# SHORT COMMUNICATION

# Chain Length is the Main Determinant of the Folding Rate for Proteins with Three-State Folding Kinetics

**Oxana V. Galzitskaya,* Sergiy O. Garbuzynskiy, Dmitry N. Ivankov, and Alexei V. Finkelstein**
*Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, Russia*

***ABSTRACT*** **We demonstrate that chain length is the main determinant of the folding rate for proteins with the three-state folding kinetics. The logarithm of their folding rate in water ($k_f$) strongly anticorrelates with their chain length $L$ (the correlation coefficient being −0.80). At the same time, the chain length has no correlation with the folding rate for two-state folding proteins (the correlation coefficient is −0.07). Another significant difference of these two groups of proteins is a strong anticorrelation between the folding rate and Baker's "relative contact order" for the two-state folders and the complete absence of such correlation for the three-state folders. Proteins 2003;51:162–166.**


**Key words: protein folding kinetics; two-state kinetics; three-state kinetics; contact order; protein size; protein topology; folding nucleus; transition state; rate of folding**

There is a significant difference in the folding behavior of small proteins with simple two-state kinetics and of larger proteins having a three-state folding kinetics.

The former have no visible intermediates in the course of folding, which therefore occurs as an "all-or-none" process under all experimental conditions; this peculiarity is demonstrated by linear dependence of logarithms of the folding (and unfolding) rates of these proteins on the denaturant concentration[1–4] [see Fig. 1(a)].

The latter fold via intermediates, which accumulate during the early stages of folding[1–4] when it occurs in denaturant-free water. This is demonstrated by a considerable decrease in the dependence of the folding rate logarithm on the denaturant concentration that occurs far off from the point of thermodynamic equilibrium between the unfolded and native states of these proteins, that is, when the native form is much more stable than the unfolded one [Fig. 1(b)].

It has been observed that the folding rate of two-state folding proteins is mainly determined by the topology of their native folds.[5] The logarithm of the folding rate of these proteins in water [$\ln(k_f)$] is strongly anticorrelated (at the level of −80–−85%) with a topological parameter called the "relative contact order,"[5] which is defined as

$$\mathrm{CO}\% = \frac{1}{L \cdot N} \sum_{i<j}^{N} \Delta L_{ij} \cdot 100\%, \qquad (1)$$

where $N$ is the total number of nonhydrogen atom–atom contacts (within 6 Å), $\Delta L_{ij}$ is the sequence separation, in residues, between contacting atoms $i$ and $j$ ($\Delta L_{ij} = 1$ when atoms $i$, $j$ belong to next-neighbor residues, etc.), and $L$ is the total number of residues in the protein. CO% is small for proteins stabilized mainly by local contacts and large when the contacts occur mainly between remote chain residues.

Despite a strong anticorrelation between CO% and folding rates reported earlier[5] and confirmed here for two-state folding proteins [circles in Fig. 2(a)], there is no trace of such a correlation for proteins with three-state kinetics [triangles in Fig. 2(a)].

Instead, the three-state folding proteins demonstrate a strong anticorrelation between their size and the folding rate logarithm [triangles in Fig. 2(b)]. This correlation is completely absent for two-state folding proteins [circles in Fig. 2(b)].

It should be noted that parameters that correlate with the chain length $L$ (e.g., the chain entropy, proportional to the total number of $\phi$, $\psi$, $\chi$ angles in the chain, the number of residues involved in secondary structure; etc.) correlate with the logarithm of folding rate for three-state folding proteins at the same level as the chain length, while no such correlation exists for two-state folding proteins.

Some basic correlations (or, better to say, anticorrelations) between the protein size and folding rate have been suggested during the last years from simulations and theoretical considerations of the folding process: $\ln(k_f) \sim L^{1/2}$ (Ref. 6), $\ln(k_f) \sim \ln(L)$ (Ref. 7), and $\ln(k_f) \sim L^{2/3}$ (Ref. 8).

*Correspondence to: Oxana V. Galzitskaya, Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia. E-mail: ogalzit@vega.protres.ru

Received 3 May 2002; Accepted 7 November 2002

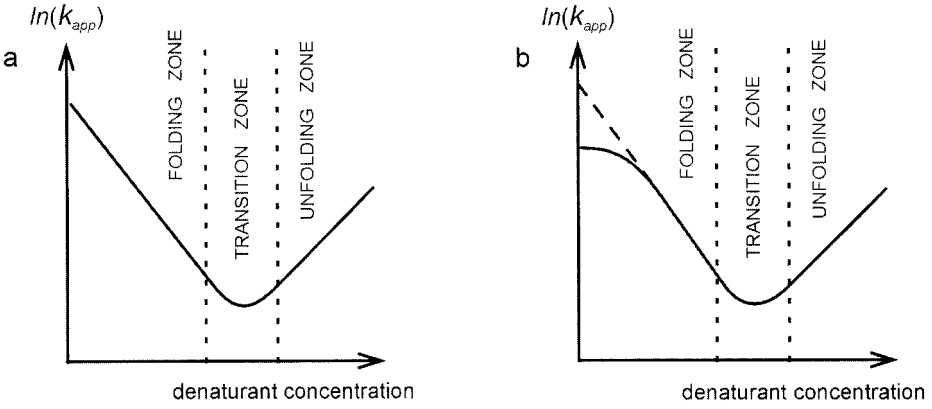


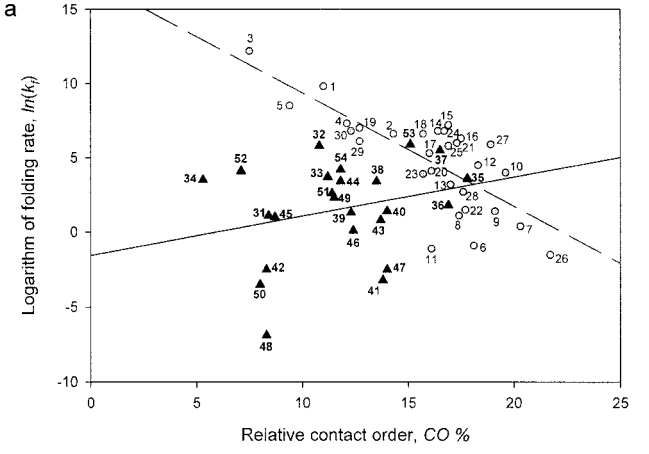

Fig. 1. Chevron plots: apparent rates of the folding/unfolding process ($k_{app}$) versus the denaturant concentrations; $k_{app} = k_f + k_{unf}$, where $k_f$ and $k_{unf}$ are the folding and unfolding rates.[60] Under the folding conditions, $k_{app} \approx k_f \gg k_{unf}$; under unfolding conditions, $k_{app} \approx k_{unf} \gg k_f$. The observed folding rate is minimal at the midtransition, where $k_f = k_{unf}$. (**a**) Typical plot for a protein having the two-state transition over the whole range of experimental conditions. (**b**) Typical plot for a protein having the three-state transition in water: It has the two-state transition only close to the zone of thermodynamic equilibrium between the native and denatured states. The broken line extrapolates the two-state folding rate to low denaturant concentrations.

All of them stress that the folding time must grow with the protein size. It seems that this tendency is in general correct when we consider the total set of proteins and peptides,[9–14] as well as the set of three-state folding proteins [Fig. 2(b)], but it cannot be traced for the pool of two-state folding proteins.[5,11]

We made an attempt to find out which of the above-mentioned dependences [$\ln(k_f) \sim L^{1/2}$, $\ln(k_f) \sim \ln(L)$, $\ln(k_f) \sim L^{2/3}$] is more consistent with the available experimental data. To this end, we calculated the correlation of $\ln(k_f)$ with $L^p$, where the power $p$ varies from 0 to 1 (it should be noted that the case $p \to 0$ corresponds to correlation with $\ln(L)$ because $L^p = \exp(p \ln(L)) = 1 + p \ln(L)$ when $p \to 0$.

The results are presented in Figure 3. Unfortunately, the observed correlation is virtually the same for $0 \le p \le 1$ (and becomes worse only when $p > 1$). This gives us no possibility to say which of the suggested theoretical dependences is more consistent with experiments.

Anyhow, the results presented in Figure 2 show that two- and three-state folding proteins have different determinants of the folding rate: contact order for the former[5] and chain length for the latter.



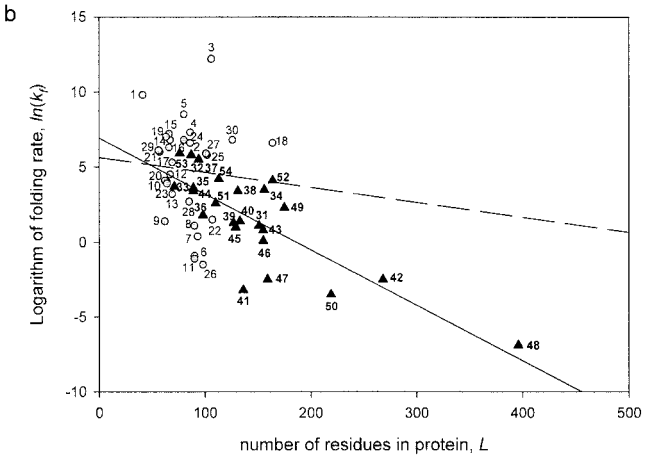




Fig. 2. (**a**) Natural logarithm of observed folding rate in water, $\ln(k_f)$, versus relative contact order, CO%, for proteins listed in Table I and numbered according to the table. Circles, proteins having two-state folding kinetics at all denaturant concentrations; triangles, proteins having three-state folding kinetics at low denaturant concentrations. If several folding rates in water are observed for some protein (see Table I), $\ln(k_f)$ is taken as the mean of their natural logarithms. The broken line represents the best linear fit for two-state folders only (the correlation coefficient is as significant as $-0.76$; the P-value associated with this correlation, $P = 0.006$, is extremely low, suggesting that the observed correlation is highly unlike to have arisen by chance). The solid line represents linear fit for three-state folders only (the correlation coefficient is 0.34, i.e., the correlation is small and, remarkably, its sign is opposite that for the two-state folders; the associated $P$-value is high, $P = 0.5$, suggesting that uncertainty in the correlation value is high). (**b**) Natural logarithm of observed folding rate in water versus chain length for the same set of proteins. The solid line represents the best linear fit for three-state folders only (triangles; the correlation coefficient is as significant as $-0.80$, the $P$-value being equal to 0.004). The broken line represents linear fit for two-state folders only (circles; the correlation coefficient is negligible, $-0.07$, $P = 0.5$).

## ACKNOWLEDGMENTS

The authors are grateful to D. Baker for presenting them their unpublished results. This work was supported by the Russian Foundation for Basic Research (Grants 01-04-48329 and 01-04-484000) and an International Research Scholar's Award to A.V.F. from the Howard Hughes Medical Institute.

**TABLE I. List of Proteins**[†]

| No. | Protein | Ref. | PDB | $L$[a] | CO% | $\ln(k_f)$ |
|---|---|---|---|---|---|---|
| Proteins with two-state kinetics | | | | | | |
| 1 | E3/E1 binding domain of dihydrolipoyl acyltransferase | 15 | 2PDD | 41 | 11.0 ± 0.4 | 9.8 |
| 2 | ACBP | 16 | 2ABD | 86 | 14.3 ± 0.3 | 6.6 |
| 3 | Cytochrome b562[b] | 17 | 256B | 106 | 7.5 | 12.2 |
| 4 | Colicin E9 immunity protein | 18 | 1IMQ | 86 | 12.1 | 7.3 |
| 5 | λ-repressor | 19 | 1LMB | 80 | 9.4 | 8.5 |
| 6 | Fibronectin 9th FN3 module | 20 | 1FNF | 90 | 18.1 | −0.9 |
| 7 | Twitchin | 21 | 1WIT | 93 | 20.3 | 0.4 |
| 8 | Tenascin (short form) | 22 | 1TEN | 90 (89) | 17.4 | 1.1 |
| 9 | SH3 domain (α-spectrin) | 23 | 1SHG | 62 (57) | 19.1 | 1.4 |
| 10 | SH3 domain (src) | 24 | 1SRL | 64 (56) | 19.6 | 4.0 |
| 11 | SH3 domain (Pl3 kinase)[c] | 25 | 1PNJ | 90 (86) | 16.1 | −1.1 |
| 12 | SH3 domain (fyn) | 26 | 1SHF | 67 (59) | 18.3 | 4.5 |
| 13 | Photosystem I accessory protein | —[d] | 1PSF | 69 | 17.0 | 3.2 |
| | | 27 | | | | 7.0 |
| 14 | CspB (*Bacillus subtilis*) | | 1CSP | 67 | 16.4 | |
| | | 28 | | | | 6.5 |
| 15 | CspB (*Bacillus caldolyticus*) | 28 | 1C9O | 66 | 16.9 | 7.2 |
| 16 | CspB (*Thermotoga maritima*) | 28 | 1G6P | 66 | 17.5 ± 0.4 | 6.3 |
| 17 | CspA | 29 | 1MJC | 69 | 16.0 | 5.3 |
| 18 | Cyclophilin A | 30 | 1LOP | 164 | 15.7 | 6.6 |
| 19 | DNA-binding protein sso7D[e] | 31 | 1C8C | 63 | 12.7 | 7.0 |
| 20 | IgG binding domain of streptococcal protein L[f] | 32 | 1HZ6 | 62 | 16.1 | 4.1 |
| 21 | Protein G | 33 | 1PGB | 57 (56) | 17.3 | 6.0 |
| 22 | FKBP12 | 34 | 1FKB | 107 | 17.7 | 1.5 |
| 23 | Cl2 | 35 | 2CI2 | 64 | 15.7 | 3.9 |
| 24 | Activation domain procarboxypeptidase A2 | 36 | 1AYE | 80 | 16.7 | 6.8 |
| 25 | Spliceosomal protein U1A[g] | 37 | 1URN | 102 (96) | 16.9 | 5.8 |
| 26 | Muscle-AcP[h] | 38 | 1APS | 98 | 21.7 ± 0.6 | −1.5 |
| 27 | S6 | 39 | 1RIS | 101 (97) | 18.9 | 5.9 |
| 28 | His-containing phosphocarrier protein | 40 | 1POH | 85 | 17.6 | 2.7 |
| 29 | N-terminal domain from L9 | 41 | 1DIV | 56 | 12.7 | 6.1 |
| 30 | Villin 14T | 42 | 2VIK | 126 | 12.3 | 6.8 |
| Proteins with three-state kinetics | | | | | | |
| 31 | Apomyoglobin[i] | 43 | 1A6N | 151 | 8.4 | 1.1 |
| 32 | Colicin E7 immunity protein | 18 | 1CEI | 87 (85) | 10.8 | 5.8 |
| 33 | Cro protein | 44 | 2CRO | 71 (65) | 11.2 | 3.7 |
| 34 | P16 protein | 45 | 2A5E | 156 | 5.3 | 3.5 |
| 35 | Twitchin Ig repeat 27 | 46 | 1TIT | 89 | 17.8 | 3.6 |
| 36 | CD2, 1st domain | 47 | 1HNG | 98 (95) | 16.9 | 1.8 |
| 37 | Fibronectin 10th FN3 module | 48 | 1FNF | 94 | 16.5 | 5.5 |
| 38 | IFABP from rat | 49 | 1IFC | 131 | 13.5 | 3.4 |
| 39 | ILBP[j] | 50 | 1EAL | 127 | 12.3 ± 0.5 | 1.3 |
| 40 | CRBP II | 49 | 1OPA | 133 | 14.0 | 1.4 |
| 41 | CRABP I | 49 | 1CBI | 136 | 13.8 | −3.2 |
| 42 | Tryptophan synthase α-subunit[k] | 51 | 1QOP | 268 (267) | 8.3 | −2.5 |
| 43 | GroEL apical domain (191–345) | 52 | 1AON | 155 | 13.7 | 0.8 |
| 44 | Barstar[l] | 53 | 1BRS | 89 | 11.8 | 3.4 |
| 45 | CheY | 54 | 3CHY | 129 (128) | 8.7 | 1.0 |
| 46 | Ribonuclease HI[m] | 55 | 2RN2 | 155 | 12.4 | 0.1 |
| 47 | DHFR (dihydrofolate reductase)[n] | 56 | 1RA9 | 159 | 14.0 | −2.5 |
| 48 | Tryptophan synthase β2-subunit[k] | 57 | 1QOP | 396 (390) | 8.3 | −6.9 |
| 49 | N-terminal domain from PGK | 58 | 1PHP | 175 | 11.5 | 2.3 |
| 50 | C-terminal domain from PGK[o] | 59 | 1PHP | 219 | 8.0 | −3.5 |
| 51 | Barnase | 60 | 1BNI | 110 (108) | 11.4 | 2.6 |
| 52 | T4 lysozyme[p] | 55 | 2LZM | 164 | 7.1 | 4.1 |
| 53 | Ubiquitin[q] | 61 | 1UBQ | 76 | 15.1 | 5.9 |
| 54 | Suc 1[r] | 62 | 1SCE | 113 (101) | 11.8 | 4.2 |

[†]Protein, name of a protein; Ref., reference to the original article on folding and unfolding kinetics; PDB, the Protein Data Bank[63] entry; $L$, the number of residues in protein used in the experimental study and (in parentheses) the number of residues that have 3D coordinates and thus contribute to the CO% calculations; CO%, relative contact order; $\ln(k_f)$, natural logarithm of the experimental folding rate ($s^{-1}$) measured in or extrapolated to pure water (i.e., at zero denaturant concentration). This is the list of single-domain proteins with no S-S bonds and covalent bonds to a ligand. If some protein was investigated at different temperatures, the experiment at the temperature closest to 25°C is presented. We took the slowest folding phase, which is not considered as the phase of *cis*/*trans* proline isomerization in the original article. If several 3D structures (i.e., several PDB entries) are available for some protein, the best-refined full-length X-ray structure is used in our calculation; in the absence of X-ray structure, the averaged NMR structure is used; in the absence of such a structure, the order parameter was averaged over all NMR models (in this case, the error of averaging is also given). The proteins numbered 1–30 exhibit the two-state folding within the whole range of experimental conditions and from 31–54 exhibit the three-state folding when the native state is much more stable than the denatured one.

[a]Length $L$ is taken as the total number of residues in the studied chain. If, instead, the total number of well-ordered residues is used, this does not significantly affect the reported correlations between length and logarithm of experimental folding time (the correlation coefficient for three-state proteins remains the same; for two-state proteins it changes only a little, from −0.07 to −0.06).

[b]Two-state folding is assumed by the long extrapolation made by authors.
[c]Although the authors reported that the SH3 domain from Pl3 kinase is 84aa long, it has been refolded with two additional N-terminal residues and four additional C-terminal residues. The latter four are absent in the PDB entry.
[d]Bowers and Baker, unpublished results.
[e]Folding of the mutant protein Y34W was studied; we used the structure of WT in our CO% calculations.
[f]Folding of the mutant protein Y47W was studied; we used the structure of this mutant in our CO% calculations.
[g]Folding of the mutant protein F56W was studied; we used the structure of mutant Y31H/Q36R in our CO% calculations.
[h]Folding of the mutant C21S of human protein was studied; we used the structure of WT of horse protein in our CO% calculations.
[i]We used the structure of holoform of myoglobin in our CO% calculations.
[j]We used the structure of mutant protein T118S of protein from pig in our CO% calculations instead of WT protein from rat.
[k]Folding of the protein from *Escherichia coli* was studied; we used the structure of the same protein from *Salmonella typhimurium* in our CO% calculations.
[l]Folding of the mutant protein C40A/C82A was studied; we used the structure of this mutant in our CO% calculations.
[m]Folding of the mutant protein C13A/C63A/C133A was studied; we used the structure of WT in our CO% calculations.
[n]We used the structure of mutant protein N37D in our CO% calculations.
[o]Folding of the mutant protein W290Y was studied; we used the structure of WT in our CO% calculations.
[p]Folding of the cys-free mutant was studied; we used structure of WT in our CO% calculations.
[q]Folding of the mutant protein F45W of bovine protein was studied; we used the structure of WT of human protein in our CO% calculations.
[r]There is only a strand-exchanged form of suc1 in PDB. We used the concatenation of fragment 2–88 of chain C and fragment 89–102 of chain A as a tentative structure of monomeric protein in our CO% calculations.

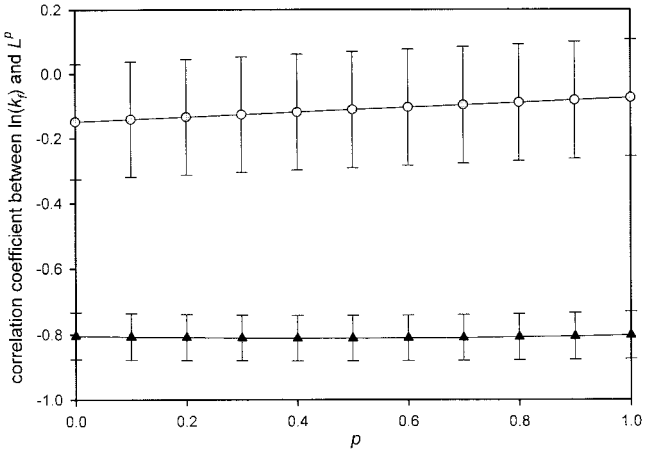



Fig. 3. Correlation between the logarithm of observed protein folding rate and scaling law $\ln(k_f) \sim L^p$ at various values of power $p$ [the case $p = 0$ corresponds to the scaling law $\ln(k_f) \sim \ln(L)$]. Errors are shown by vertical bars. Circles, two-state folders; triangles, three-state folders.

## REFERENCES


1. Jackson SE. How do small single-domain proteins fold? Fold Des 1998;3:R81–R91.
2. Fersht AR. Kinetics of protein folding. In: Hadler GL, editor. Structure and mechanism in protein science. New York: W.H. Freeman & Co.; 1999. p 540–572.
3. Fersht AR. Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism. Proc Natl Acad Sci USA 2000;97:1525–1529.
4. Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, Hofrichter J. Fast kinetics and mechanisms in protein folding. Annu Rev Biophys Biomol Struct 2000;29:327–359.
5. Plaxco KW, Simons KT, Baker D. Contact order, transition state placement and the refolding rates of single domain proteins. J Mol Biol 1998;277:985–994.
6. Thirumalai D. From minimal models to real proteins: time scales for protein folding kinetics. J Phys 1995;5:1457–1467.
7. Gutin AM, Abkevich VI, Shakhnovich EI. Chain length scaling of protein folding time. Phys Rev Lett 1996;77:5433–5436.
8. Finkelstein AV, Badretdinov AY. Rate of protein folding near the point of thermodynamic equilibrium between the coil and the most stable chain fold. Fold Des 1997;2:115–121.
9. Galzitskaya OV, Ivankov DN, Finkelstein AV. Folding nuclei in proteins. FEBS Lett 2001;489:113–118.
10. Ivankov DN, Garbuzynskiy SO, Alm E, Plaxco KW, Baker D, Finkelstein AV. Contact order revisited: influence of protein size on the folding rate. Protein Sci. Submitted.
11. Plaxco KW, Simons KT, Ruczinski I, Baker D. Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. Biochemistry 2000;39:11177–11183.
12. Zana R. On the rate-determining step for helix propagation in the helix-coil transition of polypeptides in solution. Biopolymers 1975; 14:2425–2428.
13. Gilmanshin R, Williams S, Callender RH, Woodruff WH, Dyer RB. Fast events in protein folding: relaxation dynamics of secondary and tertiary structure in native apomyoglobin. Proc Natl Acad Sci USA 1997;94:3709–3713.
14. Eaton WA, Munoz V, Thompson PA, Henry ER, Hofrichter J. Kinetics and dynamics of loops, α-helices, β-hairpins, and fast-folding proteins. Acc Chem Res 1998;31:745–753.
15. Spector S, Raleigh DP. Submillisecond folding of the peripheral subunit-binding domain. J Mol Biol 1999;293:763–768.
16. Kragelund BB, Robinson CV, Knudsen J, Dobson CM, Poulsen FM. Folding of a four-helix bundle: studies of acyl-coenzyme A binding protein. Biochemistry 1995;34:7217–7224.
17. Wittung-Stafshede P, Lee JC, Winkler JR, Gray HB. Cytochrome b562 folding triggered by electron transfer: approaching the speed limit for formation of a four-helix-bundle protein. Proc Natl Acad Sci USA 1999;96:6587–6590.
18. Ferguson N, Capaldi AP, James R, Kleanthous C, Radford SE. Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. J Mol Biol 1999;286: 1597–1608.
19. Burton RE, Huang GS, Daugherty MA, Fullbright PW, Oas TG. Microsecond protein folding through a compact transition state. J Mol Biol 1996;263:311–322.
20. Plaxco KW, Spitzfaden C, Campbell ID, Dobson CM. A comparison of the folding kinetics and thermodynamics of two homologous fibronectin type III modules. J Mol Biol 1997;270:763–770.
21. Clarke J, Cota E, Fowler SB, Hamill SJ. Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway. Struct Fold Des 1999;7:1145–1153.
22. Clarke J, Hamill SJ, Johnson CM. Folding and stability of a fibronectin type III domain of human tenascin. J Mol Biol 1997;270: 771–778.
23. Viguera AR, Serrano L, Wilmanns M. Different folding transition states may result in the same native structure. Nat Struct Biol 1996;3:874–880.
24. Grantcharova VP, Baker D. Folding dynamics of the src SH3 domain. Biochemistry 1997;36:15685–15692.
25. Guijarro JI, Morton CJ, Plaxco KW, Campbell ID, Dobson CM. Folding kinetics of the SH3 domain of PI3 kinase by real-time NMR combined with optical spectroscopy. J Mol Biol 1998;276:657–667.
26. Plaxco KW, Guijarro JI, Morton CJ, Pitkeathly M, Campbell ID, Dobson CM. The folding kinetics and thermodynamics of the Fyn-SH3 domain. Biochemistry 1998;37:2529–2537.

27. Schindler T, Herrler M, Marahiel MA, Schmid FX. Extremely rapid protein folding in the absence of intermediates. Nat Struct Biol 1995;2:663–673.
28. Perl D, Welker C, Schindler T, Schroder K, Marahiel MA, Jaenicke R, Schmid FX. Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. Nat Struct Biol 1998;5:229–235.
29. Reid KL, Rodriguez HM, Hillier BJ, Gregoret LM. Stability and folding properties of a model beta-sheet protein, *Escherichia coli* CspA. Protein Sci 1998;7:470–479.
30. Ikura T, Hayano T, Takahashi N, Kuwajima K. Fast folding of *Escherichia coli* cyclophilin A: a hypothesis of a unique hydrophobic core with a phenylalanine cluster. J Mol Biol 2000;297:791–802.
31. Guerois R, Serrano L. The SH3-fold family: experimental evidence and prediction of variations in the folding pathways. J Mol Biol 2000;304:967–982.
32. Kim DE, Fisher C, Baker D. A breakdown of symmetry in the folding transition state of protein L. J Mol Biol 2000;298:971–984.
33. McCallister EL, Alm E, Baker D. Critical role of beta-hairpin formation in protein G folding. Nat Struct Biol 2000;7:669–673.
34. Main ER, Fulton KF, Jackson SE. Folding pathway of FKBP12 and characterisation of the transition state. J Mol Biol 1999;291: 429–444.
35. Jackson SE, Fersht AR. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. Biochemistry 1991;30:10428–10435.
36. Villegas V, Azuaga A, Catasus L, Reverter D, Mateo PL, Aviles FX, Serrano L. Evidence for a two-state transition in the folding process of the activation domain of human procarboxypeptidase A2. Biochemistry 1995;34:15105–15110.
37. Silow M, Oliverberg M. High-energy channeling in protein folding. Biochemistry 1997;36:7633–7637.
38. Van Nuland NA, Chiti F, Taddei N, Raugei G, Ramponi G, Dobson CM. Slow folding of muscle acylphosphatase in the absence of intermediates. J Mol Biol 1998;283:883–891.
39. Otzen DE, Oliveberg M. Salt-induced detour through compact regions of the protein folding landscape. Proc Natl Acad Sci USA 1999;96:11746–11751.
40. Van Nuland NA, Meijberg W, Warner J, Forge V, Scheek RM, Robillard GT, Dobson CM. Slow cooperative folding of a small globular protein HPr. Biochemistry 1998;37:622–637.
41. Kuhlman B, Luisi DL, Evans PA, Raleigh DP. Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9. J Mol Biol 1998;284:1661–1670.
42. Choe SE, Matsudaira PT, Osterhout J, Wagner G, Shakhnovich EI. Folding kinetics of villin 14T, a protein domain with a central beta-sheet and two hydrophobic cores. Biochemistry 1998;37: 14508–14518.
43. Cavagnero S, Dyson HJ, Wright PE. Effect of H helix destabilizing mutations on the kinetic and equilibrium folding of apomyoglobin. J Mol Biol 1999;285:269–282.
44. Laurents DV, Corrales S, Elias-Arnanz M, Sevilla P, Rico M, Padmanabhan S. Folding kinetics of phage 434 Cro protein. Biochemistry 2000;39:13963–13973.
45. Tang KS, Guralnick BJ, Wang WK, Fersht AR, Itzhaki LS. Stability and folding of the tumour suppressor protein p16. J Mol Biol 1999;285:1869–1886.
46. Fowler SB, Clarke J. Mapping the folding pathway of an immunoglobulin domain. Structural detail from phi value analysis and movement of the transition state. Structure (Camb) 2001;9:355–366.
47. Parker MJ, Dempsey CE, Lorch M, Clarke AR. Acquisition of native beta-strand topology during the rapid collapse phase of protein folding. Biochemistry 1997;36:13396–13405.
48. Cota E, Clarke J. Folding of beta-sandwich proteins: three-state transition of a fibronectin type III module. Protein Sci 2000;9:112–120.
49. Burns LL, Dalessio PM, Ropson IJ. Folding mechanism of three structurally similar beta-sheet proteins. Proteins 1998;33:107–118.
50. Dalessio PM, Ropson IJ. Beta-sheet proteins with nearly identical structures have different folding intermediates. Biochemistry 2000;39:860–871.
51. Ogasahara K, Yutani K. Unfolding–refolding kinetics of the tryptophan synthase alpha subunit by CD and fluorescence measurements. J Mol Biol 1994;236:1227–1240.
52. Golbik R, Zahn R, Harding SE, Fersht AR. Thermodynamic stability and folding of GroEL minichaperones. J Mol Biol 1998;276: 505–515.
53. Schreiber G, Fersht AR. The refolding of *cis-* and *trans*-peptidylprolyl isomers of barstar. Biochemistry 1993;32:11195–11203.
54. Munoz V, Lopez EM, Jager M, Serrano L. Kinetic characterization of the chemotactic protein from *Escherichia coli*, Che Y. Kinetic analysis of the inverse hydrophobic effect. Biochemistry 1994;33: 5858–5866.
55. Parker MJ, Marqusee S. The cooperativity of burst phase reactions explored. J Mol Biol 1999;293:1195–1210.
56. Jennings PA, Finn BE, Jones BE, Matthews CR. A reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli:* verification and refinement of a four-channel model. Biochemistry 1993;32:3783–3789.
57. Goldberg ME, Semisotnov GV, Friguet B, Kuwajima K, Ptitsyn OB, Sugai S. An early immunoreactive folding intermediate of the tryptophan synthease beta 2 subunit is a "molten globule." FEBS Lett 1990;263:51–56.
58. Parker MJ, Spencer J, Clarke AR. An integrated kinetic analysis of intermediates and transition states in protein folding reactions. J Mol Biol 1995;253:771–786.
59. Parker MJ, Sessions RB, Badcoe IG, Clarke AR. The development of tertiary interactions during the folding of a large protein. Fold Des 1996;1:145–156.
60. Matouschek A, Kellis JT Jr, Serrano L, Bycroft M, Fersht AR. Transient folding intermediates characterized by protein engineering. Nature 1990;346:440–445.
61. Khorasanizadeh S, Peters ID, Roder H. Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. Nat Struct Biol 1996;3:193–205.
62. Schymkowitz JW, Rousseau F, Irvine LR, Itzhaki LS. The folding pathway of the cell-cycle regulatory protein p13suc1: clues for the mechanism of domain swapping. Struct Fold Des 2000;8:89–100.
63. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank. A computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.