# From Coarse-Grained to Atomic-Level Characterization of Protein Dynamics: Transition State for the Folding of B Domain of Protein A

Sebastian Kmiecik,* Dominik Gront, Maksim Kouza, and Andrzej Kolinski*

Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland

ABSTRACT: Atomic-level molecular dynamics simulations are widely used for the characterization of the structural dynamics of proteins; however, they are limited to shorter time scales than the duration of most of the relevant biological processes. Properly designed coarse-grained models that trade atomic resolution for efficient sampling allow access to much longer time-scales. In-depth understanding of the structural dynamics, however, must involve atomic details. In this study, we tested a method for the rapid reconstruction of all-atom models from α carbon atom positions in the application to convert a coarse-grained folding trajectory of a well described model system: the B domain of protein A. The results show that the method and the spatial resolution of the resulting coarse-grained models enable computationally inexpensive reconstruction of realistic all-atom models. Additionally, by means of structural clustering, we determined the most persistent ensembles of the key folding step, the transition state. Importantly, the analysis of the overall structural topologies suggests a dominant folding pathway. This, together with the all-atom characterization of the obtained ensembles, in the form of contact maps, matches the experimental results well.
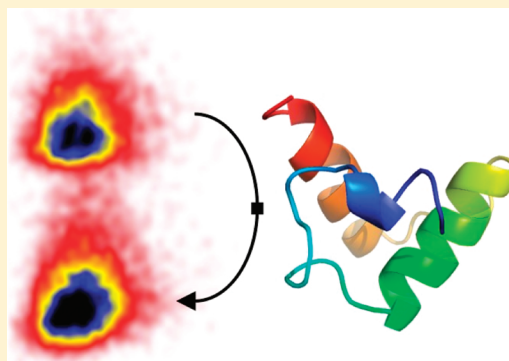
INTRODUCTION

The computational modeling of protein folding dynamics plays an important role in understanding the most important problems in molecular biology, such as the role of dynamics in biological function, protein folding and unfolding, protein interactions, etc.[1] In theory, the problem of protein structure and dynamics could be solved based on the first principles. Such attempts, initiated more than 30 years ago,[2] already brought remarkable successes in the atomic-level characterization of small and fast folding proteins and peptides.[3−5] These, however, came at an immense cost of computational power and also thanks to newly designed, customized hardware.[4,5] One of the directions to speed up the calculations is the introduction of highly efficient sampling techniques,[6] being constantly optimized to access larger time scales.[7] The most significant speed-up (by several orders of magnitude) is provided by the introduction of a simplified representation of the polypeptide chain.[1] Therefore, reduced (coarse-grained) models, often enhanced by efficient sampling techniques, are commonly used.[1] They have been proven to be useful tools for the efficient exploration of conformational space and proper characterization of protein dynamics that is consistent with experimental data, even without using any information about the experimental structure.[8,9]

In the summary of an excellent review "Protein-Folding Dynamics: Overview of Molecular Simulation Techniques" by Scheraga et al.,[1] the authors highlighted two future directions of development: force field improvement (both all-atom and reduced ones) and the need to design methods for converting coarse-grained trajectories to all-atom representation. In this work, we address the second issue of an accurate and efficient reconstruction of atomic details in the application for the analysis of protein dynamics. To achieve this, we applied a previously proposed method for the fast all-atom reconstruction and energy-ranking of protein structure prediction models.[10] The method enabled proper ranking of the quality of the models (distance from the native structure) when all-atom energy was used as the ranking criterion.

We tested the reconstruction method on a coarse-grained folding trajectory of the B domain of protein A (BdpA), one of the most extensively studied proteins in the context of the folding pathway available in the literature.[11] The folding trajectory was generated by the CABS model,[12] a well-established coarse-grained modeling tool employing stochastic dynamics and knowledge-based potentials. The model has proved to be a successful tool in the folding pathway description of a few model proteins,[9,13] including BdpA[14] (described shortly by testing the hypothetical mechanism of chaperonin action). The reconstruction method was applied to the whole 10 000 snapshot isothermal trajectory (purely de novo simulation conducted without any experimental protein-specific data), in which nearly 100 transitions could be observed from the fully unfolded to the near-native ensemble.[14] Next, we extracted the all-atom conformations of the BdpA transition state (TS) ensemble for the structural characterization of its most persistent conformers.

METHODS

CABS Model. The CABS method is based on a high-resolution coarse-grained model using a discrete lattice representation of polypeptide chains with 800 possible orientations of virtual α carbon bonds.[12] The conformational space is sampled by the Metropolis scheme (here under isothermal conditions). The force field employs statistical (knowledge-based) potentials, including generic protein-like conformational biases, short-range conformational propensities, the main-chain hydrogen bond model, and context-dependent side-chain interaction potentials. The model has already been successfully applied for high-resolution comparative modeling,[10,15] de novo structure prediction of small globular proteins,[15] and protein folding studies.[9,13,14]

The simulation was performed in the constant (reduced) temperature, near the transition temperature (assigned in the previous study[14]), and started from a random conformation prepared in a separate, short, high temperature run.

All-Atom Reconstruction. The reconstruction procedure was performed in two steps (see Figure 1). First, the protein backbone was reconstructed from alpha carbons using the Sybyl/Biopolymer module (implemented in Sybyl software, Tripos Inc. St. Louis, MO) with default settings. The method uses a "spare parts" approach[16] using backbone fragments

Figure 1. Reconstruction and minimization scheme shown on an example of a six-residue peptide. The scheme consists of the following steps: (1) protein backbone reconstruction from α carbon trace, (2) reconstruction of side chains from backbone chain, (3) all-atom minimization step (see Methods for detailed description). The orange circles in steps 2 and 3 mark the side chain which undergoes a significant conformational change during the minimization procedure.

retrieved from the protein database (PRODAT) to construct the complete polyalanine backbone (fragments of different lengths, from 4 up to 10 residues, that best match the α carbon trace and differ no more than the RMSD threshold of 0.5 Å, are used). In the second step, side chains were reconstructed on the basis of backbone atom coordinates by the SCWRL 4.0 package.[17]

The reconstruction procedure was similar to the previously applied protocol for the reconstruction and ranking of structure prediction decoys.[10] However, here a more accurate algorithm was used for side chains reconstruction than previously. According to our extensive tests, the SCWRL 4.0 approach results in a much smaller number of reconstruction inaccuracies than the Sybyl/Biopolymer side-chain reconstruction procedure. Such inaccuracies (incorrect rotamer assignment and atom−atom clashes) can hardly be corrected in the subsequent all-atom refinement step.[10]

All-Atom Minimization. The all-atom minimization (as the third step in the reconstruction and minimization procedure, see Figure 1) was performed in the same way as in the previous structure prediction study[10] using the methods implemented in the Sybyl package. The following minimization setup was used: Powell minimization method with frozen alpha carbons (1000 iterations), Amber7 f99 force field, Amber charges, and a dielectric constant of 1. Fixed positions of the alpha carbons during minimization and solvent absence significantly reduce the cost of computations for a large set of decoys (single minimization time was less than 1 min on a single CPU). As exercised by us and others (in tests of model assessment methods), there is very little added value when more rigorous molecular mechanics procedures[10] are used.
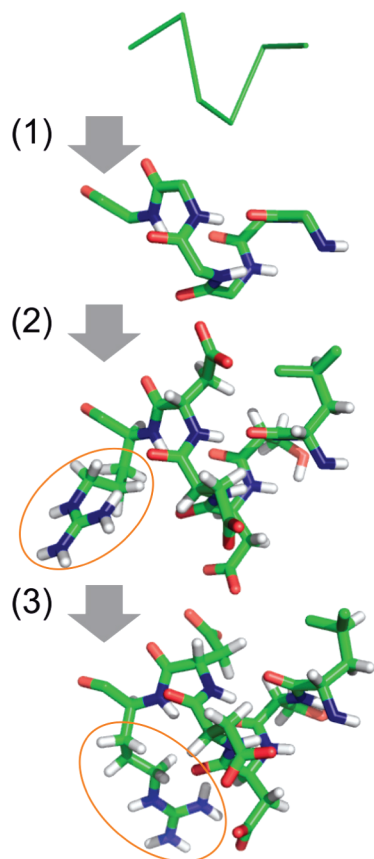
Structural Clustering and Analysis of All-Atom Contact Maps. The all-atom analysis was performed by the BioShell package.[18−20] In order to characterize the structural variety of the TS population, 1457 all-atom models were selected from a 10 000 snapshots trajectory based on the RMSD criteria (between misfolded/unfolded and near-native ensembles, see Figure 2). The selected models were subjected to structural clustering according to a complete link hierarchical algorithm[21] (Clust program from the BioShell package). The stopping criteria were set to cluster 75% of the initial objects (i.e., hierarchical clustering algorithm was merging objects into clusters until only 25% of the initial objects remained 1-element clusters designating them as the most diverse from the whole set), which resulted in 195 clusters of various size. The 10 largest clusters were subjected to further analysis. For each of them, a medoid structure was selected on the basis of the highest correlation between an all-atom contact map for a model and an average all-atom contact map for the cluster.

The accurate reconstruction algorithm used in this work enables us to rigorously define an inter-residue contact including also hydrogen atoms. Two residues are said to be in contact if any two atoms of the two residues are closer to each other than the sum of the van der Waals radii for these atoms. Such a definition revealed packing of the loose hydrophobic core in transient structures similar to the previous coarse-grained analysis.[14]
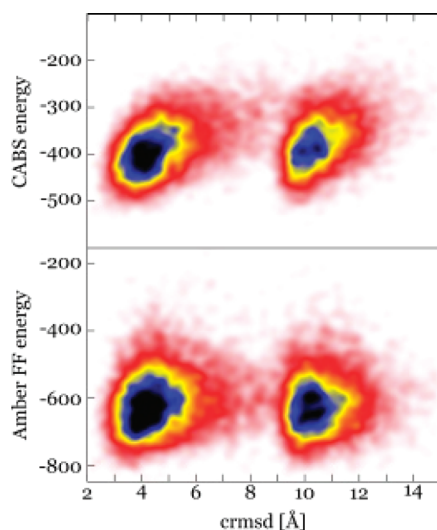
The protein pictures were generated using Pymol software.[22]

RESULTS

Performance of the Reconstruction Method. The applied rebuilding and optimization method appeared to be very efficient not only in terms of computational time (about 1

Figure 2. Distribution of the BdpA folding trajectory structures according to their cRMSD and energy values: CABS energy (upper panel) and all-atom Amber energy (lower panel). Histograms were calculated for all snapshots (10 000) of the single isothermal (near-transition temperature) trajectory. Colors denote frequencies of occurrence—each successive color from the sequence of red, yellow, blue, and black denotes twice as large appearance.

min per model on a standard CPU) but also in terms of the relatively consistent production of low-energy all-atom models. A very small number of models (0.76%) have been identified with abnormally high energy values (i.e., distinctively higher than the rest, with positive values ranging from +500 to +2500 kcal/mol), resulting from steric clashes which the structural refinement could not repack. The most substantial source of errors (2.51% of all models) was due to the inability of the backbone reconstruction method used to rebuild the complete backbone chain. This most likely resulted from the lack of appropriate short backbone fragments matching a given region of the α carbon trace (see Methods). These high-energy and incomplete models were excluded from further analysis. The performance of the reconstruction method seems to be similarly efficient in producing low-energy models as the method employed by Heath et al. to the reconstruction of folding trajectories.[23] The method by Heath et al. follows the same reconstruction steps but a different backbone reconstruction tool and a different all-atom minimization scheme including the GB/SA solvent model.

General View of BdpA Folding: Coarse-Grained vs All-Atom Characterization. Apart from providing a test of the proposed approach to convert coarse-grained dynamics trajectories into realistic all-atom ones, this work extends our results onto BdpA folding, initiated in the recent study by Kmiecik and Kolinski,[14] where BdpA was one of the systems characterized in the context of model chaperonin assistance. We chose BdpA for our rebuilding tests considering the abundance of the available experimental data and, importantly, for the sake of its CABS simulation characteristics. The CABS energy landscape for BdpA seems to be the least rugged among the model proteins we studied so far—the protein easily passes between the two ensembles (clearly separated by their resemblance to the native structure): near-native and unfolded/misfolded.[14] The results on BdpA folding, presented by Kmiecik and Kolinski, included the coarse-grained-level characterization of the TS ensemble averaged over all TS

structures, in the form of contact map, protein chain mobility profile, and secondary structure content (see Figures 2 and 4 and Table 2 of ref 14 and the discussion below). These results appeared to be in excellent agreement with experimental studies, particularly with the only experimental data sufficiently close to atomic resolution: the phi value analysis.[24]

In this work, we attempted to examine the resemblance of energy landscapes between coarse grained and all-atom structures, and to investigate thoroughly the structural variance of the TS of BdpA in all-atom resolution (see next sections). Energy landscapes, presented as two-dimensional functions depending on cRMSD and conformational energy values for coarse-grained structures versus all-atom counterparts, show good correspondence to one another (see Figure 2). The fact of the comparable energy values between near-native and denatured ensembles, reflected similarly in the CABS and AMBER99 force field, suggests general resemblance of the ensembles in the presence of attractive interactions. Indeed, the denatured ensemble contains a high number of misfolded structures with substantial content of native helices and a folded segment of H2−H3 helices. Interestingly, similar characteristics of the whole conformational ensemble in the transition temperature (being not a 50%−50% mixture of fully folded and unfolded conformations) was found in the MD simulations with the physics-based UNRES force-field.[25,26] Similarly to the CABS simulations, in the transition temperature, the UNRES simulations exhibited nearly equal populations of near-native and misfolded conformations, with similar characteristics (having native or near-native clusters of nonpolar residues in place) but different chirality.

The investigation of the correlation between the CABS and AMBER energy values has shown poor correspondence (see Figure 3). That means that the trajectory models are to some extent repositioned between each other within both energy basins (on the cRMSD vs energy maps, Figure 2), but interestingly, the basins preserve a similar probability shape. This poor correspondence is not surprising or disappointing in the light of our previous studies investigating the possibility of energy ranking of protein structure predictions based on all-atom energy values.[10] For instance, for structure prediction models in the vicinity of 3 Å from the native structure (which is the case of the near-native energy basin), the CABS force field is poorly correlated with the RMSD for the majority of proteins, just as the all-atom force-fields. The reconstruction and minimization method used here, when applied to the ranking of structure prediction results, shows very good performance, which places it among the state-of-the-art protein model assessments tool.[10] As tests showed, it is possible to discriminate only the high resolution decoys (RMSD around 1 Å from the native) from the broad set of high and moderate resolution models (1−3 Å from the native), as well as moderate resolution decoys (3 Å) from the low ones (RMSD > 4 Å and much higher).[10] Thus, one may expect the ability of the all-atom ranking to show the energy difference between near-native and misfolded ensembles. However, the above-mentioned successful rankings were performed on models built via classical methods of comparative modeling,[10] that is, maximally densely packed structures. In this work, we deal with ensembles of compact but loose structures (either near-native or misfolders) of a molten globule-like character, with a substantially lower sum of favorable interactions compared to the completely folded models, thus being much more difficult to rank.
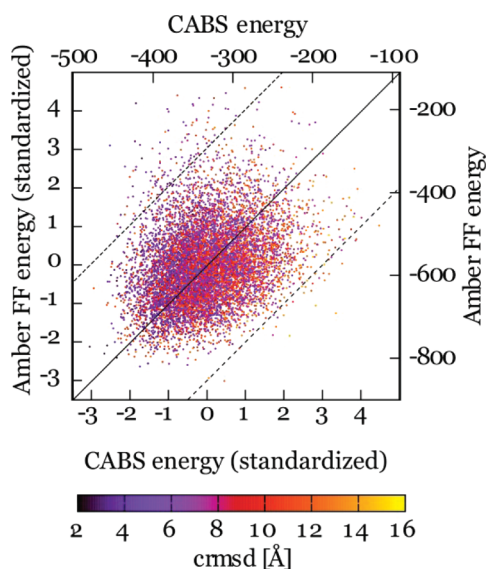
Figure 3. Distribution of coarse-grained CABS energy vs all-atom Amber FF energy for the trajectory models. The energy values are shown in appropriate units and also in standardized units. Additionally, the $y = x$ axis (solid line) and $y = x \pm 3\delta$ axes (dashed lines) are shown, where δ is the standard deviation of the data set shown in the figure (equal to 1 in the standardized units). The data points corresponding to the individual trajectory models were colored according to their cRMSD values, from deep blue to yellow (see the color legend below). The correlation coefficient for the data presented is 0.3, and after removal of the data points which differ by more than 3δ (constituting 2.5% of all models), the correlation is 0.36.

In summary, the good correspondence of CABS and AMBER folding landscapes, presented as energy values vs models resolution (RMSD) (Figure 2), reveals acceptable quality of local all-atom details of the models, since otherwise the AMBER energy values would be diffused compared to those of the CABS.

Secondary Structure: Coarse-Grained vs All-Atom Assignment. An accurate secondary structure assignment ultimately requires all-atom details. An approximate assignment based on coarse-grained representation is also possible. We computed the BdpA helical content (averaged over the whole trajectory) in order to compare coarse-grained and all-atom assignments. In the case of the CABS models, secondary structure was assigned by the P-SEA[27] algorithm based solely on the alpha-carbon trace. As for the reconstructed and refined all-atom models, the DSSP method[28] was used. The estimated percentages of secondary structure content evaluated for individual helices H1, H2, and H3 as an average over the all-atom structures were 21.1, 69.3, and 66.44%, respectively. Corresponding values computed from the CABS models[14] exhibit a similar tendency, although the actual values—13.9, 66.6, and 55.9%—were slightly lower.

These data show higher sensitivity of the all-atom recognition for alpha helices than the coarse-grained representation, which indicates a decent ability of the applied reconstruction method to retrieve realistic all-atom details of the alpha secondary structure.

Characterization of the BdpA TS in the Previous Study. The structural characterizations of the TS conformations of BdpA obtained in many simulation studies broadly fall into two classes: those highlighting the role of H1 and H2 interactions[29,30] and those stipulating the dominant role of H3

and its interactions with H2[31−36] (as summarized in ref 11). It was suggested that these inconsistencies may be caused by the existence of multiple TSs and folding pathways of BdpA, as concluded from statistical mechanical model analysis.[37] Interestingly, a later attempt to validate this hypothesis by experimental efforts toward searching for multiple BdpA folding pathways ended in a conclusion that the protein folds via a single dominant folding pathway (involving a single major TS),[38] which agrees with other (up-to-date) experimental data.

As mentioned earlier, in the recent work of Kmiecik and Kolinski,[14] we provided the general features of the BdpA TS, based on the analysis on a coarse-grained level. In order to study the effect of an external influence on the particular nucleation sites of BdpA, we classified all the TS conformations into four types (according to existence of the two previously suggested H1−H2 and H2−H3 nucleation sites) in which H1−H2 and H2−H3 are well formed, H1−H2 only, H2−H3 only, or none of them. This classification was based on the number of long-range contacts (native and non-native) between the centers of the side chain masses within H1−H2 and H2−H3. For instance (when considering a particular structure), if the number of long-range contacts within H1−H2 was significant (larger than the average number of H1−H2 contacts from the whole trajectory), while within H2−H3 it was not, the structure was then classified as the "H1−H2 only" type. The analysis revealed roughly equal populations of the four TS types (with slightly higher population of the "H1−H2 only" conformers than others). The proportions of the four TS populations varied according to temperature and external destabilization (chaperonin) cycles (for details, see Table 2 and discussion in ref 14). The average picture of all TS conformations reflected by the averaged side-chain contact map (based on the centers of side chain masses) was found to be strikingly consistent with the extended hydrophobic core identified in phi value analysis.[24]

Topological and All-Atom Characterization of the BdpA TS. Here, we extend the previously published results on the structural features of the BdpA TS to the all-atom characterization of the most persistent TS conformations along the folding pathway. This is done by structural clustering of the TS all-atom models followed by the selection of the representative models (based on all-atom contact map analysis, see Methods). Such an approach provides another level of classification for identifying similar structures (compared to the previous one, described above), including 3D information about the entire protein chain. The previous classification neglects the spatial arrangement within the subunits considered (H1−H2 and H2−H3) and between them; therefore, structures were classified solely on the basis of the number of contacts within H1−H2 and H2−H3, out of context of the entire system.

The structural clustering of TS conformers revealed a large number of highly diverse clusters (most of them very poorly populated). For further analysis, we extracted the 10 largest clusters (constituting about 37% of the whole TS population). Figure 4 presents an average contact map computed for each of the clusters numbered according to cluster size, from the largest (95 structures) to the 10th largest (28 structures), together with representative models. All the maps in Figure 4 reflect all-atom contact details of the structural diversity of the most persistent structures in the BdpA TS ensemble. The contact map for the native structure is provided for comparison in
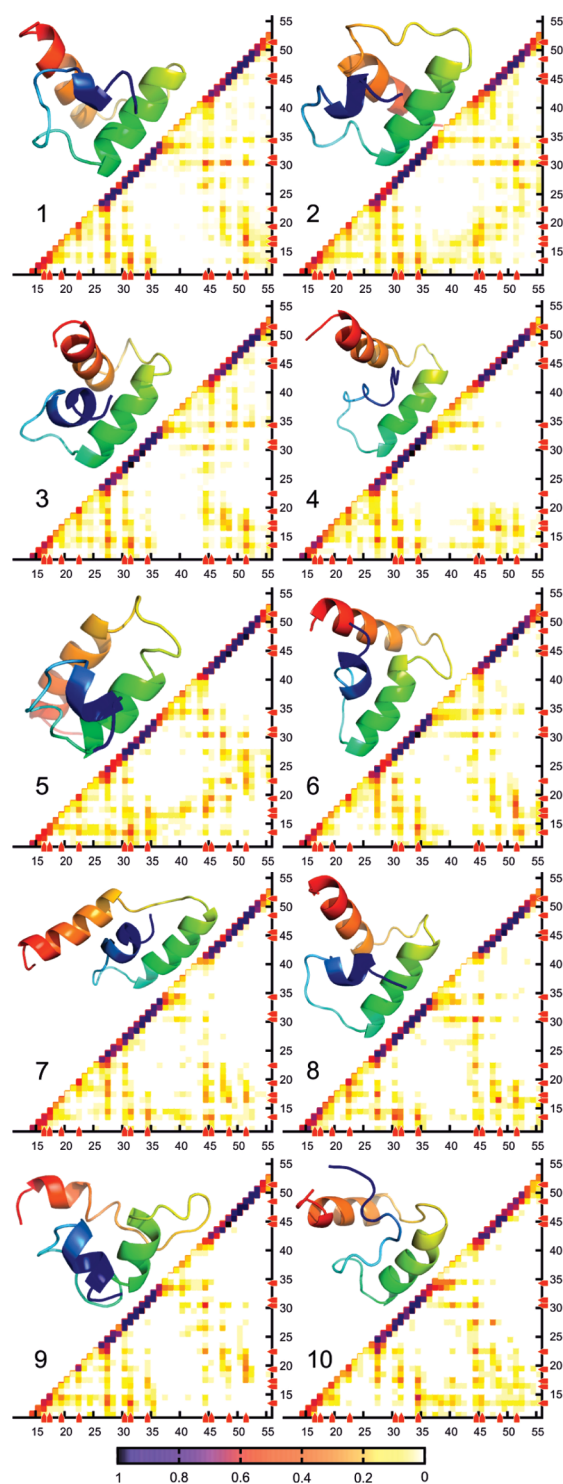
Figure 4. Structural diversity of the most persistent TS structures (obtained from structural clustering) reflected by average contact maps computed for each of the clusters and their representative models. The TS structures were numbered according to cluster size from the largest to 10th largest (the numbers of TS models in the clusters were as follows: 95, 86, 78, 58, 41, 37, 36, 31, 30, 28). The maps are colored according to the frequency of all-atom contacts in each cluster (see the color legend at the bottom of the figure). Additionally, residues identified in phi value analysis as belonging to an extended hydrophobic core in the TS are marked in red at the map borders. As a reference, the native contact map is shown in Figure 5.

Figure 5 (according to the same all-atom contact definition, given in Methods).
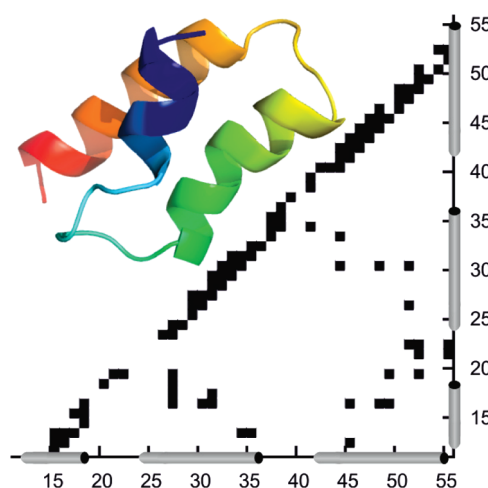
Figure 5. Native BdpA structure together with its all-atom contact map. Native helices are marked at the map borders.

Interestingly, the majority of the most persistent structures exhibit similar topology (see TSs numbered 1, 3, 4, 6, 8, 9, and 10 in Figure 4), with H2 and H3 chain fragments being perpendicular to each other (and rather in close contact) and H1 docked to H2−H3. With respect to H1 interactions, the H1−H2 contact pattern is clearly far more persistent and consistent than that of H1−H3. H1 itself is weakly structured and represents the most flexible chain fragment in the ensemble considered. Other spatial configurations found include a similar helices arrangement (as described above) but the opposite direction of H2 and H3 (see TS number 2), and near-native parallel arrangements of H2−H3 helices (see TS number 5) and H1−H2 helices (see TS number 7) with the third remaining helix drifted away.

As shown earlier, on the basis of the average contact map for all the TS structures (on a coarse-grained level),[14] the most persistent contacts forming an unstable network of tertiary interactions belong to the extended hydrophobic core identified in phi value analysis.[24] The same applies to maps representing the heterogeneity of the most persistent conformations (the hydrophobic residues identified in the experiment to form an extended core are marked at the map borders in Figure 4). For instance, H1 and H2 interactions are most frequently formed by F14, L18, F31, I32, and L35 (all having high phi values and also W15 for which phi values are not estimated[24]). With regard to H2−H3 interactions, L45 contacts with F31 and L35 play a dominant role, which perfectly agrees with the experiment.[24] The maps presented in Figure 4 suggest nearly formed helices H2 and H3; however, the secondary structure assignment by the DSSP algorithm[28] estimated the following fractions of the H1, H2, and H3 helices content (average values for all the TS structures): 0.19, 0.66, and 0.64, respectively (see also the secondary structure assignment section above). In comparison, the kinetic amide H/D isotope effect indicates that the TS has ∼70% of the native helical content.[39] The analysis presented above highlights the central role of H2 around which the TS structure is constructed and also the important role of the most disrupted helix H1 in the stabilization of the most persistent TS structures.

Each of the 10 TSs (listed in Figure 4) represents a cluster of structures that differ among one another up to 4.6 Å (in terms of cRMSD). The differences between different clusters (thus different TSs) are larger. Despite the fact that heterogeneity reflects the loose packing of the TS conformations, the majority of the TSs exhibit the same topology (numbered 1, 3, 4, 6, 8, 9, and 10 in Figure 4), that is, arrangement of the helices or chain fragments responsible for helices formation. This is a nontrivial result indicating the major route from the unfolded/misfolded to the near-native ensemble. The resulting picture of the major TSs (Figure 4) highlights the importance of the formation of both nucleation sites (H1−H2 and H2−H3). However, nucleation within H1−H2 seems to be more significant as well as homogeneous in the contact pattern than that of H2−H3 (which was shown also in the previous analysis[14]).

The loose packing of the TSs found is also reflected in the persistence of the all-atom contacts marked by colors in contact maps in Figure 4. While many local contacts resulting from helices formation exist in most of the structures, the long-range contacts responsible for the final nucleation event are present on average in half of them (contact frequencies around 0.5 marked in red). This indicates an unstable character of the nucleation sites being frequently very close to formation but not actually formed and oscillating around certain conformations. Thus, the present analysis suggesting the formation of both nucleation sites (H1−H2 and H2−H3) in the TS takes into account sometimes preformed nucleation sites, which could not be the case in the previous study[14] based solely on the number of contacts.

CONCLUSIONS

It is widely recognized that the development of new multiscale modeling methods which use different resolution techniques, both theoretical and experimental, is critical for the understanding of important macromolecular processes[40] such as protein dynamics. Such methods have already been shown to be useful in the structural dynamics characterization of fluids,[41] lipid bilayers,[42] and proteins.[23,43−46]

Here we characterize the structural dynamics of the TS of BdpA in all-atom resolution using coarse-grained simulation data converted by the fast structure reconstruction and optimization method. The results show that our method produces realistic all-atom models with a very high yield on the system tested (around 96%). Thus, this work provides a promising reference for future studies of protein dynamics aiming at larger protein systems (and of different topological classes) and method modifications (such as testing various solvation models). These studies are currently being carried out in our laboratory.

Because of the highly dynamic character of the BdpA TS conformers, the characterization of the whole ensemble requires some generalizations which we proposed in ref 14. This generalized analysis revealed more or less similar populations of the four conformer groups classified according to the existence or nonexistence of nucleation sites, H1−H2 or H2−H3 (defined without taking into account their spatial configuration). In this work, we took into consideration the spatial arrangement of the entire protein and narrowed the analysis to the most persistent BdpA structures only, providing their detailed all-atom characterization. The resulting picture of the most common TS structures is largely homogeneous at the general (topology) level, although heterogeneous in details of its structure and all-atom contact pattern, suggesting the

existence of a major folding route. This conclusion agrees with the available experimental data, since multiple folding pathways have not been evidenced experimentally, despite extensive studies by phi value analysis to test this hypothesis.[38] Moreover, our results fit well the topological requirements of the TS, which emerged from the quantification of the experimental data (phi value analysis and kinetic H/D amide isotope effects), suggesting a similar contact order to that for the native state.[39]

Numerous studies stressed the importance of BdpA folding either through an H1−H2 or an H2−H3 nucleation event. Our all-atom picture of the TS ensemble shows that both answers may be true in the sense that, if only one of the nucleation sites is formed, the remaining one is very likely to be very close to formation. This is a consequence of a loose character of the TS ensemble found, rather than competing TSs composed of either H1−H2 or H2−H3.

Finally, it is worth emphasizing that the presented TS ensemble represents the most common structures between the two major folding basins which both are collections of residually folded structures. Such a picture of the whole conformational ensemble is in excellent agreement with the recent simulation studies of BdpA folding.[25,26] Thus, the characterization of the TS described here provides structural models for transition between near-native and unfolded or, importantly, misfolded states (mainly topological mirror image of the native three-helix bundle).

AUTHOR INFORMATION

Corresponding Author
*E-mail: sekmi@chem.uw.edu.pl (S.K.); kolinski@chem.uw.edu.pl (A.K.). Phone: 48 822 02 11 (ext. 310) (S.K.); 48 822 02 11 (ext. 320) (A.K.). Fax: 48 822 02 11 (ext. 310) (S.K.); 48 822 02 11 (ext. 310) (A.K.).

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Scheraga, H. A.; Khalili, M.; Liwo, A. Annu. Rev. Phys. Chem. 2007, 58, 57−83.
(2) Levitt, M.; Warshel, A. Nature 1975, 253, 694−698.
(3) Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. Biophys. J. 2008, 94, L75−L77.
(4) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. Science 2010, 330, 341−346.
(5) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. Science 2011, 334, 517−520.
(6) Liwo, A.; Czaplewski, C.; Oldziej, S.; Scheraga, H. A. Curr. Opin. Struct. Biol. 2008, 18, 134−139.
(7) Kouza, M.; Hansmann, U. H. J. Chem. Phys. 2011, 134, 044124.
(8) Liwo, A.; He, Y.; Scheraga, H. A. Phys. Chem. Chem. Phys. 2011, 13, 16890−16901.
(9) Kmiecik, S.; Kolinski, A. Proc. Natl. Acad. Sci. U.S.A. 2007, 104, 12330−12335.
(10) Kmiecik, S.; Gront, D.; Kolinski, A. BMC Struct. Biol. 2007, 7, 43.

(11) Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6837−6838.

(12) Kolinski, A. *Acta Biochim. Pol.* **2004**, *51*, 349−371.

(13) Kmiecik, S.; Kolinski, A. *Biophys. J.* **2008**, *94*, 726−736.

(14) Kmiecik, S.; Kolinski, A. *J. Am. Chem. Soc.* **2011**, *133*, 10283−10289.

(15) Kolinski, A.; Bujnicki, J. M. *Proteins* **2005**, *61*, 84−90.

(16) Claessens, M.; Van Cutsem, E.; Lasters, I.; Wodak, S. *Protein Eng.* **1989**, *2*, 335−345.

(17) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L., Jr. *Proteins* **2009**, *77*, 778−795.

(18) Gront, D.; Kolinski, A. *Bioinformatics* **2006**, *22*, 621−622.

(19) Gront, D.; Kolinski, A. *Bioinformatics* **2008**, *24*, 584−585.

(20) Gront, D. *Bioshell - utility library for structural bioinformatics.* www.bioshell.pl.

(21) Gront, D.; Kolinski, A. *Bioinformatics* **2005**, *21*, 3179−3180.

(22) *The PyMOL Molecular Graphics System*, 1.3; Schrödinger, LLC.

(23) Heath, A. P.; Kavraki, L. E.; Clementi, C. *Proteins* **2007**, *68*, 646−61.

(24) Sato, S.; Religa, T. L.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6952−6956.

(25) Maisuradze, G. G.; Liwo, A.; Oldziej, S.; Scheraga, H. A. *J. Am. Chem. Soc.* **2010**, *132*, 9444−9452.

(26) Maisuradze, G. G.; Liwo, A.; Scheraga, H. A. *J. Chem. Theory Comput.* **2010**, *6*, 583−595.

(27) Labesse, G.; Colloc'h, N.; Pothier, J.; Mornon, J. P. *Comput. Appl. Biosci.* **1997**, *13*, 291−295.

(28) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.

(29) Boczko, E. M.; Brooks, C. L., 3rd *Science* **1995**, *269*, 393−396.

(30) Garcia, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898−13903.

(31) Alonso, D. O.; Daggett, V. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 133−138.

(32) Berriz, G. F.; Shakhnovich, E. I. *J. Mol. Biol.* **2001**, *310*, 673−685.

(33) Favrin, G.; Irback, A.; Wallin, S. *Proteins* **2002**, *47*, 99−105.

(34) Ghosh, A.; Elber, R.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10394−10398.

(35) Jang, S.; Kim, E.; Shin, S.; Pak, Y. *J. Am. Chem. Soc.* **2003**, *125*, 14841−14846.

(36) Kussell, E.; Shimada, J.; Shakhnovich, E. I. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5343−5348.

(37) Itoh, K.; Sasai, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 7298−7303.

(38) Sato, S.; Fersht, A. R. *J. Mol. Biol.* **2007**, *372*, 254−267.

(39) Baxa, M. C.; Freed, K. F.; Sosnick, T. R. *J. Mol. Biol.* **2008**, *381*, 1362−1381.

(40) Russel, D.; Lasker, K.; Phillips, J.; Schneidman-Duhovny, D.; Velazquez-Muriel, J. A.; Sali, A. *Curr. Opin. Cell Biol.* **2009**, *21*, 97−108.

(41) Ayton, G. S.; Tepper, H. L.; Mirijanian, D. T.; Voth, G. A. *J. Chem. Phys.* **2004**, *120*, 4074−4088.

(42) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750−760.

(43) Thorpe, I.; Zhou, J.; Voth, G. *J. Phys. Chem. B* **2008**, *112*, 13079−13090.

(44) Feig, M.; Karanicolas, J.; Brooks, C. L., 3rd *J. Mol. Graphics Modell.* **2004**, *22*, 377−395.

(45) Kolinski, A. *Multiscale Approaches to Protein Modeling*; Springer: New York, 2011; p 335.

(46) Whitford, P. C.; Noel, J. K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K. Y.; Onuchic, J. N. *Proteins* **2009**, *75*, 430−441.