Replica-Averaged Metadynamics

Carlo Camilloni,† Andrea Cavalli,‡,† and Michele Vendruscolo*,†

†Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW United Kingdom
‡Institute for Research in Biomedicine, 6500 Bellinzona, Switzerland

ABSTRACT: A statistical mechanics description of complex molecular systems involves the determination of ensembles of conformations that represent their Boltzmann distributions. The observable properties of these systems can be then predicted by calculating averages over such ensembles. In principle, given accurate energy functions and efficient sampling methods, these ensembles can be generated by molecular dynamics simulations. In practice, however, often the energy functions are known only approximately and the sampling can be carried out only in a limited manner. We describe here a method that enables to increase simultaneously both the quality of the energy functions and of the extent of the sampling in a system-dependent manner. The method is based on the incorporation of experimental data as replica-averaged structural restraints in molecular dynamics simulations and exploits the metadynamics framework to enhance the sampling. The application to the case of α-conotoxin SI, a 13-residue peptide that has been characterized extensively by experimental measurements, shows that the approach that we describe enables accurate free energy landscapes to be generated. The analysis of these landscapes indicates the presence of a low population state in equilibrium with the native state in which the only aromatic residue of α-conotoxin SI is exposed to the solvent, which is a feature that may predispose the peptide to interact with its partners.

INTRODUCTION

Molecular dynamics simulations represent a powerful approach to explore the behavior of molecular systems.1–4 There are, however, two main limitations in the use of this method to predict the behavior of systems of biological interest: the quality of the energy functions (force fields) and the extent of the sampling of the conformational space, which is associated with the achievable length of the simulations.1–4 These two issues are not independent, as in order to assess the quality of a given force field one needs to calculate statistical averages of specific properties of a system and compare them with those measured experimentally, for instance, by nuclear magnetic resonance (NMR) spectroscopy, small-angle X-ray scattering (SAXS), fluorescence resonance energy transfer (FRET), or cryo-electron microscopy (cryo-EM) methods.5–7 Both an accurate force field and an accurate sampling are therefore needed to establish whether the computations are consistent with the experiments.

In the case of proteins, in recent years, force fields have been improved systematically, both by changing their functional form and by choosing better values for the parameters.8–12 At the same time, to address the problem of increasing the extent of the sampling, major advances have been made to improve the software and the hardware used to run the simulations4,13–15 and to develop efficient advanced sampling techniques.1,2,16 In the latter case, a common feature is that by trading off some information about the dynamics one can sample the relevant conformations by enhancing the crossing of high-energy barriers. Among the most commonly used advanced sampling methods, we mention here those classified as collective variable based techniques (e.g., umbrella sampling17) or as extended ensemble techniques (e.g., replica-exchange).18 It has also been shown that the incorporation of experimental data as structural restraints in molecular simulations enables the structures and dynamics of proteins to be determined quite accurately.19–25 Furthermore, it has also been shown that if the restraints are averaged over several copies, or replicas, of a given protein, this approach generates ensembles of conformations compatible with the maximum entropy principle.26–28 These findings demonstrate that the incorporation of experimental data as replica-averaged structural restraints in molecular dynamics simulations provides an accurate representation of the unknown Boltzmann distribution of a system given an approximate force field and a set of experimental data.26–28 In practice, however, although the use of replica-averaged structural restraints enables one to improve the quality of the force field for the specific systems for which experimental data are available, it also makes it very challenging to sample the conformational space efficiently, because it increases the number of degrees of freedom proportionally to the number of replicas. In order to be generally applicable, therefore, this approach should be implemented together with enhanced sampling methods.

To address this problem, in this paper we introduce the replica-averaged metadynamics (RAM) method. In this
approach, experimental data are incorporated as replica-averaged structural restraints in molecular dynamics simulations, and the sampling is enhanced using the metadynamics method, in this case with replica exchange in the well-tempered ensemble (WTE) framework. This approach improves at the same time, and in a system-dependent manner, the quality of the force field by the direct addition of experimental information and the extent of the sampling by the use of metadynamics. In order to assess the quality of the distribution of structures obtained by the RAM method, we compare the results obtained with it with those obtained by sampling the unmodified force field through the use of the parallel-tempering metadynamics (PTMetaD) technique. In this sense, our goal here is not to compare the efficiency of the sampling of the RAM and PTKMetaD methods, but instead to identify the differences in the free energy landscapes associated with the incorporation of experimental information to generate replica-averaged structural restraints in the simulations. The PTKMetaD and RAM methods are represented schematically in Figure 1.

In order to illustrate the RAM method, we studied α-conotoxin SI, a neurotoxic 13-residue peptide with two disulfide bridges that strongly constraint the accessible conformational space of this molecule (Figure 2a). The small dimensions of this system, with a well-defined folded conformation and wide range of experimental data available (including a high-resolution X-ray structure, NMR chemical shifts, JHINHA-couplings, and NOE-derived distances), make it well suited for the purpose of the present work. Furthermore α-conotoxins bind selectively in a way that is not yet fully understood nicotinic acetylcholine receptors, which represent a diverse family of homo- or heteropentameric ligand-gated ion channels. A detailed characterization of the conformational space populated by this peptide is thus expected to be helpful for future studies of the mechanism of binding between α-conotoxins and its associated receptors.

### METHODS

The simulations described here were performed using the Amber03W force field with the TIP4P05 water model. All the simulations were run in GROMACS modified with PLUMED and Almost. A time step of 2 fs was used together with LINCS constraints. The van der Waals interactions were implemented with a cutoff at 0.9 nm, and long-range electrostatic effects were treated with the particle mesh Ewald method. All simulations were carried out in the canonical ensemble by keeping the volume fixed and by thermosetting the system with the Nose–Hoover thermostat. The starting conformation was taken from an available X-ray structure (PDB code 1HJE). This structure was protonated and solvated with 1224 water molecules in a dodecahedron box of 43 nm$^3$ of volume. The energy of the system was first minimized and then the temperature was increased to 283 K in two separate steps, in the first one a 50 ps simulation was performed by keeping fixed the heavy atoms of the protein, and successively a second 200 ps simulation was performed without restraints. The density of the system was relaxed by a 200 ps run using the Berendsen barostat.

In the RAM simulations, the experimental data are used to modify on the fly in a system-dependent manner the underlying force field by employing a replica averaging procedure over back-calculated experimental parameters, which are compared with their experimentally measured values. In this way, in the implementation of the RAM approach discussed here, the ensemble of structures resulting from the simulations at the lowest temperature represents the free energy of the force field modified, in this case, by NMR chemical shifts. The equilibrium values of the observables can be thus calculated directly on this ensemble without the use of a reweighting procedure. In the implementation of the RAM method described here, we augmented the force field using a term based on the replica averaging of the backbone NMR chemical shifts, where the chemical shifts themselves are back-calculated at each time step and averaged over four replicas of the system. The replica-averaged chemical shifts are restrained to be close to their experimental values using a quadratic potential. As a parallel tempering approach applied to the RAM case would require four times more replicas than in a standard parallel tempering, to reduce the number of replicas we employed the WTE technique, which is a metadynamics approach that uses the energy as the only collective variable. The six temperatures used were 283, 288, 318, 353, 394.2, and 443.1 K, so that also in this case 24 replicas were needed. Furthermore, the metadynamics bias was not used over the first four replicas, in such a way that the resulting ensemble was biased only by the chemical shifts. In this way an ensemble of structures that incorporates the information provided by the experimental chemical shifts in the sense of the maximum entropy principle can be easily obtained and analyzed. Alternatively, the RAM method could be implemented by using the metadynamics also on the first four replicas, adopting then a reweighting procedure to obtain the correct statistical weights of the conformations. The bias factor of the well-tempering was set to 10, and the exchanges were tried every 250 steps. Each replica was simulated for 60 ns, where the first 10 ns were discarded, and the convergence was tested by comparing two halves of the simulation (Figure 2b).

In the PTKMetaD simulations, the parallel tempering method, in which multiple copies of a system are simulated at different temperatures, is used to increase the acceptance rate between replicas at different temperatures, and the sampling is enhanced using the metadynamics technique, which is a metadynamics approach that uses the experimental information and the extent of the sampling by the use of metadynamics. In order to assess the quality of the distribution of structures obtained by the RAM method, we compare the results obtained with it with those obtained by sampling the unmodified force field through the use of the parallel-tempering metadynamics (PTMetaD) technique. In this sense, our goal here is not to compare the efficiency of the sampling of the RAM and PTKMetaD methods, but instead to identify the differences in the free energy landscapes associated with the incorporation of experimental information to generate replica-averaged structural restraints in the simulations. The PTKMetaD and RAM methods are represented schematically in Figure 1.

![Figure 1. Schematic representation of the PTKMetaD and RAM methods. In the PTKMetaD method, metadynamics, in the version of the RAM method discussed in this work, is used together with parallel tempering to enhance the sampling of the unmodified force field. By contrast, in the RAM method the experimental data are used as replica-averaged structural restraints to enhance on the fly in a system-dependent manner the quality of the force-field, while metadynamics is used to increase the acceptance rate between replicas at different temperature in the WTE approach to parallel tempering.](image-url)
temperatures and exchanges between neighboring simulations are performed with a Metropolis criterion, is coupled with the metadynamics method on a set of collective variables. The coupling of the two methods ensures that slow degrees of freedom not included in the collective variables are sampled correctly through parallel tempering. In the PTMetaD simulations that we carried out here, we used two collective variables. The first, called CamShift, is the root-mean-square difference between calculated and experimental (BMRB 4503) chemical shifts, where the calculated chemical shifts are obtained using the CamShift method; for this collective variable, we used a functional form described previously. The second collective variable, called Alpha-Beta, is a measure of the deviation from zero of the $\varphi$ and $\psi$ dihedral angles as implemented in PLUMED and defined as

$$AB = \frac{1}{4} \sum_{N_{res}} \left[ (1 + \cos(\varphi)) + (1 + \cos(\psi)) \right]$$

where the sum runs over all the $N_{res}$ residues of the protein. The temperatures, which were chosen using an algorithm proposed recently, were 283, 288, 293.1, 298.4, 303.8, 309.4, 315.1, 321.0, 327.0, 333.3, 339.7, 346.2, 353.0, 360.0, 367.2, 374.6, 382.2, 390.1, 398.2, 406.6, 415.2, 424.2, 433.4, and 442.9 K for a total of 24 replicas. The Gaussian functions in the metadynamics were deposited every 200 steps, with a height of 0.05 kJ/mol and a width of 5.0 and 0.1, respectively, for the CamShift and Alpha-Beta collective variables; a bias factor of 10 was used to rescale the Gaussian functions as in the well-tempered formulation of metadynamics. Exchanges were tried every 250 steps. In order to back-calculate collective variables and observables not included in the metadynamics simulations, we used a reweighting algorithm. Each individual replica was simulated for 60 ns, which we verified was sufficient to reach convergence in the free energy calculations (Figure 2c).

The exchange rules for the PTMetaD and the RAM methods follow those of the replica-exchange method. In the replica-exchange approach, multiple replicas of a system are simulated in parallel and exchanges between the replicas are tried at a given rate and accepted with a probability $p$ according to the Metropolis rule, $p = \min(1, \exp(-\Delta))$. In the PTMetaD method, independently from the collective variables employed, given two replicas $i$ and $j$ the factor $\Delta$ required to calculate the probability of exchange is

$$\Delta_{i}^{\text{PTMetaD}} = (\beta_{i} - \beta_{j})(E(X_{i}) - E(X_{j}))$$

$$- \beta_{i}(V(q(X_{i})) - V(q(X_{j}))) - \beta_{j}(V(q(X_{j})) - V(q(X_{i}))))$$

where $\beta$ is the inverse of the temperature times the Boltzmann constant, $E$ is the value of the force field for a configuration $X$, and $V(q(X))$ is the value of the bias resulting from the metadynamics rules on the collective variable $q$. In the version of the RAM method discussed here, at each temperature $\beta$, $M$ replicas (four in this case) are coupled together by a restraint imposed on the average value of multiple observables. Furthermore, all the replicas of each block $b$, except those at the reference temperature ($b = 0$), experience also a metadynamics bias on the energy. In principle, the exchange rule of two replicas belonging to two different blocks should include the back-calculation of the new average chemical shifts. A problem with this approach, however, is that the coupling of the replicas resulting from the chemical shifts reduces the exchange probability. Furthermore, one can only try one exchange between two blocks at different temperatures, thus resulting in a poor exchange rate. Alternatively, one can attempt to exchange all the replicas at two given temperatures. This second approach does not need the recalculation of the chemical shifts, and a move is accepted using the product of the four probabilities of exchange. In the present implementation, we used this second exchange move, and because of the enhancement of the probability of exchange resulting from the WTE approach, we obtained an average acceptance rate of the 10%. Concerning the factor $\Delta_{w}$ for two replicas belonging to different blocks, this factor is the same of PTMetaD with...
the difference that now $E(X)$ that is the energy of the force-field includes also the energy of the chemical shift restraints

$$E^{\alpha}_{c}(X_i) = \alpha_i \sum_{k=1}^{K} \sum_{l=1}^{L} \left( \delta_{kl}^{exp} - \frac{1}{M} \sum_{m=1}^{M} \delta_{kl}^{calc}(X_{im}) \right)^2$$

where the $\alpha_i$ are force constants, $k$ runs over the amino acids of the peptide or protein, $l$ runs over the six backbone atoms used in the simulations (C$\alpha$, C$\beta$, H$\alpha$, HN, and N), and $m$ runs over the $M$ replicas composing the $i$th block.

**RESULTS**

In order to assess the quality of the ensembles obtained by the RAM method, we applied it to determine the conformational fluctuations in the native state of $\alpha$-conotoxin SI, and compared the resulting free energy landscape with the corresponding one obtained for an unmodified force field using the PTMetaD method (Figure 3). Since in both cases we reached convergence, the differences in the free energy landscapes are a consequence of the additional restraining term. Both simulations explored a similar range of values for the collective variable Alpha-Beta (see Methods) of the backbone, while the PTMetaD explored a wider range of values of the collective variable CamShift (see Methods), which correspond to conformations with larger deviations from the experimental chemical shifts. In the PTMetaD simulations, the extension of the lower region of the free energy with respect to CamShift is larger than in the RAM case, indicating that structures with very different levels of agreement with the experimental data have similar statistical weights.

The quality of the two extensive samplings of the conformational space that we have described above has been assessed using different sets of experimental data. In the RAM ensemble, after discarding the first 10 ns, all the structures sampled by the four replicas at the lowest temperature were used for the analysis. In the PTMetaD ensemble all the
structures sampled by the lowest temperature replica have been used with a standard reweighing procedure.\textsuperscript{24} Chemical shifts were back-calculated using the SPARTA+ method,\textsuperscript{54} $J_{\text{HNHA}}$-couplings were back calculated using the Karplus equations with the “DFT2” set of parameters\textsuperscript{55} and NOE distances were back calculated by using a $1/r^6$ distance dependence.\textsuperscript{23} The average errors of the back-calculated data with respect to the experimental sets are reported in Figure 4 and Table 1. In all cases, the RAM ensemble was found to be in better agreement with the experimental data than the PTMetaD ensemble (Figure 4 and Table 1), thus indicating that the use of replica-averaged structural restraints can improve in a system-dependent manner the underlying force field. The comparison of the NMR structure\textsuperscript{33} (PDB code 1QMW, which was generated by using NOEs and $J_{\text{HNHA}}$-couplings) and the X-ray structure (PDB code 1HJE) shows that a single structure cannot fully reproduce all the available experimental data. For the NOE violations, the numbers in parentheses indicate the NOE violations larger than 0.05 nm.

Table 1. Assessment of the Quality of the Different Structural Ensembles of α-Conotoxin SI Described in This Work\textsuperscript{a}

<table>
<thead>
<tr>
<th>experiment</th>
<th>RAM</th>
<th>PTMetaD</th>
<th>RA</th>
<th>MD</th>
<th>NMR (1QMW)</th>
<th>X-ray (1HJE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Ca}$ (ppm)</td>
<td>0.69</td>
<td>0.98</td>
<td>0.72</td>
<td>0.97</td>
<td>0.93</td>
<td>0.72</td>
</tr>
<tr>
<td>$\text{C}_\beta$ (ppm)</td>
<td>0.52</td>
<td>0.88</td>
<td>0.59</td>
<td>0.90</td>
<td>1.42</td>
<td>1.15</td>
</tr>
<tr>
<td>$\text{H}_\alpha$ (ppm)</td>
<td>0.20</td>
<td>0.38</td>
<td>0.23</td>
<td>0.38</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>$\text{H}_\beta$ (ppm)</td>
<td>0.31</td>
<td>0.47</td>
<td>0.32</td>
<td>0.49</td>
<td>0.44</td>
<td>0.35</td>
</tr>
<tr>
<td>$J_{\text{HNHA}}$ (Hz)</td>
<td>1.30</td>
<td>1.90</td>
<td>1.51</td>
<td>2.10</td>
<td>1.30</td>
<td>1.60</td>
</tr>
<tr>
<td>NOE violations</td>
<td>20(10) /143</td>
<td>26(18) /143</td>
<td>23(12) /143</td>
<td>32(21) /143</td>
<td>17(2) /143</td>
<td>27(13) /143</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Average differences between the experimental values and those back-calculated from the different ensembles. The best agreement between experimental and calculated parameters is shown in bold. The RAM and PTMetaD ensembles are discussed in the main text, while the molecular dynamics (MD) ensemble is generated with a NVT 250 ns simulation; this simulation never leaves the native minimum of the PTMetaD ensemble. The higher quality of the PTMetaD ensemble with respect to the MD ensemble illustrates the importance of taking account of a broad ensemble of conformations explored by a protein through its equilibrium fluctuations when calculating NMR observables; this observation holds in particular for the $C_\beta$ chemical shifts, which are related to the averaging of the side chains. The higher quality of the RAM ensemble with respect to the PTMetaD ensemble shows how the use of the experimental data as replica-averaged structural restraints can improve in a system-dependent manner the underlying force field. The comparison of the NMR structure\textsuperscript{33} (PDB code 1QMW, which was generated by using NOEs and $J_{\text{HNHA}}$-couplings) and the X-ray structure (PDB code 1HJE) shows that a single structure cannot fully reproduce all the available experimental data. For the NOE violations, the numbers in parentheses indicate the NOE violations larger than 0.05 nm.

Figure 5. Comparison of the free energy landscapes of the RAM (panel a) and PTMetaD (panel b) ensembles as a function of the collective variable Alpha-Beta (see Methods) and the all-atom RMSD with respect to the X-ray structure (isolines at 2 kJ/mol). The ensembles span a similar range of values, but there are notable differences. The RAM ensemble shows a native minimum closer to the X-ray structure than that of the PTMetaD ensemble (the RMSD value is 0.1 nm for the RAM simulations while it is 0.2 nm for the PTMetaD simulations). The X-ray structure is shown in blue in the lower right corner of panel superposed to the central structure of most populated cluster (a). A second minimum is almost in the same place for both the ensembles, with a RMSD value of 0.3 nm and an Alpha-Beta value of 11. A third minimum at a RMSD value of 0.4 nm and an Alpha-Beta value of 13.5 are only present in the RAM ensemble, and a fourth minimum at a RMSD value of 0.5 nm is present in both the ensembles but with different Alpha-Beta values (14 for the RAM and 11 for the PTMetaD simulations).

the atoms with respect to an X-ray structure (1HJE, Figure 5). We found that in the PTMetaD ensemble the native minimum is farther from the X-ray structure (0.18 nm) than the RAM ensemble (0.11 nm). A second minimum is almost in the same place for both the ensembles, with a RMSD value of 0.3 nm and an Alpha-Beta value of 11. A third minimum at a RMSD value of 0.4 nm and an Alpha-Beta value of 13.5 is only present in the RAM ensemble, and a fourth minimum at a RMSD value of 0.5 nm is present in both the ensembles but with a different value of the Alpha-Beta collective variable (14 for the RAM and 11 for the PTMetaD ensembles).

The native state population has been estimated at 92\%,\textsuperscript{33} a value that corresponds quite accurately to the value of 90% that we obtain for the RAM ensemble by considering all the conformations in which Tyr11 is buried (first and second minima in Figure 5a). Representative conformations for these minima are shown in Figure 6a,b. Further analysis of the free energy landscapes reveals the presence of low population states.
Replica-Averaged simulations. Both the 250 ns unbiased molecular dynamics simulation (MD run) and four replica-averaged (RA, i.e. the RAM method without metadynamics) 60 ns long molecular dynamics simulations are in an overall worse agreement with all the available data than the corresponding simulations in which the sampling is enhanced (PTMetaD and RAM), thus showing the importance of enhanced sampling techniques in exploring high energy conformations and that of replica-averaging in correcting the force-field. RAM can address both the problems at the same time. By contrast, a previously determined NMR structure (PDB code 1QMW) was found to be in better agreement with NOEs and $^3J_{\text{HNHA}}$-couplings, which were the data employed to generate it, than the PTMetaD ensemble; the agreement with the chemical shifts is poorer for the 1QWM structure than the PTMetaD ensemble. The X-ray structure too is not in good agreement with the chemical shifts and the $^3J_{\text{HNHA}}$-couplings, despite having only a few NOE violations. These results show the importance of considering conformational averaging in the comparison with experimental measurements also in the case of a peptide with a well-defined folded conformation and suggest that the RAM method will be useful in the case of disordered system such as intrinsically disordered proteins.

### CONCLUSIONS

We have described the RAM method, which combines the use of experimental data as replica-averaged structural restraints in molecular dynamics simulations and the metadynamics scheme. This approach enables one to simultaneously enhance the sampling efficiency and improve the force field for specific systems, and thus represents an efficient tool to generate ensemble of structure representing the structure and the dynamics of proteins.

### AUTHOR INFORMATION

**Corresponding Author**

E-mail: mv245@cam.ac.uk

**Notes**

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

C.C. was supported by a Marie Curie Intra European Fellowship.

### REFERENCES


