Prediction of Experimental Phi Values in Protein Folding by Simulation with Knowledge-based Potentials: B Domain of Protein A

Sebastian Kmiecik, Maciej Blaszczyk, and Andrzej Kolinski

Laboratory of Theory of Biopolymers, Faculty of Chemistry, University of Warsaw
Pasteura 1, 02-093 Warsaw, Poland
E-mail: sekmi@chem.uw.edu.pl

Here we highlight our recent studies which showed that simulation of protein folding process, with the use of knowledge-based potentials and reduced representation of the polypeptide chain, can be a useful method in prediction of experimental Phi values. Apart from providing hints for experimental design, the simulation method (by delivering trajectory of conformations) enables interpretation of the Phi values. The interpretation is a non-trivial task, especially as the obtaining of reliable Phi values is not-trivial either. Here, we present a protein chain mobility profile, from simulation of B domain of protein A, consistent with a detailed Phi value analysis.

1 Introduction

Our recent works demonstrated the usefulness of protein folding simulation with knowledge based potentials for structural characterization of the folding process and denatured state ensemble[1–3]. Such a force-field approximation together with approximations of protein chain representation[4] seem to be a useful tool for qualitative studies of protein folding dynamics of large proteins (not accessible to all-atom Molecular Dynamics - a classical simulation method).

Experimental characterization of protein folding dynamics close to atomic resolution remains a big challenge. The detailed characterization of entire folding process is difficult for small proteins and usually impossible for larger. The only experimental technique sufficiently close to atomic resolution is a protein engineering method called Phi value analysis. The method sometimes leads to misleading results, requires much effort (extensive mutation scanning is needed) and reliable results may be expected only under certain conditions[5].

Giving some hints on the involvement of a particular residues in protein structure formation, Phi value analysis doesn't provide complete information on three-dimensional structure of transient conformers. In the past decade, interpretation of the experimental results leading to structural models[6] became possible thanks to simulation techniques (mainly all-atom Molecular Dynamics) utilizing experimental data and/or knowledge of an experimental (folded) structure.

2 Folding Chain Mobility of B Domain of Protein A and its Correspondence to Phi Values

Recently, while testing the hypothetical mechanism of chaperonin action, we described in detail the folding of B domain of protein A[3]. The results of this de novo simulation

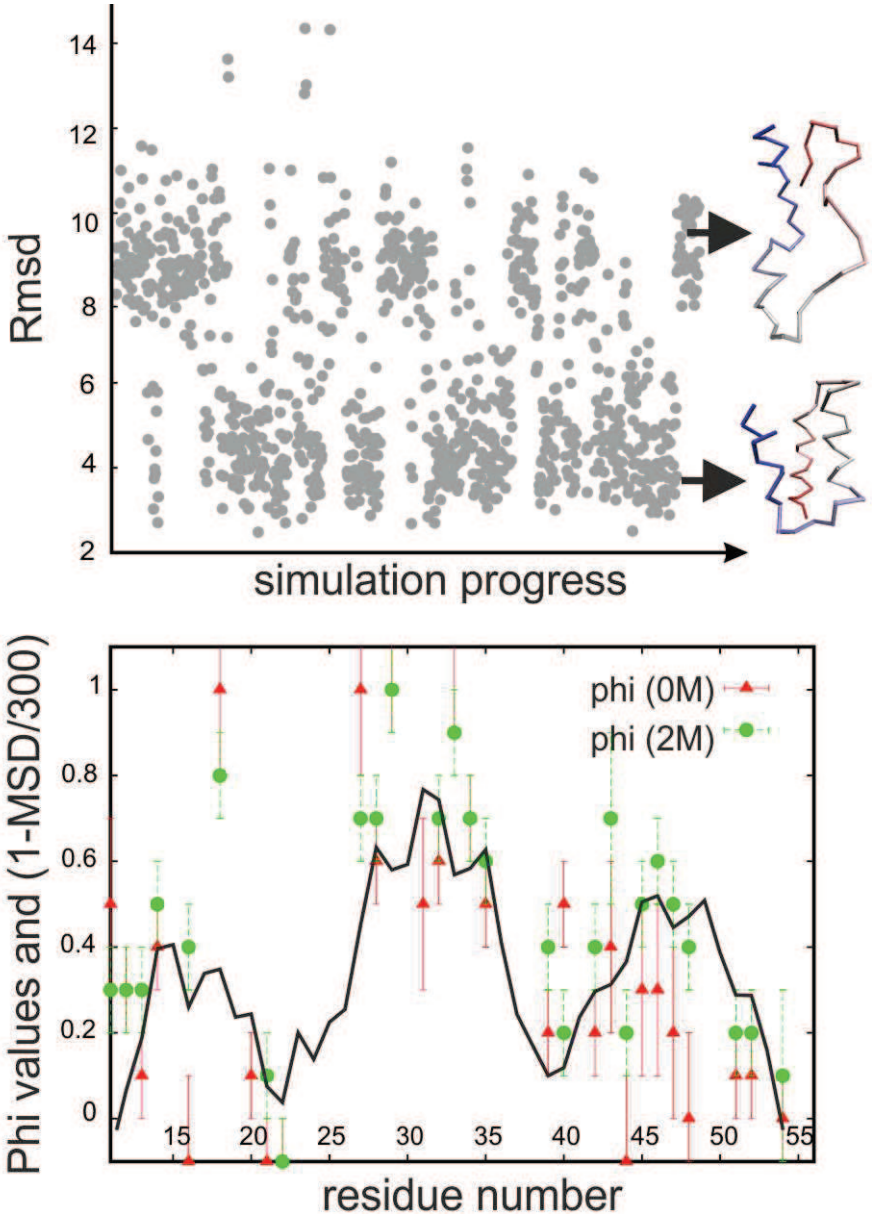

Figure 1. Upper panel: folding trajectory fragment of B domain of protein A (as a function of RMSD from native) together with example conformations from denatured and native like ensemble. Lower panel: profile of the protein chain mobility (marked with the line) superimposed onto Phi values data. The chain mobility is shown as (1-MSD/300), where MSD is a mean square displacement of each residue (average distance a given residue travels along the folding pathway)[3]. Phi values where marked in green and red according to denaturant concentration (0 and 2M GdmCl) together with error bars[7].

(not using any information about the native structure nor any experimental protein-specific data) appeared to be in perfect agreement with the detailed Phi value analysis[7]. This result is quite remarkable taking into consideration: (1) lack of utilizing any experimental data (unlike most of the studies so far), (2) other simulation results - not always agreeing with the experiment[7] and (3) coarse-grained approximations of the CABS model (force-field and protein representation).

Among the paradigm systems of protein folding we studied so far, the transition between the denatured and native like ensemble for B domain of protein A seem to be the easiest to happen along the folding trajectory. Multiple transitions from misfolded (mostly highly expanded) conformations (around 8-11 Angstroms from the native) into native-like (around 5-2 Angstroms from the native) can be observed in a reasonable computational time (in tens of minutes on a single CPU, see the upper panel of the Fig. 1).

As can be seen in the lower panel of the Fig. 1, there is a good correspondence between the protein chain mobility and Phi values – high Phi values associate with low mobility and low Phi values with high mobility along the folding simulation.

High Phi values appeared to be also very well correlated with the most frequently formed ternary interactions along the folding simulation of B domain of protein A (see the Fig. 2C in the Ref. 3). Frequency of contacts maps, derived from the simulation trajectories, enable easy interpretation of the Phi values[3] on the level of particular residue-residue interactions, as we showed also for the other paradigm systems of protein folding[1,2,8].

It is worth noting that the Phi values reflect the degree of structure nativeness around the mutation site in transition state: Phi values close to zero implies that the local structure around the mutation site is relatively unfolded (non-native), if close to 1 implies the similar structure as in the native state. The protein chain mobility measured as MSD (see description of the Fig. 1) reflects local structure involvement in any interactions, not only the native. Thus, the chain mobility profile is more straightforward measure of any structure formation in the particular region and gives the opportunity for studying impact of non-native interactions on observed Phi values.

Acknowledgments

The authors acknowledge the grant from Polish National Science Center (NN301071140).

References

1. S. Kmiecik and A. Kolinski, *Characterization of protein-folding pathways by reduced-space modeling*, Proc Natl Acad Sci USA, **104**, no. 30, 12330–5, 2007.
2. S. Kmiecik and A. Kolinski, *Folding pathway of the b1 domain of protein G explored by multiscale modeling*, Biophys J, **94**, no. 3, 726–36, 2008.
3. S. Kmiecik and A. Kolinski, *Simulation of chaperonin effect on protein folding: a shift from nucleation-condensation to framework mechanism*, J Am Chem Soc, **133**, no. 26, 10283–9, 2011.
4. A. Kolinski, *Protein modeling and structure prediction with a reduced representation*, Acta Biochim Pol, **51**, no. 2, 349–71, 2004.
5. A. R. Fersht and S. Sato, *Phi-value analysis and the nature of protein-folding transition states*, Proc Natl Acad Sci U S A, **101**, no. 21, 7976–81, 2004.

6. R. D. Schaeffer, A. Fersht, and V. Daggett, *Combining experiment and simulation in protein folding: closing the gap for small model systems*, Curr Opin Struct Biol, **18**, no. 1, 4–9, 2008.
7. S. Sato, T. L. Religa, V. Daggett, and A. R. Fersht, *Testing protein-folding simulations by experiment: B domain of protein A*, Proc Natl Acad Sci U S A, **101**, no. 18, 6952–6, 2004.
8. S. Kmiecik, M. Kurcinski, A. Rutkowska, D. Gront, and A. Kolinski, *Denatured proteins and early folding intermediates simulated in a reduced conformational space*, Acta Biochim Pol, **53**, no. 1, 131–44, 2006.