



Efficient sampling of puckering states of monosaccharides through replica exchange with solute tempering and bond softening

Lingle Wang^{1,2,a)} and B. J. Berne^{2,a)}

¹*Schrödinger, Inc., 120 West 45th Street, New York, New York 10036, USA*

²*Department of Chemistry, Columbia University, 3000 Broadway, New York, New York 10027, USA*

(Received 31 January 2018; accepted 13 March 2018; published online 8 May 2018)

A molecular-level understanding of the structure, dynamics, and reactivity of carbohydrates is fundamental to the understanding of a range of key biological processes. The six-membered pyranose ring, a central component of biological monosaccharides and carbohydrates, has many different puckering conformations, and the conformational free energy landscape of these biologically important monosaccharides remains elusive. The puckering conformations of monosaccharides are separated by high energy barriers, which pose a great challenge for the complete sampling of these important conformations and accurate modeling of these systems. While metadynamics or umbrella sampling methods have been used to study the conformational space of monosaccharides, these methods might be difficult to generalize to other complex ring systems with more degrees of freedom. In this paper, we introduce a new enhanced sampling method for the rapid sampling over high energy barriers that combines our previously developed enhanced sampling method REST (replica exchange with solute tempering) with a bond softening (BOS) scheme that makes a chemical bond in the ring weaker as one ascends the replica ladder. We call this new method replica exchange with solute tempering and bond softening (REST/BOS). We demonstrate the superior sampling efficiency of REST/BOS over other commonly used enhanced sampling methods, including temperature replica exchange method and REST. The conformational free energy landscape of four biologically important monosaccharides, namely, α -glucose, β -glucose, β -mannose, and β -xylose, is studied using REST/BOS, and results are compared with previous experimental and theoretical studies. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5024389>

INTRODUCTION

The six-membered pyranose ring is a central component of the chemical structure of biological monosaccharides, including α -glucose, β -glucose, β -mannose, and β -xylose. These monosaccharides are the essential building blocks of oligosaccharides and polysaccharides, as well as the carbohydrate moieties of glycoconjugates including glycoproteins and glycolipids. The carbohydrates, which constitute the most abundant and diverse set of biological molecules on Earth, have specific roles for a vast array of biological functions including cell adhesion and recognition, modulation of growth factor receptors, immune defense, inflammation, and viral and parasitic infections.¹

The pyranose ring of these monosaccharides has 38 canonical puckering conformations classified according to the nomenclature described by Schwarz² and adopted by the International Union of Pure and Applied Chemistry (IUPAC).³ These different puckering conformations have been suggested to play an essential role in the hydrolysis of glycosidic bonds in carbohydrates by the enzyme glycoside hydrolases (or glycosidases),⁴ but the free energy

landscape of these puckering conformations has remained elusive. The different puckering conformations of these biological monosaccharides are separated by high energy barriers making the complete sampling of their conformational space and quantitative characterization of the free energy profile very challenging.⁵⁻⁷

Previous computational studies of monosaccharides have been focused on the mechanism of the transition between different anomers, the equilibrium populations of different anomers, and the potential energies of various conformations mainly in the gas phase.^{5,8,9} Metadynamics and umbrella sampling methods, using the two spherical pseudorotation coordinates introduced by Cremer and Pople¹⁰ as the collective variables, have also been used to study the conformational free energy landscape of a pyranose ring.^{6,9,11} However, these methods rely on the prior knowledge of the slow degrees of freedom of the systems and might be difficult to generalize for other complex ring systems with many more degrees of freedom. In this article, we introduce a new enhanced sampling method that can efficiently sample the different puckering conformations of any complex ring systems including the biological monosaccharides. This method combines the previous developed enhanced sampling method replica exchange with solute tempering (REST)^{12,13} with a bond softening (BOS)

^{a)}Lingle.wang@schrodinger.com and bb8@columbia.edu

scheme^{14,15} that makes a chemical bond in the ring weaker as one ascends the replica ladder, the combination of which effectively removes the energy barrier separating the different puckering conformations, leading to efficient sampling of these puckering conformations. We call this new method replica exchange with solute tempering and bond softening (REST/BOS).

We have applied REST/BOS to four biologically most important monosaccharides in water solution, namely, α -glucose, β -glucose, β -mannose, and β -xylose, and compared the results with other commonly employed enhanced sampling methods, including TREM (temperature replica exchange method) and REST. While these sugars are trapped in their initial puckering conformations using these currently available sampling methods, REST/BOS can efficiently sample the thermally accessible puckering conformations of these sugars. We also quantify the relative populations of the low energy puckering conformations of these four monosaccharides and discuss their implication to the mechanism of glycosidic bond hydrolysis.⁴ While the current study focuses on the sampling of biologically important monosaccharides, the method introduced here can be generally applied to the sampling of any complex ring structures.

METHODS

REST/BOS sampling method

The free energy landscape of biophysical systems is very complex with many local minima separated by energy barriers much higher than $k_B T$, leading to kinetic trapping for long periods of time and thus to quasi-ergodicity problems in the simulations. The complete sampling of the conformational space of these biophysical systems remains a grand challenge in computational biophysics. The temperature replica exchange method (TREM) is one of the most commonly used enhanced sampling methods designed to overcome the problem of quasi-ergodicity.¹⁶ TREM works by performing multiple replicas of the same system each at a slightly different physical temperature, and configurations are swapped between different replicas satisfying the detailed balance condition. Atoms move and cross energy barriers faster at higher temperatures, and the different configurations sampled at higher-level replicas can propagate into the lowest level replica corresponding to the physical temperature of interest, leading to improved sampling efficiency as compared to brute force molecular dynamics simulations. However, the number of replicas required for efficient sampling in TREM scales as the square root of the number of degrees of freedom of the whole system, and the poor scaling over system size often limits the applicability of TREM for large systems, particularly for systems in aqueous solution containing a large number of water molecules.¹³

To overcome this problem, we devised the method “Replica Exchange with Solute Tempering” (REST)¹² and a recently improved version REST2,¹³ both Hamiltonian replica exchange methods, in which only a local region of the system is effectively heated more and more as the replicas climb the replica ladder, while the rest of the system remains cold in

higher replicas. In this way, the number of the replicas required in REST is greatly reduced as compared to what is needed in TREM. In REST2, the total potential energy of the system is decomposed into three components, the potential energy from a localized “hot” region E_H , the potential energy from the “cold” region E_C , and the interaction energy between the two regions E_{HC} . All of the replicas are run at the same physical temperature T_0 , and the potential energy for replica m is scaled as follows:

$$E_m(X) = \frac{\beta_m}{\beta_0} E_H(X) + \sqrt{\frac{\beta_m}{\beta_0}} E_{HC}(X) + E_C(X).$$

Here, X represents the configuration of the whole system, $\beta_m = 1/k_B T_m$, $\beta_0 = 1/k_B T_0$, and their ratios are the scaling parameters in Hamiltonian replica exchange; T_0 is the temperature of interest and $T_m > T_0$. In this way, the acceptance ratio for the attempted configuration exchange between different replicas only depends on the potential energy from the “hot” region and the interaction energy between the “hot” region and “cold” region, but does not depend on the potential energy from the “cold” region, significantly reducing the number of replicas from what is required in TREM. In the applications of REST2 to study the puckering conformations of monosaccharides in aqueous solution, the “hot” region includes the whole monosaccharide, and all the water molecules are in the “cold” region. Other divisions might prove advantageous.

REST2 has been used extensively in many biomolecular simulations,^{17–19} particularly for protein-ligand binding simulations,^{20–30} where it was shown to yield a superior sampling efficiency over other sampling methods.²³ However, the different puckering conformations of monosaccharide rings are separated by energy barriers as high as ~ 20 kcal/mol, and the brute force application of REST2 on these systems gives poor sampling of the puckering conformations.

The high energy barriers separating the different puckering conformations of monosaccharides are mainly due to the restrictions imposed by cyclization. The free energy landscapes of corresponding linear molecules without the ring closure restrictions are much smoother. Therefore, we have designed a method that combines REST2 with a bond softening scheme^{14,15,29} that efficiently samples the different puckering conformations of complex ring structures. We call this new method replica exchange with solute tempering and bond softening (REST/BOS). In REST/BOS, in addition to the higher effective temperature of the solute in the “hot” region in REST, one bonded stretch interaction in the ring of interest is also softened (interactions involving this bond are scaled by a factor smaller than unity), significantly reducing the potential energy barriers separating the different puckering conformations. The different puckering conformations sampled in the higher-level replicas propagate into the lowest level replica through replica exchange consistent with detailed balance. Thus the lowest replica with no scaling represents the thermodynamics for the system of interest.

To be specific, in REST/BOS, all of the replicas are run at the same real temperature T_0 , and the potential energy for replica m is scaled as

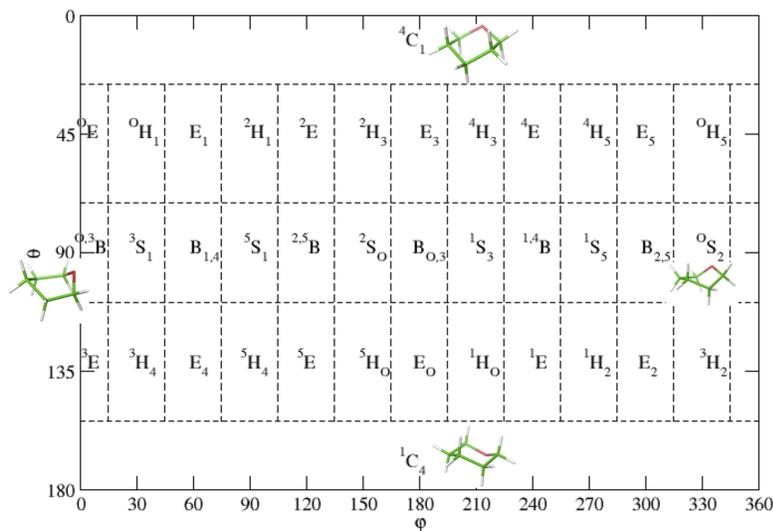


FIG. 1. The 38 canonical puckering conformations of the pyranose ring and the mapping of the puckering conformations to the 2-dimensional spherical pseudorotation coordinates.

$$E_m(X) = \frac{\beta_m}{\beta_0} \left[E'_H(X) + E_{bs}(\lambda_{sbs}^m, X) + E_{ba}(\lambda_{ba}^m, X) + E_{bd}(\lambda_{bd}^m, X) \right] + \sqrt{\frac{\beta_m}{\beta_0}} E_{HC}(X) + E_C(X).$$

Here, the potential energy of the solute in the “hot” region is further decomposed into two components, the potential energy involving a pre-selected bond in the ring of interest (which includes the bond stretch E_{bs} , bond angle E_{ba} , and bond dihedral angle E_{bd} interactions) and the other potential energy terms from the hot region of the solute E'_H . The potential energy terms involving the pre-selected bond in the ring of interest are scaled, respectively, by factors (λ_{sbs} , λ_{ba} , λ_{bd}) smaller than unity in the higher-level replicas, and these scaling factors are unity for the first replica corresponding to the physical state of interest just the same as REST2. The bond stretch interaction is scaled following the soft bond potential introduced in previous publications, which has been shown to be critical for overcoming singularity and numerical instability problems if the direct scaling of the harmonic bond stretch potential is used.^{14,15,31} To be specific, the soft bond stretch potential at coupling parameter λ_{sbs} is

$$E_{sbs}(\lambda_{sbs}, r) = \frac{\lambda_{sbs} k (r - r_0)^2}{1 + \alpha (1 - \lambda_{sbs}) (r - r_0)^2},$$

where k is the force constant and α is a positive constant called the “soft bond parameter.” When the coupling parameter is 1 (corresponding to the physical replica), the above soft bond potential recovers the regular harmonic bond stretch potential commonly used in molecular mechanics force fields. In the applications of REST/BOS to the pyranose ring in this study, the bond between C_5 and O is softened which allows the bond to fluctuate between the broken and open configurations thereby allowing the ring to pucker much more frequently in the higher replicas. The soft bond parameter α controls the effective range of distances of the soft bond sampled in higher-level replicas; the larger the α , the shorter the range of sampled distances, and the smaller the α ,

the longer the range of sampled distances. A detailed discussion about the soft bond parameter is presented in a previous publication.²⁹ In this study, the soft bond parameter is set to 2, the same as what is used in core-hopping free energy perturbation (FEP),¹⁴ which works well for all the systems tested in this study. Other values of the soft bond parameter might prove more advantageous and can be tuned for different systems.

Puckering conformations of pyranose ring

The generalized pseudorotation coordinates as introduced by Cremer and Pople¹⁰ are used to identify puckering conformations of these four monosaccharides. In the case of the pyranose ring shared by these four monosaccharides, the three pseudorotation coordinates can be converted to a set of three spherical coordinates (Q , θ , φ), and the set of spherical coordinates uniquely defines the puckering conformation for each ring configuration. The coordinate Q , the so-called puckering amplitude, does not fluctuate too much for all the accessible conformations, and thus only the two angular coordinates (θ , φ) are needed to define the puckering conformations.^{6,9,10} The conversion from the (θ , φ) coordinates to the IUPAC

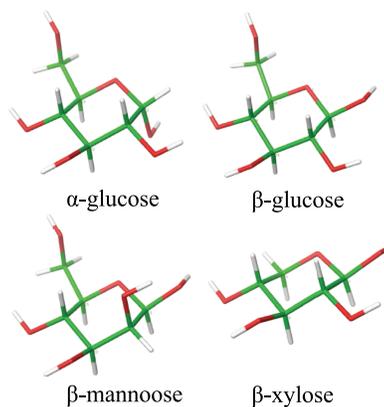


FIG. 2. The structures of the four most important monosaccharides studied in the paper.

TABLE I. The parameters of the potential energies for the 4 replicas used in REST/BOS simulations. The soft bond parameter is set to 2.

| Replica number | Effective temperature (K) | λ_{bs} | λ_{ba} | λ_{bd} |
|----------------|---------------------------|----------------|----------------|----------------|
| 1 | 300 | 1 | 1 | 1 |
| 2 | 420 | 0.4 | 0.4 | 0.4 |
| 3 | 588 | 0.15 | 0.15 | 0.15 |
| 4 | 823 | 0.03 | 0.01 | 0.01 |

defined canonical puckering states is shown in Fig. 1. The (θ, φ) plane has been partitioned into 38 regions each corresponding to a IUPAC defined canonical puckering conformation:

$\theta < 26^\circ$ is associated with 4C_1 ; $\theta > 154^\circ$ is associated with 1C_4 ; $26^\circ < \theta < 71^\circ$ and $109^\circ < \theta < 154^\circ$ regions are further divided into 24 sub-regions each with a 30° span in φ corresponding to either a half boat (H) or an envelope (E) conformation; the $71^\circ < \theta < 109^\circ$ region is also further divided into 12 sub-regions each corresponding to a boat (B) or skewed (S) conformation.

Detailed description of the simulations

All the simulations are run with DESMOND.³² The four monosaccharides, α -glucose, β -glucose, β -mannose, and β -xylose (shown in Fig. 2), in their 4C_1 conformations are built in Maestro and solvated in a water box with a 10 \AA buffer width,

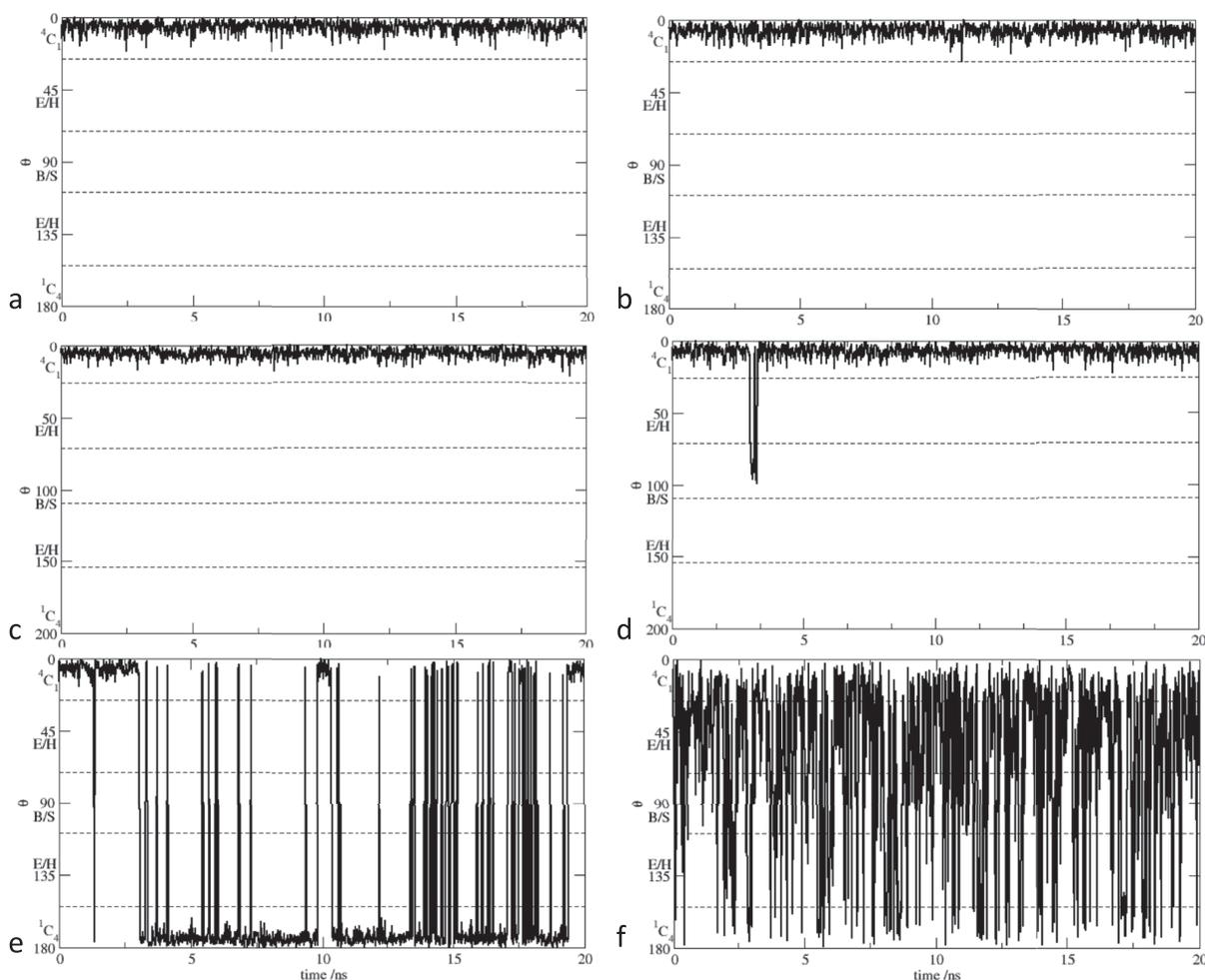


FIG. 3. The time series of puckering conformations of α -glucose sampled using TREM, REST, and REST/BOS. Upper plots (a) and (b): The θ pseudorotation coordinate as a function of time for the trajectories sampled using TREM. Plot (a) corresponds to the lowest level replica at room temperature ($T = 300 \text{ K}$), and plot (b) corresponds to the highest level replica at $T = 372 \text{ K}$. Starting from the 4C_1 puckering conformation, α -glucose was trapped in the same starting puckering conformation during the entire 20 ns simulations, even for the highest level replica. Middle plots (c) and (d): The θ pseudorotation coordinate as a function of time for the trajectories sampled using REST. Plot (c) corresponds to the lowest level replica at room temperature ($T = 300 \text{ K}$), and plot (d) corresponds to the highest level replica with the effective temperature of the solute at $T = 602 \text{ K}$. Starting from the 4C_1 puckering conformation, the highest level replica only visited the B/S conformation for one time during the entire 20 ns simulations, and only the same 4C_1 puckering conformation was sampled at the lowest level replica corresponding to the physical state. Lower plots (e) and (f): The θ pseudorotation coordinate as a function of time for the trajectories sampled using REST/BOS. Plot (e) corresponds to the lowest level replica in the physical state, and plot (f) corresponds to the highest level replica with the effective temperature of the solute increased to $T = 823 \text{ K}$ and one bond in the pyranose ring partially broken. Starting from the 4C_1 puckering conformation, the highest level replica samples many different puckering conformations during the 20 ns simulations, and the lowest level replica corresponding to the physical state also samples the 1C_4 puckering conformation in addition to the 4C_1 conformation many times.

which constitute the systems for the following simulations. The Optimized Potential for Liquid Simulations 3 (OPLS3)³³ force field is used for the sugars, and the SPC³⁴ model is used for water. For α -glucose, we ran three sets of simulations, using TREM, REST, and REST/BOS, to compare their sampling efficiency. The TREM simulation included 12 replicas with temperatures spanning between 300 K and 372 K following the exponential distribution, and each replica was run for 20 ns. The REST simulation included 4 replicas with the effective temperature of the solute ranging between 300 K and 602 K, and the temperature profile is determined by setting the expectation value of the acceptance ratio between neighboring replicas to be 0.3 following previous publications.^{27,35} Each replica in the REST simulation also lasted 20 ns, and thus the entire REST simulation only took one third of the computational time in TREM. The REST/BOS simulation also included 4 replicas with the parameters of the potential energy for each replica listed in Table I.

In REST/BOS, the additional modification of the Hamiltonian for bond softening leads to a slightly lower average

acceptance ratio (~ 0.2) than that in REST (~ 0.3). The parameters for REST/BOS simulations as listed in Table I might be further optimized to give better sampling efficiency. The soft bond parameter is set to be 2 as discussed in the section titled Methods. Same as the REST simulation, the REST/BOS simulation only took one third of computational time in TREM. After confirming the dramatically improved sampling efficiency of REST/BOS as compared to TREM and REST, a production REST/BOS simulation for each of the four systems was performed with each replica lasting 100 ns. The trajectories were saved every 20 ps, and the last 95 ns of the long REST/BOS simulations from the first replica (physical state) were used to collect the statistics of the equilibration populations of the thermally accessible puckering conformations. The free energies of the other puckering conformations as compared to the 4C_1 conformation are calculated from the observed relative populations, and errors are calculated from block average (5 blocks each with 19 ns data, and standard deviation among the 5 estimates are reported as the errors).

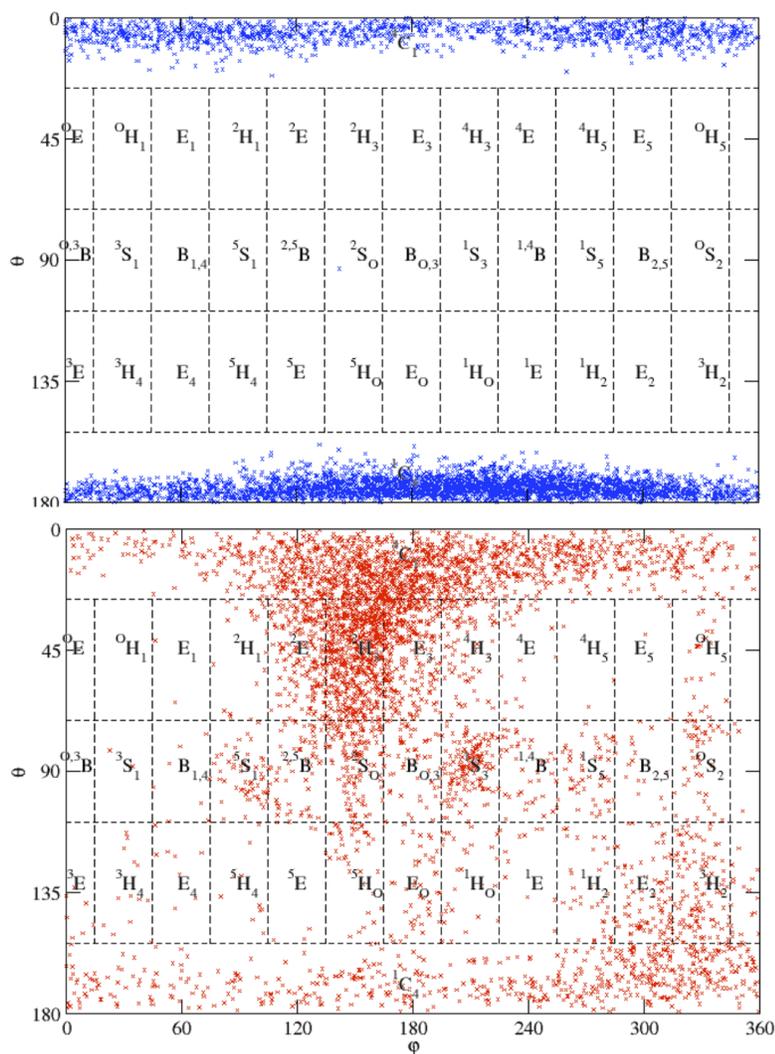


FIG. 4. The conformations of the α -glucose sampled in the production REST/BOS simulations as projected to the 2D spherical coordinates. Top panel is from the lowest replica corresponding to the physical state, and bottom panel is from the highest replica with a higher effective temperature on the solute and one bond softened. The highest replica explores all the puckering conformations, while only a few thermally accessible conformations are sampled in the physical state. In addition to the two chair conformations, the 2S_0 conformation is calculated to be the next lowest free energy conformation.

RESULTS

Comparison of the sampling efficiency among TREM, REST, and REST/BOS

Previous studies indicated that the two chair conformations are the dominant conformations of a pyranose ring. As an initial step, we first compared whether these three methods can efficiently sample the two chair conformations. The two chair conformations are located in different regions along the φ coordinate ($\theta < 26^\circ$ for 4C_1 and $\theta > 154^\circ$ for 1C_4). For α -glucose, the φ coordinate as a function of simulation time for the lowest (corresponding to the physical state) and highest replicas from TREM, REST, and REST/BOS simulations is shown in Fig. 3. Using TREM, the α -glucose stayed in the 4C_1 conformation during the entire 20 ns simulation, even for the highest replica at the highest temperature of 372 K. Similarly, using REST, the α -glucose stayed in the 4C_1 conformation during the entire 20 ns simulation in the physical replica. The highest replica only visited the 2S_0 conformation (the spike with the θ value of about 90° at ~ 3 ns) for one time for a very short period of time, and that conformation did not

propagate into the lowest replica. By comparison, using REST/BOS, starting from the 4C_1 conformation, the lowest replica visited the 1C_4 conformation many times during the 20 ns simulation, and the highest replica very efficiently explored many different conformations covering the whole conformational space, including the two chair conformations, many different boat and skewed conformations, and some half boat and envelope conformations. These results clearly demonstrate the superior sampling efficiency of REST/BOS as compared to TREM and REST.

Conformational free energy landscape of the four monosaccharides

Next, we performed a long REST/BOS simulation with 100 ns for each replica and collected the statistics of the equilibration populations of the thermally accessible puckering conformations for each of the four monosaccharides. The conformations of the pyranose ring sampled in the production REST/BOS simulations as projected to the 2D spherical coordinates are shown in Figs. 4–7 for α -glucose, β -mannose,

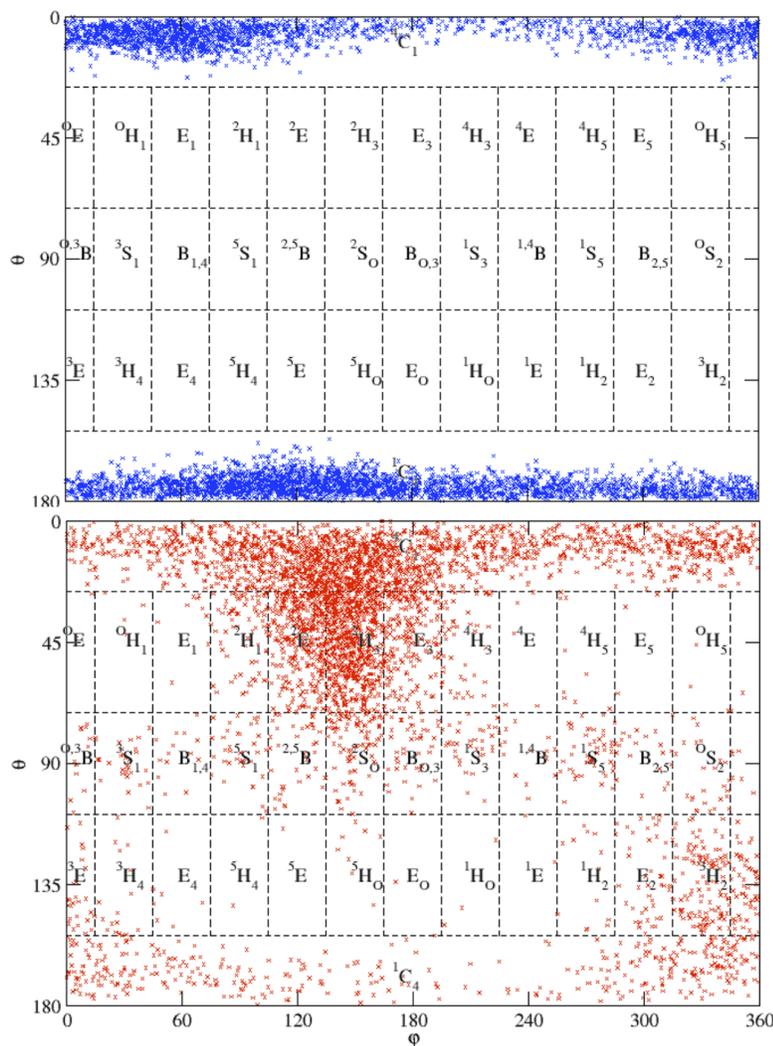


FIG. 5. The conformations of the β -mannose sampled in the production REST/BOS simulations as projected to the 2D spherical coordinates. Top panel is from the lowest replica corresponding to the physical state, and bottom panel is from the highest replica with a higher effective temperature on the solute and one bond softened. The highest replica explores all the puckering conformations, while only the two thermally accessible chair conformations are sampled in the physical state.

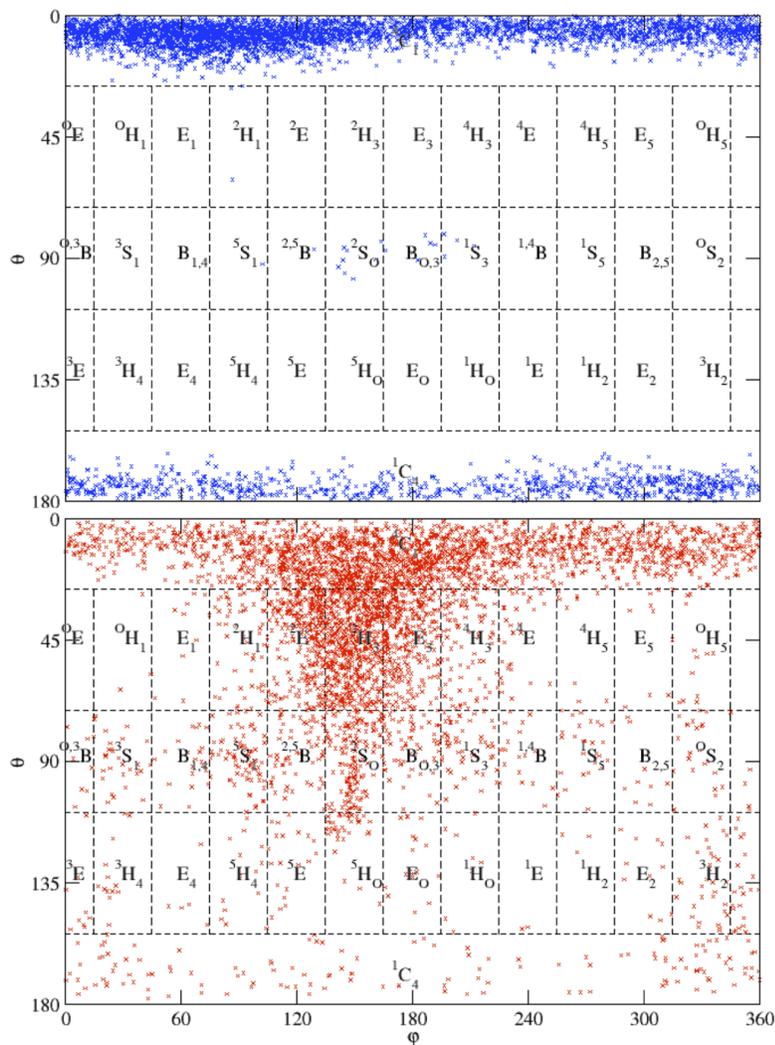


FIG. 6. The conformations of the β -xylose sampled in the production REST/BOS simulations as projected to the 2D spherical coordinates. Top panel is from the lowest replica corresponding to the physical state, and bottom panel is from the highest replica with a higher effective temperature on the solute and one bond softened. The highest replica explores all the puckering conformations, while only a few thermally accessible conformations are sampled in the physical state. In addition to the two chair conformations, the 2S_0 and $B_{O,3}/{}^1S_3$ conformations are calculated to be the next lowest free energy conformations.

β -xylose, and β -glucose, respectively. The top panels are from the lowest replicas (physical states), and the bottom panels are from the highest replicas. The highest replicas visited all conformations many times for all systems, but the lowest replicas only visited a few conformations (mainly the two chair conformations). The diffusion rates in the replica space are fast in these REST/BOS simulations, and there are many round trips for each replica. Therefore, the lowest replica has many opportunities to sample the other conformations. These other conformations sampled in higher replicas but did not successfully propagate into the lowest replica because they are located on very high energy regions for the physical states.

For all four monosaccharides, the two chair conformations, 4C_1 and 1C_4 , dominated more than 99% of the sampled frames, in agreement with previous experimental and computational studies of these systems.^{6,7} The relative populations of 4C_1 and 1C_4 , however, differ significantly among these four monosaccharides. In particular, from α -glucose, to β -mannose, to β -xylose, and to β -glucose, the populations of 4C_1 conformation increase from 27%, to 41%, to 82%,

and to 91%, respectively. This trend is in agreement with experimental data suggesting a higher population of the 4C_1 conformation in β -glucose than that in α -glucose.⁶ The calculated relative free energies of the 1C_4 conformation as compared to the 4C_1 conformation for these four systems are summarized in Table II. The 1C_4 conformation is calculated to have slightly favorable free energy than the 4C_1 conformation for α -glucose and β -mannose. However, the magnitude of the free energy difference is very small, and it does not significantly deviate from 0. The 4C_1 conformations for β -xylose and β -glucose are calculated to have much lower free energy than the 1C_4 conformation in agreement with experimental data.^{6,7} The free energy levels for different conformations for these four monosaccharides are summarized in Table II.

In addition to the two chair conformations, a few other conformations are also sampled for α -glucose, β -xylose, and β -glucose. For β -xylose, 2S_0 and $B_{O,3}/{}^1S_3$ are the next two lowest free energy conformations following the two chair conformations, each sampled with $\sim 0.2\%$ population, corresponding to an estimated free energy of 3.5 kcal/mol as compared to

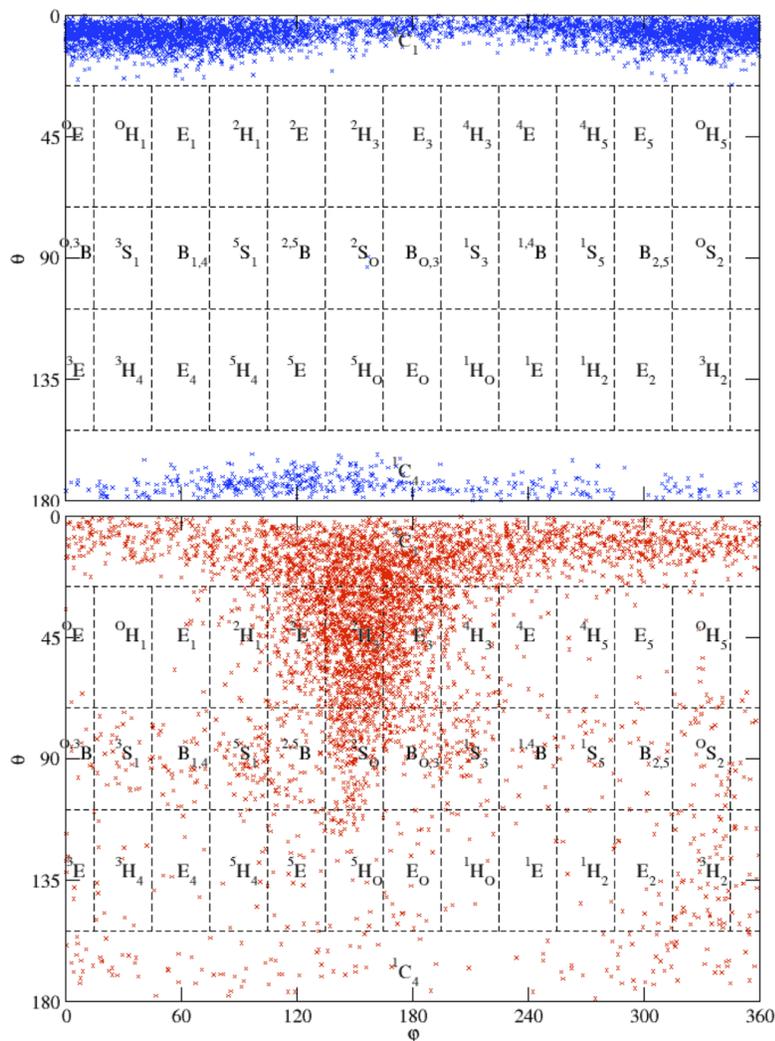


FIG. 7. The conformations of the β -glucose sampled in the production REST/BOS simulations as projected to the 2D spherical coordinates. Top panel is from the lowest replica corresponding to the physical state, and bottom panel is from the highest replica with a higher effective temperature on the solute and one bond softened. The highest replica explores all the puckering conformations, while only a few thermally accessible conformations are sampled in the physical state. In addition to the two chair conformations, the 2S_0 conformation is calculated to be the next lowest free energy conformations.

the lowest free energy 4C_1 conformation. In addition, the 2H_1 and 5S_1 are also sampled in one frame, corresponding to an estimated free energy greater than 3.5 kcal/mol for these conformations. The other conformations are not sampled, indicating that their free energies are much higher than 3.5 kcal/mol. For α -glucose, in addition to the two chair conformations, the 2S_0 conformation is also sampled in one frame, indicating that the 2S_0 conformation might be the next lowest free energy state. The estimated free energy of the 2S_0 conformation as

compared to the 4C_1 conformation is higher than 3.5 kcal/mol. For β -glucose, the 2S_0 conformation is also sampled in two frames, indicating that the 2S_0 conformation is the next lowest free energy conformation with an estimated free energy greater than 3.5 kcal/mol. This is in agreement with previous gas phase metadynamics studies and the experimental data that the majority of crystal structures of β -glycoside hydrolase with the β -glucose derivative bound have the pyranose ring in the 2S_0 conformation.⁹ For β -mannose, only the two chair

TABLE II. The populations of the 4C_1 conformations and the relative free energies (in kcal/mol) of the 1C_4 conformations (compared to the 4C_1 conformation) for the four monosaccharides as determined by the REST/BOS simulations. The next lowest free energy conformation(s) is also shown for each monosaccharide.

| Monosaccharides | 4C_1 population | 1C_4 population | 1C_4 free energy | Next lowest free energy conformation(s) |
|-------------------|----------------------|----------------------|-----------------------|---|
| α -glucose | 0.27 | 0.73 | -0.64 ± 0.38 | 2S_0 |
| β -mannose | 0.41 | 0.58 | -0.23 ± 0.50 | NA |
| β -xylose | 0.82 | 0.18 | 0.96 ± 0.35 | ${}^2S_0, B_{0.3}/{}^1S_3$ |
| β -glucose | 0.91 | 0.09 | 1.55 ± 0.71 | 2S_0 |

conformations are observed in the physical state, indicating that the other conformations have much higher free energies than these two chair conformations.

DISCUSSIONS AND CONCLUSIONS

The main result of this paper is the superior sampling efficiency of the REST/BOS method as compared to currently available enhanced sampling methods, including the most commonly employed TREM and REST methods. While the pyranose ring stays trapped in the initial puckering conformation using TREM and REST methods, very rapid exploration of the whole puckering conformational space is obtained using REST/BOS. Previously, others have used metadynamics or umbrella sampling methods with two spherical pseudorotation coordinates as the collective variables to study the conformational free energy landscape of pyranose, but these methods might be difficult to generalize to other more complex ring systems where more pseudorotation coordinates are needed to describe the whole conformational space. The REST/BOS method introduced here does not require prior knowledge of the slow degrees of freedom of the systems and can be easily applied to any complex ring systems. Softening any bond in the ring of interest would probably be sufficient to achieve efficient sampling.

Applications of the REST/BOS method to the four biologically most important monosaccharides indicated that the two chair conformations are the most dominant conformations for these systems, in agreement with previous experimental and theoretical studies. In addition, we found that the relative populations of the two chair conformations are highly dependent on the exocyclic groups on the pyranose ring, also in agreement with previous studies. In addition to the two stable chair conformations, we also observed a few other conformations that are thermally relevant, namely, the 2S_0 conformation for α -glucose and β -glucose and the 2S_0 , $B_{0,3}/1S_3$, 2H_1 , and 5S_1 conformations for β -xylose. The 2S_0 conformation for β -glucose was also calculated to be a low energy metastable state in a previous gas phase Car-Parrinello molecular dynamics (CPMD) study,⁹ in agreement with experimental crystal structures of β -glycoside hydrolase and β -glucose derivative complexes.

We anticipate that the efficient enhanced sampling method REST/BOS presented in this paper will have potential applications in many different areas. For example, accurate characterization of the conformational free energies of carbohydrates and other complex ring systems has proven to be very difficult, and it is impossible to assess the accuracy of a given set of force field parameters without a thorough sampling of the complete conformational space. The REST/BOS method introduced here thus can be used for the parameterization of accurate force fields for complex rings. In addition, macrocycles have been emerging as a very important drug class in the past few decades due to their expanded chemical diversity benefiting from advances in synthetic methods and their unique ability to balance various drug properties, including potency, selectivity, metabolic stability, and bioavailability. However, the complex conformational space of macrocycles poses a great challenge in the computational modeling of macrocycles.

The REST/BOS method can be used for efficient macrocycle conformational sampling. Furthermore, drug-like molecules in general contain rings with complex puckering conformations, and the REST/BOS method can be combined with existing tools for accurate modeling of drug-like compounds. It should be noted that we only have tested one particular bond-softening potential introduced in previous publications with the soft bond parameter set to 2 and expect that other values of the soft bond parameter or other forms of the bond-softening potentials may prove useful.

ACKNOWLEDGMENTS

B.J.B. is a consultant to Schrodinger, Inc. and is on its Scientific Advisory Board.

- ¹S. Hurtle, R. Service, and P. Szuroni, "Cinderella's coach is ready," *Science* **291**, 2337 (2001).
- ²J. C. P. Schwarz, "Rules for conformation nomenclature for five- and six-membered rings in monosaccharides and their derivatives," *J. Chem. Soc., Chem. Commun.* **14**, 505–508 (1973).
- ³IUPAC-IUBMB Joint Commission on Biochemical Nomenclature, *Eur. J. Biochem.* **111**, 295–299 (1980).
- ⁴V. A. Money, N. L. Smith, A. Scaffidi, R. V. Stick, H. J. Gilbert, and G. J. Davies, "Substrate distortion by a lichenase highlights the different conformational itineraries harnessed by related glycoside hydrolases," *Angew. Chem., Int. Ed.* **45**, 5136–5140 (2006).
- ⁵H. B. Mayes, L. J. Broadbelt, and G. T. Beckham, "How sugars pucker: Electronic structure calculations map the kinetic landscape of five biologically paramount monosaccharides and their implications for enzymatic catalysis," *J. Am. Chem. Soc.* **136**, 1008–1022 (2014).
- ⁶E. Autieri, M. Sega, F. Pederiva, and G. Guella, "Puckering free energy of pyranoses: A NMR and metadynamics-umbrella sampling investigation," *J. Chem. Phys.* **133**, 095104 (2010).
- ⁷J. R. Snyder and A. S. Serianni, "D-idose: A one- and two-dimensional NMR investigation of solution composition and conformation," *J. Org. Chem.* **51**, 2694–2702 (1986).
- ⁸S. Ha, J. Gao, B. Tidor, J. W. Brady, and M. Karplus, "Solvent effect on the anomeric equilibrium in D-glucose: A free energy simulation analysis," *J. Am. Chem. Soc.* **113**, 1553–1557 (1991).
- ⁹X. Biarnés, A. Ardèvol, A. Planas, C. Rovira, A. Laio, and M. Parrinello, "The conformational free energy landscape of β -d-glucopyranose. Implications for substrate preactivation in β -glucoside hydrolases," *J. Am. Chem. Soc.* **129**, 10686–10693 (2007).
- ¹⁰D. Cremer and J. A. Pople, "General definition of ring puckering coordinates," *J. Am. Chem. Soc.* **97**, 1354–1358 (1975).
- ¹¹M. A. Wilson and D. Chandler, "Molecular dynamics study of cyclohexane interconversion," *Chem. Phys.* **149**, 11–20 (1990).
- ¹²P. Liu, B. Kim, R. A. Friesner, and B. J. Berne, "Replica exchange with solute tempering: A method for sampling biological systems in explicit water," *Proc. Nat. Acad. Sci. U. S. A.* **102**, 13749–13754 (2005).
- ¹³L. Wang, R. A. Friesner, and B. J. Berne, "Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2)," *J. Phys. Chem. B* **115**, 9431–9438 (2011).
- ¹⁴L. Wang, Y. Deng, Y. Wu, B. Kim, D. N. LeBard, D. Wandschneider, M. Beachy, R. A. Friesner, and R. Abel, "Accurate modeling of scaffold hopping transformations in drug discovery," *J. Chem. Theory Comput.* **13**, 42–54 (2017).
- ¹⁵R. Abel and L. Wang, "Methods and systems for calculating free energy differences using a modified bond stretch potential," U.S. patent 14138186 (June 25, 2015).
- ¹⁶R. H. Swendsen and J.-S. Wang, "Replica Monte Carlo simulation of spin-glasses," *Phys. Rev. Lett.* **57**, 2607–2609 (1986).
- ¹⁷D. J. Cole, J. Tirado-Rives, and W. L. Jorgensen, "Enhanced Monte Carlo sampling through replica exchange with solute tempering," *J. Chem. Theory Comput.* **10**, 565–571 (2014).
- ¹⁸S. Jo and W. Jiang, "A generic implementation of replica exchange with solute tempering (REST2) algorithm in NAMD for complex biophysical simulations," *Comput. Phys. Commun.* **197**, 304–311 (2015).

- ¹⁹G. Bussi, "Hamiltonian replica exchange in GROMACS: A flexible implementation," *Mol. Phys.* **112**, 379–384 (2014).
- ²⁰L. Wang, B. J. Berne, and R. A. Friesner, "On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities," *Proc. Natl. Acad. Sci. U. S. A.* **109**, 1937–1942 (2012).
- ²¹L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood, D. L. Romero, C. Mase, J. L. Knight, T. Steinbrecher, T. Beuming, W. Damm, E. Harder, W. Sherman, M. Brewer, R. Wester, M. Murcko, L. Frye, R. Farid, T. Lin, D. L. Mobley, W. L. Jorgensen, B. J. Berne, R. A. Friesner, and R. Abel, "Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field," *J. Am. Chem. Soc.* **137**, 2695–2703 (2015).
- ²²D. A. Goldfeld, R. Murphy, B. Kim, L. Wang, T. Beuming, R. Abel, and R. A. Friesner, "Docking and free energy perturbation studies of ligand binding in the kappa opioid receptor," *J. Phys. Chem. B* **119**, 824–835 (2015).
- ²³R. Abel, L. Wang, E. D. Harder, B. J. Berne, and R. A. Friesner, "Advancing drug discovery through enhanced free energy calculations," *Acc. Chem. Res.* **50**, 1625–1632 (2017).
- ²⁴B. Kuhn, M. Tichý, L. Wang, S. Robinson, R. E. Martin, A. Kuglstatler, J. Benz, M. Giroud, T. Schirmeister, R. Abel, F. Diederich, and J. Hert, "Prospective evaluation of free energy calculations for the prioritization of cathepsin L inhibitors," *J. Med. Chem.* **60**, 2485–2497 (2017).
- ²⁵E. B. Lenselink, J. Louvel, A. F. Forti, J. P. D. van Veldhoven, H. de Vries, T. Mulder-Krieger, F. M. McRobb, A. Negri, J. Goose, R. Abel, H. W. T. van Vlijmen, L. Wang, E. Harder, W. Sherman, A. P. Ijzerman, and T. Beuming, "Predicting binding affinities for GPCR ligands using free-energy perturbation," *ACS Omega* **1**, 293–304 (2016).
- ²⁶J. Mondal, P. Tiwary, and B. J. Berne, "How a kinase inhibitor withstands gatekeeper residue mutations," *J. Am. Chem. Soc.* **138**, 4608–4615 (2016).
- ²⁷L. Wang, Y. Deng, J. L. Knight, Y. Wu, B. Kim, W. Sherman, J. C. Shelley, T. Lin, and R. Abel, "Modeling local structural rearrangements using FEP/REST: Application to relative binding affinity predictions of CDK2 inhibitors," *J. Chem. Theory Comput.* **9**, 1282–1293 (2013).
- ²⁸J. W. Kaus, E. Harder, T. Lin, R. Abel, J. A. McCammon, and L. Wang, "How to deal with multiple binding poses in alchemical relative protein-ligand binding free energy calculations," *J. Chem. Theory Comput.* **11**, 2670–2679 (2015).
- ²⁹H. S. Yu, Y. Deng, Y. Wu, D. Sindhikara, A. R. Rask, T. Kimura, R. Abel, and L. Wang, "Accurate and reliable prediction of the binding affinities of macrocycles to their protein targets," *J. Chem. Theory Comput.* **13**, 6290–6300 (2017).
- ³⁰N. M. Lim, L. Wang, R. Abel, and D. L. Mobley, "Sensitivity in binding free energies due to protein reorganization," *J. Chem. Theory Comput.* **12**, 4620–4631 (2016).
- ³¹S. Liu, L. Wang, and D. L. Mobley, "Is ring breaking feasible in relative binding free energy calculations?," *J. Chem. Inf. Model.* **55**, 727–735 (2015).
- ³²K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, and D. E. Shaw, "Scalable algorithms for molecular dynamics simulations on commodity clusters," in *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing* (ACM, Tampa, Florida, 2006), p. 84.
- ³³E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel, and R. A. Friesner, "OPLS3: A force field providing broad coverage of drug-like small molecules and proteins," *J. Chem. Theory Comput.* **12**, 281–296 (2016).
- ³⁴H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, "Interaction models for water in relation to protein hydration," in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981), p. 331.
- ³⁵A. Patriksson and D. van der Spoel, "A temperature predictor for parallel tempering simulations," *Phys. Chem. Chem. Phys.* **10**, 2073–2077 (2008).