Sampling small-scale and large-scale conformational changes in proteins and molecular complexes

Mi-Ran Yun
Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Institut de Biologie Physico-Chimique et Université Paris 7, 13 Rue Pierre et Marie Curie, 75005 Paris, France

N. Mousseau\textsuperscript{a,b}
Département de Physique et Centre de Bioinformatique Robert-Cedergren, Université de Montréal, CP 6128, Succursale Centre-Ville, Montréal (Québec) H3C 3J7, Canada

P. Derreumaux\textsuperscript{a,c}
Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Institut de Biologie Physico-Chimique et Université Paris 7, 13 Rue Pierre et Marie Curie, 75005 Paris, France

(Received 6 December 2006; accepted 25 January 2007; published online 8 March 2007)

Sampling of small-scale and large-scale motions is important in various computational tasks, such as protein-protein docking and ligand binding. Here, we report further development and applications of the activation-relaxation technique for internal coordinate space trajectories (ARTIST). This method generates conformational moves of any complexity and size by identifying and crossing well-defined saddle points connecting energy minima. Simulations on two all-atom proteins and three protein complexes containing between 70 and 300 amino acids indicate that ARTIST opens the door to the full treatment of all degrees of freedom in dense systems such as protein-protein complexes. © 2007 American Institute of Physics. [DOI: 10.1063/1.2710270]

I. INTRODUCTION

Ensuring an efficient sampling of the energy landscape of dense systems is a challenge that arises in many fields of science. In the context of computational biology, such problems include protein folding and aggregation, protein-protein docking, ligand binding, DNA-protein binding, and many others. While proteins adopt well-defined structures in solution, they are not static objects, but are in constant motion and structural deformations are often observed during binding.\textsuperscript{1,2} In the case where the protein displacement is limited to a few residues within the catalytic pocket and the structure of the target protein with a similar ligand is available, a rigid treatment of the protein backbone is not a severe limitation, simplifying considerably the theoretical treatment.\textsuperscript{3} However, many crystal structures of complexes display large-scale motion of the protein upon ligand binding that are generally very difficult to characterize numerically. For instance, the EF-hand calcium-binding proteins, belonging to one of the most populated domains in the human genome, adopt very different bundles of four $\alpha$ helices, depending on the presence and concentration of ions.\textsuperscript{4} Similarly, protein kinases display an inherent plasticity in response to interactions with specific partners and can reveal atomic displacements of specific regions by 10–15 Å from their unbounded states.\textsuperscript{5} Experimental and theoretical studies also show that protein domains are, in fact, not rigidly packed, but are rather dynamic with movements including the hinge bending and the shear motions.\textsuperscript{6}

Over the last two decades, there has been considerable work on the design of theoretical methods for sampling conformational changes in large proteins and molecular complexes. We can group most approaches into three main types of methods. The first group of methods involves the use of conventional molecular dynamics (MD) simulations.\textsuperscript{7–9} Since these simulations typically cover the nanosecond time scale, large conformational changes cannot be explored. Similar approaches can also use real-space Monte Carlo (MC) based procedures, but these are as costly to perform and lack information about the time scale.\textsuperscript{10}

The second type of methods consists in reducing the total number of degrees of freedom of the system by restricting motions to subset of the phase space. Several solutions have been tested, restricting motion to flexible side chains of the ligand,\textsuperscript{11} flexible side chains of both the receptor and the ligand,\textsuperscript{12} and flexible ligand with flexible side chains of the active site combined to flexible or selected main chain atoms of the receptor. These simplifications, generally coupled to implicit solvent representations, have opened the door to many MC based methods,\textsuperscript{13} multicopy mean-field approaches,\textsuperscript{14,15} and discretizations of the backbone fragment prior to docking.\textsuperscript{16} Identification of hinge bending regions followed by rigid-body docking of the resulting subdomains is also an interesting approach to reduce the total number of degrees of freedom.\textsuperscript{17} These algorithms have led to improved docking results in a number of cases, but are limited by the number of flexible loops that can be considered within the complexes.

The third group of approaches focuses on identifying the most relevant protein motions in order to move the protein structure in the subspace spanned by the first few lowest-
frequency vibrational modes. These soft modes can be determined by a normal-mode analysis of the interaction matrix constructed geometrically using the position of Ca (Refs. 18 and 19) or they can be obtained by principal component analysis of MD simulations. It is possible to go beyond the harmonic description of the protein motion by using MD simulations to introduce anharmonicity in the calculation of the protein vibrational modes. This procedure is costly, however, as it is necessary to sample the appropriate ensemble of states, which might involve many hundreds of minima.

Here, we present the development of a method in line with this third group of approaches and consider the full anharmonicity of the energy landscape by moving over well-defined saddle points. Along this idea, we recently combined the use of internal coordinates with a first-order saddle point search method and proposed the activation-relaxation technique for internal coordinate space trajectories (ARTIST) which samples the energy landscape by identifying well-defined transition states connecting two local minima. ARTIST is the result of merging two programs: LIGAND, an internal coordinate minimization program, which defines the internal degrees of freedom and the intermolecular rotations and translations, and ART NOUVEAU, an activation-relaxation algorithm. As a first step, we showed on four model peptides with 4–47 amino acids that ARTIST gives good results in finding folding paths from an unfolded state to the ground state and in generating native-like conformations characterized by the simultaneous motions of the loops, secondary structures, and side chains. In our previous work, we had limited the application of ARTIST to small proteins as the algorithm could not handle the frequent steric clashes that occur in large and dense proteins.

The focus of this article is to show that it is possible to circumvent this limitation by refining the activation process. The most recent version of ARTIST, dubbed Version 2.0, is modified to minimize steric collision so that it can deal with short-scale and large-scale motions within proteins and protein complexes. This article is organized as follows. We present the model proteins and the force fields used. We then give a detailed description of the modified algorithm. Finally, we illustrate its properties by using two proteins of 73 and 158 amino acids, two heterodimers of (10, 57) and (89, 96) amino acids, and an assembly of four polypeptide fragments of 291 residues, and by focusing essentially on the exploration of the conformation space surrounding the initial state, as a result of force field limitations.

II. MATERIALS AND METHODS
A. Protein structures and energy models

The initial structures for simulations are taken from the protein data bank. The two proteins in solution are the C-terminal domain of calmodulin [amino acids 76–148, code 1CMF (Ref. 25)] and the 158-amino acid enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) [code 1HKA (Ref. 26)]. The selected heterodimers consist in the 96-amino acid ribonuclease SA complexed by the 89-amino acid barstar [code 1AY7 (Ref. 27)] and the 57-amino acid GRB2 N-terminal SH3 domain complexed by a 10-amino acid inhibitor [code 2GBQ (Ref. 28)]. The final model deals with the interaction between the 62-amino acid eglin C inhibitor and bovine alpha-chymotrypsin, which is composed of three polypeptide segments spanning residues 1–13, 16–146, and 149–245 [code 1ACB (Ref. 29)].

All systems use an all-atom representation with all polar and nonpolar hydrogen atoms included. Energy calculations are carried out using either the AMBER force field and the parm98 parameter set or the FLEX force field. Solvent electrostatics is treated using either a sigmoidal distance-dependent dielectric function or the continuum solvent generalized Born (GB), and no cutoff distances are used. In an earlier report, we show that the diversity of the structures generated by a single ARTIST event is independent of the implicit solvation model used. Minimization at the relaxation step of ARTIST is performed with a quasi-Newton algorithm using analytic first derivatives of the internal and solvent energy contributions with respect to the internal degrees of freedom, and is considered to have converged when the predicted energy gain at the next step falls below 10−3 kcal mol−1.

B. Moving on the energy surface with ARTIST v2

Following ART NOUVEAU, each ARTIST event consists of four steps, steps (i) and (ii) constituting the activation phase and steps (iii) and (iv) the relaxation phase. (i) Leave the present local minimum along a random direction and find a negative eigenvalue in the Hessian matrix using the Lanczos technique. (ii) Push the system along the corresponding eigenvector, while minimizing the forces in the perpendicular hyperplane until the total force is close to zero. (iii) Push the system over the saddle point and minimize the conformation in the new local minimum. (iv) Accept the newly generated conformation using a Metropolis criterion: 

\[ p = \min(e^{-\beta \Delta E}, 1) \]

where \( \beta = 1/k_B T \) and \( \Delta E = E_{\text{final}} - E_{\text{initial}} \), the energy difference between the final and initial local minima. Since ARTIST moves from local minimum to local minimum, the Metropolis temperature cannot be simply related to the physical temperature. We typically select a value that ensures a reasonable acceptance rate. A similar approach to ARTIST was also introduced in Ref. 33 and applied to several small proteins.

In contrast to Cartesian coordinate space where the displacement of a variable generates to local deformations, small torsional moves usually result in large structural effects at remote regions, leading, in some cases, to collisions, and necessitating often sophisticated compensatory moves in order to create a structure with a localized change. This steric hindrance, particular acute in compact conformation, is encountered regularly in protein docking and was a major limitation in the original version of ARTIST.

To control or limit atomic collisions, it is common to use soft-core Lennard-Jones potentials or reduced representations. Here, we prefer to keep the full-atom representation and not to use soft-core nonbonded potentials because if such interactions are acceptable for generating dis-
continuous trajectories, their impact on transition points is uncertain. Instead, we modify the activation step used in Ref. 23 and in the first version of ARTIST, 22 by adding an adaptive approach to offer a better relaxation in the hyperplane perpendicular to the direction of activation, while moving along the direction of negative curvature with an adaptive step to avoid head-on collisions between atoms.

The new algorithm implemented in ARTIST V2.0 responds much more smoothly to the change in the energy landscape, allowing the particles to go around each other instead of colliding heads on. This is done by using a two-step approach. (1) The size of the displacement is adaptive: as the energy increases, the steps are reduced to improve the responsiveness to the rapid changes of the landscape. (2) A conjugate-gradient minimization, rather than a quasi-Newton minimization, is now used in the hyperplane perpendicular to the eigenvector followed toward the saddle point, leading to a better overall minimization. With these modifications, collisions are now fully under control and are handled, as they should be, by the activation algorithm.

In all simulations, activation and relaxation of the protein models are carried out using all torsion angles of the backbone and of the side chains in addition to the valence angles of the backbone. All protein bond lengths and all valence angles of the side chains are kept fixed. Starting from the NMR- or x-ray-minimized structure, the system is moved away from its local minimum by applying perturbations, of the order of 5°–10°, to a total of 20 variables randomly chosen from all ϕ, ψ, and χ dihedral angles, which are free to move. The valence angles of the backbone are not selected at this stage, but are free to move during the activation and relaxation steps. More sophisticated perturbations, based on normal mode analysis, 18 for instance, could be attempted, but this is beyond the scope of this work. In what follows, the Metropolis temperature used to accept or reject each move is set at 600 K.

III. RESULTS

While the original version of ARTIST works well on Lys-Phe-Phe-Glu (KFFE) tetrapeptide, dodecalanine, a 27-residue α-helical hairpin, and the 47-residue N-terminal domain of RNase II, its application to large proteins and molecular assemblies was not successful. Specifically, we find that the activation phase does not properly handle collisions resulting from changes in the dihedral angles in dense environments. Using the original approach for the present protein-protein complexes, 70% of the attempted activations lead to head-on collisions between atoms causing significant increase in the total energy and 10% of the attempted activations loose a negative eigenvalue in the Hessian, leading to a very low acceptance rate for locating a well-defined transition state. In what follows, we present the results with ARTIST Version 2.0, which succeeds in lifting this limitation, opening the door to applications to larger systems.

A. Single-domain protein: C-terminal domain of calmodulin

We first examine the properties of the algorithm in sampling the rigid-body motion of the C-terminal domain of calmodulin. This domain is known to display different bundles of helices upon calcium binding.35–37 This independent unit is subjected to a series of 100 events. The calcium ion is not treated here. Two test cases are studied here using the AMBER-GB force field. In the first test case, the four helices are kept rigid and only the ϕ and ψ torsional angles of the loop residues (namely, residues 20–25, 34–40, and 54–60) are free to move during ARTIST events. In the second test case, all torsional angles are flexible.

In Figs. 1(a) and 1(b), we show the variations of the six interhelical angles calculated using the procedure described in Babini et al. 4 as a function of all generated ARTIST events with the helices held rigid. At a Metropolis temperature of 600 K, the acceptance rate is 58%. The protein moves through a wide range set of states during the simulation, with interhelical angles varying between 30° and 160°. For instance, helices I and IV are nearly orthogonal at event 8, while they are nearly antiparallel at event 31. Figure 1(b) shows the effects of ARTIST moves on the topology. The conformational changes during this 100-event simulation are considerable. The structure generated after one single accepted event (S1) is compact, but deviates from the NMR-minimized structure S0 by 6.4 Å root-mean-square deviation (RMSD). Structure S100, stabilized by 37 kcal/mol with respect to S0, is completely unfolded, with a RMSD of 15.4 Å.

To understand what happens if all torsional angles are flexible, we run three 100-event simulations (R01–R03), starting from the NMR structure. Figure 1(c) shows the lowest-energy structures obtained by each simulation using AMBER-GB and different random seeds. Again, ARTIST unfolds the C-terminal domain of calmodulin with RMSDs of 4.9 Å (R01), 10.7 Å (R02), and 5.9 Å (R03) from the NMR-minimized structure. Such a high mobility is also seen with the FLEX force field (data not shown).

To our surprise, the C-terminal domain of calmodulin is a problematic target for the AMBER-GB and FLEX force fields. Because it is possible that AMBER, developed for treatment of biomolecules using Cartesian coordinates, cannot be used in internal coordinates without relaxing the valence angles of the side chains and the bond/bond angles terms, we minimize the structures in Cartesian space using both AMBER-GB and CHARMM-EEF1. 38 Again, we find that the force fields cannot discriminate the native (S0) from the unfolded state (S100), although the generated structures are realistic and pass successfully all the PROCHECK tests. 39 This result highlights the limitations of implicit solvent force fields when used in combination with methods, such as ARTIST, which are not limited to the exploration of small-scale motions around the starting structure.

To avoid moving into unrealistic conformations as in the calmodulin case, we will limit ARTIST simulations to 20 events, focusing on the exploration of the conformational space surrounding the initial state. For the systems below, we
will show the first accepted structure (S1) and the lowest-energy structure generated after 20 events (S20). All these simulations are performed with FLEX, a potential adjusted for internal coordinate description.

B. Single-domain protein: HPPK enzyme

Using the protocol presented above, we study a second single-domain protein in solution and focus on the plasticity of three loops observed between the structures of the apo-HPPK enzyme and the HPPK-MgATP complex. Comparison of the apoenzym and binary complex crystal structures points to a significant motion of three surface loops: loop 1 (residues 9–14), loop 2 (residues 42–53), and loop 3 (residues 80–94). Here, even though we are interested in the motion of three loops, the full protein is flexible during ARTIST events. At a Metropolis temperature of 600 K, the acceptance rate is 35%.

Figure 2 shows the superposition of the first accepted structure (S1) [Fig. 2(a)] and the lowest-energy structure (S20) [Fig. 2(b)] on the NMR-minimized structure (S0). S1 and S20 deviate by 0.5 and 0.7 Å RMSDs from S0. The displacement motions are further analyzed in panels (c) and (d) which report the changes in the Ca positions as a function of the residue numbers between the S1 (S20) structures and S0 structures. The variations of the side-chain torsional angles of the loop residues in going from S0 to S20 are also shown in panels (e)–(g).

While the motion of loop 1 is relatively confined, with a maximum change in Ca positions of about 1.0 Å, loops 2 and 3 show larger displacements, reaching 3.5 Å in S20. Even though the total displacement of loop 1 is smaller, the torsional motion of the side chains from their initial states is of similar amplitude in all loops, as shown in panels (e)–(g), with maximum changes in the side chain torsional angles reaching 180°. This result clearly illustrates the capability of ARTIST to move simultaneously main-chain and side-chain atoms in large proteins.

C. Protein/protein docking

We now investigate whether ARTIST can deal with protein-protein docking. For the three models studied here, the main-chain and side-chain atoms of all partners are flexible, along with each protein with respect to the other(s). The first two models start from the bound complex structure, whereas the third model starts from the unbound structure of eglin C and the bound structure of alpha-chymotrypsin superimposed on the complex.

N-terminal SH3 domain with inhibitor. This binary system contains a 57-residue protein with an inhibitor of ten amino acids. In Fig. 3, panels (a)–(c) show three states: the minimized bound complex structure (S0), the first accepted event (S1), and the lowest-energy event (S20). At 600 K, the Metropolis acceptance rate is 40%. The structures S1 and S20 deviate by 2.9 and 7.5 Å RMSDs from S0. We see that the inhibitor moves significantly, exploring distinct ϕ and ψ regions and rotating by about 90° with respect to the protein. However, this motion is not independent of that of the protein; it is followed in part by the N-terminal extremity of the SH3 domain at event 1 and even by the unfolding of the C-terminal region of the SH3 domain at event 19. This motion is clearly visible in the panels (d) and (e), which show the angular (ϕ, ψ) changes of all residues from their initial values in S0. As seen, ARTIST easily handles collective motions involving both partners simultaneously, as expected for protein-protein docking. It is also clear from panel (g) that the lowest-energy structure S20 contains only few contacts.

FIG. 1. Calmodulin plasticity. (a) Evolution of the six interhelical angles (in deg) as a function of all generated structures starting from the NMR-minimized structure. (b) NMR structure and ARTIST-generated structures at events 1 (S1) and 100 (S100) with the helices held fixed. (c) Final ARTIST structures with all particles free to move.
between the partners compared to the nativelike S1 structure [panel (f)].

Ribonuclease SA with barstar. Here, we examine the interaction of two proteins of 96 and 89 residues. The panels (a) and (b) of Fig. 4 show the superposition of the minimized bound complex structure (S0) on the first accepted (S1) and the lowest-energy structure (S20). At 600 K, the Metropolis acceptance rate is 42%. S1 and S20 are found to deviate by 5.9 and 7.0 Å RMSD from S0. Panels (c) and (d) of Fig. 4 show the per-residue Cα RMSD, calculated over each chain separately, as the second chain rotates with respect to the first. Using a best-fit procedure on each chain, chain 1 deviates by 1.2 and 2.3 Å RMSDs in going from S0 to S1, and from S0 to S20, respectively, while chain 2 moves by 0.7 and 1.2 Å RMSDs. In chain 1, we see a significant motion of two loops (residues 40–46 and 60–68), after only 1 event, with Cα displacements of 2 Å [Fig. 4(c)]. Continuing the simulation extends the sampling of the conformational space. In S20, all the loops in both chains have moved by more than 2 Å, with some amino acids in chain 2 moving by as much as 4 Å [Fig. 4(d)]. As seen in Fig. 4(e), this internal motion is not only limited to the loops, but is fully delocalized on the protein backbones. As for proteins in solution, flexibility also takes place within the side chains; analysis of the generated structures shows that the side chain torsional angles can be modified by more than 70° in a few ARTIST events.

Alpha-chymotrypsin—eglin C. Finally, we consider the recognition of alpha-chymotrypsin (241 residues) by eglin C (62 residues) to illustrate the likely usage of ARTIST for sys-

FIG. 2. Characterization of the HPPK enzyme flexibility. Superposition of the NMR-minimized S0 structure (black) on the first accepted (S1, gray, panel a) and the lowest-energy structure (S20, gray, panel b). The backbone and side-chain atoms of Ala12 (loop 1), Glu48 (loop 2), and Arg88 (loop 3) are also indicated. [(c) and (d)] Changes in the Cα positions (in Å) between the S1 (S20) and S0 structures using a best-fit superposition procedure. [(e)–(g)] Changes in the side-chain torsional angles in going from S0 to S20 within loop 1 (e), loop 2 (f), and loop 3 (g).

FIG. 3. Evolution of the N-terminal SH3 domain–inhibitor structure. Superposition of the S0 (a), S1 (b), and S20 (c) structures. In each chain, the N-terminal and C-terminal ends are also indicated. Changes in the φ and ψ dihedral angles (in degrees) of S1 (d) and S20 (e) as a function of the residue numbers with respect to the minimized S0 structure. Δψ angles in dotted lines. For simplicity, the inhibitor sequence is numbered from S8 to 67 and is separated from the receptor sequence by a vertical line. Contact maps of the S1 (f) and S20 (g) structures.
tems of 300 residues or more. In contrast to previous complexes, we use the unbound structure of eglin C superimposed on the complex structure (we therefore move one step toward the real world scenario). Figures 5(a) and 5(b) shows the superposition of the unbound NMR structure [code 1EGL (Ref. 40)] and the bound complex structure (code 1ACB) on the lowest-energy structures generated by two independent 20-event runs. Figure 5(c) shows the changes in the Ca positions of the inhibitor between the lowest-energy ARTIST structures and the unbound starting structure, and between the unbound and bound complex structures. We find that ARTIST moves affect all amino acids, with larger displacements within loops 40–47, in good agreement with the experimental motion between the free and bound inhibitor structures. However, in run 1, the interface is native, whereas in run 2, it is non-native.

IV. DISCUSSION

The results presented above show that it is possible, using a careful activation procedure such as the one introduced in ARTIST2, to handle a large variety of short-scale and long-range motions within single-domain proteins and complexes.
As in any Monte Carlo procedures or sampling techniques, the lowest-energy structure cannot be located within 20 or 100 ARTIST steps and the outcome will change from one run to another (see calmodulin test case). However, if the simulations were extended, we expect that ARTIST2, as its variant in Cartesian coordinate space, will converge to the lowest-energy structure, independently of the initial conformation.24

Because of the limitations of the AMBER-GB and FLEX force fields in discriminating native from non-native conformations, we could not apply ARTIST2 to a real scenario where we start from the unbound structures nor show any correlation to experimentally observed movements. As we have shown, the C-terminal domain of calmodulin unfolds with these potentials, independently of the number of internal degrees of freedom used. Similarly, the SH3-peptide and the alpha-chymotrypsin-eglin C systems show that ARTIST with the FLEX force field favors conformations that contain only few native contacts between the partners. We emphasize that a different calibration of the rigid body degrees of freedom relative to the intramolecular degrees of freedom preserves the native interface, but the generated structures are of higher energy than those expected. It follows that the motion presented here does not necessarily occur in vitro. They correspond to physically and mechanically feasible transitions on the AMBER-GB and FLEX energy surfaces, however. It remains to be determined to what extent the sampling of the energy landscapes (minima and saddle points) will be affected by other energy functions.

The averaged CPU times on a 3 GHz microprocessor are 15, 5, and 35 min for generating a single activation-relaxation event on the (1CMF, 1131 atoms using AMBER +GB; 1HKA, 2535 atoms using FLEX; and the heterodimer 1AY7, 2877 atoms using FLEX). This time may appear too slow to cover appropriate ensembles of states and suggest a principle limitation of the method for large systems. This is not the case. First, because the number of force evaluations to locate a transition state is independent of the number of variables, the CPU cost can be alleviated by truncating the nonbonded interactions. We find that the CPU time drops from 35 to 5 min for 1AY7 if we truncate the interactions at 14 Å. Second, the CPU cost will also be substantially reduced if we use the standard “united-atom” representations for the CH2 and CH3 groups, rather than the all-atom representation.

V. CONCLUSIONS

We have modified the activation-relaxation technique for internal coordinate space trajectories (ARTIST) to handle properly atomic collisions in dense environments. ARTIST V2.0 now includes an adaptive step size and the use of conjugate-gradient minimization during the activation phase. The revised algorithm is then applied to two proteins of 73 and 158 amino acids and three protein complexes in settings varying from dimer to tetramer and a total number of amino acids varying between 67 and 291.

For all cases, we show that collective effects are fully included as ARTIST allows for moves of any complexity and size: localized on a few residues or distributed over all main chain atoms and side chain atoms of both molecules simultaneously with the translation and rotation of one protein with respect to another. This result is significant because ARTIST uses an all-atom representation and a physically based molecular mechanics energy function in contrast to many methods, where a simplified chain model or a soft-core nonbonded potential is used.41 In addition, no preliminary knowledge of the final state is needed in contrast to double-ended saddle point techniques43 or nonlinear conformational change path methods.44

Comparing the performances of ARTIST2 with respect to previously published protein-protein and protein-drug algorithms is beyond the scope of the present work. What is clear, however, is that this study demonstrates the efficiency of our revised algorithm to generate short-scale and large-scale motions in single-domain protein or during protein-protein docking.

ACKNOWLEDGMENTS

One of the authors (N.M.) is supported in part by grants from NSERC and the Canada Research Chair program, and was supported by a CNRS post rouge in 2006. Another author (P.D.) thanks CNRS and Université de Paris 7 for financial support, and the IDRIS and CINES centers for computational support.

32 V. Tsui and D. A. Case, Biopolymers 56, 275 (2000).
41 J. Cherfils, T. Bizebard, M. Knossow, and J. Janin, Proteins 18, 8 (1994).
42 M. Zacharias, Protein Sci. 12, 1271 (2003).