



ELSEVIER

Biophysical Chemistry 101–102 (2002) 57–65

Biophysical
Chemistry

www.elsevier.com/locate/bpc

Submolecular cooperativity produces multi-state protein unfolding and refolding

S. Walter Englander*, Leland Mayne, Jon N. Rumbley

Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6059, USA

Received 11 January 2002; received in revised form 27 February 2002; accepted 27 February 2002

Abstract

Hydrogen exchange experiments show that cytochrome *c* and other proteins under native conditions reversibly unfold in a multi-step manner. The step from one intermediate to the next is determined by the intrinsically cooperative nature of secondary structural elements, which is retained in the native protein. Folding uses the same pathway in the reverse direction, moving from the unfolded to the native state through relatively discrete intermediate forms by the sequential addition of native-like secondary structural units.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Protein folding; Protein unfolding; Protein cooperativity; Hydrogen exchange

1. Introduction

John Schellman was the first to consider the cooperative nature of polypeptide helices and the implications for their dynamic unfolding behavior [1–3]. Schellman's groundbreaking work at the Carlsberg Laboratories, where Kai U. Linderstrøm-Lang and his co-workers were pioneering the development of hydrogen exchange approaches, provided the basis for Lang's visionary explanation of the dynamic basis of protein hydrogen exchange (HX) behavior. Linderstrøm-Lang and John Schellman together published the first discussion

of protein dynamics in 1959 [4], anticipating by many years the development of modern molecular dynamics and paving the way for the experimental study of protein and nucleic acid dynamics by hydrogen exchange techniques.

As can be expected from equilibrium thermodynamic considerations, protein molecules must unfold and refold all the time, even under native conditions, continually searching through all of their possible higher-energy forms and populating each according to its Boltzmann factor. This behavior is invisible to most measurements, which are dominated by signals from the predominant native state. Uniquely, HX rates receive no contribution from the undistorted native state, but are wholly determined by the transiently populated higher-energy distorted forms. Thus, we can

*Corresponding author. Tel.: +1-215-898-8042; fax: +1-215-898-2415.

E-mail address: walter@hxiris.med.upenn.edu (S.W. Englander).

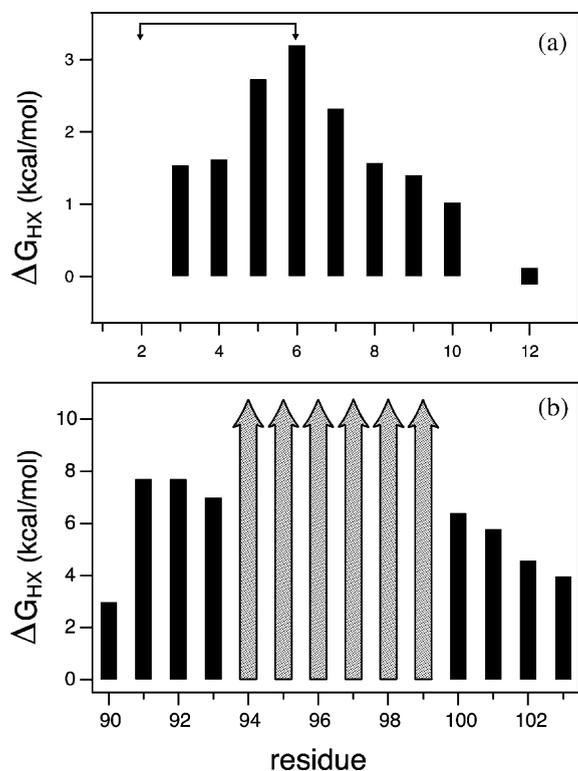


Fig. 1. Helix end fraying shown by HX protection. (a) Fraying is apparent from the C-terminal end of a 12-residue synthetic helix stabilized by an internal cross-linking lactam bridge (K2–D6) and an N-capping acetyl group. Data are from [35]. (b) Residues in the last turn of the C-terminal helix of Cyt *c* show similar fraying. Subsequent residues are buried and exchange far more slowly, by way of global unfolding. The data in (a) were obtained using 1D NMR, together with H–D exchange or saturation transfer H–H exchange, and in (b) by H–D exchange analyzed by 2D NMR.

exploit HX measurements to study the nature of dynamic, reversible unfolding reactions.

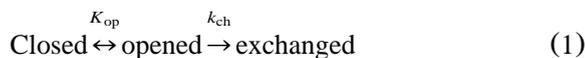
Recent evidence, organized here, confirms the continued role of helical cooperativity in globular proteins, generalizes this to the cooperativity of Ω -loops and groupings thereof, and suggests how this kind of unit cooperativity can determine the way in which protein molecules fold.

2. Cooperativity detected by hydrogen exchange

Fig. 1a shows an example of end fraying in an isolated polypeptide helix, monitored by H–D

exchange at an amino acid-resolved level. End fraying is a consequence of helical cooperativity and formally depends upon the coil-to-helix propagation factor [5–7]. As expected, residues at increasing depth into the helix experience progressively less unfolding, and therefore slower HX rates, because the exchange rate of each amide NH hydrogen depends on the fraction of time that each protecting H-bond is transiently broken.

Eq. (1), Eq. (2) and Eq. (3) formalize the relationship between HX rate and structural protection in the chemical thermodynamic view used by Linderstrøm-Lang [8]. In most cases (the so-called EX2 limit), HX rates can be interpreted in terms of an opening equilibrium constant (K_{op}) and the unprotected chemical exchange rate (k_{ch}) [9,10]. The HX rate measured then provides the free energy for the operative opening reaction, as used in Fig. 1:



$$k_{ex} = (\text{fraction open})k_{ch} \\ = K_{op}k_{ch}/(K_{op} + 1) \approx K_{op}k_{ch} \quad (2)$$

$$\Delta G_{op} = -RT \ln K_{op} \approx -RT \ln k_{ex}/k_{ch} \quad (3)$$

Fig. 1b shows the HX protection observed for the slowly exchanging NHs in the C-terminal helix of Cyt *c*. Here again the HX pattern at the end of the helix suggests fraying behavior, as expected if helix cooperativity is maintained in the protein. Farther into the helix where the fraying-dependent HX pathway is suppressed, other non-fraying pathways can dominate exchange. Exchange may occur by way of rare fluctuations that are able to break protecting H-bonds, perhaps one or two at a time in local fluctuations, or through cooperative whole-helix unfolding, or by way of even larger unfolding reactions up to global unfolding. For any given hydrogen, the pathway that provides the fastest exchange will control the HX rate observed.

3. Multi-state cooperativity in equilibrium protein unfolding

These fluctuational and unfolding modes can be analyzed by a technique known as native-state HX. Fig. 2 exhibits the H-exchange protection of all of

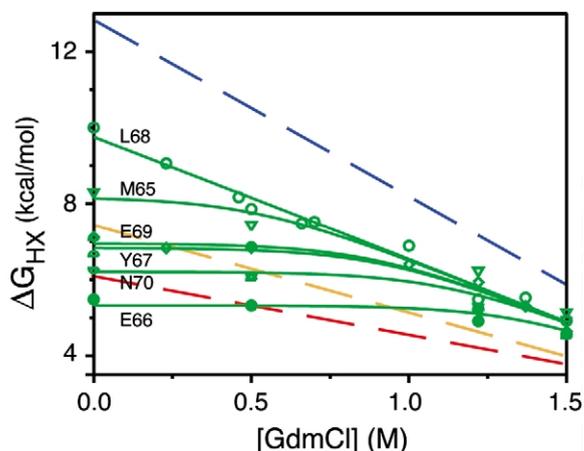


Fig. 2. Native state HX results for amide NHs in the 60s helix of Cyt *c*. At low denaturant concentration, the NHs exchange by way of local fluctuational pathways, indicated by the lack of denaturant dependence and the fact that neighboring residues exchange at very different rates. When denaturant concentration is increased, exchange becomes dominated by a larger cooperative unfolding, in which the entire 60s helix transiently unfolds. The dashed lines refer to other cooperative unfolding reactions, noted in Fig. 3. These data were obtained using H–D exchange measured at an amino acid-resolved level by 2D NMR (pH 7.0, 30 °C) [11].

the H-bonded amide hydrogens in the 60s helix of the globular protein Cyt *c*. The various hydrogens exchange at very different rates plotted in Fig. 2 in terms of the free energy of the opening reactions

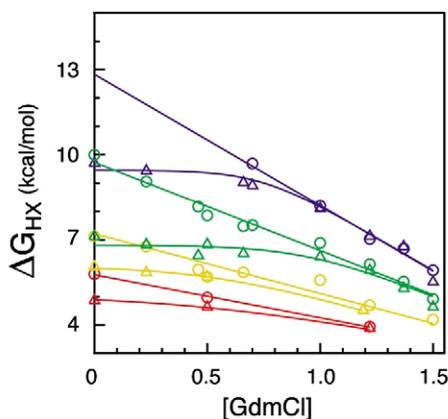
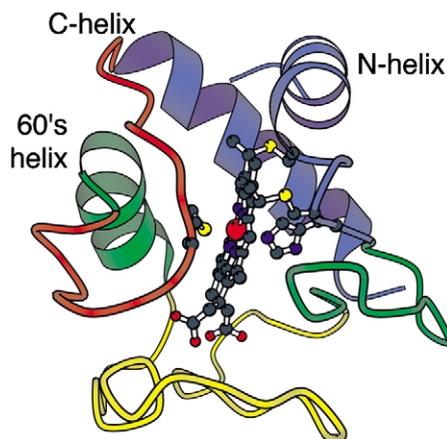


Fig. 3. Summary depiction of the several cooperative unfolding units in Cyt *c*. The color-coded foldons identified in this way are shown in the Cyt *c* model.

that break protecting H-bonds and allow each NH to exchange [Eq. (3)]. Most of the hydrogens are insensitive to increasing denaturant concentration ($m \sim 0$), apparently because the structural distortion that allows them to exchange is small and exposes little new denaturant binding surface. In contrast, the Leu68 NH is so well protected that it exchanges only by way of a much larger unfolding, indicated by its sharp denaturant dependence and large ΔG_{op} . As denaturant concentration is increased, the large unfolding is promoted and eventually comes to dominate the exchange of all of the NHs in the 60s helix. This behavior identifies the entire 60s helix as a cooperative unfolding–refolding unit (foldon).

The same approach reveals several submolecular cooperative unfolding units in Cyt *c*, summarized in Fig. 3. They are, in order of increasing unfolding energy, the red Ω loop, the yellow Ω loop, the Green 60s helix together with the green Ω loop, and the N- and C-terminal helices (*blue*), the unfolding of which brings the protein to the globally unfolded U state. The progressively larger denaturant dependence (serially increasing size) of the unfoldings is due to the fact that they occur in a sequential manner, with each unfolding including the lower lying units (see below).

In the initial decomposition of Cyt *c* by native-state HX [11], the three major Cyt *c* helices were



well defined by their many slowly exchanging H-bonded amide NHs. It could be clearly observed that each helix unfolds in a concerted way. The inference that the three Ω loops also act as concerted unfolding units was less definitive, because loops contain fewer H-bonded NHs than helices and only a few NHs in each loop exchanged slowly enough to measure. More recent work confirms the cooperative nature of the entire red Ω loop (16 residues with all six protected NHs measured; Linh Hoang, unpublished work), and shows that the yellow loop contains a nested ‘sub-yellow’ unfolding unit, consisting of the interior residues in the loop (18 residues with all nine protected NHs measured; Mallela M.G. Krishna and Yan Lin, unpublished results).

These results show that unfolding occurs in large unit steps involving entire secondary-structural elements of 10–20 amino acids in length rather than, for example, one or two amino acids at a time. Helices and Ω -loops, or groupings thereof, behave as cooperative unfolding units even when they are built into globular proteins. Native-state HX provides analogous results when applied to other proteins, using temperature and pressure as well as denaturant concentration to selectively promote and identify large unfolding reactions [12–18]. The behavior of β -structures remains to be determined.

4. Hidden intermediates mimic two-state behavior

The ‘crossover curves’ in Fig. 4a summarize the native-state HX results for Cyt *c*. Similar curves can be drawn for other proteins that have been similarly analyzed. These results broaden our current understanding of apparent two-state protein unfolding and show why intermediates have been so difficult to detect.

Fig. 4b shows the common equilibrium melting transition. Intermediates are not obvious. It is often concluded that intermediates do not exist. Melting curves are only measurable in a region where N has very little residual stability and the partially unfolded forms (PUFs) are at higher free energy than either N or U (Fig. 4a). The N and U forms predominate and melting appears to be two-state

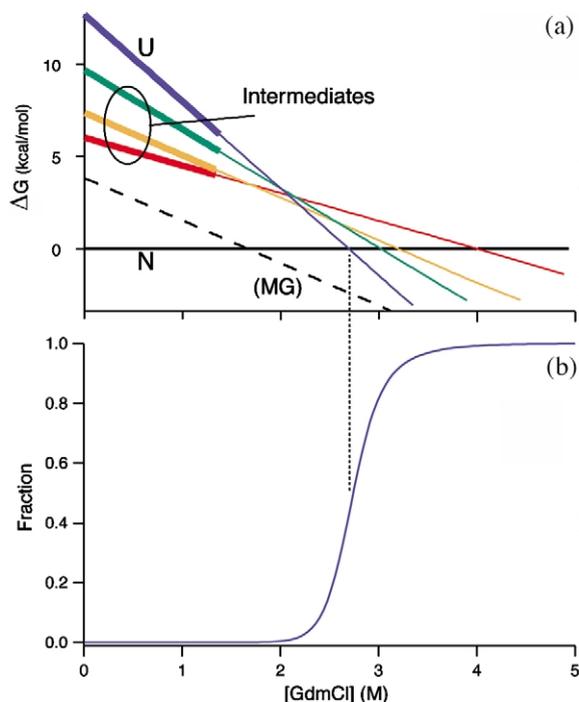


Fig. 4. The standard melting curve measured for Cyt *c*, commonly interpreted in terms of two-state unfolding, and crossover curves showing multi-state unfolding. The free energy axis is referenced to the native state ($\Delta G=0$) and the intermediate unfoldings are from Fig. 3. The melting midpoint occurs where the globally unfolded state, the blue open state, crosses N. The thickened regions show where native state HX data were actually measured [11]. The dashed line suggests how some PUF might become the most stable form when it is selectively promoted by some mildly destabilizing condition (e.g. low pH), as for molten globules.

(although careful analysis of Cyt *c* melting demonstrates the presence of the marginally populated intermediates [19]). The picture changes at more native conditions, where the PUFs can be more stable than the fully unfolded U state (Fig. 4a). Nevertheless, N predominates and the presence of the intermediates is still undetectable by most measurements.

In many proteins, global stability is much lower than for Cyt *c*. In this case, structured intermediates may not be more stable than U, even under native conditions. Here too, partially unfolded forms, although they must exist, will be invisible, or will be considered to be part of the U state. An

exception can occur in the region of mild destabilization where, in special cases, some units may be selectively unfolded, causing some PUF to populate, as suggested by the dashed line in Fig. 4a. These equilibrium PUFs (perhaps equivalent to ‘molten globules’) present the possibility of studying folding intermediates directly at equilibrium.

Intermediate accumulation also tends to be non-obvious in kinetic folding, especially for small proteins, which generally fold in an apparent two-state manner [20]. Only N and U are observed. We have suggested elsewhere that kinetic intermediates, although they may be important for the folding process, tend not to accumulate in kinetic folding because they occur after the rate-limiting step, usually the initial on-pathway barrier [21].

Finally, constructive on-pathway intermediates often fail to appear in theoretical folding simulations. This can be expected. The simplified theoretical models used in many computer simulations do not incorporate the separately cooperative units proposed here to determine intermediate formation. Accordingly, the models appear to act as single, globally cooperative entities. An independent cooperative unit, whether global or sub-global, is likely to organize itself one amino acid at a time in no particular order. This random search process accounts for the large part of the funnel landscape (Fig. 5).

In all of these cases, the usual measurements do not detect intermediate forms. Only N and U states are observed. This situation has produced the widespread conception that proteins are rigorously two-state all-or-none entities. Hydrogen exchange observations indicate otherwise.

5. Multi-state cooperativity in kinetic protein folding

It has been exceedingly difficult to define the nature and properties of partially folded intermediates in protein folding processes. The entire process is often over in less than 1 s, intermediates often fail to accumulate, and in any case kinetic intermediates cannot be isolated for study by crystallography or NMR. Various spectroscopic

techniques can detect fast signals during folding, but these are essentially one-parameter methods that do not provide detailed structural information.

Given these fundamental uncertainties, divergent views have been advanced concerning how proteins might fold from U to N. The different views differ fundamentally on the nature and role of folding intermediates [22,23], as suggested in Fig. 5. In the classical pathway view, folding proceeds by way of discrete, constructive pathway intermediates. Given the experimental results just described and other supporting evidence, it can no longer be doubted that relatively discrete partially folded forms occur. In the original funnel view (Fig. 5), derived from computer simulations of simplified models, folding proceeds one amino acid at a time in no particular order, as expected for a single cooperative unit. However, more recent theoretical progress does suggest distinct intermediates and preferential pathways.

How do these intermediates relate to the folding process? A quantity of evidence favors the view that partially folded native-like forms occur as constructive on-pathway intermediates [24–29]. Intermediates detected in kinetic folding appear to be essentially identical to intermediates described at equilibrium in native-state HX experiments, or in equilibrium molten globules, or that have been prepared by synthesis [28,30–32]. The intermediate forms are almost always seen to be partial replicas of the native structure at greater or lesser degree of advancement.

Cyt *c* currently provides the most fully worked out example. The HX pulse labeling experiment, which can define the structure of populated kinetic intermediates, showed an initial intermediate with the N- and C-terminal helices formed and the rest of the protein still unfolded [33]. Bai demonstrated that the N/C bihelical structure is on the kinetic folding pathway [25]. Exploration of the high-energy conformations available to Cyt *c* by native-state HX revealed a small set of intermediates. Among these, the highest free-energy blue unfolding unit is identical to the initial, kinetically observed N/C helix intermediate.

These results raise a fundamental question [38]. The PUFs are seen as equilibrium intermediates

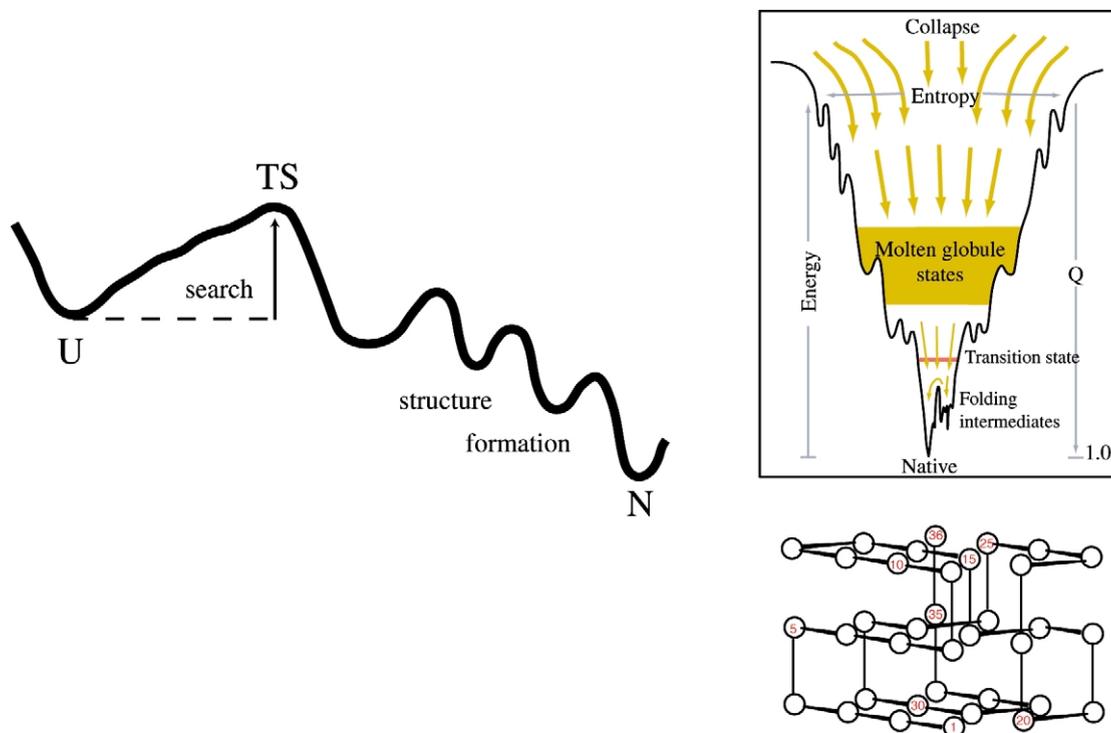


Fig. 5. Extreme alternative views of the protein folding process. Folding might be viewed as a well-ordered biochemical pathway (left) with distinct intermediates and barriers, or as a chaotic energetically downhill slide, folding one amino acid at a time in no particular order (right, adapted from [36]). The simplified computer models that appear to fold in this way do not have separately cooperative structural elements (e.g. the lattice model shown, adapted from [37]).

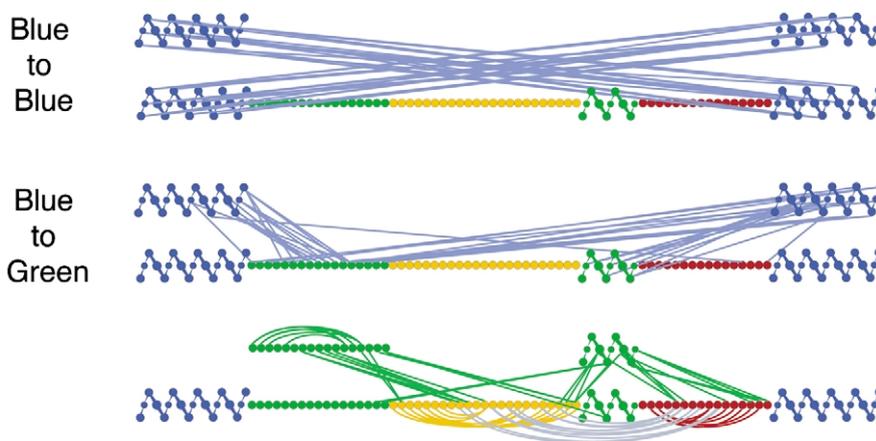
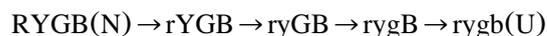


Fig. 6. Contact diagram for Cyt *c* showing atom–atom contacts between the color-coded foldons illustrated in Fig. 3. In the native context, the two blue helices stabilize each other and they fold first. The folded blue unit supports the subsequent folding of the green unit. The green unit can then guide and stabilize the folding of the yellow and red units.

by native-state HX. Do they represent the intermediates in a kinetic unfolding/refolding pathway, or do the various unfoldings occur more or less independently? A ‘stability labeling’ experiment with Cyt *c* makes this distinction [34]. The red unfolding unit could be stabilized simply by reducing the heme iron, which increases the stability of the Met80S–heme iron ligation by 3.2 kcal/mol (independently measured with small molecule models). A repeat of the native-state experiment with reduced Cyt *c* showed that the red unfolding moved to a free energy level higher by 3.2 kcal/mol, as expected, since unfolding of the red loop is likely to include the breakage of the weak S–Fe ligation. If the other unfoldings are independent of the red unit, they should experience little if any difference between the reduced and oxidized states (which differ only by a single electron on the heme iron). However, in a pathway situation (equation below), all of the higher-lying unfoldings should be raised by the same free energy increment found for red. The stability-labeling native-state HX experiment showed that both the yellow and green unfoldings are also raised by close to 3.2 kcal/mol (and the global unfolding by somewhat more) [34].

The apparent interpretation of these results is that the series of intermediate PUFs detected by the native-state HX experiment represents the states: red unfolded, red + yellow unfolded, red + yellow + green unfolded, and finally all unfolded to yield the U state [34]. These are just the forms necessary for a sequential unfolding pathway, as follows:



If so, then the reverse pathway must represent the major refolding pathway because, at equilibrium where the native-state HX experiment was carried out, each unfolding arrow in the reaction scheme must be matched by an equivalent reverse arrow in the refolding direction.

6. Pathway construction

This particular ordering of steps suggests a dramatically simple mechanism for pathway construction. Earlier-formed native-like structural ele-

ments create a docking surface that is able to guide and stabilize the formation of subsequent native-like elements. This can be inferred by inspection of the Cyt *c* model in Fig. 3 and in more detail by the contact diagram in Fig. 6. The N/C bihelix intermediate (the two blue helices in Fig. 6) is the first to form, according to independent kinetic pulse-labeling and equilibrium native-state HX results. These elements are far apart in the polypeptide sequence, but in the native context they are in intimate, mutually stabilizing contact. The green units are seen to be the next to form. Why? In the native context, the first-formed blue helices associate only with the green units (Fig. 6), and thus will naturally promote their formation. The green units, in the native context, can then promote the formation of the yellow and red units.

Thus, the native protein structure rationalizes the independently derived pathway sequence. The folding pathway is automatically created by a process of sequential stabilization, progressively adding native-like, intrinsically cooperative, secondary structural building blocks in a small number of steps that leads to the final native protein.

7. Conclusions

The results organized here argue that proteins fold by the sequential accretion of discrete units of the native structure. The crucial determinants are the intrinsically cooperative nature of the secondary structural building blocks and their sequential stabilization by the same tertiary interactions that knit them together in the native protein. The same structures and interactions that constitute the native state naturally generate the pathway for getting there.

Acknowledgments

This work was supported by research grants from the National Institutes of Health and the Mathers Charitable Foundation.

References

- [1] J.A. Schellman, The thermodynamics of urea solutions and the heat of formation of the peptide hydrogen bond,

- C. R. Trav. Lab. Carlsberg Ser. Chim. 29 (1955) 223–229.
- [2] J.A. Schellman, The stability of hydrogen-bonded peptide structures in aqueous solution, C. R. Trav. Lab. Carlsberg Ser. Chim. 29 (1955) 230–259.
- [3] J.A. Schellman, The factors affecting the stability of hydrogen-bonded polypeptide structures in solution, J. Phys. Chem. 62 (1958) 1485–1494.
- [4] K.U. Linderstrøm-Lang, J.A. Schellman, Protein structure and enzyme activity, in: P.D. Boyer, H. Lardy, K. Myrback (Eds.), *The Enzymes*, Academic Press, New York, 1959, pp. 443–510.
- [5] B.H. Zimm, J.K. Bragg, Theory of the phase transition between helix and random coil in polypeptide chains, J. Chem. Phys. 31 (1959) 526–535.
- [6] S. Lifson, A. Roig, On the theory of the helix-coil transition in polypeptides, J. Chem. Phys. 34 (1961) 1963–1974.
- [7] H. Qian, J.A. Schellman, Helix coil theories: a comparative study for finite length polypeptides, J. Phys. Chem. 96 (1992) 3987–3994.
- [8] K.U. Linderstrøm-Lang, Deuterium exchange and protein structure, in: A. Neuberger (Ed.), *Symposium on Protein Structure*, Methuen, London, 1958.
- [9] A. Hvidt, S.O. Nielsen, Hydrogen exchange in proteins, *Adv. Protein Chem.* 21 (1966) 287–386.
- [10] S.W. Englander, N.R. Kallenbach, Hydrogen exchange and structural dynamics of proteins and nucleic acids, *Q. Rev. Biophys.* 16 (1983) 521–655.
- [11] Y. Bai, T.R. Sosnick, L. Mayne, S.W. Englander, Protein folding intermediates studied by native state hydrogen exchange, *Science* 269 (1995) 192–197.
- [12] A.K. Chamberlain, T.M. Handel, S. Marqusee, Detection of rare partially folded molecules in equilibrium with the native conformation of RNase H, *Nat. Struct. Biol.* 3 (1996) 782–787.
- [13] R. Hiller, Z.H. Zhou, M.W.W. Adams, S.W. Englander, Stability and dynamics in a hyperthermophilic protein with melting temperature close to 200 °C, *Proc. Natl. Acad. Sci. USA* 94 (1997) 11329–11332.
- [14] A.K. Chamberlain, S. Marqusee, Touring the landscapes: partially folded proteins examined by hydrogen exchange, *Structure* 5 (1997) 859–863.
- [15] E.J. Fuentes, A.J. Wand, Local stability and dynamics of apocytochrome *b562* examined by the dependence of hydrogen exchange on hydrostatic pressure, *Biochemistry* 37 (1998) 9877–9883.
- [16] A.K. Bhuyan, J.B. Udgaonkar, Two structural subdomains of barstar detected by rapid mixing NMR measurement of amide hydrogen exchange, *Proteins* 30 (1998) 295–308.
- [17] M. Llinas, B. Gillespie, F.W. Dahlquist, S. Marqusee, The energetics of T4 lysozyme reveal a hierarchy of conformations, *Nat. Struct. Biol.* 6 (1999) 1072–1078.
- [18] J. Hollien, S. Marqusee, A thermodynamic comparison of mesophilic and thermophilic ribonucleases H, *Biochemistry* 38 (1999) 3831–3836.
- [19] L. Mayne, S.W. Englander, Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange, *Protein Sci.* 9 (2000) 1873–1877.
- [20] S.E. Jackson, How do small single-domain proteins fold?, *Fold. Des.* 3 (1998) R81–91.
- [21] T.R. Sosnick, L. Mayne, S.W. Englander, The barriers in protein folding, *Biophys. J.* 70 (1996) 2.
- [22] R.L. Baldwin, The nature of protein folding pathways: the classical versus the new view, *J. Biomol. NMR* 5 (1995) 103–109.
- [23] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, *Nat. Struct. Biol.* 4 (1997) 10–19.
- [24] S.W. Englander, Protein folding intermediates and pathways studied by hydrogen exchange, *Annu. Rev. Biophys. Biomol. Struct.* 29 (2000) 213–238.
- [25] Y. Bai, Kinetic evidence for an on-pathway intermediate in the folding of cytochrome *c*, *Proc. Natl. Acad. Sci. USA* 96 (1999) 477–480.
- [26] M. Jamin, S.R. Yeh, D.L. Rousseau, R.L. Baldwin, Submillisecond unfolding kinetics of apomyoglobin and its pH 4 intermediate, *J. Mol. Biol.* 292 (1999) 731–740.
- [27] D.V. Laurents, M. Bruix, M. Jamin, R.L. Baldwin, A pulse-chase-competition experiment to determine if a folding intermediate is on or off-pathway: application to ribonuclease A, *J. Mol. Biol.* 283 (1998) 669–678.
- [28] M. Arai, K. Kuwajima, Rapid formation of a molten globule intermediate in refolding of alpha-lactalbumin, *Fold. Des.* 1 (1996) 275–287.
- [29] C. Garcia, C. Nishimura, S. Cavagnero, H.J. Dyson, P.E. Wright, Changes in the apomyoglobin folding pathway caused by mutation of the distal histidine residue, *Biochemistry* 39 (2000) 11227–11237.
- [30] P.A. Jennings, P.E. Wright, Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin, *Science* 262 (1993) 892–896.
- [31] F.M. Hughson, P.E. Wright, R.L. Baldwin, Structural characterization of a partly folded apomyoglobin intermediate, *Science* 249 (1990) 1544–1548.
- [32] T.M. Raschke, S. Marqusee, The kinetic folding intermediate of ribonuclease H resembles the acid molten globule and partially unfolded molecules detected under native conditions, *Nat. Struct. Biol.* 4 (1997) 298–304.
- [33] H. Roder, G.A. Elöve, S.W. Englander, Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR, *Nature* 335 (1988) 700–704.
- [34] Y. Xu, L. Mayne, S.W. Englander, Evidence for an unfolding and refolding pathway in cytochrome *c*, *Nat. Struct. Biol.* 5 (1998) 774–778.
- [35] H.X. Zhou, L.A. Hull, N.R. Kallenbach, L. Mayne, Y. Bai, S.W. Englander, Quantitative evaluation of stabilizing interactions in a pre-nucleated alpha helix by hydro-

- gen exchange, *J. Am. Chem. Soc.* 116 (1994) 6482–6483.
- [36] P.G. Wolynes, J.N. Onuchic, D. Thirumalai, Navigating the folding routes, *Science* 267 (1995) 1619–1620.
- [37] V.I. Abkevich, A.M. Gutin, E.I. Shakhnovich, Specific nucleus as the transition state for protein folding: evidence from the lattice model, *Biochemistry* 33 (1994) 10026–10036.
- [38] J. Clarke, A.R. Fersht, An evaluation of the use of hydrogen exchange at equilibrium to probe intermediates on the protein folding pathway, *Fold. Des.* 1 (1996) 243–254.