



# Unified understanding of folding and binding mechanisms of globular and intrinsically disordered proteins

Munehito Arai<sup>1</sup>

Received: 17 October 2017 / Accepted: 13 November 2017 / Published online: 6 January 2018  
© The Author(s) 2018. This article is an open access publication

## Abstract

Extensive experimental and theoretical studies have advanced our understanding of the mechanisms of folding and binding of globular proteins, and coupled folding and binding of intrinsically disordered proteins (IDPs). The forces responsible for conformational changes and binding are common in both proteins; however, these mechanisms have been separately discussed. Here, we attempt to integrate the mechanisms of coupled folding and binding of IDPs, folding of small and multi-subdomain proteins, folding of multimeric proteins, and ligand binding of globular proteins in terms of conformational selection and induced-fit mechanisms as well as the nucleation–condensation mechanism that is intermediate between them. Accumulating evidence has shown that both the rate of conformational change and apparent rate of binding between interacting elements can determine reaction mechanisms. Coupled folding and binding of IDPs occurs mainly by induced-fit because of the slow folding in the free form, while ligand binding of globular proteins occurs mainly by conformational selection because of rapid conformational change. Protein folding can be regarded as the binding of intramolecular segments accompanied by secondary structure formation. Multi-subdomain proteins fold mainly by the induced-fit (hydrophobic collapse) mechanism, as the connection of interacting segments enhances the binding (compaction) rate. Fewer hydrophobic residues in small proteins reduce the intramolecular binding rate, resulting in the nucleation–condensation mechanism. Thus, the folding and binding of globular proteins and IDPs obey the same general principle, suggesting that the coarse-grained, statistical mechanical model of protein folding is promising for a unified theoretical description of all mechanisms.

**Keywords** Protein folding · Intrinsically disordered protein · Ligand binding · Conformational selection · Induced-fit · Nucleation–condensation

## Introduction

Elucidation of the mechanisms of protein folding and function remains an outstanding challenge in biophysics. Extensive experimental and theoretical studies have greatly advanced our understanding of the folding mechanisms of globular proteins (Dill and Chan 1997; Arai and Kuwajima 2000; Daggett and Fersht 2003; Englander and Mayne 2014; Takahashi et al. 2016), although many fundamental problems remain

unsolved (Dill et al. 2008; Sosnick and Barrick 2011). Recent studies have also revealed that proteins disordered in isolation fold into specific structures upon binding to their partners. Because more than 30% of eukaryotic proteins have disordered regions of over 30 residues in length and participate in critical cellular control mechanisms, including transcription, translation, and cell cycle control, these proteins are categorized as intrinsically disordered proteins (IDPs) (Wright and Dyson 1999; Dunker et al. 2001). The mechanisms of coupled folding and binding of IDPs have been extensively studied (Dyson and Wright 2005; Wright and Dyson 2009, 2015; Tompa 2012; Mollica et al. 2016). In addition, recent advances in nuclear magnetic resonance (NMR) and fluorescence techniques have provided a detailed understanding of the dynamic motions of globular proteins during ligand binding and catalysis over multiple time scales, ranging from picoseconds to seconds or longer (Mittermaier and Kay 2006; Henzler-Wildman and Kern 2007; Banerjee and Deniz 2013).

---

This article is part of a Special Issue on ‘Biomolecules to Bio-nanomachines — Fumio Arisaka 70th Birthday’ edited by Damien Hall, Junichi Takagi and Haruki Nakamura.

---

✉ Munehito Arai  
arai@bio.c.u-tokyo.ac.jp

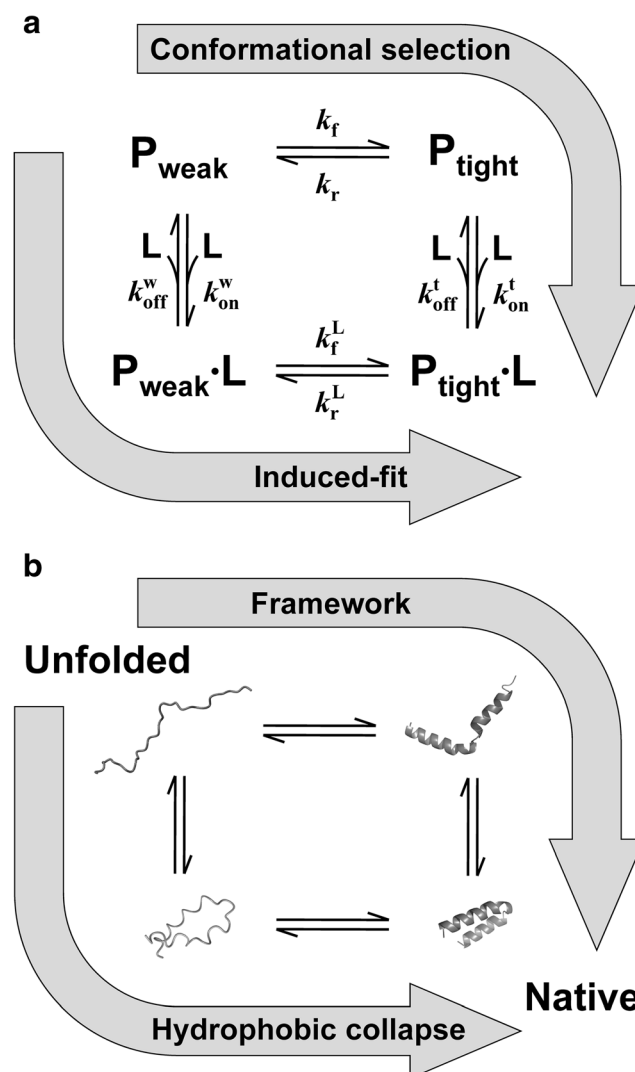
<sup>1</sup> Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8902, Japan

However, the folding and binding mechanisms of globular proteins and IDPs are typically discussed separately, and few studies have attempted to integrate these mechanisms (Kumar et al. 2000; Tsai et al. 2001; Liu et al. 2012; Chen et al. 2015). Although folding reactions of globular proteins and IDPs are induced by intramolecular and intermolecular interactions respectively, the forces responsible for conformational changes and binding are common to both proteins. Therefore, it may be possible to comprehensively understand the mechanisms of coupled folding and binding of IDPs, folding of small and multi-subdomain proteins, folding of multimeric proteins, and ligand binding of globular proteins.

In this review, we attempt to integrate the folding and binding mechanisms of globular proteins and IDPs in terms of conformational selection and induced-fit mechanisms (Boehr et al. 2009; Csermely et al. 2010). The two mechanisms have been widely used to understand the mechanisms of coupled folding and binding of IDPs and ligand binding of globular proteins, but have not been used to describe the folding mechanisms of globular proteins. We also reinterpret and synthesize folding mechanisms of small and multi-subdomain proteins, regarding the protein-folding reaction as the binding reaction of intramolecular segments accompanied by secondary structure formation. In the following, we describe conformational selection and induced-fit mechanisms and discuss mechanisms of coupled folding and binding of IDPs, folding of monomeric globular proteins, folding of multimeric proteins, and ligand binding of globular proteins. Finally, we aim to obtain a unified understanding of these mechanisms.

## Conformational selection and induced-fit mechanisms

The conformational selection and induced-fit mechanisms are two common mechanisms that explain binding reactions accompanied by conformational changes of proteins (Fig. 1a). Recent studies have demonstrated the existence of dynamic motions of proteins both in ligand-free and ligand-bound forms, underlying the dynamic binding mechanisms rather than the lock-and-key mechanism. The mechanism in which conformational change precedes binding is known as the *conformational selection* mechanism (or population-shift mechanism, pre-existing mechanism, folding-before-binding mechanism), while the mechanism in which binding precedes conformational change is known as the *induced-fit* mechanism (or binding-before-folding mechanism) (Monod et al. 1965; Koshland et al. 1966; Ma et al. 1999; James and Tawfik 2003; Onitsuka et al. 2008; Boehr et al. 2009; Csermely et al. 2010; Changeux 2012; Vogt et al. 2014). The induced-fit mechanism assumes that after weak binding to a ligand, a protein undergoes conformational change from the weakly bound conformation ( $P_{\text{weak}} \cdot L$ ) to the tightly bound



**Fig. 1** **a** Conformational selection and induced-fit mechanisms.  $P_{\text{weak}}$  and  $P_{\text{tight}}$  denote weakly and tightly binding conformations respectively, and  $L$  denotes a ligand.  $k_{\text{on}}$  and  $k_{\text{off}}$  denote the second-order binding rate constant and first-order dissociation rate constant respectively.  $k_f$  and  $k_r$  denote the forward and reverse rate constants respectively. **b** Two representative mechanisms of protein folding. The framework model corresponds to the conformational selection mechanism, while the hydrophobic collapse model corresponds to the induced-fit mechanism

conformation ( $P_{\text{tight}} \cdot L$ ) to fit the ligand. Recent studies have shown that many IDPs bind their partners through the induced-fit mechanism (Wright and Dyson 2009; Mollica et al. 2016). In contrast, the conformational selection mechanism assumes that equilibrium between the weakly binding conformation ( $P_{\text{weak}}$ ) (or binding-incompetent conformation) and tightly binding conformation ( $P_{\text{tight}}$ ) pre-exists, and that a ligand selectively binds  $P_{\text{tight}}$ . Recent advances in experimental studies using NMR spectroscopy and theoretical studies using molecular dynamics simulations have revealed that many globular proteins have low-populated, excited states that can bind ligands (Henzler-Wildman and Kern 2007; Boehr et al. 2009). In this mechanism, the population of  $P_{\text{tight}}$  is not

necessarily lower than that of  $P_{\text{weak}}$  if conformational equilibrium exists.

At one extreme, the “*ideal*” *induced-fit* mechanism assumes that a ligand must bind  $P_{\text{weak}}$  before the formation of  $P_{\text{tight}}L$ . Examples of this mechanism are a protein that cannot fold without a ligand, a protein with an extremely fast binding rate, and a protein that sequesters a ligand-binding site in  $P_{\text{tight}}$ , as a ligand cannot bind an inaccessible site buried deep inside of a protein. At the other extreme, the “*ideal*” *conformational selection* mechanism assumes that conformational changes of a protein must precede ligand binding, and that equilibrium pre-exists between the binding-incompetent and binding-competent conformations. Examples of this mechanism are a protein that cannot form a ligand-binding site before conformational change, a protein that cannot fold after binding, and a protein that has an extremely fast rate of conformational change, exceeding the apparent binding rate of the diffusion-controlled limit, even at high ligand concentrations.

Between these two extremes, the observed reaction mechanism is determined by competition of the fluxes of the conformational selection and induced-fit pathways (Hammes et al. 2009; Daniels et al. 2014; Greives and Zhou 2014) (Fig. 1). A flux is determined by both the rate constants and concentrations of all species involved in a reaction pathway. Thus, simple comparison of the rate of conformational change and second-order binding rate constant may not accurately reveal which pathway is dominant. Instead, the rate of conformational change and apparent binding rate of interacting elements determine the reaction mechanisms (Hammes et al. 2009; Greives and Zhou 2014). The former is affected by both the forward rate ( $P_{\text{weak}}$  to  $P_{\text{tight}}$ ) and reverse rate ( $P_{\text{tight}}$  to  $P_{\text{weak}}$ ), while the latter is affected by the second-order binding rate constant, first-order dissociation rate constant, and protein and ligand concentrations. The flux description suggests that if protein and ligand concentrations are low or if the conformational change from  $P_{\text{weak}}$  to  $P_{\text{tight}}$  is fast, the flux of the induced-fit pathway is small. Consequently, most protein molecules go through the conformational selection pathway, while a small fraction of protein molecules can go through the induced-fit pathway. Thus, the observed mechanism should be referred to as the “*apparent*” *conformational selection* mechanism. In contrast, if protein and ligand concentrations are high or if the conformational change from  $P_{\text{weak}}$  to  $P_{\text{tight}}$  is slow, the flux of the induced-fit pathway is large. Consequently, most protein molecules go through the induced-fit pathway, while a small fraction of protein molecules can go through the conformational selection pathway. Thus, the observed mechanism should be referred to as the “*apparent*” *induced-fit* mechanism. Therefore, the conformational selection and induced-fit mechanisms can coexist. Because few examples of the “*ideal*” mechanisms have been described, the reaction mechanisms correspond to the “*apparent*” mechanisms in many cases and depend on protein

and ligand concentrations. This indicates that assignment of a single mechanism to a single protein is impossible in many cases. However, if experiments are carried out under similar conditions for different proteins, comparison of the “*apparent*” reaction mechanisms can reveal the intrinsic characters of the proteins.

A more complicated reaction mechanism, known as the *extended conformational selection* mechanism (Csermely et al. 2010), involves initial ligand binding by the conformational selection mechanism followed by subsequent conformational change by the induced-fit mechanism (James and Tawfik 2003, 2005; Tang et al. 2007; Boehr et al. 2009; Espinoza-Fonseca 2009; Wlodarski and Zagrovic 2009; Wang et al. 2013a; Schneider et al. 2015). Furthermore, different regions of a single protein can adopt different reaction mechanisms (Arai et al. 2015).

## Mechanisms of coupled folding and binding of IDPs

Coupled folding and binding reactions of IDPs have been studied experimentally using various methods, including NMR relaxation dispersion measurements which reveal dynamic conformational changes on microsecond to millisecond time scales with a residue-specific-level spatial resolution (Mittermaier and Kay 2006; Gibbs and Showalter 2015). Additionally, many theoretical studies have been performed to explain experimental results and provide insight into the reaction mechanisms (Chen et al. 2015). The results revealed that many IDPs bind their partners by the induced-fit mechanism (Wright and Dyson 2009; Shammas et al. 2016; Mollica et al. 2016). The best studied example is the kinase-inducible domain (KID) of cAMP response element binding (CREB) protein, which binds the KIX domain of a CREB-binding protein (CBP) (Dyson and Wright 2005, 2016; Wright and Dyson 2015). Post-translational modifications of IDPs, such as phosphorylation, can modulate interactions with their partners by changing electrostatic interactions and/or by inducing the folding of IDPs (Forman-Kay and Mittag 2013; Bah and Forman-Kay 2016). Phosphorylation of KID (pKID) slightly stabilizes its free form, but largely increases its affinity for KIX by enhancing electrostatic attractions (Radhakrishnan et al. 1998). NMR relaxation dispersion experiments showed that intrinsically disordered pKID binds KIX by the induced-fit mechanism with accumulation of an intermediate (Sugase et al. 2007). Coupled folding and binding reactions by the induced-fit mechanism have also been reported for many other IDPs, including the transactivation domains (TADs) of c-Myc, Gal4, and VP16 (Ferreira et al. 2005), CBD of WASP binding to Cdc42 (Lu et al. 2007), IA<sub>3</sub> binding with YPrA (Narayanan et al. 2008), S-peptide binding with S-protein (Kiefhaber et al. 2012), N<sub>TAIL</sub> domain from Sendai virus

nucleoprotein binding to phosphoprotein PX (Wang et al. 2013b; Dosnon et al. 2015; Schneider et al. 2015), PUMA binding with MCL-1 (Rogers et al. 2014), BimBH3 binding with BAX (Jhong et al. 2016), and STAT2 TAD binding with the TAZ1 domain of CBP (Lindstrom and Dogan 2017).

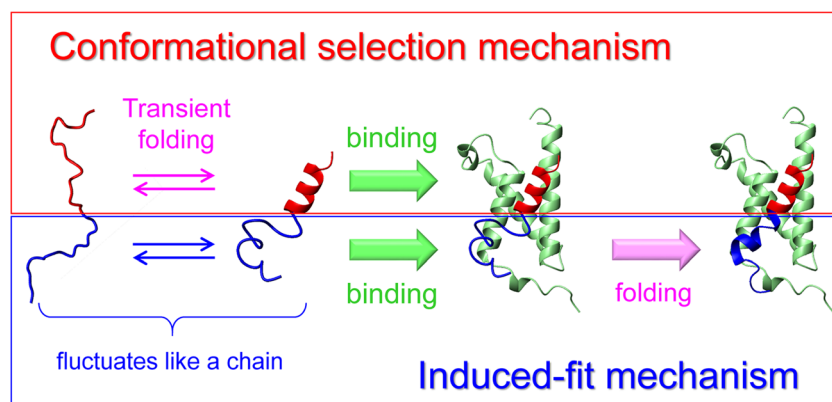
In contrast, few studies have reported IDPs that bind partners by the conformational selection mechanism (Onitsuka et al. 2008; Song et al. 2008; Iešmantavičius et al. 2014; Schneider et al. 2015). However, a combination of NMR and mutational analyses showed that the TAD of c-Myb binds the KIX domain of CBP mainly via the conformational selection mechanism (Giri et al. 2013; Arai et al. 2015). Mutation data were interpreted based on the following assumptions. In the case of conformational selection, mutations that stabilize the tightly binding conformation of an IDP should increase its population and accelerate the overall reaction rate. In contrast, in the case of the induced-fit mechanism, mutations that destabilize the tightly binding conformation of an IDP should increase the population of the weakly binding conformation, which can bind a partner, and accelerate the overall reaction rate. When mutations were introduced to stabilize the helical structure in the N-terminal region of c-Myb TAD, the overall reaction rate linearly increased with predicted helix stability (Arai et al. 2015). These results indicate that the N-terminal region of c-Myb TAD binds KIX through the conformational selection mechanism. Furthermore, mutations in the C-terminal region of c-Myb TAD that destabilize the helical structure increased the overall reaction rate, indicating that the C-terminal region interacts with KIX by the induced-fit mechanism (Arai et al. 2015). Thus, c-Myb provides an interesting example of an IDP in which two different reaction mechanisms coexist in a single protein (Fig. 2).

The above results show that pKID and c-Myb bind the same site on the same target protein KIX, but with different reaction mechanisms. Although both have similar second-order binding rate constants of  $10^6$ – $10^7$  M<sup>-1</sup> s<sup>-1</sup>, their rates of folding (conformational change) were different; it takes ~1 ms for pKID folding in the KIX-bound form and < 60 μs for c-Myb folding in the free form (Sugase et al. 2007; Arai

et al. 2015). Therefore, the difference in the reaction mechanism is probably related to differences in the folding rate, and a faster (or slower) folding rate results in the conformational selection (or induced-fit) mechanism. This is consistent with theoretical studies, which suggested that the rate of conformational change can determine the reaction mechanisms (Greives and Zhou 2014). Prediction of helical propensity indicated that c-Myb and pKID have high and low helical propensities (Arai et al. 2015), which are consistent with the fast and slow folding rates respectively. Such conformational propensities may depend on protein function. c-Myb is a constitutive transcriptional activator expected to bind its partner as soon as it is synthesized, and thus, it has high helical propensity to fold and bind quickly. In contrast, pKID is an inducible transcriptional activator that tightly binds KIX only after it is phosphorylated. The high degree of disorder and low propensity for secondary structure formation of pKID may facilitate interactions with and phosphorylation by protein kinase A, which binds peptide substrates in a relatively extended conformation (Zor et al. 2002). Therefore, function determines conformational propensity, conformational propensity determines folding rate, and folding rate determines the reaction mechanism of c-Myb and pKID (Arai et al. 2015).

Laser-induced temperature jump experiments showed that α-helices and β-hairpins are formed with a time constant of 0.2–2 and 0.1–6 μs respectively (Kubelka et al. 2004; Muñoz and Cerminara 2016). However, because these measurements were conducted using stable secondary structure elements, the folding of intrinsically disordered regions (IDRs) will occur over longer time scales, as indicated by the marginal stability of isolated secondary structure elements (Sadqi et al. 2003; Sugase et al. 2007; Arai et al. 2015). Moreover, some IDPs form single β-strand or irregular structures upon binding to their partners, but such structures cannot be stabilized without partners. Thus, few IDPs have stable secondary structure elements that can fold rapidly. Consequently, coupled folding and binding reactions of IDPs are dominated by the induced-fit mechanism.

**Fig. 2** Mechanisms of the coupled folding and binding reaction of intrinsically disordered c-Myb TAD upon binding to KIX (green). The N-terminal region of c-Myb TAD (red) binds KIX by the conformational selection mechanism (upper), while the C-terminal region of c-Myb TAD (blue) interacts with KIX by the induced-fit mechanism (lower)





Second-order binding rate constants of IDPs have been reported to be  $10^5$ – $10^{10}$   $\text{M}^{-1} \text{s}^{-1}$  (Sugase et al. 2007; Arai et al. 2012, 2015; Zhou and Bates 2013; Dogan et al. 2014; Milles et al. 2015; Shammass et al. 2016). We recently found that the intrinsically disordered N-terminal activation domain 2 of tumor suppressor p53 interacts with the TAZ2 domain of CBP with an extremely fast binding rate of  $1.7 \times 10^{10}$   $\text{M}^{-1} \text{s}^{-1}$  (Ferreon et al. 2009; Lee et al. 2010; Arai et al. 2012). This is the fastest rate among all previously known protein–protein associations. Although extended structures of IDPs were predicted to enhance binding rates by the “fly-casting mechanism” (Shoemaker et al. 2000), the large capture radius of IDPs also leads to slower translational diffusion, which opposes rapid binding (Huang and Liu 2009). Instead, favorable electrostatic attractions can dramatically accelerate binding reactions (Berg and von Hippel 1985; Schreiber 2002). p53 activation domain 2 and TAZ2 are negatively (−8) and positively (+14.3) charged respectively, and interact with each other through strong electrostatic attractions, leading to a binding rate close to the diffusion-controlled limit (Berg and von Hippel 1985; Arai et al. 2012). Consistently, theoretical studies showed that long-range electrostatic interactions are necessary for the rapid association and formation of initial encounter complexes (Ganguly et al. 2012; Wong et al. 2013; Chu et al. 2017; Ou et al. 2017). Although IDPs tend to be deficient in hydrophobic residues, the regions that directly interact with their partners often contain hydrophobic residues (Meszaros et al. 2007; Arai et al. 2010, 2012; Forman-Kay and Mittag 2013). Therefore, whereas long-range electrostatic interactions are important for attracting an IDP close to its partner, hydrophobic interactions stabilize direct contacts between them (Arai et al. 2012; Ganguly et al. 2012; Wong et al. 2013). These results also suggest that although hydrophobic interactions are reported to be long-range (Meyer et al. 2006), electrostatic interactions are more effective at longer ranges than are hydrophobic interactions.

## Folding mechanisms of globular proteins

### Multi-subdomain proteins

Experimental studies on protein folding have shown that folding behaviors differ between small single-domain proteins of less than 100 residues and multi-subdomain proteins of more than 100 residues (Jackson 1998; Arai and Kuwajima 2000; Daggett and Fersht 2003; Englander and Mayne 2014; Takahashi et al. 2016). Kinetic folding mechanisms of multi-subdomain proteins have been well studied for  $\alpha$ -lactalbumin ( $\alpha$ -LA) (Kuwajima 1989; Arai and Kuwajima 1996; Chaudhuri et al. 2000; Yoda et al. 2001; Arai et al. 2002; Saeki et al. 2004), non- $\text{Ca}^{2+}$ -binding lysozymes (hen and human lysozymes) (Matagne and Dobson 1998; Arai et al.

2000),  $\text{Ca}^{2+}$ -binding lysozymes (equine and canine lysozymes) (Mizuguchi et al. 1998; Nakamura et al. 2010), apomyoglobin (Dyson and Wright 2017; Nishimura 2017), barnase (Fersht 1993), dihydrofolate reductase (DHFR) (Jennings et al. 1993; Arai et al. 2003b, c, 2007, 2011; Arai and Iwakura 2005),  $\beta$ -lactoglobulin (Kuwajima et al. 1996; Arai et al. 1998; Fujiwara et al. 1999; Forge et al. 2000; Kuwata et al. 2001), cytochrome *c* (Takahashi et al. 1997; Akiyama et al. 2000; Winkler 2004; Goldbeck et al. 2009; Kathuira et al. 2014; Hu et al. 2016), ribonuclease A (Kim and Baldwin 1982; Neira and Rico 1997; Wedemeyer et al. 2000; Kimura et al. 2005), ribonuclease H (Raschke and Marqusee 1997; Hu et al. 2013; Rosen et al. 2014), and tryptophan synthase  $\alpha$ -subunit (Wu et al. 2008). These proteins accumulate a kinetic intermediate(s), resembling a molten globule state, during the folding reaction from the unfolded state to the native state (Kuwajima 1989; Kim and Baldwin 1990; Matthews 1993; Ptitsyn 1995; Arai and Kuwajima 2000; Bilsel and Matthews 2000, 2006; Baldwin 2008; Englander and Mayne 2014; Takahashi et al. 2016). Thus, the folding reaction of multi-subdomain proteins consists of at least three states and two steps. The molten globule state has a compact, globular structure (“globule”) with a pronounced secondary structure, but has little, if any, tertiary structure, as exemplified by the presence of only a small amount of tight packing of side chains (“molten”) (Kuwajima 1989; Ptitsyn 1995; Arai and Kuwajima 2000). For some proteins, including  $\alpha$ -LA,  $\text{Ca}^{2+}$ -binding lysozymes, apomyoglobin, cytochrome *c*, and ribonuclease H, the molten globule states are observed under equilibrium conditions, such as at low pH, moderate concentrations of denaturants, and moderate temperatures, and have been shown to be equivalent to the kinetic folding intermediates (Kuwajima 1989; Ptitsyn 1995; Raschke and Marqusee 1997; Mizuguchi et al. 1998; Arai and Kuwajima 2000; Nakao et al. 2005). Because the hydrophobic core present in the molten globule state is “wet” and exposed to solvent, water-separated hydrophobic interactions can exist, in addition to direct contacts between hydrophobic residues (Pratt and Chandler 1986; Arai and Kuwajima 1996). Protein folding from the molten globule to the native state occurs by the exclusion of water molecules through the “dry” molten globule state (Baldwin et al. 2010).

Two representative models have been postulated as the folding mechanisms of globular proteins (Arai and Kuwajima 1996): the *framework* model (or secondary structure coalescence model) (Kim and Baldwin 1982, 1990) and *hydrophobic collapse* model (Dill 1990; Dill et al. 1995). In the framework model, the formation of a secondary structure precedes compaction of a protein molecule through interactions between secondary structure elements, followed by formation of tight packing of side chains. In contrast, in the hydrophobic collapse model, compaction of a protein molecule by hydrophobic interactions precedes the

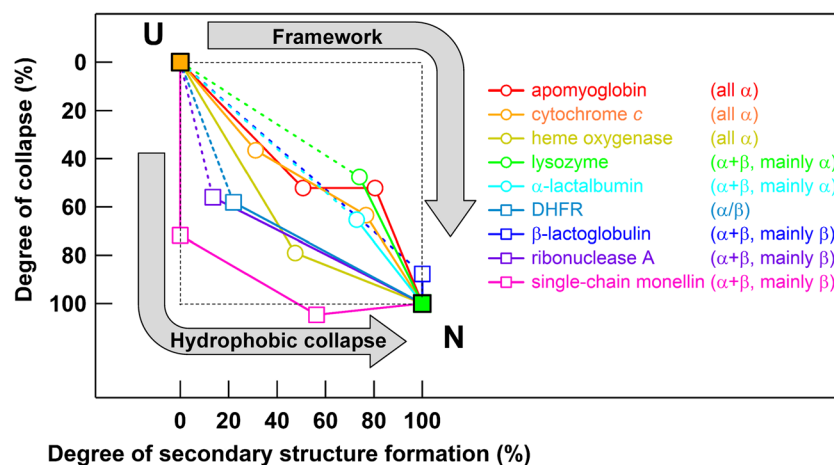
formation of secondary structure elements and tight packing of side chains. The driving force of protein folding is the local secondary structure propensities in the former and non-local hydrophobic interactions in the latter. In terms of conformational selection and induced-fit mechanisms (Tompa 2012; Chen et al. 2015), the framework model corresponds to a conformational selection mechanism, as secondary structure formation precedes binding of intramolecular segments (Fig. 1b). In contrast, the hydrophobic collapse model corresponds to an induced-fit mechanism, as compaction by binding of intramolecular segments precedes secondary and tertiary structure formation (Fig. 1b). Analogously, determinants of the folding mechanisms are the rate of secondary structure formation (i.e., conformational change) and binding rate of intramolecular segments.

One method for discriminating which mechanism better explains protein folding reactions is to draw folding trajectories in the simplified folding landscape, in which one axis is the degree of secondary structure formation while the other axis is the degree of collapse (Arai et al. 2007) (Fig. 3). Here, the folding trajectory of the ideal framework model involves an expanded intermediate with a substantial amount of secondary structure, while that of the ideal hydrophobic collapse model involves a compact intermediate without any secondary structure (Fig. 1b). Experimentally characterized folding trajectories of multi-subdomain proteins showed that almost all folding pathways were in the lower left half of the landscape, indicating that multi-subdomain proteins fold by the hydrophobic collapse (induced-fit) mechanism rather than by the framework (conformational selection) mechanism (Arai et al. 2007) (Fig. 3). These results demonstrate that non-local hydrophobic interactions are more important than local secondary structure propensities early in the folding of

multi-subdomain proteins. Furthermore, the results suggest that binding of intramolecular segments is faster than secondary structure formation, and competition between the rate of secondary structure formation and binding rate of intramolecular segments can determine the folding mechanisms. The rationale for these observations is that connections of interacting segments as a single chain increase the effective concentration between them and enhance their apparent binding rate. Experimental studies reported that protein hydrophobic collapse can occur within 60 ns (Sadqi et al. 2003), which is much faster than secondary structure formation (Kubelka et al. 2004; Muñoz and Cerminara 2016). If the second-order binding rate constant of intramolecular segments is  $10^6 \text{ M}^{-1} \text{ s}^{-1}$  and the effective concentration between them is 10 M (Robinson and Sauer 2000), hydrophobic collapse may occur on a time scale of 100 ns, which is consistent with experimental observations. Thus, although equilibrium and kinetic folding intermediates often have native-like secondary structures in a part of the molecule (Arai and Kuwajima 2000; Uversky and Fink 2002), non-specific hydrophobic collapse can occur before the formation of specific structures.

In addition to hydrophobic interactions, electrostatic interactions such as salt bridges can stabilize folding intermediates (Oliveberg and Fersht 1996). However, hydrophobic interactions are more important than electrostatic interactions in the folding of globular proteins. The presence of surface charge–charge interactions even decelerates a folding reaction, probably by restricting the ability to collapse (Kumik et al. 2012).

The folding trajectories depicted in Fig. 3 also show that secondary structure contents in the native structure are related to the folding mechanisms. Proteins composed of mainly  $\alpha$ -helices have folding trajectories involving concomitant compaction and secondary structure formation (Fig. 3), indicating



**Fig. 3** Folding trajectories of multi-subdomain proteins drawn in the simplified folding landscape. The horizontal and vertical axes show the degree of secondary structure formation, estimated from the change in circular dichroism intensity during folding, and degree of collapse, estimated from the change in the radius of gyration during folding, respectively. The unfolded state (*U*) and native state (*N*) are located at the *upper*

*left* and *lower right* respectively. Open circles and squares show the location of folding intermediates of  $\alpha$ -rich and  $\beta$ -rich proteins respectively. Continuous and dotted lines show the folding trajectories. Folding trajectories for the ideal conformational selection (*Framework*) mechanism and ideal induced-fit (*Hydrophobic collapse*) mechanism are indicated with *arrows*. Adapted with permission from Arai et al. (2007)

rapid formation of  $\alpha$ -helices. In contrast, proteins composed of mainly  $\beta$ -sheets display folding trajectories close to those of the ideal hydrophobic collapse model (Fig. 3), indicating slow formation of  $\beta$ -sheets. Therefore, secondary structure contents in native structure are closely related to the rate of secondary structure formation and thus determine the detailed folding mechanisms of multi-subdomain proteins (Arai et al. 2007).

In principle, protein folding reactions coupled with disulfide bond formation are essentially the same as those described above for disulfide-uncoupled folding. Experimental studies have shown that only native tertiary structures develop during oxidative folding if the refolding conditions are optimized (Wedemeyer et al. 2000), indicating that a native pair of cysteine residues comes in proximity resulting from hydrophobic collapse and native-like secondary structure formation during folding. Consistently, a disulfide-reduced protein can form an overall structure similar to the native state of the oxidized form (Redfield et al. 1999).

### Small single-domain proteins

Small globular proteins of less than 100 residues are typically composed of a single domain and fold in a two-state manner from the unfolded state to the native state without accumulation of an intermediate (Jackson 1998). The transition state between the unfolded and native state can be analyzed by  $\Phi$ -value analysis (Fersht et al. 1992; Fersht and Sato 2004). Comprehensive  $\Phi$ -value analysis has been performed for many small proteins, including chymotrypsin inhibitor 2 (CI2) (Itzhaki et al. 1995; Jackson 1998) and the B domain of protein A (Sato et al. 2004). The results showed that in the transition state, several non-local residues have high  $\Phi$ -values and form specific interactions, while others have low  $\Phi$ -values. The transition state had neither stable secondary structures nor compact structures (Plaxco et al. 1998), which is inconsistent with both the framework and hydrophobic collapse models. Because the number of hydrophobic residues is limited in small proteins, the binding rate between intramolecular segments is expected to be low. Additionally, the rate of secondary structure formation is expected to be low, as secondary structure elements have a small size and are less stable in small proteins. Thus, both hydrophobic collapse and secondary structure formation are less likely to occur early in the folding of small proteins. Rather, small proteins fold by a mechanism between the framework (conformational selection) and hydrophobic collapse (induced-fit) mechanisms, known as the *nucleation–condensation* mechanism (Fersht 1997). In this mechanism, a small number of non-local residues separated in an amino acid sequence form a specific hydrophobic interaction known as “a folding nucleus”. Formation of a folding nucleus is probably accompanied by formation of overall native-like secondary structure and

backbone topology, although hydrogen bonds and tight packing of side chains have not yet formed. Once the critical interactions are formed, the remaining structure condenses rapidly around the nucleus to fold into a stable native structure.

Statistical analysis of the folding kinetics of small single-domain proteins showed that the logarithm of the folding rate, which corresponds to the free energy difference between the unfolded and transition state,  $\Delta G_{U-TS}$ , is negatively well-correlated with the contact order, a parameter that represents the native backbone topology of a protein (Jackson 1998; Plaxco et al. 1998, 2000; Ivankov et al. 2003; Kamagata et al. 2004). This correlation indicates that proteins in the all- $\alpha$  class, which have many local contacts and a lower contact order, fold faster, while those in the  $\alpha/\beta$  and all- $\beta$  classes, which have many non-local contacts and a higher contact order, fold more slowly. In other words, proteins that show a smaller (or larger) decrease in conformational entropy during folding have smaller (or larger)  $\Delta G_{U-TS}$ , suggesting that the free-energy barrier in the folding of small proteins corresponds to a decrease in the conformational entropy to form both a folding nucleus and overall native-like (but unstable) secondary structure and backbone topology in the transition state. Thus, the folding rate of small proteins is closely related to the rate of secondary structure formation.

Remarkably, a correlation between the folding rate and contact order has been observed, even for multi-subdomain proteins; both the folding rates from the unfolded state to the intermediate and from the intermediate to the native state were significantly correlated with native backbone topology, as represented by the absolute contact order (Kamagata et al. 2004; Kamagata and Kuwajima 2006). This suggests that the folding mechanisms of both small and multi-subdomain proteins are essentially identical. In fact, similar to the situation in multi-subdomain proteins, the rate of secondary structure formation and binding rate of intramolecular segments can determine the detailed folding mechanisms of small proteins. As described above, the folding rate of small proteins indicates the rate of secondary structure formation. In addition, the molecular size of the transition state indicates the binding rate of intramolecular segments. Statistical analysis showed that proteins in the all- $\alpha$  class have expanded molecular sizes in the transition state (Plaxco et al. 1998), indicating that  $\alpha$ -helical proteins have faster rates of secondary structure formation but slower binding rates of intramolecular segments. This corresponds to the conformational selection (framework) mechanism. In contrast, proteins in the all- $\beta$  class have compact dimensions in the transition state (Plaxco et al. 1998), indicating that  $\beta$ -sheet proteins have slower rates of secondary structure formation but faster binding rates of intramolecular segments. This corresponds to the induced-fit (hydrophobic collapse) mechanism. Thus, although small proteins generally fold by the nucleation–condensation mechanism, the detailed folding mechanism approaches the conformational selection (framework)

mechanism when the rate of secondary structure formation is large and induced-fit (hydrophobic collapse) mechanism when it is small. These observations are consistent with those for multi-subdomain proteins (Arai et al. 2007). Thus, folding mechanisms of both small and multi-subdomain proteins can be integrated using the rate of secondary structure formation and binding rate of intramolecular segments.

The above considerations suggest that small proteins with many hydrophobic residues fold by the induced-fit (hydrophobic collapse) mechanism, as observed in the three-state folding of ubiquitin (Khorasanizadeh et al. 1996). Moreover, it is suggested that small proteins containing stable  $\alpha$ -helices may fold by the conformational selection (framework) mechanism, as observed for the engrailed homeodomain that accumulates a folding intermediate with  $\alpha$ -helices (Mayor et al. 2003).

By analogy, the nucleation–condensation mechanism, which is between the conformational selection and induced-fit mechanisms, has been observed in the coupled folding and binding of small IDPs. Whereas pKID has a low helical propensity particularly in the  $\alpha$ B region, which makes a dominant contribution to the free-energy barrier of binding and binds KIX by induced-fit, c-Myb has a high helical propensity and binds KIX by conformational selection (Arai et al. 2015). This suggests that IDPs with helical propensities between those of pKID and c-Myb bind their partners through the nucleation–condensation mechanism. In addition, stabilization of helical structures by mutations may change the binding mechanism from nucleation–condensation to conformational selection. An interesting example is the synergistic folding of ACTR and the nuclear cofactor binding domain (NCBD) of CBP, which exist in the unfolded and molten globule states in isolation respectively, but fold into specific structures upon mutual binding (Dogan et al. 2013; Haberz et al. 2016). The  $\Phi$ -values of the transition state in the coupled folding and binding of ACTR and NCBD are low, indicating that ACTR and NCBD synergistically fold via the nucleation–condensation mechanism. Indeed, helical propensities predicted by the AGADIR server (Muñoz and Serrano 1994) were 5.9% and 5.0% for ACTR and NCBD respectively, which are between those of the  $\alpha$ B region of pKID (0.6%) and c-Myb TAD (40.6%). In addition, stabilization of ACTR by mutations resulted in a binding reaction by the conformational selection mechanism (Iešmantavičius et al. 2014), supporting the conclusion that wild-type ACTR binds NCBD by nucleation–condensation. Therefore, mechanisms of coupled folding and binding of IDPs that form helices upon binding may be predicted by their helical propensities.

### Subdomain-wise folding of large proteins

Multi-subdomain proteins are composed of several subdomains corresponding to small single-domain proteins.

There are two ways to connect two small proteins as subdomains in a multi-subdomain protein. One is end-to-end tandem connection. Here, if two subdomains do not have mutual interactions and are stable enough to fold independently, both subdomains may fold by the nucleation–condensation mechanism (Arora et al. 2006). In contrast, if two subdomains interact with each other directly or indirectly through a rigid linker, both subdomains may affect the mutual folding reactions (Batey et al. 2008; Steward et al. 2012). The enzyme rhodanese, which has two similar subdomains tightly packed with each other, tends to misfold and requires molecular chaperones to correctly fold (Mendoza et al. 1991).

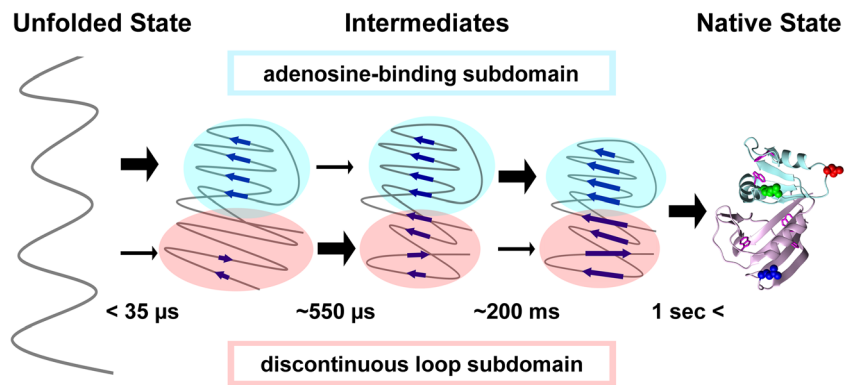
Another way of connecting two small proteins is to insert one subdomain (continuous insert) into a loop region of another subdomain (discontinuous parent). Domain insertion has been observed for 9% of domain combinations in the non-redundant structure database (Aroul-Selvam et al. 2004). If the insert subdomain is more stable than the parent subdomain, the insert may fold faster than the parent. In contrast, if the insert is unstable, the parent may fold faster than the insert. If both the insert and parent have similar amino acid sequences, folding reactions are more complicated; insertion of one CI2 into another CI2 resulted in destabilization of the double CI2, leading to the existence of two native conformers that folded and unfolded through two parallel pathways (Inaba et al. 2000).

For both ways of connecting two subdomains, subdomain-wise folding can occur, accumulating a folding intermediate in which one subdomain is partially folded and another is not. Such consideration is consistent with experimental results showing that one of the subdomains has a more organized structure than the others in the molten globule intermediate observed during the folding reaction of multi-subdomain proteins, including  $\alpha$ -LA, lysozyme, apomyoglobin, barnase, and DHFR (Fersht 1993; Matagne and Dobson 1998; Arai and Kuwajima 2000; Arai et al. 2011; Dyson et al. 2017). However, because the number of hydrophobic residues is larger in multi-subdomain proteins than in small proteins, hydrophobic collapse rather than nucleation of a small number of hydrophobic residues tend to occur in multi-subdomain proteins, and thus compact molten globules with a localized native-like structure are observed.

Recent experimental studies of DHFR folding support the subdomain-wise folding mechanism (Arai et al. 2011) (Fig. 4). DHFR consists of two subdomains, a discontinuous loop subdomain (DLD) and continuous, adenosine-binding subdomain (ABD). Previous stopped-flow studies showed that the folding intermediate ( $I_5$ ) that formed within the dead time ( $\sim 5$  ms) of the measurement had a more ordered structure in the DLD than in the ABD. However, continuous-flow experiments combined with fluorescence resonance energy transfer measurement showed that the burst-phase intermediate ( $I_6$ ) that formed within the dead time (35  $\mu$ s) of the



**Fig. 4** Subdomain-wise folding of DHFR. *Blue arrows* in the folding intermediates represent  $\beta$ -strands. *Thick and thin black arrows* respectively show that large and small conformational changes occur in the indicated subdomain during each phase. Adapted with permission from Arai et al. (2011)



measurement had a more compact structure in the ABD than in the DLD, and that compaction of the DLD to form the  $I_5$  intermediate occurred with a time constant of  $550 \mu\text{s}$  (Arai et al. 2011). Thus, hierarchical assembly of DHFR was observed in which each subdomain independently folds, subsequently docks, and then anneals into the native conformation after an initial heterogeneous global collapse. This observation suggests that proteins with kinetic molten globule intermediates, in which discontinuous subdomains are more organized than continuous subdomains when observed by stopped-flow techniques, may fold through initial compaction of a continuous subdomain when observed by techniques with a shorter dead time. Progressive folding, beginning with a continuous subdomain and spreading to distal regions, shows that chain entropy is a significant organizing principle in the folding of multi-subdomain proteins and single-domain proteins (Arai et al. 2011).

Notably, recent studies of *in vivo* folding have shown that codon translation rates can profoundly impact the cotranslational folding process on the ribosome, indicating that protein folding is guided not only by the amino acid sequence but also by the RNA sequence (O'Brien et al. 2014). The presence of rare codon clusters at domain boundaries of proteins, which may efficiently enable domain-wise folding of large multi-domain proteins, is controversial (Deane et al. 2011).

### Theory of folding mechanisms of globular proteins

Theoretical studies of protein folding have proposed that the energy landscape of a protein has a funnel shape, and that an unfolded protein molecule folds into its native state by sliding down the surface of the energy landscape (Bryngelson et al. 1995; Dill and Chan 1997; Dinner et al. 2000; Shakhnovich 2006). Development of theoretical methods for predicting a folding energy landscape and native structure using only an amino acid sequence is one of the major goals of theoretical studies of protein folding (Dill et al. 2008). One of the most promising theoretical models describing protein folding mechanisms is the *Wako–Saitō–Muñoz–Eaton (WSME)*

model (or island model) (Wako and Saitō 1978a, b; Muñoz and Eaton 1999; Sasai et al. 2016). The WSME model is a coarse-grained, statistical mechanical model of proteins and enables one to draw a free-energy landscape of a protein-folding reaction using information from the native structure. The model assumes that each residue adopts only the unfolded and native states and that two residues form a native-like contact only when the intervening residues between them are all in the native state. Thus, folding starts from local interactions between neighboring residues, and spreads to distal regions by the growth and coalescence of native-like segments. Moreover, this model considers only native contacts and thus reproduces folding reactions of hypothetical idealized proteins that can always fold into their native state (Gō 1983). Consequently, the WSME model guarantees that both the principle of minimal frustration and consistency principle that locally stable structure is consistent with the final folded, globally stable structure (Gō 1983; Bryngelson et al. 1995). Previous studies showed that the WSME model accurately explains the experimentally observed folding mechanism, i.e., the nucleation–condensation mechanism, of small single-domain proteins (Muñoz and Eaton 1999; Itoh and Sasai 2006; Sasai et al. 2016). These results suggest that real small proteins behave as ideal foldable proteins and that the consistency principle holds for small proteins.

Application of the WSME model to multi-subdomain proteins with end-to-end tandem connections has been successful in drawing free-energy landscapes of folding (Itoh and Sasai 2008, 2009). However, during the folding reactions of multi-subdomain proteins containing a subdomain insert, intermediates with localized native-like structures are frequently observed, in which a discontinuous parent subdomain is more organized than a continuous insert subdomain. Such folding behaviors cannot be described by the WSME model, as the model assumes that a discontinuous subdomain folds only after folding of the intervening subdomain. To solve this problem, a virtual linker between the N- and C-termini of DHFR, separated by  $15 \text{ \AA}$  in the native state, was introduced in the *extended WSME (eWSME)* model (Inanami et al. 2014). Experimental studies in which both termini were connected

by a linker showed that “circular” DHFR was more stable than wild-type “linear” DHFR, but that the folding behaviors under native conditions were unchanged by circularization (Arai et al. 2003b; Takahashi et al. 2007). Therefore, the eWSME model with a virtual linker may predict a free-energy landscape of folding that is applicable to a protein without a linker. Application of the eWSME model to DHFR predicted a folding pathway involving initial folding of the ABD and subsequent folding of the DLD, which is consistent with the experimental results (Inanami et al. 2014). Therefore, the eWSME model has been successfully used to explain the folding of a multi-subdomain protein containing a subdomain insert.

Although the original Ising-like Hamiltonian of the WSME model can explain the nucleation–condensation mechanism of folding for small single-domain proteins, the requirement of an additional Hamiltonian for linker introduction suggests that the original Hamiltonian is not sufficient to describe the cooperative hydrophobic collapse (induced-fit) mechanism of folding for multi-subdomain proteins, and that an additional term to enhance non-local (hydrophobic) interactions is necessary to formulate the consistency principle for multi-subdomain proteins. Future studies applying the eWSME model to many other proteins, optimizing linker introduction, and/or formulating variants of the WSME model may enable the prediction of free-energy landscapes of folding of all proteins. Moreover, because the forces responsible for folding and binding are common in globular proteins and IDPs, the WSME model is promising for developing a unified theoretical description of the mechanisms of folding and binding of all proteins. Indeed, the *allosteric WSME* model has been successfully applied to explain conformational selection and induced-fit mechanisms involved in allosteric transitions of proteins coupled with effector binding (Itoh and Sasai 2010, 2011; Sasai et al. 2016). Although future improvement in computer simulations may enable the reproduction of folding and binding reactions of all proteins, coarse-grained, statistical mechanical models are still required to determine the physics underlying these biological phenomena.

## Folding mechanisms of multimeric proteins

Elucidation of the folding mechanisms of multimeric proteins is also important, as most natural proteins exist as oligomeric complexes (Goodsell and Olson 2000). Multimer formation from fully unfolded monomers involves both folding and binding. Thus, overall folding mechanisms of multimeric proteins correspond to a combination of the folding mechanisms of monomeric globular proteins and coupled folding and binding mechanisms of IDPs. Many multimeric proteins fold by the induced-fit mechanism in which inter-subunit interactions induce subunit folding (Jaenicke 1987; Gloss and Matthews 1998; Jaenicke and Lilie 2000; Rumfeldt et al. 2008). This

mechanism is observed when monomers are unstable and when segment-swapped dimers are formed (Rentzeperis et al. 1999; Topping and Gloss 2004). Large stabilization energies and highly cooperative folding transitions of multimeric proteins primarily result from inter-subunit interactions (Neet and Timm 1994; Arai et al. 2003a). There are also many examples of multimeric proteins that fold by the conformational selection mechanism, in which subunit assembly occurs after the formation of molten globule-like monomeric intermediates (Jaenicke 1987; Jaenicke and Lilie 2000; Svensson et al. 2006; Rumfeldt et al. 2008; Galvagnion et al. 2009; Noel et al. 2009). This mechanism tends to be observed when monomers are stable. Some multimers fold by an ideal conformational selection mechanism, in which formation of a binding-competent structure is a prerequisite for inter-subunit interactions (Topping and Gloss 2004). In such cases, monomeric intermediates are not kinetic traps, but rather productive, on-pathway intermediates. Both hydrophobic interactions and electrostatic interactions are important for the folding of multimeric proteins, as the reactions are accelerated by strengthening hydrophobic interactions and by removing electrostatic repulsion (Waldburger et al. 1996; Jelesarov et al. 1998; Dürr et al. 1999; Rentzeperis et al. 1999).

Reflecting the complexity of the native structures, larger multimers generally fold with more complicated folding mechanisms, involving both conformational selection and induced-fit mechanisms. Examples include sequential formation of monomeric and dimeric intermediates and the presence of parallel folding channels (Rumfeldt et al. 2008). Multimers larger than dimers may fold through formation of lower-order multimers and their assembly (Dürr and Bosshard 2000; Ali et al. 2003; Riechmann et al. 2005; Rumfeldt et al. 2008). Gp57A is a molecular chaperone for tail fiber formation of bacteriophage T4, and folds into a native hexameric structure by rapidly forming a trimeric coiled-coil intermediate (Matsui et al. 1997; Ali et al. 2003).

One of the best-studied multimers is the homodimeric coiled-coil peptide GCN4-p1. In the folding of GCN4-p1, partial helix formation precedes dimerization (Zitzewitz et al. 2000), indicating a conformational selection mechanism. Consistent with this, stabilization of helical structures accelerated the folding reaction (Zitzewitz et al. 2000). However, destabilization of the helices resulted in the induced-fit mechanism, in which the initial step is the binding of two unstructured monomers (Meisner and Sosnick 2004). These results support the notion that competition between the rate of conformational change and apparent rate of binding determines apparent reaction mechanisms. Interestingly, cross-linked variants of GCN4-p1 accumulate a folding intermediate formed by collision of two strands (Wang et al. 2005), supporting the view that connections of interacting segments increase the effective concentration between them and enable rapid association.

## Mechanisms of ligand binding

### Ligand-induced folding

Folding of globular proteins and IDPs can be coupled with binding to various ligands, including DNA/RNA (Spolar and Record 1994; Rentzeperis et al. 1999; Boehr et al. 2009; van der Vaart 2015), metals (Wittung-Stafshede 2002; Wilson et al. 2004; Li et al. 2015), osmolytes (Henkels et al. 2001), and cations/anions (Hagihara et al. 1993; Daniels et al. 2014, 2015). Many metalloproteins retain strong metal binding even in the unfolded state, indicating the induced-fit mechanism, but some metalloproteins must form well-defined metal-binding sites before metal binding, indicating the conformational selection mechanism (Wilson et al. 2004).

Ligand-induced folding is also observed for proteins used in folding studies. Binding of a heme to apomyoglobin induces folding of the F-helix (Eliezer and Wright 1996). *Escherichia coli* DHFR has four native structures (Jennings et al. 1993), one of which binds the cofactor NADPH by conformational selection (Dunn et al. 1978).  $\alpha$ -LA and  $\text{Ca}^{2+}$ -binding lysozymes form molten globule states in the absence of metals (Kuwajima 1989; Arai and Kuwajima 2000; Nakao et al. 2005). Because ligand binding is necessary for their folding, these proteins can be classified as IDPs. The  $\text{Ca}^{2+}$ -binding site of  $\alpha$ -LA is organized at the transition state of folding from the molten globule to the native state (Kuwajima 1989).

### Ligand binding to globular proteins

Studies of free proteins and protein–ligand complexes have demonstrated that proteins can fluctuate into conformations resembling those of the bound form, even in the absence of ligands, suggesting the conformational selection mechanism (Boehr et al. 2009). Ligand-binding reactions have been reported to occur by the conformational selection mechanism for many globular proteins, including DHFR, cyclophilin A, adenylate kinase, ribonuclease A, triose phosphate isomerase, ubiquitin, calmodulin, *lac* repressor, immunoglobulin, NCBD, maltose binding protein, and trypsin-like proteases (James and Tawfik 2003, 2005; Henzler-Wildman and Kern 2007; Tang et al. 2007; Lange et al. 2008; Loria et al. 2008; Boehr et al. 2009; Kjaergaard et al. 2010; Ma and Nussinov 2010; Nashine et al. 2010; Changeux 2012; Vogt et al. 2014). Adenylate kinase was previously thought to bind a ligand by the induced-fit mechanism, but was recently shown to occur by conformational selection using state-of-the-art methodologies to observe protein dynamics (Henzler-Wildman et al. 2007b). Even characteristic enzyme motions detected during catalysis were found to be already present in the free enzyme with frequencies corresponding to the catalytic turnover rates, suggesting that the protein motions necessary for catalysis are

an intrinsic property of the enzyme (Eisenmesser et al. 2005; Boehr et al. 2006b). Furthermore, proteins from thermophilic bacteria have high optimal temperatures and function through flexible motions at high temperatures; however, catalytic activity is reduced at low temperatures by the repression of fluctuations (Závodszy et al. 1998; Wolf-Watz et al. 2004). These results support the importance of fluctuations and dynamics in protein function.

The dominance of the conformational selection mechanism indicates that the apparent ligand-binding rate of the weakly binding conformation is lower than the fast rate of conformational change, occurring on microsecond to millisecond time scales. Specific complementarity in shape and polarity between a protein and ligand is typically necessary for tight binding, and slight changes in protein conformations by fluctuations cause steric hindrance of binding.

The enzyme reaction cycle of *E. coli* DHFR goes through five intermediate states (Fierke et al. 1987; Schnell et al. 2004; Boehr et al. 2006a, b; Hammes-Schiffer and Benkovic 2006). NMR relaxation dispersion experiments showed that all five intermediate states can access conformations resembling those of the next step in the reaction cycle, indicating that all steps in the reaction cycle occur by the conformational selection mechanism (Boehr et al. 2006b). Surprisingly, the next state in the reaction cycle is already prepared as an excited state, and this efficient reaction cycle is encoded in the amino acid sequence of only 159 residues. Human DHFR has more flexible structures and higher activity than *E. coli* DHFR (Bhabha et al. 2013), supporting the link between protein dynamics and enzymatic catalysis. Thus, the intrinsically flexible nature of DHFR is essential for the enzymatic reaction, which is manifested in multiple native structures in the apo form of *E. coli* DHFR (Jennings et al. 1993).

Although conformational selection is the prevailing mechanism, the presence of high concentrations of a ligand can shift the binding mechanism toward the apparent induced-fit mechanism (Hammes et al. 2009; Greives and Zhou 2014). Moreover, the ideal induced-fit mechanism is observed for the ligand binding of HIV-1 protease, in which the conformational change from the semi-open to the closed form occurs after ligand binding, as the ligand-binding site is sequestered in the closed form and access of a ligand to the buried binding site is sterically prohibited (Hornak and Simmerling 2007). Recently, an increasing number of reports showed that proteins bind ligands by the sequential conformational-selection and induced-fit mechanism, in which primary binding by conformational selection is followed by conformational optimization by induced-fit (James and Tawfik 2003, 2005; Tang et al. 2007; Wlodarski and Zagrovic 2009; Wang et al. 2013a).

The physical origin of catalytically important collective motions in slow time scales (microsecond to millisecond) is the local hinge motions in fast time scales (picosecond to nanosecond) (Henzler-Wildman et al. 2007a), and is related

to the average time required to sample the configurations that are conducive to chemical reactions, which occur on femto-second to picosecond time scales (Ma and Nussinov 2010; Hammes et al. 2011).

### Stability–activity trade-off

The conformational selection mechanism assumes that equilibrium exists between the binding-incompetent and binding-competent conformations, while the induced-fit mechanism assumes that conformational change occurs from the weakly binding conformation to the tightly binding conformation. In both cases, the conformation with lower affinity for a ligand often corresponds to the native structure of the free protein. Thus, introduction of mutations that stabilize the native state of a free protein decelerates the transition to the conformation with higher affinity for a ligand, leading to a decrease in the ligand-binding rate and, for an enzyme, a decrease in catalytic activity. This observation is known as the “stability–activity trade-off” (Siddiqui 2017). From a structural perspective, the ligand-binding sites of proteins often contain conformational strain, hydrophobic surface exposure, and/or electrostatic repulsion caused by residues with the same charges. Thus, interactions in the native state of a protein are not optimized but can confer flexibility and fluctuation to lower-affinity conformations. Consequently, mutations at ligand-binding sites can optimize the stability of a native protein by removing unfavorable interactions, but concomitantly reduce the affinity for a ligand, leading to a stability–activity trade-off (Shoichet et al. 1995; Wang et al. 2002; Thomas et al. 2010; Yokota et al. 2010; Klesmith et al. 2017). In contrast, mutations that stabilize the binding-competent conformation and/or destabilize the binding-incompetent conformation—that is, the native structure of a free protein—should increase affinity and activity. IDPs are extreme cases of protein destabilization for tight binding.

The stability–activity trade-off can be viewed as competition between intramolecular and intermolecular interactions. By analogy, this corresponds to the interplay of local and non-local interactions in the folding of monomeric proteins. Previous studies reported that whereas favorable native-like local interactions can accelerate folding reactions (Viguera et al. 1997), further stabilization of secondary structure decelerates the folding rates (Chiti et al. 1999). Moreover, if non-native local structures are stable, they can be observed in kinetic folding intermediates, but not in the final native state (Forge et al. 2000). Thus, both the balance between local and non-local interactions and minimization of non-native local interactions are important for optimizing protein stability and folding speed (Muñoz and Serrano 1996). This suggests that the balance between intramolecular and intermolecular interactions is important for protein stability and activity.

## Unified mechanisms of protein folding and binding

We have described the mechanisms of coupled folding and binding of IDPs, folding of small and multi-subdomain proteins, folding of multimeric proteins, and ligand binding of globular proteins. All mechanisms are well-explained using the conformational selection and induced-fit mechanisms as well as the nucleation–condensation mechanism, which is intermediate between them, suggesting that we can integrate the understanding of folding and binding mechanisms of globular proteins and IDPs.

In general, reaction mechanisms are determined by competition of the fluxes of the conformational selection and induced-fit pathways. Accumulating evidence has shown that both the rate of conformational change and apparent rate of binding between interacting elements can determine reaction mechanisms. The former is affected by both the forward rate ( $P_{\text{weak}}$  to  $P_{\text{tight}}$ ) and reverse rate ( $P_{\text{tight}}$  to  $P_{\text{weak}}$ ), while the latter is affected by the second-order binding rate constant, first-order dissociation rate constant, and protein and ligand concentrations (see Fig. 1a). If the conformational change from  $P_{\text{weak}}$  to  $P_{\text{tight}}$  is faster than the ligand binding of  $P_{\text{weak}}$ , the reaction tends to occur through the apparent conformational selection mechanism. In contrast, if the ligand binding of  $P_{\text{weak}}$  is faster than the conformational change, the reaction tends to occur through the apparent induced-fit mechanism. Because the folding of monomeric proteins can be regarded as the binding of intramolecular segments accompanied by secondary structure formation, the rate of conformational change corresponds to the rate of secondary structure formation and the binding rate corresponds to the rate of binding between intramolecular segments. Instead of protein and ligand concentrations, the effective concentration between intramolecular segments affects the apparent binding rate in protein folding.

Because most IDPs contain unstable secondary structure elements, the secondary structure formation of free IDPs is slower than binding with partners. Therefore, coupled folding and binding reactions of IDPs are dominated by the induced-fit mechanism. In contrast, ligand-binding reactions of globular proteins are dominated by the conformational selection mechanism, as the conformational change is fast and because the weakly binding conformation of a protein has low affinity for a ligand because of the lack of specific complementarity between them. For both IDPs and globular proteins, if the rate of conformational change and rate of binding are comparable, the nucleation–condensation mechanism is observed.

In the case of folding reactions of monomeric proteins, the rate of secondary structure formation and binding rate of intramolecular segments can determine the folding mechanisms. Multi-subdomain proteins fold by the induced-fit (hydrophobic collapse) mechanism because the connection of interacting segments increases the effective concentration between them



and accelerates their mutual binding. Indeed, protein hydrophobic collapse can occur much faster than secondary structure formation. In contrast, small proteins typically fold by the nucleation–condensation mechanism. Secondary structure elements with short length and small numbers of hydrophobic residues in small proteins preclude the conformational selection (framework) mechanism and induced-fit (hydrophobic collapse) mechanism respectively. However, for both small and multi-subdomain proteins, the folding mechanism approaches the conformational selection (framework) mechanism or induced-fit (hydrophobic collapse) mechanism if secondary structure formation is fast or slow respectively. Generally,  $\alpha$ -helical proteins fold faster than  $\beta$ -sheet proteins. Thus, the secondary structure contents in native structures can determine the detailed folding mechanisms. Therefore, the folding mechanisms of both small and multi-subdomain proteins are essentially identical, except that lower or higher numbers of hydrophobic residues lead to nucleation or collapse of hydrophobic residues in small or multi-subdomain proteins respectively.

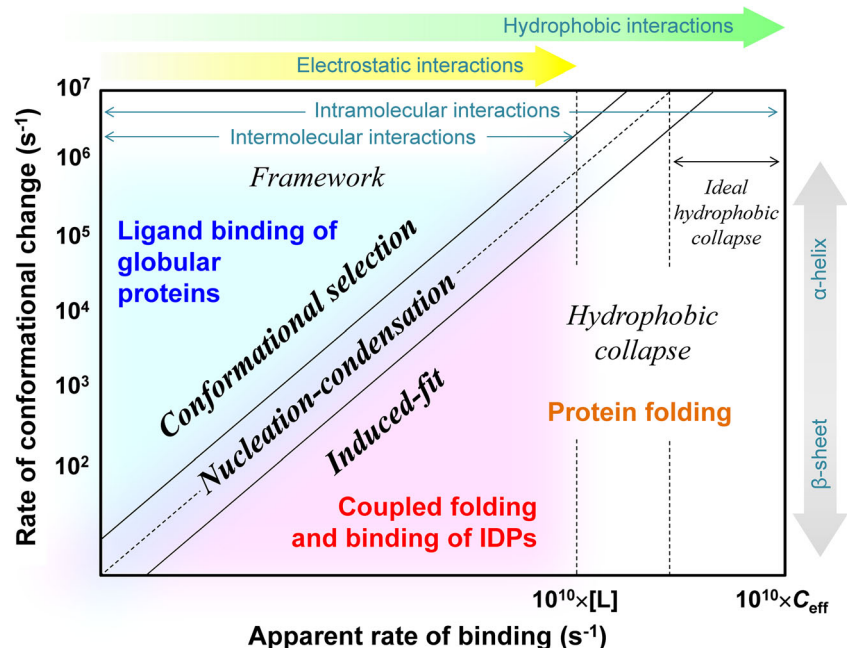
Both long-range electrostatic interactions and non-local hydrophobic interactions are essential for determining the binding rate, while local secondary structure propensities are important for determining the rate of conformational change. For intermolecular interactions, electrostatic attractions are more effective than hydrophobic interactions in diminishing the distance between interacting elements. In protein folding, non-local hydrophobic interactions are more important than local secondary structure propensities early in the folding reaction.

Folding mechanisms of multimeric proteins are consistent with those of monomeric globular proteins and IDPs. Depending on the stability of subunit structures, many

multimeric proteins fold by both the conformational selection mechanism, in which subunit assembly occurs after formation of molten globule-like monomeric intermediates, and induced-fit mechanism, in which inter-subunit interactions induce subunit folding.

Figure 5 summarizes the apparent dependence of the folding and binding mechanisms of globular proteins and IDPs on both the rate of conformational change (vertical axis) and apparent binding rate of interacting segments (horizontal axis). The rate limit of conformational change is the fastest folding rate of stable  $\alpha$ -helices and  $\beta$ -hairpins ( $\sim 10^7 \text{ s}^{-1}$ ). The upper limit of the apparent binding rate is estimated as  $\sim 10^{10} \text{ M}^{-1} \text{ s}^{-1}$  multiplied by a free ligand concentration  $[L]$  for intermolecular interactions and  $\sim 10^{10} \text{ M}^{-1} \text{ s}^{-1}$  multiplied by an effective concentration  $C_{\text{eff}}$  for intramolecular interactions. The reaction mechanism will be the induced-fit if the apparent binding is fast and conformational change is slow. In contrast, the reaction mechanism will be conformational selection if the apparent binding is slow and conformational change is fast. If both rates are comparable, the reaction mechanism will be the nucleation–condensation mechanism. Whereas coupled folding and binding reactions of IDPs are located on the lower right side in Fig. 5, ligand-binding reactions of globular proteins are located on the upper left side. Folding reactions of monomeric and multimeric proteins can be located over a wide region, but are mainly located on the lower right side. Fast binding (compaction) may require large numbers of hydrophobic residues, resulting in folding mainly by the induced-fit (hydrophobic collapse) mechanism. When the intramolecular binding rate exceeds the rate limit of conformational change, the folding reaction will occur by the ideal induced-fit (hydrophobic collapse) mechanism.

**Fig. 5** Apparent dependence of the folding and binding mechanisms of globular proteins and IDPs on both the rate of conformational change and apparent binding rate of interacting elements.  $[L]$  and  $C_{\text{eff}}$  denote the free ligand concentration and effective concentration for intramolecular interactions respectively. See text for details



In summary, the folding and binding mechanisms of globular proteins and IDPs obey the same general principle and can be integrated into a unified understanding. Elucidation of the mechanisms of folding and function of proteins and development of theoretical models to explain these mechanisms will profoundly impact the rational design of proteins for medicine and industry. The coarse-grained, statistical mechanical models explaining the mechanisms of folding reactions and allosteric transitions are promising for a unified theoretical description of folding and binding mechanisms of globular proteins and IDPs. Although many studies are needed to clarify these mechanisms, comparison of the mechanisms between various reactions will improve the understanding of each mechanism, and may enable a unified understanding to be developed that is applicable to all proteins.

**Acknowledgements** We thank Dr. Yuuki Hayashi (The University of Tokyo) for critical reading of the manuscript.

### Compliance with ethical standards

**Funding** This study was funded by Grants-in-Aids for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

**Conflict of interest** Munehito Arai declares that he has no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by the author.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

### References

- Akiyama S, Takahashi S, Ishimori K, Morishima I (2000) Stepwise formation of  $\alpha$ -helices during cytochrome *c* folding. *Nat Struct Biol* 7: 514–520
- Ali SA, Iwabuchi N, Matsui T, Hirota K, Kidokoro S, Arai M, Kuwajima K, Schuck P, Arisaka F (2003) Reversible and fast association equilibria of a molecular chaperone, gp57A, of bacteriophage T4. *Biophys J* 85:2606–2618
- Arai M, Dyson HJ, Wright PE (2010) Leu628 of the KIX domain of CBP is a key residue for the interaction with the MLL transactivation domain. *FEBS Lett* 584:4500–4504
- Arai M, Ferreon JC, Wright PE (2012) Quantitative analysis of multisite protein-ligand interactions by NMR: binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP. *J Am Chem Soc* 134:3792–3803
- Arai M, Hamel P, Kanaya E, Inaka K, Miki K, Kikuchi M, Kuwajima K (2000) Effect of an alternative disulfide bond on the structure, stability, and folding of human lysozyme. *Biochemistry* 39: 3472–3479
- Arai M, Ikura T, Semisotnov GV, Kihara H, Amemiya Y, Kuwajima K (1998) Kinetic refolding of  $\beta$ -lactoglobulin. Studies by synchrotron X-ray scattering, and circular dichroism, absorption and fluorescence spectroscopy. *J Mol Biol* 275:149–162
- Arai M, Inobe T, Maki K, Ikura T, Kihara H, Amemiya Y, Kuwajima K (2003a) Denaturation and reassembly of chaperonin GroEL studied by solution X-ray scattering. *Protein Sci* 12:672–680
- Arai M, Ito K, Inobe T, Nakao M, Maki K, Kamagata K, Kihara H, Amemiya Y, Kuwajima K (2002) Fast compaction of  $\alpha$ -lactalbumin during folding studied by stopped-flow X-ray scattering. *J Mol Biol* 321:121–132
- Arai M, Iwakura M (2005) Probing the interactions between the folding elements early in the folding of *Escherichia coli* dihydrofolate reductase by systematic sequence perturbation analysis. *J Mol Biol* 347:337–353
- Arai M, Iwakura M, Matthews CR, Bilsel O (2011) Microsecond subdomain folding in dihydrofolate reductase. *J Mol Biol* 410: 329–342
- Arai M, Kataoka M, Kuwajima K, Matthews CR, Iwakura M (2003b) Effects of the difference in the unfolded-state ensemble on the folding of *Escherichia coli* dihydrofolate reductase. *J Mol Biol* 329: 779–791
- Arai M, Kondrashkina E, Kayatekin C, Matthews CR, Iwakura M, Bilsel O (2007) Microsecond hydrophobic collapse in the folding of *Escherichia coli* dihydrofolate reductase, an  $\alpha/\beta$ -type protein. *J Mol Biol* 368:219–229
- Arai M, Kuwajima K (1996) Rapid formation of a molten globule intermediate in refolding of  $\alpha$ -lactalbumin. *Fold Des* 1:275–287
- Arai M, Kuwajima K (2000) Role of the molten globule state in protein folding. *Adv Protein Chem* 53:209–282
- Arai M, Maki K, Takahashi H, Iwakura M (2003c) Testing the relationship between foldability and the early folding events of dihydrofolate reductase from *Escherichia coli*. *J Mol Biol* 328: 273–288
- Arai M, Sugase K, Dyson HJ, Wright PE (2015) Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proc Natl Acad Sci U S A* 112:9614–9619
- Arora P, Hammes GG, Oas TG (2006) Folding mechanism of a multiple independently-folding domain protein: double B domain of protein A. *Biochemistry* 45:12312–12324
- Aroul-Selvam R, Hubbard T, Sasidharan R (2004) Domain insertions in protein structures. *J Mol Biol* 338:633–641
- Bah A, Forman-Kay JD (2016) Modulation of intrinsically disordered protein function by post-translational modifications. *J Biol Chem* 291:6696–6705
- Baldwin RL (2008) The search for folding intermediates and the mechanism of protein folding. *Annu Rev Biophys* 37:1–21
- Baldwin RL, Frieden C, Rose GD (2010) Dry molten globule intermediates and the mechanism of protein unfolding. *Proteins* 78: 2725–2737
- Banerjee PR, Deniz AA (2013) Shedding light on protein folding landscapes by single-molecule fluorescence. *Chem Soc Rev* 43:1172–1188
- Batey S, Nickson AA, Clarke J (2008) Studying the folding of multidomain proteins. *HFSP J* 2:365–377
- Berg OG, von Hippel PH (1985) Diffusion-controlled macromolecular interactions. *Annu Rev Biophys Biophys Chem* 14:131–160
- Bhabha G, Ekiert DC, Jennewein M, Zmasek CM, Tuttle LM, Kroon G, Dyson HJ, Godzik A, Wilson IA, Wright PE (2013) Divergent evolution of protein conformational dynamics in dihydrofolate reductase. *Nat Struct Mol Biol* 20:1243–1249
- Bilsel O, Matthews CR (2000) Barriers in protein folding reactions. *Adv Protein Chem* 53:153–207

- Bilsel O, Matthews CR (2006) Molecular dimensions and their distributions in early folding intermediates. *Curr Opin Struct Biol* 16:86–93
- Boehr DD, Dyson HJ, Wright PE (2006a) An NMR perspective on enzyme dynamics. *Chem Rev* 106:3055–3079
- Boehr DD, McElheny D, Dyson HJ, Wright PE (2006b) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* 313:1638–1642
- Boehr DD, Nussinov R, Wright PE (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* 5:789–796
- Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* 21:167–195
- Changeux JP (2012) Allostery and the Monod–Wyman–Changeux model after 50 years. *Annu Rev Biophys* 41:103–133
- Chaudhuri TK, Arai M, Terada TP, Ikura T, Kuwajima K (2000) Equilibrium and kinetic studies on folding of the authentic and recombinant forms of human  $\alpha$ -lactalbumin by circular dichroism spectroscopy. *Biochemistry* 39:15643–15651
- Chen T, Song J, Chan HS (2015) Theoretical perspectives on nonnative interactions and intrinsic disorder in protein folding and binding. *Curr Opin Struct Biol* 30:32–42
- Chiti F, Taddei N, Webster P, Hamada D, Fiaschi T, Ramponi G, Dobson CM (1999) Acceleration of the folding of acylphosphatase by stabilization of local secondary structure. *Nat Struct Biol* 6:380–387
- Chu WT, Clarke J, Shammas SL, Wang J (2017) Role of non-native electrostatic interactions in the coupled folding and binding of PUMA with Mcl-1. *PLoS Comput Biol* 13:e1005468
- Csermely P, Palotai R, Nussinov R (2010) Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci* 35:539–546
- Daggett V, Fersht A (2003) The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol* 4:497–502
- Daniels KG, Suo Y, Oas TG (2015) Conformational kinetics reveals affinities of protein conformational states. *Proc Natl Acad Sci U S A* 112:9352–9357
- Daniels KG, Tonthat NK, McClure DR, Chang YC, Liu X, Schumacher MA, Fierke CA, Schmidler SC, Oas TG (2014) Ligand concentration regulates the pathways of coupled protein folding and binding. *J Am Chem Soc* 136:822–825
- Deane CM, Saunders R (2011) The imprint of codons on protein structure. *Biotechnol J* 6:641–649
- Dill KA (1990) Dominant forces in protein folding. *Biochemistry* 29:7133–7155
- Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS (1995) Principles of protein folding—a perspective from simple exact models. *Protein Sci* 4:561–602
- Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4:10–19
- Dill KA, Ozkan SB, Shell MS, Weikl TR (2008) The protein folding problem. *Annu Rev Biophys* 37:289–316
- Dinner AR, Sali A, Smith LJ, Dobson CM, Karplus M (2000) Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem Sci* 25:331–339
- Dogan J, Gianni S, Jemth P (2014) The binding mechanisms of intrinsically disordered proteins. *Phys Chem Chem Phys* 16:6323–6331
- Dogan J, Mu X, Engstrom A, Jemth P (2013) The transition state structure for coupled binding and folding of disordered protein domains. *Sci Rep* 3:2076
- Dosnon M, Bonetti D, Morrone A, Erasles J, di Silvio E, Longhi S, Gianni S (2015) Demonstration of a folding after binding mechanism in the recognition between the measles virus NTAIL and X domains. *ACS Chem Biol* 10:795–802
- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, JS O, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. *J Mol Graph Model* 19:26–59
- Dunn SM, Batchelor JG, King RW (1978) Kinetics of ligand binding to dihydrofolate reductase: binary complex formation with NADPH and coenzyme analogues. *Biochemistry* 17:2356–2364
- Dürr E, Bosshard HR (2000) Folding of a three-stranded coiled coil. *Protein Sci* 9:1410–1415
- Dürr E, Jelesarov I, Bosshard HR (1999) Extremely fast folding of a very stable leucine zipper with a strengthened hydrophobic core and lacking electrostatic interactions between helices. *Biochemistry* 38:870–880
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208
- Dyson HJ, Wright PE (2016) Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. *J Biol Chem* 291:6714–6722
- Dyson HJ, Wright PE (2017) How does your protein fold? Elucidating the apomyoglobin folding pathway. *Acc Chem Res* 50:105–111
- Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 438:117–121
- Eliezer D, Wright PE (1996) Is apomyoglobin a molten globule? Structural characterization by NMR. *J Mol Biol* 263:531–538
- Englander SW, Mayne L (2014) The nature of protein folding pathways. *Proc Natl Acad Sci U S A* 111:15873–15880
- Espinoza-Fonseca LM (2009) Reconciling binding mechanisms of intrinsically disordered proteins. *Biochem Biophys Res Commun* 382:479–482
- Ferreira ME, Hermann S, Prochasson P, Workman JL, Berndt KD, Wright AP (2005) Mechanism of transcription factor recruitment by acidic activators. *J Biol Chem* 280:21779–21784
- Ferreon JC, Lee CW, Arai M, Martinez-Yamout MA, Dyson HJ, Wright PE (2009) Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proc Natl Acad Sci U S A* 106:6591–6596
- Fersht AR (1993) The sixth Datta lecture. Protein folding and stability: the pathway of folding of barnase. *FEBS Lett* 325:5–16
- Fersht AR (1997) Nucleation mechanisms in protein folding. *Curr Opin Struct Biol* 7:3–9
- Fersht AR, Matouschek A, Serrano L (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J Mol Biol* 224:771–782
- Fersht AR, Sato S (2004)  $\Phi$ -value analysis and the nature of protein-folding transition states. *Proc Natl Acad Sci U S A* 101:7976–7981
- Fierke CA, Johnson KA, Benkovic SJ (1987) Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from *Escherichia coli*. *Biochemistry* 26:4085–4092
- Forge V, Hoshino M, Kuwata K, Arai M, Kuwajima K, Batt CA, Goto Y (2000) Is folding of  $\beta$ -lactoglobulin non-hierarchical? Intermediate with native-like  $\beta$ -sheet and non-native  $\alpha$ -helix. *J Mol Biol* 296:1039–1051
- Forman-Kay JD, Mittag T (2013) From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* 21:1492–1499
- Fujiwara K, Arai M, Shimizu A, Ikeguchi M, Kuwajima K, Sugai S (1999) Folding-unfolding equilibrium and kinetics of equine  $\beta$ -lactoglobulin: equivalence between the equilibrium molten globule state and a burst-phase folding intermediate. *Biochemistry* 38:4455–4463
- Galvagnion C, Smith MT, Broom A, Vassall KA, Meglei G, Gaspar JA, Stathopoulos PB, Cheyne B, Meiering EM (2009) Folding and association of thermophilic dimeric and trimeric DsrEFH proteins: Tm0979 and Mth1491. *Biochemistry* 48:2891–2906
- Ganguly D, Otieno S, Waddell B, Iconaru L, Kriwacki RW, Chen J (2012) Electrostatically accelerated coupled binding and folding of intrinsically disordered proteins. *J Mol Biol* 422:674–684



- Gibbs EB, Showalter SA (2015) Quantitative biophysical characterization of intrinsically disordered proteins. *Biochemistry* 54:1314–1326
- Giri R, Morrone A, Toto A, Brunori M, Gianni S (2013) Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. *Proc Natl Acad Sci U S A* 110:14942–14947
- Gloss LM, Matthews CR (1998) Mechanism of folding of the dimeric core domain of *Escherichia coli* trp repressor: a nearly diffusion-limited reaction leads to the formation of an on-pathway dimeric intermediate. *Biochemistry* 37:15990–15999
- Gō N (1983) Theoretical studies of protein folding. *Annu Rev Biophys Bioeng* 12:183–210
- Goldbeck RA, Chen E, Klinger DS (2009) Early events, kinetic intermediates and the mechanism of protein folding in cytochrome *c*. *Int J Mol Sci* 10:1476–1499
- Goodsell DS, Olson AJ (2000) Structural symmetry and protein function. *Annu Rev Biophys Biomol Struct* 29:105–153
- Greives N, Zhou HX (2014) Both protein dynamics and ligand concentration can shift the binding mechanism between conformational selection and induced fit. *Proc Natl Acad Sci U S A* 111:10197–10202
- Haberz P, Arai M, Martinez-Yamout MA, Dyson HJ, Wright PE (2016) Mapping the interactions of adenoviral E1A proteins with the p160 nuclear receptor coactivator binding domain of CBP. *Protein Sci* 25:2256–2267
- Hagihara Y, Aimoto S, Fink AL, Goto Y (1993) Guanidine hydrochloride-induced folding of proteins. *J Mol Biol* 231:180–184
- Hammes GG, Benkovic SJ, Hammes-Schiffer S (2011) Flexibility, diversity, and cooperativity: pillars of enzyme catalysis. *Biochemistry* 50:10422–10430
- Hammes GG, Chang YC, Oas TG (2009) Conformational selection or induced fit: a flux description of reaction mechanism. *Proc Natl Acad Sci U S A* 106:13737–13741
- Hammes-Schiffer S, Benkovic SJ (2006) Relating protein motion to catalysis. *Annu Rev Biochem* 75:519–541
- Henkels CH, Kurz JC, Fierke CA, Oas TG (2001) Linked folding and anion binding of the *Bacillus subtilis* ribonuclease P protein. *Biochemistry* 40:2777–2789
- Henzler-Wildman K, Kern D (2007) Dynamic personalities of proteins. *Nature* 450:964–972
- Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D (2007a) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* 450:913–916
- Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Hubner CG, Kern D (2007b) Intrinsic motions along an enzymatic reaction trajectory. *Nature* 450:838–844
- Hornak V, Simmerling C (2007) Targeting structural flexibility in HIV-1 protease inhibitor binding. *Drug Discov Today* 12:132–138
- Hu W, Kan ZY, Mayne L, Englander SW (2016) Cytochrome *c* folds through foldon-dependent native-like intermediates in an ordered pathway. *Proc Natl Acad Sci U S A* 113:3809–3814
- Hu W, Walters BT, Kan ZY, Mayne L, Rosen LE, Marqusee S, Englander SW (2013) Stepwise protein folding at near amino acid resolution by hydrogen exchange and mass spectrometry. *Proc Natl Acad Sci U S A* 110:7684–7689
- Huang Y, Liu Z (2009) Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the “fly-casting” mechanism. *J Mol Biol* 393:1143–1159
- Iešmantavičius V, Dogan J, Jemth P, Teilum K, Kjaergaard M (2014) Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew Chem Int Ed Engl* 53:1548–1551
- Inaba K, Kobayashi N, Fersht AR (2000) Conversion of two-state to multi-state folding kinetics on fusion of two protein foldons. *J Mol Biol* 302:219–233
- Inanami T, Terada TP, Sasai M (2014) Folding pathway of a multidomain protein depends on its topology of domain connectivity. *Proc Natl Acad Sci U S A* 111:15969–15974
- Itoh K, Sasai M (2006) Flexibly varying folding mechanism of a nearly symmetrical protein: B domain of protein A. *Proc Natl Acad Sci U S A* 103:7298–7303
- Itoh K, Sasai M (2008) Cooperativity, connectivity, and folding pathways of multidomain proteins. *Proc Natl Acad Sci U S A* 105:13865–13870
- Itoh K, Sasai M (2009) Multidimensional theory of protein folding. *J Chem Phys* 130:145104
- Itoh K, Sasai M (2010) Entropic mechanism of large fluctuation in allosteric transition. *Proc Natl Acad Sci U S A* 107:7775–7780
- Itoh K, Sasai M (2011) Statistical mechanics of protein allostery: roles of backbone and side-chain structural fluctuations. *J Chem Phys* 134:125102
- Itzhaki LS, Otzen DE, Fersht AR (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation–condensation mechanism for protein folding. *J Mol Biol* 254:260–288
- Ivankov DN, Garbuzynskiy SO, Alm E, Plaxco KW, Baker D, Finkelstein AV (2003) Contact order revisited: influence of protein size on the folding rate. *Protein Sci* 12:2057–2062
- Jackson SE (1998) How do small single-domain proteins fold? *Fold Des* 3:R81–R91
- Jaenicke R (1987) Folding and association of proteins. *Prog Biophys Mol Biol* 49:117–237
- Jaenicke R, Lilie H (2000) Folding and association of oligomeric and multimeric proteins. *Adv Protein Chem* 53:329–401
- James LC, Tawfik DS (2003) Conformational diversity and protein evolution—a 60-year-old hypothesis revisited. *Trends Biochem Sci* 28:361–368
- James LC, Tawfik DS (2005) Structure and kinetics of a transient antibody binding intermediate reveal a kinetic discrimination mechanism in antigen recognition. *Proc Natl Acad Sci U S A* 102:12730–12735
- Jelesarov I, Dürr E, Thomas RM, Bosshard HR (1998) Salt effects on hydrophobic interaction and charge screening in the folding of a negatively charged peptide to a coiled coil (leucine zipper). *Biochemistry* 37:7539–7550
- Jennings PA, Finn BE, Jones BE, Matthews CR (1993) A reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli*: verification and refinement of a four-channel model. *Biochemistry* 32:3783–3789
- Jhong SR, Li CY, Sung TC, Lan YJ, Chang KJ, Chiang YW (2016) Evidence for an induced-fit process underlying the activation of apoptotic BAX by an intrinsically disordered BimBH3 peptide. *J Phys Chem B* 120:2751–2760
- Kamagata K, Arai M, Kuwajima K (2004) Unification of the folding mechanisms of non-two-state and two-state proteins. *J Mol Biol* 339:951–965
- Kamagata K, Kuwajima K (2006) Surprisingly high correlation between early and late stages in non-two-state protein folding. *J Mol Biol* 357:1647–1654
- Kathuria SV, Kayatekin C, Barrea R, Kondrashkina E, Graceffa R, Guo L, Nobrega RP, Chakravarthy S, Matthews CR, Irving TC, Bilsel O (2014) Microsecond barrier-limited chain collapse observed by time-resolved FRET and SAXS. *J Mol Biol* 426:1980–1994
- Khorasanizadeh S, Peters ID, Roder H (1996) Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nat Struct Biol* 3:193–205
- Kiefhaber T, Bachmann A, Jensen KS (2012) Dynamics and mechanisms of coupled protein folding and binding reactions. *Curr Opin Struct Biol* 22:21–29



- Kim PS, Baldwin RL (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu Rev Biochem* 51:459–489
- Kim PS, Baldwin RL (1990) Intermediates in the folding reactions of small proteins. *Annu Rev Biochem* 59:631–660
- Kimura T, Akiyama S, Uzawa T, Ishimori K, Morishima I, Fujisawa T, Takahashi S (2005) Specifically collapsed intermediate in the early stage of the folding of ribonuclease A. *J Mol Biol* 350:349–362
- Kjaergaard M, Teilum K, Poulsen FM (2010) Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc Natl Acad Sci U S A* 107:12535–12540
- Klesmith JR, Bacik JP, Wrenbeck EE, Michalczyk R, Whitehead TA (2017) Trade-offs between enzyme fitness and solubility illuminated by deep mutational scanning. *Proc Natl Acad Sci U S A* 114:2265–2270
- Koshland DE Jr, Nemethy G, Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* 5:365–385
- Kubelka J, Hofrichter J, Eaton WA (2004) The protein folding ‘speed limit’. *Curr Opin Struct Biol* 14:76–88
- Kumar S, Ma B, Tsai CJ, Sinha N, Nussinov R (2000) Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* 9:10–19
- Kurnik M, Hedberg L, Danielsson J, Oliveberg M (2012) Folding without charges. *Proc Natl Acad Sci U S A* 109:5705–5710
- Kuwajima K (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* 6:87–103
- Kuwajima K, Yamaya H, Sugai S (1996) The burst-phase intermediate in the refolding of  $\beta$ -lactoglobulin studied by stopped-flow circular dichroism and absorption spectroscopy. *J Mol Biol* 264:806–822
- Kuwata K, Shastry R, Cheng H, Hoshino M, Batt CA, Goto Y, Roder H (2001) Structural and kinetic characterization of early folding events in  $\beta$ -lactoglobulin. *Nat Struct Biol* 8:151–155
- Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* 320:1471–1475
- Lee CW, Ferreon JC, Ferreon AC, Arai M, Wright PE (2010) Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc Natl Acad Sci U S A* 107:19290–19295
- Li W, Wang J, Zhang J, Wang W (2015) Molecular simulations of metal-coupled protein folding. *Curr Opin Struct Biol* 30:25–31
- Lindstrom I, Dogan J (2017) Native hydrophobic binding interactions at the transition state for association between the TAZ1 domain of CBP and the disordered TAD-STAT2 are not a requirement. *Biochemistry* 56:4145–4153
- Liu SQ, Ji XL, Tao Y, Tan DY, Zhang KQ, Fu YX (2012) Protein folding, binding and energy landscape: a synthesis. In: Kaumaya P (ed) *Protein Engineering*. In Tech, Rijeka, pp 207–252
- Loria JP, Berlow RB, Watt ED (2008) Characterization of enzyme motions by solution NMR relaxation dispersion. *Acc Chem Res* 41:214–221
- Lu Q, HP L, Wang J (2007) Exploring the mechanism of flexible biomolecular recognition with single molecule dynamics. *Phys Rev Lett* 98:128105
- Ma B, Kumar S, Tsai CJ, Nussinov R (1999) Folding funnels and binding mechanisms. *Protein Eng* 12:713–720
- Ma B, Nussinov R (2010) Enzyme dynamics point to stepwise conformational selection in catalysis. *Curr Opin Chem Biol* 14:652–659
- Matagne A, Dobson CM (1998) The folding process of hen lysozyme: a perspective from the ‘new view’. *Cell Mol Life Sci* 54:363–371
- Matsui T, Griniuvienė B, Goldberg E, Tsugita A, Tanaka N, Arisaka F (1997) Isolation and characterization of a molecular chaperone, gp57A, of bacteriophage T4. *J Bacteriol* 179:1846–1851
- Matthews CR (1993) Pathways of protein folding. *Annu Rev Biochem* 62:653–683
- Mayor U, Guydosh NR, Johnson CM, Grossmann JG, Sato S, Jas GS, Freund SM, Alonso DO, Daggett V, Fersht AR (2003) The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* 421:863–867
- Meisner WK, Sosnick TR (2004) Fast folding of a helical protein initiated by the collision of unstructured chains. *Proc Natl Acad Sci U S A* 101:13478–13482
- Mendoza JA, Rogers E, Lorimer GH, Horowitz PM (1991) Unassisted refolding of urea unfolded rhodanese. *J Biol Chem* 266:13587–13591
- Meszaros B, Tompa P, Simon I, Dosztanyi Z (2007) Molecular principles of the interactions of disordered proteins. *J Mol Biol* 372:549–561
- Meyer EE, Rosenberg KJ, Israelachvili J (2006) Recent progress in understanding hydrophobic interactions. *Proc Natl Acad Sci U S A* 103:15739–15746
- Milles S, Mercadante D, Aramburu IV, Jensen MR, Banterle N, Koehler C, Tyagi S, Clarke J, Shammas SL, Blackledge M, Gräter F, Lemke EA (2015) Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* 163:734–745
- Mittermaier A, Kay LE (2006) New tools provide new insights in NMR studies of protein dynamics. *Science* 312:224–228
- Mizuguchi M, Arai M, Ke Y, Nitta K, Kuwajima K (1998) Equilibrium and kinetics of the folding of equine lysozyme studied by circular dichroism spectroscopy. *J Mol Biol* 283:265–277
- Mollica L, Bessa LM, Hanouille X, Jensen MR, Blackledge M, Schneider R (2016) Binding mechanisms of intrinsically disordered proteins: theory, simulation, and experiment. *Front Mol Biosci* 3:52
- Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* 12:88–118
- Muñoz V, Cerminara M (2016) When fast is better: protein folding fundamentals and mechanisms from ultrafast approaches. *Biochem J* 473:2545–2559
- Muñoz V, Eaton WA (1999) A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc Natl Acad Sci U S A* 96:11311–11316
- Muñoz V, Serrano L (1994) Elucidating the folding problem of helical peptides using empirical parameters. *Nat Struct Biol* 1:399–409
- Muñoz V, Serrano L (1996) Local versus nonlocal interactions in protein folding and stability—an experimentalist’s point of view. *Fold Des* 1:R71–R77
- Nakamura T, Makabe K, Tomoyori K, Maki K, Mukaiyama A, Kuwajima K (2010) Different folding pathways taken by highly homologous proteins, goat  $\alpha$ -lactalbumin and canine milk lysozyme. *J Mol Biol* 396:1361–1378
- Nakao M, Maki K, Arai M, Koshiba T, Nitta K, Kuwajima K (2005) Characterization of kinetic folding intermediates of recombinant canine milk lysozyme by stopped-flow circular dichroism. *Biochemistry* 44:6685–6692
- Narayanan R, Ganesh OK, Edison AS, Hagen SJ (2008) Kinetics of folding and binding of an intrinsically disordered protein: the inhibitor of yeast aspartic proteinase YPrA. *J Am Chem Soc* 130:11477–11485
- Nashine VC, Hammes-Schiffer S, Benkovic SJ (2010) Coupled motions in enzyme catalysis. *Curr Opin Chem Biol* 14:644–651
- Neet KE, Timm DE (1994) Conformational stability of dimeric proteins: quantitative studies by equilibrium denaturation. *Protein Sci* 3:2167–2174
- Neira JL, Rico M (1997) Folding studies on ribonuclease A, a model protein. *Fold Des* 2:R1–R11
- Nishimura C (2017) Folding of apomyoglobin: analysis of transient intermediate structure during refolding using quick hydrogen deuterium exchange and NMR. *Proc Jpn Acad Ser B Phys Biol Sci* 93:10–27

- Noel AF, Bilsel O, Kundu A, Wu Y, Zitzewitz JA, Matthews CR (2009) The folding free-energy surface of HIV-1 protease: insights into the thermodynamic basis for resistance to inhibitors. *J Mol Biol* 387:1002–1016
- O'Brien EP, Ciryam P, Vendruscolo M, Dobson CM (2014) Understanding the influence of codon translation rates on cotranslational protein folding. *Acc Chem Res* 47:1536–1544
- Oliveberg M, Fersht AR (1996) A new approach to the study of transient protein conformations: the formation of a semiburied salt link in the folding pathway of barnase. *Biochemistry* 35:6795–6805
- Onitsuka M, Kamikubo H, Yamazaki Y, Kataoka M (2008) Mechanism of induced folding: both folding before binding and binding before folding can be realized in staphylococcal nuclease mutants. *Proteins* 72:837–847
- Ou L, Matthews M, Pang X, Zhou HX (2017) The dock-and-coalesce mechanism for the association of a WASP disordered region with the Cdc42 GTPase. *FEBS J* 284(20):3381–3391
- Plaxco KW, Simons KT, Baker D (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277:985–994
- Plaxco KW, Simons KT, Ruczinski I, Baker D (2000) Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. *Biochemistry* 39:11177–11183
- Pratt LR, Chandler D (1986) Theoretical and computational studies of hydrophobic interactions. *Methods Enzymol* 127:48–63
- Ptitsyn OB (1995) Molten globule and protein folding. *Adv Protein Chem* 47:83–229
- Radhakrishnan I, Perez-Alvarado GC, Dyson HJ, Wright PE (1998) Conformational preferences in the Ser133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Lett* 430:317–322
- Raschke TM, Marqusee S (1997) The kinetic folding intermediate of ribonuclease H resembles the acid molten globule and partially unfolded molecules detected under native conditions. *Nat Struct Biol* 4:298–304
- Redfield C, Schulman BA, Milhollen MA, Kim PS, Dobson CM (1999)  $\alpha$ -Lactalbumin forms a compact molten globule in the absence of disulfide bonds. *Nat Struct Biol* 6:948–952
- Rentzperis D, Jonsson T, Sauer RT (1999) Acceleration of the refolding of arc repressor by nucleic acids and other polyanions. *Nat Struct Biol* 6:569–573
- Riechmann L, Lavenir I, de Bono S, Winter G (2005) Folding and stability of a primitive protein. *J Mol Biol* 348:1261–1272
- Robinson CR, Sauer RT (2000) Striking stabilization of *Arc* repressor by an engineered disulfide bond. *Biochemistry* 39:12494–12502
- Rogers JM, Oleinikovas V, Shammass SL, Wong CT, De Sancho D, Baker CM, Clarke J (2014) Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *Proc Natl Acad Sci U S A* 111:15420–15425
- Rosen LE, Kathuria SV, Matthews CR, Bilsel O, Marqusee S (2014) Non-native structure appears in microseconds during the folding of *E. coli* RNase H. *J Mol Biol* 427:443–453
- Rumfeldt JA, Galvagnion C, Vassall KA, Meiering EM (2008) Conformational stability and folding mechanisms of dimeric proteins. *Prog Biophys Mol Biol* 98:61–84
- Sadqi M, Lapidus LJ, Muñoz V (2003) How fast is protein hydrophobic collapse? *Proc Natl Acad Sci U S A* 100:12117–12122
- Saeki K, Arai M, Yoda T, Nakao M, Kuwajima K (2004) Localized nature of the transition-state structure in goat  $\alpha$ -lactalbumin folding. *J Mol Biol* 341:589–604
- Sasai M, Chikenji G, Terada TP (2016) Cooperativity and modularity in protein folding. *Biophys Physicobiol* 13:281–293
- Sato S, Religa TL, Daggett V, Fersht AR (2004) Testing protein-folding simulations by experiment: B domain of protein A. *Proc Natl Acad Sci U S A* 101:6952–6956
- Schneider R, Maurin D, Communie G, Kragelj J, Hansen DF, Ruigrok RW, Jensen MR, Blackledge M (2015) Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion NMR. *J Am Chem Soc* 137:1220–1229
- Schnell JR, Dyson HJ, Wright PE (2004) Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu Rev Biophys Biomol Struct* 33:119–140
- Schreiber G (2002) Kinetic studies of protein-protein interactions. *Curr Opin Struct Biol* 12:41–47
- Shakhnovich E (2006) Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. *Chem Rev* 106:1559–1588
- Shammass SL, Crabtree MD, Dahal L, Wicky BI, Clarke J (2016) Insights into coupled folding and binding mechanisms from kinetic studies. *J Biol Chem* 291:6689–6695
- Shoemaker BA, Portman JJ, Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci U S A* 97:8868–8873
- Shoichet BK, Baase WA, Kuroki R, Matthews BW (1995) A relationship between protein stability and protein function. *Proc Natl Acad Sci U S A* 92:452–456
- Siddiqui KS (2017) Defying the activity-stability trade-off in enzymes: taking advantage of entropy to enhance activity and thermostability. *Crit Rev Biotechnol* 37:309–322
- Song J, Guo LW, Muradov H, Artemyev NO, Ruoho AE, Markley JL (2008) Intrinsically disordered  $\gamma$ -subunit of cGMP phosphodiesterase encodes functionally relevant transient secondary and tertiary structure. *Proc Natl Acad Sci U S A* 105:1505–1510
- Sosnick TR, Barrick D (2011) The folding of single domain proteins—have we reached a consensus? *Curr Opin Struct Biol* 21:12–24
- Spolar RS, Record MT, Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA. *Science* 263:777–784
- Steward A, Chen Q, Chapman RI, Borgia MB, Rogers JM, Wojtala A, Wilmanns M, Clarke J (2012) Two immunoglobulin tandem proteins with a linking  $\beta$ -strand reveal unexpected differences in cooperativity and folding pathways. *J Mol Biol* 416:137–147
- Sugase K, Dyson HJ, Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447:1021–1025
- Svensson AK, Bilsel O, Kondrashkina E, Zitzewitz JA, Matthews CR (2006) Mapping the folding free energy surface for metal-free human Cu,Zn superoxide dismutase. *J Mol Biol* 364:1084–1102
- Takahashi H, Arai M, Takenawa T, Sota H, Xie QH, Iwakura M (2007) Stabilization of hyperactive dihydrofolate reductase by cyanocysteine-mediated backbone cyclization. *J Biol Chem* 282:9420–9429
- Takahashi S, Kamagata K, Oikawa H (2016) Where the complex things are: single molecule and ensemble spectroscopic investigations of protein folding dynamics. *Curr Opin Struct Biol* 36:1–9
- Takahashi S, Yeh SR, Das TK, Chan CK, Gottfried DS, Rousseau DL (1997) Folding of cytochrome *c* initiated by submillisecond mixing. *Nat Struct Biol* 4:44–50
- Tang C, Schwieters CD, Clore GM (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* 449:1078–1082
- Thomas VL, McReynolds AC, Shoichet BK (2010) Structural bases for stability–function tradeoffs in antibiotic resistance. *J Mol Biol* 396:47–59
- Tomba P (2012) Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci* 37:509–516
- Topping TB, Gloss LM (2004) Stability and folding mechanism of mesophilic, thermophilic and hyperthermophilic archaeal histones: the importance of folding intermediates. *J Mol Biol* 342:247–260

- Tsai CD, Ma B, Kumar S, Wolfson H, Nussinov R (2001) Protein folding: binding of conformationally fluctuating building blocks via population selection. *Crit Rev Biochem Mol Biol* 36:399–433
- Uversky VN, Fink AL (2002) The chicken–egg scenario of protein folding revisited. *FEBS Lett* 515:79–83
- van der Vaart A (2015) Coupled binding–binding–folding: the complex conformational dynamics of protein–DNA binding studied by atomistic molecular dynamics simulations. *Biochim Biophys Acta* 1850:1091–1098
- Viguera AR, Villegas V, Aviles FX, Serrano L (1997) Favourable native-like helical local interactions can accelerate protein folding. *Fold Des* 2:23–33
- Vogt AD, Pozzi N, Chen Z, Di Cera E (2014) Essential role of conformational selection in ligand binding. *Biophys Chem* 186:13–21
- Wako H, Saitô N (1978a) Statistical mechanical theory of the protein conformation. I. General considerations and the application to homopolymers. *J Phys Soc Jpn* 44:1931–1938
- Wako H, Saitô N (1978b) Statistical mechanical theory of the protein conformation. II. Folding pathway for protein. *J Phys Soc Jpn* 44:1939–1945
- Waldburger CD, Jonsson T, Sauer RT (1996) Barriers to protein folding: formation of buried polar interactions is a slow step in acquisition of structure. *Proc Natl Acad Sci U S A* 93:2629–2634
- Wang Q, Zhang P, Hoffman L, Tripathi S, Homouz D, Liu Y, Waxham MN, Cheung MS (2013a) Protein recognition and selection through conformational and mutually induced fit. *Proc Natl Acad Sci U S A* 110:20545–20550
- Wang T, Lau WL, DeGrado WF, Gai F (2005) T-jump infrared study of the folding mechanism of coiled-coil GCN4-p1. *Biophys J* 89:4180–4187
- Wang X, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol* 320:85–95
- Wang Y, Chu X, Longhi S, Roche P, Han W, Wang E, Wang J (2013b) Multiscaled exploration of coupled folding and binding of an intrinsically disordered molecular recognition element in measles virus nucleoprotein. *Proc Natl Acad Sci U S A* 110:E3743–E3752
- Wedemeyer WJ, Welker E, Narayan M, Scheraga HA (2000) Disulfide bonds and protein folding. *Biochemistry* 39:7032
- Wilson CJ, Apiyo D, Wittung-Stafshede P (2004) Role of cofactors in metalloprotein folding. *Q Rev Biophys* 37:285–314
- Winkler JR (2004) Cytochrome *c* folding dynamics. *Curr Opin Chem Biol* 8:169–174
- Wittung-Stafshede P (2002) Role of cofactors in protein folding. *Acc Chem Res* 35:201–208
- Wlodarski T, Zagrovic B (2009) Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin. *Proc Natl Acad Sci U S A* 106:19346–19351
- Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D (2004) Linkage between dynamics and catalysis in a thermophilic–mesophilic enzyme pair. *Nat Struct Mol Biol* 11:945–949
- Wong ET, Na D, Gsponer J (2013) On the importance of polar interactions for complexes containing intrinsically disordered proteins. *PLoS Comput Biol* 9:e1003192
- Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: reassessing the protein structure–function paradigm. *J Mol Biol* 293:321–331
- Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* 19:31–38
- Wright PE, Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* 16:18–29
- Wu Y, Kondrashkina E, Kayatekin C, Matthews CR, Bilsel O (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein. *Proc Natl Acad Sci U S A* 105:13367–13372
- Yoda T, Saito M, Arai M, Horii K, Tsumoto K, Matsushima M, Kumagai I, Kuwajima K (2001) Folding–unfolding of goat  $\alpha$ -lactalbumin studied by stopped-flow circular dichroism and molecular dynamics simulations. *Proteins* 42:49–65
- Yokota A, Takahashi H, Takenawa T, Arai M (2010) Probing the roles of conserved arginine-44 of *Escherichia coli* dihydrofolate reductase in its function and stability by systematic sequence perturbation analysis. *Biochem Biophys Res Commun* 391:1703–1707
- Závodszky P, Kardos J, Svingor, Petsko GA (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc Natl Acad Sci U S A* 95:7406–7411
- Zhou HX, Bates PA (2013) Modeling protein association mechanisms and kinetics. *Curr Opin Struct Biol* 23:887–893
- Zitzewitz JA, Ibarra-Molero B, Fishel DR, Terry KL, Matthews CR (2000) Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system. *J Mol Biol* 296:1105–1116
- Zor T, Mayr BM, Dyson HJ, Montminy MR, Wright PE (2002) Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. *J Biol Chem* 277:42241–42248