Eur. Phys. J. E **24**, 311–316 (2007)
DOI 10.1140/epje/i2007-10241-1

THE EUROPEAN PHYSICAL JOURNAL E

Optimized folding simulations of protein A

S. Trebst[1] and U.H.E. Hansmann[2,3,a]

[1] Microsoft Research, Station Q, University of California, Santa Barbara, CA 93106, USA
[2] Department of Physics, Michigan Technological University, Houghton, MI 49931, USA
[3] John-von-Neumann Institute for Computing, Forschungszentrum Jülich, D-52425 Jülich, Germany

Received 9 August 2007
Published online: 11 December 2007 –

Abstract. We describe optimized parallel tempering simulations of the 46-residue B-fragment of protein A. Native-like configurations with a root-mean-square deviation of ≈ 3 Å to the experimentally determined structure (Protein Data Bank identifier 1BDD) are found. However, at biologically relevant temperatures such conformations appear with only $\approx 10\%$ frequency in our simulations. Possible shortcomings in our energy function are discussed.

PACS. 87.14.Ee Proteins – 87.15.Aa Theory and modeling; computer simulation – 87.15.He Dynamics and conformational changes – 87.15.Cc Folding and sequence analysis

1 Introduction

Rational drug design or the pathology of amyloid diseases are only two problems whose solutions require a detailed understanding of the relation between chemical composition and structure (and function) of proteins. Exploring this relationship through numerical simulations is a computationally hard problem. Two major factors limit our ability to efficiently simulate large proteins and study their folding transitions. First, statistically sampling the rough energy landscape of a protein can be extremely slow even at room temperature. Second, the present energy functions are often insufficiently accurate in describing the interactions between the atoms within a protein, and between protein and their surrounding solvent. It is often not clear whether the failure of a computer experiment to find the known structure of a protein results from poor sampling or lack of accuracy in the energy function.

To overcome some of the limitations of statistical sampling in the simulation of small proteins, sophisticated simulation schemes such as parallel tempering [1,2] or generalized ensemble methods [3,4] are now widely employed numerical methods [5–7] In a recent line of research feedback-optimized algorithms have been developed that aim at further improving the statistical sampling of these methods by systematically improving the simulated statistical ensemble [8], *e.g.* the exact placement of replicas in temperature space [9–11]. Recent simulations of the 36-residue villin headpiece sub-domain HP-36 in reference [10] demonstrated that optimizing the sampled temperature distribution leads to qualitatively different results for the same force field, in this case a combination of the ECEPP/3 force field [12] with an implicit solvent [13]. Previous simulations of HP-36 in reference [14] had indicated that the native structure is *not* the global free-energy minimum at room temperature for this force field. However, with the optimized temperature distribution it was later found that the correct structure is sampled with $\approx 90\%$ frequency at room temperature [10]. Clearly, the earlier numerical simulations suffered from a sampling problem and resulted in misleading conclusions on the force field, while the energy function in fact accurately described the interactions for HP-36 as shown in the latter study with improved sampling. While this result is promising, the observed sampling difficulties nevertheless suggest that the energy landscape for HP-36 modeled by this force field may be more rough than the one experienced by the "real" protein. Given the complexity of the energy landscape for this small protein, one would expect that with increasing size and complexity of the molecule the accuracy of the energy function will further decrease as the energy landscape gains further complexity. This would render it even more difficult (or simply impossible) to pick the correct structure in the numerical simulation of larger proteins. As stable domains in proteins usually consists of 50–200 residues it is therefore important to test the accuracy of current energy functions for proteins of this size. As a first step in this direction we have applied the feedback-optimized parallel tempering scheme to simulate the 46-residue fragment 10–55 of the B domain of protein A (Protein Data Bank identifier 1BDD) which forms the three-helix bundle displayed in Figure 1 as determined in experiments [15]. This is one of the

[a] e-mail: hansmann@mtu.edu

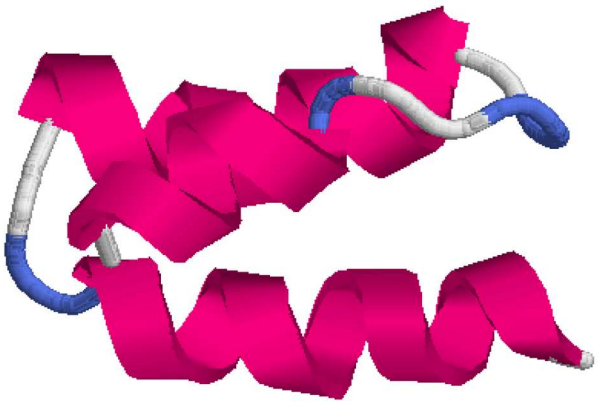

Fig. 1. The experimentally determined structure of the 46-residue B-fragment of protein A as stored in the Protein Data Bank (identifier 1BDD).

few small proteins with experimentally well-characterized states. For this reason, it has raised some interest as a model to test folding algorithms or energy functions [16–22]. While our high-statistics simulations do find native-like configurations with a root-mean-square deviation of $\approx 3\,\text{Å}$ to the experimentally determined structure, these configurations do not correspond to the lowest-energy configuration and appear with only around 10% probability at biologically relevant temperatures. Our results indicate that this deviation from experiment is due to a bias in the ECEPP/3 energy function toward helical structures. Another contributing factor is the use of an implicit solvent model in our simulations. These models reduce dramatically the numerical costs of protein simulations but can lead to distorted free energy landscapes. This effect has been observed earlier [23,24] even for more sophisticated implicit solvents than used by us.

2 Methods

Our simulations of the protein-A fragment utilize the ECEPP/3 force field [12] as implemented in the 2005 version of the program package SMMP [25,26]. Here, the interactions between the atoms within the protein are approximated by a sum $E_{\text{ECEPP/3}}$ consisting of electrostatic energy E_{C}, a Lennard-Jones term E_{LJ}, a hydrogen-bonding term E_{hb} and a torsion energy E_{tor}:

$$\begin{aligned} E_{\text{ECEPP/3}} &= E_{\text{C}} + E_{\text{LJ}} + E_{\text{hb}} + E_{\text{tor}} \\ &= \sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}} \\ &\quad + \sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \\ &\quad + \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ &\quad + \sum_{l} U_l (1 \pm \cos(n_l \xi_l)) \,, \end{aligned} \tag{1}$$

where r_{ij} is the distance between the atoms i and j, ξ_l is the l-th torsion angle, and energies are measured in kcal/mol. The protein-solvent interactions are approximated by a solvent-accessible surface term

$$E_{\text{solv}} = \sum_i \sigma_i A_i \,. \tag{2}$$

The sum goes over the solvent-accessible areas A_i of all atoms i weighted by solvation parameters σ_i as determined in reference [13], a common choice when the ECEPP/3 force field is utilized.

The above-defined energy function leads to an energy landscape that is characterized by a multitude of minima separated by high barriers. In order to improve statistical sampling and speed up equilibration at low temperatures we utilize a parallel tempering scheme [1,2] which was first used in protein science in reference [27]. In this scheme N non-interacting copies, or "replicas", of the protein are simultaneously simulated at a range of temperatures $\{T_1, T_2, \ldots, T_N\}$. After a fixed number of Monte Carlo sweeps (or a molecular dynamics run for a fixed time interval) a sequence of swap moves, the exchange of two replicas at neighboring temperatures, T_i and T_{i+1}, is suggested and accepted with a probability

$$p(E_i, T_i \to E_{i+1}, T_{i+1}) = \min\left(1, \exp(\Delta\beta \Delta E)\right) \,, \tag{3}$$

where $\Delta\beta = 1/T_{i+1} - 1/T_i$ is the difference between the inverse temperatures and $\Delta E = E_{i+1} - E_i$ is the difference in energy of the two replicas. The exchange of conformations considerably improves equilibration for all replicas, especially those at low temperatures which can have extremely long equilibration times for conventional canonical simulations (at a fixed temperature).

The improved equilibration of the replica exchange scheme is due to the random walks that individual replicas can perform in temperature space allowing them to move to higher temperatures where equilibration is fast and then move back down to lower temperatures thereby escaping barriers in the energy landscape. Obviously, the number of round trips n_{rt} between the lowest and highest temperature, T_1 and T_N, respectively, is a lower bound for the statistically independent visits at the lowest temperature, and therefore a good measure for the equilibration of the parallel tempering simulation. It is therefore desirable to maximize the number of round trips by optimizing the position of temperature points in the interval $[T_1, T_N]$. This can indeed be achieved in a systematic way by feeding back the local diffusivity of the described random walk using the feedback algorithm described in references [9,10]. Technically, replicas are "labeled" according to which of the two extremal temperatures, T_1 or T_N, they have visited last. Using this label we can define the number of replicas $n_{\text{up}}(i)$ ($n_{\text{down}}(i)$) which at temperature T_i have come from T_1 (T_N). Then the fraction of replicas moving in one direction,

$$f_{\text{up}}(i) = \frac{n_{\text{up}}(i)}{n_{\text{up}}(i) + n_{\text{down}}(i)} \,, \tag{4}$$

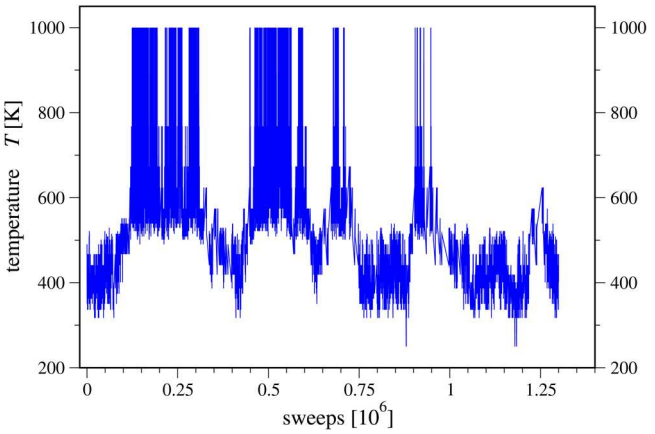

Fig. 2. Time series of visited temperatures for one of the replicas in the parallel-temperature simulation. Time is measured in MC sweeps and starts with the beginning of measurements.

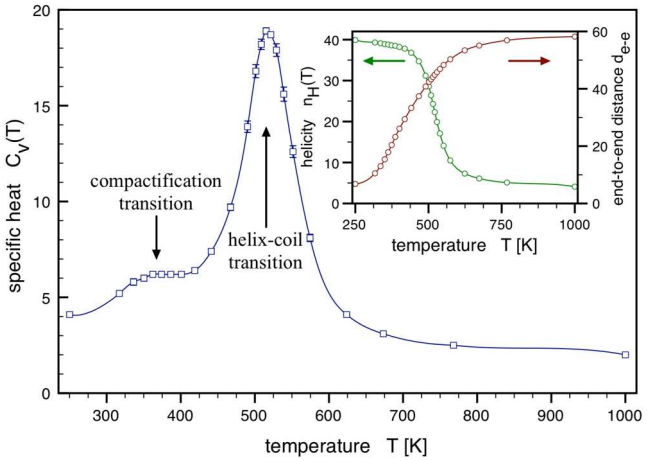

Fig. 3. The specific heat $C(T)$ as a function of temperature T. A peaked signal is found at the helix-coil transition around $T_c = 515\,\mathrm{K}$ where the helicity $n_H(T)$ increases, see the inset. Below the transition a shoulder emerges over a broad temperature regime in which the average end-to-end distance $d_{e\text{-}e}$ of the protein decreases rapidly as shown in the inset, an indication that the helices form a secondary structure.

describes a stationary distribution of probability flow between temperatures T_1 and T_N with boundary conditions $f_{up}(1) = 1$ and $f_{up}(N) = 0$. The local diffusivity $D(T)$ of the random walk which a single replica performs in temperature is then given by $D(T) \propto \Delta T \cdot (\mathrm{d}f/\mathrm{d}T)^{-1}$, where ΔT is the size of the temperature interval around T. Feeding back this local diffusivity it was shown in references [9,10] that for the optimal temperature distribution, *i.e.* the one that maximizes the number of round trips n_{rt}, the fraction decreases linearly

$$f_{up}^{(opt)}(i) = \frac{N-i}{N-1}\,. \tag{5}$$

For our simulations of the protein-A fragment such an optimized distribution can be found by applying the iterative procedure described in references [9,10] and is given for the 24 replicas in our simulations by 1000, 768, 673, 624, 574, 552, 539, 529, 521, 515, 509, 501, 490, 467, 441, 419, 401, 386, 374, 362, 350, 336, 317, 250. Our data are taken from a simulation with 1000000 sweeps for each replica. Here, a sweep consists of a sequential series of attempts to update each of the 276 dihedral angles (the true degrees of freedom in our model) once. After each sweep, we attempt an exchange move (swap) of configurations between neighboring temperatures which is accepted with probability (3). The walk of one replica along the ladder of temperatures is displayed in Figure 2. Measurements are taken every ten sweeps and stored for further analysis. These include the energy E, the radius of gyration r_{gy} as a measure of the geometrical size, and the number of helical residues n_H, *i.e.* residues where the pair of dihedral angles (ϕ, ψ) takes values in the range $(-70° \pm 30°, -37° \pm 30°)$. Finally we recorded the configurations with overall lowest energy obtained in our simulation.

3 Results and discussions

The biologically active state of a protein is thought to be the global minimum of the free energy at room temperature. Heating leads to unfolding that is reversible after cooling. Hence, the folding transition should be marked by a signal in the specific heat

$$C(T) = \beta^2(\langle E^2\rangle - \langle E\rangle^2)/N, \tag{6}$$

with $\beta = 1/k_B T$ (k_B is the Boltzmann constant) and N the number of residues. For protein A we indeed find a pronounced peak in the specific heat, displayed in Figure 3, at a temperature $T_c = 515\,\mathrm{K}$. This peak is related to the formation of α-helices as one can see from the inset where the average number n_H of residues which are part of an α-helix is plotted *versus* temperature. Around the transition temperature T_c the helicity rapidly increases. The corresponding formation of hydrogen bond between residues $(i, i+4)$ leads to a much lower energy of such configurations, and the resulting fluctuation in the average energy as a function of temperature is measured by the specific heat. Below this helix-coil transition a shoulder is observed in the specific heat in a temperature range between 330–430 K. Since the helicity varies only little in this temperature range, but the typical end-to-end distance $d_{e\text{-}e}$ of the protein-A configurations decreases rapidly —as shown in the inset of Figure 3— this temperature regime is marked by the formation of a secondary structure of the helical segments which we will discuss in detail below. As the temperature is further decreased below $T \approx 330\,\mathrm{K}$, the specific heat lowers again and the end-to-end distance approaches a constant value.

A more detailed picture of the formation of a compact secondary structure emerges when looking at the frequency of configurations with typical end-to-end distances $d_{e\text{-}e}$ and their respective radius of gyration r_{gy}, a measure for the compactness of a folded protein structure, which are both plotted *versus* temperature in Figure 4. At temperatures directly below the helix coil transition, *e.g.* our data point at $T = 440\,\mathrm{K}$, the histograms indicate

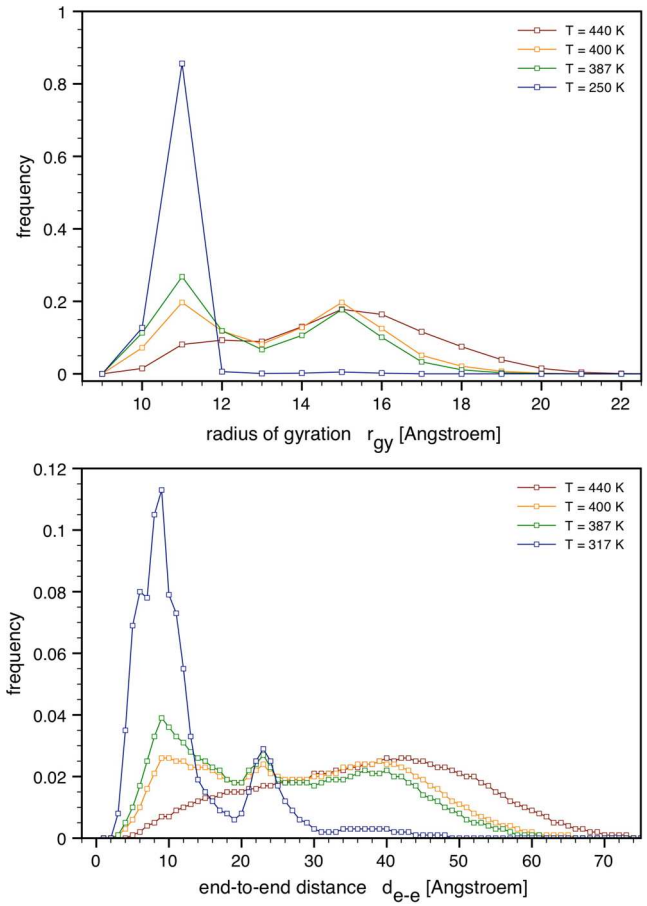

Fig. 4. Frequency of (a) the radius of gyration $r_{\rm gy}$ and (b) the end-to-end distance $d_{\rm e\text{-}e}$ for various temperatures. Each data point samples entries of a bin with 1 Å width. The histograms are normalized.

that configurations still differ widely as seen from the broad distributions found in both histograms, but with a clear preference for extended structures with a radius of gyration centered around $r_{\rm gy} = 15\,$Å. In the shoulder region of the specific heat, *e.g.* at our data points $T = 400\,$K and $T = 387\,$K, a double peak appears in histogram of the radius of gyration indicating the competion between extended structures ($r_{\rm gy} \approx 15\,$Å) and compact ones ($r_{\rm gy} \approx 11\,$Å). Further decreasing the temperature we are finally left with compact configurations only (with $r_{\rm gy} \leq 11\,$Å). The histograms in the radius of gyration seem to indicate a compactification transition where distinct secondary structures are formed at around 387–400 K that separates compact helical structures from extended helical structures immediately below the helix-coil transition around $T_{\rm c} = 515\,$K. Indeed, we find that two different compact structures are primarily formed below this compactification transition as becomes evident from our measurements of frequencies of typical end-to-end distance for various temperatures. Above the compactification transition and below the helix-coil transition the histogram of typical end-to-end distances is first centered around $d_{\rm e\text{-}e} = 43\,$Å for $T = 440\,$K and an almost flat histogram is observed

Fig. 5. Typical extended low-energy structure with high helical content (Type I).

at $T = 400\,$K for $d_{\rm e\text{-}e} = 9$–$40\,$Å. Below $T = 387\,$K two additional peaks form around $d_{\rm e\text{-}e} = 9\,$Å and $d_{\rm e\text{-}e} = 23\,$Å, while there is still a broad feature around $d_{\rm e\text{-}e} = 40\,$Å. Further lowering the temperature, the two peaks $d_{\rm e\text{-}e} = 9\,$Å and $d_{\rm e\text{-}e} = 23\,$Å further proliferate and become the dominant feature in the histogram. Finally, towards the lowest temperature $T = 250\,$K which we sampled, only configurations with small end-to-end distance prevail with more than 90% of all sampled configurations having a typical end-to-end distance of $d_{\rm e\text{-}e} \leq 10\,$Å.

The above analysis of the histograms of typical radius of gyration and end-to-end distance indicate that below the helix-coil transition temperature of $T_{\rm c} = 515\,$K there exist *three* different types of configurations which all have high helix-content ($n_{\rm H} \approx 32$–40) but differ strongly in their respective arrangement of the helical segments. Two of these structures are compact, while one is extended and found only above the compactifcation transition around $T \approx 387\,$K. A typical example for the extended helical structure —named "structure I"— is shown in Figure 5. The illustrated configuration has a radius of gyration of $r_{\rm gy} = 13.5.\,$Å, an end-to-end distance of $d_{\rm e\text{-}e} = 34.0\,$Å, and a solvent-accessible surface area of $A_{\rm SAS} = 4440\,$Å^2. All three helical segments are formed, but the total helicity is with 39 helical residues higher than the one found for the PDB structure where only 34 residues are part of an α-helix. Accordingly, its root-mean square deviation (rmsd) to the native structure (as deposited in the Protein Data Bank under the identifier 1BDD) is 11.5 Å (calculated over backbone atoms only). The total energy after minimization is $E_{\rm tot} = -643\,$kcal/mol, of which $E_{\rm solv} = -186\,$kcal/mol result from the solvation energy and $E_{\rm ECEPP/3} = -457\,$kcal/mol from the intramolecular interactions. Figure 6 displays the first of the two compact structures found below the compactification transition. Named "structure II" it is the configuration with lowest overall energy found in our extensive simulations. After minimization we have $E_{\rm tot} = -692\,$kcal/mol, with the energy difference to the previous structure resulting from the more favorable intramolecular interactions ($E_{\rm ECEPP/3} = -510\,$kcal/mol) while there is little difference in the solvation term, $E_{\rm solv} = -182\,$kcal/mol. The small end-to-end distance of $d_{\rm e\text{-}e} = 5.7\,$Å reflects a triangle

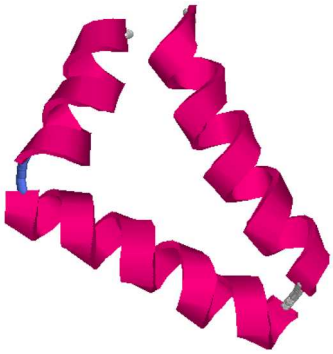

Fig. 6. One of the typical triangle-shaped compact structures characterized by small end-to-end distance $d_{\text{e-e}}$ (Type II).

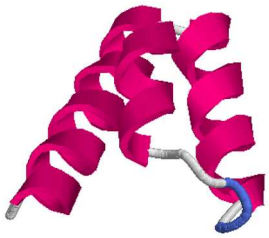

Fig. 7. Another typical compact structure, but with larger end-to-end distance $d_{\text{e-e}}$ (Type III). This structure resembles the native structure as deposited in the PDB and shown in Figure 1.

shape of this configuration which differs from the native one by a rmsd of 7.5 Å and is with 41 residues maximal helical. The configuration is with $r_{\text{gy}} = 11.3$ Å and a solvent accessible surface area $A_{\text{SAS}} = 3950$ Å^2 more compact than the extended structure (type I) but not as compact as the native one which has $r_{\text{gy}} = 9.7$ Å, and a solvent-accessible surface area $A_{\text{SAS}} = 3333$ Å^2. A closer resemblance to the native structure has the second type of compact configurations found in our simulations, named "structure III" and illustrated in Figure 7. This structure differs from the other compact, triangle-shaped structure ("structure II") in that it is more compact but has a smaller helicity with only 37 residues which are part of an α-helix. For this configuration we have measured a radius of gyration of $r_{\text{gy}} = 10.4$ Å, a solvent-accessible surface area $A_{\text{SAS}} = 3780$ Å^2, and an end-to-end distance $d_{\text{e-e}} = 23.0$ Å. Strikingly, the rmsd to the native structure is only 3.3 Å. However, the total energy $E_{\text{tot}} = -662$ kcal/mol is by about 30 kcal/mol higher than the one of the other compact structure shown in Figure 6. This is similar to reference [22] where the best sampled structure has a rmsd of 2.33 Å to the native one, but the lowest-energy configuration differs by 6.41 Å. In our case, the difference for the two compact structures is primarily due to the ECEPP/3 energy which for the second compact structure is found to be $E_{\text{ECEPP/3}} = -484$ kcal/mol while both compact structure have similar solvation energy, $E_{\text{solv}} = -178$ kcal/mol for the second compact structure. This may indicate shortcomings of our implicit solvent. This conjecture is supported by reference [19] where native-like configurations

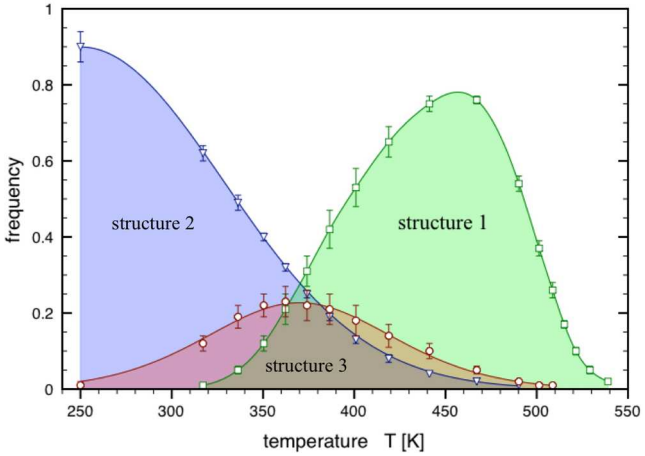

Fig. 8. Frequency of the three dominating helical low-energy structures as a function of temperature.

are observed as free-energy global minimum in simulations with an explicit solvent.

Note that the triangle-shaped compact configuration ("structure II") in Figure 6 has a total energy that is ≈ 30 kcal/mol lower than the lowest-energy configuration found in previous simulations reference [28] with the same force field. These simulations reported a lowest-energy configuration where the middle helix is broken at residue *Gly21*, *i.e.* the configuration is built out of four helical segments. Similar configurations are also observed in our simulations and are found to have energies comparable to the ones found previously in reference [28]. Depending on the arrangement of the helices they either resemble the triangle shaped structure of type "II" or the other compact structure of type "III", and are grouped together with those structures in our analysis of the distribution of varying structures as a function of temperature.

Finally, we note that all three structures are built out of helical segments and therefore only observed with substantial frequency at and below the helix-coil transition at T_{c}. The extended structure (type I) is most frequent at temperatures between 441–467 K. Its probability decreases with temperature and at $T = 317$ K only about 1% of the configurations belong to type I. Compact configurations (type II and III) apear with a frequency of more than 1% only for temperatures below $T = 490$ K. The native-like configurations of type III are slightly more frequent than the triangle-shaped of type II for a small temperature window above $T = 362$ K where their frequency reaches with 23% a maximum, see Figure 8. Further decreasing the temperature the frequency of configurations of type III diminishes while configurations of type II become more frequent. At $T = 317$ K only 12% of the configurations are native-like, but already 62% belong to type II. While the entropy of the triangle shaped structure (type II) is certainly lower than the entropy of the native structure (which might explain the suppression of type-II structures at higher temperatures), the entropic gain of the native-like structure is compensated by an energetic gain of the maximally helical triangle structure at lower

temperatures making the triangle shaped structure the predominant one. This observation clearly reflects a bias in the ECEPP/3 energy function toward helical structures. Growth of helices is energetically favored which in turn strongly restricts the possible topologies the folded structures can assume. This leads to the observed dominance of structures that while less compact and larger solvent-accessible surface are more helical than the one observed in experiment. Note that such a thermodynamic bias has not been observed in reference [21] where also configurations similar to the ones in Figures 5–7 were observed during the simulation. This indicates that this coarse-grained model employed in reference [21] (and the one of Ref. [18]) captures the effective interactions in protein A better than our all-atom energy function. Hence, we should contemplate possible corrections to our energy functions which should decrease the helix-forming tendencies. We are currently exploring this conjecture with a variant of the ECEPP force field proposed by Abagyan and co-workers [29].

4 Conclusions

We have performed parallel tempering simulations of the 46-residue B-fragment of protein A with an optimized temperature distribution and high statistics. Our goal was to test the limitations set on protein simulations by our energy function, the ECEPP/3 force field with an implicit solvent. While we find native-like structures with ≈ 3 Å rmsd, these structures appear at room temperature with only $\approx$ 10% probability. Energetically favored are compact structures that are maximal helical, and more exposed to the surrounding solvent while less compact than the native structure. This observation suggests the need for correction terms to the ECEPP/3 force field which decrease the helix-forming bias when using the force field in combination with an implicit solvent.

This work is supported, in part, by a research grant (CHE-0313618) of the National Science Foundation (USA).

References

1. C.J. Geyer, A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
2. K. Hukushima, K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
3. U.H.E. Hansmann, Y. Okamoto, Phys. Rev. E **56**, 2228 (1997).
4. P.M.C. Oliveira, T.J.P. Penna, H.J. Herrmann, Eur. Phys. J. B **1**, 205 (1998).
5. S. Kmiecik, A. Kolinski, Proc. Natl. Acad. Sci. U.S.A. **104**, 12330 (2007).
6. K. Hamacher, Europhys. Lett. **74**, 944 (2006).
7. H. Li, M. Fajer, W. Yang, J. Chem. Phys. **126**, 24106 (2007).
8. S. Trebst, D.A. Huse, M. Troyer, Phys. Rev. E **70**, 046701 (2004).
9. H.G. Katzgraber, S. Trebst, D.A. Huse, M. Troyer, J. Stat. Mech. P03018 (2006).
10. S. Trebst, M. Troyer, U.H.E. Hansmann, J. Chem. Phys. **124**, 174903 (2006).
11. D. Gront, A. Kolinski, J. Phys.: Condens. Matter **19**, 036225 (2007).
12. M.J. Sippl, G. Némethy, H.A. Scheraga, J. Phys. Chem. **88**, 6231 (1984) and references therein.
13. T. Ooi, M. Obatake, G. Nemethy, H.A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 3086 (1987).
14. C.-Y. Lin, C-K. Hu, U.H.E. Hansmann, Proteins **52**, 436 (2003).
15. H. Gouda, H. Torigoe, A. Saito, M. Sato, Y. Arata, I. Shimada, Biochemistry **31**, 9665 (1992).
16. E.M. Boczko, C.L. Brooks III, Science **269**, 393 (1995).
17. A. Ghosh, R. Elber, H. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **99**, 10394 (2002).
18. G. Favrin, A. Irbäck, S. Wallin, Proteins: Struct. Funct. Genet. **47**, 99 (2002).
19. A. Garcia, J.N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **100**, 13898 (2003).
20. T. Herges, W. Wenzel, Biophys. J. **87**, 3100 (2004).
21. M. Khalili, A. Liwo, H. Scheraga, J. Mol. Biol. **355**, 536 (2006).
22. J.S. Yang, W.W. Chen, J. Skolnick, E.I. Shakhnovich, Structure **15**, 53 (2007).
23. R. Zhou, B.J. Berne, Proc. Natl. Acad. Sci. U.S.A. **99**, 12777 (2002).
24. N. Nymeyer, A.E. Garcia, Proc. Natl. Acad. Sci. U.S.A. **100**, 13934 (2003).
25. F. Eisenmenger, U.H.E. Hansmann, Sh. Hayryan, C.-K. Hu, Comput. Phys. Commun. **138**, 192 (2001).
26. F. Eisenmenger, U.H.E. Hansmann, Sh. Hayryan, C.-K. Hu, Comput. Phys. Commun. **174**, 422 (2006).
27. U.H.E. Hansmann, Chem. Phys. Lett. **281**, 140 (1996).
28. W. Kwak, U.H.E. Hansmann, Phys. Rev. Lett. **95**, 138102 (2005).
29. R.A. Abagyan, M. Trotov, J. Comput. Phys. **151**, 405 (1999).