Protein structure prediction: Do hydrogen bonding and water-mediated interactions suffice?

Vanessa Oklejas [c,*,1], Chenghang Zong [b,1], Garegin A. Papoian [a], Peter G. Wolynes [c]

[a] Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599, United States
[b] Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, United States
[c] Department of Chemistry and Biochemistry, University of California, La Jolla, CA 92093-0371, United States

ARTICLE INFO

ABSTRACT

The many-body physics of hydrogen bond formation in alpha-helices of globular proteins was investigated using a simple physics-based model. Specifically, a context–sensitive hydrogen bond potential, which depends on residue identity and degree of solvent exposure, was used in the framework of the Associated Memory Hamiltonian codes developed previously but without using local-sequence structure matches ("memories"). Molecular dynamics simulations employing the energy function using the context–sensitive hydrogen bond potential alone (the "amnesiac" model) were used to generate low energy structures for three alpha-helical test proteins. The resulting structures were compared to both the X-ray crystal structures of the test proteins and the results obtained using the full Associated Memory Hamiltonian previously used. Results show that the amnesiac Hamiltonian was able to generate structures with reasonably high structural similarity ($Q \sim 0.4$) to that of the native protein but only with the use of predicted secondary structure information encoding local steric signals. Low energy structures obtained using the amnesiac Hamiltonian without any *a priori* secondary structure information had considerably less similarity to the native protein structures ($Q \sim 0.3$). Both sets of results utilizing the amnesiac Hamiltonian are poorer than when local-sequence structure matches are used.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The formation of hydrogen bonds in the backbone has long been recognized as a key aspect in stabilizing the native structures of globular proteins. This idea was the basis of the early predictions about protein structure made by Pauling and Kauzmann which pre-dated crystal structure determination [1,2]. Hydrogen bonds between polar amine and carboxyl groups alleviate the desolvation penalty of those groups as they become buried in protein's native structure. This, in turn, gives rise to the familiar protein secondary structures, such as alpha-helices and beta-sheets. Furthermore, formation of the secondary structures in the molten globular state substantially diminishes the search space for the folded state, accelerating the folding kinetics [3–5]. Despite the clear importance of hydrogen bonding in protein folding thermodynamics and kinetics, and the enormous body of prior work in this area, a significant number of outstanding questions remain about the precise energetics of hydrogen bond formation among various pairs of amino-acid residues in the context of the local environment (e.g.,

whether there is a high density of neighboring residues or significant exposure to solvent.) Certain aspects of hydrogen bond energetics, owing to the solvent involvement should lead to non-trivial cooperative or anti-cooperative effects. One therefore wonders whether encoding the physics of hydrogen bond formation upon collapse might actually be sufficient to predict protein structure. In this work, we probe the many-body physics of hydrogen bond formation in alpha-helices of globular proteins by introducing corresponding context–sensitive hydrogen bond potentials as additional terms to a protein structure prediction Hamiltonian already developed in our group and evaluating the performance of this simple physics model.

Protein structure prediction potentials that incorporate knowledge of local-sequence structure patterns have become remarkably successful over the last decade. There are a number of features which are common to the potentials used by various groups. In addition to chain connectivity, usually both local and tertiary interactions enter as distinct terms in these Hamiltonians. In prior work by Saven and Wolynes, it was shown that local and tertiary interactions likely contribute nearly equally to the overall specificity of native folds [6]. There are a number of ways by which local interactions are treated in structure prediction Hamiltonians. In both fragment assembly-like methods [7,8] and the Associated Memory Hamiltonian [9–12], local structural signals are inferred by finding

* Corresponding author.
  E-mail addresses: voklejas@ucsd.edu (V. Oklejas), pwolynes@ucsd.edu (P.G. Wolynes).
[1] These authors contributed equally to this work.

distant or close sequence similarity to corresponding local segments of many other proteins in the structural database. In recent developments of the Associated Memory Hamiltonian, called the AMH, memory based guidance of local structure formation is supplemented by direct and water-mediated interactions which are burial-specific [13,14]. We call this the AMW model. Interestingly, although no direct hydrogen bonding potential was included in the alpha-helical AMW prediction code, highly native-like alpha-helices form in successful prediction runs, driven by the secondary structure bias from the memory proteins [13]. In this work, conversely we turn off these memories, but instead add direct many-body hydrogen bonding potential to the resulting memory-less ("amnesiac") Hamiltonian. We colloquially use the latter phrase to distinguish and compare the present results with the alpha-helical AMW results published previously.

Collapse of a random coil to a molten globule has a number of important consequences for the thermodynamics and kinetics of protein folding. In particular, local steric constraints and excluded volume interactions reduce configurational entropy and thus facilitate the formation of flickering alpha-helical segments which become partially aligned, reminiscent of a liquid crystal [4]. When local structural signals are added for either helix formation, helix capping, or a turn formation, via specific local-sequence information, the resulting bias towards native-like secondary structures contributes to the minimal frustration of the folded conformation and significantly diminishes the entropy of the globule, which allows for more efficient folding kinetics [6]. As explained in Section 2, some of these local signals are included in the alpha-helical hydrogen bonding term of the amnesiac Hamiltonian, by modulating the specific hydrogen bonding strength depending on the specific amino-acid residue pair. This approach could also potentially be used to model some of the helix capping effects involving side chain-backbone hydrogen bonds. On the other hand, the present hydrogen bonding potential is non-pairwise only insofar as local polypeptide chain density is concerned, while additional cooperative effects may potentially also be important. Explicit turn signals, on the other hand, are also completely absent in the amnesiac Hamiltonian we study here. We address the question of relative importance of these effects in the present study, by comparing the amnesiac Hamiltonian results with the AMW calculations.

Another interesting question addressed in the present work is the effect of modulating the strength of hydrogen bonding based on burial. According to the original Kauzmann arguments, when amine and carboxyl groups are well solvated, there should be little driving force for hydrogen bond formation. Upon desolvation in the protein core or partial desolvation within an alpha-helix, the energetics of hydrogen bond formation becomes important. Since local polypeptide chain density and hydrogen bond formation become coupled under this scenario, this is expected to introduce additional cooperativity into the folding process. In the amnesiac Hamiltonian, the alpha-helical hydrogen bonding term is modulated by the local surrounding density, so we can independently vary the strengths of core and surface hydrogen bonds. While we keep the core hydrogen bonding always stabilizing, we compare the quality of structure prediction runs for various proteins as the surface hydrogen bonding energies are either energetically favorable, neutral or unfavorable.

In summary, in this work we have developed a many-body alpha-helical hydrogen bonding potential in the context of a structure prediction potential. We have compared folding runs and free energy profiles obtained from the amnesiac Hamiltonian with AMW results, where in the latter protein memories are used to guide local alpha-helical and turn formation processes. This comparison indicates that the performance of the new amnesiac Hamiltonian, when combined with secondary structure prediction bias,

is although not at par with AMW, which includes complex sequence-specific local interactions, still rather predictive of native-like conformations but that hydrogen bonding and water-mediated interactions alone are insufficient to predict protein structure. These findings are somewhat unexpected, suggesting that introducing additional physical interactions may allow one to somewhat simplify current knowledge-based structure prediction Hamiltonians. This may consequently allow application of this potential to polypeptide chains where truly novel structural patterns exist locally, for example to study conformational dynamics of intrinsically-disordered proteins.

2. Methods

The infrastructure of our structure prediction program is based on the Associative Memory Hamiltonian with water mediation (AMW) [13]. There are four components in the Hamiltonian: a backbone term, a water-mediated term for medium/long range interactions (sequence distance larger than 8) and two hydrogen bonding potential terms.

$$E_0 = E_{\text{backbone}} + E_{\text{med-long}} + E_{\text{helical}} + E_{\text{Kauzmann}}. \qquad (1)$$

In our Hamiltonian, the atoms C_α, C_β and O are explicitly represented with chain connectivity. The backbone angles are controlled by a Ramachandran potential and chirality potential. Excluded volume is treated for the explicitly represented atoms in the system. The exposure or buriedness of any residue is calibrated based on the residue density surrounding it, which is calculated on-the-fly. The medium/long range interactions between residues depend on the residue density surrounding the interaction as well. Here we present the details of helical hydrogen bond potential and how water mediation is integrated into the potential.

We focus on implementing the water mediation into the hydrogen bond potential that controls the formation of helices. As a result, the exposure or degree of burial of the residue is used to characterize the environment of the forming hydrogen bond as was done earlier in a study by Takada without water-mediated interactions [15,16]. The memory term used in previous studies [13] is replaced by the water-mediated helical potential in our Hamiltonian, which we denote as the "amnesiac" Hamiltonian.

We use the following hydrogen bonding potential to control helical formation:

$$E_{\text{helical}} = \gamma_p \times V_{\text{helical}}^{i,i+4}(r_{\text{ON}}, r_{\text{OH}}) \times \sigma(i, i+4) \qquad (2)$$

where E_{helical} is our helical hydrogen bond potential. And $\sigma(i, i+4) = H_{\text{burial}}(\rho_i) \times H_{\text{burial}}(\rho_{i+4})$ is the burial profile term of the pairwise interaction. ρ is the local density of each residue calculated in the water potential.

$$V_{\text{helical}}^{i,i+4}(r_{\text{ON}}, r_{\text{OH}}) = -(f(i) + f(i+4))$$
$$\times \exp\left[\frac{-\left(r_{i,i+4}^{\text{ON}} - <r^{\text{ON}}>\right)^2}{2\sigma_{\text{NO}}^2} - \frac{\left(r_{i,i+4}^{\text{OH}} - <r^{\text{OH}}>\right)^2}{2\sigma_{\text{HO}}^2}\right]$$
$$(3)$$

The term $f(i)$ describes the probability of finding the corresponding residue type in helices from the protein database [17]. The hydrogen bond stability is proportional to the sum of helical propensities of the two interacting residues.

The overall energy scale of helical hydrogen bond potential is denoted by γ_p, which has been optimized using training proteins based on the minimal frustration principle [18].

The above potential directs the hydrogen bonding between different pairs of residues when both are buried. As residues become exposed, the free energy gain due to hydrogen bond formation will
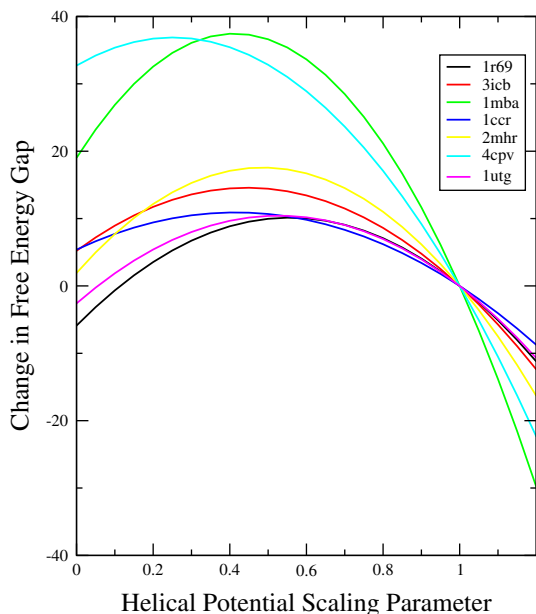
Fig. 1. Optimization results of the overall energy scaling parameter for $E_{helical}$.

Fig. 2. Crystal structure of 1UTG protein with buried (red) and exposed (blue) hydrogen bonds colored according to the amnesiac potential.

decrease to reflect the fact that water molecules can also participate in hydrogen bond formation. For $\gamma_w < 0$, this free energy for intra-helical hydrogen bond formation can become negative (anti-helical), which we dub as the "Kauzmann effect". The following potential is given:

$$E_{Kauzmann} = \gamma_w \times V_{helical}^{i,i+4}(r_{ON}, r_{OH}) \times (1 - \sigma(i, i + 4)) \qquad (4)$$

where $V_{helical}$ was described by (3).

Seven alpha-helical proteins were used for training the new parameters. In this procedure, we first simply set γ_p and γ_w as 2.5 and −2.5 respectively. Simulations of seven training proteins were performed to generate ensembles of native-like structures and protein structures in the molten-globule state using the amnesiac Hamiltonian both with and without a secondary structure bias. Before the simulations were started, a secondary structure prediction for each protein was obtained [19]. For simulations employing a secondary structure bias, the results of the secondary structure predictions were used to bias the Ramachandran potential in favor of the predicted secondary structure. The cumulant expansion of free energy with respect to parameter γ_p and γ_w was performed. The optimal value is chosen at the position where the free energy gap between native-like structure ensemble and molten-globule structure ensemble is maximized. Based on the optimization result, γ_p is approximately 2.0 and γ_w is −1.0 (see Fig. 1).

In the force calculations, the derivative of the potential is calculated for every explicit atom involved in the potential. The derivative follows the chain rule calculation in the hydrogen bond potential. The forces are transferred from the implicit nitrogen atom to the explicit atoms C_α, C_β and O_x in the planar scaffold of the amino acid. For the H_{burial} terms, the forces will be applied to the relevant residues in proximity. Therefore our helical potential is a non-additive potential, instead of a simple two-body potential.

3. Results and discussion

Molecular dynamics annealing simulations were performed using the amnesiac Hamiltonian to generate low energy conformations of target proteins. Three relatively small alpha-helical test proteins were used to test the performance of the amnesiac Ham-
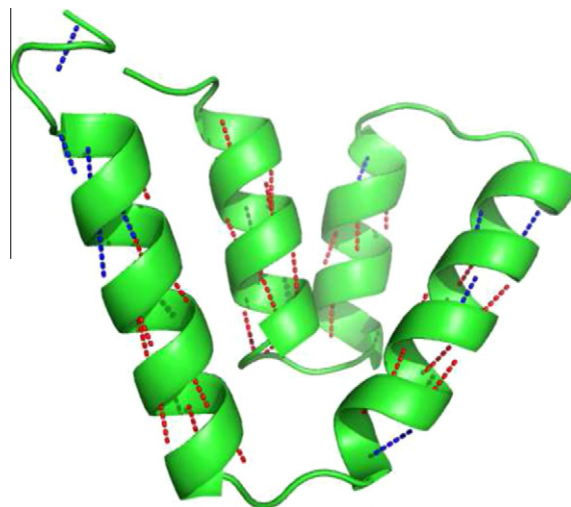
iltonian: amino-terminal domain of phage 434 repressor, uteroglobin, and vitamin D-dependent calcium-binding protein (PDB accession ID numbers 1R69, 1UTG and 3ICB, respectively). For comparison, molecular dynamics annealing simulations were performed on all three targets using the standard AMW Hamiltonian, a knowledge-based potential that has previously been shown to predict structures with high structural similarity to the native fold [13]. In this case, of course, no homologs were included in the memory set. To illustrate how the amnesiac Hamiltonian categorizes buried and exposed hydrogen bonds, Fig. 2 shows the native structure of 1UTG protein with hydrogen bonds colored according to the amnesiac potential. The exposed and buried hydrogen bonds depicted in Fig. 2 as determined by the amnesiac potential are consistent with our expectations: all exposed hydrogen bonds are limited to alpha-helical segments facing out from the protein interior or located near loops, while buried hydrogen bonds are mostly confined to alpha-helical segments located near the protein's interior.

For each target, two separate sets of 14 annealing simulations with the amnesiac model were performed in which one set uses a secondary structure bias and the other does not. The best Q score for each run is reported in Fig. 3. The Q score is defined as $2\Sigma_{i<j-2} \exp[-(r_{ij} - r_{ij}^N)^2/2\sigma_{ij}^2]/(N-1)(N-2)$, Q serves as a quantitative measure for all pairwise distances within a given structure [20]. For example, a structure with $Q = 1.0$ corresponds to the native structure, while a structure with a Q score <0.2 bears little resemblance to the native form. The Z score calculated using the combinatorial extension (CE) algorithm [21] was used as another similarity measure with which to compare predicted structures to their crystallographic structures. This score identifies general topological similarities irrespective of protein sequence. For example, Z scores larger than 4.0 indicate strong similarities between protein structures.

The results in Fig. 3 were obtained using values of $\gamma_p = 2.0$ and $\gamma_w = -1.0$ for buried and exposed hydrogen bonds, respectively. These values should strongly favor the formation of buried hydrogen bonds and discourage formation of exposed intra-helical hydrogen bonds. The best Q score obtained for simulations using a secondary structure bias are substantially improved over simulations carried out without a secondary structure bias. While this is not surprising, it is interesting to observe the degree of improvement in the prediction. For each target protein, the best overall Q scores for simulations using a secondary structure bias approaches
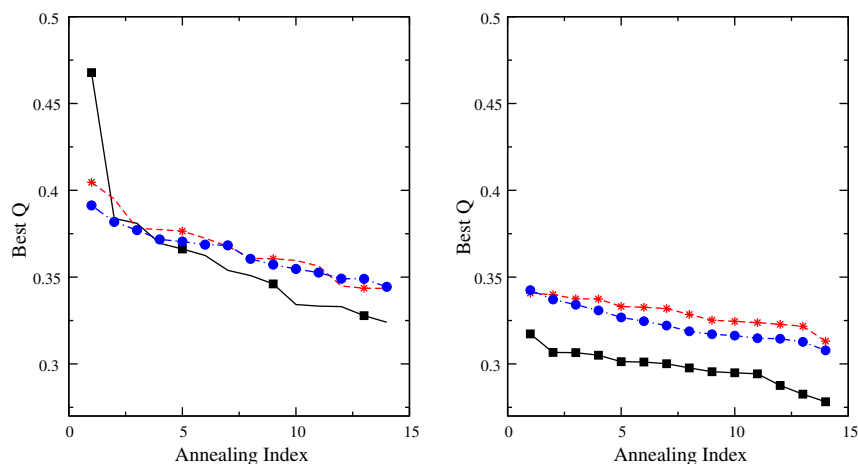
Fig. 3. Best Q sampled in 14 annealing simulations for proteins 1UTG (black squares), 1R69 (red stars) and 3ICB (blue circles) in descending order for simulations using a secondary structure bias (left) and simulations without using a secondary structure bias (right).
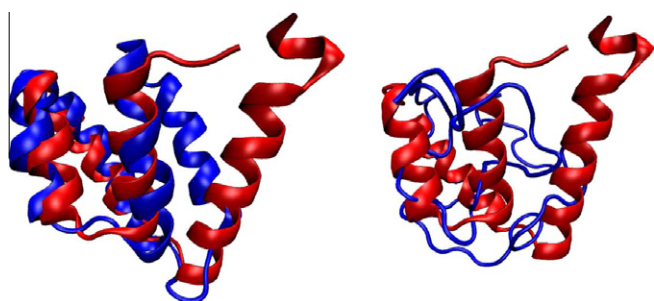
Fig. 4. Left: predicted structure from an annealing simulation using a secondary structure bias for protein 1UTG with Q = 0.468; CE: Z = 4.1 (blue) overlaid onto the X-ray crystal structure of protein 1UTG (red). Right: predicted structure from an annealing simulation without using a secondary structure bias for protein 1UTG with Q = 0.305; CE: Z = 3.1 (blue) overlaid onto the X-ray crystal structure of protein 1UTG (red).
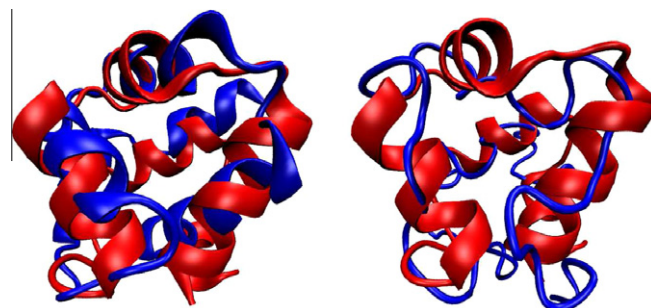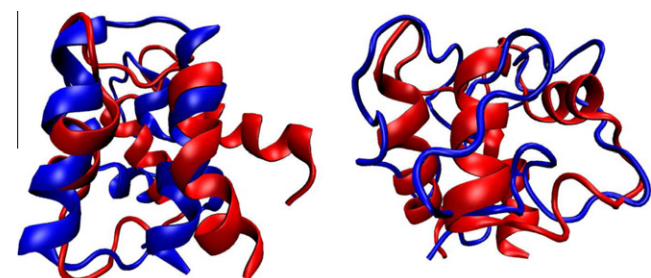

Fig. 5. Left: predicted structure from an annealing simulation using a secondary structure bias for protein 1R69 with Q = 0.344; CE: Z = 3.3 (blue) overlaid onto the X-ray crystal structure of protein 1R69 (red). Right: predicted structure from an annealing simulation without using a secondary structure bias for protein 1R69 with Q = 0.333; CE: Z = 2.6 (blue) overlaid onto the X-ray crystal structure of protein 1R69 (red).

Fig. 6. Left: predicted structure from an annealing simulation using a secondary structure bias for protein 3ICB with Q = 0.344; CE: Z = 3.5 (blue) overlaid onto the X-ray crystal structure of protein 3ICB (red). Right: predicted structure from an annealing simulation without using a secondary structure bias for protein 3ICB with Q = 0.343; CE: Z = 3.1 (blue) overlaid onto the X-ray crystal structure of protein 3ICB (red).

0.4, which is typically characterized by a better than 6 Å RMSD fit to the native structure. In contrast, the best overall Q scores found for simulations that do not use a secondary structure bias do not exceed 0.35. This trend is mirrored by the Z scores calculated for each of the structures giving the best Q score. The highest Z score for simulations biased by their predicted secondary structure was 4.1, while the highest value for the unbiased simulations was merely 3.1.

The structures obtained from annealing simulations with and without the secondary structure bias for protein 1UTG are displayed in Fig. 4 (corresponding contact maps can be found in the Appendix, Figs. A10–A12). The predicted structure obtained for protein 1UTG from simulations using the secondary structure bias represents the best predicted structure generated by all simulations employing the amnesiac potential. Comparison with the X-ray crystal structure shows that structural similarity is very high with some discrepancy between the packing of one of the helical segments. The best predicted structure obtained for protein 1UTG from simulations without the secondary structure bias appears quite different from the native structure due to the lack of secondary structural elements. However, the contact map (see Appendix) illustrates that a reasonable number of native pairwise contacts are present in this structure, which indicates that the amnesiac potential is able to recapitulate some of the structural features of the protein even without a secondary structure bias.

Results obtained for proteins 1R69 and 3ICB (Figs. 5 and 6) mirror those discussed for 1UTG above. Again, the best predicted

structures for each protein were obtained from simulations employing the secondary structure bias. A brief visual inspection of the left side of Figs. 5 and 6 shows that the overall similarity between the predicted structure and that of the X-ray structure bias was high, with some mis-packing of secondary structure elements (particularly for protein 3ICB). The best predicted structure for proteins 1R69 and 3ICB for simulations without a secondary structure bias, shown on the right side of Figs. 5 and 6, exhibit poorly developed secondary structure. The contact map for protein 1R69 (see
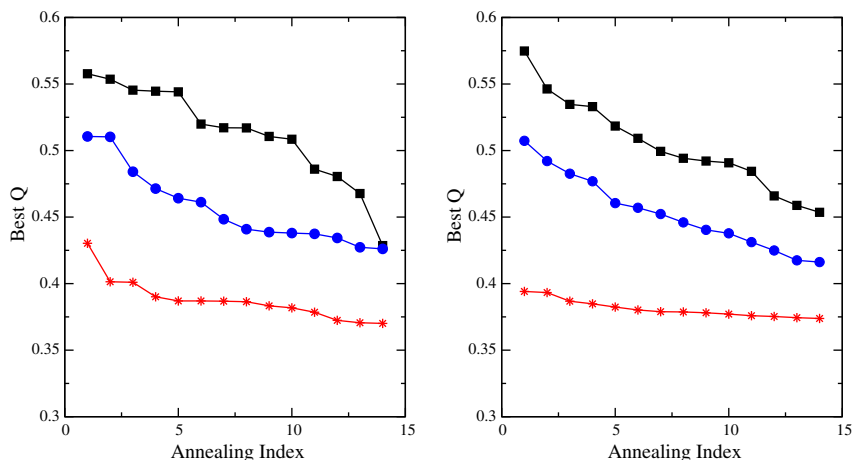
Fig. 7. Best Q sampled using AMW Hamiltonian in 14 annealing simulations for proteins 1UTG (black squares), 1R69 (red stars) and 3ICB (blue circles) in descending order for simulations using a secondary structure bias (left) and for simulations without using a secondary structure bias (right).
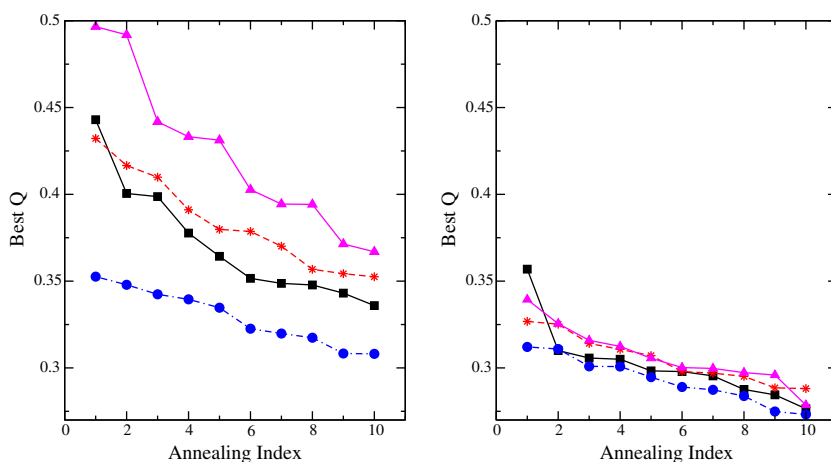
Fig. 8. Best Q sampled in 10 annealing simulations for protein 1UTG with varying values of the γ_w: 0.0 (black squares), 1.0 (red stars), −3.0 (blue circles) and 3.0 (magenta triangles) in descending order for simulations using a secondary structure bias (left) and simulations without using a secondary structure bias (right).
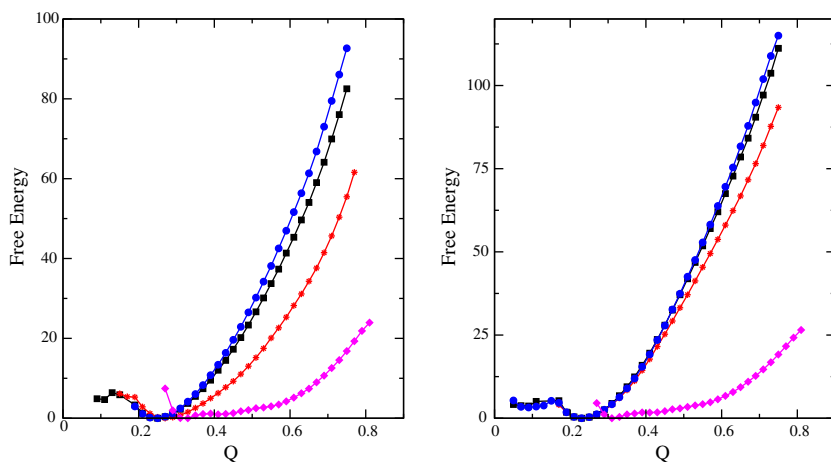
Fig. 9. Comparison of free energy calculations for protein 1UTG using the amnesiac Hamiltonian at $T = 0.7$ using varying values of γ_w: −1.0 (black squares), 3.0 (red stars), −3.0 (blue circles) and the AMW Hamiltonian (pink diamonds) for simulations using a secondary structure bias (left) and for simulations without using a secondary structure bias (right).

Appendix) indicates reasonably large regions where native contacts are maintained, while the number native contacts shown in the contact map for protein 3ICB is quite low.

Fig. 7 shows the best overall Q scores for annealing simulations using the AMW Hamiltonian with and without a secondary structure bias. In contrast to the results obtained using the amnesiac

Hamiltonian, the addition of a secondary structure bias does not lead to any substantial improvement in the prediction results. This can be attributed to the fact that associative memory potential term already biases the secondary structure according to its alignment with a set of non-homologous memory proteins. The best Q scores measured for the AMW annealing simulations are considerably better than those achieved with amnesiac Hamiltonian.

Fig. 8 displays the best Q sampled in 10 annealing simulations for protein 1UTG using varying values of the Pauling–Kauzmann coefficient for exposed hydrogen bonds (γ_w) to investigate the effect of favoring or disfavoring exposed hydrogen bonds on structure prediction. The results for simulations employing a secondary structure bias clearly exhibit a dependence on the value of this parameter, while simulations without a secondary structure bias exhibit a weaker dependence. One possible explanation for this trend could be that the secondary structure bias influences the residue density, which in turn influences the E_{helical}, which depends on local residue density. In both cases, larger positive values of γ_w appear to result in predicted structures with higher Q scores and suggests that annealing simulations using larger positive values of γ_w would have yielded structures with higher structural similarity that those reported in Fig. 3.

To further understand the role of exposed hydrogen bonds in protein folding, free energy profiles for 1UTG were calculated as a function of Q score using umbrella sampling along the reaction coordinate. Umbrella sampling was carried out with the Q score bias potential $V_i(Q) = 5000 \times (Q - Q_i)^4$, where Q_i = 0.15, 0.20, 0.25, 0.30,...,0.95. Free energy plots as a function of Q score were calculated using the weighted histogram analysis method [22]. Fig. 9 shows free energy profiles calculated from simulations using γ_w values of 3.0, 0.0 and −3.0 and both with and without the secondary structure bias. The trends observed in the free energy calculations are consistent with the trends observed in the annealing simulation results. For those simulations using a secondary structure bias, the free energy curve minima are shifted to larger values of Q compared to that of simulations without the secondary structure bias. Interestingly, the free energy minima for simulations using γ_w values of 0.0 and −3.0 both occur at $Q \sim 0.30$, while for simulations using a γ_w value of 3.0 the free energy minimum moves to a noticeably higher Q score ($Q > 0.4$). In contrast, the free energy vs. Q plots for corresponding simulations without the secondary structure bias all exhibit minima at $Q \sim 0.23$. This low value is also consistent with results observed in the annealing simulations. The results in Fig. 9 suggest that the use of the secondary structure bias strongly influences the E_{helical} term of the amnesiac potential and, in turn, its ability to sample structures more similar to the native state. These results also suggest that the role of intrahelical exposed hydrogen bonds plays an important role in stabilizing the native protein conformation.

4. Conclusions

We have investigated the ability of a coarse-grained microphysics-based model to predict the structures of globular proteins. The microphysics incorporated in the simplest "amnesiac" model emphasizes the importance of backbone buried hydrogen bond and sequence dependent compaction with the inclusion of water-mediated interactions. While this model captures many of the common themes of protein structures in its predictions, it performs considerably more poorly than schemes that also incorporate local in sequence interactions, like the corresponding associative memory Hamiltonian. Surprisingly, adding a rather simple local secondary structure potential term already improves performance considerably, although not to the level of the full AMW method or the most powerful hybrid of AMH and fragment assembly. The amplification of local signals, by a roughly organized collapse, predicted by Saven and Wolynes' analytical theory [6] seems to be at the core of this performance.

The value of the current amnesiac code is that it is applicable to the prediction of alpha-helical proteins where any local structural signals have a distinct origin from those typically seen in globular proteins. Examples include "intrinsically-disordered" proteins, which have become increasingly recognized for their biological significance. It will also be of interest to use the amnesiac code to examine and predict membrane protein structures, many of which exhibit only alpha-helical structure, where again the main organizing microscopic forces should be captured by the amnesiac model. Extension of this model to include proteins with alpha-beta and all-beta secondary structure will require a more complicated potential for hydrogen bonding, as discussed in [14], and is a problem for the future. Refinement of the hydrogen bonding potential in the amnesiac model to allow more extensive hydrogen bonding will allow us to investigate a broader range of "intrinsically-disordered" proteins, as well as explore the possibility of whether these proteins adopt fluctuating super-secondary structures not yet observed.

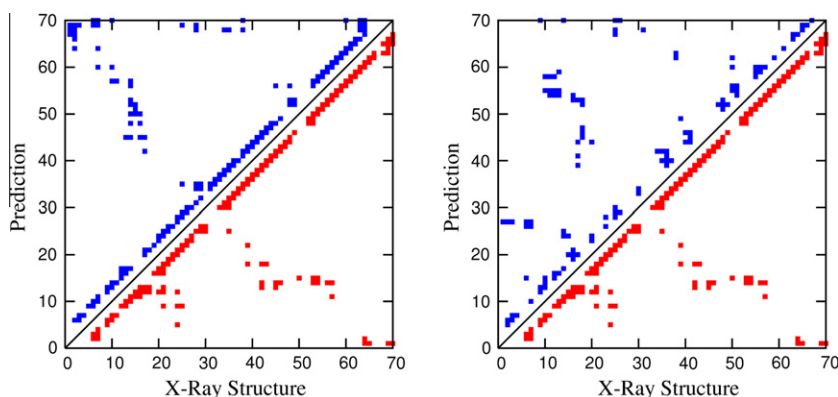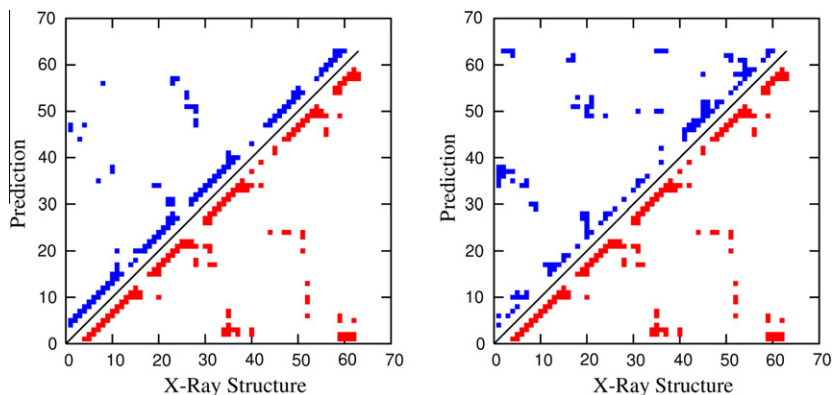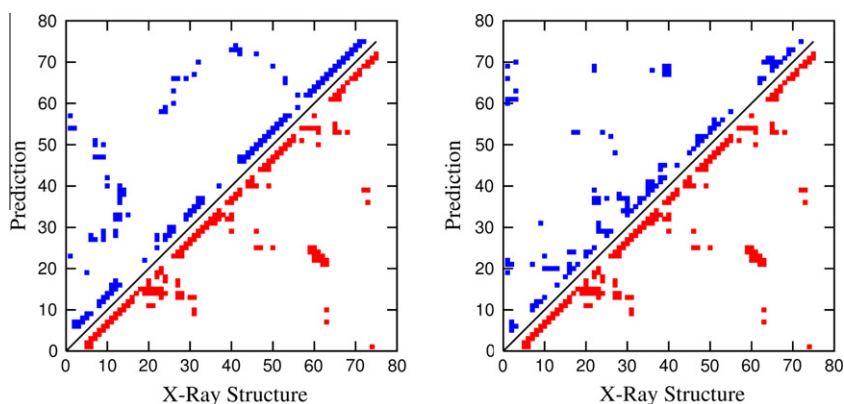## Acknowledgment

Appendix A

Fig. A10. Left: contact map for a predicted structure from an annealing simulation using a secondary structure bias for protein 1UTG. Right: contact map for a predicted structure from an annealing simulation without using a secondary structure bias for protein 1UTG.

Fig. A11. Left: contact map for a predicted structure from an annealing simulation using a secondary structure bias for protein 1R69. Right: contact map for a predicted structure from an annealing simulation without using a secondary structure bias for protein 1R69.

Fig. A12. Left: contact map for a predicted structure from an annealing simulation using a secondary structure bias for protein 3ICB. Right: contact map for a predicted structure from an annealing simulation without using a secondary structure bias for protein 3ICB.

References

[1] L. Pauling, R.B. Corey, H.R. Branson, Proc. Natl. Acad. Sci. USA 37 (4) (1951) 205–211.
[2] W. Kauzmann, Adv. Protein Chem. 14 (1959) 1–63.
[3] J.N. Onuchic, Z. Luthey-Schulten, P.G. Wolynes, Ann. Rev. Phys. Chem. 48 (1997) 545–600.
[4] Z. Lutheyschulten, B.E. Ramirez, P.G. Wolynes, J. Phys. Chem. 99 (7) (1995) 2177–2185.
[5] J.N. Onuchic, P.G. Wolynes, Z. Lutheyschulten, N.D. Socci, Proc. Natl. Acad. Sci. USA 92 (8) (1995) 3626–3630.
[6] J.G. Saven, P.G. Wolynes, J. Mol. Biol. 257 (1996) 199–216.
[7] C.A. Rohl, C.E.M. Strauss, K.M.S. Misura, D. Baker, Numer. Comput. Methods D 383 (2004) 66–93.
[8] J.A. Hegler, J. Latzer, A. Shehu, C. Clementi, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 106 (36) (2009) 15302–15307.
[9] M.S. Friedrichs, P.G. Wolynes, Science 246 (1989) 371–373.
[10] C. Hardin, M.P. Eastwood, Z. Luthey-Schulten, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 97 (26) (2000) 14235–14240.
[11] K.K. Koretke, Z. Luthey-Schulten, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 95 (1998) 2932–2937.

[12] M.C. Prentiss, C. Hardin, M.P. Eastwood, C.H. Zong, P.G. Wolynes, J. Chem. Theory Comput. 2 (3) (2006) 705–716.
[13] G.A. Papoian, J. Ulander, M.P. Eastwood, Z. Luthey-Schulten, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 101 (10) (2004) 3352–3357.
[14] C.H. Zong, G.A. Papoian, J. Ulander, P.G. Wolynes, J. Am. Chem. Soc. 128 (15) (2006) 5168–5176.
[15] S. Takada, Z. Luthey-Schulten, P.G. Wolynes, J. Chem. Phys. 110 (23) (1999) 11616–11629.
[16] Y. Fujitsuka, S. Takada, Z.A. Luthey-Schulten, P.G. Wolynes, Proteins-struct. Funct. Genet. 54 (1) (2004) 88–103.
[17] C.N. Pace, J.M. Scholtz, Biophys. J. 75 (1) (1998) 422–427.
[18] M.P. Eastwood, C. Hardin, Z.L. Schulten, P.G. Wolynes, J. Chem. Phys. 117 (2002) 4602–4615.
[19] J.A. Cuff, M.E. Clamp, A.S. Siddiqui, M. Finlay, G.J. Barton, Bioinformatics 14 (10) (1998) 892–893.
[20] S.S. Cho, Y. Levy, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 103 (3) (2006) 586–591.
[21] I.N. Shindyalov, P.E. Bourne, Protein Eng. 11 (9) (1998) 739–747.
[22] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, J.M. Rosenberg, J. Comput. Chem. 13 (8) (1992) 1011–1021.