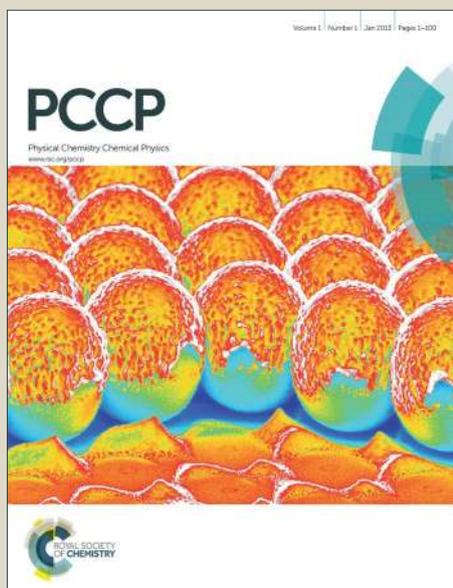


PCCP

Accepted Manuscript



This article can be cited before page numbers have been issued, to do this please use: S. Gopi, N. Rajasekaran, A. Singh, S. Ranu and A. N. Naganathan, *Phys. Chem. Chem. Phys.*, 2015, DOI: 10.1039/C5CP04765J.



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



Journal Name

COMMUNICATION

Energetic and topological determinants of a phosphorylation-induced disorder-to-order protein conformational switch

Received 00th January 20xx,
Accepted 00th January 20xxSoundhararajan Gopi,^a Nandakumar Rajasekaran,^a Animesh Singh,^b Sayan Ranu^b and Athi N. Naganathan^{a*}

DOI: 10.1039/x0xx00000x

www.rsc.org/

We show that the phosphorylation of 4E-BP2 acts as a trigger event to shape its folding-function landscape that is delicately balanced between conflicting favorable energetics and intrinsically unfavorable topological connectivity. We further provide first evidence that the fitness landscapes of proteins at the threshold of disorder can differ considerably from ordered domains.

Post-translational modifications (PTMs) such as phosphorylation, acetylation, methylation *etc.* determine the functional state of proteins. Particularly, protein (de-)phosphorylation is known to promote or eliminate protein-protein associations through generating or occluding a binding site, shifts in conformational equilibrium and modulation of local secondary structure.¹⁻⁴ A recent work has widened the role of phospho-regulation by demonstrating that the degree of phosphorylation in the four-stranded beta sheet protein 4E-BP2⁵ (Figure 1a) dramatically alters its conformational stability to range from disordered (non-phosphorylated; WT) to weakly stable (phosphorylation at a single site) to ordered (multisite phosphorylation).⁶ The dissociation constant (K_D) of 4E-BP2 to its binding partner eIF4E is shown to proportionately vary from ~3.2 nM to ~12.3 μ M thus abrogating the binding event and promoting cap-dependent translation initiation in mammalian systems.⁶

This counter-intuitive observation of a protein folding process (i.e. a gain in structural order guided by PTMs) eliminating binding has challenged the conventional 'structure-function paradigm'. Given these novel experimental observations, there are two open questions: a) how do the various energetic and topological factors intertwine to create this unique system, and b) how do these evolutionary design principles in a disordered domain differ from those of a typical well-folded globular protein? An understanding of these

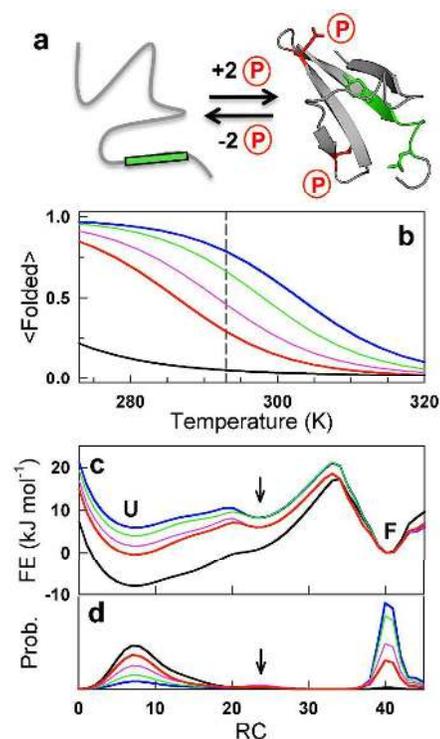


Fig. 1 Thermodynamics and Free Energy Profiles from the Exact Solution to the WSME Model. (a) Cartoon of the conformational switch in 4E-BP2 from a non-phosphorylated disordered state to a phosphorylated ordered state. Green shaded area represents the functional YXXXXLΦ motif occluded upon folding. (b) Unfolding curves of the doubly phosphorylated (blue), singly phosphorylated (pT29, green and pT20, magenta) and non-phosphorylated 4E-BP2 (red). Black: doubly phosphorylated mutant AXXXXAΦ. The vertical dashed line represents the experimental temperature of 293 K. (c & d) Predicted 1D free-energy profiles (panel c) and probabilities (panel d) as a function of the reaction coordinate (RC), the number of structured residues, for the variants shown in panel b. The arrow highlights the partially structured state C.

^a Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India.

^b Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai 600036, India.

*Email: athi@iitm.ac.in

Electronic Supplementary Information (ESI) available: Model parameters and additional figures supporting the results. See DOI: 10.1039/x0xx00000x

features would aid in the design of similar conformational switches taking advantage of the diverse physico-chemical nature of the amino acid residues. To answer these questions, we employ the structure-based statistical mechanical Wako-Saitō-Muñoz-Eaton model that assumes an ensemble of 2^N

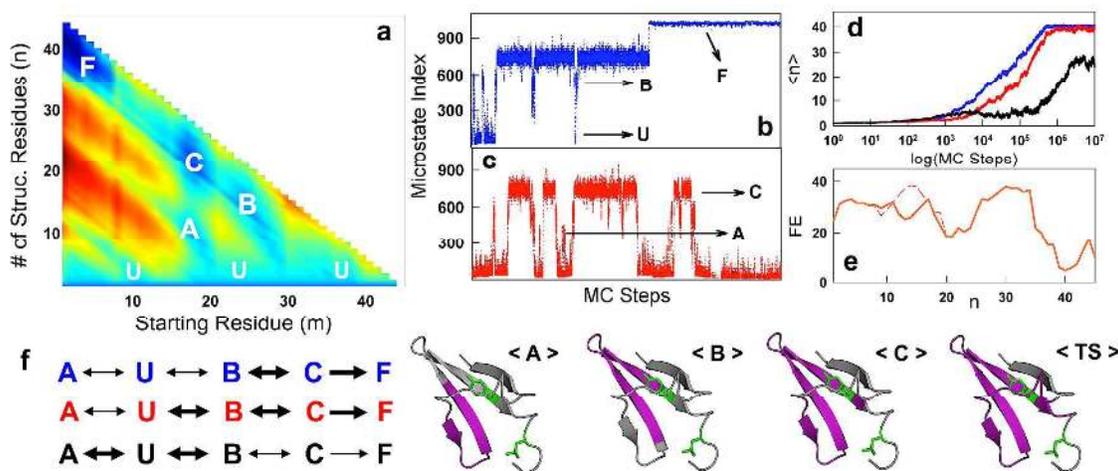


Fig. 2 Folding Mechanism as inferred from the SSA Landscape. (a) The SSA landscape highlighting the various partially structured states (A, B, C) apart from the fully unfolded state (U). Spectral color-coding is employed going from low free energy (blue) to high (red). The coordinate, say (25, 15), should be read as a microstate containing 15 folded residues (n) starting from the 25th residue (m) – here, it represents the state B. (b & c) Representative examples of reasonably long Monte-Carlo (MC) runs of 10^5 steps on the 2D SSA landscape as a function of the microstate index for the doubly phosphorylated (blue and panel b) and non-phosphorylated 4E-BP2 (red and panel c) highlighting the conformational sub-states. (d) Ensemble behaviour from 100 MC runs with the ordinate denoting the mean number of structured residues (black is the doubly phosphorylated mutant AXXXXAΦ). (e) The path $U \rightarrow A \rightarrow C \rightarrow F$ exhibits a higher free energy (FE in kJ mol^{-1} ; dashed red) compared to $U \rightarrow B \rightarrow C \rightarrow F$ (continuous red) for the doubly phosphorylated variant. (f) Conformational heterogeneity in the folding mechanism and dynamics. Thicker arrow represents a higher flux path. A cartoon of the various partially structured states extracted from the 2D landscape – grey and magenta represent unfolded and folded regions, respectively. The residues Y and L of the functional motif YXXXXLΦ are coloured in green.

microstates for an N -residue protein.^{7, 8} It has been successful in characterizing the folding mechanistic behaviours of several proteins,⁹⁻¹¹ predicting folding rates from 3D structures,⁸ and engineering stabilities.^{12, 13} Importantly, it has been shown to capture the average conformational properties of disordered protein ensembles (unfolded states and natively disordered proteins) and their heterogeneity in very good agreement with experiments and molecular simulations.¹⁴⁻¹⁶

We start with the known structure of the doubly phosphorylated variant of 4E-BP2 (pT20pT29, phosphorylated at T20 and T29; PDB id: 2MX4) and employ parameters similar to that used earlier for a beta-sheet protein CspB: a 6 Å cut-off to identify van der Waals (vdW) interactions excluding nearest neighbours, all-to-all electrostatics, solvation and secondary-structure dependent conformational entropy (see Electronic Supporting Information; ESI).¹² The phosphate groups are assigned an effective charge of -2 at pH 7.0. The model is therefore Gō-like¹⁷ with the native-bias introduced through the vdW and electrostatics terms; the landscape is funnelled towards the folded state in accordance to the minimal frustration principle of the energy landscape theory.¹⁸ All calculations are performed at 293 K, pH 7.0 and 0.16 M ionic strength thus simulating the experimental conditions, unless mentioned otherwise.

Because experimental equilibrium unfolding curves are not available for the different variants, we adjust the vdW term to achieve a minimum native state population (F) of 80% at 293 K for the pT20pT29 4E-BP2 (blue in Figure 1b and ESI); employing identical parameters, we predict that the non-phosphorylated WT exhibits only ~26% folded population and

hence disordered (U ; red in Figure 1b) under the same conditions in agreement with experiments. We also find a graded stability difference between the different singly phosphorylated variants (pT20 and pT29, magenta and green, respectively, in Figure 1b) similar to experimental observations.⁶ Mutating the canonical binding motif YXXXXLΦ located at the C-terminus of the 4th strand to AXXXXAΦ destabilizes the protein dramatically resulting in a folded population of ~5% at 293 K (Figure 1b). All variants exhibit sigmoidal unfolding curves suggestive of an apparent two-state-like behaviour (Figure 1b). However, one-dimensional (1D) free energy profiles as a function of the order parameter, the number of structured residues, point to a minor intermediate-like state (I) at a reaction coordinate value of 24 structured residues (arrows in Figure 1c-1d). Effectively, we show that the WSME model energetics, calibrated by quantitatively characterizing the folding behaviour of ordered proteins,¹² is robust enough to capture the observed large changes in stability induced by phosphorylation or mutations in 4E-BP2.

We provide a first look at the possible folding mechanisms of 4E-BP2 by constructing a minimalistic landscape with single stretches of folded residues (Single-Sequence Approximation (SSA) landscape⁸) with 1035 microstates, i.e. $N*(N+1)/2$ species, where $N=45$ here (ESI). This approach has been previously successful in identifying intermediates in good agreement with experiments and simulations in several proteins with varied topologies including RNase H,¹⁹ CytR,¹⁴ bovine lactalbumin¹² and IkBα.²⁰ Under native conditions, the SSA landscape is funnelled towards the folded state with

multiple partially structured macro-states (labelled A, B, C in Figure 2a) populated on the unfolded side of the main barrier. The free energy gradient immediately suggests that there could potentially be three different microscopic folding paths towards F: $U \rightarrow A \rightarrow C \rightarrow F$ (UACF) or UBCF or UABCF. The direct path from U to F, without sampling any of the above states, is seen to be free-energetically unfavourable (dark red in Figure 2a).

To identify the path with the highest flux, we perform 100 Monte-Carlo runs of 10^7 steps each on the SSA landscapes¹⁹ starting from the unfolded state for the three different variants: pT20pT29, non-phosphorylated WT and the AXXXXAΦ mutant (illustrative examples are provided in Figure 2b-2c; see ESI text and Figures S1 and S2). The ensemble behaviour, i.e. the average from 100 molecules (Figure 2d), displays an expectation consistent with the 1D free energy profile: 1) the non-phosphorylated variant takes longer to fold compared to the fully phosphorylated variant due to the smaller gradient towards the folded state (Figures 1c, S2, S3), and 2) the mutant despite folding is unable to consistently stay folded due to the lower free energy of the unfolded state resulting in frequent unfolding events that effectively result in ~21-23 structured residues at the end of 10^7 steps. Analysing the trajectories, we find the path UBCF to be the most favourable for the variant pT20pT29 (98% flux; continuous orange in Figure 2e) because of the relatively higher barrier of ~4.1 kJ mol⁻¹ in the path UACF (2% flux; dashed orange in Figure 2e). The macro-state A can therefore be seen as an off-pathway intermediate in a linear chemical-reaction mechanism or a low flux path in a triangular mechanism; macro-states B and C, on the other hand, appear to be obligate intermediate-like states from a chemical view of folding (scheme on top in Figure 2f). The nature of the partially structured states observed in this simplified landscape is in accordance with various other representations calculated from both the exact solution and multiple-sequence approximations (Figures S3-S5).

The non-phosphorylated WT 4E-BP2 predominantly samples states U, B and C following which it folds (scheme at the centre in Figure 2f). This is consistent with the 1D free-energy profile that displays a near identical free energies for the folded and unfolded states (Figure 1c). The equilibrium native ensemble of the disordered WT can therefore be viewed as consisting of distinct states B and C apart from U, that contribute to the intermediate-like state on the 1D free energy profile (arrow in Figure 1c). However, the landscape is very different for the AXXXXAΦ mutant that shuttles between states U, A and B with states C and F being energetically unfavourable (scheme at the bottom in Figure 2f). Our results suggest that the phosphorylation status and the nature of mutations in 4E-BP2 has a dramatic effect on the observed equilibrium ensembles and the underlying dynamics, indicative of a system that is highly sensitive to perturbations (Figures S3-S5).

As the model is structure-based it is straightforward to obtain the structures of the different intermediate-like states observed in the SSA landscape (Figure 2f): (State A)

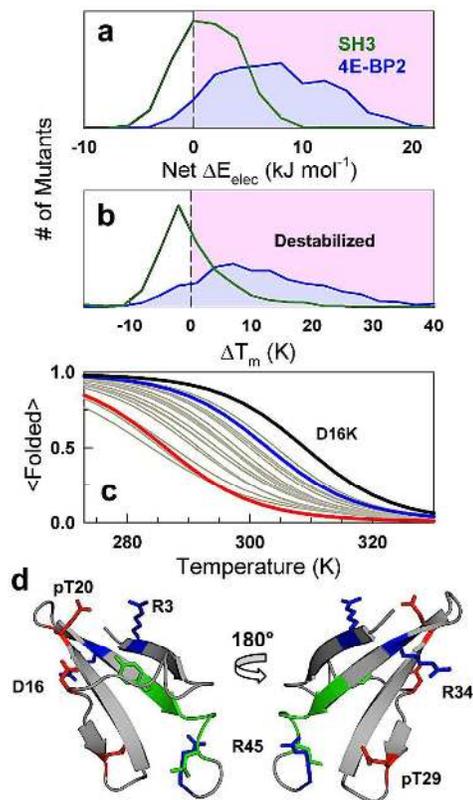


Fig. 3 Energetics and Mutational Fitness Landscape. (a) Distribution of the relative net charge-charge interaction energies (ΔE_{elec}) in the 1562 single-, double- and triple-point mutants of src-SH3 (green) and phosphorylated 4E-BP2 (blue) calculated from the charge-shuffling procedure as $E_{\text{elec}}(\text{Mutant}) - E_{\text{elec}}(\text{WT})$. The vertical dashed line represents the reference WT. (b) Same as in panel a, but for the differences in melting temperatures calculated as $T_m(\text{WT}) - T_m(\text{Mutant})$, employing the exact solution to the WSME model. (c) Representative unfolding curves for the single-point mutants (grey) of phosphorylated 4E-BP2 (blue). The unfolding curves of the non-phosphorylated 4E-BP2 and the mutant D16K are shown in red and black, respectively. (d) Structure of phosphorylated 4E-BP2 highlighting the various residues contributing to the 'electrostatic staples' with respect to the functional motif (green).

partial structure in hairpin 1, strands 2 and 3; (State B) partial structure in hairpin 2, strands 3 and 4; (State C) partial structure in hairpins 1, 2 and strands 2, 3 and 4; (Transition State; TS) same as state C but with the long loop connecting strands 1 and 2 exhibiting native-like conformation. The formation of this long loop hastens the condensation of the strand 1 on the folding side of the main barrier in a downhill-like fashion (Figure S5). This process simultaneously occludes the binding motif YXXXXLΦ that is nearly covered in its entirety by the loop in the fully folded conformation (Figure 1a). NMR experiments on singly phosphorylated variants point to partially folded structures with significant order near the phosphorylated turn regions⁶ in accord with our predictions on the complex inter-connected landscape (i.e. states A and B) and early hairpin formation. Large experimental NOE violations are also observed in the folded structure corresponding to the loop region and in one case the dissolution of the parallel beta-sheet formed by strands 1 and

COMMUNICATION

Journal Name

4 is shown to reduce these violations⁶ – this observation is consistent with the structure of state C which is ~0.8 % populated under these conditions (arrows in Figure 1c-1d).

The ability of the model to reproduce trends in experimental data, albeit qualitatively, enables us to probe further for the molecular determinants of this unique mechanism. Particularly, the fitness landscape of ordered domains is expected to be flat to buffer the predominantly destabilizing effect of random mutations that accumulate during Evolution.^{21, 22} Does 4E-BP2 exhibit a behaviour similar to an ordered domain and if it does not, how much does it differ? To answer this question, we first perform the computational charge-shuffling procedure on the pT20pT29 variant by iteratively shuffling the charges around in the native structure to identify stabilizing and de-stabilizing mutants (single-, double- or triple-mutants involving charge residues) from Debye-Hückel interaction energy alone.¹³ We perturb only the surface charged residues for three reasons: 1) residues on the protein surface, that are generally polar or charged, exhibit an enhanced probability of mutation during the course of Evolution,²³ 2) it is straightforward to capture the changes in stability due to the elimination or addition of a charge-charge interactions without introducing complex entropy-enthalpy compensations associated with polar → non-polar (or *vice versa*) substitutions,^{12, 13} and 3) this approach reduces the phase-space to a more amenable number of mutations (see below) compared to innumerable 20^N when considering all possible mutations at all sites.

In the ideal-case scenario of charge shuffling with 11 charges in 4E-BP2, we find that only a very small fraction of charged-residue substitutions are stabilizing (~5% of the possible 1562 single-, double- and triple-mutations; Figure 3a) – *i.e.* a distribution that is skewed towards the right with respect to the starting structure. To further confirm this observation, we generate structures of each of the 1562 mutants and predict their unfolding curves (representative examples of 22 single mutants are presented in Figure 3c) and hence the melting temperatures (T_m) from the WSME model. Only a small fraction of mutants are identified to be stabilizing in this direct approach (~10%) in agreement with the charge-shuffling procedure – the difference between the numbers predicted by the two methods arises because of the unavoidable differences in packing introduced while generating mutant structures.

The situation is reversed in the well-folded protein src-SH3 that also contains 11 charged residues. We find that ~26% (or ~49%) of the mutants are predicted to be more stable than the WT src-SH3 from the charge-shuffling procedure (or by generating mutants). In fact, the distribution of destabilization obtained for the WT src-SH3 is very similar to those reported from the ProTherm mutational database indicating that this is a generic observation expected of ordered domains.^{24, 25} This comparative study clearly highlights that the fitness landscape of 4E-BP2 is naturally tuned towards disorder ($\sigma_{\Delta T_m} \sim 10.4$ K and a broad range from -10 K - 35 K) compared to a well-folded protein src-SH3 that is more resistant to protein surface

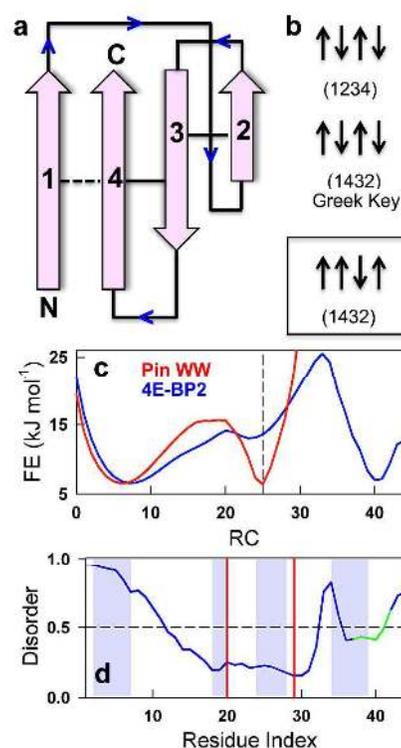


Fig. 4 Topological Determinants and Sequence-Based Disorder. (a) The topology of the 4-stranded 4E-BP2. The dashed line (between strands 1 and 4) and the solid lines between strands 4, 3 and 3, 2 represent weak and strong hydrogen bonds, respectively. (b) The topologies of the most common 4-stranded motifs with a long loop connecting strands 1 and 2. 4E-BP2 is characterized by the least common arrangement (within box and panel a) with a strand ordering of 1432 and strand topology of up-up-down-up. (c) The 1D free energy profiles near the respective denaturation midpoints of Pin WW domain (red) and 4E-BP2 (blue) from the exact solution to the WSME model. The vertical dashed line indicates the position of the Pin WW domain folded state that is in agreement with the partially structured state C of 4E-BP2. (d) Sequence based disorder from the PONDR-FIT server.²⁶ The shaded area represents the position of the strands. The phosphorylation sites and functional binding region are shown as red and green, respectively.

mutations ($\sigma_{\Delta T_m} \sim 5.9$ K and a narrow range from -10 K - 12 K). Not only is the distribution for 4E-BP2 atypical of a globular protein but also mutations of just two residues D9 and D16 (to positive or neutral residues) account for >90% of the stable mutants while in the case of src-SH3 there are at least five independent sites (E17, E21, D23, R31 and D41).

Phosphorylation induced stabilization is frequently governed by strong electrostatic interactions between the phosphate and the amino acid arginine.¹ A closer look at the 4E-BP2 structure reveals the same interactions observed in ordered domains: interactions between pT20 and R3 and between pT29 and R45/R34 (Figure 3d). These long-range 'electrostatic staples' hold the structure together and contribute a significant proportion of the stabilization free energy and that is exploited by Nature for functional purposes. 4E-BP2 does not, however, possess the optimal distribution of charge-charge interactions, as a highly stable state could effectively eliminate the switching behaviour arising from phosphorylation. This is possibly the reason for the presence

of the two negatively charged residues D9 and D16 (in the long loop that connects strands 1 and 2) that slightly destabilize the folded structure (Figure 3c) to counter-balance the other favourable interactions, thus maintaining the structure at the threshold of disorder.

The topology or the relative arrangement of beta-strands in 4E-BP2 also serves an interesting case. Figure 4a shows that the strand 1 of 4E-BP2 is connected by a long loop to strand 2 in a convoluted manner; strand 2 continues to strand 3 and 4 aided by two tight turns. In effect, 4E-BP2 structure is characterized by one parallel and two anti-parallel beta-sheets with a strand ordering of 1, 4, 3, and 2 and a strand topology of up-up-down-up (Figure 4a and the cartoon within the box in Figure 4b).⁶ Extensive PDB database analyses^{27, 28} has shown that such a strand ordering and topology with a long loop connecting strands 1 and 2 is the least common among all four-stranded motifs in all- β proteins (L5 in Figure 3b of reference 28; also see Figure 4b here). This unique fact is in accordance with model predictions that indicate long loop and the first strand is the last to form during folding (Figure 2 and Figure S5). The formation of native-like structure in the long loop is not adequately compensated by a gain in stabilization energy resulting in a large free energy barrier that is primarily entropic in origin (that separates states C and F in Figure 1c) and that Nature has avoided on most other 4-stranded β -proteins owing to the tortuous connectivity.

Experiments, model predictions and topological considerations point to the state C being a well-structured unit that is populated to different extents depending on the degree of phosphorylation (Figures 1c, 1d, 2f). The topology of state C with 3 folded strands is very similar to that of WW-domains. In fact, the predicted 1D free energy profile of Pin WW domain is in good agreement both in terms of the barrier magnitude and the location of the folded minima (Figure 4c). Experiments and simulations on WW domains have demonstrated that the folding mechanism is complex with the possibility of nucleation in either one of the β -turns²⁹ (similar to a triangular three-state mechanism; see Figure S6 for the SSA landscape of the WW domain that displays an identical behaviour); this conforms to the early stage folding mechanism of 4E-BP2 we predict from the SSA landscape that shows the possibility of two parallel paths towards the state C that go through either of the states A or B starting from U. In other words, the complex 4E-BP2 system seems to be built on a simple WW-domain-like scaffold with the long-range stabilization induced by phosphorylation acting as a trigger event to compensate for the unfavourable entropic cost of ordering the long loop. Interestingly, PONDR-FIT²⁶ predicts a large degree of structural disorder in the first strand and the loop connecting it, hinting that the disorder tendency of these two elements is encoded even at the level of sequence (Figure 4d). The two β -turns with the high propensity PGGT motifs³⁰ is also identified to be less disordered aligning well with both experiments and model predictions. Collectively, we find that partially structured (or intermediate) states encompassing different combinations of β -hairpin-like structures could be a generic feature of proteins

rich in anti-parallel β -sheets especially when they are linked by β -turns of varying intrinsic propensity.

Conclusions

The acquisition of native structure in the loop is the event that blocks the site YXXXXL Φ in 4E-BP2 from binding its partner eIF4E. Our detailed analysis above shows that the formation of this loop is unfavourable both at the level of sequence (Figure 4d) and topology (Figure 4a, 4b). Even under conditions that promote folding, i.e. phosphorylation, the loop and the strand preceding it is the last to form (Figure 2), despite the structure exhibiting relatively optimized energetics (Figure 3) and high propensity β -turn motifs. These observations coupled with experiments that point to the possibility of several conformational sub-states effectively suggest that the folding-functional landscape of 4E-BP2 has evolved to be at the threshold of disorder for a rapid functional regulation. It also highlights the various sequence, structure-energetic and topological factors that come together to create this unique system.

Our findings presented here open up the possibility of designing protein-based conformational switches by simply working backwards from the disorder prediction algorithms, simple energetic considerations and by employing basic structural scaffolds such as the three-helix bundles or three-stranded structures, an avenue that is yet to be exploited in protein design. Finally, we report first evidence to the possibility of fundamentally different fitness landscapes for proteins that undergo order-to-disorder (or *vice versa*) transition that could be further tested by experiments and simulations.

Notes and references

1. L. N. Johnson and R. J. Lewis, *Chem. Rev.*, 2001, **101**, 2209.
2. C. Antz, T. Bauer, H. Kalbacher, R. Frank, M. Covarrubias, H. R. Kalbitzer, J. P. Ruppertsberg, T. Baukrowitz and B. Fakler, *Nat. Struct. Biol.*, 1999, **6**, 146.
3. C. Löw, N. Homeyer, U. Weiningner, H. Sticht and J. Balbach, *ACS Chem. Biol.*, 2008, **4**, 53.
4. D. Kast, L. M. Espinoza-Fonseca, C. Yi and D. D. Thomas, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 8207.
5. C. M. Fletcher, A. M. McGuire, A. C. Gingras, H. J. Li, H. Matsuo, N. Sonenberg and G. Wagner, *Biochemistry*, 1998, **37**, 9.
6. A. Bah, R. M. Vernon, Z. Siddiqui, M. Krzeminski, R. Muhandiram, C. Zhao, N. Sonenberg, L. E. Kay and J. D. Forman-Kay, *Nature*, 2015, **519**, 106.
7. H. Wako and N. Saito, *J. Phys. Soc. Japan*, 1978, **44**, 1939.
8. V. Muñoz and W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 11311.
9. M. M. Garcia-Mira, M. Sadqi, N. Fischer, J. M. Sanchez-Ruiz and V. Muñoz, *Science*, 2002, **298**, 2191.
10. J. Kubelka, E. R. Henry, T. Cellmer, J. Hofrichter and W. A. Eaton, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 18655.
11. P. Bruscolini and A. N. Naganathan, *J. Am. Chem. Soc.*, 2011, **133**, 5372.
12. A. N. Naganathan, *J. Chem. Theory Comput.*, 2012, **8**, 4646.
13. A. N. Naganathan, *J. Phys. Chem. B*, 2013, **117**, 4956.

COMMUNICATION

Journal Name

14. A. N. Naganathan and M. Orozco, *J. Phys. Chem. B*, 2013, **117**, 13842.
15. A. N. Naganathan, J. M. Sanchez-Ruiz, S. Munshi and S. Suresh, *J. Phys. Chem. B*, 2015, **119**, 1323.
16. S. Munshi and A. N. Naganathan, *Phys. Chem. Chem. Phys.*, 2015, **17**, 11042.
17. H. Taketomi, Y. Ueda and N. Go, *Inter. J. Prot. Pep. Res.*, 1975, **7**, 445.
18. J. D. Bryngelson, J. N. Onuchic, N. D. Socci and P. G. Wolynes, *Proteins*, 1995, **21**, 167.
19. A. Narayan and A. N. Naganathan, *J. Phys. Chem. B*, 2014, **118**, 5050.
20. S. Sivanandan and A. N. Naganathan, *PLOS Comp. Biol.*, 2013, **9**, e1003403.
21. N. Tokuriki and D. S. Tawfik, *Curr. Opin. Struct. Biol.*, 2009, **19**, 596.
22. A. W. R. Serohijos and E. I. Shakhnovich, *Curr. Opin. Struct. Biol.*, 2014, **26**, 84.
23. Y.-S. Lin, W.-L. Hsu, J.-K. Hwang and W.-H. Li, *Mol. Biol. Evol.*, 2007, **24**, 1005.
24. M. M. Gromiha, J. An, H. Kono, M. Oobatake, H. Uedaira and A. Sarai, *Nuc. Acids Res.*, 1999, **27**, 286.
25. E. I. Shakhnovich and B. E. Shakhnovich, *Curr. Opin. Struct. Biol.*, 2008, **18**, 375.
26. B. Xue, R. L. Dunbrack, R. W. Williams, A. K. Dunker and V. N. Uversky, *Biochim. Biophys. Acta*, 2010, **1804**, 996.
27. C. Zhang and S. H. Kim, *J. Mol. Biol.*, 2000, **299**, 1075.
28. I. Ruczinski, C. Kooperberg, R. Bonneau and D. Baker, *Proteins*, 2002, **48**, 85.
29. A. J. Wirth, Y. Liu, M. B. Prigozhin, K. Schulten and M. Gruebele, *J. Am. Chem. Soc.*, 2015, **137**, 7152.
30. B. Madan, S. Y. Seo and S.-G. Lee, *Proteins*, 2014, **82**, 1721.