PCCP

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: S. Munshi and A. N. Naganathan, *Phys. Chem. Chem. Phys.*, 2015, DOI: 10.1039/C4CP06102K.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

View Article Online
DOI: 10.1039/C4CP06102K

PCCP

RSCPublishing

ARTICLE

Imprints of Function on the Folding Landscape: Functional Role for an Intermediate in a Conserved Eukaryotic Binding Protein

Sneha Munshi and Athi N. Naganathan*

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

In the computational characterization of single domain protein folding, the effective free energy of numerous microstates are projected onto few collective degrees of freedom that in turn serve as well-defined reaction coordinates. In this regard, one-dimensional (1D) free energy profiles are widely used mainly for their simplicity. Since folding and functional landscapes are interlinked, how well can these reduced representations capture the structural and dynamic features of functional states while being simultaneously consistent with experimental observables? We investigate this issue by characterizing the folding of the four-helix bundle bovine acyl-CoA binding protein (bACBP), which exhibits complex equilibrium and kinetic behaviours, employing an Ising-like statistical mechanical model and molecular simulations. We show that the features of the 1D free energy profile are sufficient to quantitatively reproduce multiple experimental observations including millisecond chevron-like kinetics and temperature dependence, a microsecond fast phase, barrier heights, unfolded state movements, intermediate structure and average ϕ-values. Importantly, we find that the structural features of the native-like intermediate (partial disorder in helix 1) is intricately linked to a unique interplay between packing and electrostatics in this domain. By comparison with available experimental data, we propose that this intermediate determines the promiscuous functional behaviour of bACBP that exhibits broad substrate specificity. Our results present evidence to the possibility of employing the statistical mechanical model and the resulting 1D free energy profile to not just understand folding mechanisms but to even extract features of functionally relevant states and their energetic origins.

Introduction

Protein folding involves a subtle interplay between stabilizing energetic terms that promote folding and conformational entropic contributions that disfavour folding.[1, 2] Given the weak non-covalent nature of the stabilizing interactions and the large phase space available for the main-chain, it is now well recognized that folding processes and even the equilibrium ensembles are best described by a very large number of microstates among which one or a small subset correspond to the folded or functional state.[3, 4] Computational treatments of folding have to therefore account for the large conformational freedom of the main chain and this is well captured in simulations.[5-9] All-atom simulations now capable of approaching longer and functionally relevant timescales and hence a better sampling of the conformational space primarily through the development of specialized hardware[10] and more accurate force-fields.[11]

The large number of degrees of freedom of the main chain effectively means that the free-energy surface on which the protein molecule diffuses towards the folded state is hyper-dimensional.[12] However, analysis of cubic lattice simulations has shown that even a single effective coordinate if aptly chosen is sufficient to reproduce the dynamics of folding in many cases.[13] It is therefore common in atomistic or coarse-grained simulations to integrate out non-relevant degrees of freedom such that only a handful of order parameters remain that in turn could serve as good reaction coordinates (RC). It is important to emphasize that a one-dimensional (1D) treatment is not restricted to molecular simulations and statistical mechanical models of folding have frequently taken recourse to the same. For example, the variable-barrier model inspired by the Landau theory of phase transitions employs enthalpy as the 1D order parameter, and predicts the thermodynamic folding barriers from the heat capacity profile[14] in very good agreement with the experimental folding rates.[15] A related phenomenological 1D model based on analytical energy functionals is remarkably adept in not only explaining several complex fast folding-related phenomena, but also in estimating thermodynamic barriers from the shape of temperature-relaxation rate plot.[16]

ARTICLE View Article Online DOI: 10.1039/C4CP06152K Journal Name

The dimensionality reduction using 1D approaches are particularly advantageous as they provide a simplified picture of the folding process that can be interpreted in a straightforward manner and compared to experiments directly. Despite their effectiveness, questions remain on their feasibility to quantitatively characterize the folding complexity of larger single-domain proteins at different conditions (temperature and/or co-solvent). An interesting case is that of the designed 73-residue helical protein α_3D whose folding kinetics is incompatible with diffusion on a 1D free energy profile.[17] In other words, in going from a detailed energy landscape to a free energy surface and finally to a free energy profile, relevant coordinates that carry important information on folding motions are possibly lost thus contributing to this discord.[18] It is however not clear at this point if this is a generic expectation or just the unique behaviour of a designed protein.

Of the various models employed by the protein folding community, the Wako-Saitô-Muñoz-Eaton (WSME) model or the ME model has been constantly gaining ground as an attractive alternative to molecular simulations.[19-21] It has been shown to be robust enough to quantitatively reproduce the heat capacity profiles of several proteins,[21-26] capture the phenomenon of cold denaturation,[21, 27, 28] and the rates and amplitudes of different kinetic phases observed in experiments[21, 23] – three critical and often overlooked folding thermodynamic/dynamic features that are still beyond the reach of both coarse-grained and molecular simulations.[29] Apart from these, the one-state global downhill folding behaviour in BBL was first ascertained from a global thermodynamic analysis of four different spectroscopic probes employing the ME model.[22] A follow-up study applied the same model to the multi-probe data on its homolog PDD revealing a marginal barrier to folding ($\sim k_BT$)[26] consistent with various other thermodynamic models[24, 30] and kinetic arguments.[30] In the case of villin headpiece folding, the model's predictive powers, particularly in the distribution of possible folding mechanisms, approach that of atomistic simulations.[31] The energetic terms in the latest version of the model[25] is sound enough to predict stabilizing or destabilizing mutations of charged residues from 16 different proteins and enzymes thus providing a tool to engineer protein stabilities.[32]

The natural RC of the versatile WSME model is the number of structured residues, which is very different from that employed in molecular simulations and other statistical methods. Though this approach tends to slightly enhance the folding cooperativity,[33] it has already been shown to be successful in capturing the complex folding kinetics of α-helices,[34] β-hairpin[35] and in predicting folding rates from solely the 3D structure of proteins.[20, 36] Given that the 1D free energy profiles from the WSME model is competent enough to delineate folding mechanisms, it is natural to expect that it

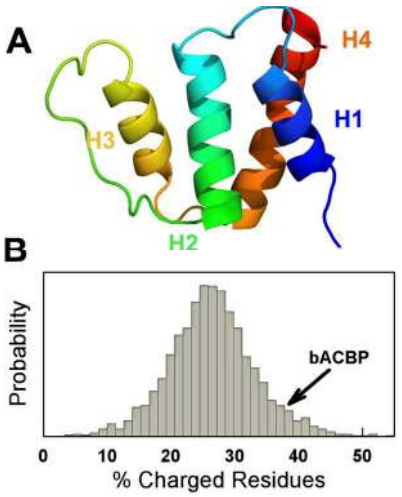

Fig. 1 (A) Structure of bovine ACBP. (B) Distribution of the fraction of charged residues in folded single-domain proteins without disulphide bridges or prosthetic groups and in the size range of 30-150 residues with a sequence similarity of at most 30%. The fractional charged residue composition of bACBP is indicated.

should also be possible to discern the relevant functional states or intermediates from the same, since folding and functional phenomena are inter-linked. In fact, one of the main results of several experimental-computational studies[4, 28, 37-42] is that the conformational behaviour of proteins is not optimized for large stability or rapid folding or high cooperativity but that functional constraints are the primary driving force that determine each of the above features. If this is indeed the case, then it should be possible to extract the functional imprints on the folding process from appropriately designed experiments or modelling or a combination of both. This is particularly possible in the WSME model as it offers a direct connection between sequence, structure, energetics (packing and electrostatics), and 1D free energy profiles together with the ability to quantitatively characterize experimental data.

We explore these issues by studying the experimentally well-characterized 86-residue four-helix bundle bovine acyl-CoA binding protein (bACBP; Fig. 1A).[43] Numerous experiments have shown that bACBP populates an intermediate,[44] possesses residual structure in the unfolded state[45, 46] and exhibits non-trivial folding relaxation kinetics.[47] Interestingly, bACBP has a very large fraction of charged residues in its sequence unlike an average single-domain protein (Fig. 1B) - ~38% of charged residues is nearly 2 standard deviations higher than the mean behaviour. We show below that it is possible to unite all of these disparate observations through a simple 1D free-energy profile that provides both quantitative mechanistic and functional insights into the conformational complexity of this domain.

Methods

WSME Model. The details of the model, including the energetic terms, the procedure to calculate the partition function, exact solution energy profiles and residue

Journal Name
View Article Online
DOI: 10.1039/C4CP06102K

probabilities are discussed in recent works[25,32]. Briefly, the WSME model defines a large instantaneous ensemble of 2^N microstates where N is the protein length, considering that each residue can sample either folded-like (binary code *1*) or unfolded-like (binary code *0*) conformations.[19, 20] Moreover, native-like islands are assumed to interact with each other only when all of the intervening residues are also folded. As the number of unfolded microstates out-weighs the folded ones, a large entropic cost (ΔS_{conf}) per residue is introduced at a reference temperature of 385 K[48] that determines the energy-scale. Since the model is structure based (i.e. Gō-like[49]), only those interactions present in the native state of apo-bACBP (PDB id: 2ABD) are treated as favourable. Van der Waals interactions (E_{vdW}) are identified with a 5 Å heavy-atom cut-off while excluding the nearest neighbour contribution with mean-field interaction energy of ξ. The model explicitly considers all-to-all electrostatics (E_{elec}) with an effective dielectric constant (ε_{eff}) of 29 obtained by a previous calibration of the electrostatic term against homologous proteins and single point mutations.[25, 32] Charged residues are protonated according to the pH - lysine and arginine are charged at pH 7.0, while histidine gets additionally protonated at pH 5.0. An ionic strength value of 0.02 M and pH 5 is employed in all simulations mimicking the experimental conditions,[50] unless otherwise mentioned. The solvation free energy (ΔG_{solv}) is introduced through a heat capacity term that is proportional to the number of contacts with the constant being the heat capacity change upon forming a native contact ΔC_p^{cont}. The effective free energy of each and every microstate therefore includes contributions from E_{vdW}, E_{elec}, and ΔG_{solv} while the destabilizing effect arises from the cost of fixing residues within that microstate. The effect of denaturants are introduced by a linear term similar to what is observed experimentally,[51] with the proportionality term being a product of the constant m_{cont} and the total number of interactions within that microstate.

As no heat capacity profiles or cold-denaturation studies are available on bACBP, we fix the ΔC_p^{cont} to -0.36 J mol^{-1} K^{-1} per native contact from a previous analysis.[25] The final parameters required to reproduce the chemical denaturation midpoint (C_m ~2.3 M) and the sharpness of the unfolding transition (m-value ~14.8 kJ mol^{-1} M^{-1}) at 278 K are: ΔS_{conf} = -13.5 J mol^{-1} K^{-1} per residue, ξ = -82.0 J mol^{-1} per native contact, and m_{cont} = -5.1 J mol^{-1} M^{-1} per native contact. Rate constants were estimated by diffusive kinetic calculations employing the discretized version of the 1D diffusion equation that involves the construction of a rate matrix as prescribed before.[52] Eigenvalues and eigenvectors are obtained by the diagonalization of the 86 x 86 rate matrix. The eigenvector corresponding to the eigenvalue of zero provides information on the final equilibrium population. The other non-zero eigenvalues report on relaxation rates with the corresponding eigenvectors pointing to the origin of these rates through population redistributions on the 1D free energy profile. A temperature-dependent 1D diffusion constant of 2.3x10^5 n^2 s^{-1} at 278 K (k_0; n is the reaction coordinate value) with an activation energy of ~1.08 kJ mol^{-1} per residue (E_a; a measure of landscape roughness[16]) is employed to match the experimental chemical midpoint rates at 278, 299 and 313 K. Effectively, the model in its current version requires five adjustable parameters - 3 thermodynamic (ΔS_{conf}, ξ, m_{cont}) and 2 dynamic (k_0, E_a). The dynamic term k_0 sets the scale for the rate *vs*. temperature plot while the activation term E_a defines its shape on top of the temperature dependent heat capacity term – a more detailed discussion is provided in Ref. 16.

We algorithmically quantify the statistical weights of each and every microstate with single- and two-stretches of folded residues, SSA (single sequence approximation) and DSA (double sequence approximation), respectively. This required an enumeration of 3741 SSA microstates and 2,225,895 DSA microstates;[20] the effective free energy of these states is then projected on to the 1D reaction coordinate for comparison with the free energy profile from the exact solution calculation. As SSA microstates are characterized by single stretches of folded residues, their structures are obtained by editing the PDB file.[22] The number of contacts made by tryptophan 55 and 58 in each of the SSA microstates are tabulated in a two-dimensional map and then projected onto the same RC to estimate a signal for tryptophan fluorescence similar to our calculations before.[42, 53]

All-Atom MD Simulations. Virtual hydrogen sites were added on to the bACBP structure (PDB: 2ABD) assuming a pH 7.0 protonation state using the Amber-99SB* ILDN force-field[11, 54] in GROMACS 4.5[55] and placed in a dodecahedral volume with a minimum distance of 12 Å from the sides. The system was solvated with 7036 TIP3P waters, a sodium ion to maintain charge neutrality, energy minimized and finally relaxed for 4 nanoseconds at 310 K using 4-femtosecond time-step. A Langevin thermostat with a damping coefficient of 1/(1 picosecond) was employed to maintain temperature with the long-range electrostatics calculated using the particle mesh Ewald method (PME) at a grid spacing of 1.2 Å and a 10Å cut-off for non-bonded interactions. Five independent constant volume simulations were performed at 370 K with different starting velocities.

Results and Discussion

Thermodynamics. We calibrate the model by adjusting the basic parameters through exactly reproducing the experimental chemical denaturation midpoint (C_m ~2.3 M) and the sharpness of the unfolding transition (m-value ~14.8 kJ mol^{-1} M^{-1}) of bACBP at 278 K (see Methods). Though this methodology is guaranteed to reproduce the features of the unfolding curve, it is not clear if the temperature dependence of the experimental C_m can also be captured. Such a comparative exercise will provide a stringent test to the performance of the model and the robustness of the thermodynamic parameters within the experimental temperature range. To this end, we merely substitute the experimental temperatures (299 and 313 K) and predict the unfolding curves by keeping all other parameters constant. The resulting unfolding curves shift to lower C_m values as the conformational entropy term starts dominating that in turn decreases the stability with increasing temperatures

ARTICLE View Article Online DOI: 10.1039/C4CP06152K Journal Name

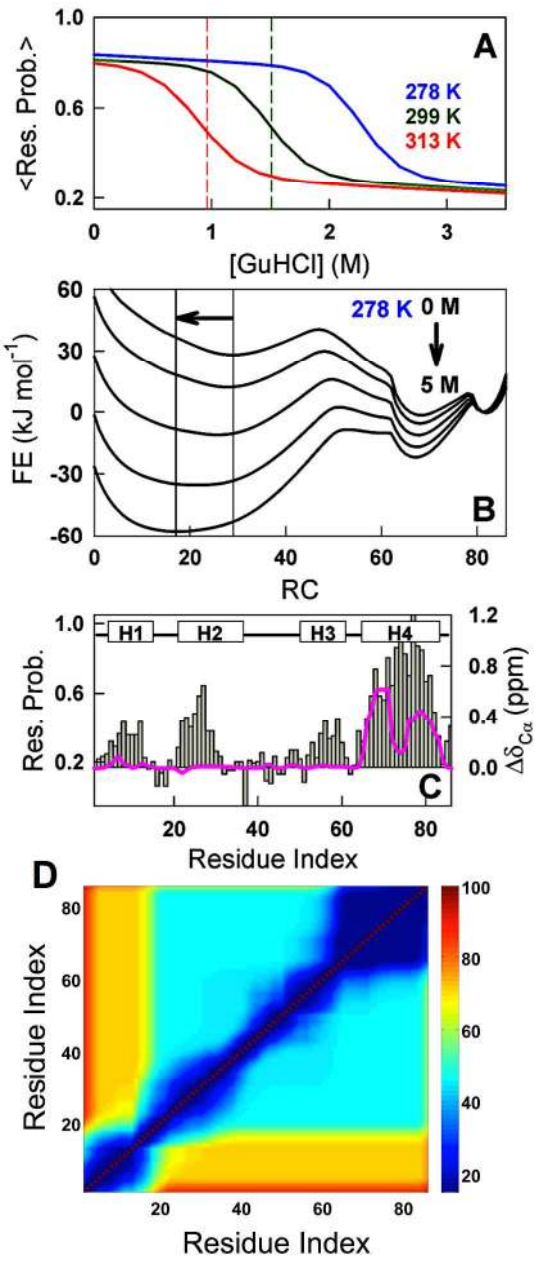

Fig. 2 (A) Mean residue-unfolding probabilities as a function of denaturant at different temperatures. The vertical dashed lines are the experimental chemical denaturation midpoint values. (B) 1D free energy profiles as a function of the number of structured residues as the reaction coordinate. The horizontal arrow highlights the continuous movement of the unfolded state upon changes in stability. (C) Comparison of the residue unfolding probabilities under destabilizing conditions (393 K; magenta and left axis) to the experimental chemical shift differences at low pH (bars and right axis). (D) Pair-wise residue contact probabilities represented as $-RT\ln(p_{ij})$ at 393 K as predicted by the WSME model. A spectral colour coding is employed going from low (blue) to high (red) free energy.

(Fig. 2A). The predicted C_m values are in very good agreement with experiments with a slope of one and intercept near zero (-0.11), attesting to the robustness of the method. In addition, the mean residue probabilities at zero denaturant concentration display a distinct trend wherein they decrease from 0.84 at 278 K to 0.79 at 313 K. Though small, this behaviour is not expected in a two-state system at low temperatures because the probability of the unfolded state is negligible. This suggests that the 'native state' of bACBP is not a distinct ensemble but changes with temperature and/or solvent conditions.

To identify if movement of the native state is continuous (i.e. in the absence of macroscopic free energy barrier) or if there is an exchange of population between two alternate native-like states separated by a barrier, we calculate one-dimensional free energy profiles as a function of the number of structured residues as the reaction coordinate. We find several interesting features simply from this projection that highlight the conformational plasticity of this protein (Fig. 2B). First, the profiles are characterized by three macrostates – native (*N*), partially structured intermediate (*I*) and unfolded (*U*) – indicative of a three-state-like behaviour with *I* as populated as *N* at 0 M GuHCl. This observation therefore suggests that the changes in the mean residue probability at low temperatures originate from an inter-conversion between the states *N* and *I*. The folding barrier height is predicted to be ~12.5 kJ mol^{-1} in the absence of denaturant while a free energy barrier of ~10 kJ mol^{-1} separates *I* and *N*. At midpoint conditions, the intermediate with ~68 structured residues is more populated than the native state. Second, a Hammond-like movement of the barrier top located at ~47 structured residues is evident and at ~5 M GuHCl, the barrier is quite broad when compared to the same at 0 M GuHCl. Third, the unfolded state possesses significant residual structure with ~30 residues adopting native-like conformations at 0 M. The position of the unfolded state is not constant and it progressively becomes more unstructured with increasing denaturant concentration, a feature that has been commonly observed in single-molecule FRET experiments of both ordered and disordered proteins.[56]

We calculate residue probabilities under highly destabilizing conditions from the partition function, i.e. the overall probability of finding a residue adopting a native-like conformation, to identify the regions that are structured in the unfolded state. Fourth helix is predicted to be the most structured under these conditions in agreement with chemical shift perturbation experiments under destabilizing conditions (Fig. 2C). Pair-wise residue contact probabilities (represented in Fig. 2D as $-RT*\ln(p_{ij})$) at these conditions imply large relative propensities for forming local contacts within the helical region (blue in Fig. 2D) while hinting at the possibility of transient inter-helix interactions involving helices 2, 3 and 4 consistent with paramagnetic relaxation enhancement measurements[57] and mutational studies in unfolded bACBP.[45] The 2D map in fact resembles similar estimates from both coarse-grained[58] and atomistic simulations[59]. The only difference between the previous treatments and the current depiction is that the interaction probabilities in the unfolded state are quite uniform for the three C-terminal helices in the latter case. This is mainly because of the mean-field nature of the van der Waals energy function, i.e. the interaction energies are assumed to be uniform and independent of the atom type.

Journal Name

View Article Online
DOI: 10.1039/C4CP06102K ARTICLE

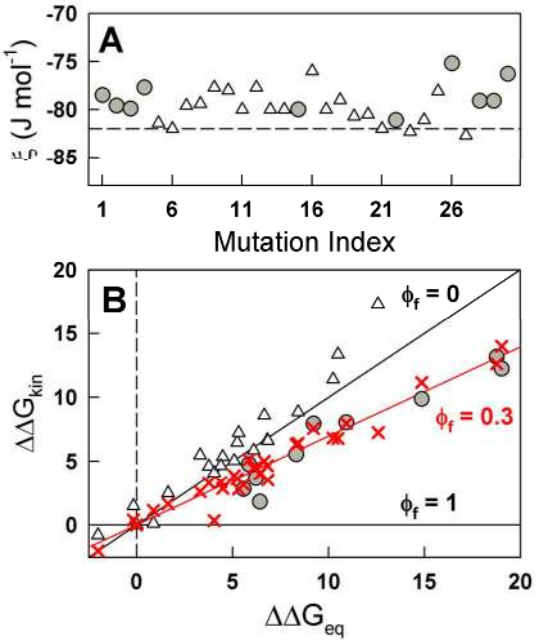

Fig. 3 (A) Mean-field vdW interaction energies that reproduce the experimental changes in stabilities upon point mutations.[50] (B) Predicted unfolding activation energies (red) shown as a Brønsted plot (energy units in kJ mol^{-1}). Mutations that result in classical and non-classical/zero ϕ-values experimentally are shown as filled circles and open triangles, respectively. The correlation line (red) indicates the predicted average ϕ-value of 0.3.

Kinetics. For the 1D free energy profile to be physically reasonable, it should also capture more relevant residue-level quantities and relaxation kinetics. To test this, we take advantage of protein engineering studies on bACBP wherein 30 mutations have been engineered throughout the protein to quantify the effect on both the equilibrium and kinetics, i.e. ϕ-values.[50] Out of 30 mutations, just 10 show classical phi-values ($0 < \phi < 1$), with a mean of folding ϕ-value 0.3, while the rest are near zero or negative. The mutational effects are introduced into the model by adjusting the vdW interaction energy (ξ) to exactly reproduce the magnitude of the equilibrium perturbation, i.e. $\Delta\Delta G_{eq}$, as prescribed before.[20] The resulting changes in ξ are quite small with a mean value of -79.5 J mol^{-1} and a standard deviation of just ~1.85 J mol^{-1} (Fig. 3A). A diffusive calculation with a denaturant-independent diffusion coefficient is then performed on this mutant free-energy surface to predict the relaxation rates (approximated by the largest non-zero eigenvalue; see Methods) and hence the changes in activation free energies. The resulting Brønsted plot presents a good agreement between the mean ϕ-value from the model and the classical experimental ϕ-values (compare red crosses and filled circles in Fig. 3B). However, none of the non-classical effects are reproduced. This can arise from either the simplistic nature of the model itself or due to the varied origins of non-classical phi-values that are challenging to interpret from the viewpoint of a Gō-model: non-native effects during folding,[60] in the denatured state[60, 61] or experimental noise.[62, 63]

Fluorescence kinetic experiments on bACBP at 299 K indicate that the refolding relaxation profiles is bi-phasic with a fast denaturant-independent rate constant of ~1/(100 μs) and a slow chevron-like denaturant dependent rate.[47] As discussed before, the one-dimensional free energy surface points to a continuous loss of structure in the unfolded state with increasing destabilization: but does the rate of this gradual process match with the experimental numbers? We approach this question by first estimating a fluorescence signal as a function of the one-dimensional reaction coordinate by using the mean number of contacts formed by W55 and W58 as a proxy (see Methods). The projected signal is predicted to be insensitive to the population of the native-like intermediate but sensitive to the changes in the unfolded state structure as can be seen from the dashed line in Fig. S1A (ESI†).

Diffusive kinetic calculations are then performed on the denaturant-dependent free-energy surface mimicking the experimental protocol to estimate the rates at 278, 299 and 313 K; the resulting population redistributions are weighted by the tryptophan signal to obtain kinetic decays (see Methods). The model predicts that the relaxation profiles are bi-phasic in the refolding limb and monophasic in the unfolding limb (Fig. S1B), an observation that matches exactly with experiments. [47] The faster rates vary between ~1/(100 μs) at 278 K to ~1/(15 μs) at 299 K while the chevron-like denaturant dependence of the slower rate is reproduced reasonably well (Fig. 4A). However, the slow rates predicted by the model under low denaturant concentrations systematically deviate from experiments. This could arise from the approximate nature of the modelled signal that is predicted to be insensitive to small population redistributions on the native side of the barrier or from the thermodynamic nature of the free energy profile that does not explicitly account for kinetic barrier effects that is expected to be more dominant under native-like conditions (i.e. underlying landscape roughness).[13, 64]

While an estimate of the tryptophan signal is required to simulate the experimental observations, the rate matrix constructed from the 1D probability distribution alone has sufficient information to reproduce the rates and their origins (Fig. S3). As an example, we plot the eigenvectors corresponding to the first three non-zero eigenvalues in the absence of denaturants at 278 in Fig. 4B. They nicely reproduce the three possible scenarios intuitively expected for a 1D free-energy profile with a single rate determining barrier and flexible end-states – population redistributions across the main barrier (black in Fig. 3B), on the folded side (green) and on the unfolded side (pink).

The good agreement between predicted changes in rates and experiments prompted us to look closer into the magnitude of the underlying barriers from the 1D free energy profiles. The rate determining folding barrier height of bACBP is predicted to be ~12.5 kJ mol^{-1} that increases to ~26.8 kJ mol^{-1} under chemical midpoint conditions at 278 K. This relative change in barrier height is in very good agreement with predictions from an independent structure-free approach that classifies protein folding behaviours based on the differences in barrier heights between two different stability conditions (see Fig. 4B in Ref. [16]). This unexpected correspondence suggests that the absolute

View Article Online
DOI: 10.1039/C4CP06152K
Journal Name

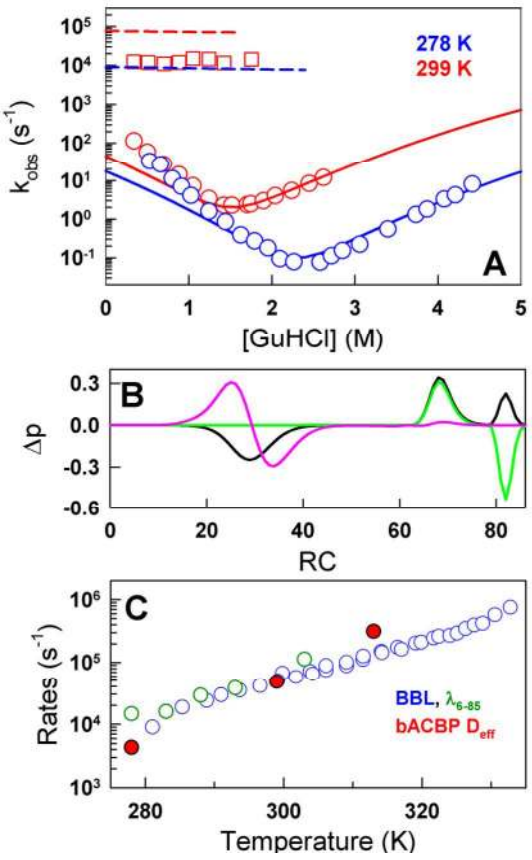

Fig. 4 (A) Comparison of experimental (circles) and predicted (continuous lines) chevron plots calculated solely from the 1D free energy profiles and an assumed fluorescence signal at two different temperatures. Squares and dashed lines represent the measured and predicted microsecond rates, respectively. (B) The eigenvectors corresponding to the first three non-zero eigenvalues at 0 M GuHCl, highlighting the origin of the slow (black), fast (green) and ultra-fast phases (pink). (C) Experimental folding relaxation dynamics of BBL (blue) and $\lambda_{6\text{-}85}$ (green) compared against the estimated D_{eff} for bACBP (red) using the rate equation.

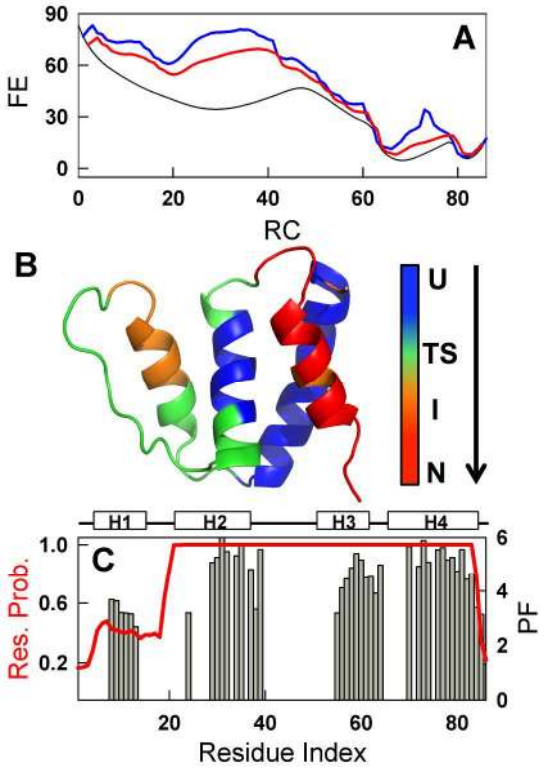

Fig. 5 (A) 1D free energy profiles from the exact solution (black), single- (blue) and double-sequence (red) approximations at 278 K. (B) Folding mechanism of bACBP mapped on to the structure shown in blue, green, orange and red in that order of formation; U - unfolded, TS – transition state, I – intermediate, N – native. (C) Experimental protection factors (bars) compared against the model residue unfolding probability (red). A cartoon of the secondary structures is shown on top.

barrier heights are well determined providing us an avenue to estimate the folding speed-limit (D_{eff}) from the perspective of a slow-folding protein. We combine the barrier height calculated from the model (ΔG) with the experimental midpoint relaxation rates at the midpoint (k) using the rate equation, i.e. from $k=2D_{eff}*exp(-\Delta G/RT)$, to estimate the speed-limit. The resulting D_{eff} and its temperature dependence agree very well with the dynamic relaxation rates of the one-state folding BBL over a large temperature range[65] and of the downhill 80-residue lambda repressor fragment $\lambda_{6\text{-}85}$ at low temperatures[66] (Fig. 4C). The implications of this observation are discussed below.

Folding Mechanism. The ability of the model to capture multiple experimental observations in a quantitative manner gives us sufficient confidence to provide a detailed picture of the folding mechanism from solely the statistical weights of the different microstates. However, a picture of the folding mechanism requires a map of the protein regions that are structured at different points along the reaction coordinate. To do so, we supplement the exact solution analysis by identifying the structure of microstates, *i.e.* the positions of 0s and 1s along the primary sequence, from both the single- and double-sequence approximations (SSA and DSA; see Methods). The 1D free energy profiles from these approximations are in agreement with the expectation from the exact solution profile (Fig. 5A). The microstates are ranked according to their probability at every value of the reaction coordinate and the most probable states are then mapped on to the native structure (Fig. 5B). It can be the seen that the unfolded state under stabilizing conditions possesses significant residual structure in both H2 and H4 (blue in Fig. 5B), though characterized by low probability, as they are high in free energy. The residual structure in the unfolded state now expands to encompass part of the long loop connecting H2 and H3, and the C-terminal portion of H3 including the loop following it (green in Fig. 5B) to reach the barrier top (transition-state structure; blue + green in Fig. 5B). The N-terminus of H3 now zips up in the downhill region of the main barrier (orange in Fig. 5B) to an intermediate state with well-folded helices H2, H3 and H4 (blue + green + orange in Fig. 5B). The helix H1 and the loop following it (residues 1-20) fold last, thus defining the native state.

The folding mechanism is in good agreement with several experimental measures particularly from NMR-based approaches. For example, PRE[58] and mutational studies[45] under destabilizing low pH conditions point to the possibility of H2 and H4 transiently interacting in the unfolded state. It is

View Article Online
DOI: 10.1039/C4CP06102K ARTICLE

important to emphasize here that we merely predict that the two regions (H2 and H4) are structured and not that they interact as the model construction allows for an interaction between two islands of structured residues if and only when all the intervening residues are also folded.[19, 20] Such an interaction could still require transient and reversible structure formation in the intervening regions under *folding* conditions that is yet to be tested experimentally. Hydrogen exchange experiments point to a very low protection factor for H1 compared to the other helices.[67] Interestingly, the equilibrium residue folding probability from the model is alone sufficient to capture this observation (Fig. 5C). Taken together with the 1D free energy profile, these results confirm that H1 is partially structured under stabilizing conditions and is the last to latch onto the rest of the protein during folding.

Electrostatics and Functional Implications. The fact that we have an experimentally constrained free energy profile allows us to ask some specific questions that are not possible otherwise. What structural-energetic factors determine the relative intermediate population and the height of the folding barrier? And, is this intermediate required for the functioning of bACBP, i.e. the binding to acyl CoA?

Differences in protein packing density (the number of van der Waals contacts per residue) correlate with the speed of folding and hence the barrier.[68] On the other hand, differences in the number and distribution of charged residues on the protein surface have also been shown to affect barriers,[25, 69] stability[69] and intermediate population[69] when comparing homologous members of the same family. The latter is particularly relevant given that the bACBP possesses a large percentage of charged residues (Fig. 1B). Therefore, to identify the dominant energy term (packing *vs.* electrostatics), we calculate the 1D free energy profile by setting the electrostatic term to zero. This exercise reveals that though vdW interactions alone can point to the possibility of intermediates, it significantly underestimates the population (green in Fig. 6A) while also simultaneously underestimating the folding barrier height. The implication is that charge-charge interactions are major players in determining the intermediate population and thermodynamic folding barrier height of bACBP. We delve deeper by estimating the pair-wise charge-charge interaction energy ($\Delta G_{q\text{-}q}$) from the Tanford-Kirkwood algorithm using the native structure as input.[70, 71] We find that helices 1 and 4 (H1 and H4) are involved in a complex array of favourable and unfavourable interactions (blue and red in Fig. 6B within the circled region and Fig. S2). However, intra-helical electrostatic interactions within H1 and H4 are stronger than inter-helical interactions - -8.2 kJ mol^{-1} for intra-H1, -49.8 kJ mol^{-1} for intra-H4 and a mere -2.1 kJ mol^{-1} for inter-H1-H4 charge-charge interactions – indicating that H1 is packed loosely against H4 despite their proximity. Interestingly, AGADIR[72] predicts that H1 can form a helix in isolation at pH 7.0 with a helicity that ranges from ~10% at 278 K to ~3% at 310 K which is significantly higher than the intrinsic helical propensity of H2 or H3 that are near zero.

We effectively deduce that a combination of weak packing, unfavourable electrostatics and intrinsic helical propensity all contribute to the presence of partial disorder in helix H1. This in turn manifests as an intermediate that is in dynamic equilibrium with the fully folded state on the folding side of the main barrier (Fig. 6C). It is however possible that the magnitude of electrostatic interactions and packing density could be correlated – for example, attractive interactions could pull the structural elements closer while repulsive interactions could push them apart (weak packing), or vice versa. This is possibly why the recent coarse-grained treatment of bACBP folding is able to identify a similar intermediate without any electrostatic term in its energetics (though the free energy

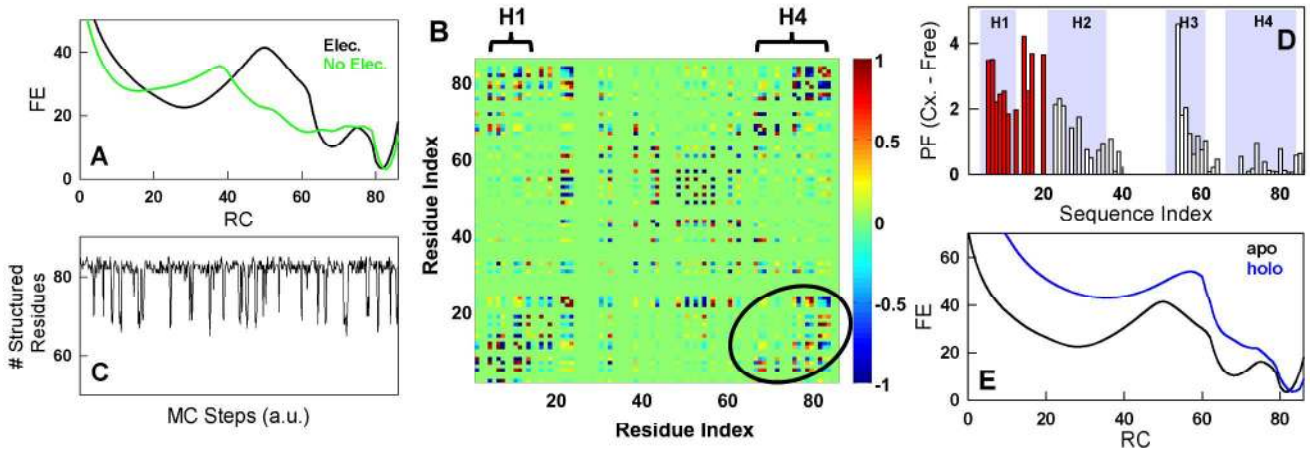

Fig. 6 All relevant energy units are in kJ mol^{-1}. (A) Folding free energy (FE) profiles of bACBP at pH 7 and 310 K with and without electrostatic term in the energy function. (B) Symmetric 2D map of the charge-charge interaction energy calculated from the TK algorithm. The circled area highlights the complex network of favourable (blue) and unfavourable (red) interactions between H1 and H4. The scale on the right is shown to vary from -1 to 1 for ease of visualization. (C) Results of a Metropolis Monte Carlo simulation run on the 1D free energy profile shown in black in panel A. (D) Experimental difference in protection factors between ligand bound (Cx./holo) and free/apo-form of bACBP (bars) highlighting that the major difference between the two is concentrated in the first helix and the loop (red bars) following it.[67] The shaded areas are the helical regions H1 to H4. (E) Free energy profiles of the apo- and holo-forms of bACBP calculated using identical model parameters from the respective native structures.

View Article Online
DOI: 10.1039/C4CP06102K

PCCP

RSCPublishing

ARTICLE

profiles have been generated only near the apparent midpoint).[58] This agreement between different methods (HX experiments, coarse-grained simulations and WSME model) on the identity of the intermediate suggests that it is a robust feature of the folding landscape and raises the question if this partial disorder is required for the non-covalent binding to acyl-CoA.

In this regard, the majority of protein-ligand interactions in holo-bACBP (PDB id: 1ACA) are grouped in two regions: between the C-terminal half of helix H1 to the N-terminal half of H2 (residues ~8-32) and the N-terminal half of helix H3 (residues ~49-55).[67] HX experiments that measure the accessibility of backbone amides to the solvent point to a large gain in rigidity of helix 1 and the loop following it upon binding the ligand in bACBP (red in Fig. 6D). In fact, there is very little change in the protection factor in the rest of the structure except for residue K54 that strongly interacts with the phosphate of the CoA.[67]

All this experimental evidence strongly indicates that the conformational flexibility of helix H1 is essential for binding to acyl-CoA. We propose that this conformational flexibility is imprinted on the sequence through unfavourable charge-charge interactions and/or weak packing of helix H1 with the rest of the structure providing evidence here that folding intermediates possibly arise from functional constraints. Interestingly, we find that the structure of holo-bACBP is more compact with ~5% more interactions and slightly different orientation of charged groups despite exhibiting a Cα-RMSD of ~1.6 Å with respect to apo-ACBP. This can be seen in the corresponding 1D free energy profile (Fig. 6E; calculated using parameters identical to that of apo-ACBP) that is highly tilted towards the native state with a miniscule population of the intermediate, reminiscent of conformational selection binding mechanism.

Folding Intermediate from Molecular Simulations. We further validate the structure of the predicted intermediate through all-atom explicit solvent MD simulations (see Methods). Five independent 500 nanosecond simulations were performed at 370 K with different starting velocities to identify regions of structure that are thermodynamically less stable. The expectation is that any weakly coupled region would unfold first at this high temperature thus promoting the sampling of partially structured states that can in turn be compared to WSME model predictions and experiments.

No major changes are observed in two of the five independent trajectories (grey in in Fig. 7A, 7B); two show reversible fluctuations with respect to the starting structure (blue and green in Fig. 7A, 7B) and one other trajectory points to a large irreversible change in Cα-RSMD within the simulated

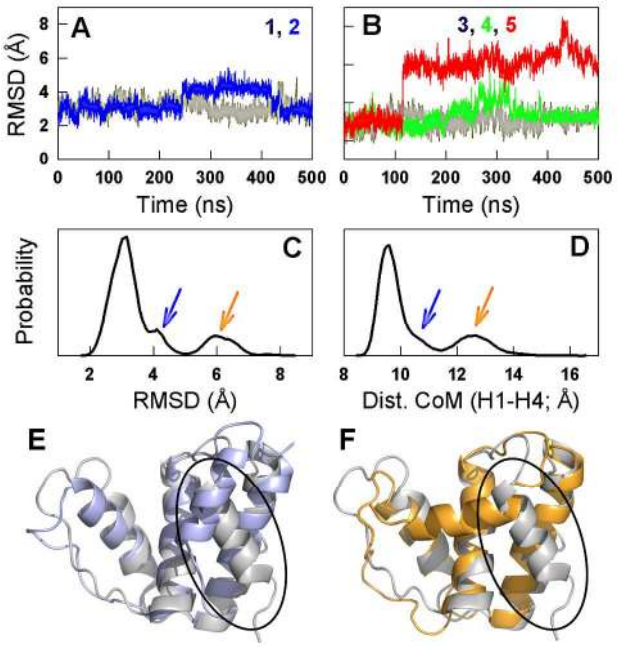

Fig. 7 (A & B) Cα-RMSD with respect to the PDB structure 2ABD from 5 independent explicit solvent MD simulation runs at 370 K. (C & D) Distribution of Cα-RMSD and centre of mass (CoM) distances between helices H1 and H4. (E & F) Comparison of the partially structured states corresponding to the arrows indicated in panels C and D with respect to the starting structure (grey). H1 in the native structure is marked with a black ellipse.

timescale of 500 nanoseconds (red in Fig. 7B). A simple binning analysis indicates that the mean Cα-RMSD is ~3 Å (Fig. 7C); a major contribution to this comes from the long unstructured loop connecting the structural elements H2 and H3 that undergoes enhanced thermodynamic fluctuations. It also points to the possibility of two partially structured states, one at ~4-4.5 Å and the other at ~6-6.5 Å (arrows in Fig. 7C, 7D). A centre of mass (CoM) distance calculation between H1 and H4 is able to reproduce the shape of the global RMSD plot indicating that major structural changes are concentrated in these regions.

Structural analysis reveals that the helix H1 partially unfolds particularly at the N-terminus and samples additional conformational states that either interacts only weakly with H4 (RMSD of 4 Å; Fig. 7E) or completely detaches from H4 (RMSD 6 Å; Fig. 7F) with the centre of mass distance ranging from 11 Å to as much as 14 Å (compared to the starting CoM distance of ~10 Å). This direct observation is in very good agreement with the nature of the intermediate predicted from the WSME model, the expectation from the electrostatic and intrinsic helical propensity calculations and HX experiments. It is important to note that these simulations do not provide

View Article Online
DOI: 10.1039/C4CP06102K ARTICLE

accurate equilibrium probabilities, as they are quite short compared to the overall folding time, but clearly hint that an unfolding pathway involving detachment of H1 and H4 exists in agreement with the model predictions.

Conclusions

The relative organization of residues in the folded state is entirely determined by functional constraints because of which both topological and energetic frustrations arise.[73, 74] In a recent work, we have shown strong evidence from multiple computational methods that the folding mechanism of Barstar is proportionately influenced by functional constraints to range from multi-state to two-state folding upon just single point mutations.[28] Accordingly, the partial disorder in H1 of bACBP inscribed onto the sequence through both local (helical propensity) and nonlocal effects (weak packing) could point to a possible avenue taken by Nature to fine-tune the binding affinity of bACBP to acyl-CoA. This could be a feature selected through Natural Selection as ACBP is a highly conserved protein across all eukaryotes and is known to play an important role in lipid metabolism though its precise functions and the mechanism of action are unclear.[75, 76] Experimentally, it is known that bACBP is highly promiscuous and binds an array of acyl-CoA esters with the carbon chain length varying from 2 to 16 with dissociation constants from micromolar to sub-picomolar, respectively.[77] We propose that this promiscuous binding of varied carbon chain lengths is an offshoot of the partially disordered nature of H1. Direct evidence for this comes from the observation of a gain in rigidity of H1 from HX experiments[67] and an indirect evidence from the more compact nature of the holo-bACBP compared to apo-bACBP; it should be possible to experimentally verify this by quantifying the degree of protection of backbone amides for different substrates.

From a folding mechanistic viewpoint, a variety of models, ranging from chemical models to all-atom simulations are generally employed to characterize protein-folding processes. There is unfortunately a large trade-off between the predictive powers of different models and the ability to reproduce multiple experimental observables in a (semi-)quantitative manner that is a difficult test for the model parameters and assumptions. For example, chemical models, though simple are rarely predictive and are usually employed to extract relevant thermodynamic or kinetic parameters. Simulations, on the other hand, are capable of providing detailed pictures of folding mechanisms but lag behind in their ability to reproduce the global unfolding thermodynamics (for example, heat capacity profiles) and are generally time-intensive. Here, we have shown that the Ising-like statistical mechanical model provides a good compromise between the two requirements above. The 1D free energy profile as a function of the natural coordinate of the model, the number of structured residues, is able to predict several experimental observations in bACBP – the height of the folding barrier, nature of the intermediate, moving unfolded state and its structure, phi-values, chevron kinetics and the fast-phase. Apart from these, the model is able to capture the temperature dependence of the changes in stability induced by denaturant in this domain. It is important to emphasize that the WSME model is constrained by a single equilibrium-unfolding curve (Fig. 2A) while the rest are predictions that naturally follow. These results attest to both the robustness of the 1D free energy profile, the magnitude of the Gō-like energetic terms and hence the overall approach.

The rate-determining folding barrier height is predicted to be well above *3RT* for bACBP under all conditions, thus placing it in the two-state-like regime. We find a very good agreement between the bACBP folding pre-exponential at different temperatures and the relaxation rates of proteins that have been shown to fold downhill, BBL and $\lambda_{6\text{-}85}$, thus validating the magnitude of the main barrier in bACBP (Fig. 4C). The magnitude of the activation energy on the D_{eff} of bACBP (~1.08 kJ mol^{-1}) is similar to the ~1 kJ mol^{-1} per residue estimate extracted by characterizing the temperature-dependent folding kinetics of several fast-folding proteins using a phenomenological free energy surface model.[16] This unexpected agreement between proteins that are characterized by different structures, functions, and that fold at different speeds hints that the underlying dynamic factors that determine the activation energy are universal. It is of interest note that this dynamic term, also termed internal friction or landscape roughness, has recently been revealed to originate from microscopic barriers to peptide bond rotations.[78, 79]

In the analysis of mutational data, the correspondence between the average ϕ-value estimated from the 1D free energy profile without explicitly modelling the mutations, and the mean number obtained from a global analysis of experimental mutational effects (~0.3[63]) provides additional evidence to the physically sound nature of the WSME model. In parallel, this agreement hints at a possible universal parameter that the mean ϕ-value signifies - the average stabilization free energy gained by the protein at the barrier top during folding – that has now been observed from four different approaches including the current treatment.[2, 80, 81]

Though the main barrier of bACBP is large we find that the folding behaviour is far from a two-state phenomenon. Apart from the slowest rate that follows a chevron-like features, experiments indicate an additional microsecond phase during refolding and this has been interpreted as originating from the population of an on- or off-pathway intermediate.[47] We show here that the unfolded state of bACBP is flexible and changes its structure in a continuous manner thus giving rise to a faster phase in response to rapid change in conditions. This is in agreement with a combined experimental-simulation study that points to a structural reorganization in the unfolded state at similar time-scales.[82] On the other hand, we identify a well-defined intermediate characterized by partial structure in helix H1 on the folded side of the main barrier whose structure is in agreement with HX experiments on bACBP with and without the ligand (Fig. 6) and all-atom MD simulations (Fig. 7). Using the 1D free energy profile from the WSME model, we estimate that the rate constant for the transition between the fully folded

View Article Online
DOI: 10.1039/C4CP06102K
Journal Name

and partially structured state should be ~630 s^{-1} at 278 K (Fig. S3), a prediction that can be tested kinetically employing a fluorescent reporter that probes the packing between the helices H1 and H4.

In summary, we show that coarse features of the 1D free energy profile are sufficient to capture several macroscopic features of the folding behaviour in a large single-domain protein bACBP. A more detailed picture of the structure of intermediates and folding mechanism is obtained by a small increase in dimensionality – in the current work this is highlighted by the enumeration of single- and two-stretches of folded residues. We further underline the fact that global equilibrium thermodynamic observables are sufficient to constrain computational methods, an aspect that we are increasingly observing for several systems from the perspective of the WSME model.[21, 27, 42] We believe that the detailed WSME model with multiple energetic terms is now poised to address questions on not just folding mechanisms but also to quantify the energetic consequences of small functionally-driven differences in primary sequences that is otherwise not apparent in structural comparisons alone. This might require a more subtle correction to the underlying energy function by introducing residue-specific or secondary-structure-specific conformational entropy terms guided by experimental observables.

Acknowledgements

We thank Martin Gruebele for critical comments on the manuscript.

Notes and References

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

†Electronic Supplementary Information (ESI) available: Additional figures supporting the conclusions. See DOI: 10.1039/b000000x/

1. J. D. Bryngelson, J. N. Onuchic, N. D. Socci and P. G. Wolynes, *Proteins*, 1995, **21**, 167.
2. A. Akmal and V. Muñoz, *Proteins*, 2004, **57**, 142.
3. V. Muñoz, *Ann. Rev. Biophys. Biomol. Struct.*, 2007, **36**, 395.
4. D. D. Boehr, R. Nussinov and P. E. Wright, *Nat. Chem. Biol.*, 2009, **5**, 789.
5. K. A. Dill and H. S. Chan, *Nat. Struct. Biol.*, 1997, **4**, 10.
6. M. Oliveberg and P. G. Wolynes, *Q. Rev. Biophys.*, 2005, **38**, 245.
7. Y. Levy, S. S. Cho, J. N. Onuchic and P. G. Wolynes, *J. Mol. Biol.*, 2005, **346**, 1121.
8. L. Mirny and E. Shakhnovich, *Ann. Rev. Biophys. Biomol. Struct.*, 2001, **30**, 361.
9. C. Hyeon and D. Thirumalai, *Nat. Commun.*, 2011, **2**, 487.
10. D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. B. Shan and W. Wriggers, *Science*, 2010, **330**, 341.
11. R. B. Best and G. Hummer, *J. Phys. Chem. B*, 2009, **113**, 9004.
12. J. N. Onuchic, Z. LutheySchulten and P. G. Wolynes, *Ann. Rev. Phys. Chem.*, 1997, **48**, 545.
13. N. D. Socci, J. N. Onuchic and P. G. Wolynes, *J. Chem. Phys.*, 1996, **104**, 5860.
14. V. Muñoz and J. M. Sanchez-Ruiz, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 17646.
15. A. N. Naganathan, J. M. Sanchez-Ruiz and V. Muñoz, *J. Am. Chem. Soc.*, 2005, **127**, 17970.
16. A. N. Naganathan, U. Doshi and V. Muñoz, *J. Am. Chem. Soc.*, 2007, **129**, 5673.
17. F. Liu, C. Dumont, Y. J. Zhu, W. F. DeGrado, F. Gai and M. Gruebele, *J. Chem. Phys.*, 2009, **130**, 061101.
18. M. Gruebele, *Curr. Opin. Struct. Biol.*, 2002, **12**, 161.
19. H. Wako and N. Saito, *J. Phys. Soc. Japan*, 1978, **44**, 1939.
20. V. Muñoz and W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 11311.
21. P. Bruscolini and A. N. Naganathan, *J. Am. Chem. Soc.*, 2011, **133**, 5372.
22. M. M. Garcia-Mira, M. Sadqi, N. Fischer, J. M. Sanchez-Ruiz and V. Muñoz, *Science*, 2002, **298**, 2191.
23. J. Kubelka, E. R. Henry, T. Cellmer, J. Hofrichter and W. A. Eaton, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 18655.
24. A. N. Naganathan, R. Perez-Jimenez, V. Muñoz and J. M. Sanchez-Ruiz, *Phys. Chem. Chem. Phys.*, 2011, **13**, 17064.
25. A. N. Naganathan, *J. Chem. Theory Comput.*, 2012, **8**, 4646.
26. A. N. Naganathan and V. Muñoz, *J. Phys. Chem. B*, 2014, **118**, 8982.
27. A. N. Naganathan and M. Orozco, *J. Phys. Chem. B*, 2013, **117**, 13842.
28. A. N. Naganathan, J. M. Sanchez-Ruiz, S. Munshi and S. Suresh, *J. Phys. Chem. B*, 2015, **119**, 1323.
29. A. N. Naganathan, *WIREs Comput. Mol. Sci.*, 2013, **3**, 504.
30. A. N. Naganathan, P. Li, R. Perez-Jimenez, J. M. Sanchez-Ruiz and V. Muñoz, *J. Am. Chem. Soc.*, 2010, **132**, 11183.
31. E. R. Henry, R. B. Best and W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, 17880.
32. A. N. Naganathan, *J. Phys. Chem. B*, 2013, **117**, 4956.
33. H. S. Chan, Z. Zhang, S. Wallin and Z. Liu, in *Ann. Rev. Phys. Chem.*, eds. S. R. Leone, P. S. Cremer, J. T. Groves and M. A. Johnson2011, vol. 62, pp. 301.
34. U. Doshi and V. Muñoz, *Chem. Phys.*, 2004, **307**, 129.
35. V. Muñoz, P. A. Thompson, J. Hofrichter and W. A. Eaton, *Nature*, 1997, **390**, 196.
36. E. R. Henry and W. A. Eaton, *Chem. Phys.*, 2004, **307**, 163.
37. G. Schreiber, A. M. Buckle and A. R. Fersht, *Structure*, 1994, **15**, 945.
38. M. Jager, Y. Zhang, J. Bieschke, H. Nguyen, M. Dendle, M. E. Bowman, J. P. Noel, M. Gruebele and J. W. Kelly, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 10648.
39. H. R. Kalbitzer, M. Spoerner, P. Ganser, C. Hozsa and W. Kremer, *J. Am. Chem. Soc.*, 2009, **131**, 16714.
40. A. N. Naganathan and M. Orozco, *J. Am. Chem. Soc.*, 2011, **133**, 12154.
41. J. A. Lamboy, H. Kim, K. S. Lee, T. Ha and E. A. Komives, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 10178.
42. S. Sivanandan and A. N. Naganathan, *PLOS Comp. Biol.*, 2013, **9**, e1003403.

View Article Online
DOI: 10.1039/C4CP06102K ARTICLE

43. B. B. Kragelund, C. V. Robinson, J. Knudsen, C. M. Dobson and F. M. Poulsen, *Biochemistry*, 1995, **34**, 7217.
44. K. Teilum, F. M. Poulsen and M. Akke, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 6877.
45. S. W. Bruun, V. Iesmantavicius, J. Danielsson and F. M. Poulsen, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 13306.
46. W. Fieber, S. Kristjansdottir and F. M. Poulsen, *J. Mol. Biol.*, 2004, **339**, 1191.
47. K. Teilum, K. Maki, B. B. Kragelund, F. M. Poulsen and H. Roder, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 9807.
48. A. D. Robertson and K. P. Murphy, *Chem. Rev.*, 1997, **97**, 1251.
49. H. Taketomi, Y. Ueda and N. Go, *Inter. J. Prot. Pep. Res.*, 1975, **7**, 445.
50. B. B. Kragelund, P. Osmark, T. B. Neergaard, J. Schiodt, K. Kristiansen, J. Knudsen and F. M. Poulsen, *Nat. Struc. Biol.*, 1999, **6**, 594.
51. J. K. Myers, C. N. Pace and J. M. Scholtz, *Protein Sci.*, 1995, **4**, 2138.
52. L. J. Lapidus, P. J. Steinbach, W. A. Eaton, A. Szabo and J. Hofrichter, *J. Phys. Chem. B*, 2002, **106**, 11628.
53. A. Narayan and A. N. Naganathan, *J. Phys. Chem. B*, 2014, **118**, 5050.
54. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins: Struct. Funct. Bioinf.*, 2010, **78**, 1950.
55. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845.
56. S. Müller-Späth, A. Soranno, V. Hirschfeld, H. Hofmann, S. Ruegger, L. Reymond, D. Nettels and B. Schuler, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 14609.
57. S. Kristjansdottir, K. Lindorff-Larsen, W. Fieber, C. M. Dobson, M. Vendruscolo and F. M. Poulsen, *J. Mol. Biol.*, 2005, **347**, 1053.
58. V. Ozenne, J. K. Noel, P. O. Heidarsson, S. Brander, F. M. Poulsen, M. R. Jensen, B. B. Kragelund, M. Blackledge and J. Danielsson, *J. Mol. Biol.*, 2014, **426**, 722.
59. K. Lindorff-Larsen, S. Kristjansdottir, K. Teilum, W. Fieber, C. M. Dobson, F. M. Poulsen and M. Vendruscolo, *J. Am. Chem. Soc.*, 2004, **126**, 3291.
60. D. P. Raleigh and K. W. Plaxco, *Prot. Pept. Lett.*, 2005, **12**, 117.
61. J. H. Cho and D. P. Raleigh, *J. Am. Chem. Soc.*, 2006, **128**, 16492.
62. I. E. Sanchez and T. Kiefhaber, *J. Mol. Biol.*, 2003, **334**, 1077.
63. A. N. Naganathan and V. Muñoz, *Proc. Natl. Acad. Sci. U.S.A*, 2010, **107**, 8611.
64. J. Chahine, R. J. Oliveira, V. B. P. Leite and J. Wang, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 14646.
65. P. Li, F. Y. Oliva, A. N. Naganathan and V. Muñoz, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 103.
66. S. J. DeCamp, A. N. Naganathan, S. A. Waldauer, O. Bakajin and L. J. Lapidus, *Biophys. J.*, 2009, **97**, 1772.
67. B. B. Kragelund, J. Knudsen and F. M. Poulsen, *J. Mol. Biol.*, 1995, **250**, 695.
68. G. Zuo, J. Wang and W. Wang, *Proteins*, 2006, **63**, 165.
69. O. Halskau, R. Perez-Jimenez, B. Ibarra-Molero, J. Underhaug, V. Muñoz, A. Martinez and J. M. Sanchez-Ruiz, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, 8625.
70. C. Tanford and J. G. Kirkwood, *J. Am. Chem. Soc.*, 1957, **79**, 5333.
71. B. Ibarra-Molero, V. V. Loladze, G. I. Makhatadze and J. M. Sanchez-Ruiz, *Biochemistry*, 1999, **38**, 8138.
72. V. Muñoz and L. Serrano, *Nat. Struct. Biol.*, 1994, **1**, 399.
73. D. U. Ferreiro, E. A. Komives and P. G. Wolynes, *Q. Rev. Biophys.*, 2014, **47**, 285.
74. D. U. Ferreiro, J. A. Hegler, E. A. Komives and P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 19819.
75. N. J. Faergeman, M. Wadum, S. Feddersen, M. Burton, B. B. Kragelund and J. Knudsen, *Mol. Cell. Biochem.*, 2007, **299**, 55.
76. M. Burton, T. M. Rose, N. J. Faergeman and J. Knudsen, *Biochem. J.*, 2005, **392**, 299.
77. N. J. Faergeman, B. W. Sigurskjold, B. B. Kragelund, K. V. Andersen and J. Knudsen, *Biochemistry*, 1996, **35**, 14118.
78. D. de Sancho, A. Sirur and R. B. Best, *Nat. Commun.*, 2014, **5**, 4307.
79. I. Echeverria, D. E. Makarov and G. A. Papoian, *J. Am. Chem. Soc.*, 2014, **136**, 8708.
80. V. Muñoz, M. Sadqi, A. N. Naganathan and D. de Sancho, *HFSP Journal*, 2008, **2**, 342.
81. A. N. Naganathan and M. Orozco, *Phys. Chem. Chem. Phys.*, 2011, **13**, 15166.
82. V. A. Voelz, M. Jager, S. H. Yao, Y. J. Chen, L. Zhu, S. A. Waldauer, G. R. Bowman, M. Friedrichs, O. Bakajin, L. J. Lapidus, S. Weiss and V. S. Pande, *J. Am. Chem. Soc.*, 2012, **134**, 12565.