JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Insights on the Role of (Dis)order from Protein−Protein Interaction Linear Free-Energy Relationships

Meher K. Prakash*

Chemistry and Applied Biosciences Division, ETH Zürich, USI Campus, CH 6900 Lugano, Switzerland

Supporting Information

ABSTRACT: Protein−protein interactions (PPIs) are remarkably diverse and form the basis for various cellular functions. PPIs can be classified as ordered or disordered; the disordered ones do not have a well-defined structure prior to association, which is an exception to the conventional structure−function relationship. The occurrence of disordered proteins in functional roles is not explained by the conventional structure−function paradigm, and at present there is no clear understanding of the differences between the natures of these two PPIs. In this work, we studied the relationship between the kinetics and thermodynamics in PPIs to provide insights into the latter, with possible implications for the former. Analyzing the experimental data for various protein complexes, we found linear free-energy behavior with a striking kinetic difference between these two types of interactions. Binding affinities of (dis)ordered proteins are correlated with their (association) dissociation rates. Our observation, combined with the correspondence between biological activity and affinity, suggests that selection pressure on the dissociation or association kinetics in a functional context necessitates the presence of (dis)order in the structure.

Most cellular functions in control and regulation, such as intracellular signaling, transcription, and replication, are mediated by protein−protein interactions (PPIs).[1−3] Because of their importance, PPIs have been studied experimentally and theoretically at various levels of detail. At the systems level, these studies include the networks of interactions that cascade into different cellular events.[2] At the level of individual proteins, the focus is on the atomic details of interaction from the structure of the complex or the effects of site-directed mutagenesis on the kinetics and thermodynamics. The idea underlying all these efforts is that a better understanding of PPIs can provide fundamental insights into the physicochemical interactions governing molecular recognition in biological processes in health and disease and also offer sufficient guidance for the design of new proteins/enzymes. These considerations make studies of PPIs interesting from the biological, biochemical, and biophysical perspectives.

Although it is commonly believed that a protein has to be in the native fold for it to be functional, this is not true in general. About 30% of eukaryotic proteins are composed of proteins that do not have a well-defined structure prior to interaction with their binding partner. These intrinsically disordered proteins (IDPs),[4−6] which attain their structures during the course of interacting with their partners, participate promiscuously in critical cellular functions such as signaling and regulation. Only in recent years has the combination of folding and binding in IDPs been investigated experimentally in NMR studies[7] and at the single-molecule level[8] as well as in theoretical[9] and computational works.[10,11]

Ordered and disordered proteins have been distinguished in terms of several physical and chemical features, such as geometry and the residue composition of interfaces.[12] Ordered proteins have a higher content of hydrophobic residues, while the disordered ones are rich in polar and charged groups. Experimental studies have suggested that IDPs have several advantages, including specificity without excessive binding strength, increased speed of interaction, and binding promiscuity, among others.[13,14] Theoretical studies such as the "fly-casting" mechanism have suggested that the increased flexibility of the polypeptide chain provides a better capture radius for interactions.[9] However, understanding of the occurrence of disordered proteins and the nature of both ordered PPIs and disordered ones (wherein at least one of the interacting partners is disordered) is not complete.

The binding affinity of interacting proteins corresponds to their biological activity. However, the interactions between the complex-forming proteins are dynamic and go on continuously inside a cell. Therefore, the nature of interacting surfaces and thermodynamic affinities cannot give a complete description of these interactions. Kinetics combined with mutagenesis has been instrumental in providing insights into the mechanisms of protein folding.[15] Therefore, here we studied the relationship between the kinetics of PPIs and their affinities. We show a qualitative difference in the nature of these PPIs from the viewpoint of kinetics. Using this observation, we comment on the role of (dis)order in biological contexts and in the evolution of PPIs.

We first studied the kinetics of disordered PPIs by examining the effects of mutations on three experimental observables: the rate constants for association (k_{on}) and dissociation (k_{off}) and the binding constant $K_D = k_{off}/k_{on}$. The values of k_{on} and k_{off} reflect the kinetics of the process, and K_D governs the thermodynamics of association. The SH2 domain is a structurally conserved domain in several intracellular signaling proteins. It interacts with tyrosine-phosphorylated proteins to achieve cellular communication. Figure 1A shows the kinetics of SH2 domain interactions with disordered peptides containing phosphotyrosine.[16] Bone morphogenetic proteins (BMPs) regulate many developmental processes during embryogenesis and tissue homeostasis. BMP

Received: February 17, 2011
Published: June 07, 2011

ACS Publications © 2011 American Chemical Society

dx.doi.org/10.1021/ja201500z | J. Am. Chem. Soc. 2011, 133, 9976–9979

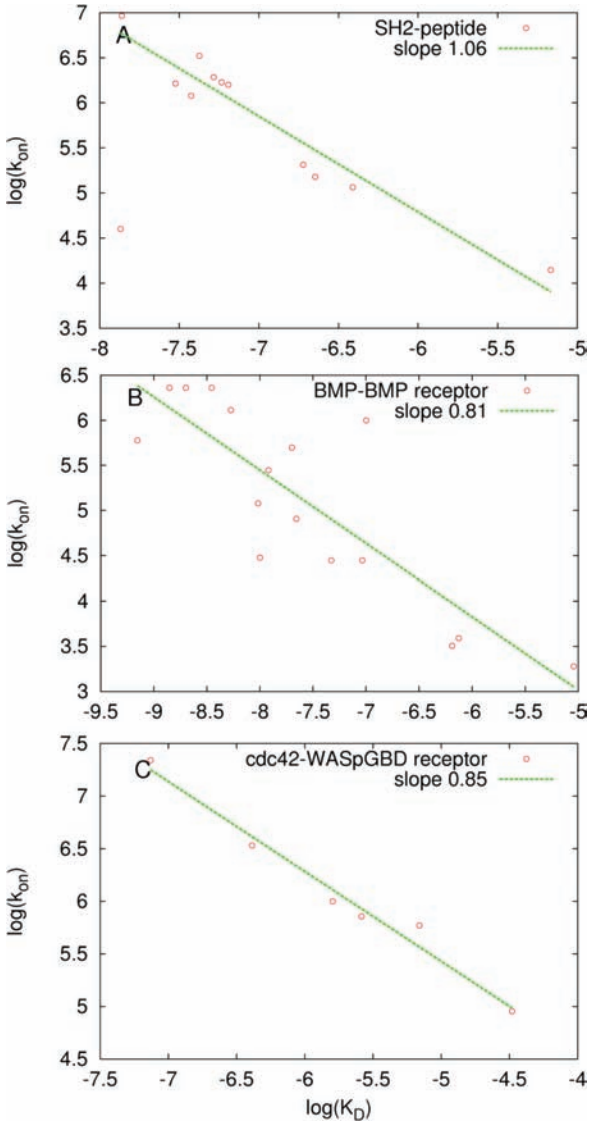

Figure 1. Plots of $\log(k_{on})$ vs $\log(K_D)$ for disordered interactions: (A) SH2 domain interacting with disordered peptides[16] (r^2 = 0.93); (B) BMP−BMP receptor interactions[17] (r^2 = 0.71); (C) WASp−Cdc42 interactions (r^2 = 0.98).[18] Logarithms are to the base 10, and k_{on}, k_{off}, and K_D have units of M^{-1} s^{-1}, s^{-1}, and M, respectively.

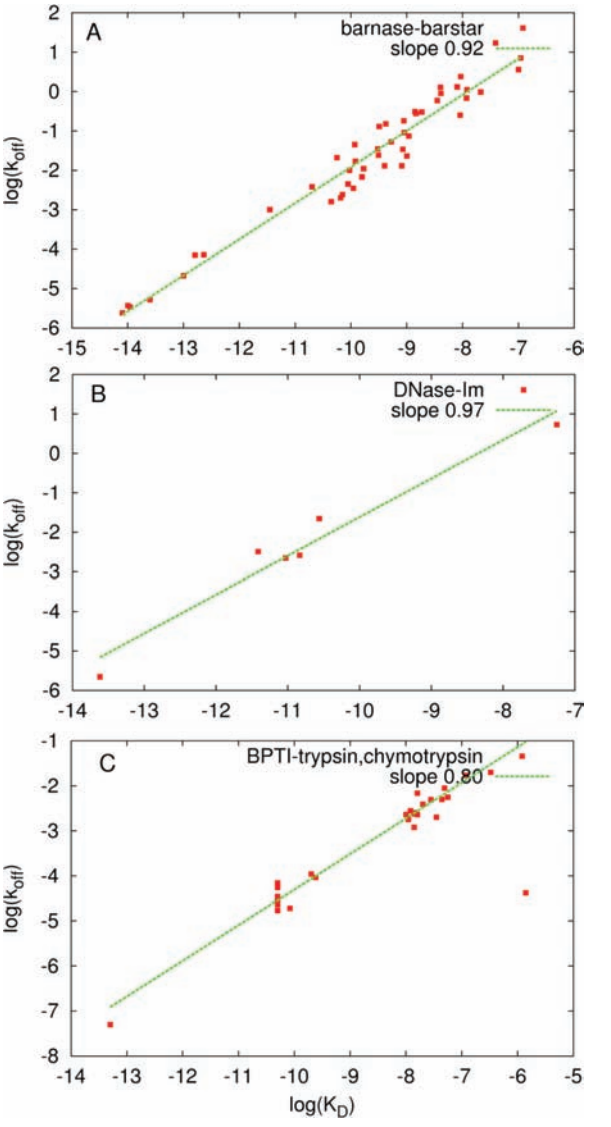

Figure 2. Plots of $\log(k_{off})$ vs $\log(K_D)$ for ordered interactions: (A) barnase−barstar binding[22] (r^2 = 0.95); (B) DNase−Im binding (r^2 = 0.96); (C) BPTI interactions with trypsin and chymotrypsin (r^2 = 0.97).[24] Units are the same as in Figure 1.

receptors are highly flexible, and several of their loops are disordered in the free state, making the interactions in the BMP superfamily promiscuous. Figure 1B shows the kinetics of BMP−BMP receptor binding complexes.[17] Wiskott−Aldrich Syndrome protein (WASp) interacts with Rho family GTPases in the transduction of signals from the cell surface to the actin cytoskeleton.[18] The interacting region of WASp is disordered, and the kinetics of its interactions with Cdc42 are shown in Figure 1C. The data in Figure 1show strong correlations between $\log(k_{on})$ and $\log(K_D)$, with slopes of −1.06, −0.81, and −0.85 for the SH2 domain,[19] BMP,[20] and WASp,[21] which form complexes through disordered interactions. These slopes surprisingly suggest that K_D is, to *zeroth-order*, independent of k_{off} in disordered PPIs [plots of $\log(k_{on})$ vs $\log(K_D)$ for disordered proteins are given in the Supporting Information (SI)].

On the other hand, data for ordered complexes are shown in Figure 2. Barnase is an extracelluar RNase. Its intracellular activity can be lethal to the cell and is halted by its binding to barstar. The kinetics of binding to barstar are shown in Figure 2A.[22] The kinetics of colicin endonucleases (DNases) binding to immunity receptors (Im) is shown in Figure 2B.[23] Cell suicide from DNases produced in bacteria under environmental stress is prevented by binding to Im proteins. Bovine pancreatic trypsin inhibitor (BPTI) acts on serine proteases that are prematurely activated in the pancreas, preventing self-digestion.[24] The kinetics of BPTI binding to its receptors trypsin and chymotrypsin are shown in Figure 2C. The plots of $\log(k_{off})$ versus $\log(K_D)$ in Figure 2A−C have slopes of 0.92, 0.97, and 0.80, respectively.

The observations from Figures 1 and 2 are quite striking. First, these data show linear free-energy relationships (LFERs)[25] correlating the kinetics and thermodynamics of protein−protein interactions. Second, K_D has a selective dependence on k_{on} and k_{off} in the disordered and ordered PPIs, respectively. This qualitative difference between the ordered and disordered interactions

dx.doi.org/10.1021/ja201500z | J. Am. Chem. Soc. 2011, 133, 9976–9979

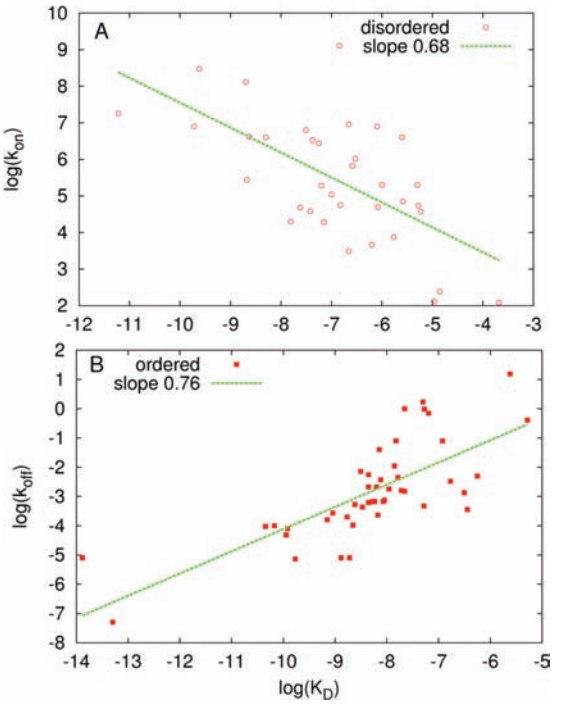

Figure 3. Plots showing correlations of (A) $\log(k_{on})$ vs $\log(K_D)$ for disordered proteins ($r^2 = 0.39$) and (B) $\log(k_{off})$ vs $\log(K_D)$ for ordered proteins ($r^2 = 0.54$). Units are the same as in Figure 1.

is surprising. LFERs have been observed in physical organic chemistry and in studies of binding of enzymes and small ligands,[15] and they are helpful in understanding the mechanisms of reactions. In a different context, a differential dependence in the kinetics was observed for ligands binding to hemoglobin.[26] The binding affinity of hemoglobin for small ligands such as NO, CO, and O_2 is governed by k_{off}, while that with bulky isocyanides (for which conformational changes are required) is dependent upon k_{on}. However, to our knowledge, LFERs in protein−protein interactions with a differential dependence of K_D on k_{on} or k_{off} have not been noted previously.

We also analyzed the experimental data for the kinetics of protein−protein complex formation compiled in ref 11. These data are represented in Figure 3, which shows slopes of 0.68 for the $\log(k_{on})-\log(K_D)$ plot for disordered proteins and 0.75 for the $\log(k_{off})-\log(K_D)$ plot for ordered complexes. This data set comprises values for protein−protein complexes that are not related in either sequence or structure, and we checked the effect of the length of the disordered interacting region (l_d). The rate of folding (related to k_{on}) is expected to have a weak dependence on l_d,[27] while $\log(k_{off})$ should have a stronger dependence on the number of bonds to be broken before the transition state is reached. The slope of the plot of $\log(k_{off}/l_d)$ versus $\log(K_D)$ is 0.31, as is the one without l_d (see the SI). While it is tempting to interpret this to mean that detaching the disordered region is not the slow step in k_{off}, the variation in k_{off} among the complexes is too large to draw such conclusions. Nevertheless, the slopes for these data spanning 8 orders of magnitude in K_D are different from the trivial value of 0.5 resulting from a random distribution of barriers, in alignment with the general spirit of correlations noted above.

In fact, in the light of the present observation, we revisited the data showing the effect of the α-helical stability (order) on the interactions of disordered S-peptide with S-protein.[28] It was seen that as the net α-helicity of the S-peptide analogues increased, the affinity switched from association-controlled to dissociation-controlled.[28] However, it is hard to compare the slopes from this data set, as only a few mutants of S-peptide were studied. All of the data presented so far suggests that as the proteins switch from ordered to disordered, the determinant of the affinity switches from the rate of dissociation to that of association.

The correlations among different mutants noted above arise from a *zeroth-order* similarity between k_{on} for ordered PPIs and k_{off} for disordered PPIs. Several factors govern the rates of association and dissociation. The rate of association is dictated by diffusion, conformational changes in the interacting proteins, and Coulombic electrostatic forces. Conversely, dissociation is dictated by the strength of short-range interactions between the proteins (i.e., van der Waals, hydrogen-bonding, and hydrophobic interactions and salt bridges). In the case of ordered proteins, the association rates can be predicted using diffusion models.[3,29] Thus, for ordered proteins, where large-scale conformational changes are not required, diffusional approach of the two interacting partners is the rate-limiting step for association, making k_{on} identical for different mutants to zeroth-order. This mechanism of association is similar to the diffusion−collision model of protein folding, where pre-existing secondary structural elements form a tertiary structure upon collision.[30] However, the stability of the ensuing protein−protein complex depends on the interfacial interactions, which govern the dissociation rate. In the case of disordered proteins, large conformational changes are required for association, and the strong $k_{on}-K_D$ relationship observed in the present work suggests that most of these structural changes occur prior to the formation of the transition state. The disordered proteins are enriched in polar and charged residues, facilitating penetration of the interface by water and fast dissociation, making the zeroth-order k_{off} of mutants identical. Although theoretical[9] and computational studies[10,11] have predicted increases in k_{on} by factors of up to 1.6 due to entropic factors in disordered proteins, several studies have shown an overall increase in k_{on}[28] and stronger binding[31,32] with an increase in helicity, in line with the present arguments.

The slopes in the LFER plots are commonly used to interpret the position of the transition state.[25,26] We follow a similar analysis to rephrase the PPI data in terms of transition-state theory. The binding reaction can be represented as

$$\mathbf{A} + \mathbf{B} = [\mathbf{A}\cdots\mathbf{B}]^{\ddagger} = \mathbf{AB} \qquad (1)$$

where $[\mathbf{A}\cdots\mathbf{B}]^{\ddagger}$ denotes the transition-state complex. With changes in the thermodynamic stabilities of reactants (ΔG_r) and products (ΔG_p), the free-energy changes in the transition state ($\Delta G^{\ddagger}$) can be approximated as[26]

$$\Delta G^{\ddagger} = \alpha\Delta G_p + (1-\alpha)\Delta G_r \qquad (2)$$

where α ($0 \le \alpha \le 1$) denotes the degree of closeness of the transition state to the reactants. From eq 2, the slope of a plot of $\log(k_{on})$ versus $\log(K_D)$ should be equal to α: $\partial[\log(k_{on})]/\partial[\log(K_D)] = (\Delta G^{\ddagger} - \Delta G_r)/(\Delta G_p - \Delta G_r) = \alpha$. If it is assumed that the mutations studied in the present work affect mostly the thermodynamic stabilities, the observed LFERs can be interpreted in structural terms. The $k_{on}-K_D$ relationship for disordered proteins suggests that $\alpha \approx 1$, so the transition state for complex formation should be structurally similar to the products (late transition state), that is, the disordered protein is nearly

dx.doi.org/10.1021/ja201500z | J. Am. Chem. Soc. 2011, 133, 9976–9979

ordered at the transition state itself. On the other hand, the k_{off}–K_D relationship suggests that $\alpha \approx 0$, so the transition state should be closer to the reactants (early transition state).[26] In certain PPIs, three-state behavior involving an encounter complex as an intermediate has been noted. One of the commonly discussed issues in such a context is the relative order of folding and binding. All of the above analysis can be extended to the folding step in the three-state disordered PPIs as well, whether it happens before or after binding.

The qualitative kinetic distinction made in this work can provide a better perspective on understanding the dynamic activity in the cell and its relation to disorder. Depending on the functional context, it may be important to achieve the binding affinity required for function through k_{on} or k_{off}. For proteins involved in signaling, mostly disordered, slow dissociation may not be an option, as it implies a long-lasting bound state. Thus, the affinity in this case can be association-controlled. For proteins involved in cell survival, such as in the case of the barnase inhibitor barstar, the association should be fast, and happens rapidly with $k_{on} = 10^8\ M^{-1}\ s^{-1}$. The differences in affinity among barnase–barstar complexes arise from rates of dissociation.[22] The rate–affinity relationships noted in our work thus show a correlation between these functional requirements on the rates and the presence of order in the proteins.

During evolution, proteins are subject to a constant pressure for change resulting from random mutations and subsequent selection. A selection pressure for rapid binding was also seen in the evolution of Im9, an inhibitor of the bacterial SOS-induced DNA toxins known as colicins (ColE). Starting from Im9, an inhibitor of ColE9, successive stages of mutagenesis and selection produced an inhibitor of ColE7 with a 10^6-fold increase in affinity and a 10^8-fold increase in selectivity. The criteria for successive stages of evolution were mutagenesis and selectivity for ColE7 inhibition, as judged by the digestion of ColE7.[23] These ordered protein–protein interactions evolved with a behavior showing k_{off} dependence on the stability, as shown in Figure 2B. This analysis suggests that the selection pressure reflects on the relative roles of k_{on} and k_{off} in defining the affinity and possibly in the choice of order for achieving the affinities. Future studies of the relationship between selection pressure and order can provide further understanding of the evolution of PPIs.

In conclusion, by studying linear free-energy plots for PPIs, we have shown that the affinity in disordered protein–protein interactions is controlled by the rate of association and that in ordered protein–protein complexes by the rates of dissociation. The presence of (dis)order may be interpreted as a consequence of the selection criterion imposed on the protein's kinetics by its functional role in the cell. The validity of LFERs and the relationship between functional kinetics and (dis)order can both be verified by extensive mutagenesis studies on ordered and disordered PPIs.

■ ASSOCIATED CONTENT

Supporting Information. Figures S1–S4; a table showing k_{on}, k_{off}, K_D, and l_d values; and complete ref 2. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
meher@phys.chem.ethz.ch

■ ACKNOWLEDGMENT

We thank Prof. Michele Parrinello, Prof. Rudolph Marcus, Dr. Alessandro Barducci, and Hagai Eshet for helpful discussions.

■ REFERENCES

(1) Alberts, B. *Molecular Biology of the Cell*; Garland Science: New York, 2002.
(2) Rual, J. F.; et al. *Nature* **2005**, *437*, 1173.
(3) Schreiber, G.; Haran, G.; Zhou, H.-X. *Chem. Rev.* **2009**, *109*, 839.
(4) Dunker, A. K.; Garner, E.; Guillot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z. *Pac. Symp. Biocomput.* **1998**, 473.
(5) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321.
(6) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. *Proteins: Struct., Funct., Genet.* **2000**, *41*, 415.
(7) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021.
(8) Wang, J.; Lu, Q.; Lu, H. P. *PLOS Comput. Biol.* **2006**, *2*, No. e78.
(9) Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8868.
(10) Turjanski, A. G.; Gutkind, J. S.; Best, R. B.; Hummer, G. *PLOS Comput. Biol.* **2008**, *4*, No. e1000060.
(11) Huang, Y.; Liu, Z. *J. Mol. Biol.* **2009**, *393*, 1143.
(12) Mészáros, B.; Tompa, P.; Simon, I.; Dosztanyi, Z. *J. Mol. Biol.* **2007**, *372*, 549.
(13) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197.
(14) Tompa, P. *FEBS Lett.* **2005**, *579*, 3346.
(15) Fersht, A. R. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; Freeman: New York, 1999.
(16) Panayotou, G.; Gish, G.; End, P.; Truong, O.; Gout, I.; Dhand, R.; Fry, M. J.; Pawson, T.; Waterfield, M. D. *Mol. Cell. Biol.* **1993**, *13*, 3567.
(17) Umulis, D.; O'Connor, M. B.; Blair, S. S. *Development* **2009**, *136*, 3715.
(18) Hesmath, L.; Dvorsky, R.; Fiegen, D.; Carlier, M.-F.; Ahmadian, M. R. *Mol. Cell* **2005**, *20*, 313.
(19) Siegal, G.; Davis, B.; Kristensen, S. M.; Sankar, A.; Linacre, J.; Stein, R. C.; Panayotou, G.; Waterfield, M. D.; Driscoll, P. C. *J. Mol. Biol.* **1998**, *276*, 461.
(20) Klages, J.; Kotzsch, A.; Coles, M.; Sebald, W.; Nickel, J.; Mueller, T.; Kessler, H. *Biochemistry* **2008**, *47*, 11930.
(21) Cheng, H.-C.; Skehan, B. M.; Campellone, K. G.; Leong, J. M.; Rosen, M. K. *Nature* **2008**, *454*, 1009.
(22) Schreiber, G.; Fersht, A. *J. Mol. Biol.* **1995**, *248*, 478.
(23) Levin, K. B.; Dym, O.; Albeck, S.; Magdassi, S.; Keeble, A. H.; Kleanthous, C.; Tawfik, D. S. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1049.
(24) Castro, M. J. M.; Anderson, S. *Biochemistry* **1996**, *35*, 11435.
(25) Cohen, A. O.; Marcus, R. A. *J. Phys. Chem.* **1968**, *72*, 4249.
(26) Szabo, A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2108.
(27) Ivankov, D. N.; Finkelstein, A. V. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8942.
(28) Goldberg, J. M.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2019.
(29) Alsallaq, R.; Zhou, H.-X. *Structure* **2007**, *15*, 215.
(30) Myers, J. K.; Oas, T. G. *Nat. Struct. Biol.* **2001**, *8*, 552.
(31) Zor, T.; Mayr, B. M.; Dyson, H. J.; Montminy, M. R.; Wright, P. E. *J. Biol. Chem.* **2002**, *277*, 42241.
(32) Tan, R.; Chen, L.; Buettner, J. A.; Hudson, D.; Frankel, A. D. *Cell* **1993**, *73*, 1031.