Accelerating Convergence in Molecular Dynamics Simulations of Solutes in Lipid Membranes by Conducting a Random Walk along the Bilayer Normal

Chris Neale,[†,‡] Chris Madill,[†,‡,§] Sarah Rauscher,[†,‡,∥] and Régis Pomès*[,†,‡]

[†]Molecular Structure and Function, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, M5G 1X8, Canada
[‡]Department of Biochemistry, University of Toronto, 101 College Street, Toronto, Ontario, M5G 1L7, Canada

Ⓢ _Supporting Information_

ABSTRACT: All molecular dynamics simulations are susceptible to sampling errors, which degrade the accuracy and precision of observed values. The statistical convergence of simulations containing atomistic lipid bilayers is limited by the slow relaxation of the lipid phase, which can exceed hundreds of nanoseconds. These long conformational autocorrelation times are exacerbated in the presence of charged solutes, which can induce significant distortions of the bilayer structure. Such long relaxation times represent hidden barriers that induce systematic sampling errors in simulations of solute insertion.

To identify optimal methods for enhancing sampling efficiency, we quantitatively evaluate convergence rates using generalized ensemble sampling algorithms in calculations of the potential of mean force for the insertion of the ionic side chain analog of arginine in a lipid bilayer. Umbrella sampling (US) is used to restrain solute insertion depth along the bilayer normal, the order parameter commonly used in simulations of molecular solutes in lipid bilayers. When US simulations are modified to conduct random walks along the bilayer normal using a Hamiltonian exchange algorithm, systematic sampling errors are eliminated more rapidly and the rate of statistical convergence of the standard free energy of binding of the solute to the lipid bilayer is increased 3-fold. We compute the ratio of the replica flux transmitted across a defined region of the order parameter to the replica flux that entered that region in Hamiltonian exchange simulations. We show that this quantity, the transmission factor, identifies sampling barriers in degrees of freedom orthogonal to the order parameter. The transmission factor is used to estimate the depth-dependent conformational autocorrelation times of the simulation system, some of which exceed the simulation time, and thereby identify solute insertion depths that are prone to systematic sampling errors and estimate the lower bound of the amount of sampling that is required to resolve these sampling errors. Finally, we extend our simulations and verify that the conformational autocorrelation times estimated by the transmission factor accurately predict correlation times that exceed the simulation time scale—something that, to our knowledge, has never before been achieved.
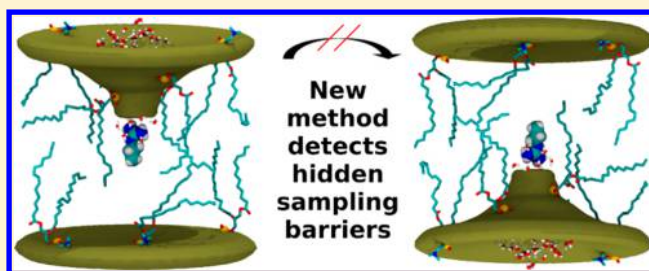
INTRODUCTION

Lipid bilayers are semipermeable barriers that surround cells and their compartments and passively control the movement of nutrients, waste, drugs, infectious particles, and ions.[1] The great biological and pharmaceutical importance of this selective permeability has motivated many attempts to characterize the molecular details underlying these processes, which at present remain poorly understood.

To characterize the interactions of solutes with lipid bilayers, unrestrained molecular dynamics (MD) simulations have been conducted for a variety of molecular solutes.[2−8] Such simulations, however, require long simulation times to surmount energetic barriers whose magnitudes exceed the thermal energy. When the underlying energy landscape is rugged, this limitation may preclude the sampling of low-energy states on accessible simulation time scales.[4,9−12] Furthermore, the relative populations of low-energy states may be incorrectly estimated by unrestrained simulations when there are

insufficient sampling transitions between local minima.[3,7,13,14] In other words, all simulation studies are susceptible to sampling errors,[15] particularly when simulation time scales are less than or similar to the autocorrelation times of the degrees of freedom on which observed values depend.[14] In such cases, statistical measures such as the mean, μ, and standard deviation, σ, of values observed in simulations change significantly with increased sampling. This time dependence is a hallmark of insufficient sampling.

Two types of sampling errors may arise from insufficient sampling: statistical and systematic errors. Statistical sampling errors primarily affect a value's perceived precision, causing oscillations about an average value as the simulation time is increased. Conversely, systematic sampling errors directly affect a value's perceived accuracy. When systematic sampling errors

arise from insufficient sampling, statistical values drift toward the infinite-sampling values as the equilibration time is increased and the systematic sampling errors become smaller.

A variety of techniques have been developed to address the sampling insufficiencies of brute-force MD simulations.[14,16] One of these techniques is umbrella sampling (US).[17,18] US simulations employ a set of restraining potentials to divide phase space into overlapping segments that collectively span a predetermined degree of freedom along which sampling is nearly uniform. The degrees of freedom along which US restraining potentials (umbrellas) exert biasing forces are referred to as order parameters or reaction coordinates. After the US simulations are completed, the biased probability distributions along the order parameter generated under the influence of each umbrella are debiased, combined, and converted into a free energy profile or potential of mean force (PMF).

The most common application of US to studies of solute insertion in a lipid bilayer uses the displacement of the solute along the bilayer normal as the order parameter to compute the PMF for the insertion of the solute from bulk water to the bilayer center.[19,20] Such studies can be used to predict the conformations associated with the lowest-energy insertion depths as well as on-pathway intermediates or transition states for solute insertion or translocation.[13]

This approach has been applied to a long list of solutes of increasing size and complexity: H^+,[21,22] Na^+,[19,21,23] Cl^-,[19,23] K^+,[24] OH^-,[21] and the Na^+Cl^- ion pair;[19] water,[21,25] methanol,[26] ethanol,[27] propane,[27] neopentane,[27] hexane,[27,28] oxygen,[29] ammonia,[27,29] nitric oxide,[27] chloroform,[26,30] 1,2-dimethoxy-ethane,[31] 1,2-dimethoxypropane,[31] hexafluoroethane,[32] chlorobenzene,[26] and 1,3,5-trichlorobenzene;[26] liquid-crystal forming molecules;[33] dimannose;[34] lipids,[35-39] plant waxes,[40] and the lignin argenteane;[41] the steroids cholesterol[39,42-44] and cortisone;[45] drugs including the anticonvulsive valproic acid,[46] the stimulant caffeine[47] and its metabolic derivative para-xanthine,[47] the muscle relaxant chlorzoxazone,[47] the anti-hypertensive debrisoquine,[47] various antiviral adamantanes,[48] the nonsteroidal anti-inflammatories acetylsalicylic acid and ibuprofen,[47,49,50] the anesthetics 1,1,2-trifluoroethane[32] and lidocaine,[26] the antibiotic ciprofloxacin,[51] the beta-blocker atenolol,[50] the chemotherapeutic doxorubicin,[50] the neurotransmitter serotonin,[52] and isomers of the anticoagulant warfarin;[53] the drug-like molecules tyramine,[54] phenethyl-amine,[54] 4-ethylphenol,[54] 2,4,5-trichloroaniline,[26] coumar-in,[47,55] 7-hydroxycoumarin,[47] 3-hydroxyibuprofen,[47] 4-hydroxydebrisoquine,[47] and 6-hydroxychlorzoxazone;[47] the dye Di-4-ASPBS and its derivatives,[56] as well as 2-(2′-pyridyl)-5-methylindole and 2-[2′-(4′,6′-dimethylpyrimidyl)]-indole;[57] the flavonoid antioxidant quercetin and its metabolic derivatives;[58] poly(amidoamine) dendrimers;[59] poly-(alkylthiophene) conductive polymers;[60] a carbon-cage nano-particle;[61] amino acid side chain analogs[13,62-69] and a blocked tryptophan amino acid;[70] and peptides/proteins including a model hexapeptide,[71] a cyclic arginine nonamer,[72] 10-residue antimicrobial oligomers and copolymers of β-amino acids,[73] the 13-, 18-, and 26-residue cationic antimicrobial peptides indolicidin,[74] protegrin-1,[75] and melittin,[76] 16-27 residue synthetic α-helical WALP peptides,[77] the 30-residue protein heteropodatoxin-2,[78] a 42-residue amyloid-β peptide,[79] a ≥80-residue polyleucine α-helix with a central tryptophan[80] or charged or neutral arginine[65,66,81] residue, and the 84-residue protein mechanotoxin-4.[78]

The statistical convergence of PMF calculations depends on the equilibration of all degrees of freedom. Even though the application of a sampling bias along the bilayer normal circumvents problems of convergence that are related to slow relaxation of the solute's bilayer insertion depth, it does nothing to speed up the convergence of sampling in degrees of freedom that do not overlap with (are orthogonal to) the order parameter. In particular, it may remain difficult to obtain converged values because the bilayer structure relaxes slowly.[13] As such, the difficulties of sampling a rugged free energy landscape are only partially alleviated by the US method, which does not eliminate the sampling problem.

Following the work of others,[82] we have previously shown that the addition of equilibrium exchange of umbrella potentials enhances the convergence of US simulations in aqueous solution where systematic sampling errors are introduced by slow relaxation in degrees of freedom orthogonal to the selected order parameter.[83] Others have reached the same conclusion.[84,85] Furthermore, we[11] and others[86] have demonstrated that convergence can also be enhanced by conducting a random walk in temperature,[87-91] although the efficiency of this method depends on the magnitude of the potential energy barriers between states of interest.[92,93] To achieve a random walk in biomolecular simulations, one may use the replica exchange (RE)[87-91] technique or the computationally more efficient virtual replica exchange (VREX)[11] approach (see Methods).

In our previous work, we identified rare and abrupt transitions in bilayer structure as a function of solute insertion depth.[13] These transitions correspond to the slow reorganization of ionic interactions involving zwitterionic phospholipid headgroups when a solute penetrates the lipid−water interface and when n-propylguanidinium, the cationic side chain analog of arginine, is forced through the bilayer center.[13] These rare lipid-reorganization events constitute sampling barriers in degrees of freedom orthogonal to the US order parameter. Such sampling barriers often limit the rate of convergence of equilibrium properties and result in systematic sampling errors but usually remain hidden.[13] This is a good system with which to assess methodological efficiency because, in spite of the solute's small size and few internal degrees of freedom, it can require >100 ns per umbrella to attain convergence in US simulations.[13] Furthermore, the interaction of arginine with lipid bilayers is relevant to membrane protein topology[94] and localization,[95,96] as well as to the activity of antimicrobial peptides.[97]

In this work, we compare the efficiencies of US simulations with and without equilibrium exchange of umbrellas, addressing the statistical convergence of PMF calculations using a system consisting of n-propylguanidinium in a lipid bilayer. We show that conducting a random walk along the bilayer normal speeds up the convergence of US simulations, reducing both systematic and statistical sampling errors. We proceed to introduce a new metric, the transmission factor, that detects hidden free energy barriers by quantifying their impact on the random walks conducted by US-VREX simulations. In this way, we are able to identify solute insertion depths that are prone to systematic sampling errors and quantitatively estimate the amount of time required to attain convergence, even when the required time greatly exceeds the simulation time scale. Finally, we extend the time scale of nonexchanging simulations in which the solute is restrained at the bilayer center and verify

that the conformational autocorrelation times estimated by the transmission factor are quantitatively accurate.

■ METHODS

Simulation Protocol. The simulation system consisted of *n*-propylguanidinium in a hydrated 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC) bilayer. MD simulations were conducted with version 4.0.7 of the GROMACS simulation package.[98] The water model was TIP4P.[99] POPC was modeled by the Berger parameters.[100] *n*-Propylguanidinium was modeled by the OPLS-AA/L parameters[101,102] for the side chain of arginine, where the α-carbon was replaced by a hydrogen atom and the charge on the β-carbon was adjusted to yield an integral molecular charge.[13,62] Hydrogen atoms in *n*-propylguanidinium were modeled as virtual sites that have regular Lennard-Jones (LJ) and Coulombic interactions but no internal degrees of freedom.[103] Berger and OPLS-AA/L parameter sets were combined self-consistently using the half-ε double-pairlist method.[104] The simulation system was neutralized with a single OPLS-AA/L chloride ion. Water molecules were rigidified with SETTLE,[105] and bond lengths between nonvirtual atoms in both *n*-propylguanidinium and POPC were constrained with P-LINCS.[106] LJ interactions were evaluated using a group-based cutoff and truncated at 1 nm without a smoothing function. This treatment of LJ interactions is computationally more expedient than the 1.0/1.8 nm twin-range cutoff that was originally used with these lipid parameters[107] but yields the same area per lipid (APL; see below). Coulomb interactions were calculated using the smooth particle-mesh Ewald method[108,109] with a real-space cutoff of 1 nm and a Fourier grid spacing of 0.12 nm. Simulation in the NpT ensemble was achieved by semi-isotropic coupling to Berendsen barostats[110] at 1 bar with coupling constants of 4 ps and temperature-coupling of the simulation system using velocity Langevin dynamics[111] at 300 K with a coupling constant of 1 ps. The integration time step was 4 fs. The nonbonded pairlist was updated every 20 fs.

System Setup and Umbrella Sampling. A configuration containing a POPC lipid bilayer with 64 lipids per leaflet was obtained from Tieleman et al.[107] The simulation box was elongated along the global bilayer normal (the Cartesian z dimension) to 10.5 nm. Excess water was added to the newly created cavity, resulting in a total of 9162 water molecules in the entire system. This POPC bilayer was simulated for 50 ns in the absence of any solute, during which the APL did not drift systematically (data not shown). Over these 50 ns, the APL was 0.64 ($\sigma = 0.01$) nm² and the dimension of the box along z was 10.8 ($\sigma = 0.1$) nm. (Throughout this article, standard deviations of the sample, σ, are shown in parentheses while the ± symbol is reserved for standard deviations of the mean, σ_M.)

The all-*trans* conformation of *n*-propylguanidinium was embedded in the bilayer using the inflategro routine.[112] Specifically, Cartesian coordinates x and y of the center of mass (COM) of each lipid molecule were scaled by a factor of 4. The solute was then placed with its COM at the desired depth along z and near the middle of the central unit cell in x and y, adjusted by up to ±0.2 nm along x and y to avoid atomic overlap with lipids. Position restraints were applied to solute heavy atoms during system setup. Cycles of steepest descent energy minimization and inflategro lipid coordinate contraction in the xy plane by a factor of 0.95 were conducted until the bilayer APL ≤ 0.7 nm², after which the contraction factor was changed to 0.99 and cycles were continued until the APL ≤

0.65 nm². The structure from the penultimate cycle was used for further simulation. This solute insertion process was repeated 79 times, varying the axial position of the COM of the solute relative to that of the lipid bilayer (the solute insertion depth, z) in 0.1 nm increments in the range $-3.9 \leq z \leq 3.9$ nm. This entire procedure was conducted for three initial solute orientations, with the $C_\delta C_\zeta$ vector of the guanidino group pointing either up or down along the bilayer normal or in the bilayer plane. Thus, the initial angle between the solute dipole and the positive bilayer normal, θ, was either 0°, 90°, or 180°. Note that solute atomic nomenclature is based on the cognate amino acid side chain.

US simulations were conducted for 1.2 μs for each of three initial conformations and under the influence of each of 79 umbrellas using the GROMACS pull code. In these simulations, the solute's insertion depth, z, was harmonically restrained at a specified value, z_i^0, in each umbrella i, with a force constant, $k_u = 500$ kJ/mol/nm², and was stored every 1 ps. Considering that each bilayer has two leaflets, these simulations yield a total of six distinct sets of US simulations for separate evaluations of the PMF from bulk water to the center of the bilayer. These US simulations comprised 284 μs of simulation time.

To evaluate the effects of using larger umbrella force constants for $-0.3 < z_i^0 < 0.3$ nm, we conducted additional restrained simulations for 1.2 μs at each umbrella and for each initial $\theta = 0°$, 90°, or 180° using the close spacing of z_i^0 and large values of k_u as outlined below. These additional US simulations comprised 83 μs of simulation time and will be referred to as the US-B set of simulations.

To verify our estimates of conformational correlation times that exceed the time scale of US-VREX replicas (see below), we conducted 10 additional 5-μs simulations in which the solute was restrained at the bilayer center using $k_u = 3000$ kJ/mol/nm². These simulations were conducted using GROMACS version 4.5.3 and comprised 50 μs of simulation time.

Free Energies and Standard States. The values of z sampled in US simulations were converted to PMFs using Alan Grossfield's implementation[113] of the weighted histogram analysis method (WHAM).[114] To this end, recorded values of the solute insertion depth in the range $-3.9 \leq z \leq 3.9$ nm were distributed among 3120 histogram bins, and the WHAM calculation was performed with a tolerance of 1×10^{-5}. In spite of the congruency of phase space when the solute is in bulk water at $z_i^0 = \pm 3.9$ nm, we did not use a periodic implementation of WHAM (i.e., we did not enforce identical free energies at $z = \pm 3.9$ nm) because it is our intention to identify systematic sampling errors, and it is therefore important that any such errors remain localized at their point of origin and are not artificially distributed across the order parameter, as occurs in periodic WHAM. (A representative PMF computed successively with linear and periodic WHAM is included in Figure S1.) This procedure was repeated for each set of US simulations. Each resulting PMF describes the free energy as a function of solute insertion depth, ΔG_z, from bulk water ($z = -3.9$ nm) across the bilayer center ($z = 0$) to bulk water ($z = 3.9$ nm). For presentation, each PMF was then shifted such that the average value of ΔG_z in the range $-3.9 \leq z \leq -3$ nm, ΔG_{bulk}, equaled zero. Finally, the binding free energy, ΔG_{bind}, was determined by trapezoid integration of the PMF, as outlined previously,[13] separately for each leaflet in the range $0 \leq |z| \leq 3$ nm (after shifting each leaflet-specific PMF such that $\Delta G_{bulk} = 0$), thus placing the boundary of the bound state just

3688

dx.doi.org/10.1021/ct301005b | *J. Chem. Theory Comput.* 2013, 9, 3686−3703

beyond the point where the mean force becomes zero (see Results). The standard binding free energy, $\Delta G_{\text{bind}}^{\circ}$, was then computed by adjusting ΔG_{bind} such that the available volume is equal in the bound and unbound states,[115] thus calculating $\Delta G_{\text{bind}}^{\circ}$ based on a volume-fraction partition coefficient,[116] as outlined previously.[13]

Error Estimates Using Block Averaging. To assess systematic sampling errors in free energy estimates from standard US simulations, we used the block averaging technique.[117] To this end, we used WHAM to compute free energy profiles based on data sets comprising sampling from t to $t + 20$ ns from each umbrella independently for $t = 0$ to 1.18 μs. We then used visual inspection to determine the amount of initial sampling to discard as equilibration. Specifically, we discarded the sampling prior to the value of t at which the standard binding free energy and the difference between the free energy at the global maximum and the minimum in each leaflet stopped drifting systematically with increasing t (see Results).

US with Exchange of Umbrellas. Among the techniques that have been used to achieve a random walk in biomolecular simulations, replica exchange (RE)[87−91] is a generalized ensemble algorithm that comprises a set of simulations (replicas), each of which samples a target ensemble that is restrained at a different position along a globally defined order parameter. Pairs of simulations are periodically subjected to a reciprocal exchange attempt based on their potential energies and their instantaneous positions along this order parameter using a Metropolis Monte Carlo test that preserves equilibrium populations within each ensemble.

Virtual replica exchange (VREX) is another generalized ensemble algorithm that we recently introduced to improve the computational efficiency of RE and used to conduct a random walk of temperatures (T-VREX).[11] VREX improves computational efficiency by removing the need for synchronicity of the two simulations in exchange attempts. This asynchronicity is accomplished by storing the instantaneous value on which exchange is based on a list that is associated with the target sampling ensemble. These stored values are later used in virtual reverse exchange attempts while evaluating the probability of a transition for a replica in a neighboring ensemble.[11]

In this work, we trivially modified T-VREX to conduct a random walk along the bilayer normal by exchanging umbrellas (US-VREX). This extension is similar to replica-exchange umbrella sampling (REUS)[82] and distributed replica umbrella sampling (DRUS).[83]

We conducted three sets of US-VREX simulations in the range $-3.9 \leq z_i^0 \leq 3.9$ nm. Initial conformations were the same as in the US simulations. Umbrellas employed $k_{\text{u}} = 500$ kJ/mol/nm² and were distributed every 0.1 nm, except near the bilayer center where, in an attempt to improve the quality of the random walk, umbrellas were more closely spaced and had stronger restraining potentials than in the first set of US simulations. Specifically, we used $k_{\text{u}} = 1046$ kJ/mol/nm² with $z_i^0 = \pm 0.275$ nm, $k_{\text{u}} = 2092$ kJ/mol/nm² with $z_i^0 = \pm 0.250$ nm, and $k_{\text{u}} = 4184$ kJ/mol/nm² with $z_i^0 = \pm0.225, \pm0.200, \pm0.175$, ..., $\pm0.025, 0.000$ nm. There were thus 97 umbrellas in each set of US-VREX simulations. Stochastic jumps between umbrellas with adjacent z_i^0 were attempted every 4 ps for a total of 1.2 μs per replica (3×10^5 exchange attempts per replica). Sampled values were stored every 0.2 ps.

US-VREX was implemented as a derivative of the client−server architecture of the distributed replica (DR) sampling

software,[118] although VREX does not use the distributed replica potential energy function.[11] Specifically, US-VREX was implemented as follows. After each 4 ps simulation segment, the instantaneous value of z and a unique identifier for the current replica were appended to a dynamic list corresponding to the current value of z_i^0 (the primary list). To determine the transition probability, we computed, in addition to the change in potential energy for the current move attempt, the change in potential energy for a virtual reverse transition. To carry out this virtual reverse transition, we selected the value in the primary list of $z_{i\pm1}^0$ (where the $\pm$ symbol corresponds to the randomly selected direction of the move attempt) that was generated most recently but not by the same replica. The selected value was then removed from the $z_{i\pm1}^0$ primary list and placed on the $z_{i\pm1}^0$ secondary list, which does not track the parent replica. If a value could not be drawn from a primary list according to the rules outlined above, then a value was randomly selected from the secondary list (and not removed). The primary lists were unlimited in length, whereas the secondary lists contained a maximum of 500 values each, beyond which the oldest value was removed prior to adding a new value. The purpose of the secondary lists is to ensure that the VREX set of simulations never runs out of values for virtual reverse moves. In practice, the secondary lists were rarely used (data not shown). No values were stored on virtual-exchange lists in the first 50 simulation segments of any replicas, and no exchanges were attempted for the first 100 simulation segments. To rapidly change the value of z_i^0, we used in-house software to directly modify the GROMACS binary input file. Together, these US-VREX simulations comprised 349 μs of simulation time.

Analysis of Time Correlations. The amount of simulation time required to obtain a new sample of z that is, as far as can be determined, statistically independent of the previous sample of z at a given umbrella during US simulations can be derived from the normalized depth-autocorrelation function

$$C_z(\Delta t) = \left\langle \frac{[z(t) - \mu] \times [z(t + \Delta t) - \mu]}{\sigma^2} \right\rangle_t \quad (1.1)$$

where μ and σ are the mean the standard deviation of z, respectively, and angular brackets indicate averaging over all possible initial-sample times, t. The resulting profile of $C_z(\Delta t)$ is then used to fit an exponential decay curve of the form $\exp(-\Delta t/\tau_{\text{ACF}})$ to obtain τ_{ACF}, the exponential fit to the insertion depth autocorrelation time, which is approximately equal to half of the amount of time required, on average, to obtain a statistically independent (decorrelated) sample of z.[119] The profile of $C_z(\Delta t)$, and the exponential fit, were computed with the GROMACS analysis tool g_analyze.

Back-Exchanges in US-VREX Simulations. Given a replica transition from $z_i^0 \rightarrow z_{i\pm1}^0$, an immediate back-exchange is the occurrence of the reverse transition, $z_{i\pm1}^0 \rightarrow z_i^0$, on the subsequent transition attempt.[11,86] The probability of a back-exchange is denoted P_{back}. For comparison, we also computed $P_{\text{back}}^{\text{random}}$ based on three sets of 97 entirely random walks of 300,000 move-attempts with values of z_i^0 distributed as in the US-VREX simulations, based on the acceptance probabilities obtained from the actual US-VREX simulations. The ratio of P_{back} to $P_{\text{back}}^{\text{random}}$ is $P_{\text{back}}^{\text{ratio}}$.

Transmission Factor from US-VREX Simulations. To quantify the extent and randomness of replica diffusion in US-VREX simulations, we computed the transmission factor along

z, T_z, which generalizes the concept of replica back-exchange[11,86] to nonadjacent umbrellas by evaluating the ratio of the number of times a replica was transmitted across a defined region of the order parameter to the number of times that a replica entered that region. Regions were defined by two constants: x, which represents the umbrella spacing, here 0.1 nm, and N, which is selected such that $2Nx$ is the width of the region under consideration. Specifically, we computed T_z according to

$$T_{z(Nx)}(z_i^0) = \frac{\eta_T(z_i^0)}{\eta_E(z_i^0)} \qquad (1.2)$$

for each umbrella, i, where $\eta_T(z_i^0)$ is the number times a replica moved across a defined region from $z_i^0 \mp Nx \to z_i^0 \pm Nx$ and $\eta_E(z_i^0)$ is the number of times that a replica entered a defined region, moving from $z_i^0 \mp Nx \to z_i^0 \mp (N-1)x$. Because the US-VREX umbrella spacing was nonuniform, eq 1.2 evaluates T_z based on a constant spatial displacement along z, as opposed to a constant number of umbrellas. Values of T_z were not computed for central values of z_i^0 that are not a multiple of 0.1 nm. We also computed T_z^{random} and T_z^{ratio} analogously to our computation of $P_{\text{back}}^{\text{random}}$ and $P_{\text{back}}^{\text{ratio}}$. By default, we use the symbol T_z to represent $T_{z(0.5\text{ nm})}$, explicitly including the value of Nx in the T_z symbol otherwise (T_z^{random} and T_z^{ratio} are notated analogously).

■ RESULTS

To compare the efficiency of US and US-VREX simulations, we used each method to compute the PMF for the water-to-lipid partition of n-propylguanidinium along the normal to a POPC bilayer. Using each method, we conducted three independent sets of simulations across the entire lipid bilayer with different initial orientations of the solute to obtain six independent evaluations of the PMF for moving the solute from bulk water to the bilayer center.

PMFs from US Simulations. The three PMFs across the entire lipid bilayer computed from US simulations are shown together in Figure 1A. In all three cases, partitioning of the solute into the bilayer is favorable (Figure 1A). The leaflet-specific free energy minimum, ΔG_b, is the difference between the minimum value of ΔG_z, $\min(\Delta G_z)$, and the average value of ΔG_z in bulk water, ΔG_{bulk}, on the same side of the bilayer, $\Delta G_b \equiv \min(\Delta G_z) - \Delta G_{\text{bulk}}$. On average, $\Delta G_b = -6.12 \pm 0.19$ kcal/mol and the free energy minimum is located at $|z| = 1.36 \pm 0.04$ nm (Figure 1A). A representative snapshot of the system at the free energy minimum is shown in Figure 1B. Here, the solute's hydrocarbon tail is in the bilayer core, and its charged guanidino group is solvated by lipid headgroups and water molecules.

There is a large free energy barrier at the bilayer center, corresponding to a large invagination of the bilayer that is due to strong interactions between the guanidino group and zwitterionic lipid headgroups (Figure 1C). The global free energy maximum is asymmetric about the bilayer center in all three sets of US simulations (Figure 1A), indicating that all of these PMFs contain systematic sampling errors in this region. Nevertheless, these sampling errors do not significantly affect the computed binding free energy, ΔG_{bind}, because solute residence at the bilayer center is very unlikely (Figure 1A).

To determine the amount of initial sampling to discard as equilibration, t_{eq}, we used block averaging[117] to assess systematic sampling errors according to two values: (i) the
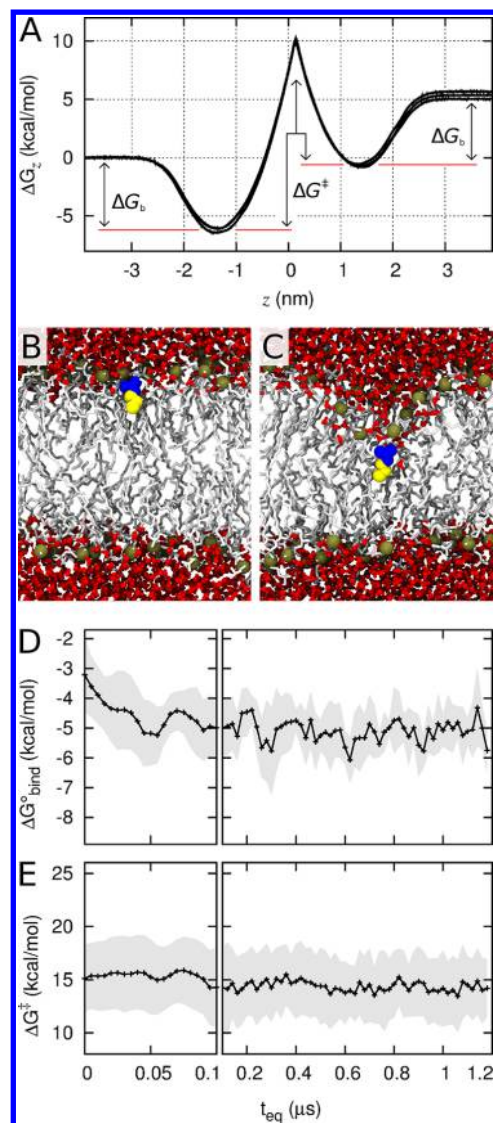

Figure 1. Systematic convergence of US simulations. (A) PMF profiles for insertion of n-propylguanidinium along the bilayer normal independently derived from three sets of US simulations spanning the entire bilayer. In each case, PMFs were computed using the data from 0.1 to 1.2 μs per umbrella. Red lines are used to schematically indicate the free energy differences from bulk water to the leaflet-specific free energy minimum, ΔG_b, and the difference between ΔG_b and the global free energy maximum, $\Delta G^{\ddagger}$. (B and C) Representative conformations at the free energy minimum (B) and maximum (C). Lipid headgroup phosphorus atoms are shown as large brown spheres, water is shown as small red spheres, lipids are shown in white licorice, and n-propylguanidinium heavy atoms are shown in a surface representation with the acyl chain in yellow (C^β,C^γ,C^δ) and the charged guanidino group in blue (N^ε, C^ζ, $N^{\eta 1}$, $N^{\eta 2}$). Molecular visualizations were prepared with VMD.[146] (D) 20-ns block-averaged values of the standard binding free energy, $\Delta G_{\text{bind}}^{\circ}$, as a function of equilibration time, t_{eq}. The mean of six independent evaluations of $\Delta G_{\text{bind}}^{\circ}$ is shown as plus signs, which are connected by a black line to guide the eye. The standard deviation of the mean is shown as a gray contour. To assist the demarcation of systematic sampling errors, separate scales of t_{eq} are used to represent the data for $t_{\text{eq}} < 0.1$ μs and $t_{\text{eq}} > 0.1$ μs. (E) 20-ns block-averaged values of $\Delta G^{\ddagger}$ as a function of t_{eq}. All error bars represent σ_M, the standard deviation of the mean.

standard binding free energy, $\Delta G_{\text{bind}}^{\circ}$, and (ii) the difference between the values of ΔG_z at the global maximum and

minimum in each leaflet, $\Delta G^{\ddagger} \equiv \max(\Delta G_z) - \Delta G_b$ (see Methods). The mean value of ΔG_{bind}° exhibited a large initial systematic drift until $t_{eq} \geq 45$ ns (Figure 1D). Likewise, the mean value of $\Delta G^{\ddagger}$ drifted systematically until $t_{eq} = 95$ ns (Figure 1E). We therefore defined the initial equilibration period as 100 ns, and the PMFs depicted in Figure 1A were computed using the sampling from 0.1 to 1.2 μs per umbrella.

Autocorrelation Times from US Simulations. To assess the amount of simulation time required to obtain a decorrelated sample of solute insertion depth, z, we computed the depth-autocorrelation function, C_z, at each umbrella (omitting the first 0.1 μs as equilibration). We fit these functions to single exponential decays and extracted the autocorrelation times, τ_{ACF}, at each value of z_i^0 (see Methods). These estimates are shown in Figure 2. The bilayer is conceptually divided into four
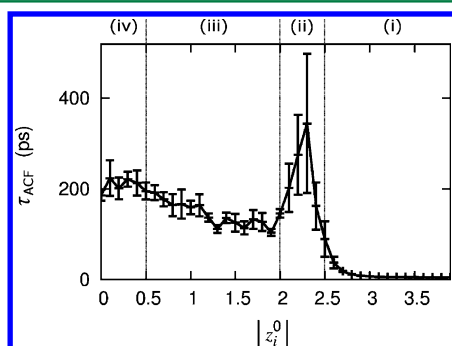
Figure 2. Autocorrelation of solute depth along the bilayer normal in US simulations. The autocorrelation time of z, τ_{ACF}, was computed using the data from 0.1 to 1.2 μs per umbrella. Error bars represent the standard deviation among six independent evaluations of τ_{ACF}, drawn from six US simulations. The bilayer normal is divided into four regions based on the profile of τ_{ACF}: (i) bulk water, (ii) bilayer interface, (iii) hydrophobic core, excluding (iv) the bilayer center.

regions. Region i ($|z_i^0| > 2.5$) corresponds to bulk water, where τ_{ACF} is estimated to be 10 ± 10 ps. Region ii ($2.0 \leq |z_i^0| \leq 2.5$ nm) is the interface between water and the zwitterionic headgroups of the lipid bilayer. Estimates of τ_{ACF} reach a global maximum of 300 ± 100 ps for $2.2 \leq |z_i^0| \leq 2.3$ nm. Region iii ($0.5 < |z_i^0| < 2.0$ nm) is the hydrophobic core of the bilayer, excluding the central region. Estimates of τ_{ACF} are larger in the bilayer core than in bulk water and increase with deeper insertion, averaging 150 ± 30 ps. Finally, region iv ($|z_i^0| \leq 0.5$ nm) is the central region of the bilayer core. Estimates of τ_{ACF} in this region are similar to those in region iii. However, this region is distinct because it contains unresolved systematic sampling errors (Figures 1A,E), and thus τ_{ACF} is grossly underestimated. It is not possible to estimate the convergence time of conformational transitions that were not observed in US simulations,[120] although we will show later that this is indeed possible in US-VREX simulations.

PMFs from US-VREX Simulations. We also conducted three US-VREX simulations to compute the free energy profile for *n*-propylguanidinium along the bilayer normal. Together, these PMFs yield $\Delta G_b = -6.12 \pm 0.06$ kcal/mol with a free energy minimum at $|z| = 1.37 \pm 0.01$ nm (Figure 3A). For ease of comparison, Figure 4 depicts the PMF profiles from both US and US-VREX simulations as a function of $|z|$, exploiting the symmetry of the system with respect to the xy plane at $z = 0$ and shifting each PMF profile such that $\Delta G_{bulk} = 0$.
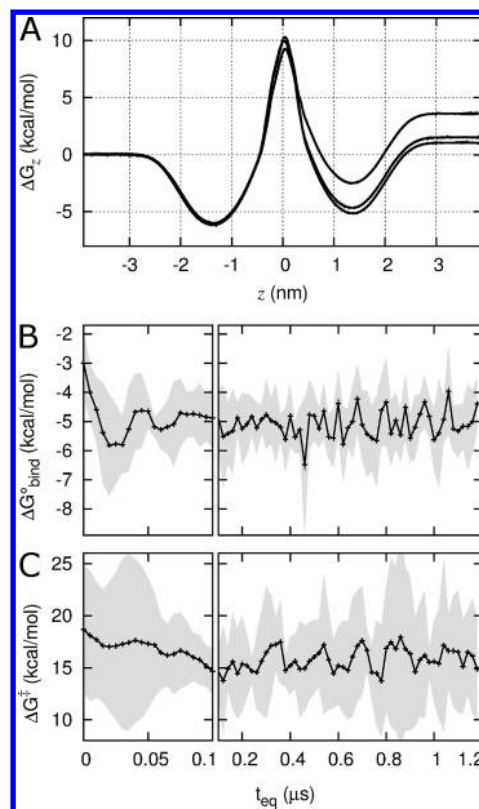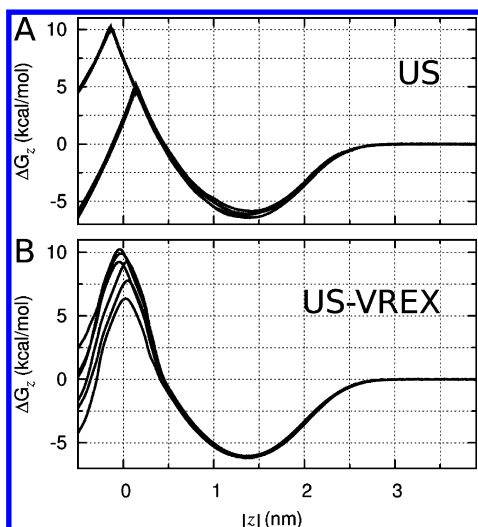

Figure 3. Systematic convergence of US-VREX simulations. (A) PMF profiles for the insertion of *n*-propylguanidinium along the bilayer normal independently derived from three sets of US-VREX simulations spanning the entire bilayer. In each case, PMF profiles were computed using the data from 0.1 to 1.2 μs per replica. (B) 20-ns block-averaged values of the standard binding free energy, ΔG_{bind}°, as a function of equilibration time, t_{eq}. The mean of six independent evaluations of ΔG_{bind} is shown as plus signs, which are connected by a black line to guide the eye. The standard deviation of the mean is shown as a gray contour. To assist the demarcation of systematic sampling errors, separate scales of t_{eq} are used to represent the data for $t_{eq} < 0.1$ μs and $t_{eq} > 0.1$ μs. (C) 20-ns block-averaged values of the difference between the leaflet-specific free energy minimum and the global free energy maximum, $\Delta G^{\ddagger}$, as a function of t_{eq}.

In the US-VREX simulations, the values of ΔG_{bind}° and $\Delta G^{\ddagger}$ exhibited large systematic drifts until $t_{eq} = 20$ ns (Figure 3B and C). Therefore, US-VREX simulations required slightly less equilibration than their US counterparts to eliminate large systematic sampling errors affecting ΔG_{bind}°. Nevertheless, for consistent comparison to US simulation results, we used the data from 0.1 to 1.2 μs for each replica to generate the PMFs depicted in Figure 3A.

Although the central free energy barrier is more symmetric about the bilayer center in PMFs from US-VREX (Figure 3A) than from US (Figure 1A) simulations, as is perhaps most evident in Figure 4, sampling errors persisted throughout 1.2 μs of US-VREX sampling per replica for umbrellas near the bilayer center (Figures 3A and 4B).

Comparing the Statistical Efficiency of US and US-VREX Simulations. To evaluate the relative statistical sampling efficiency of US and US-VREX simulations, we compared the values of the average and the standard deviation of ΔG_{bind}° as a function of increasing simulation time, t, as shown in Figure 5. Values of ΔG_{bind}° from individual PMF profiles are shown in Figure S2.

Figure 4. PMFs as a function of $|z|$. Profiles for insertion of n-propylguanidinium along the bilayer normal from (A) three sets of US simulations and (B) three sets of US-VREX simulations, each spanning the entire bilayer. In each case, PMFs were computed using the data from 0.1 to 1.2 μs per umbrella. These PMFs are identical to those depicted in Figures 1A and 3A but are here displayed as a function of $|z|$ to highlight the increased extent of convergence in US-VREX simulations.
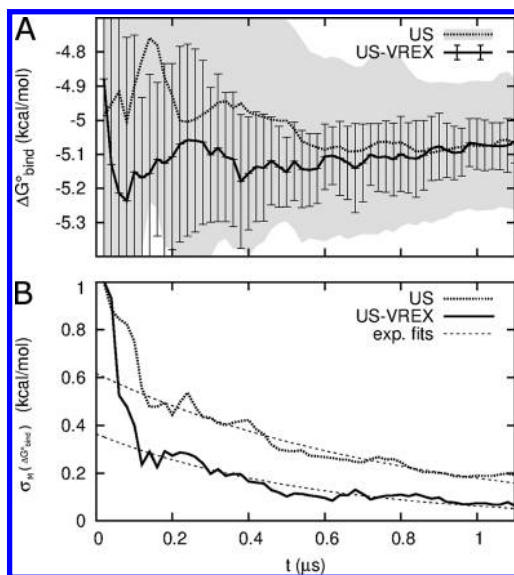
Figure 5. Statistical convergence of US and US-VREX simulations. (A) Value of ΔG°_{bind} as a function of simulation time, t, starting from 0.1 μs per umbrella or replica for (dotted line) mean and (gray contours) standard deviation of the mean from US simulations and (solid line) mean and (error bars) standard deviation of the mean from US-VREX simulations. For individual evaluations of ΔG°_{bind}, see Figure S2. (B) Standard deviation of the mean, $\sigma_M(\Delta G^\circ_{bind})$, as a function of t for (dotted line) US and (solid line) US-VREX simulations. Exponential fits for $0.15 \leq t \leq 1.1$ μs were computed by gnuplot[147] and are shown as broken lines.

The mean values of ΔG°_{bind} obtained from US and US-VREX simulations differ by less than their respective standard deviations for all sampling times (Figure 5A). We therefore evaluated the rate of statistical convergence of ΔG°_{bind} for each method based on the standard deviation of ΔG°_{bind}, $\sigma_M(\Delta G^\circ_{bind})$ (Figure 5B). US-VREX simulations attained a set degree of

statistical precision more rapidly than US simulations for all $\sigma_M(\Delta G^\circ_{bind}) \leq 0.8$ kcal/mol. The time evolution of $\sigma_M(\Delta G^\circ_{bind})$ is not accurately represented by single exponential decay functions, decreasing much more rapidly for $t \leq 0.15$ μs than for $t > 0.15$ μs (Figure 5B). Therefore, we fit these profiles to single exponential decay functions in the range $0.15 \leq t \leq 1.1$ μs. The value of $\sigma_M(\Delta G^\circ_{bind})$ is reliably twice as large when using US as when using US-VREX for all $t \geq 0.1$ μs. Because the decay of $\sigma_M(\Delta G^\circ_{bind})$ is complex, it is not possible to obtain a linear scaling value that describes methodological efficiency so that one might, for example, claim that one method is N times more efficient than the other method. We therefore chose to describe sampling efficiency in terms of the sampling per umbrella or replica that was required to obtain a particular value of $\sigma_M(\Delta G^\circ_{bind})$. This analysis is shown in Table 1. Both US and

Table 1. The Amount of Sampling Required to Obtain Particular Values of $\sigma_M(\Delta G^\circ_{bind})$ from US and US-VREX Simulations[a]

$\sigma_M^{\Delta G^\circ_{bind}}$ (kcal/mol)	sampling per umbrella (μs)		
	US	US-VREX	US/US-VREX
1.0	0.02	0.02	1.0
0.8	0.1	0.06	1.7
0.5	0.14	0.08	1.8
0.4	0.42	0.1	4.2
0.2	1.02	0.34	3
0.1	b	0.6	
0.064		1.1	

[a]Sampling times are listed per umbrella or replica, after discarding the first 0.1 μs as equilibration, and represent the shortest time at which a particular value of $\sigma_M(\Delta G^\circ_{bind})$ was observed. [b]The value of $\sigma_M(\Delta G^\circ_{bind})$ remained >0.185 kcal/mol throughout these US simulations.

US-VREX simulations yield $\sigma_M(\Delta G^\circ_{bind}) = 1.0$ kcal/mol with similar amounts of sampling, but US requires twice as much sampling as US-VREX to reach $\sigma_M(\Delta G^\circ_{bind}) = 0.5$ kcal/mol and three times as much sampling to reach $\sigma_M(\Delta G^\circ_{bind}) = 0.2$ kcal/mol (Table 1).

Sampling Equality and Detailed Balance in US-VREX Simulations. In replica exchange US simulations, reciprocal replica exchanges ensure sampling uniformity along the order parameter. That is, the amount of sampling for each umbrella exactly matches the amount of sampling for each replica. This is not necessarily true for US-VREX simulations, in which the replica exchanges are temporally decoupled. We therefore evaluated the distribution of sampling along the order parameter in US-VREX simulations.
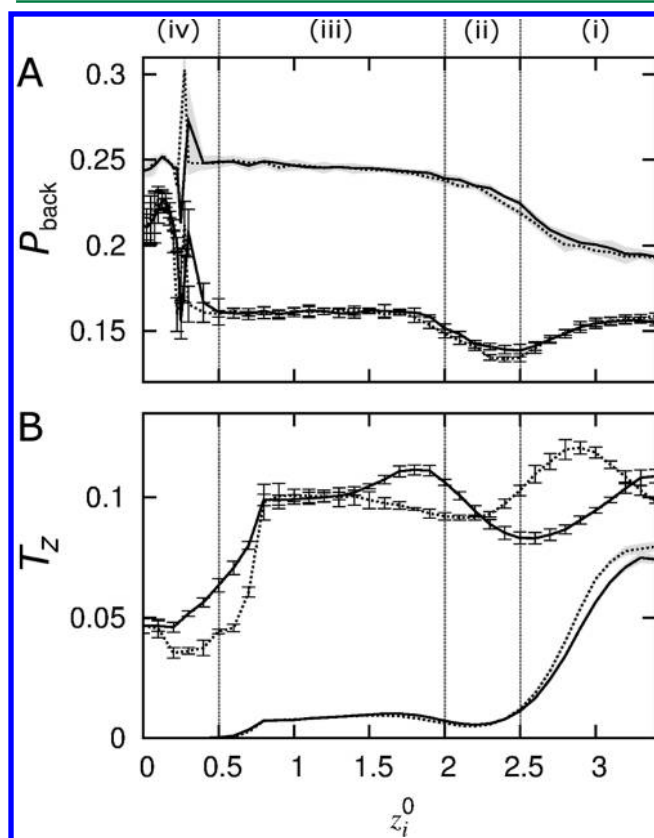
In our US-VREX simulations, the amount of sampling for each umbrella closely matched the amount of sampling for each replica, with 80% and 100% of z_i^0 locations sampled within 50 and 100 ns of the target of 1.2 μs per umbrella, respectively (Figure S3A). Furthermore, replica flux occurred near equilibrium, with the ratio of forward and reverse transitions between z_i^0 and z_{i+1}^0 averaging 1.0 ($\sigma = 0.02$; Figure S3D).

Evaluating the Quality of the Random Walk in US-VREX Simulations. US-VREX is more efficient than US (Figure 5B and Table 1), and yet these methods differ only in their sampling mobility along the order parameter. Therefore, the efficiency of US-VREX must depend on the replica mobility in the so-called random-walk. Generally, the diffusive character of the random walk in Hamiltonian exchange simulations is

affected by free energy barriers that lie along the order parameter ("explicit barriers") and those that lie along orthogonal degrees of freedom ("hidden barriers"). Explicit barriers affect the random walk because changes in the mean force along the order parameter affect sampling overlap between neighboring umbrellas. Hidden barriers affect the random walk because they increase autocorrelation times along the order parameter.[83,121,122]

We evaluated the extent and randomness of replica mobility along z_i^0 by computing the probability of replica back-exchange, P_{back}, and a new metric called the transmission factor, T_z, which measures the ratio of the replica flux that was transmitted across a defined region of the order parameter to the replica flux that entered that region in Hamiltonian exchange simulations (see Methods). Profiles of P_{back} and T_z as a function of z_i^0 computed from US-VREX simulations are shown in Figure 6. In each case, we computed these profiles separately for replica migration toward and away from the bilayer center.

To isolate the effect of explicit barriers on the profiles of P_{back} and T_z, we also computed these profiles from truly random walks based on the exchange-acceptance rates from US-VREX simulations (see Methods). The resulting profiles, P_{back}^{random} and

T_z^{random}, are shown together with P_{back} and T_z in Figure 6. To extract the impact of hidden barriers on the profiles of P_{back} and T_z, we divided the values of P_{back} and T_z obtained from US-VREX simulations by those obtained from truly random walks to obtain P_{back}^{ratio} and T_z^{ratio}. These ratios are shown as a function of z_i^0 in Figure 7. P_{back}^{ratio} reaches a maximum at $1.9 \leq |z_i^0| \leq 2.6$ nm
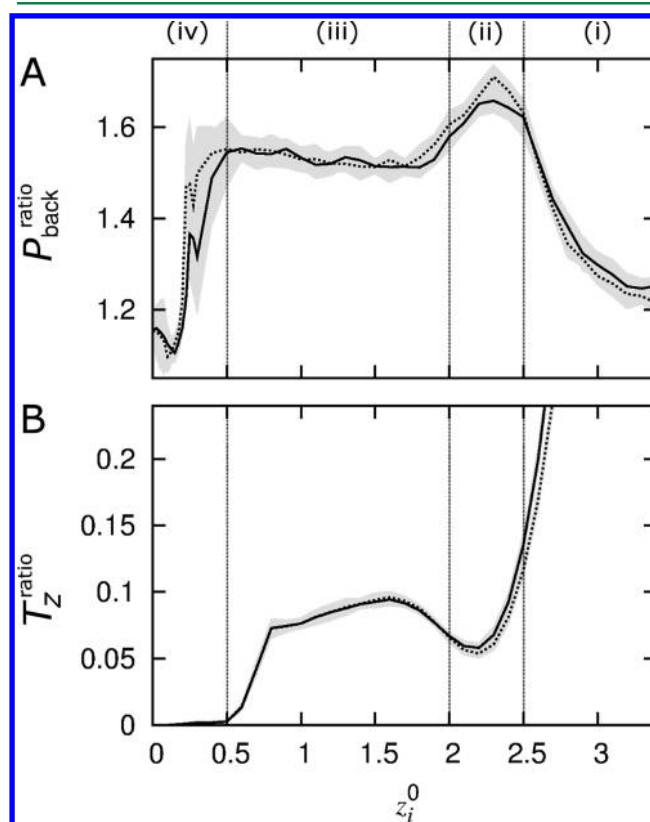
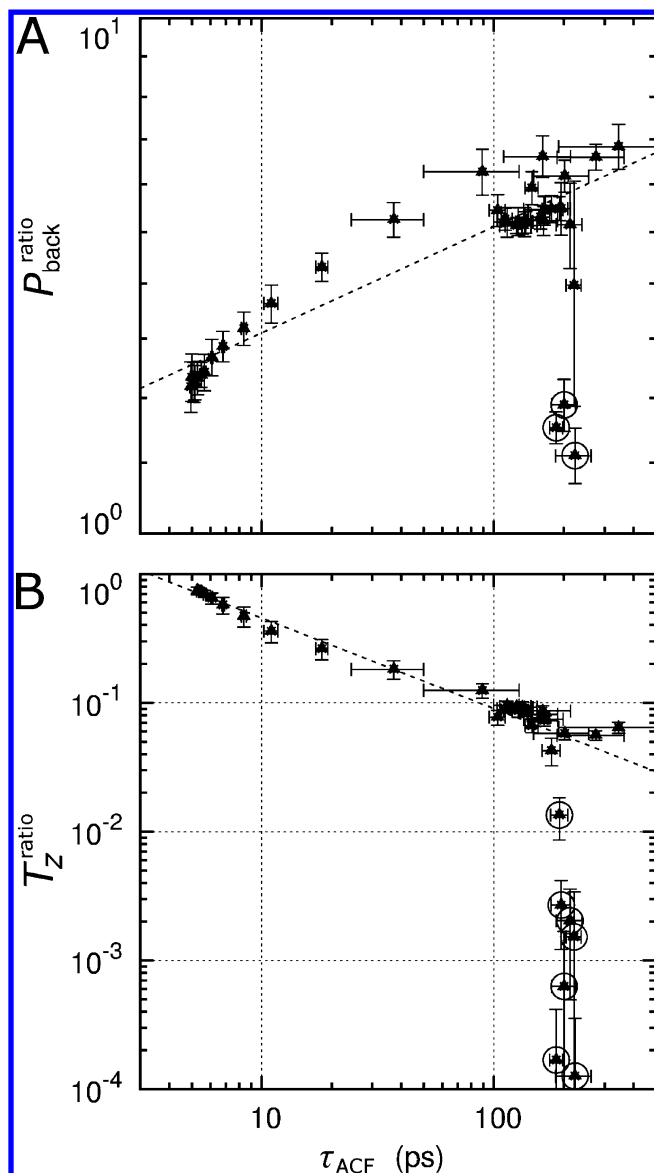**Figure 7.** The quality of the random walk in US-VREX simulations. (A) P_{back}^{ratio} and (B) T_z^{ratio} are plotted along the bilayer normal at center of restraint z_i^0. P_{back}^{ratio} is the ratio of the probability at which back-exchanges occurred in the US-VREX simulation to that probability obtained from truly random walks computed based on the acceptance ratios from US-VREX simulations; T_z was computed with $Nx = 0.5$ nm (see Methods). In each plot, the solid line represents the values obtained for replica migration toward the bilayer center and the dotted line represents replica migration away from the bilayer center. The standard deviation of the mean is shown as gray contours. The bilayer normal is divided into four regions based on the profile of τ_{ACF}, as outlined in Figure 2.

(Figure 7A), and T_z^{ratio} reaches a local minimum at $1.7 \leq |z_i^0| \leq$ 2.4 nm (Figure 7B). The location of both of these extrema overlaps with a local maximum of τ_{ACF} ($2.0 \leq |z_i^0| \leq 2.5$ nm; Figure 2). In addition, T_z^{ratio} reaches a global minimum at $|z_i^0| \leq$ 0.7 nm (Figure 7B).

Next, values of P_{back}^{ratio} and T_z^{ratio} obtained from US-VREX simulations were plotted against values of τ_{ACF} obtained from US simulations (Figure 8). In each case, a good fit was obtained to a power-law relation of the form $a \times (\tau_{ACF})^b$. The fitted power-law constants were $P_{back}^{ratio} = 1.1 \times (\tau_{ACF})^{0.06}$ and $T_z^{ratio} = 2.3 \times (\tau_{ACF})^{-0.7}$ (Figure 8). (This analysis is general to different widths of the transition region, see Figure S4.) Values of P_{back}^{ratio} and T_z^{ratio} obtained at $|z_i^0| \leq 0.2$ nm and $|z_i^0| \leq 0.6$ nm, respectively, were omitted from these fits because their standard deviations did not fit the power-law relation when they were included. It is reasonable to exclude values of τ_{ACF}

Figure 6. The quality of the random walk in US-VREX simulations. (A) P_{back} and (B) T_z are plotted along the bilayer normal at center of restraint z_i^0. Lines and shaded contours represent the overall mean and standard deviations of the mean for P_{back} and T_z, which were obtained from US-VREX simulations. Lines and error bars represent P_{back}^{random} and T_z^{random}, which were obtained from truly random walks based on the exchange-acceptance rates from US-VREX. In all cases, solid and dotted lines represent replica mobility toward and away from the bilayer center, respectively. T_z was computed with $Nx = 0.5$ nm (see Methods). The bilayer normal is divided into 4 regions based on the profile of τ_{ACF}, as outlined in Figure 2.
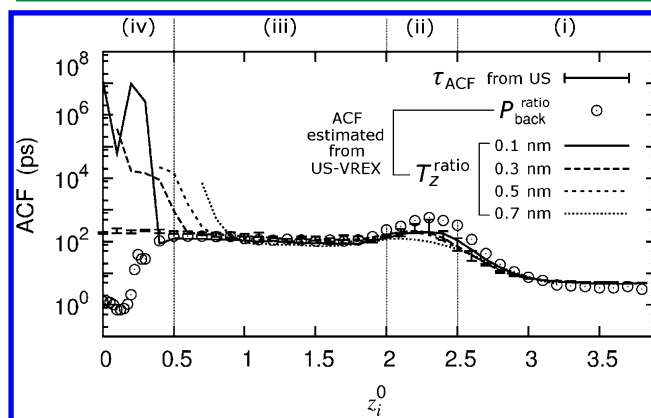
Figure 8. Correlation of US depth-autocorrelation times with the quality of the local random walk in US-VREX simulations. Values of τ_{ACF} computed from US simulations are plotted against values of (A) P_{back}^{ratio} and (B) T_z^{ratio} computed from US-VREX simulations at the same value of z_i^0. Note the log−log scale. Power-law fits to these data are shown as broken lines. Values obtained near the bilayer center whose standard deviations do not fit the power-law relation are circled and were omitted from the fit. These values correspond to $|z_i^0| \leq 0.2$ nm for P_{back}^{ratio} and $|z_i^0| \leq 0.6$ nm for T_z^{ratio}. T_z was computed with $Nx = 0.5$ nm (see Methods).

obtained from US simulations near the bilayer center because, although values of τ_{ACF} computed from US simulations at $|z_i^0| \leq 0.7$ nm were on the order of 200 ps (Figure 2), the PMFs remained asymmetric about the bilayer center after 1.2 μs per umbrella (Figure 1A), indicating that unresolved systematic sampling errors persisted throughout US simulations near the bilayer center and that the associated estimates of τ_{ACF} are incorrect.

To estimate the depth autocorrelation times from US-VREX simulations, we used the fits of τ_{ACF} to P_{back}^{ratio} and T_z^{ratio} to back-calculate estimates of τ_{ACF}. These profiles are shown in Figure

9. Importantly, profiles of T_z^{ratio} indicate that τ_{ACF} may be on the order of 10 μs for some umbrella(s) near the bilayer center.

Figure 9. Depth-autocorrelation times computed from US simulations and estimated from US-VREX simulations. Values of τ_{ACF} from US simulations are shown as error bars. Estimates of τ_{ACF} from US-VREX simulations based on P_{back}^{ratio} are shown as circles. Estimates of τ_{ACF} based on T_z^{ratio} are shown as lines for transmission factors computed using $Nx = 0.1$ nm (solid line), $Nx = 0.3$ nm (broken line with small gaps), $Nx = 0.5$ nm (broken line with large gaps), and $Nx = 0.7$ nm (dotted line; see Methods). The bilayer normal is divided into four regions based on the profile of τ_{ACF}, as outlined in Figure 2.

Testing the Transmission Factor Prediction. To improve the accuracy of the estimates of τ_{ACF} obtained from US simulations with 1.2 μs per umbrella, we conducted 10 additional 5-μs simulations in which the solute was restrained to the bilayer center. These simulations were initiated with a defect in the upper bilayer leaflet (similar to the conformation depicted in Figure 1C). The time-series of z values, averaged over 100-ns blocks, $\langle z \rangle$, are shown in Figure 10.

When n-propylguanidinium is restrained at the bilayer center ($z_i^0 = 0.0$ nm), water may penetrate either the upper or lower leaflet, or both.[13] When this cationic solute is solvated by a defect in the upper leaflet, surface tension draws the solute upward on average.[13] Conversely, when n-propylguanidinium is
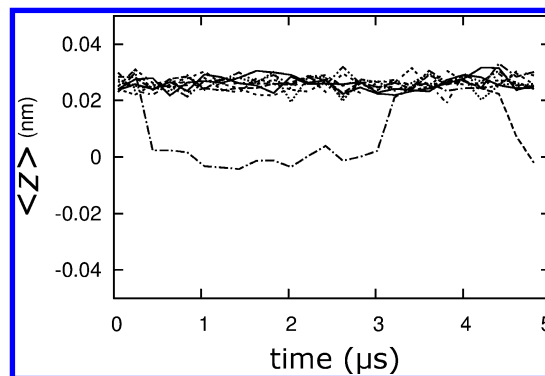
Figure 10. Bilayer reorganization in longer simulations. Results depict z values averaged over 100 ns, $\langle z \rangle$, as a function of simulation time. Each line represents one of ten independent simulations initiated with a single defect in the upper leaflet of the bilayer, with n-propylguanidinium restrained at $z_i^0 = 0.0$ nm. Values of $\langle z \rangle$ greater than 0.02 nm represent a defect in the upper bilayer defect. Values of $\langle z \rangle$ near 0.00 nm represent a double bilayer defect, and values of $\langle z \rangle$ less than −0.02 nm, which were not sampled, represent a lower bilayer defect.

solvated by a defect in the lower leaflet, it tends to be drawn downward.[13] Finally, n-propylguanidinium may be solvated by a double bilayer defect in which the integrity of both leaflets is compromised. In this case, the time-averaged position of the solute is at the bilayer center.

It is apparent in Figure 10 that, out of a total of 50 μs of simulation, there were two transitions from an upper to a double bilayer defect, one transition from a double to an upper bilayer defect, and no sampling of a lower bilayer defect. Contrary to the estimate of τ_{ACF} obtained from the 1.2 μs of restrained simulation (Figure 2), these results corroborate the estimate of $\tau_{ACF} \geq 10$ μs obtained from our analysis of T_z^{ratio} (Figure 9).

Water Pores. In the simulations, large water-filled pores spanning the entire bilayer occasionally formed. These pores were transient and invariably originated from conformations in which n-propylguanidinium was restrained near the bilayer center (data not shown). Large water pores, with more than 20 water molecules within 0.5 nm of the bilayer COM, formed on four separate occasions in US-VREX simulations, but never in the US simulations used to generate PMFs. When water pores occurred, they were restricted to $|z_i^0| < 0.2$ nm. An example of the transient formation of a water pore is provided in Figure 11.

Because the US-VREX simulations utilized stronger force constants for $|z_i^0| < 0.3$ nm, it was unclear whether the apparent increase in the propensity for bilayers to form pores in US-VREX simulations was due to replica migration or to the stronger restraining potentials. We thus conducted an additional set of US simulations for $|z_i^0| < 0.3$ nm using the same umbrella spacing and force constants as the US-VREX simulations. We refer to these additional simulations as US-B. Water pores formed in three of the 23 US-B simulations, all with $|z_i^0| < 0.1$ nm, demonstrating that the formation of water pores is indeed facilitated by stronger force constants.

Bilayer Leaflet Asymmetry and Lipid Flip-Flop. Throughout the US and US-VREX simulations used to generate PMFs in this work, the phosphorus atom of a lipid headgroup sometimes crossed the bilayer center to sample z values within the opposing leaflet. In these events, however, fluctuation of errant lipid headgroups across the bilayer center never exceeded a displacement of 0.2 nm (Table 2). An example of such an event, taken from US simulations, is depicted in Figure 12. In this simulation, the solute was restrained near the center of the bilayer at $z_i^0 = 0.1$ nm. The errant lipid, initially not in contact with the solute, diffused into the conical bilayer defect around n-propylguanidinium, upon which the phosphate oxygen atoms formed two salt-links with the cationic solute headgroup (Figure 12C) as the lipid transiently crossed the bilayer center (Figure 12E). This state was short-lived, with the lipid returning to its original leaflet after 3 ns (Figures 12D,E). The transient leaflet headgroup asymmetry event depicted in Figure 12 was the only occurrence of such an event in the US simulations. Such transient leaflet crossings were more common in the US-VREX simulations, occurring in 11 US-VREX replicas (Table 2).

Complete lipid migration from one leaflet to the other (lipid flip-flop) did not occur in our US or US-VREX simulations (Table 2). It did, however, occur in the US-B simulations. The phosphorus atom of a lipid headgroup crossed the bilayer center in 10 of the 23 US-B simulations and, in three of these simulations, a total of 14 lipid molecules completely crossed the bilayer and joined the opposing leaflet (Table 2). Nine of the 14 lipid flip-flop events occurred in a single simulation at $z_i^0 =$
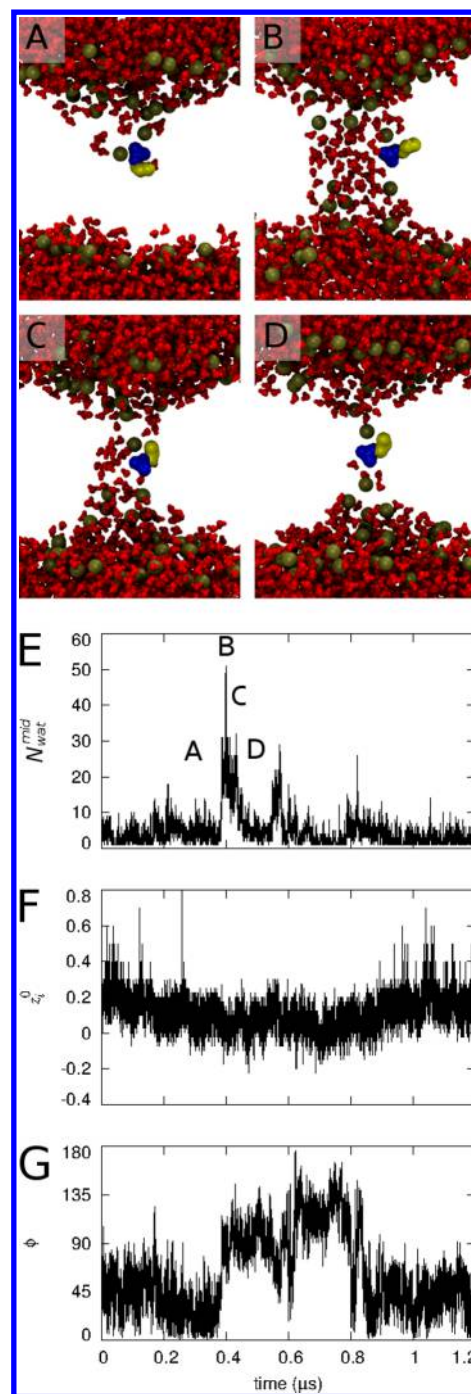
Figure 11. Formation of a water pore. The data were taken from one US-VREX replica. (A−D) Snapshots taken from one simulation at (A) 0.3 μs, (B) 0.4 μs, (C) 0.43 μs, and (D) 0.5 μs. (E) Number of water molecules within 0.5 nm of the bilayer center, N_{wat}^{mid}, (F) value of z_i^0, and (G) angle between the solute dipole and the positive bilayer normal, θ, sampled by this replica as a function of simulation time. In parts A−D, lipid headgroup phosphorus atoms are shown as large brown spheres, water is shown as small red spheres, and n-propylguanidinium heavy atoms are shown in a surface representation with the acyl chain in yellow (C^β, C^γ, C^δ) and the charged guanidino group in blue (N^ε, C^ζ, $N^{\eta 1}$, $N^{\eta 2}$).

0.05 nm (Supporting Information Movie S1). Note that lipid bilayer asymmetry would contribute to systematic errors in the calculation of solute insertion. For this reason, US-B simulations were not used to generate PMFs.

Table 2. Bilayer Leaflet Asymmetry from US and US-VREX Simulations

	US	US-VREX	US-B[a]
no. simulations in which the lipid headgroup phosphorus atom crossed the $z = 0$ nm plane	1	11	10
max. z_i^0 that sampled a crossing event (nm)	0.1	0.5	0.175
total simulation time for $z_i^0 \leq 0.2$ nm (μs)	18	61.2	61.2
no. lipids that flip-flop	0	0	14
no. simulations with flip-flop	0	0	3
max. displacement across $z = 0$ nm (nm)	0.2	0.2	>2
max. duration across $z = 0$ nm without dissociating from the bilayer defect (ns)	20	38	400
max. duration after flip-flop (ns)	0	0	800

[a]These are the additional US simulations conducted near the bilayer center using stronger force constants for the harmonic restraints (see Methods).
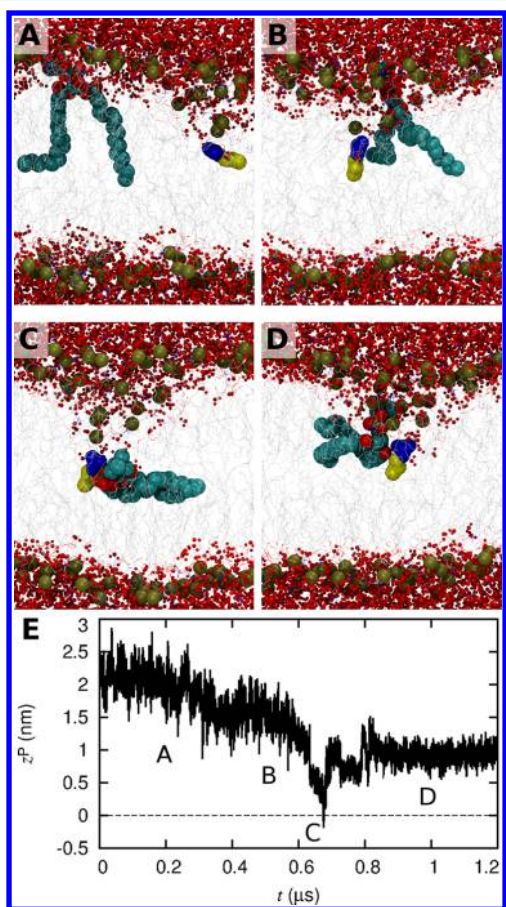
Figure 12. Solute-induced fluctuations of a lipid headgroup across the bilayer center. The plots depict the only event in which a lipid headgroup phosphorus atom crossed the bilayer center to sample z-values within the opposing leaflet in US simulations with low force constants. The solute was restrained at $z_i^0 = 0.1$ nm. (A−D) Snapshots depicting this simulation at (A) 0.2 μs, (B) 0.5 μs, (C) 0.67 μs, and (D) 1 μs. The snapshot in part C corresponds to the minimum value of z^P, the insertion depth of the phosphorus atom of the errant lipid. (E) z^P as a function of simulation time. In parts A−D, the representation is the same as in Figure 11, with the following changes: water molecules are smaller, the errant lipid is highlighted in a space-filling representation, and the remaining lipids are depicted by gray lines.

Localization of the Cl⁻ Ion. The neutralization of solute charge by a single Cl⁻ ion could in principle result in sampling errors based on quasi-nonergodic localization of the Cl⁻ ion. However, throughout all of the sampling conducted in this study, which exceeds 0.7 ms of simulation time, the single Cl⁻ ion remained solvated by bulk or interfacial water and never penetrated the surface formed by lipid headgroups.

■ **DISCUSSION**

The primary objective of the present study is to evaluate the rate at which the insertion free energy of a molecular solute into a membrane attains convergence in US simulations. Therefore, although there has been much debate about the protonation state of arginine within a lipid bilayer[97,123] we do not address the pK_a shift that a bilayer imposes on n-propylguanidine, and we focus instead on the sampling convergence of the bilayer interacting with n-propylguanidinium in its cationic state. Computational evaluation of pK_a requires converged sampling of both ionic and neutral species,[62,63,66,124,125] and our evaluation of sampling convergence thus remains relevant to computational pK_a calculations whether or not arginine is charged in cell membranes.

The Results show that US-VREX simulations attained convergence more rapidly than US simulations (Figures 1, 3, 4, and 5 and Table 1). Here, we discuss the factors contributing to this relative efficiency in terms of systematic and statistical sampling errors.

Systematic Sampling Efficiency of US-VREX Simulations. The reduction of systematic sampling errors by US-VREX is likely to be assisted by the ability of replicas to locally sort themselves along the order parameter according to their relative suitabilities at different values of z_i^0. Replica sorting has also been identified as an important mechanism by which sampling efficiency is enhanced when using temperature as an order parameter.[86]

As an example of the benefits of replica sorting in US-VREX simulations, consider a situation in which one replica at $z_i^0 = 0.1$ nm has a defect in the lower leaflet and another replica at $z_i^0 = -0.1$ nm has a defect in the upper leaflet. It is energetically favorable for these two replicas to exchange values of z_i^0 and thereby in each case reduce the size of the bilayer invagination that is required for the solvation of n-propylguanidinium by water and lipid headgroups.[13] In US simulations, z_i^0 is constant and this exchange can only be accomplished by defect interconversion at each static value of z_i^0. Such defect interconversion requires the system to pass through a transition state consisting of either a solute that is completely surrounded by acyl chains (unperturbed bilayer) or one in which both leaflets invaginate (double bilayer defect). That is, we postulate that the progression from an upper-leaflet defect to a lower-leaflet defect in the lipid bilayer must occur according to one of the following two defect progressions: upper→none→lower or upper→double→lower. Our previous results indicate that n-propylguanidinium is in constant contact with water and lipid headgroups along the bilayer normal,[13] suggesting that the transition state does not involve complete nonpolar solvation of the solute and that the progression is upper→double→lower. Furthermore, the present simulations indicate that double defects can occur for values of $z_i^0 < 0.2$ nm (Figure 11). The spontaneous formation of a double defect, which requires a bilayer invagination exceeding 1 nm along the bilayer normal,[13] represents a hidden barrier to sampling convergence in US simulations. This hidden barrier is large, as indicated by the

systematic sampling errors that persisted throughout our US simulations for $z_i^0 < 0.2$ nm (Figures 1A and 4A) and the long autocorrelation times of insertion depth in this region estimated from US-VREX simulations (Figure 9). In US-VREX simulations, however, z_i^0 is variable, and the exchange of upper and lower defects can occur by way of random walks along the order parameter, given that the appropriate states exist at some values of z_i^0.

As another example, consider the sampling at $z_i = 2.5$ nm. At this solute location, the most favorable arrangement involves local lipid ordering and extension to form a bilayer protrusion that can directly interact with *n*-propylguanidinium.[13] This state is stabilized by the formation of salt bridges between the guanidino group of *n*-propylguanidinium and the zwitterionic headgroups of the lipids.[13] Because our simulations were initiated by placing the solute at z_i^0 with respect to an unperturbed bilayer, this distorted state can only be attained in an US simulation following a spontaneous bilayer fluctuation that may require tens of nanoseconds to occur.[13] US-VREX simulations provide an alternative route by which this state can be attained. It is possible, for instance, for a replica restrained at $z_i^0 < 2.5$ nm to rapidly attain a state in which *n*-propylguanidinium intercalates between lipid headgroups and then, by way of a random walk along z_i^0, to reach $z_i^0 = 2.5$ nm. Replica sorting in US-VREX simulations may thus provide an important source of rapid equilibration and reduction of systematic sampling errors.

Both of the above examples of replica sorting outline ways in which Hamiltonian-exchange simulations can circumvent free energy barriers orthogonal to the reaction coordinate, a feature that we have previously demonstrated for the relaxation of the backbone dihedral angles of the alanine dipeptide in aqueous solution in US simulations undergoing an equilibrium exchange of umbrellas.[83]

In spite of this inherent ability of US-VREX simulations to assist conformational relaxation by replica sorting, rearrangements of z_i^0 did not encompass the entire order parameter but were blocked by the long depth-autocorrelation times at the bilayer center (Figure 9). This limitation is another consequence of replica sorting, here an undesirable one. Once the replicas near $z_i^0 = 0$ nm have sorted themselves locally about the bilayer center, they remain trapped on the side of the bilayer in which the invagination occurs. Replica mobility across the bilayer center then requires the US-VREX simulations to surmount the same hidden free energy barrier to invagination of the opposing leaflet that must be surmounted in US simulations. This general effect has been observed by others while conducting random walks along umbrellas,[84,85,126] alchemical transformations involving a λ coupling parameter[127−129] or an extra spatial dimension,[130] and other Hamiltonian modifications.[131]

It is a general conclusion that Hamiltonian exchange helps to alleviate systematic sampling errors. Consider, for instance, simulations aimed at determining the binding free energy of a flexible and/or orientationally anisotropic ligand in a protein binding pocket. Standard US simulations and alchemical transformations involving a λ coupling parameter[132,133] must include component simulations in which the ligand is fully bound, or coupled, to the protein. Here, the van der Waals surface of the protein may forbid significant solute reorientation or conformational change, potentially abrogating sampling of binding modes with low free energies. Hamiltonian exchange simulations reintroduce sampling mobility along the order parameter while maintaining sampling homogeneity along the order parameter. Thus, in this example, Hamiltonian exchange allows the solute to unbind, or decouple, from the protein's binding pocket, find new orientations/configurations, and then rebind in a new mode.[130] Generally, Hamiltonian exchange simulations facilitate the occurrence of transitions between local minima orthogonal to the order parameter by allowing trajectories to explore the order parameter and find lower energy pathways.[83]

It is important to note that generalized ensemble methods relying on a random walk along an order parameter are not panaceas. Although these random walks reintroduce relaxation pathways that are disallowed by umbrella sampling methods, orthogonal degrees of freedom must still be crossed by brute force.[133] If the simulations are too short to permit the remaining barriers to be crossed by brute force, then conformational basins of low free energy may remain unsampled. That is, Hamiltonian exchange is no substitute for the ideal order parameter which, in general, is unknown.

Statistical Sampling Efficiency of US-VREX Simulations. US-VREX simulations led to marked improvements in statistical sampling over US simulations for equal amounts of data (Figures 4 and 5B and Table 1). The addition of equilibrium exchange to US simulations of lipid bilayers provides a substantial computational advantage, as demonstrated by the 3-fold reduction in sampling times, from 1.02 to 0.34 μs per umbrella, to achieve a statistical precision of 0.2 kcal/mol (Table 1).

The reduction of statistical sampling errors in US-VREX simulations stems from the fact that multiple replicas contribute to the sampling at each value of z_i^0. Similarly, statistical sampling errors can be reduced in US simulations by conducting many repeat simulations at each value of z_i^0. In US simulations, however, this procedure does not lead to an increase in statistical sampling efficiency, which is inversely proportional to the amount of sampling. Conversely, by conducting a random walk along z_i^0, US-VREX permits multiple replicas to contribute to the sampling at each umbrella without increasing the total amount of sampling. Whenever US-VREX exchanges are accepted at a rate faster than the longest autocorrelation time encountered in US simulations at a given umbrella (which is quite generally the case), the sampling efficiency of US-VREX is expected to be better than that of US. Moreover, the transient migration of replicas to values of z_i^0 for which the system is poorly equilibrated can contribute meaningfully to the tails of the population distributions at each value of z_i^0 without necessitating that any individual replica sample all of these regions extensively, as is required in US simulations.

Detecting and Quantifying Hidden Sampling Barriers. In our previous work, we detected the presence of hidden barriers hindering the convergence of solute insertion depth from the analysis of abrupt transitions in bilayer organization as a function of z.[13] The identification of systematic sampling errors in that work, as in the present study, was facilitated by the symmetry of the lipid bilayer. However, the coordinates in which hidden barriers exist are not known *a priori*, and in complex systems it is not in general obvious which orthogonal degrees of freedom may present sampling barriers.

If the conformational decorrelation orthogonal to the order parameter is very rapid, then there are no hidden free energy barriers, enthalpic or entropic. In this case, replica migration along the order parameter approaches a true random walk (with deviations from randomness being entirely ascribed to

variations in sampling overlap between neighboring umbrellas as a consequence of heterogeneous umbrella spacing and/or free energy differences along the order parameter). In complex systems, hidden barriers increase the time required for conformational decorrelation orthogonal to the order parameter. When relaxation of the order parameter is coupled to relaxation of the orthogonal degrees of freedom presenting hidden barriers, hysteresis in the order parameter is increased.

The situation in which relaxation of degrees of freedom orthogonal to the order parameter lags behind the relaxation along the order parameter is referred to as Hamiltonian-lagging,[121] a concept that was generalized in a recent article.[122] Hamiltonian-lagging reduces the efficiency of Hamiltonian-exchange algorithms by impeding the random walk. It does so whenever a replica is in a region of orthogonal space for which moves (in one or both directions) along the order parameter are less likely than they are in some other region of orthogonal space. If the two above-mentioned regions of orthogonal space are separated by a free energy barrier, extensive sampling is required to move between them. Importantly, it follows that the extent of the Hamiltonian lag, and thus the extent of local replica mobility along the order parameter, contains information about the magnitudes of hidden free energy barriers. We can therefore use the local quality of the random walk to quantify the size of the hidden barriers or their impact on the order parameter autocorrelation function in each umbrella. In this way, the transmission factor turns Hamiltonian-lagging to our advantage by revealing the existence of hidden barriers wherever T_z is smaller than T_z^{random} (wherever $T_z^{ratio} < 1$).

In this regard, T_z^{ratio} outperformed P_{back}^{ratio}, which failed to identify hidden barriers at the bilayer center (Figure 9). An alternative definition of P_{back}, requiring that a reverse transition occur on the subsequent transition, rather than on the subsequent transition attempt, also failed to identify hidden barriers at the bilayer center (data not shown).

The reason why T_z^{ratio} is superior to P_{back}^{ratio} in detecting hidden barriers in this system is due to noise resulting from collective fluctuations of the bilayer, which can spread the apparent width of the hidden barrier along the order parameter beyond the finite width of the transition region. Specifically, when lipids far from the solute move along the bilayer normal, they affect z but not the local environment of the solute. Thus, in a given microstate, a hidden barrier may be located away from its average position along the bilayer normal. The superior hidden-barrier detection of T_z^{ratio} is not, however, limited to systems containing a lipid bilayer. In US simulations of protein−ligand interaction, for example, a distance-based order parameter may also be degenerate due to protein breathing motions, side chain rotamerization, and conformational reorganization of the ligand.

An important finding of this paper is that the transmission factor can be used not only to identify hidden barriers but also to estimate the magnitude of autocorrelation times that exceed the simulation time scale. Thus, we estimated that the depth autocorrelation time at $z_i^0 = 0.0$ nm is on the order of 10 μs based on the analysis of US-VREX simulations of only 1.2 μs per replica (Figure 9). We then confirmed this estimate using long-time US simulations (Figure 10). It was possible to accurately estimate the depth autocorrelation time from US-VREX simulations because many replicas had the opportunity to migrate across the bilayer center, each contributing to the transmission factor, whereas in each set of US simulations, only one simulation can be used to evaluate the depth

autocorrelation time at $z_i^0 = 0.0$ nm. This rationale sets the total simulation time of the entire set of replicas as the upper limit on the value of the autocorrelation time along the order parameter that can be estimated from Hamiltonian-exchange US simulations. Hidden barriers of sufficient magnitude may preclude replica migration across some regions of the order parameter, in which case the transmission factor can only estimate the lower bound of the depth autocorrelation time, and the actual autocorrelation time may be substantially longer. Note that although hidden barriers can be identified and ranked by magnitude based on US-VREX simulations alone, a separate set of standard US simulations is required to obtain the correlation between T_z^{random} and τ_{ACF}, which is necessary to estimate the autocorrelation times from US-VREX simulations.

The above analysis shows that transmission factors computed from US-VREX simulations (or other methods based on equilibrium exchange of biasing potentials) are a useful new tool to identify regions along the order parameter where hidden barriers may lead to quasi-nonergodicity and systematic sampling errors. US-VREX is thus an excellent method for efficiently constructing a PMF while concurrently identifying the putative locations of hidden sampling barriers along the order parameter, toward which significantly more computational resources should be directed.

Novelty and Usefulness of the Transmission Factor. We were first alerted to the presence of hidden barriers when we plotted replica z_i^0 values as a function of time and observed inhomogeneous diffusion of replicas along the order parameter (Figure S5). We rationalized this behavior in terms of replica sorting (see above). Generally, the replicas of Hamiltonian exchange simulations tend to remain sorted along the order parameter based on their sampling of orthogonal degrees of freedom that (a) contain large hidden free energy barriers and (b) influence the mean force along the order parameter. We are not the first to examine replica mobility as a means to gauge sampling convergence. In temperature replica exchange simulations,[11,89,129] it is common to qualitatively evaluate the diffusion of replicas in temperature to ensure that all replicas have sampled at all temperatures and that the random walks have moved between the lowest and highest temperature many times.[92,93,134−138] Estimates of the first passage times that characterize replica migration across the temperature range have also been used to optimize temperature selection.[139−141] Furthermore, Wyczalkowski and Pappu have noted the ability of Hamiltonian exchange to reduce the hysteresis error in free-energy perturbation calculations.[142] The novelty of the present approach consists of quantifying the local replica diffusion along the order parameter and of using that information to learn about sampling overlap in orthogonal degrees of freedom (and therefore about the presence of hidden free energy barriers). To extract this information, we developed a measure of the attenuation of replica diffusion based on an analogy to optics, wherein the transmission factor is defined as the ratio of the light transmitted by an object to the light that is incident upon that object. A schematic depiction of how a hidden free energy barrier can lead to a decrease in the transmission factor is included in Figure S6.

Importantly, we expect that the transmission factor will also provide useful information about λ coupling parameter values at which alchemical transformations are prone to systematic sampling errors. Furthermore, the transmission factor may be capable of identifying phase transitions in temperature exchange simulations.

Lipid Asymmetry. Water pores sometimes formed in the bilayer when *n*-propylguanidinium was restrained near the bilayer center, specifically at $|z_i^0| \leq 0.2$ nm (Figure 11). A previous MD study indicated that water pores enhance the propensity for lipid flip-flop.[35] In the context of the current simulations, lipid flip-flop can change the location of the bilayer COM and affect the solute's PMF by perturbing the order parameter.

Although no *bona fide* lipid flip-flop events were observed in our US and US-VREX simulations (Table 2), we expect that lipid flip-flop would eventually occur if the simulation time was increased. Replica mobility in US-VREX simulations may then permit the migration of asymmetric bilayers across the entire range of the order parameter, exacerbating the potential degeneracy of the order parameter and potentially increasing statistical sampling errors in these simulations. Accordingly, US-VREX simulations sampled transient leaflet headgroup asymmetry for $|z_i^0| \leq 0.5$ nm whereas US-B simulations sampled this transient asymmetry only for $|z_i^0| < 0.2$ nm (Table 2).

One solution to this problem is to abrogate lipid asymmetry by adding a planar restraint on the average position of the *sn*-2 carbon of the glycerol moiety in each lipid, as done by Nymeyer et al.[143] in order to maintain bilayer symmetry and integrity at high temperatures. This approach, however, is not feasible in the presence of solutes such as *n*-propylguanidinium, whose insertion induces significant bilayer defects. During deep insertion of this solute, the bilayer invaginates[13] because the free energy cost of bilayer deformation is less than the free energy cost of entirely immersing this cation among the hydrophobic acyl-chains of an otherwise unperturbed lipid bilayer. The inhibition of bilayer deformation therefore precludes the accurate evaluation of the free energy barrier to translocation of this solute. An elegant solution is to conduct simulations in the P2$_1$ space group, which allows lipids to move from one leaflet to the other across periodic boundary conditions.[144] In such simulations, bilayer asymmetry can be rapidly remediated by intraleaflet diffusion because the lower leaflet becomes the upper leaflet as it crosses a periodic boundary.

CONCLUSIONS

We have investigated the rate of convergence of the PMF for the insertion of the cationic side chain analog of arginine in a lipid bilayer using US MD simulations. The addition of Hamiltonian exchange in US-VREX simulations enhanced the statistical convergence of the standard binding free energy by a factor of 3. However, the magnitude and shape of the free energy barrier to the solute crossing the bilayer center remained unconverged after 1.2 μs of simulation per umbrella, both with and without Hamiltonian exchange. To identify regions of the US order parameter that remained susceptible to systematic sampling errors on the simulation time scale, we analyzed the transmission factor of US-VREX replicas through predetermined regions of the order parameter. There is a hidden sampling barrier at the bilayer center, which is caused by rare conformational transitions between lipid defects in either the upper or the lower leaflet, or both. Using the transmission factor, we estimated that this hidden barrier leads to autocorrelation times on the order of 10 μs. We then verified this prediction with lengthy US simulations. Therefore, not only is US-VREX more statistically efficient than standard US but it also allows the detection of correlation times that exceed the simulation time scale as well as their quantitative

prediction—something that, to our knowledge, has never been achieved before. Importantly, the approach introduced here, which consists of computing transmission factors from simulations with Hamiltonian exchange, is general and can be applied to other systems with different order parameters.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Six additional figures and one movie are available. This information is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

*Tel.: 1-416-813-5686. Fax: 1-416-813-5022. E-mail: pomes@sickkids.ca.

Present Addresses

§Arches Computing Systems, 260 Spadina Avenue, Suite 301, Toronto, Ontario, M5T 2E4, Canada
∥Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Yeagle, P. L. *The membranes of cells*; 2nd ed.; Academic Press, Inc.: San Diego, CA, 1993.
(2) Alper, H. E.; Stouch, T. R. Orientation and diffusion of a drug analog in biomembranes: molecular dynamics simulations. *J. Phys. Chem.* **1995**, *99*, 5724−5731.
(3) Bassolino-Klimas, D.; Alper, H. E.; Stouch, T. R. Mechanism of solute diffusion through lipid bilayer membranes by molecular dynamics simulation. *J. Am. Chem. Soc.* **1995**, *117*, 4118−4129.
(4) Tu, K.; Tarek, M.; Klein, M. L.; Scharf, D. Effects of anesthetics on the structure of a phospholipid bilayer: molecular dynamics investigation of halothane in the hydrated liquid crystal phase of dipalmitoylphosphatidylcholine. *Biophys. J.* **1998**, *75*, 2123−2134.
(5) Pitman, M. C.; Suits, F.; MacKerell, A. D., Jr.; Feller, S. E. Molecular-level organization of saturated and polyunsaturated fatty acids in a phosphatidylcholine bilayer containing cholesterol. *Biochemistry* **2004**, *43*, 15318−15328.
(6) Mukhopadhyay, P.; Vogel, H. J.; Tieleman, D. P. Distribution of pentachlorophenol in phospholipid bilayers: a molecular dynamics study. *Biophys. J.* **2004**, *86*, 337−345.

(7) Bemporad, D.; Luttmann, C.; Essex, J. W. Behaviour of small solutes and large drugs in a lipid bilayer from computer simulations. *Biochim. Biophys. Acta, Biomembr.* **2005**, *1718*, 1−21.

(8) Norman, K. E.; Nymeyer, H. Indole localization in lipid membranes revealed by molecular simulation. *Biophys. J.* **2006**, *91*, 2046−2054.

(9) Marrink, S.-J.; Berendsen, H. J. C. Simulation of water transport through a lipid membrane. *J. Phys. Chem.* **1994**, *98*, 4155−4168.

(10) Shinoda, W.; Mikami, M.; Baba, T.; Hato, M. Molecular dynamics study on the effects of chain branching on the physical properties of lipid bilayers. 2. Permeability. *J. Phys. Chem. B* **2004**, *108*, 9346−9356.

(11) Rauscher, S.; Neale, C.; Pomès, R. Simulated tempering distributed replica sampling, virtual replica exchange, and other generalized-ensemble methods for conformational sampling. *J. Chem. Theory Comput.* **2009**, *5*, 2640−2662.

(12) Grossfield, A.; Feller, S. E.; Pitman, M. C. Convergence of molecular dynamics simulations of membrane proteins. *Proteins* **2007**, *67*, 31−40.

(13) Neale, C.; Bennett, W. F. D.; Tieleman, D. P.; Pomès, R. Statistical convergence of equilibrium properties in simulations of molecular solutes embedded in lipid bilayers. *J. Chem. Theory Comput.* **2011**, *7*, 4175−4188.

(14) Zuckerman, D. M. Equilibrium sampling in biomolecular simulations. *Annu. Rev. Biophys.* **2011**, *40*, 41−62.

(15) van Gunsteren, W. F.; Dolenc, J.; Mark, A. E. Molecular simulation as an aid to experimentalists. *Curr. Opin. Struct. Biol.* **2008**, *18*, 149−153.

(16) Adcock, S. A.; McCammon, J. A. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.* **2006**, *106*, 1589−1615.

(17) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(18) Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **1995**, *91*, 275−282.

(19) Khavrutskii, I. V.; Gorfe, A. A.; Lu, B.; McCammon, J. A. Free energy for the permeation of Na+ and Cl- ions and their ion-pair through a zwitterionic dimyristoyl phosphatidylcholine lipid bilayer by umbrella integration with harmonic fourier beads. *J. Am. Chem. Soc.* **2009**, *131*, 1706−1716.

(20) MacCallum, J. L.; Tieleman, D. P. Chapter 8 Interactions between small molecules and lipid bilayers. In *Curr. Top. Membr.*; Scott, E. F., Ed.; Academic Press: New York, 2008; Vol. *60*, pp 227−256.

(21) Tepper, H. L.; Voth, G. A. Mechanisms of passive ion permeation through lipid bilayers: insights from simulations. *J. Phys. Chem. B* **2006**, *110*, 21327−21337.

(22) Tepper, H. L.; Voth, G. A. Protons may leak through pure lipid bilayers via a concerted mechanism. *Biophys. J.* **2005**, *88*, 3095−3108.

(23) Wilson, M. A.; Pohorille, A. Mechanism of unassisted ion transport across membrane bilayers. *J. Am. Chem. Soc.* **1996**, *118*, 6580−6587.

(24) Gambu, I.; Roux, B. Interaction of K+ with a phospholipid bilayer: a molecular dynamics study. *J. Phys. Chem. B* **1997**, *101*, 6066−6072.

(25) Zahn, D.; Brickmann, J. Molecular dynamics study of water pores in a phospholipid bilayer. *Chem. Phys. Lett.* **2002**, *352*, 441−446.

(26) Jambeck, J. P. M.; Lyubartsev, A. P. Implicit inclusion of atomic polarization in modeling of partitioning between water and lipid bilayers. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4677−4686.

(27) Wennberg, C. L.; van der Spoel, D.; Hub, J. S. Large influence of cholesterol on solute partitioning into lipid membranes. *J. Am. Chem. Soc.* **2012**, *134*, 5351−5361.

(28) MacCallum, J. L.; Tieleman, D. P. Computer simulation of the distribution of hexane in a lipid bilayer: spatially resolved free energy, entropy, and enthalpy profiles. *J. Am. Chem. Soc.* **2006**, *128*, 125−130.

(29) Marrink, S.-J.; Berendsen, H. J. C. Permeation process of small molecules across lipid membranes studied by molecular dynamics simulations. *J. Phys. Chem.* **1996**, *100*, 16729−16738.

(30) Vorobyov, I.; Bennett, W. F. D.; Tieleman, D. P.; Allen, T. W.; Noskov, S. The role of atomic polarization in the thermodynamics of chloroform partitioning to lipid bilayers. *J. Chem. Theory Comput.* **2012**, *8*, 618−628.

(31) Samanta, S.; Hezaveh, S.; Milano, G.; Roccatano, D. Diffusion of 1,2-dimethoxyethane and 1,2-dimethoxypropane through phosphatidycholine bilayers: a molecular dynamics study. *J. Phys. Chem. B* **2012**, *116*, 5141−5151.

(32) Pohorille, A.; Cieplak, P.; Wilson, M. A. Interactions of anesthetics with the membrane-water interface. *Chem. Phys.* **1996**, *204*, 337−345.

(33) Kim, E. B.; Lockwood, N.; Chopra, M.; Guzmán, O.; Abbott, N. L.; de Pablo, J. J. Interactions of liquid crystal-forming molecules with phospholipid bilayers studied by molecular dynamics simulations. *Biophys. J.* **2005**, *89*, 3141−3158.

(34) Tian, J.; Sethi, A.; Swanson, B. I.; Goldstein, B.; Gnanakaran, S. Taste of sugar at the membrane: thermodynamics and kinetics of the interaction of a disaccharide with lipid bilayers. *Biophys. J.* **2013**, *104*, 622−632.

(35) Tieleman, D. P.; Marrink, S.-J. Lipids out of equilibrium: energetics of desorption and pore mediated flip-flop. *J. Am. Chem. Soc.* **2006**, *128*, 12462−12467.

(36) Xing, C.; Faller, R. Density imbalances and free energy of lipid transfer in supported lipid bilayers. *J. Chem. Phys.* **2009**, *131*, 175104.

(37) Sapay, N.; Bennett, W. F. D.; Tieleman, D. P. Thermodynamics of flip-flop and desorption for a systematic series of phosphatidylcholine lipids. *Soft Matter* **2009**, *5*, 3295−3302.

(38) Bennett, W. F. D.; MacCallum, J. L.; Tieleman, D. P. Thermodynamic analysis of the effect of cholesterol on dipalmitoylphosphatidylcholine lipid membranes. *J. Am. Chem. Soc.* **2009**, *131*, 1972−1978.

(39) Bennett, W. F. D.; Tieleman, D. P. Molecular simulation of rapid translocation of cholesterol, diacylglycerol, and ceramide in model raft and nonraft membranes. *J. Lipid Res.* **2012**, *53*, 421−429.

(40) Coll, E. P.; Kandt, C.; Bird, D. A.; Samuels, A. L.; Tieleman, D. P. The distribution and conformation of very long-chain plant wax components in a lipid bilayer. *J. Phys. Chem. B* **2007**, *111*, 8702−8704.

(41) Podloucká, P.; Berka, K.; Fabre, G.; Paloncýová, M.; Duroux, J.-L.; Otyepka, M.; Trouillas, P. Lipid bilayer membrane affinity rationalizes inhibition of lipid peroxidation by a natural lignan antioxidant. *J. Phys. Chem. B* **2013**, *117*, 5043−5049.

(42) Zhang, Z.; Lu, L.; Berkowitz, M. L. Energetics of cholesterol transfer between lipid bilayers. *J. Phys. Chem. B* **2008**, *112*, 3807−3811.

(43) Jo, S.; Rui, H.; Lim, J. B.; Klauda, J. B.; Im, W. Cholesterol flip-flop: insights from free energy simulation studies. *J. Phys. Chem. B* **2010**, *114*, 13342−13348.

(44) Bennett, W. F. D.; MacCallum, J. L.; Hinner, M. J.; Marrink, S.-J.; Tieleman, D. P. Molecular view of cholesterol flip-flop and chemical potential in different membrane environments. *J. Am. Chem. Soc.* **2009**, *131*, 12714−12720.

(45) Vijayan, R.; Biggin, P. C. A steroid in a lipid bilayer: localization, orientation, and energetics. *Biophys. J.* **2008**, *95*, L45−L47.

(46) Ulander, J.; Haymet, A. D. J. Permeation across hydrated DPPC lipid bilayers: simulation of the titrable amphiphilic drug valproic acid. *Biophys. J.* **2003**, *85*, 3475−3484.

(47) Paloncýová, M.; Berka, K.; Otyepka, M. Molecular insight into affinities of drugs and their metabolites to lipid bilayers. *J. Phys. Chem. B* **2013**, *117*, 2403−2410.

(48) Chew, C. F.; Guy, A.; Biggin, P. C. Distribution and dynamics of adamantanes in a lipid bilayer. *Biophys. J.* **2008**, *95*, 5627−5636.

(49) Boggara, M. B.; Krishnamoorti, R. Partitioning of nonsteroidal antiinflammatory drugs in lipid membranes: a molecular dynamics simulation study. *Biophys. J.* **2010**, *98*, 586−595.

(50) Meng, F.; Xu, W. Drug permeability prediction using PMF method. *J. Mol. Model.* **2013**, *19*, 991−997.

(51) Cramariuc, O.; Rog, T.; Javanainen, M.; Monticelli, L.; Polishchuk, A. V.; Vattulainen, I. Mechanism for translocation of fluoroquinolones across lipid membranes. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818*, 2563−2571.

(52) Peters, G. H.; Wang, C.; Cruys-Bagger, N.; Velardez, G. F.; Madsen, J. J.; Westh, P. Binding of serotonin to lipid membranes. *J. Am. Chem. Soc.* **2013**, *135*, 2164−2171.

(53) Karlsson, B. C. G.; Olsson, G. D.; Friedman, R.; Rosengren, A. M.; Henschel, H.; Nicholls, I. A. How warfarin's structural diversity influences its phospholipid bilayer membrane permeation. *J. Phys. Chem. B* **2013**, *117*, 2384−2395.

(54) Tejwani, R. W.; Davis, M. E.; Anderson, B. D.; Stouch, T. R. Functional group dependence of solute partitioning to various locations within a DOPC bilayer: a comparison of molecular dynamics simulations with experiment. *J. Pharm. Sci.* **2011**, *100*, 2136−2146.

(55) Paloncýová, M.; Berka, K.; Otyepka, M. Convergence of free energy profile of coumarin in lipid bilayer. *J. Chem. Theory Comput.* **2012**, *8*, 1200−1211.

(56) Hinner, M. J.; Marrink, S.-J.; de Vries, A. H. Location, tilt, and binding: a molecular dynamics study of voltage-sensitive dyes in biomembranes. *J. Phys. Chem. B* **2009**, *113*, 15807−15819.

(57) Kyrychenko, A.; Wu, F.; Thummel, R. P.; Waluk, J.; Ladokhin, A. S. Partitioning and localization of environment-sensitive 2-(2'-pyridyl)- and 2-(2'-pyrimidyl)-indoles in lipid membranes: a joint refinement using fluorescence measurements and molecular dynamics simulations. *J. Phys. Chem. B* **2010**, *114*, 13574−13584.

(58) Košinová, P.; Berka, K.; Wykes, M.; Otyepka, M.; Trouillas, P. Positioning of antioxidant quercetin and its metabolites in lipid bilayer membranes: implication for their lipid-peroxidation inhibition. *J. Phys. Chem. B* **2012**, *116*, 1309−1318.

(59) Kelly, C. V.; Leroueil, P. R.; Nett, E. K.; Wereszczynski, J. M.; Baker, J. R., Jr.; Orr, B. G.; Banaszak Holl, M. M.; Andricioaei, I. Poly(amidoamine) dendrimers on lipid bilayers I: free energy and conformation of binding. *J. Phys. Chem. B* **2008**, *112*, 9337−9345.

(60) Widge, A. S.; Matsuoka, Y.; Kurnikova, M. Computational modeling of poly(alkylthiophene) conductive polymer insertion into phospholipid bilayers. *Langmuir* **2007**, *23*, 10672−10681.

(61) Choe, S.; Chang, R.; Jeon, J.; Violi, A. Molecular dynamics simulation study of a pulmonary surfactant film interacting with a carbonaceous nanoparticle. *Biophys. J.* **2008**, *95*, 4102−4114.

(62) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. Distribution of amino acids in a lipid bilayer from computer simulations. *Biophys. J.* **2008**, *94*, 3393−3404.

(63) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. Partitioning of amino acid side chains into lipid bilayers: results from computer simulations and comparison to experiment. *J. Gen. Physiol.* **2007**, *129*, 371−377.

(64) Grossfield, A.; Woolf, T. B. Interaction of tryptophan analogs with POPC lipid bilayers investigated by molecular dynamics calculations. *Langmuir* **2002**, *18*, 198−210.

(65) Vorobyov, I.; Li, L.; Allen, T. W. Assessing atomistic and coarse-grained force fields for protein−lipid interactions: the formidable challenge of an ionizable side chain in a membrane. *J. Phys. Chem. B* **2008**, *112*, 9588−9602.

(66) Li, L.; Vorobyov, I.; Allen, T. Potential of mean force and pKa profile calculation for a lipid membrane-exposed arginine side chain. *J. Phys. Chem. B* **2008**, *112*, 9574−9587.

(67) Johansson, A. C. V.; Lindahl, E. Position-resolved free energy of solvation for amino acids in lipid membranes from molecular dynamics simulations. *Proteins* **2008**, *70*, 1332−1344.

(68) Johansson, A. C. V.; Lindahl, E. Correction to the article "Position-resolved free energy of solvation for amino acids in lipid membranes from molecular dynamics simulations". *Proteins* **2008**, *70*, 1655−1656.

(69) Ou, S.; Lucas, T. R.; Zhong, Y.; Bauer, B. A.; Hu, Y.; Patel, S. Free energetics and the role of water in the permeation of methyl guanidinium across the bilayer-water interface: insights from molecular dynamics simulations using charge equilibration potentials. *J. Phys. Chem. B* **2013**, *117*, 3578−3592.

(70) Cardenas, A. E.; Jas, G. S.; DeLeon, K. Y.; Hegefeld, W. A.; Kuczera, K.; Elber, R. Unassisted transport of N-acetyl-l-tryptophanamide through membrane: experiment and simulation of kinetics. *J. Phys. Chem. B* **2012**, *116*, 2739−2750.

(71) Babakhani, A.; Gorfe, A. A.; Kim, J. E.; McCammon, J. A. Thermodynamics of peptide insertion and aggregation in a lipid bilayer. *J. Phys. Chem. B* **2008**, *112*, 10528−10534.

(72) Huang, K.; García, A. E. Free energy of translocating an arginine-rich cell-penetrating peptide across a lipid bilayer suggests pore formation. *Biophys. J.* **2013**, *104*, 412−420.

(73) Mondal, J.; Zhu, X.; Cui, Q.; Yethiraj, A. Sequence-dependent interaction of *β*-peptides with membranes. *J. Phys. Chem. B* **2010**, *114*, 13585−13592.

(74) Yeh, I.-C.; Ripoll, D. R.; Wallqvist, A. Free energy difference in indolicidin attraction to eukaryotic and prokaryotic model cell membranes. *J. Phys. Chem. B* **2012**, *116*, 3387−3396.

(75) Vivcharuk, V.; Kaznessis, Y. Free energy profile of the interaction between a monomer or a dimer of protegrin-1 in a specific binding orientation and a model lipid bilayer. *J. Phys. Chem. B* **2010**, *114*, 2790−2797.

(76) Irudayam, S. J.; Berkowitz, M. L. Binding and reorientation of melittin in a POPC bilayer: computer simulations. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818*, 2975−2981.

(77) Kim, T.; Im, W. Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation. *Biophys. J.* **2010**, *99*, 175−183.

(78) Chen, R.; Chung, S.-H. Effect of gating modifier toxins on membrane thickness: implications for toxin effect on gramicidin and mechanosensitive channels. *Toxins* **2013**, *5*, 456−471.

(79) Davis, C. H.; Berkowitz, M. L. Interaction between amyloid-*β* (1−42) peptide and phospholipid bilayers: a molecular dynamics study. *Biophys. J.* **2009**, *96*, 785−797.

(80) de Jesus, A. J.; Allen, T. W. The role of tryptophan side chains in membrane protein anchoring and hydrophobic mismatch. *Biochim. Biophys. Acta, Biomembr.* **2013**, *1828*, 864−876.

(81) Dorairaj, S.; Allen, T. W. On the thermodynamic stability of a charged arginine side chain in a transmembrane helix. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 4943−4948.

(82) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **2000**, *113*, 6042−6051.

(83) Neale, C.; Rodinger, T.; Pomès, R. Equilibrium exchange enhances the convergence rate of umbrella sampling. *Chem. Phys. Lett.* **2008**, *460*, 375−381.

(84) Park, S.; Kim, T.; Im, W. Transmembrane helix assembly by window exchange umbrella sampling. *Phys. Rev. Lett.* **2012**, *108*, 108102.

(85) Wolf, M. G.; Jongejan, J. A.; Laman, J. D.; de Leeuw, S. W. Rapid free energy calculation of peptide self-assembly by REMD umbrella sampling. *J. Phys. Chem. B* **2008**, *112*, 13493−13498.

(86) Periole, X.; Mark, A. E. Convergence and sampling efficiency in replica exchange simulations of peptide folding in explicit solvent. *J. Chem. Phys.* **2007**, *126*, 014903.

(87) Ferrenberg, A. M.; Swendsen, R. H. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.* **1988**, *61*, 2635−2638.

(88) Hansmann, U. H. E. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* **1997**, *281*, 140−150.

(89) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(90) Tesi, M. C.; van Rensburg, E. J. J.; Orlandini, E.; Whittington, S. G. Monte carlo study of the interacting self-avoiding walk model in three dimensions. *J. Stat. Phys.* **1996**, *82*, 155−181.

(91) Hukushima, K.; Nemoto, K. Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604−1608.

(92) Zuckerman, D. M.; Lyman, E. A Second Look at Canonical Sampling of Biomolecules Using Replica Exchange Simulation. *J. Chem. Theory Comput.* **2006**, *2*, 1200−1202.

(93) Zuckerman, D. M.; Lyman, E. A Second Look at Canonical Sampling of Biomolecules Using Replica Exchange Simulation [*J. Chem. Theory Comput.* 2, 1200−1202 (2006)]. *J. Chem. Theory Comput.* **2006**, *2*, 1693−1693.

(94) Dalbey, R. E. Positively charged residues are important determinants of membrane protein topology. *Trends Biochem. Sci.* **1990**, *15*, 253−257.

(95) Boyd, D.; Beckwith, J. The role of charged amino acids in the localization of secreted and membrane proteins. *Cell* **1990**, *62*, 1031−1033.

(96) Boyd, D.; Beckwith, J. Positively charged amino acid residues can act as topogenic determinants in membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 9446−9450.

(97) Hristova, K.; Wimley, W. C. A look at arginine in membranes. *J. Membr. Biol.* **2011**, *239*, 49−56.

(98) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.

(99) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(100) Berger, O.; Edholm, O.; Jähnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **1997**, *72*, 2002−2013.

(101) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(102) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(103) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **1999**, *20*, 786−798.

(104) Chakrabarti, N.; Neale, C.; Payandeh, J.; Pai, E. F.; Pomès, R. An iris-like mechanism of pore dilation in the CorA magnesium transport system. *Biophys. J.* **2010**, *98*, 784−792.

(105) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952−962.

(106) Hess, B. P-LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116−122.

(107) Tieleman, D. P.; Sansom, M. S. P.; Berendsen, H. J. C. Alamethicin helices in a bilayer and in solution: molecular dynamics simulations. *Biophys. J.* **1999**, *76*, 40−49.

(108) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N•log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(109) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(110) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(111) van Gunsteren, W. F.; Berendsen, H. J. C. A leap-frog algorithm for stochastic dynamics. *Mol. Sim.* **1988**, *1*, 173−185.

(112) Kandt, C.; Ash, W. L.; Tieleman, D. P. Setting up and running molecular dynamics simulations of membrane proteins. *Methods* **2007**, *41*, 475−488.

(113) Grossfield, A. WHAM: the weighted histogram analysis method. http://membrane.urmc.rochester.edu/content/wham (accessed August 3, 2011).

(114) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(115) Ben-Tal, N.; Ben-Shaul, A.; Nicholls, A.; Honig, B. Free-energy determinants of alpha-helix insertion into lipid bilayers. *Biophys. J.* **1996**, *70*, 1803−1812.

(116) General, I. J. A note on the standard state's binding free energy. *J. Chem. Theory Comput.* **2010**, *6*, 2520−2524.

(117) Flyvbjerg, H.; Petersen, H. G. Error estimates on averages of correlated data. *J. Chem. Phys.* **1989**, *91*, 461−466.

(118) Rodinger, T.; Howell, P. L.; Pomès, R. Distributed replica sampling. *J. Chem. Theory Comput.* **2006**, *2*, 725−731.

(119) Janke, W. Statistical analysis of simulations: data correlations and error estimation. In *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*; Grotendorst, J., Marx, D., Murmatsu, A., Eds.; John von Neumann Institute for Computing: Jülich, Germany, 2002; Vol. 10, pp 423−445.

(120) Romo, T. D.; Grossfield, A. Block Covariance Overlap Method and Convergence in Molecular Dynamics Simulation. *J. Chem. Theory Comput.* **2011**, *7*, 2464−2472.

(121) Pearlman, D. A.; Kollman, P. A. The lag between the Hamiltonian and the system configuration in free energy perturbation calculations. *J. Chem. Phys.* **1989**, *91*, 7831−7839.

(122) Zheng, L.; Chen, M.; Yang, W. Simultaneous escaping of explicit and hidden free energy barriers: application of the orthogonal space random walk strategy in generalized ensemble based conformational sampling. *J. Chem. Phys.* **2009**, *130*, 234105.

(123) Roux, B. Lonely arginine seeks friendly environment. *J. Gen. Physiol.* **2007**, *130*, 233−236.

(124) Li, L.; Vorobyov, I.; MacKerell, A. D.; Allen, T. W. Is arginine charged in a membrane? *Biophys. J.* **2008**, *94*, L11−L13.

(125) Yoo, J.; Cui, Q. Does arginine remain protonated in the lipid membrane? Insights from microscopic pKa calculations. *Biophys. J.* **2008**, *94*, L61−L63.

(126) Lou, H.; Cukier, R. I. Molecular dynamics of apo-adenylate kinase: a distance replica exchange method for the free energy of conformational fluctuations. *J. Phys. Chem. B* **2006**, *110*, 24121−24137.

(127) Jiang, W.; Roux, B. t. Free energy perturbation Hamiltonian replica-exchange molecular dynamics (FEP/H-REMD) for absolute ligand binding free energy calculations. *J. Chem. Theory Comput.* **2010**, *6*, 2559−2565.

(128) Meng, Y.; Sabri Dashti, D.; Roitberg, A. E. Computing alchemical free energy differences with Hamiltonian replica exchange molecular dynamics (H-REMD) simulations. *J. Chem. Theory Comput.* **2011**, *7*, 2721−2727.

(129) Woods, C. J.; Essex, J. W.; King, M. A. The development of replica-exchange-based free-energy methods. *J. Phys. Chem. B* **2003**, *107*, 13703−13710.

(130) Rodinger, T.; Howell, P. L.; Pomès, R. Calculation of absolute protein-ligand binding free energy using distributed replica sampling. *J. Chem. Phys.* **2008**, *129*, 155102−12.

(131) Vreede, J.; Wolf, M. G.; de Leeuw, S. W.; Bolhuis, P. G. Reordering hydrogen bonds using Hamiltonian replica exchange enhances sampling of conformational changes in biomolecular systems. *J. Phys. Chem. B* **2009**, *113*, 6484−6494.

(132) Resat, H.; Mezei, M. Studies on free energy calculations. I. Thermodynamic integration using a polynomial path. *J. Chem. Phys.* **1993**, *99*, 6052−6061.

(133) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. Alchemical free energy calculations: ready for prime time? In *Annu. Rep. Comput. Chem.*; Spellmeyer, D. C., Wheeler, R., Eds.; Elsevier: Amsterdam, 2007; Vol. 3, pp 41−59.

(134) Mitsutake, A.; Sugita, Y.; Okamoto, Y. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Peptide Sci.* **2001**, *60*, 96−123.

(135) Mitsutake, A.; Okamoto, Y. Replica-exchange simulated tempering method for simulations of frustrated systems. *Chem. Phys. Lett.* **2000**, *332*, 131−138.

(136) Yoshida, K.; Yamaguchi, T.; Okamoto, Y. Replica-exchange molecular dynamics simulation of small peptide in water and in ethanol. *Chem. Phys. Lett.* **2005**, *412*, 280−284.

(137) Cecchini, M.; Rao, F.; Seeber, M.; Caflisch, A. Replica exchange molecular dynamics simulations of amyloid peptide aggregation. *J. Chem. Phys.* **2004**, *121*, 10748−10756.

(138) Rao, F.; Caflisch, A. Replica exchange molecular dynamics simulations of reversible folding. *J. Chem. Phys.* **2003**, *119*, 4035−4042.

(139) Nadler, W.; Meinke, J. H.; Hansmann, U. H. E. Folding proteins by first-passage-times-optimized replica exchange. *Phys. Rev. E* **2008**, *78*, 061905.

(140) Kone, A.; Kofke, D. A. Selection of temperature intervals for parallel-tempering simulations. *J. Chem. Phys.* **2005**, *122*, 206101−2.

(141) Trebst, S.; Troyer, M.; Hansmann, U. H. E. Optimized parallel tempering simulations of proteins. *J. Chem. Phys.* **2006**, *124*, 174903−6.

(142) Wyczalkowski, M. A.; Pappu, R. V. Satisfying the fluctuation theorem in free-energy calculations with Hamiltonian replica exchange. *Phys. Rev. E* **2008**, *77*, 026104.

(143) Nymeyer, H.; Woolf, T. B.; García, A. E. Folding is not required for bilayer insertion: replica exchange simulations of an alpha-helical peptide with an explicit lipid bilayer. *Proteins* **2005**, *59*, 783−790.

(144) Dolan, E. A.; Venable, R. M.; Pastor, R. W.; Brooks, B. R. Simulations of membranes and other interfacial systems using P21 and Pc periodic boundary conditions. *Biophys. J.* **2002**, *82*, 2317−2325.

(145) Loken, C.; Gruner, D.; Groer, L.; Peltier, R.; Bunn, N.; Craig, M.; Henriques, T.; Dempsey, J.; Yu, C.-H.; Chen, J.; Dursi, L. J.; Chong, J.; Northrup, S.; Pinto, J.; Knecht, N.; Van Zon, R. SciNet: lessons learned from building a power-efficient top-20 system and data centre. *J. Phys. Conf. Ser.* **2010**, *256*, 012026.

(146) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(147) Williams, T.; Kelly, C. gnuplot 4.4: in interactive plotting program. http://gnuplot.sourceforge.net/ (accessed March 20, 2012).