ELSEVIER

FEBS Letters 341 (1994) 15–18

FEBS 13726

FEBS LETTERS

The molten globule is a third thermodynamical state of protein molecules

O.B. Ptitsyn[a,b,*], V.N. Uversky[a]

[a]*Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142292, Russian Federation*
[b]*NIH, National Cancer Institute, Laboratory of Mathematical Biology, WSC, 6010 Executive Blvd., Bethesda, MD 20892, USA*

Received 14 January 1994

Abstract

Analysis of published data on conformational transitions in relatively small proteins shows that the slopes of these transitions are proportional to the protein molecular weight. It is true both for transitions from the native (N) to the unfolded (U) states (when protein denaturation is coupled to its unfolding) and for transitions from the native to the molten globule (MG) states and from the molten globule to the unfolded state (when protein denaturation is decoupled from protein unfolding). This is precisely the behaviour predicted by thermodynamics for first order phase transitions ('all-or-none' transitions) in small systems. It follows that N→U, N→MG and MG→U transitions in proteins are all of the 'all-or-none' type. Thus the molten globule state of protein molecules is separated by an 'all-or-none' transition both from the native and the unfolded state, i.e. the molten globule state is a third thermodynamic state of protein molecules in addition to the two previously established states – the native and the unfolded.

Key words: Protein denaturation; Protein folding; Molten globule state

1. Introduction

The molten globule (MG) is an equilibrium state of many proteins under mild denaturing conditions and a universal kinetic intermediate of protein folding [1–7]. Its properties are intermediate between those of native (N) and unfolded (U) protein molecules: it is compact [8,9] and has a number of important features of the native secondary structure and of the native overall architecture [10–14] but has only traces of the detailed native tertiary structure. These properties mean that the molten globule is structually quite different both from the native protein and from the unfolded chain. The question therefore arises as to whether these three states of protein molecules are different also from the thermodynamic point of view. To answer this question we have studied urea- and GdmCl-induced denaturation of proteins since this type of denaturation normally leads to their more or less complete unfolding (see e.g. [1,2,15]). This unfolding can occur either in one step (an N→U transition, see [15,16]), or in two separate steps (N→MG and MG→U transitions, see [1–3,6]) which take place at different concentrations of denaturing agents, although they can partially overlap. This provides a good opportunity to study the N→MG and MG→U transitions separately. We show in this paper that for all small proteins studied to date both the N→MG and MG→U transitions belong to 'all-or-none' type. This means that the molten globule represents a third thermodynamical state of protein molecules in addition to two previously known states, the native and unfolded.

2. Materials and methods

Experimental curves relating the degree of urea- and GdmCl-induced protein unfolding to the urea and GdmCl activity were collected from the literature for proteins with molecular weights $M < 30$ kDa. Degrees of unfolding were defined by the equations $\Theta_{N\to U} = (X - X_N)/(X_U - X_N)$, $\Theta_{N\to MG} = (X - X_N)/(X_{MG} - X_N)$ and $\Theta_{MG\to U} = (X - X_{MG})/(X_U - X_{MG})$, respectively, where X is the parameter by which unfolding is monitored, while X_N, X_{MG} and X_U are the values of this parameter in the native, molten globule and unfolded states, respectively. Near UV circular dichroism, ^{1}H-NMR and biochemical activity data have been used as these parameters for N→MG transitions. Far UV circular dichroism, fluorescence, ^{1}H-NMR, absorbance, chromatography and viscosity data have been used for MG→U transitions, while data of all types have been used for N→U transitions. Activities of urea and GdmCl were calculated from their concentrations using empirical equations [17,18]. Values of $\Delta\nu^{eff}$ were obtained from the equation $\Delta\nu^{eff} = 4/(\ln a_2 - \ln a_1)$ which is practically equivalent to Equation (3). Here a_1 and a_2 are the activities of urea or GdmCl at the beginning and at the end of transitions, defined as the intersection of the tangent to the transition curve at the middle of the transition with the lines $\Theta = 0$ and $\Theta = 1$. The full set of proteins used was as follows:

GdmCl-induced N→U transitions (Fig. 1A): epidermal growth factor and its fragment, chicken pancreatic peptide, coat protein of f_d phage, bovine pancreatic trypsin inhibitor, thermolysin and its fragments, carp parvalbumin, cytochrome *c*, thioredoxin, ribonucleases T1 and A, bovine and human α-lactalbumins, lysozyme, interleukin 2, myoglobins (sperm whale, alligator, horse), β-lactoglobulin, T4 lysozyme, γ-crystallin, human growth hormone, glutathione S-transferase, phosphoribosyl anthranilate isomerase and its mutant, Trp synthase (α-subunit), alanine racemase.

*Corresponding author. Fax: (1) (301) 402 4724;
E-mail: <ptitsyn@lmmb.nci.nih.gov>.

Abbreviations: GdmCl, guanidinium chloride; N, native state; MG, molten globule state; U, unfolded state; UV, ultraviolet.

SSDI 0014-5793(94)00110-H

Urea-induced N→U transitions (Fig. 1A): proinsulin and its peptide, insulin, ferredoxin, caldinin, γII-crystallin, thioredoxin, Trp aporepressor, barnase, ribonucleases T1 and A, lysozyme, *Staphylococcal* nuclease A, bovine trypsinogen, myoglobins (sperm whale and horse), dihydropholate reductase, β-lactoglobulin, T4 lysozyme, bovine growth hormone, glutathione S-transferase, penicillin G acylase (α-peptide), α-chymotrypsin, bacteriorodopsin, Trp synthase (α-subunit).
GdmCl-induced N→MG→U transitions (Fig. 1A,B,C): bovine and human α-lactalbumins, bovine and reduced human growth hormones, bovine and human carbonic anhydrases B, rhodanese.
Urea-induced N→MG→U transitions (Fig. 1A,B,C): β-lactamases from *Staphylococcus aureus* and *Bacillus cereus*, bovine carbonic anhydrase B and rhodanese.
GdmCl-induced MG→U transition (Fig. 1C): carp parvalbumin, ribonuclease A, *Staphylococcal* nuclease, β-lactoglobulin, bovine and human carbonic anhydrases B.
Urea-induced MG→U transitions (Fig. 1C): human α-lactalbumin, bovine growth hormone and bovine carbonic anhydrase B.

Full details and a complete set of references will be published elsewhere.

3. Results

Consider protein molecules which undergo a solvent-induced 'all-or-none' transition between two states A and B. This means that the addition of some transition-inducing component ('denaturing agent') to a solvent provokes the A→B transition in each protein molecule as a whole. The slope of this transition can be characterized by the dependence of the equilibrium constant between the A and B states

$$K = \Theta/(1 - \Theta) \quad (1)$$

on the activity of the denaturing agent (where Θ is the fraction of protein molecules in one of these states). Strict thermodynamical treatment shows that at small concentrations of a protein

$$(\delta \ln K/\partial \ln a)_{a=a_t} = \Delta\nu^{eff}(a_t) \quad (2)$$

where a is the activity of the denaturing agent and a_t its activity at the middle of the transition. Strictly speaking $\Delta\nu^{eff}(a_t)$ is the number of molecules of denaturing agent which must be added (or removed) to keep its activity constant following the A→B transition of one protein molecule. For a simplified model in which all molecules of denaturing agent are divided between 'free' and 'bound' [19], $\Delta\nu^{eff}$ is the difference in the number of denaturant molecules 'bound' to one protein molecule in its A- and B-states.

It follows from (1) and (2) that

$$\Delta\nu^{eff}(a_t) = 4\, a_t\, (\partial\Theta/\partial a)_{a_t} \quad (3)$$

which emphasizes the fact that $\Delta\nu^{eff}$ at the middle point of the transition is directly proportional to the slope of the transition at this point.

A general method to check whether the given transition is of the 'all-or-none' type is to compare effective thermodynamic values obtained from the slope of the transition with those directly measured per mol of protein. A well-known example of this approach is the comparison of the effective enthalpy of a temperature-induced transition (measured from the temperature dependence of Θ) with its value measured directly by microcalorimetry [16]. For solvent-induced transitions this approach can only be used for pH-induced changes since in this case $\Delta\nu^{eff}(a_t)$ can be directly measured by potentiometric titration (see e.g. [20]), but not for denaturant-induced changes.

Here we study the stages of proteins unfolding induced by urea or GdmCl and use a different approach to decide whether these transitions belong to the 'all-or-none' type.

It has been demonstrated that the slope of the 'all-or-none' transition in a small system must be proportional to the number of units in this system [21]. This means that the slope of an 'all-or-none' transition in a macromolecule as a whole must be proportional to its molecular weight. However, parameters which refer to just a part of a macromolecule (a 'cooperative unit' or 'domain') will be independent of molecular weight if the cooperative unit is smaller than the whole macromolecule. It follows that $\Delta\nu^{eff}$ will increase with protein molecular weight only when the whole molecule is the cooperative unit, i.e. when urea- or GdmCl-induced transitions belong to the 'all-or-none' type.

Fig. 1 shows molecular weight dependences of $\Delta\nu^{eff}$ for urea- and GdmCl-induced N→U (Fig. 1A), N→MG (Fig. 1B) and MG→U (Fig. 1C) transitions. Values of $\Delta\nu^{eff}$ were calculated using equation (3) and using virtually all the available data in the literature for relatively small proteins. Large proteins are not considered so as to avoid complications relating to their domain structure. The figures show that in all three cases the values $\Delta\nu^{eff}$ depend linearly on molecular weight and can be approximated by the equations:

$$\Delta\nu^{eff}_{N\to U} = 0.63\, M + 1.59 \quad (4)$$

$$\Delta\nu^{eff}_{N\to MG} = 0.35\, M + 0.19 \quad (5)$$

$$\Delta\nu^{eff}_{MG\to U} = 0.26\, M + 1.05 \quad (6)$$

(M being a molecular weight in kDa).

Since each $\Delta\nu^{eff}$ value is measured at the middle of the corresponding transition region, it may depends on the activity of urea or GdmCl at this point. We have therefore checked whether the linear molecular weight dependence of $\Delta\nu^{eff}$ simply reflects a systematic dependence of $\Delta\nu^{eff}$ on a_t together with a corresponding dependence of a_t on M. For this we used the values of $\Delta\nu^{eff}$ for N→U transitions (for which many data are available) but did not observe any systematic dependence in either relashionship (data not shown). In addition, we have separately plotted $\Delta\nu^{eff}(M)$ curves for groups of proteins have

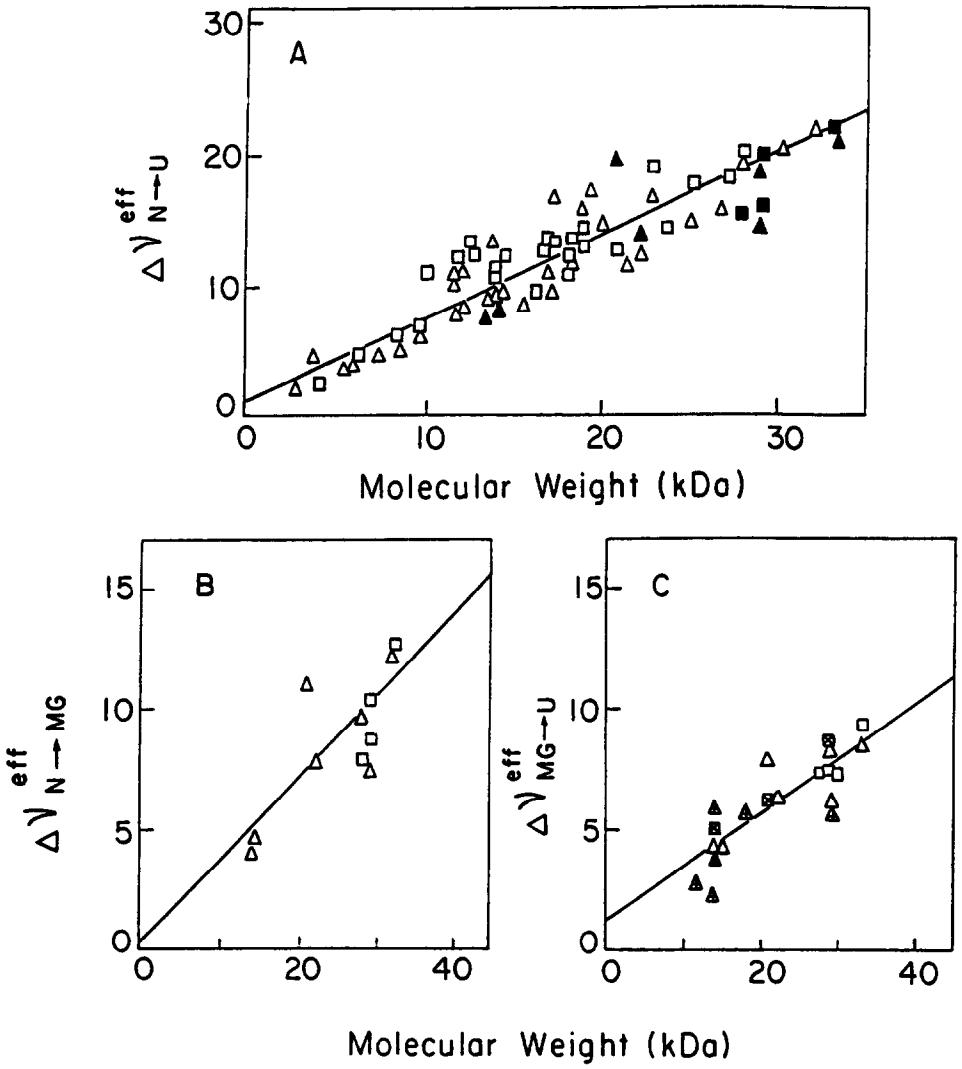

Fig. 1. Molecular weight dependences of the slopes of urea- and GdmCl-induced transitions between: A, native and unfolded state (N→U); B, native and molten globule state (N→MG); and C, molten globule and unfolded states (MG→U) of protein molecules. The data for about 60 proteins are included having molecular weights of no more than 30 kDa. The slopes were characterized by the numbers of urea or GdmCl molecules $\Delta\nu^{eff}$, which have to be added to preserve the same activity following the transition of one protein molecule from one state into the other. $\Delta\nu^{eff}$ values were calculated using Equation (3) from the slopes of unfolding curves at the middle of the transition. Triangles and squares refer to GdmCl- and urea-induced unfolding of proteins, respectively. Open symbols in Fig. 1A correspond to $\Delta\nu^{eff}{}_{N\to U}$ for proteins which follow a single sigmoidal curve, while black symbols are the sums $\Delta\nu^{eff}{}_{N\to MG} + \Delta\nu^{eff}{}_{MG\to U}$ for proteins for which denaturation takes place via the molten globule state. Crossed symbols in Fig. 1C refer to the unfolding of molten globule states stabilized by low pH, while other symbols refer to the unfolding of the molten globule states stabilized by intermediate concentrations of urea or GdmCl.

similar a_t values. In all cases the same linear dependence was seen as shown for the full data set in Fig. 1A.

It follows that the slopes at the mid-points of K against a curves for urea- and GdmCl-induced N→U, N→MG and MG→U-transitions in relatively small globular proteins are almost proportional to their molecular weights. The only possible explanation of this behaviour is that urea- or GdmCl-induced N→U, N→MG and MG→U transitions in relatively small globular proteins are of the 'all-or-none' type.

This conclusion is confirmed by the additivity of the $\Delta\nu^{eff}$ values for these transitions. Equations (4), (5) and (6) have been obtained from quite independent experiments with different proteins and under a wide range of conditions. It is striking therefore that the sum of equations (5) and (6) gives

$$\Delta\nu^{eff}{}_{N\to MG} + \Delta\nu^{eff}{}_{MG\to U} = 0.61\ M + 1.24 \qquad (7)$$

which is very close to equation (4). For those proteins, for which both N→MG and MG→U transitions have been studied, the sums of their $\Delta\nu^{eff}{}_{N\to MG}$ and $\Delta\nu^{eff}{}_{MG\to U}$ fit the same molecular weight dependence as obtained for proteins with a single N→U transition (see Fig. 1A).

Recalculation of equations (4)–(6) in terms of the number of residues in the proteins $n = M/M_0$ (taking the

molecular weight of a residue as $M_0 = 110$) leads to the conclusion that $\Delta\nu^{eff}$ is approximately equal to 0.07 n, 0.04 n and 0.03 n for N→U, N→MG and MG→U transitions, respectively. This means that one molecule of denaturing agent should be added for every 15, 25 and 35 residues, respectively, to keep the activity of urea or GdmCl constant for these transitions. In terms of the binding model [19] this corresponds to the additional binding of one urea or GdmCl molecule per 15, 25 and 35 residues, respectively.

It is striking that these values are the same for both urea and GdmCl as $\Delta\nu^{eff}$ for both denaturants fit the same linear dependences in Fig. 1. This may be important for further understanding of the mechanism of urea and GdmCl denaturation of proteins.

4. Discussion

The 'all-or-none' character of solvent-induced N→U transitions was suggested a long time ago [15,16] and confirmed more recently for three protein by size-exclusion chromatography [22–24]. Fig. 1A shows that this is indeed the case for all 60 relatively small proteins for which data are available. The key point demonstrated here is that N→MG and MG→U transitions in relatively small proteins are also of the 'all-or-none' type. Since the existence of 'all-or-none' transitions between MG and U states of proteins is rather unexpected, we have demonstrated this fact for two proteins by a direct method, viz: the observation of a bimodal distribution of molecular dimensions in the transition region [25]. The analysis presented here demonstrates that the molten globule state is quite different from the native and unfolded states not only structurally but also thermodynamically and therefore represents a third thermodynamic state of protein molecules. We propose that the existence of two 'all-or-none' transitions (N→MG and MG→U) reflects the cooperative formation (or breakdown) of two levels of protein organisation – the 'tertiary structure' (i.e. the detailed 3D structure of the protein molecule at atomic resolution) for the N→MG transition and the 'tertiary fold' (i.e. the crude mutual positions of α- and β-regions) for the MG→U transition.

The existence of the 'all-or-none' transitions between molten globules and unfolded chains can be of biological importance. It has been shown previously that the native structure of relatively small proteins can be destroyed only by an 'all-or-none' transition [16] and this preserves the active centre from thermal fluctuations [6,26] and ensure the physiological activity of native proteins. In addition, it has been suggested [27] and confirmed experimentally (see e.g. [28–31]) that the molten globule state can play a substantial role in a number of physiological processes; see [32] for a review. It follows that the molten globule must also be protected from large fluctuations by an 'all-or-none' transition.

Acknowledgement: We are grateful to A.V. Finkelstein for the thermodynamic analysis of equation (2) and to C. Crane-Robinson for carefully reading the manuscript and helpful advice. We also thank A.M. Gutin and E.I. Shakhnovich for valuable discussions.

References

[1] Ptitsyn, O.B. (1987) J. Protein Chem. 6, 272–293.
[2] Kuwajima, K. (1989) Proteins: Structure, Function and Genetics 6, 87–103.
[3] Christiensen, H. and Pain, R.H. (1991) Eur. Biophys. J. 19, 221–229.
[4] Baldwin, R.L. (1991) Chemtracts-Biochem. Mol. Biol. 2, 379–389.
[5] Dobson, C.M. (1992) Current Opinion Struct. Biol. 2, 6–12.
[6] Ptitsyn, O.B. (1992) in: Protein Folding (T.E. Creighton, Ed.) pp. 243–300, Freeman, New York.
[7] Baldwin, R.L. (1993) Current Opinion Struct. Biol. 3, 84–91.
[8] Dolgikh, D.A., Gilmanshin, R.I. Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov, S.Yu. and Ptitsyn, O.B. (1981) FEBS Lett. 136, 311–315.
[9] Gast, K., Zirwer, D., Welfle, H., Bychkova, V.E. and Ptitsyn, O.B. (1986) Int. J. Biol. Macromol. 8, 231–236.
[10] Baum, J., Dobson, C.M., Evans, P.A. and Hanley, C. (1989) Biochemistry 28, 7–13.
[11] Jeng, M.F., Englander, S.W., Elöve, G.A., Wang, A.J. and Roder, H. (1990) Biochemistry 29, 10433–10437.
[12] Hughson, F.M., Wright, P.E. and Baldwin, R.L. (1990) Science 249, 1544–1548.
[13] Baldwin, R.L. and Roder, H. (1991) Current Opinion Struct. Biol. 1, 218–220.
[14] Peng, Z. and Kim, P.S. (1994), submitted for publication.
[15] Tanford, C. (1968) Adv. Prot. Chem. 23, 121–282.
[16] Privalov, P.L. (1979) Adv. Prot. Chem. 33, 167–244.
[17] Bower, V.E. and Robinson, R.A. (1963) J. Phys. Chem. 67, 1524–1527.
[18] Pace, C.N. (1963) Methods Enzymol. 131, 266–280.
[19] Aune, K.C. and Tanford, C. (1969) Biochemistry 8, 4586–4590.
[20] Ptitsyn, O.B. (1972) Pure Appl. Chem. 31, 227–244.
[21] Hill, T. (1968) Thermodynamics of Small Systems, Wiley, New York.
[22] Corbett, R.J.T. and Roche, R.S. (1984) Biochemistry 8, 4586–4590.
[23] Withka, J., Moncuse, R., Baziotis, A. and Maskiewicz, R. (1987) J. Chromatography 398, 73–202.
[24] Uversky, V.N. (1993) Biochemistry 32, 13288–13298.
[25] Uversky, V.N., Semisotnov, G.V., Pain, R.H. and Ptitsyn, O.B. (1992) FEBS Lett., 314, 89–92.
[26] Karplus, M. and Shakhnovich, E. (1992) in: Protein Folding (T.E. Creighton, Ed.) pp. 127–195, Freeman, New York.
[27] Bychkova, V.E., Pain, R.H. and Ptitsyn, O.B. (1988) FEBS Lett. 238, 231–234.
[28] Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L. and Hartl, F.U. (1991) Nature 352, 36–42.
[29] Van der Goot, F.G., Gonzales-Manas, J.M., Lakey, J.H. and Pattos, F. (1991) Nature 354, 408–410.
[30] De Jongh, H.H.J., Killian, J.A. and de Kruijff, B. (1992) Biochemistry 31, 1631–1643.
[31] Hua, Q.X., Ladbury, T.E. and Weiss, M.A. (1993) Biochemistry 32, 1433–1442.
[32] Bychkova, V.E. and Ptitsyn, O.B. (1993) Chemtracts-Biochem. Mol. Biol. 4, 133–163.