Evidence for a molten globule state as a general intermediate in protein folding

O.B. Ptitsyn, R.H. Pain+, G.V. Semisotnov, E. Zerovnik+ and O.I. Razgulyaev

Institute of Protein Research, Academy of Sciences of the USSR, 142292 Pushchino, Moscow Region, USSR and *+Department of Biochemistry and Genetics, University of Newcastle upon Tyne NE2 4HH, England*

Received 28 November 1989

The folding of globular proteins occurs through intermediate states whose characterisation provides information about the mechanism of folding. A major class of intermediate states is the compact 'molten globule', whose characteristics have been studied intensively in those conditions in which it is stable (at acid pH, high temperatures and intermediate concentrations of strong denaturants). In studies involving bovine carbonic anhydrase, human α-lact-albumin, bovine β-lactoglobulin, yeast phosphoglycerate kinase, β-lactamase from *Staphylococcus aureus* and recombinant human interleukin 1β, we have demonstrated that a transient intermediate which accumulates during refolding is compact and has the properties of the 'molten globule' state. We show that it is formed within 0.1–0.2 s. These proteins belong to different structural types (β, α+β and α/β), with and without disulphide bridges and they include proteins with quite different times of complete folding (from seconds to decades of minutes). We propose that the formation of the transient molten globule state occurs early on the pathway of folding of all globular proteins.

Protein folding; Folding intermediate; Folding kinetics; Framework model

The stable molten globule states obtained under mild denaturing conditions have been shown to be consistently much more compact than state U by viscosity, sedimentation, diffuse X-ray scattering, quasielastic light scattering and urea gradient electrophoresis for CAB [4], βLase [5,6] and αLA [1,2,7]. It is almost as compact as state N [1–7] and has a pronounced secondary structure [1–4,6,8]. This secondary structure can be N-like and the molten globule may have some features of the N fold [9]. However, this state differs from state N by the absence of close packing throughout the molecule and by a substantial increase of fluctuations in side chains as well as of larger parts of the molecule [1–3]. In agreement with these data, the equilibrium molten globule states for βLase [10] and CAB both have Ves on FPLC gel exclusion that are intermediate between the Ves for N and U states. This permits the use of FPLC not only for the evaluation of the compactness of these states but also to monitor the kinetics of the formation of state N in refolding experiments [10].

Correspondence address: R.H. Pain, Department of Biochemistry and Genetics, University of Newcastle upon Tyne NE2 4HH, England

Abbreviations: GdmCl, guanidinium chloride; MES, 2[*N*-Morpholino]ethanesolphonic acid; near UV c.d., near ultraviolet circular dichroism; DTT, dithiothreitol; CAB, bovine carbonic anhydrase; αLA, α-lactalbumin; BLG, β-lactoglobulin; PGK, phosphoglycerate kinase; βLase, β lactamase; IL-1β, recombinant human interleukin 1β; Ve, elution volume; N, native state; U, fully unfolded state; ANS, 8-anilino-1-naphthalene-sulfonate; S, secondary structure; G, intermediate globular state

Fig.1 presents the results of the refolding of 3 proteins studied using FPLC. The elution profiles of βLase in the N and U states as well as the molten globule state stable at intermediate concentrations of urea, earlier termed the 'H-forms' [5], are shown in fig.1a. The H-forms of βLase and CAB (not shown) have closely similar elution volumes that are substantially larger and therefore considerably more compact than the fully unfolded forms.

During refolding of βLase and CAB from state U, using conditions under which state N is stable, aliquots applied to the FPLC column at times following initiation of refolding are eluted as two peaks. One of these peaks corresponds to state N and the other to the kinetical intermediate which has been termed state I by analogy with the compact intermediate state identified by urea gradient gel electrophoresis [12,23]. Fig.1b,c shows that Ve for state I lies between those for states N and H, and that therefore the kinetical intermediate is even more compact than the equilibrium one. Kinetical intermediates with the same elution volumes have also been observed for βLase and CAB in the cases where protein renaturation starts from state H rather than from state U.

Fig.1b,c shows that the rates of the I → N transition, measured by change of elution volume for all 3 proteins under the stated conditions, are slow enough to be monitored by FPLC techniques. It is important to note that the rates of formation of state N for these proteins are similar to the rates of regain of the molar ellipticity in the near UV region as well as of enzyme activity. In

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50 © 1990 Federation of European Biochemical Societies

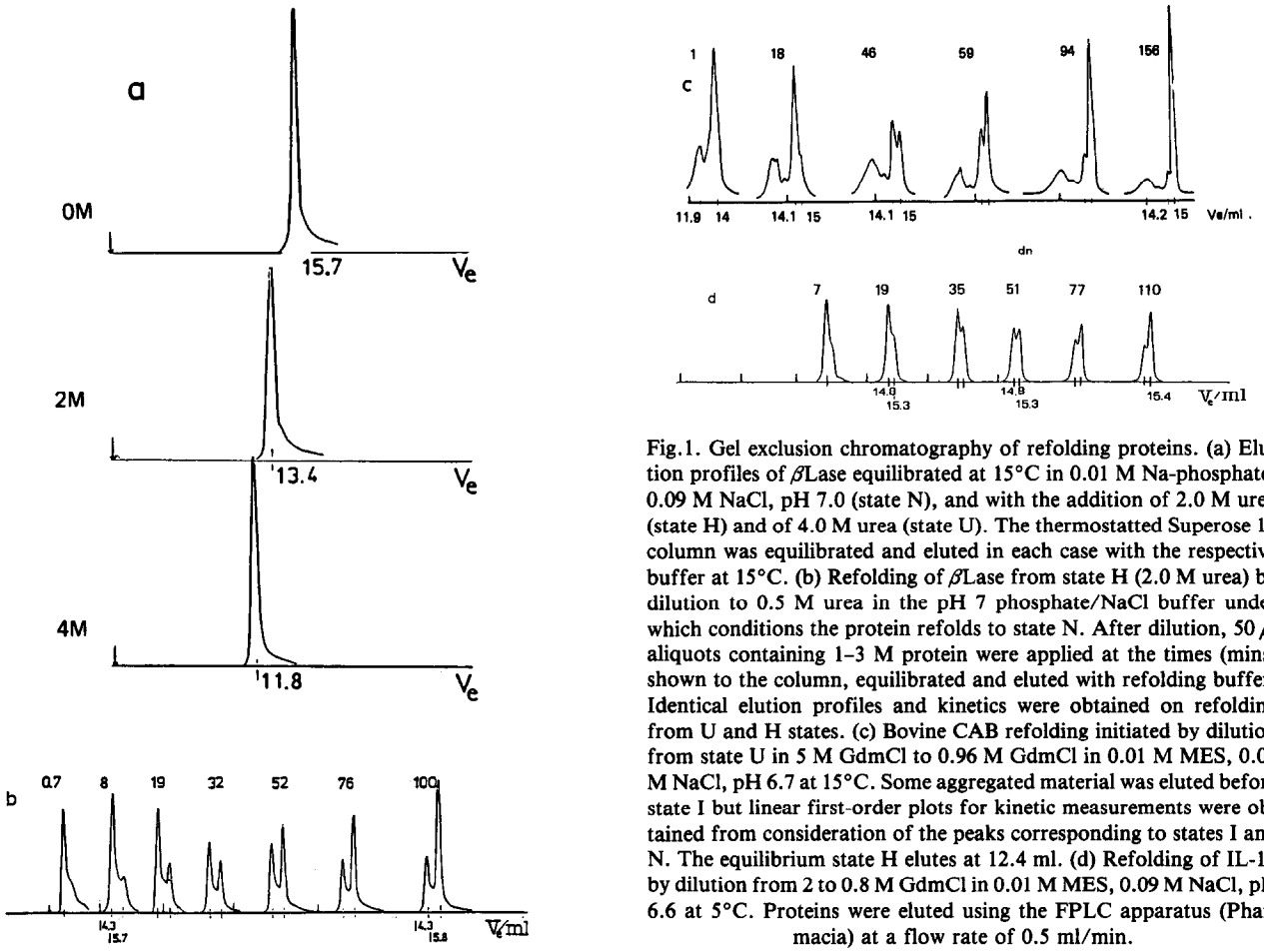

Fig.1. Gel exclusion chromatography of refolding proteins. (a) Elution profiles of βLase equilibrated at 15°C in 0.01 M Na-phosphate, 0.09 M NaCl, pH 7.0 (state N), and with the addition of 2.0 M urea (state H) and of 4.0 M urea (state U). The thermostatted Superose 12 column was equilibrated and eluted in each case with the respective buffer at 15°C. (b) Refolding of βLase from state H (2.0 M urea) by dilution to 0.5 M urea in the pH 7 phosphate/NaCl buffer under which conditions the protein refolds to state N. After dilution, 50 μl aliquots containing 1–3 M protein were applied at the times (mins) shown to the column, equilibrated and eluted with refolding buffer. Identical elution profiles and kinetics were obtained on refolding from U and H states. (c) Bovine CAB refolding initiated by dilution from state U in 5 M GdmCl to 0.96 M GdmCl in 0.01 M MES, 0.09 M NaCl, pH 6.7 at 15°C. Some aggregated material was eluted before state I but linear first-order plots for kinetic measurements were obtained from consideration of the peaks corresponding to states I and N. The equilibrium state H elutes at 12.4 ml. (d) Refolding of IL-1β by dilution from 2 to 0.8 M GdmCl in 0.01 M MES, 0.09 M NaCl, pH 6.6 at 5°C. Proteins were eluted using the FPLC apparatus (Pharmacia) at a flow rate of 0.5 ml/min.

these experiments and also when the protein in state U or state H is applied directly to the FPLC column equilibrated with refolding buffer, only state I and state N are seen in proportions depending on the column conditions. No peak corresponding to state H is found under any refolding conditions. Thus, the collapse from state U (or state H) into state I must occur within the first minute of refolding. Further, state H is not significantly populated during refolding.

Thus, the refolding of CAB and βLase involves the rapid collapse and accumulation of a state that is more compact than state H but more expanded than state N. That this type of folding intermediate is not restricted to proteins that exhibit a stable equilibrium type H intermediate is demonstrated by the folding of IL-1β. The reversible equilibrium unfolding of this protein is a two-state process [13] but the kinetics of its refolding show that it rapidly collapses to state I which accumulates and transforms more slowly to the native state (fig.1d).

Approximately 85% of the far UV ellipticity of CAB [14,15], βLase and IL-1β [13] is regained within the usual dead time for measuring ellipticity (< 1 min). The remainder of the native ellipticity is regained with the same rate constants as for near UV c.d., activity and native elution volume. The results obtained by stopped-flow techniques for ribonuclease [16], αLA [17–19], lysozyme [17], cytochrome *c* [20], βLG [20] and parvalbumin [21] show that secondary structure is formed early in protein folding before N structure formation. These results demonstrate that a substantial amount of N-like, secondary structure is present in the compact state I.

On the other hand, compact kinetical intermediates of CAB [14,15,22], βLase and IL-1β [13] have no near UV c.d.. Moreover the high-resolution NMR spectrum of state I, shown here for βLase (fig.2), demonstrates a dramatic reduction in the stable tertiary interactions present in state N. This is confirmed in the aromatic region also. The high field methyl resonances are particularly noticeable by their absence in state I in both βLase (fig.2) and CAB [22]. In addition, the activity of CAB [14,15,22] and βLase [12] is completely absent in state I.

It follows that the early compact kinetical intermediate in protein folding (state I) has all the properties of the molten globule state. The existence of the molten globule state as a kinetical intermediate was first described for CAB [15,22]. It is now established also for

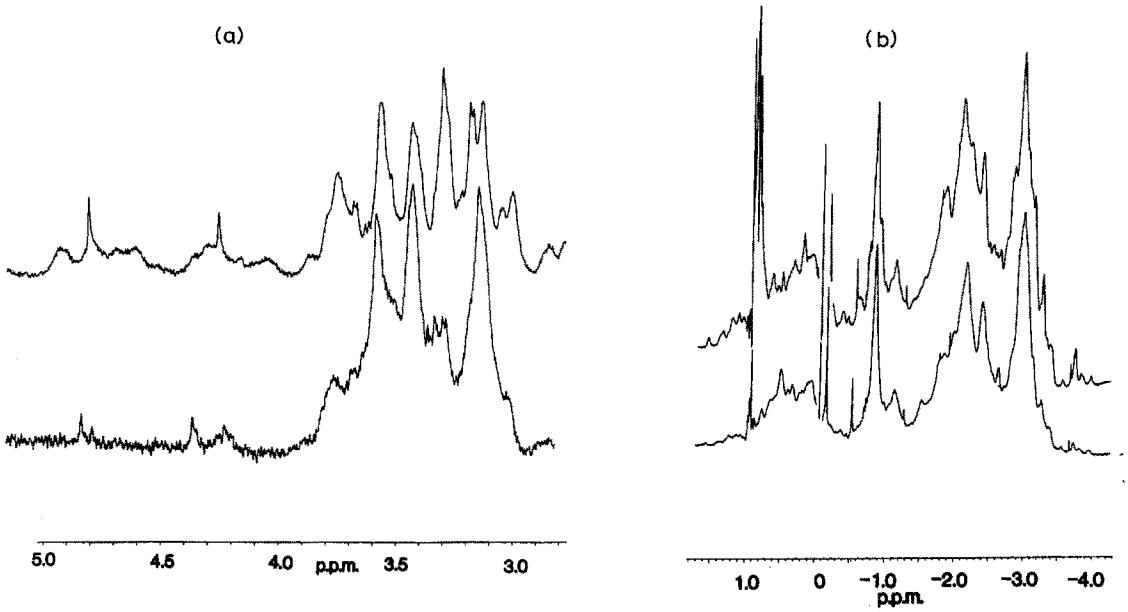

Fig.2. High resolution NMR spectra of the kinetic intermediate I state. Refolding of βLase was initiated by diluting a 45 mg/ml solution in 2.0 M urea, 0.01 M Na-phosphate, 0.09 M NaCl to 0.7 M urea in buffer at 0°C. The solution was immediately transferred to a cooled NMR tube and the spectrum measured at 15°C within 3–4 min. During this time, there was only a small accumulation of state N judged by gel exclusion and by near UV c.d. More than 90% enzyme activity was recovered on allowing refolding under these conditions to proceed to completion. Upper spectrum, N state; lower, I state. (a) Aromatic region; (b) aliphatic region.

βLase and IL-1β, as well as for αLA, βLG and PGK (see below).

Unfortunately, FPLC like all other non-optical methods cannot be used for investigating fast processes and therefore one cannot apply this method to measure the time of formation of the molten globule state. We have therefore used a hydrophobic probe ANS whose fluorescence sharply increases upon its binding to a protein molecule. It has been shown that the fluorescence of ANS in the presence of αLA and CAB [24] as well as of βLase is much stronger under conditions in which these proteins are in the molten globule state, as compared with their N and U states. We have suggested [22,24] that ANS binds much better to the I and H states because non-polar groups in the molten globule are more accessible to solvent than in state N. Fig.3 shows the changes of intensity of ANS fluorescence during the refolding of 5 proteins. The existence of maxima on these curves reflects the formation of the molten globule state and its subsequent slower transition to state N. ANS binding to a protein in the molten globule state occurs within the dead time of the stopped flow experiment (0.002–0.02 s) [24]. Therefore, the observed increase in the intensity of ANS fluorescence reflects formation of the molten globule I state, i.e. the formation of a non-rigid hydrophobic core. This point of view is confirmed by the kinetics of CAB refolding monitored by other methods.

The rate of increase of ANS fluorescence has been shown to be essentially the same as the rate of molecular compactization of CAB measured by another empirical method, the intensity of electron spin resonance signals of spin labels, as well as by the more direct method of energy transfer from tryptophan residues to dansyl labels randomly attached to the protein molecule [22]. On the other hand, the rate of decrease of the intensity of ANS fluorescence, reflecting its desorbtion from the protein, is close to the rates of formation of state N measured by the regain of N Ve (this paper), near UV c.d., enzyme activity and the slow phase of the regain of N absorbance [22].

Thus the data presented in fig.3 show that a kinetical intermediate identified as the compact molten globule state is formed within 0.1–0.2 s. They show also that not only CAB, βLase and IL-1β, but also αLA, βLG and PGK go through this kinetical intermediate during their folding to the native structures. These proteins belong to different structural types including β, α+β and α/β, both with and without disulphide bridges. They include proteins which exhibit a molten globule state at equilibrium in mild denaturing conditions and those which do not exhibit an equilibrium molten globule state. The rates of their complete refolding have been shown to differ by orders of magnitude. These data suggest that the transient molten globule I-type compact state occurs generally in the folding of globular proteins.

In 1973, one of us suggested [25] that protein folding involves at least 3 steps: (i) the formation of the native secondary structure in an unfolded chain; (ii) the collapse of these embryos into an intermediate compact state which is stabilized mainly by hydrophobic interac-

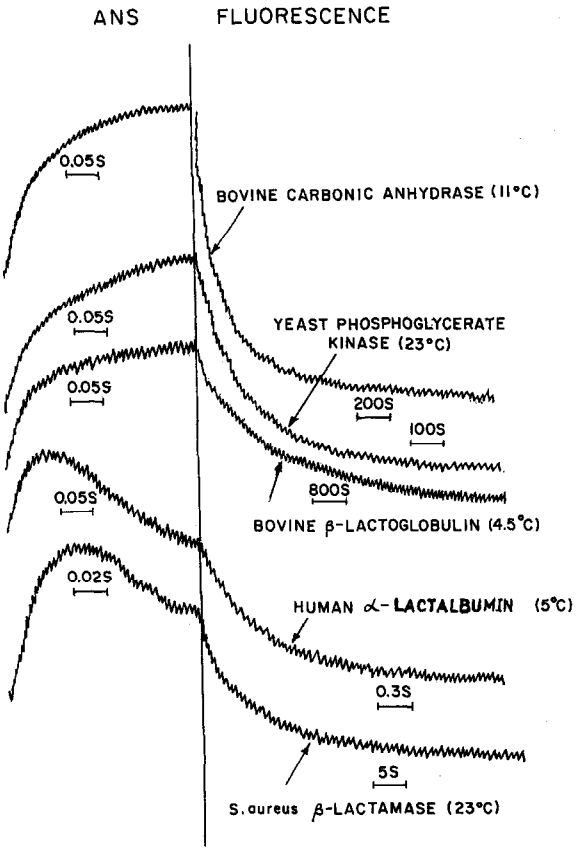

Fig.3. Refolding of globular proteins monitored by ANS fluorescence. Refolding from state U was initiated by 2- or 6-fold dilution of denaturant solution containing protein and ANS. Change of ANS fluorescence was monitored with a 'Durrum' stopped-flow apparatus (dead time ~ 2 ms) or with a stopped flow fluorescence spectrometer at 480 nm with excitation at 380 nm. Final concentration of protein was 5×10^{-6} M and of ANS 10^{-4} M. Refolding of bovine CAB from 8.0 to 4.0 M urea (similar results were obtained from 4.0 to 0.7 M GdmCl) in 0.05 M Tris-HCl, pH 8.0 at 23°C; yeast PGK from 2.7 to 1.35 M urea or 2.0 to 0.4 M GdmCl in 0.02 M Na-phosphate 0.002 M DTT, pH 7.0 at 23°C; β-LG from 4.0 to 0.07 M GdmCl in 0.1 M NaCl, pH 3.2 at 4.5°C; human αLA from 9.0 to 1.5 M urea in 0.05 M Tris-HCl, pH 8.0 at 23°C; βLase from 1.9 to 0.35 M GdmCl in 0.1 M Na-phosphate, pH 7.0 at 23°C.

tions; (iii) the local adjustment of this intermediate state to the tightly packed native structure with close van der Waals interactions. This scheme implies that both S and state G exhibit essential features present in state N. This scheme can thus be specified as $U \rightarrow S_N \rightarrow G_N \rightarrow N$.

The suggestion that native secondary structure is formed before the tertiary structure ($U \rightarrow S \rightarrow N$) has since been termed a 'framework model' [26]. Many authors (see e.g. [5,14–21]) have shown that secondary structure is regained during protein refolding much faster than its rigid tertiary structure. Recently it has been shown that amide hydrogen bonds corresponding to native secondary structures in ribonuclease A [27] and cytochrome *c* [28] are formed early in folding, thus directly confirming the framework model.

This paper presents evidence for the general existence of another kinetical intermediate, a transient molten globule state which corresponds to state G predicted in 1973 [25]. Moreover, there is some evidence which suggests that the formation of secondary structure precedes the formation of the molten globule state ($U \rightarrow S \rightarrow G \rightarrow N$).

Almost all the far UV molar ellipticity of ferricytochrome *c* [20], βLG [20] and parvalbumin [21] and half of the molar ellipticity of αLA [19] are regained within the dead time of stopped flow circular dichroism experiments (0.02 s). On the other hand, ANS binding, which reflects the formation of the molten globule state, reaches its maximum in 0.1–0.2 s, i.e. an order of magnitude slower. These results provide a strong indication that secondary structure is formed before the molten globule state I.

In another, yet earlier, suggestion [29], it was proposed that secondary structure is formed before tertiary structure and also that folding proceeds through a compact, mainly non-ordered state. However, in this model, it was suggested that helices form within a non-polar globule and, by implication, after collapse. It seems likely now that collapse is initiated by interactions between elements of secondary structure as predicted [25]. It remains to be seen whether the elements of secondary structure that form more slowly [19,28] form within the molten globule or by a sequential collapse onto the globule. We can conclude, however, that many of the important features of protein folding which have been suggested earlier [25,29] are now confirmed by experiment and that we have now demonstrated the existence of at least one general kinetic intermediate of protein folding – the molten globule state.

REFERENCES

[1] Dolgikh, D.A., Gilmanshin, R.I., Braznikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venjaminov, S. Yu. and Ptitsyn, O.B. (1981) FEBS Lett. 136, 311–315.
[2] Dolgikh, D.A., Abaturov, L.V., Bolotina, I.A., Braznikov, E.V., Bushuyev, V.N., Bychkova, V.E., Gilmanshin, R.I., Lebedev, Yu.O., Semisotnov, G.V., Tiktopulo, E.I. and Ptitsyn, O.B. (1985) Eur. Biophys. J. 13, 109–121.
[3] Ptitsyn, O.B. (1987) J. Protein Chem. 6, 277–293.
[4] Wong, K.-P. and Hamlin, L.M. (1974) Biochemistry 13, 2678–2683.
[5] Robson, B. and Pain, R.H. (1973) in: Conformation of Biological Molecules and Polymers (Bergmann, E.D. and Pullman, A. eds). pp. 161–172, Academic, London; (1976) Biochem. J. 15, 331–344.
[6] Carrey, E.A. and Pain, R.H. (1978) Biochim. Biophys. Acta 533, 12–22.
[7] Gast, H., Zirver, D., Welfe, H., Bychkova, V.E. and Ptitsyn, O.B. (1986) Int. J. Biol. Macromol. 8, 231–236.
[8] Kuwajima, K., Nitta, H., Yonegama, M. and Sugai, S. (1976) J. Mol. Biol. 106, 359–373.
[9] Baum, J., Dobson, C.M., Evans, P.A. and Hanley, C. (1989) Biochemistry 28, 7–13.
[10] Zerovnik, E. and Pain, R.H. (1987) Protein Eng. 1, 248.
[11] Creighton, T.E. and Pain, R.H. (1980) J. Mol. Biol. 137, 431–436.

[12] Mitchinson, C. and Pain, R.H. (1987) J. Mol. Biol. 184, 331–342.
[13] Craig, S., Schmeissner, U., Wingfield, P. and Pain, R.H. (1987) Biochemistry 26, 3570–3576.
[14] McCoy, L.F., Rowe, E.S. and Wong, K.-P. (1980) Biochemistry 19, 4738–4743.
[15] Dolgikh, D.A., Kolomiets, A.P., Bolotina, I.A. and Ptitsyn, O.B. (1984) FEBS Lett. 165, 88–92.
[16] Labhardt, A.M., (1984) Proc. Natl. Acad. Sci. USA 81, 7674–7678.
[17] Kuwajima, K., Hiraoka, Y., Ikeguchi, M. and Sugai, S. (1985) Biochemistry 24, 874–891.
[18] Ikeguchi, M., Kuwajima, K., Mitani, M. and Sugai, S. (1986) Biochemistry 25, 6965–6972.
[19] Gilmanshin, R.I. and Ptitsyn, O.B. (1987) FEBS Lett. 223, 327–329.
[20] Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S. and Nagamura, T. (1987) FEBS Lett. 221, 115–118.
[21] Kuwajima, K., Sakuroaka, A., Fueki, S., Yoneyama, M., and Sugai, S. (1988) Biochemistry 27, 7419–7428.
[22] Semisotnov, G.V., Rodionova, N.A., Kutyshenko, V.P., Ebert, B., Blanck, J. and Ptitsyn, O.B. (1987) FEBS Lett. 224, 9–13.
[23] Thomas, R.M., Feeney, J., Nicholson, R.B., Pain, R.H. and Roberts, G.C.K. (1983) Biochem. J. 215, 525–529.
[24] Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripasj, A.K. and Gilmanshin, R.I., Biopolymers, submitted.
[25] Ptitsyn, O.B. (1973) Dokl. Akad. Nauk SSSR 213, 473–475.
[26] Kim, P.S. and Baldwin, R.L. (1982) Annu. Rev. Biochem. 51, 459–489.
[27] Udgaonkar, J.B. and Baldwin, R.L. (1988) Nature 335, 694–699.
[28] Roder, H.R., Elove, G.A. and Englander, S.W. (1988) Nature 335, 700–704.
[29] Robson, B. and Pain, R.H. (1971) J. Mol. Biol. 58, 237–259.