatic contributions to the stability of the GCN4 leucine zipper structure", *J. Mol. Biol.* 374, 206-219.
- Watson, E., Matousek, W.M., Irimies, E.L. & Alexandrescu, A.T. (2007) "Partially folded states of staphylococcal nuclease highlight the conserved structural hierarchy of OB-fold proteins", *Biochemistry* 46, 9484-9494.
- Sallum, C.O., Martel, D.M., Fournier, R.S., Matousek, W.M & Alexandrescu, A.T. (2005) "Sensitivity of NMR residual dipolar couplings to perturbations in folded and unfolded staphylococcal nuclease". *Biochemistry* 44, 6392-6403.
- Alexandrescu, A.T. (2004) "Strategy for supplementing structure calculations using limited
- Alexandrescu, A.T., & Kammerer, R.A. (2003) "Structure and disorder in the ribonuclease S-peptide probed by NMR residual dipolar couplings". *Protein Sci.* 12, 2132-2140.
- Jaravine, V.A., Alexandrescu A. T., & Grzesiek, S. (2001) "Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide". *Protein Science* 10, 943-950.

References:

1. Anfinsen CB: **Principles that govern the folding of protein chains.** *Science* 1973, **181**:223-230.
2. Wang Y, Alexandrescu AT, Shortle D: **Initial studies of the equilibrium folding pathway of staphylococcal nuclease.** *Philos Trans R Soc Lond B Biol Sci* 1995, **348**:27-34.
3. Dobson CM: **Protein folding and misfolding.** *Nature* 2003, **426**:884-890.
4. Wuthrich K: **NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture).** *Angew Chem Int Ed Engl* 2003, **42**:3340-3363.
5. Murzin AG: **OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences.** *Embo J* 1993, **12**:861-867.

6. Alexandrescu AT, Gittis AG, Abeygunawardana C, Shortle D: **NMR structure of a stable "OB-fold" sub-domain isolated from staphylococcal nuclease.** *J Mol Biol* 1995, **250**:134-143.
7. Alexandrescu AT, Jaravine VA, Dames SA, Lamour FP: **NMR hydrogen exchange of the OB-fold protein LysN as a function of denaturant: the most conserved elements of structure are the most stable to unfolding.** *J Mol Biol* 1999, **289**:1041-1054.
8. Jaravine VA, Rathgeb-Szabo K, Alexandrescu AT: **Microscopic stability of cold shock protein A examined by NMR native state hydrogen exchange as a function of urea and trimethylamine N-oxide.** *Protein Sci* 2000, **9**:290-301.
9. Dames SA, Kammerer RA, Wiltschek R, Engel J, Alexandrescu AT: **NMR structure of a parallel homotrimeric coiled coil.** *Nat Struct Biol* 1998, **5**:687-691.
10. Dames SA, Kammerer RA, Moskau D, Engel J, Alexandrescu AT: **Contributions of the ionization states of acidic residues to the stability of the coiled coil domain of matrilin-1.** *FEBS Lett* 1999, **446**:75-80.
11. Kammerer RA, Jaravine VA, Frank S, Schulthess T, Landwehr R, Lustig A, Garcia-Echeverria C, Alexandrescu AT, Engel J, Steinmetz MO: **An intrahelical salt bridge within the trigger site stabilizes the GCN4 leucine zipper.** *J Biol Chem* 2001, **276**:13685-13688.
12. Jaravine VA, Alexandrescu AT, Grzesiek S: **Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide.** *Protein Sci* 2001, **10**:943-950.
13. Alexandrescu AT, Rathgeb-Szabo K, Rumpel K, Jahnke W, Schulthess T, Kammerer RA: **¹⁵N backbone dynamics of the S-peptide from ribonuclease A in its free and S-protein bound forms: toward a site-specific analysis of entropy changes upon folding.** *Protein Sci* 1998, **7**:389-402.
14. Alexandrescu AT, Kammerer RA: **Structure and disorder in the ribonuclease S-peptide probed by NMR residual dipolar couplings.** *Protein Sci* 2003, **12**:2132-2140.
15. Bai Y, Sosnick TR, Mayne L, Englander SW: **Protein folding intermediates: native-state hydrogen exchange.** *Science* 1995, **269**:192-197.
16. Dyson HJ, Wright PE: **Unfolded proteins and protein folding studied by NMR.** *Chem Rev* 2004, **104**:3607-3622.
17. Fieber W, Kristjansdottir S, Poulsen FM: **Short-range, long-range and transition state interactions in the denatured state of ACBP from residual dipolar couplings.** *J Mol Biol* 2004, **339**:1191-1199.
18. Redfield C: **Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins.** *Methods* 2004, **34**:121-132.
19. Shortle DR: **Structural analysis of non-native states of proteins by NMR methods.** *Curr Opin Struct Biol* 1996, **6**:24-30.
20. Alexandrescu AT, Jahnke W, Wiltschek R, Blommers MJ: **Accretion of structure in staphylococcal nuclease: an ¹⁵N NMR relaxation study.** *J Mol Biol* 1996, **260**:570-587.
21. Ye K, Wang J: **Self-association reaction of denatured staphylococcal nuclease fragments characterized by heteronuclear NMR.** *J Mol Biol* 2001, **307**:309-322.

22. Alexandrescu AT, Rathgeb-Szabo K: **An NMR investigation of solution aggregation reactions preceding the misassembly of acid-denatured cold shock protein A into fibrils.** *J Mol Biol* 1999, **291**:1191-1206.
23. Alexandrescu AT, Lamour FP, Jaravine VA: **NMR evidence for progressive stabilization of native-like structure upon aggregation of acid-denatured LysN.** *J Mol Biol* 2000, **295**:239-255.
24. Shortle D, Ackerman MS: **Persistence of native-like topology in a denatured protein in 8 M urea.** *Science* 2001, **293**:487-489.
25. Alexandrescu AT: **Strategy for supplementing structure calculations using limited data with hydrophobic distance restraints.** *Proteins* 2004, **56**:117-129.