Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Structural Biology

Molten globules, entropy-driven conformational change and protein folding

Robert L Baldwin[1] and George D Rose[2]

The exquisite side chain close-packing in the protein core and at binding interfaces has prompted a conviction that packing selectivity is the primary mechanism for molecular recognition in folding and/or binding reactions. Contrary to this view, molten globule proteins can adopt native topology and bind targets tightly and specifically in the absence of side chain close-packing. The *molten globule* is a highly dynamic form with native-like secondary structure and a loose protein core that admits solvent. The related (but still controversial) *dry molten globule* is an expanded form of the native protein with largely intact topology but a tighter protein core that excludes solvent. Neither form retains side chain close-packing, and therefore both structure and function must result from other factors, assuming that the reality of the dry molten globule is accepted. This simplifying realization calls for a re-evaluation of established models.

Addresses
[1] Department of Biochemistry, Stanford University Medical Center, Beckman Center, School of Medicine, 279 Campus Drive West, Stanford, CA 94305-5307, United States
[2] Jenkins Department of Biophysics, Johns Hopkins University, Jenkins Hall, 3400N. Charles Street, Baltimore, MD 21218, United States

Corresponding author: Baldwin, Robert L (baldwinb@stanford.edu)

Current Opinion in Structural Biology 2012, **23**:xx–yy

This review comes from a themed issue on **Folding and binding**

Edited by **Jayant Udgaonkar** and **Susan Marqusee**

0959-440X/$ – see front matter, Published by Elsevier Ltd.

http://dx.doi.org/10.1016/j.sbi.2012.11.004

Introduction

The two-state approximation in protein folding [1] dates back almost half a century [2] and has conditioned the field's thinking for the duration. Specifically, a wealth of experimental evidence is consistent with the view that the folding of sufficiently small proteins is highly cooperative, akin to a first order phase transition [1]. As folding proceeds, the equilibrium population partitions into two predominant species, U(nfolded) and N(ative), with only a negligible contribution from intermediate species. Accordingly, the folding reaction can be written as U ⇌ N, with an equilibrium constant K_{eq} = [N]/[U] that can be obtained under any given set of conditions by measuring experimentally-accessible probes of the native state, such as Trp fluorescence or far-UV circular dichroism.

The two-state model has led to the implicit assumption that conspicuous features in N are prerequisite to its formation, none more so than side chain close-packing. Packing efficiency is assessed by the packing density, a dimensionless measure of a protein's summed atom volumes normalized by its molecular volume. Typically, proteins are well packed with few internal voids of atomic dimension and with mean packing densities around ⟨0.75⟩, similar to the packing densities of small organic solids [3]. The supposition that side chain close-packing and native conformation are linked events is the basis for knowledge-based potentials, Gō models, and reaction coordinates based on counting native contacts.

Consequently, the surprising discovery of the molten globule (MG) conformation, which is not close-packed, led to much discussion about its role – if any – in the mechanism of protein folding [4,5]. A related consideration regarding function is whether or not MGs can participate in specific binding reactions.

Initially, the term *molten globule* referred to the 'wet' molten globule, a highly dynamic state with native-like secondary structure and a protein core that admits solvent and lacks close-packing [4]. Later, there has been a growing realization that the 'dry' molten globule [6] is another distinct state along a graduated MG spectrum. The dry molten globule (DMG) is an expanded form of the native protein in which close-packing has been released at the start of unfolding, but native-like conformation persists and water does not penetrate the core [7•].

A paradigm-shifting insight from both types of molten globules, MGs and DMGs, is the enormously simplifying realization that both native topology and biological function can emerge in states that precede side chain close-packing and in the absence of ligand/substrate binding. Remarkably, formation of native topology and side chain close-packing are separable, unlinked events in the DMG model for folding/unfolding [7•], contrary to well-entrenched expectations arising from a strict two-state model.

This review focuses on specific binding by MGs (Part I) and entropy-driven conformational change (Parts II and III). Part I discusses two important systems in detail, highlighting the surprising finding that MGs can bind their substrates selectively and even tightly.

Parts II and III explore the pervasive influence of side chain conformational entropy. Many studies assume, if

Please cite this article in press as: Baldwin RL, Rose GD. Molten globules, entropy-driven conformational change and protein folding, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.11.004

only tacitly, that conformational entropy does not play a guiding role in protein folding because it has already been dissipated when the native state emerges. According to this view, favorable interactions between and among side chains select the native state from other conceivable conformers while the population entropy diminishes haphazardly [8]. The studies in part II challenge this assumption. When a DMG is formed at the start of unfolding, the process is driven by the increase in conformational entropy. Similarly, when two proteins form a complex, conformational entropy is not a passive bystander but instead plays an active role in driving complex formation. Recent evidence from NMR shows that side chain conformational entropy in solution is complex, again reinforcing the realization that the thermodynamics of close-packing cannot be inferred from inspection of protein X-ray structures.

Some thermodynamic preliminaries

The energetics of close-packing in DMGs can be modeled by data from the crystal-melting equilibria of alkane hydrocarbons [7•,9,10]. Key thermodynamic features that govern the release of close-packing are: (1) a favorable change in conformational entropy that drives the process, (2) a compensating unfavorable change in enthalpy, and (3) an overall change in free energy that is much smaller than the change in $-T\Delta S$. A central question is: how cooperative is the release of close-packing and what controls its cooperativity? Thermodynamic and kinetic data are available for villin headpiece [11], and the thermodynamic results are consistent with the crystal-melting model. In this case at least, the release of close-packing is highly cooperative. Fortunately, the villin headpiece DMG is stable, and therefore it can be analyzed experimentally.

The defining difference between a dry molten globule (DMG) and a conventional molten globule (MG) is that water has been squeezed from the core of a DMG. Initially, the DMG was proposed as the transition state ($I^{\ddagger}$, which cannot be observed directly) of a two-state ($N = I^{\ddagger} = U$) folding reaction [6]. However, later experiments indicated that the DMG is a stable unfolding intermediate that can be studied experimentally [7•]. Operationally, the DMG can be unambiguously distinguished from a wet MG by the fact that its peptide backbone NH protons are well protected against exchange [7•]. Another standard test for distinguishing a DMG from an MG is that the fluorescent indicator dye ANS (8-anilino-1-naphthalene sulfonate) can penetrate and bind within the hydrophobic core of an MG but not a DMG. ANS-binding corroborated other results using circular dichroism and fluorescence resonance energy transfer, which suggest that a DMG is formed when single-chain monellin starts to unfold [12]. A related study based on molecular dynamics simulations found that urea, but not water, can penetrate the protein core of a presumed DMG that forms at the initial stage of unfolding of hen lysozyme [13].

Part I: Specific binding by molten globules

Provocative early studies of secondary structure in MGs by NMR-hydrogen exchange suggested specific binding between helices. In detail, helices A, G and H of myoglobin (Mb), which have little tendency to form in isolation, are present in the pH 4 molten globule of apoMb, as reported by backbone NH protons that are protected against exchange [14]. How could A and GH, situated at opposite ends of the polypeptide chain, stabilize each other without specific binding? A related result was reported in an early folding kinetics study of horseheart ferri-cyt *c* [15]. There, the N-terminal and C-terminal helices develop rapidly with similar kinetics, apparently forming a complex between these distant chain segments before the central helix forms. A later mutational study of yeast iso-1-cyt *c* confirmed that native-like interactions between the N-terminal and C-terminal helices are present when the acid MG is formed [16].

These studies imply that specific binding interactions can occur when MGs are formed. In retrospect, this might have been anticipated. Specific binding is also observed between complementing protein fragments such as the disordered S-peptide and its cognate S-protein, which together form ribonuclease S [17]. The basic question is: can specific binding occur in the absence of close packing?

The physical–chemical mechanism of specific binding is a relevant question in all of these systems. It is often supposed that tight, specific binding is not possible in the absence of close-packing, but clearly this is not the case for the S-peptide:S-protein system [17]. Still, it is surprising to find that an MG apparently can function as an enzyme, where tight, specific binding is needed but typically is tied to close-packing, as discussed next (Figure 1).

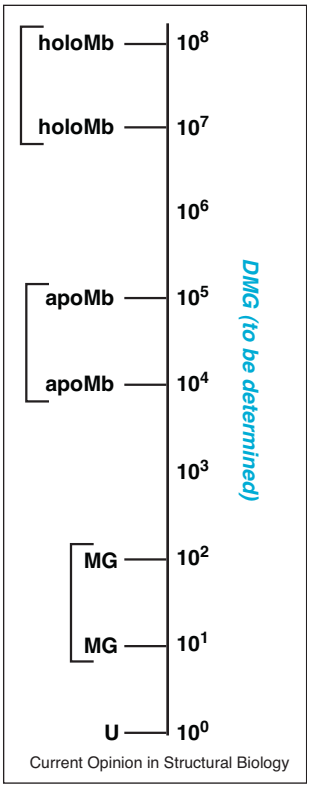
An active enzyme that appears to be an MG

In 2004, Hilvert and co-workers obtained a 109-residue designed enzyme that fortuitously has the properties of an MG [18••]. The authors' intention was simply to convert a dimeric enzyme (chorismate synthase, CM, from *methanococcus*) to a monomeric form by inserting a hinge-loop peptide within the dimer interface. Their resultant monomeric enzyme, mCM, had high specific enzymatic activity, like that of CM, but unlike the parental dimer, other properties of mCM were characteristic of an MG: rapid H/D exchange (HX) measured by mass spectrometry, thermal unfolding that shows little or no cooperativity, poor NMR signal dispersion, and the ability to bind the MG-indicator dye ANS.

The authors conducted a thorough investigation of both binding and activity for the MG enzyme, made possible

Please cite this article in press as: Baldwin RL, Rose GD. Molten globules, entropy-driven conformational change and protein folding, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.11.004

Figure 1



Hydrogen exchange protection factors in myoglobin. Protection factors (PF) measure amide proton exchange rates ($k_{observed}$) normalized by the exchange rate expected for an unprotected proton ($k_{unprotected}$): PF = $k_{unprotected}/k_{observed}$. For example, a PF of 10^2 means that hydrogen exchange of a given amide proton is slowed by a factor of 100 in comparison to the corresponding exchange in a peptide model [47]. Small PFs (10^1 to 10^2) are characteristic of wet molten globules, while at the opposite extreme, large PFs (10^7 to 10^8) are characteristic of native globular proteins. Data shown here are from Hughson *et al.* [14]. Notably, PFs for the dry molten globule have yet to be determined.

by the fact that the protein undergoes close-packing on binding a substrate analog (TSA = transition state analog). It is informative to review their results in detail. Specifically, binding TSA converts the MG into a close-packed ordered form, as judged by multiple criteria: (i) the well-resolved 2D-NMR spectrum, (ii) the appearance of a near-UV CD spectrum, (iii) a thermal unfolding transition that is moderately cooperative, and (iv) slowing of the previously rapid HX.

Furthermore, there appears to be a rapid two-state equilibrium between the ordered and disordered forms of mCM, made apparent by the fact that on titration of the MG with TSA, resolved resonances are observed only for the ordered form. Consequently, the intensity of any given NMR resonance is proportional to the bound TSA fraction. If the equilibrium pre-exists before adding TSA, it must be shifted predominantly toward the disordered form because the ordered form is not observable. Thus, the binding of TSA by mCM appears to follow an induced-fit mechanism (or, more precisely, a binding-then-folding mechanism, like that of the S-peptide:S-protein system [17]).

The enzymatic activity (k_{cat}/K_m) of mCM is only three-fold lower than that of the parental enzyme CM, and the difference appears entirely in K_m. The authors suggest that published studies of variants of two other enzyme systems have some similarity to the MG-like behavior of mCM.

Residue flexibility of the MG studied by NMR

The ordered complex of mCM with TSA has a large number of well-resolved resonances, none in the engineered hinge-loop peptide. It was possible to measure residue flexibility throughout the protein by NMR relaxation techniques, and also to monitor structural changes following ligand binding [19]. ^{2}S order parameters, internal correlation times, and line broadening caused by conformational exchange provided information about motion on both fast (nanosecond–picosecond, ns–ps) and slower (millisecond–microsecond, ms–μs) timescales. The fast timescale results are fairly uniform throughout the protein; their values indicate only limited motion, characteristic of normal, well-folded proteins. However, the ms–μs timescale results indicate considerable flexibility throughout the protein, definitely greater than for normal proteins; such results are those expected for an intermediate conformation, somewhere between an MG and a close-packed form. Rate constants for exchange between conformations were estimated to be in the range 10–1000 s^{-1}.

Binding of a transition-state analog

Fast binding kinetics, before attaining steady-state conditions, were measured for both the transition state analog and prephenate, the product of the enzymatic reaction. TSA binding shows biphasic kinetics: initial fast binding of TSA to the MG form followed by a slower intramolecular conversion of the MG to an ordered form. Formation of the ordered form follows an induced-fit mechanism and formation of the active enzyme, in the presence of substrate, is presumed to follow a similar mechanism. The kinetics of binding product (prephenate) show only a single kinetic phase for fast binding, comparable to the fast binding of TSA to the MG form.

Thermodynamics of ligand binding

Using isothermal titration calorimetry (ITC), the thermodynamics of ligand binding to both mCM and CM were analyzed and compared [20]. TSA binding to either mCM or CM is entropically favored, CM more strongly favored than mCM, with $-T\Delta S$ values (20 °C) of −1.5 kcal/mol and −6.6 kcal/mol, respectively. Changes in conformational entropy on binding are expected to localize to the enzyme because TSA is a rigid molecule, and its

Please cite this article in press as: Baldwin RL, Rose GD. Molten globules, entropy-driven conformational change and protein folding, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.11.004

conformation is not expected to change appreciably on binding. The less favorable entropy change for TSA binding to mCM was attributed to the non-localized loss of conformational entropy experienced by the MG on shifting to the ordered form. Both mCM and CM bind TSA fairly tightly (K_d = 7 and 13 μM, respectively, at 20 °C); the authors attribute tight binding to hydrophobic interactions (desolvation of nonpolar surfaces) and to release of ordered water molecules from the charged TSA. The binding of TSA is favored enthalpically, not just entropically, with a ΔH that is more favorable for mCM ($\Delta H = -5.6$ kcal/mol) than for CM ($\Delta H = -2.1$ kcal/mol). The difference between these two values (−3.5 kcal/mol) was attributed to structural tightening of mCM following initial ligand binding in the biphasic reaction.

The authors find a striking thermodynamic result for TSA binding to mCM: based on the temperature dependence of ΔH, the $-\Delta C_p$ of TSA binding for mCM is nearly four times larger than for CM. They interpret this result as arising from a protein disorder to order transition with a complex temperature dependence and argue that this behavior might be expected. For example, the average size of the partially ordered structures in the MG form of α-lactalbumin depends on temperature in a complex manner when the MG undergoes thermal unfolding [21].

Kinetics of ligand binding

Tight binding of TSA by the parental enzyme, CM, was expected, but tight TSA binding by the MG, mCM, was surprising. More surprising yet was the observation that mCM, a confirmed molten globule, binds TSA three-fold faster than CM. The kinetics of binding and dissociation of the enzymatic product, prephenate, were too fast to measure by the stopped-flow method in water at 20 °C. When measured in conditions that slow down binding (glycerol at 7.5 °C), the prephenate kinetics were still consistent with faster ligand binding to the MG (mCM) than to the ordered form (CM).

Summary

Specific and fast ligand binding to an MG is possible, as seen in the system developed by Hilvert and co-workers. Upon binding the large TSA ligand, the MG is converted to an ordered, close-packed form. Residue flexibility in this bound form has a dual character, resembling an MG on an ms–μs timescale but a normal, close-packed protein on an ns-ps timescale. As such, the system is a promising choice for dissecting factors involved in the cooperativity of close-packing.

Specific ligand binding by MGs of periplasmic binding proteins

Varadarajan and co-workers studied ligand binding by MG forms of four proteins from the periplasm of *E. coli* [22••] whose function is to bind and transport specific ligands: leucine, isoleucine, maltose and ribose. These proteins have been well characterized in numerous studies. They are large, more than twice the size (28–41 kDa) of the MG (12 kDa) studied by Hilvert and co-workers. Their X-ray structures are available in both the free and ligand-bound forms, and they show a ligand-binding cleft situated between two domains. Like most proteins known to form molten globules, the MG appears under low pH in conditions where the native protein (N) is unstable. Specifically, pH values of 3.0–3.4 were used, at a buffer concentration of 30 mM.

Thermodynamics of ligand binding

Are key features of the ligand binding reactions the same for MG and N? The thermodynamics of ligand binding were measured using isothermal titration calorimetry (ITC). The results evade generalization, apart from the observation that N has a somewhat higher affinity for its ligand than MG in every case: $-\Delta G$ values vary from 7.0 to 8.7 kcal/mol for N and from 6.0 to 7.2 kcal/mol for MG. ITC measurements were made at pH 8.0 or 7.0 for N and between pH 3.4 and 3.0 for MG. Ligand binding was entropy-driven in three of the MG proteins; two of them also exhibit entropy-driven ligand binding to N, but the remaining two have qualitatively different thermodynamic binding signatures. As such, generalities about differences in ligand binding between MG and N cannot be inferred from these thermodynamic data.

Failure of the MG to become close-packed upon ligand binding

In all four proteins, MG conformations persist in the ligand-bound forms, unlike the MG studied by Hilvert and co-workers. Although Hilvert's MG is converted to an ordered form on binding ligand, nevertheless the MG does bind ligand and the reaction is fast. For the periplasmic binding proteins, MG vs. N conformations were differentiated using standard tests: binding of the MG-indicator dye ANS and the appearance of a characteristic near-UV CD spectrum in N but not MG. ANS binding indicates that the MGs are wet, not dry molten globules. In one surprising respect, the MGs of these four proteins differ strikingly from the MG studied by Hilvert and co-workers: both the free and ligand-bound forms show cooperative thermal unfolding transitions. Ligand binding stabilizes the MG, which then releases its ligand as thermal unfolding occurs. A possible explanation for the cooperative thermal unfolding transitions is that these proteins contain two domains, a folded core and a MG domain with the ligand binding site. A structure of this kind has been proposed by Poulsen and co-workers [23] for the nuclear coactivator binding domain of CREB binding protein.

Part II: Relation between entropy-driven conformational change and folding

NMR measurements that provide information about protein conformational entropy in atom-level detail are

Please cite this article in press as: Baldwin RL, Rose GD. Molten globules, entropy-driven conformational change and protein folding, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.11.004

becoming feasible, based on the development of an 'entropy meter' [24,25••], a catchy name for a collective, residue-weighted measure of representative methyl symmetry axis order parameters, O^2_{axis} [26]. A protein's methyl groups are highly dynamic, even those buried in the close-packed core [27]. Order parameters quantify such motions on a scale ranging from 0 (high mobility/(*S*) to 1 (low mobility/(*S*). In essence, the entropy meter reports a scalar average, like a thermometer, and Wand and co-workers use it to extract side chain entropy changes throughout the molecule when a protein binds a substrate.

The entropy meter measures detailed side chain motions on the nanosecond–picosecond (ns–ps) timescale, and in the aggregate, the overall side chain conformational entropy. Concomitantly, the total entropy change (ΔS°) upon complex formation can be measured by a classical procedure, isothermal titration calorimetry (ITC), which gives the total changes in enthalpy (ΔH°) and free energy (ΔG°) upon binding; $-T\Delta S^\circ$ is found from $\Delta G^\circ - \Delta H^\circ$. These two measures – one lumped, the other exquisitely detailed – allow the total entropy change to be dissected into contributions from individual side chains and other sources, provided that a quantitative functional relationship between the overall entropy change and the summed side chain conformational entropy changes can be established. In fact, such a relationship has now been determined successfully [24,25••].

As described below, the approach has been applied incisively in two different allosteric systems: calmodulin [24], which binds a very large number (~hundreds) of target proteins and catabolite activating protein [28•], which binds a specific DNA target sequence. These two systems include multiple examples of protein complex formation; each reports that upon binding a specific target molecule, side chain conformational entropy equilibrates throughout the protein, giving rise to a broad – but varying – distribution of methyl group order parameters.

The regulatory protein calmodulin (CaM) interacts with a large and conformationally-dissimilar assortment of target proteins. The CaM binding site features a prominent cluster of methionine side chains that are aptly dubbed 'methonine puddles' [29]. Typically, peptides from a target protein fold into helices upon binding to CaM, and six such peptide–calmodulin complexes were analyzed by Wand and co-workers [24,25••].

A striking aspect of the result is that when a target peptide binds, the changes in side chain conformational entropy propagate far beyond the immediate peptide–calmodulin interface. Remarkably, in at least five of these complexes, the side chain conformational entropy of a given complex (i.e. the value given by the entropy meter) was found to be a linear function of ΔS°, the total entropy change on complex formation. This gratifying outcome was unanticipated [24,25••,30], and protein chemists are only beginning to comprehend the consequences [30]. The authors caution that this direct relationship may hold for calmodulin-peptide complexes only because CaM evolved to exert regulatory control via changes in side chain conformational entropy.

However, fully equivalent results have now been reported by Kalodimos and co-workers [28•,31] (reviewed by [32]) for an entirely different allosteric system. These authors analyzed changes in conformational entropy when catabolite activating protein (CAP) binds to a specific DNA target sequence. The dimeric form of CAP is activated for DNA binding by cyclic AMP or cyclic GMP. Again, a conspicuous feature of the result is that changes in side chain conformational entropy upon DNA binding are widespread, providing further evidence that upon binding its target, side chain conformational entropy re-equilibrates throughout the protein, far beyond the binding site.

Recent developments using pressure dependence add another promising methodology for investigating such states [33•,34–37]. In contrast to the relatively rigid backbone, the order parameters of methyl-bearing side chains have a large and variable pressure dependence. Results from an initial study using this approach indicate that the conformational fluctuations giving rise to the fast side chain motions are volume fluctuations [33•]. In related work, a previously reported allosteric change in a PDZ domain [38] is accompanied by an increase in solution molecular volume, and 2D NMR spectra rule out formation of a wet MG (A.L. Lee, personal communication, submitted for publication).

Part III: The dry molten globule mechanism of protein unfolding

The preceding parts of this review discuss tight binding by a molten globule enzyme and four binding proteins (part I) and entropy-driven conformational change in allosteric proteins (part II). These two systems share essential features with the formation of a dry molten globule at the start of protein unfolding. The main feature in common to all three systems is an absence of side chain close-packing as revealed by such diverse measures as hydrogen exchange protection factors, binding of the MG-indicator dye ANS, and methyl rotor order parameters. The critical and surprising attribute of these systems is that the absence of side chain close-packing does not abolish protein:target recognition or intra-molecular recognition in the case of protein folding/unfolding.

Background

The first strong suggestion that unfolding might be initiated by release of close-packing came with the 1984 report from Segawa and Sugihara that there is a large (~50 kcal/mol), temperature-independent activation

Please cite this article in press as: Baldwin RL, Rose GD. Molten globules, entropy-driven conformational change and protein folding, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.11.004

enthalpy for unfolding [39]. What could it be? It has long been thought that unfolding might be initiated by the penetration of water into the native structure, but the activation enthalpy results offered no support for this suggestion [39]. Had the nonpolar groups in the protein interior become solvated by water, the large activation enthalpy for unfolding would have been accompanied by a large value of ΔC_p, the change in protein heat capacity. A large ΔC_p was indeed found, but accompanying the activation enthalpy for refolding, not unfolding [39]. Moreover, water penetration was expected to – but did not – produce a substantial kinetic *m*-value for unfolding, calculated from the dependence of the unfolding rate constant on denaturant molarity.

In 1989 Shakhnovich and Finkelstein proposed an unfolding mechanism in which the native protein expands, as close-packing is released and increased conformational motion propagates throughout the protein interior [6]. In their proposal, this expanded state was hypothesized to be the highly unstable transition state ($I^{\ddagger}$) in a two-state unfolding reaction, rather than an observable intermediate. However, in later reports, it became apparent that this type of unfolding gives rise to unfolding intermediates that are amenable to experimental characterization [7•,11,12,40–42].

Lessons for folding/unfolding

Illuminating parallels exist between protein folding and tight binding by molten globule proteins or entropy-driven allostery, as described above. The monomeric enzyme, mCM, has characteristic properties of an MG, yet binds substrate tightly and with high specificity [19,20]. Clearly, molecular recognition can survive in the absence of side chain close-packing. Similarly, in a folding protein, the emergence of fluctuating elements of secondary structure establish the components of a folding framework [8,43] which then self-associate with demonstrated specificity, as seen for the A and GH helices of the apomyoglobin MG [14]. Further intra-molecular recognition among these nascent subassemblies promotes ever larger assemblies– supersecondary structure and domains – resulting in hierarchic self-assembly under suitable folding conditions [44–46].

In the unfolding direction, overall topology remains intact in dry molten globule intermediates at the start of unfolding, although the side chains are no longer close-packed and water has yet to penetrate. The DMG is an authentic intermediate [7•], and specific association of its framework elements remains in effect. The transition from N to DMG is accompanied by a widespread re-equilibration in conformational entropy as side chains unlock, suggesting parallels with entropy-driven changes in CAP [28•,31].

Acknowledgement

Support from the Alexander von Humboldt Foundation (GDR) is gratefully acknowledged.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Barrick D: **What have we learned from the studies of two-state folders, and what are the unanswered questions about two-state protein folding?** *Phys Biol* 2009, **6**:15001.

2. Ginsburg A, Carroll WR: **Some specific ion effects on the conformation and thermal stability of ribonuclease**. *Biochemistry* 1965, **4**:2159-2174.

3. Richards FM: **Areas, volumes, packing, and protein structure**. *Annu Rev Biophys Bioeng* 1977, **6**:151-176.

4. Ptitsyn OB: **Structures of folding intermediates**. *Curr Opin Struct Biol* 1995, **5**:74-78.

5. Kuwajima K: **The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure**. *Proteins* 1989, **6**:87-103.

6. Shakhnovich EI, Finkelstein AV: **Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition**. *Biopolymers* 1989, **28**:1667-1680.

7. • Baldwin RL, Frieden C, Rose GD: **Dry molten globule intermediates and the mechanism of protein unfolding**. *Proteins* 2010, **78**:2725-2737.
Background of the dry molten globule mechanism of initiating protein unfolding.

8. Rose GD, Fleming PJ, Banavar JR, Maritan A: **A backbone-based theory of protein folding**. *Proc Natl Acad Sci USA* 2006, **103**:16623-16633.

9. Nicholls A, Sharp KA, Honig B: **Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons**. *Proteins* 1991, **11**:281-296.

10. Lazaridis T, Archontis G, Karplus M: **Enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics**. *Adv Protein Chem* 1995, **47**:231-306.

11. Reiner A, Henklein P, Kiefhaber T: **An unlocking/relocking barrier in conformational fluctuations of villin headpiece subdomain**. *Proc Natl Acad Sci USA* 2010, **107**:4955-4960.

12. Jha SK, Udgaonkar JB: **Direct evidence for a dry molten globule intermediate during the unfolding of a small protein**. *Proc Natl Acad Sci USA* 2009, **106**:12289-12294.

13. Hua L, Zhou R, Thirumalai D, Berne BJ: **Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding**. *Proc Natl Acad Sci USA* 2008, **105**:16928-16933.

14. Hughson FM, Wright PE, Baldwin RL: **Structural characterization of a partly folded apomyoglobin intermediate**. *Science* 1990, **249**:1544-1548.

15. Roder H, Elove GA, Englander SW: **Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR**. *Nature* 1988, **335**:700-704.

16. Marmorino JL, Lehti M, Pielak GJ: **Native tertiary structure in an A-state**. *J Mol Biol* 1998, **275**:379-388.

17. Bachmann A, Wildemann D, Praetorius F, Fischer G, Kiefhaber T: **Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction**. *Proc Natl Acad Sci USA* 2011, **108**:3952-3957.

18. •• Vamvaca K, Vogeli B, Kast P, Pervushin K, Hilvert D: **An enzymatic molten globule: efficient coupling of folding and catalysis**. *Proc Natl Acad Sci USA* 2004, **101**:12860-12864.
A molten globule that functions as an enzyme after specific ligand binding.

19. Pervushin K, Vamvaca K, Vogeli B, Hilvert D: **Structure and dynamics of a molten globular enzyme**. *Nat Struct Mol Biol* 2007, **14**:1202-1206.

Please cite this article in press as: Baldwin RL, Rose GD. Molten globules, entropy-driven conformational change and protein folding, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.11.004

20. Vamvaca K, Jelesarov I, Hilvert D: **Kinetics and thermodynamics of ligand binding to a molten globular enzyme and its native counterpart**. *J Mol Biol* 2008, **382**:971-977.

21. Griko YV, Freire E, Privalov PL: **Energetics of the alpha-lactalbumin states: a calorimetric and statistical thermodynamic study**. *Biochemistry* 1994, **33**:1889-1899.

22. •• Prajapati RS, Indu S, Varadarajan R: **Identification and thermodynamic characterization of molten globule states of periplasmic binding proteins**. *Biochemistry* 2007, **46**:10339-10352.
Molten globules at low pH that can bind their specific ligands tightly.

23. Kjaergaard M, Teilum K, Poulsen FM: **Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP**. *Proc Natl Acad Sci USA* 2010, **107**:12535-12540.

24. Frederick KK, Marlow MS, Valentine KG, Wand AJ: **Conformational entropy in molecular recognition by proteins**. *Nature* 2007, **448**:325-329.
Molten globules at low pH that can bind their specific ligands tightly.

25. •• Marlow MS, Dogan J, Frederick KK, Valentine KG, Wand AJ: **The role of conformational entropy in molecular recognition by calmodulin**. *Nat Chem Biol* 2010, **6**:352-358.
Measuring side chain conformational entropy from NMR line broadening and methyl axis order parameters.

26. Lipari G, Szabo A: **Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules 1. Theory and range of validity**. *J Am Chem Soc* 1982, **104**:4546-4559.

27. Kossiakoff AA, Shteyn S: **Effect of protein packing structure on side-chain methyl rotor conformations**. *Nature* 1984, **311**:582-583.

28. • Tzeng SR, Kalodimos CG: **Protein activity regulation by conformational entropy**. *Nature* 2012, **488**:236-240.
Role of side chain conformational entropy in CAP binding to DNA.

29. O'Neil KT, DeGrado WF: **How calmodulin binds its targets: sequence independent recognition of amphiphilic alpha-helices**. *Trends Biochem Sci* 1990, **15**:59-64.

30. Schwalbe H, Rinnenthal J: **Thermodynamics: the world is flat**. *Nat Chem Biol* 2010, **6**:312-313.

31. Popovych N, Sun S, Ebright RH, Kalodimos CG: **Dynamically driven protein allostery**. *Nat Struct Mol Biol* 2006, **13**:831-838.

32. Baldwin AJ, Kay LE: **Structural biology: dynamic binding**. *Nature* 2012, **488**:165-166.

33. • Fu Y, Kasinath V, Moorman VR, Nucci NV, Hilser VJ, Wand AJ: **Coupled motion in proteins revealed by pressure perturbation**. *J Am Chem Soc* 2012, **134**:8543-8550.
Pressure dependence of order parameters shows side chain volume fluctuations and backbone rigidity under pressure.

34. Barrett DG, Minder CM, Mian MU, Whittington SJ, Cooper WJ, Fuchs KM, Tripathy A, Waters ML, Creamer TP, Pielak GJ: **Pressure perturbation calorimetry of helical peptides**. *Proteins* 2006, **63**:322-326.

35. Tsamaloukas AD, Pyzocha NK, Makhatadze GI: **Pressure perturbation calorimetry of unfolded proteins**. *J Phys Chem B* 2010, **114**:16166-16170.

36. Rouget JB, Schroer MA, Jeworrek C, Puhse M, Saldana JL, Bessin Y, Tolan M, Barrick D, Winter R, Royer CA: **Unique features of the folding landscape of a repeat protein revealed by pressure perturbation**. *Biophys J* 2010, **98**:2712-2721.

37. Roche J, Caro JA, Norberto DR, Barthe P, Roumestand C, Schlessman JL, Garcia AE, Garcia-Moreno BE, Royer CA: **Cavities determine the pressure unfolding of proteins**. *Proc Natl Acad Sci USA* 2012, **109**:6945-6950.

38. Petit CM, Zhang J, Sapienza PJ, Fuentes EJ, Lee AL: **Hidden dynamic allostery in a PDZ domain**. *Proc Natl Acad Sci USA* 2009, **106**:18249-18254.

39. Segawa S, Sugihara M: **Characterization of the transition state of lysozyme unfolding. I. Effect of protein-solvent interactions on the transition state**. *Biopolymers* 1984, **23**:2473-2488.

40. Kiefhaber T, Labhardt AM, Baldwin RL: **Direct NMR evidence for an intermediate preceding the rate-limiting step in the unfolding of ribonuclease A**. *Nature* 1995, **375**:513-515.

41. Hoeltzli SD, Frieden C: **Stopped-flow NMR spectroscopy: real-time unfolding studies of 6-19F-tryptophan-labeled Escherichia coli dihydrofolate reductase**. *Proc Natl Acad Sci USA* 1995, **92**:9318-9322.

42. Vidugiris GJ, Markley JL, Royer CA: **Evidence for a molten globule-like transition state in protein folding from determination of activation volumes**. *Biochemistry* 1995, **34**:4909-4912.

43. Kim PS, Baldwin RL: **Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding**. *Annu Rev Biochem* 1982, **51**:459-489.

44. Rose GD: **Hierarchic organization of domains in globular proteins**. *J Mol Biol* 1979, **134**:447-470.

45. Baldwin RL, Rose GD: **Is protein folding hierarchic? I. Local structure and peptide folding**. *Trends Biochem Sci* 1999, **24**:26-33.

46. Baldwin RL, Rose GD: **Is protein folding hierarchic? II. Folding intermediates and transition states**. *Trends Biochem Sci* 1999, **24**:77-83.

47. Molday RS, Englander SW, Kallen RG: **Primary structure effects on peptide group hydrogen exchange**. *Biochemistry* 1972, **11**:150-158.

Please cite this article in press as: Baldwin RL, Rose GD. Molten globules, entropy-driven conformational change and protein folding, Curr Opin Struct Biol (2012), http://dx.doi.org/10.1016/j.sbi.2012.11.004