Commentary

Why is protein folding so fast?

Robert L. Baldwin

Biochemistry Department, Beckman Center, Stanford University Medical Center, Stanford, CA 94305-5307

In the 1960s the question was: Why is protein folding so slow? Formation of an α-helix was shown to occur in the 10^{-7} s time range (1) or faster (2), by using ultrasonic absorption to monitor the helix-coil transition, whereas the refolding reactions of the few proteins studied (3–5) were in the seconds time range or slower. Today, several examples are known of small proteins that fold considerably faster, in the 10^{-3} s time range (ref. 6 and references therein).

In 1968 Levinthal (7, 8) turned the question around. Because the evidence at that time (3, 4, 9) suggested that small proteins fold without observable intermediates, Levinthal computed how long folding would take if there really were no intermediates. The answer was: longer than the lifetime of the universe. After that, the question changed to: why is folding so fast? Experimentalists searched for folding intermediates and theorists proposed mechanisms for achieving rapid folding.

In this issue, Sauer and coworkers (10) show that Arc repressor, a small, fast-refolding dimeric protein, can be made to refold much faster, apparently at the diffusion-controlled limit, by mutating three residues. The wild-type protein refolds in a second-order reaction at low protein concentrations (16 μM); the kinetics become complex at higher concentrations (11). At 25°C the second-order refolding rate constant is 8×10^6 M^{-1}·s^{-1} for wild type and 3×10^8 M^{-1}·s^{-1} for the mutant. The mutant rate constant is close to the estimated diffusion-limited value ($\approx 10^9$ M^{-1}·s^{-1}) for a reaction in which the rate-limiting step is the diffusion together and collision of two protein monomers. The refolding rate constant for the mutant, but not for wild type, is inversely proportional to the solvent viscosity when sucrose is added (10), in agreement with expectation for a diffusion-controlled reaction.

The mutations that make Arc repressor super-fast-folding are as follows. The wild-type protein contains a triad of buried ionizing residues (Arg-31, Glu-36, and Arg-40) which take part in ion pair interactions. Replacing any one of them with alanine destabilizes the protein, but replacing all three with nonpolar residues (Met-31, Tyr-36, and Leu-40; MYL) significantly increases the stability of the protein (12) as well as causing the MYL

mutant to refold much more rapidly than wild type. The ion-pair interactions occur within each monomer rather than between monomers (12). They might be expected to increase the folding rate by providing a structural nucleus, but instead they decrease the folding rate. In the structure of the dimer, the two monomers are intertwined and interact over much of the sequence both by hydrophobic interactions and by main-chain hydrogen bonding in a two-stranded antiparallel β-sheet (11, 12).

When a reaction becomes diffusion controlled, it is no longer possible to increase its rate by improving the chemistry of the reaction, for example by mutating more amino acid residues of Arc repressor. It will be interesting to apply this test in future work, to find out if the folding rate can nevertheless be increased by mutation. Sauer and coworkers conclude (10) that the MYL mutations have lowered the free energy barrier in the transition state for folding of wild type; as a result, the free energy barrier for the diffusion together of two monomers becomes the new rate-limiting barrier. Their result is particularly interesting in the context of the diffusion-collision model for the folding of monomeric proteins (13, 14). In this model, folded microdomains diffuse together, collide, and merge. If the MYL mutations enabled the mutant to form a stable monomeric folding intermediate, then it would be reasonably easy to understand how partly folded monomers could undergo diffusion-controlled folding into a native dimer. Sauer and coworkers (10) present good evidence, however, that this is not the case: they find that monomers of the MYL mutant are unfolded in refolding conditions, and therefore folding begins only when two unfolded monomers collide. There is, however, evidence from amide proton exchange rates that a partly folded monomeric intermediate may be present (15).
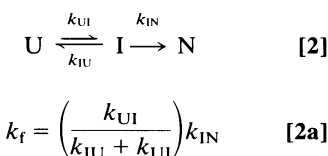
Recently there has been renewed interest in the hypothesis that small proteins might fold by a nucleation mechanism (16). The hallmark of a nucleation reaction is the absence of detectable intermediates, and early studies (3–5, 9) suggested that small proteins fold without observable intermediates. Later work, especially NMR-hydrogen exchange studies (see ref. 17 for review), showed that structural

intermediates are often well populated during the kinetics of refolding, and the nucleation mechanism fell into disfavor. It was revived recently when some particularly small proteins were found to refold rapidly (milliseconds) without detectable intermediates (6, 18, 19).

There are two basically different nucleation mechanisms. In the first mechanism, the nucleus is formed slowly, after which folding to the native conformation occurs rapidly. In a monomolecular folding reaction, the rate of folding (k_f) is the rate of forming the nucleated species (k_{UI}).

$$U \overset{\text{slow}}{\underset{k_{UI}}{\longrightarrow}} I \underset{k_{IN}}{\longrightarrow} N \qquad [1]$$

$$k_f = k_{UI}. \qquad [1a]$$
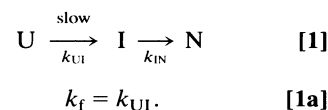
In the second mechanism, the nucleus is formed rapidly but is unstable, and the rate of folding is proportional to the small equilibrium fraction of I and to the rate constant k_{IN} for forming the native species from the nucleated species.

$$U \underset{k_{IU}}{\overset{k_{UI}}{\rightleftharpoons}} I \overset{k_{IN}}{\longrightarrow} N \qquad [2]$$

$$k_f = \left(\frac{k_{UI}}{k_{IU} + k_{UI}} \right) k_{IN} \qquad [2a]$$

Chymotrypsin inhibitor 2 has been postulated to refold in a nucleation-limited folding reaction (16) because it is small and fast folding and shows no detectable folding intermediate. Likewise, Arc repressor is small (the monomer has 53 residues), very fast folding (at 5 μM, the MYL mutant refolds with a half-time of 1 ms), and shows no folding intermediates by several standard tests in the concentration range studied (below 16 μM) (11). Thus, the next question becomes: Is the diffusion-controlled folding reaction of Arc repressor compatible with a nucleation-limited mechanism of folding? At first sight, the answer is no. Folding begins when two monomers diffuse together and collide. If folding must wait for a slow nucleation reaction in each monomer (nucleation mechanism 1), the two monomers would diffuse apart before folding begins and most collisions between two monomers would be nonproductive. If nucleation is rapid but occurs with a small equilibrium constant (nucleation mechanism 2), the probability that each of two

colliding monomers contains the folding nucleus is small, and the refolding rate constant would be much smaller than the diffusion-limited value. Thus, a nucleation mechanism seems to be incompatible with diffusion-limited folding. This is particularly true when nucleation must occur separately in each monomer. What if nucleation occurs just once, after the dimer is formed? The problem is that, if folding begins only after nucleation occurs, there is no folded structure to hold the dimer together before nucleation, and the unfolded monomers would diffuse apart. Consequently, the observation of a diffusion-controlled bimolecular folding reaction again focuses our attention on the question posed by Levinthal's calculation: Why is protein folding so fast?

1. Schwarz, G. (1965) *J. Mol. Biol.* **11,** 64–77.
2. Hammes, G. G. & Roberts, P. B. (1969) *J. Am. Chem. Soc.* **91,** 1812–1816.
3. Pohl, F. M. (1968) *Eur. J. Biochem.* **4,** 373–377.
4. Pohl, F. M. (1968) *Eur. J. Biochem.* **7,** 146–152.
5. Pohl, F. M. (1969) *FEBS Lett.* **3,** 60–64.
6. Schindler, T., Herrler, M., Marahiel, M. A. & Schmid, F. X. (1995) *Nat. Struct. Biol.* **2,** 663–673.
7. Levinthal, C. (1968) *J. Chim. Phys.* **65,** 44–45.
8. Levinthal, C. (1969) in *Mössbauer Spectroscopy in Biological Systems,* pp. 22–24.
9. Lumry, R., Biltonen, R. & Brandts, J. F. (1966) *Biopolymers* **4,** 917–944.
10. Waldburger, C. D., Jonsson, T. & Sauer, R. T. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2629–2634.
11. Milla, M. E. & Sauer, R. T. (1994) *Biochemistry* **33,** 1125–1133.
12. Waldburger, C. D., Schildbach, J. F. & Sauer, R. T. (1995) *Nat. Struct. Biol.* **2,** 122–128.
13. Karplus, M. & Weaver, D. L. (1976) *Nature (London)* **260,** 404–406.
14. Karplus, M. & Weaver, D. L. (1979) *Biopolymers* **18,** 1421–1437.
15. Burgering, M. J. M., Hald, M., Boelens, R., Breg, J. W. & Kaptein, R. (1995) *Biopolymers* **35,** 217–226.
16. Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10869–10873.
17. Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol.* **3,** 84–91.
18. Jackson, S. E. & Fersht, A. R. (1991) *Biochemistry* **30,** 10428–10435.
19. Huang, G. S. & Oas, T. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6878–6882.