JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

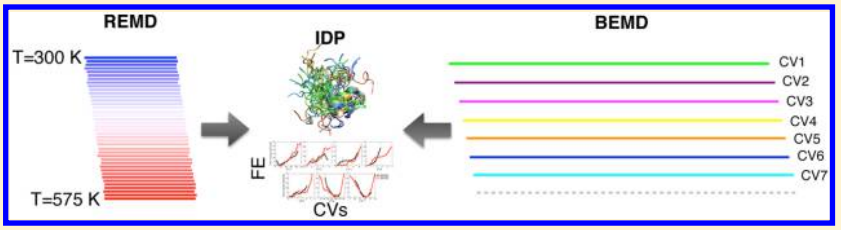
# Free Energy Surface of an Intrinsically Disordered Protein: Comparison between Temperature Replica Exchange Molecular Dynamics and Bias-Exchange Metadynamics

Gül H. Zerze,[†] Cayla M. Miller,[†] Daniele Granata,*[,‡] and Jeetain Mittal*[,†]

[†]Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, Pennsylvania 18015, United States

[‡]Institute of Computational and Molecular Science, Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States

Supporting Information




**ABSTRACT:** Intrinsically disordered proteins (IDPs), which are expected to be largely unstructured under physiological conditions, make up a large fraction of eukaryotic proteins. Molecular dynamics simulations have been utilized to probe structural characteristics of these proteins, which are not always easily accessible to experiments. However, exploration of the conformational space by brute force molecular dynamics simulations is often limited by short time scales. Present literature provides a number of enhanced sampling methods to explore protein conformational space in molecular simulations more efficiently. In this work, we present a comparison of two enhanced sampling methods: temperature replica exchange molecular dynamics and bias exchange metadynamics. By investigating both the free energy landscape as a function of pertinent order parameters and the per-residue secondary structures of an IDP, namely, human islet amyloid polypeptide, we found that the two methods yield similar results as expected. We also highlight the practical difference between the two methods by describing the path that we followed to obtain both sets of data.

## INTRODUCTION

Many proteins naturally exist in unstructured states and are hence referred to as intrinsically disordered proteins (IDPs).[1−3] Despite their lack of a well-defined structure, many of these proteins perform various biological functions, and many of these proteins are implicated in human diseases.[4,5] Heterogeneous nature of disordered proteins makes their experimental characterization challenging. Therefore, IDPs are ideal candidates for molecular dynamics (MD) simulations, which provide the ability to probe molecular characteristics on length and time scales that are not easily accessible to current experimental methods. However, many molecular events take place on much longer time scales than those that are feasible using standard brute-force MD simulations. As a result, relevant regions of the configuration space often remain unexplored in a given simulation time. To alleviate this problem, several enhanced sampling methods have been introduced,[6−10] including temperature replica exchange molecular dynamics (T-REMD)[11] and bias exchange metadynamics (BEMD).[12]

T-REMD involves simultaneously running multiple simulations of a system at different temperatures. At high temperatures, energetic barriers are more easily overcome, and regularly attempted exchanges between adjacent replicas enhance the sampling of lower temperature replicas. The canonical probability distribution at each temperature is maintained by an acceptance probability of exchange between replicas, based on the Metropolis criteria, $P \propto \exp(\Delta\beta\Delta E)$, where $\beta = (k_B T)^{-1}$ and $\Delta E$ is the potential energy difference between the two states. Because of the dependence of $E$ on system size, a disadvantage of the T-REMD method is that larger system sizes require closer temperature spacing, and therefore a greater number of replicas, to maintain sufficient exchange probability. The T-REMD method has been successfully employed in the investigation of various IDPs including amyloid-$\beta$,[13,14] $\alpha$-synuclein,[15,16] NCBD,[17,18] p53 peptides,[19] and islet amyloid polypeptide (IAPP).[20−23]

The recently introduced BEMD technique is designed to both accelerate the conformational transitions and reconstruct the free energy landscapes of complex biological systems.[12,24,25] As in standard metadynamics,[8] sampling is enhanced by a bias that acts on a set of order parameters. These parameters are referred to as collective variables (CVs) and provide a low-dimensional description of the system. The bias is a history-dependent potential constructed as a sum of Gaussian distributions deposited along the trajectory of the CVs,[8] filling

Received: January 19, 2015
Published: April 28, 2015




DOI: 10.1021/acs.jctc.5b00047
J. Chem. Theory Comput. 2015, 11, 2776−2782

the CV space. In standard metadynamics, the computational cost increases exponentially with the number of CVs, and the method usually becomes inefficient for more than three variables; however, BEMD overcomes this problem with a replica-exchange approach.[11] The bias is applied on several CVs, which act singularly on separate replicas of the same temperature, providing the framework to study larger systems and more complex behavior, such as protein folding[25] and aggregation[26] as well as conformational sampling of IDPs.[27]

In this paper, we present a comparative analysis of human islet amyloid polypeptide (hIAPP), an IDP, in aqueous solution using two different enhanced sampling techniques: T-REMD and BEMD. hIAPP is a 37 residue-long polypeptide that normally exists as a soluble IDP under physiological conditions but forms amyloid deposits in the pancreas of patients with type II diabetes.[5,28−30] In vitro, hIAPP can aggregate very quickly,[31] making its monomeric properties difficult to characterize experimentally. As a result, the hIAPP monomer is the subject of many simulation studies.[20−23,32−35] Here, comparison of the free energy profiles and secondary structures populated in the hIAPP ensemble given by each method reveals very similar results. BEMD simulations show the capability to assess the free energy for larger regions of conformation space with less simulation time than T-REMD. However, BEMD simulations require the expense of initial benchmark runs to find the best set of CVs and CV parameters as well as computational expense of CV evaluation and bias potential calculation. Additionally, BEMD simulations require additional steps during analysis, while T-REMD simulations are straightforward to set up and analyze.

## METHODS

In this work, we compare two different enhanced sampling methods, T-REMD and BEMD, to obtain equilibrium properties of an IDP. Thirty seven residue-long hIAPP was simulated in each case using the Amber03w force field[36] and TIP4P/2005 water model.[37] The Amber03w force field has previously been shown to accurately model intrinsically disordered proteins.[18,19,23,38−40] The hIAPP monomer was solvated in an octahedron box with 3165 water molecules, with five sodium and seven chloride ions added to neutralize the system charge. Electrostatic interactions were calculated using the particle-mesh Ewald method[41] with a real space cutoff of 0.9 nm, and a 1.2 nm cutoff was used for van der Waals interactions. Both T-REMD and BEMD simulations of hIAPP were performed in an NVT ensemble, employing the GROMACS 4.5.3 package as the MD engine.[42,43] Further details specific to our T-REMD and BEMD methods are provided in the following sections.

**Temperature Replica Exchange Molecular Dynamics.** The enhanced sampling in the case of T-REMD is achieved by high-temperature replicas, which can easily overcome barriers within the configurational landscape. The temperature spacing of replicas is chosen such that the potential energy distributions overlap, and exchange attempts are accepted with a sufficient probability of ∼10−20%.[11] Because the distribution of potential energy $E$ broadens with increasing temperature, geometric spacing of the replica temperatures, given by eq 1, is a useful way of assigning replica tempearatures.[44]

$$T_i = T_0\left(\frac{T_{N-1}}{T_0}\right)^{i/N-1} \quad (1)$$

Here the indices of replicas are $i = 0,1,...,N-1$; and $N$ is the total number of replicas, so that $T_0$ is the lowest replica temperature, and $T_{N-1}$ is the highest replica temperature.

Forty replicas from 300 to 575 K were used, with the temperature spacing between them adjusted according to eq 1. The system dynamics were propagated using stochastic Langevin dynamics with a 1/ps friction coefficient. Each replica was simulated for 200 ns, for a total simulation time of 8 $\mu$s, using a time step of 2 fs, and exchanges attempted every 1 ps. We find that, on average, ∼12−13% of exchange attempts are accepted. The 300 K trajectory was analyzed, discarding the first 50 ns of the trajectory as equilibration time.

**Bias-Exchange Metadynamics.** In bias exchange metadynamics simulations, multiple metadynamics simulations are run in parallel at the same temperature (300 K), each one biased using a different CV. Exchanges between the bias potentials of different replicas are periodically attempted according to a Metropolis scheme: if the exchange move is accepted, the trajectory that was previously biased in the direction of the first variable continues its evolution biased by the second, and vice versa. In this manner, a relatively large number of different variables can be biased, and a high-dimensional space can be explored with a computational cost that increases linearly with the number of CVs.[45] Once the bias potentials become stable, a multidimensional free energy surface can be reconstructed, following an accurate reweighting scheme, described by Marinelli et al.,[24] as a function of the CVs used in the sampling. The convergence is typically assessed by considering the difference between the one-dimensional free energy profiles of the first and second halves of the simulation after the equilibration time.[35,46]

*Choice of the Collective Variables.* As in other methods that reconstruct the free energy as a function of a set of generalized coordinates, in BEMD the choice of the CVs plays an essential role in determining the convergence and efficiency of the free energy calculation. If the chosen set of CVs is not able to capture the slowest degrees of freedom to trigger the conformational transitions or does not distinguish between different metastable states of the system, the simulation will be affected by hysteresis, and not all of the important regions of the conformational space will be explored. To choose an appropriate set, one needs to exploit some basic knowledge of the topological, chemical, and physical properties of the system. Although there is no a priori recipe for finding the correct set of CVs, in the BEMD method the number of variables can be relatively large, making the selection less critical.

To develop a preliminary idea of the conformational space of the system explored by BEMD, we first performed BEMD of 600 ns per replica at 300 K with only four replicas. This informed us about our final selection of CVs and relevant parameters for production (parameter details given in Supporting Information). Given the intrinsically disordered nature of hIAPP, we chose CVs that are able to rapidly sample the secondary structure of the system: $\alpha$-RMSD, para$\beta$-RMSD, and anti$\beta$-RMSD[47] on the full length chain, as well as the coordination number of hydrophobic side-chain contacts,[25] to rapidly open or collapse the protein conformation. In the given simulation time, all CVs except the $\alpha$-RMSD CV converged, while this CV displayed an evident hysteresis (see Supporting Information, Figure S1). While the $\beta$-sheet content is expected to be low,[23] this IDP has been shown to populate primarily helical conformations in varied and independent regions of the sequence. This heterogeneous behavior, together with the


DOI: 10.1021/acs.jctc.5b00047
J. Chem. Theory Comput. 2015, 11, 2776−2782

relatively short simulation length (2.4 $\mu$s of total simulation time), did not allow the system to explore all relevant states along the $\alpha$-RMSD CV.

Accordingly, we decided to divide this CV into two different residue spans in two separate replicas. We employed the region spanning residues 8−22 in one and residues 22−36 in the other. We excluded the first seven residues because this region is characterized by a disulfide bond between cysteine residues 2 and 7, and we expected this region would not sample extended helical conformations. We then exploited all the structures collected in the previous BEMD run to understand the proper set of collective variables to use. On the basis of this analysis, we decided to introduce two additional CVs: dihedral correlation, which measures the similarity between adjacent $\psi$ dihedral angles, and the coordination number of $C_\alpha$−$C_\alpha$ contacts. The first was chosen to enhance the exploration of local structures (for example, PPII), which are quite common in IDPs,[48] while also helping the convergence of the $\alpha$-RMSD CVs and the sampling of the first seven residues. The $C_\alpha$−$C_\alpha$ coordination was introduced because the number of hydrophobic side-chain contacts alone may not effectively sample the compactness of the molecule, given the relatively low content of hydrophobic residues in IDPs compared to natively folded proteins.[49] Correlation plots for these two CVs are shown in Figure 1. Neither $\alpha$- nor anti$\beta$-RMSD are correlated with

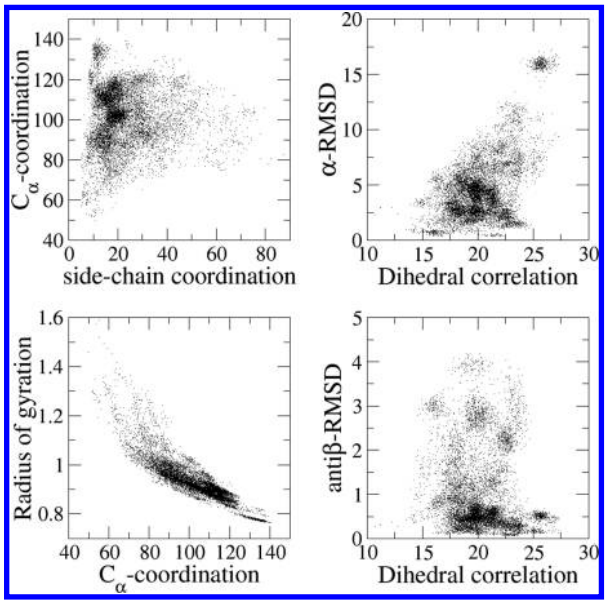



**Figure 1.** Correlation of collective variables including coordination number of hydrophobic side-chain contacts, $\alpha$-RMSD, antiparallel-$\beta$-RMSD, coordination number of $C_\alpha$−$C_\alpha$ contacts, and correlation of dihedral $\psi$ angles.

dihedral correlation (Figure 1, right). Remarkably, while the $C_\alpha$ coordination is highly correlated with the radius of gyration (a less ambiguous measure of compactness), it appears to be uncorrelated with hydrophobic side-chain contacts. It is important to stress that noncorrelation among a set of chosen CVs ensures the rapid exploration of larger regions of the phase space and allows discrimination between different structural states by projecting the system on orthogonal directions in the free energy reconstruction. Therefore, we decided to perform a BEMD simulation employing these seven collective variables (anti$\beta$-, para$\beta$- and two $\alpha$-RMSD, hydrophobic and $C_\alpha$ coordination numbers, and dihedral correlation) and an additional unbiased (neutral) replica to compare with the final free energy reconstruction and with the T-REMD simulation (see Supporting Information). The system dynamics is propagated using a leapfrog algorithm with a 2 fs time step. The temperature is kept constant at 300 K using the Nosé−Hoover thermostat[50,51] with a 1 ps coupling constant. The exchanges between the replicas were attempted every 20 ps, and each replica was simulated for 650 ns, for a total simulation time of 4.5 $\mu$s (excluding the neutral replica, which is used only for monitoring purposes). The initial 300 ns from all replicas are discarded as an equilibration time. All the metadynamics parameters and the explicit functional forms of the collective variables are reported in Supporting Information, as implemented in the plugin for free energy calculation, PLUMED v1.3.[52]

## RESULTS AND DISCUSSION

To compare the T-REMD and BEMD methods, we first calculated the free energies of the system as a function of individual collective variables. For the analysis of the BEMD simulation data we divided the seven-dimensional CV space by applying a grid on each direction (see Supporting Information, Table S1 for grid dimensions). In this way, all states after the equilibration time are grouped into sets (microstates), using the METAGUI[46] tool, a graphical user interface for Visual Molecular Dynamics (VMD).[53] The free energy of each microstate was estimated by the weighted-histogram analysis method[54] as described by Marinelli et al.,[24] considering both the metadynamics bias potential and the number of times that portion of the CV space was visited. We then integrated the free energies of the microstates over each collective variable to obtain one-dimensional free energy profiles projected onto the corresponding direction. This integration also takes into account the error associated with the free energy of each microstate as obtained from the METAGUI plugin, estimated as described by Marinelli et al.[24] Conversely, for the T-REMD simulation we evaluated the collective variables for the 300 K trajectory, calculating the potential of mean force (PMF) for the equilibrated 150 ns data on each CV based on the same grid dimensions used for the BEMD reconstruction (Table S1). In this case the corresponding errors are calculated by dividing the data into two blocks (see Supporting Information, eq S6).

As plotted in Figure 2, the two methods show free energy minima on each collective variable at the same values for most of the CVs: 15, 0.5, 0.5, 0.25, 0.25, 95, and 21, respectively, for coordination number of hydrophobic side-chain contacts, $\alpha$-RMSD of residues 8−22, $\alpha$-RMSD of residues 22−36, parallel $\beta$-RMSD, antiparallel $\beta$-RMSD (here, BEMD shows a minimum 0.4 kJ/mol lower at 0.75 than at 0.25), coordination number of $C_\alpha$−$C_\alpha$ contacts (here T-REMD shows a minimum 0.7 kJ/mol lower at 115 than at 95), and correlation of $\psi$ dihedral angles. In general, the free energy profiles of both methods show agreement within their error ranges. In particular, the shapes of the one-dimensional profiles indicate the same behavior up to ∼10 kJ/mol for all the CVs. Only CV3 shows a slight discrepancy over the range of values from 1.5 to 3.5, which is due to a computational detail in the calculation of secondary structure as discussed later. The neutral replica also shows similar profiles (Supporting Information, Figure S2). It is worth noting that BEMD usually allows the free energy to be determined for CV values of higher free energy and, consequently, over a greater part of the conformational space


DOI: 10.1021/acs.jctc.5b00047
J. Chem. Theory Comput. 2015, 11, 2776−2782

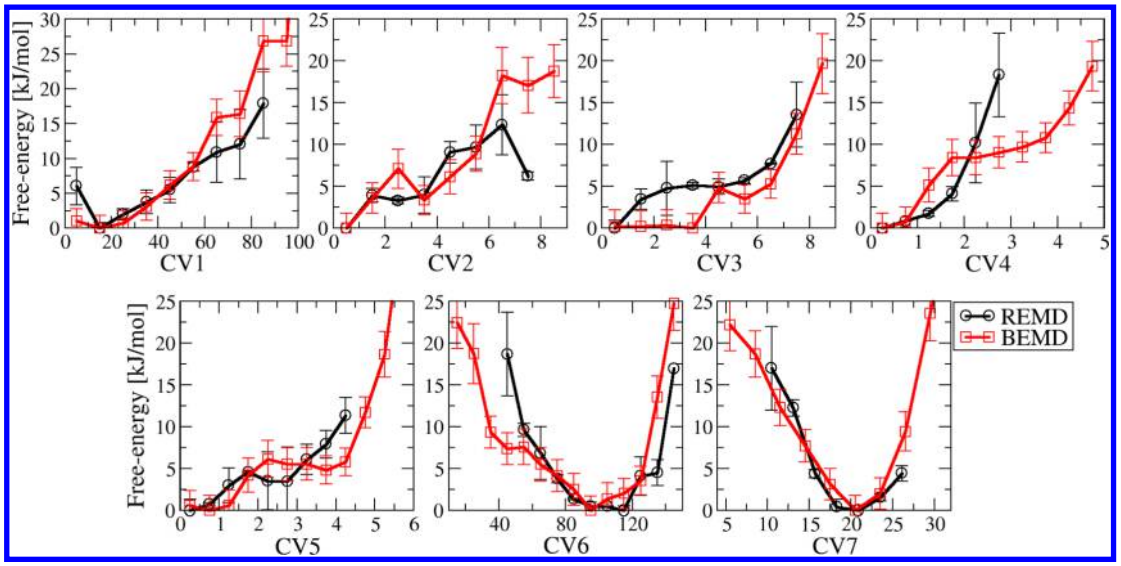



**Figure 2.** Comparison of free energies projected on each collective variable obtained from T-REMD and BEMD. Collective variables are CV1: coordination number of hydrophobic side-chain contacts, CV2: $\alpha$-RMSD of residues 8−22, CV3: $\alpha$-RMSD of residues 22−36, CV4: parallel-$\beta$-RMSD, CV5: antiparallel-$\beta$-RMSD, CV6: Coordination number of $C_\alpha-C_\alpha$ contacts, CV7: correlation of dihedral $\psi$ angles. Error bars for BEMD are calculated as described by Marinelli et al.[24] and for REMD by dividing the data into two blocks.

of the peptide than with T-REMD. This also explains why we see larger deviation between the T-REMD and BEMD free energy profiles at higher free energies, where the PMF evaluation of the T-REMD trajectory is based on very few events.

For IDPs, which have a flat free energy landscape, one cannot assess convergence based on the free energy between two states (characterized by different CV values), as in the case of a two-state folder. However, one can still obtain useful information from the convergence of free energy surface as a function of simulation time. As shown in Supporting Information, Figure S3 free energy profiles for both T-REMD and BEMD converge rapidly for most CVs, but the slowest CVs require significant additional time.

Naturally, secondary structures are of great importance for an IDP such as hIAPP, which has an amyloidogenic nature, as amyloid fibrils are described by a high content of $\beta$-sheets.[55] Moreover, helical intermediates are proposed to serve as an intermediate preceding the formation of $\beta$-sheets.[56,57] Therefore, one must be able to obtain secondary structures accurately, independent of the simulation technique, to definitively evaluate biophysical properties.

Secondary structure assignments for each frame of the output trajectories were determined using the DSSP algorithm.[58] Both total and per-residue secondary structures were averaged over the equilibrated T-REMD trajectory. To determine the correct ensemble averages for the BEMD, the number of residues assigned to each secondary structure was reweighted based upon the free energy of the microstate to which each frame belongs (see eq S5 in Supporting Information), taking the relative error into account as well. Figure 3A shows the average number of residues of hIAPP in each particular structure; $\pi$-helix, which was negligibly populated, is not shown here. For the 37 residue-long hIAPP, both methods compare extremely well with each other for all structure types, usually within one residue. Consistent with the disordered nature of the peptide, we observe a high preference for unstructured motifs (coil, bend, and turn), moderate residual helical conformations (~20%), and very low $\beta$-sheet-like structures (6%). To further

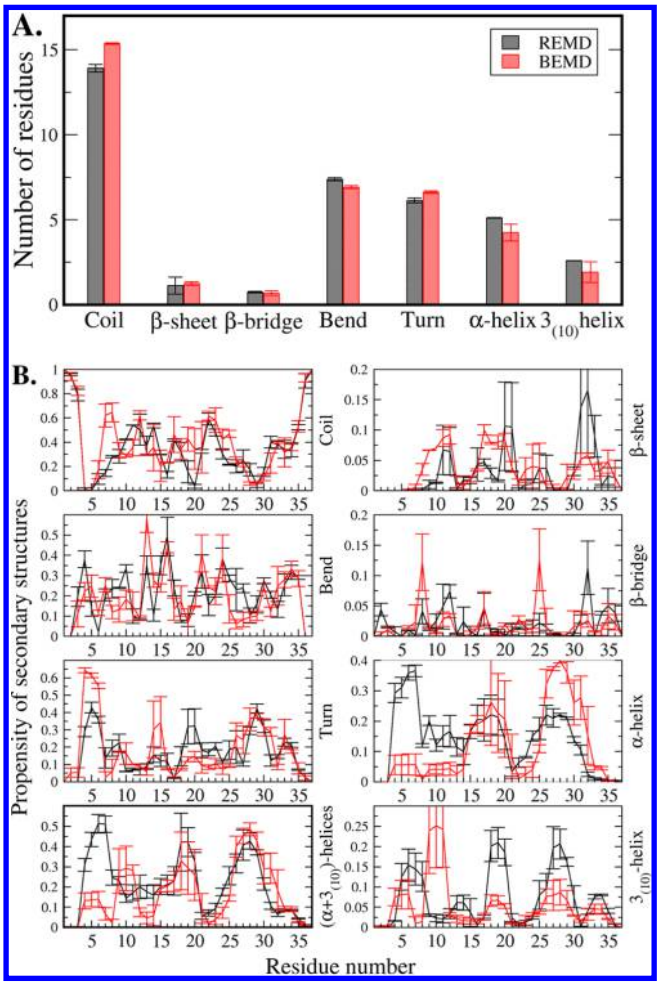



**Figure 3.** (A) Average number of residues in various types of secondary structures determined from both methods. (B) Per-residue fractions of secondary structures found from both methods. Error bars are calculated from block averages for REMD and from errors of microstate free energies for BEMD.



DOI: 10.1021/acs.jctc.5b00047
J. Chem. Theory Comput. 2015, 11, 2776−2782

elaborate on the secondary structure comparison, assignment of per-residue secondary structures is shown in Figure 3B. While propensities for particular structures agree very well for the majority of the structures at per-residue level, appreciable differences appear in $\alpha$- and $3_{(10)}$-helical structures. When summed together as the total helix fraction, $\alpha$+$3_{(10)}$-helices (Figure 3B) shows fairly strong agreement between BEMD and T-REMD, especially over residues 8−37. This further explains the discrepancy between the one-dimensional profiles along CV3 ($\alpha$-RMSD for residues 22−36), shown in Figure 2. This collective variable detects only the presence of standard $\alpha$-helices, whereas the T-REMD simulation shows a lower $\alpha$-helix propensity than BEMD over this region in favor of $3_{(10)}$-helices. As a result, the T-REMD simulation is assigned a higher free energy for more extended helices. Given the close match of the probability distributions for the total helix fraction in this specific region (Figure 3B), the relative free energy projections for $\alpha$+$3_{(10)}$-helices would also be the same. Conversely, in the same panel of Figure 3B there is a poor agreement over residues 2−7. Here, helical propensity is reduced in favor of increased turn propensity for BEMD (Figure 3B). This is likely because the first seven residues were excluded from the $\alpha$-RMSD bias (see Methods Section). Consequently, this exclusion slows the convergence of helical fraction in that region and causes the mismatch between the results from BEMD and T-REMD. The formation of $\alpha$-helices is a relatively simple process (compared to $\beta$-sheet formation), yet including these residues in CV7 (dihedral angle correlation) was not sufficient. This finding clearly highlights the importance of extending all relevant collective variables in the biasing scheme to the entire sequence for higher efficiency and, in particular, indicates the value of secondary structure collective variables for appropriate sampling.

The multidimensional reconstruction of the free energy in BEMD allows for the more meaningful representation of free energy surfaces in higher dimensions, such as in the three-dimensional projection along the two $\alpha$-RMSD variables and the coordination of $C_\alpha$ contacts in Figure 4A. The red, blue, and yellow isosurfaces correspond to free energy surfaces at 1, 2, and 3 kcal/mol, respectively. Representative structures from the lowest free energy microstates of the BEMD simulation are also shown. As expected given the intrinsically disordered nature of the peptide, the system shows a shallow, flat region of low free energy, which is characterized by unstructured conformations or very short secondary structure elements. Beyond this region, the free energy increases rapidly, as shown previously in the one-dimensional profiles of Figure 2. For comparison with the BEMD simulations, we also performed backbone clustering analysis of the equilibrated 300 K T-REMD trajectory using the gromos method[59] with a 0.3 nm cutoff distance. The representative structures of the top five populated clusters are shown along with their corresponding average CV values (Figure 4B). Though the two clustering algorithms are completely different, the low free energy regions of the BEMD simulation and top populated clusters of the T-REMD simulation also show similar structures. Both methods capture similar disordered and partially helical states, as shown by collective variable values. Again, the small visual differences are mainly due to the lower helicity over residues 2−7 in BEMD.

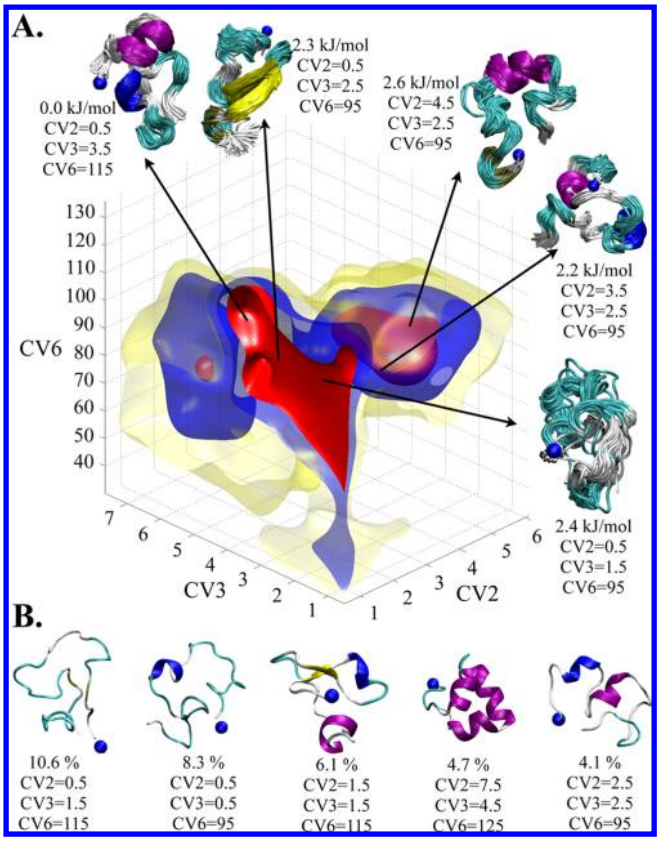



**Figure 4.** (A) The free energy landscape of hIAPP as a function of three collective variables—CV2: $\alpha$-RMSD of residues 8−22, CV3: $\alpha$-RMSD of residues 22−36, and CV6: coordination number of $C_\alpha$−$C_\alpha$ contacts. Isosurfaces are shown at 1 (red), 2 (blue), and 3 (yellow) kcal/mol. Low free-energy clusters from BEMD simulation are also shown with arrows; their weights are indicated as free energies (kJ/mol). The C-terminus of the peptide is indicated with a blue sphere. (B) Central structures of the five most populated backbone clusters from T-REMD simulation within a 0.3 nm cutoff distance. The population percentages and average values of collective variables are indicated below each structure.

## CONCLUSIONS

The exploration of the configuration space, while possible,[60] is not always feasible by brute-force MD. Here we investigate an IDP ensemble employing two enhanced sampling methods, T-REMD and BEMD. We perform a comparative analysis of all-atom explicit solvent simulations of hIAPP and describe the steps we took to perform T-REMD and BEMD simulations for appropriate sampling and analyze the respective data. Using the data, we determine the free energy landscape and secondary structures. Our results show that both methods yield highly consistent results, based on free energies projected along several different reaction coordinates (especially at lower free energy values) and on secondary structure analysis. Results from the two methods agree not only in the average secondary structure values of the entire peptide but also on the local per-residue level of secondary structure analysis. Though BEMD yields lower helical propensity than T-REMD, it is due to the selection of CVs, as we excluded residues 2−7 from $\alpha$-RMSD biasing. This finding, however, reveals the necessity of exercising great caution in the selection of CVs to obtain the most efficient sampling. We have also employed an unbiased (neutral) replica in BEMD that yields comparable free energy


DOI: 10.1021/acs.jctc.5b00047
J. Chem. Theory Comput. 2015, 11, 2776−2782

surfaces and secondary structures (see Supporting Information, Figures S2 and S4, respectively).

Fundamentally, both methods appear to be highly suitable for studying the structural properties of intrinsically disordered proteins. In particular, BEMD simulations show the capability to assess the free energy for larger regions of the conformational space, with less simulation time. However, an additional computational cost is associated with evaluating CVs and can be high depending on the complexity of the CV calculation. Initial benchmarking of the system to decide on appropriate CVs, as well as additional steps for data analysis, are also needed in BEMD. Additionally, T-REMD inherently provides temperature-dependent data on the system. Though we do not use temperature-dependent data in this work, such information can provide the means to investigate the temperature dependence of IDPs properties[61,62] or folding thermodynamics for globular proteins.[63,64] We note that it is feasible to extract temperature-dependent properties over a small temperature range from BEMD simulations performed at a single temperature using the approach described by Marinelli et al.[24] The suitability of BEMD for IDP sampling was also shown in a recent work by Do et al. for a 20 residue-long peptide[27] together with a comparison with a long unbiased MD simulation and well-tempered metadynamics. Here, we present comparison of BEMD results for a longer IDP (37 residues) with another well-established enhanced sampling method, T-REMD. In addition, we highlight the practical differences in both performing and analyzing simulations with these two popular methods and provide useful guidelines to investigate the properties of IDPs. We can suggest a similar set of CVs for IDPs of similar or even greater length, but depending on the length, regional division of secondary structure collective variables may need to be reconsidered. As mentioned in the introduction, enhanced sampling methods are not limited to T-REMD and BEMD. Promising methods include parallel tempering in the well-tempered ensemble[10,40,65−69] and solute tempering,[9,70−72] both of which can alleviate the computational expense of T-REMD and will also be worth studying.

## ■ ASSOCIATED CONTENT

### Supporting Information

Detailed equations describing collective variables and parameters for equations and metadynamics calculations, as well as ranges and bin sizes for the free energy reconstruction of BEMD and PMF calculations of T-REMD and the neutral replica of BEMD. The free energy projected on the $\alpha$-RMSD CV for the initial benchmark run of hIAPP and free energies projected on each collective variable and secondary structures obtained from T-REMD, reweighted BEMD, and the neutral replica of BEMD. This material is available free of charge via the Internet at http://pubs.acs.org/.The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00047.

## ■ AUTHOR INFORMATION


### Corresponding Authors

*E-mail: daniele.granata@gmail.com. (D.G.)
*E-mail: jeetain@lehigh.edu. (J.M.)


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

Use of the high-performance computing capabilities of the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation (NSF) grant no. TG-MCB-120014, is gratefully acknowledged.

## ■ REFERENCES


(1) Dunker, A. K.; et al. *J. Mol. Graphics Modell.* **2001**, *19*, 26−59.
(2) Fisher, C. K.; Stultz, C. M. *Curr. Opin. Struct. Biol.* **2011**, *21*, 426−431.
(3) van der Lee, R.; et al. *Chem. Rev.* **2014**, *114*, 6589−6631.
(4) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197−208.
(5) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *Annu. Rev. Biophys.* **2008**, *37*, 215−246.
(6) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(7) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042−6051.
(8) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(9) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13749−13754.
(10) Bonomi, M.; Parrinello, M. *Phys. Rev. Lett.* **2010**, *104*, 190601.
(11) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(12) Piana, S.; Laio, A. *J. Phys. Chem. B* **2007**, *111*, 4553−4559.
(13) Sgourakis, N. G.; Yan, Y.; McCallum, S. A.; Wang, C.; Garcia, A. E. *J. Mol. Biol.* **2007**, *368*, 1448−1457.
(14) Sgourakis, N. G.; Merced-Serrano, M.; Boutsidis, C.; Drineas, P.; Du, Z.; Wang, C.; Garcia, A. E. *J. Mol. Biol.* **2011**, *405*, 570−583.
(15) Yoon, J.; Park, J.; Jang, S.; Lee, K.; Shin, S. *J. Biomol. Struct. Dyn.* **2008**, *25*, 505−515.
(16) Wu, K.-P.; Weinstock, D. S.; Narayanan, C.; Levy, R. M.; Baum, J. *J. Mol. Biol.* **2009**, *391*, 784−796.
(17) Zhang, W.; Ganguly, D.; Chen, J. *PLoS Comput. Biol.* **2012**, *8*, e1002353.
(18) Knott, M.; Best, R. B. *PLoS Comput. Biol.* **2012**, *8*, e1002605.
(19) Mittal, J.; Yoo, T. H.; Georgiou, G.; Truskett, T. M. *J. Phys. Chem. B* **2012**, *117*, 118−124.
(20) Dupuis, N. F.; Wu, C.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2009**, *131*, 18283−18292.
(21) Reddy, A. S.; Wang, L.; Singh, S.; Ling, Y. L.; Buchanan, L.; Zanni, M. T.; Skinner, J. L.; de Pablo, J. J. *Biophys. J.* **2010**, *99*, 2208−2216.
(22) Wu, C.; Shea, J.-E. *PLoS Comput. Biol.* **2013**, *9*, e1003211.
(23) Miller, C.; Zerze, G. H.; Mittal, J. *J. Phys. Chem. B* **2013**, *117*, 16066−16075.
(24) Marinelli, F.; Pietrucci, F.; Laio, A.; Piana, S. *PLoS Comput. Biol.* **2009**, *5*, e1000452.
(25) Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 6817−6822.
(26) Baftizadeh, F.; Biarnes, X.; Pietrucci, F.; Affinito, F.; Laio, A. *J. Am. Chem. Soc.* **2012**, *134*, 3886−3894.
(27) Do, T. N.; Choy, W.-Y.; Karttunen, M. *J. Chem. Theory Comput.* **2014**, *10*, 5081−5094.
(28) Lorenzo, A.; Razzaboni, B.; Weir, G. C.; Yankner, B. A. *Nature* **1994**, *368*, 756−760.
(29) Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C. *Endocr. Rev.* **2008**, *29*, 303−316.
(30) Abedini, A.; Schmidt, A. M. *FEBS Lett.* **2013**, *587*, 1119−1127.
(31) Cao, P.; Marek, P.; Noor, H.; Patsalo, V.; Tu, L.-H.; Wang, H.; Abedini, A.; Raleigh, D. P. *FEBS Lett.* **2013**, *587*, 1106−1118.
(32) Laghaei, R.; Mousseau, N.; Wei, G. *J. Phys. Chem. B* **2010**, *114*, 7071−7077.
(33) Andrews, M. N.; Winter, R. *Biophys. Chem.* **2011**, *156*, 43−50.
(34) Murphy, R. D.; Conlon, J.; Mansoor, T.; Luca, S.; Vaiana, S. M.; Buchete, N.-V. *Biophys. Chem.* **2012**, *167*, 1−7.




DOI: 10.1021/acs.jctc.5b00047
*J. Chem. Theory Comput.* 2015, 11, 2776−2782

(35) Chiu, C.-c.; Singh, S.; de Pablo, J. J. *Biophys. J.* **2013**, *105*, 1227−1235.
(36) Best, R. B.; Mittal, J. *J. Phys. Chem. B* **2010**, *114*, 14916−14923.
(37) Abascal, J. L.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.
(38) Zerze, G. H.; Best, R. B.; Mittal, J. *Biophys. J.* **2014**, *107*, 1654−1660.
(39) Gianni, S.; Camilloni, C.; Giri, R.; Toto, A.; Bonetti, D.; Morrone, A.; Sormanni, P.; Brunori, M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 14141−14146.
(40) Palazzesi, F.; Prakash, M. K.; Bonomi, M.; Barducci, A. *J. Chem. Theory Comput.* **2015**, *11*, 2−7.
(41) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(42) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(43) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(44) Garcia, A. E.; Herce, H.; Paschek, D. *Annu. Rep. Comput. Chem.* **2006**, *2*, 83−95.
(45) Baftizadeh, F.; Cossio, P.; Pietrucci, F.; Laio, A. *Curr. Phys. Chem.* **2012**, *2*, 79−91.
(46) Biarnés, X.; Pietrucci, F.; Marinelli, F.; Laio, A. *Comput. Phys. Commun.* **2012**, *183*, 203−211.
(47) Pietrucci, F.; Laio, A. *J. Chem. Theory Comput.* **2009**, *5*, 2197−2201.
(48) Kjaergaard, M.; Nørholm, A.-B.; Hendus-Altenburger, R.; Pedersen, S. F.; Poulsen, F. M.; Kragelund, B. B. *Protein Sci.* **2010**, *19*, 1555−1564.
(49) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. *Proteins: Struct., Funct., Bioinf.* **2000**, *41*, 415−427.
(50) Nosé, S. *Mol. Phys.* **1984**, *52*, 255−268.
(51) Hoover, W. G. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695.
(52) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R.; Parrinello, M. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(53) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(54) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(55) Luca, S.; Yau, W.-M.; Leapman, R.; Tycko, R. *Biochemistry* **2007**, *46*, 13505−13522.
(56) Williamson, J. A.; Loria, J. P.; Miranker, A. D. *J. Mol. Biol.* **2009**, *393*, 383−396.
(57) Abedini, A.; Raleigh, D. P. *Phys. Biol.* **2009**, *6*, 015005.
(58) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.
(59) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.
(60) Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; Shaw, D. E. *J. Am. Chem. Soc.* **2012**, *134*, 3787−3791.
(61) Nettels, D.; Müller-Späth, S.; Küster, F.; Hofmann, H.; Haenni, D.; Rüegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 20740−20745.
(62) Wuttke, R.; Hofmann, H.; Nettels, D.; Borgia, M. B.; Mittal, J.; Best, R. B.; Schuler, B. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 5213−5218.
(63) Best, R. B.; Mittal, J. *J. Phys. Chem. B* **2010**, *114*, 8790−8798.
(64) English, C. A.; García, A. E. *Phys. Chem. Chem. Phys.* **2014**, *16*, 2748−2757.
(65) Deighan, M.; Bonomi, M.; Pfaendtner, J. *J. Chem. Theory Comput.* **2012**, *8*, 2189−2192.
(66) Barducci, A.; Bonomi, M.; Prakash, M. K.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E4708−E4713.
(67) Deighan, M.; Pfaendtner, J. *Langmuir* **2013**, *29*, 7999−8009.
(68) Sutto, L.; Gervasio, F. L. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 10616−10621.
(69) Papaleo, E.; Sutto, L.; Gervasio, F. L.; Lindorff-Larsen, K. *J. Chem. Theory Comput.* **2014**, *10*, 4169−4174.
(70) Camilloni, C.; Provasi, D.; Tiana, G.; Broglia, R. A. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 1647−1654.
(71) Wang, L.; Friesner, R. A.; Berne, B. *J. Phys. Chem. B* **2011**, *115*, 9431−9438.
(72) Bussi, G. *Mol. Phys.* **2014**, *112*, 379−384.


DOI: 10.1021/acs.jctc.5b00047
*J. Chem. Theory Comput.* 2015, 11, 2776−2782