Generalized Langevin models of molecular dynamics simulations with applications to ion channels

Dan Gordon,[1,a)] Vikram Krishnamurthy,[2] and Shin-Ho Chung[1]
[1]*Computational Biophysics Group, Research School of Biological Science, The Australian National University, Canberra, Australian Capital Territory 0200, Australia*
[2]*Department of Electrical Engineering, The University of British Columbia, Vancouver, British Columbia, Canada V6T1Z4*

(Received 19 March 2009; accepted 30 August 2009; published online 2 October 2009)

We present a new methodology, which combines molecular dynamics and stochastic dynamics, for modeling the permeation of ions across biological ion channels. Using molecular dynamics, a free energy profile is determined for the ion(s) in the channel, and the distribution of random and frictional forces is measured over discrete segments of the ion channel. The parameters thus determined are used in stochastic dynamics simulations based on the nonlinear generalized Langevin equation. We first provide the theoretical basis of this procedure, which we refer to as "distributional molecular dynamics," and detail the methods for estimating the parameters from molecular dynamics to be used in stochastic dynamics. We test the technique by applying it to study the dynamics of ion permeation across the gramicidin pore. Given the known difficulty in modeling the conduction of ions in gramicidin using classical molecular dynamics, there is a degree of uncertainty regarding the validity of the MD-derived potential of mean force (PMF) for gramicidin. Using our techniques and systematically changing the PMF, we are able to reverse engineer a modified PMF which gives a current-voltage curve closely matching experimental results. © *2009 American Institute of Physics*. [doi:10.1063/1.3233945]

I. INTRODUCTION

Biological ion channels are protein-based structures that span the cell membrane and provide a passage for the controlled transport of ions between the inside and outside of the cell. They are responsible for processes such as the propagation of nerve impulses and muscle contraction. In recent years, x-ray crystallographic structures have become available for ion channels,[1] thus making it possible to develop theoretical models that provide a link between the detailed structure of a channel and its function. Molecular dynamics (MD) is one technique that appears to give the appropriate level of resolution to allow detailed structural modeling of ion channels.

In order to provide a link to experiment, a quantitative, experimentally achievable means is needed to characterize the function of channels. This is usually provided by current-voltage and current-concentration curves, by measuring the inhibitory effect that classical channel blockers such as tetraetylammonium or 4-aminopyridine have on ion currents, by investigating the effect of site-directed mutagenesis on ion currents, or by determining the location of ion binding sites. Such multidimensional data provide a yardstick against which theoretical models can be assessed.

A theoretical model should therefore ideally be capable of simulating a channel long enough to provide reliable current-voltage and current-concentration curves and the like, or a means should be available to extrapolate theoretical results in order to predict currents. In order to simulate the channel long enough to produce a reliable set of simulation results to compare to experimental measurements, we might want to observe several hundred conduction events, which, under physiological conditions for a channel such as the bacterial KcsA potassium channel, translates to a simulation time of around 10–100 μs. Currently, MD simulations of 0.1 μs are about the limit of what is practical for realistic models of ion channels containing tens of thousands of atoms [see Crozier *et al.*[2,3] for an actual example of a MD current calculation on a simplified system consisting of around 1000 atoms or Aksimentiev and Schulten[4] for a calculation of ionic current in a large (nanometer scale) pore]. So, while there is room for optimism about the possibility of using MD to provide a direct and routine link between theory and experiment in the future, there is still some way to go.

To make further progress, there are a number of possibilities. One strategy is to turn to higher level physical theories, such as the Poisson–Nernst–Planck (PNP) theory[5] or Brownian dynamics (BD) (Ref. 6) or to higher level versions of MD, such as implicit solvent MD or coarse-grained MD. Another strategy, which we shall follow in this paper, is to use a two-tiered approach. MD is used to measure various properties of the system, such as free energy and diffusion, and the results of the measurements are fed into a higher level simulation that is able to extend the achievable simulation time horizon beyond what is practical using MD.

This strategy has been applied in the past by several groups. For example, transition state theory assigns discrete states to system configurations and then models the transition rates between these states as a Markov chain, potentially using data from MD.[7] More recently, several authors[8,9] have

[a)]Electronic mail: dan.gordon@anu.edu.au.

extended these ideas to models that describe the motion of the ions as the limit of a random walk, typically using a single spatial dimension and a reaction coordinate of several dimensions. The dynamics are governed by the free energy profile, or potential of mean force (PMF), and the position dependent diffusion coefficient. The motion is assumed to be first-order Markovian (memoryless), and inertial and frictional terms are neglected. The entry and exit of ions to and from the channel are assumed to be a first-order process, and there is some uncertainty regarding the prefactor for the entry and exit rates. These models have been used to describe complex interactions, such as proton transport in gramicidin including a description of the water defect[8] and multi-ion conduction pathways in KcsA,[9] providing a good link to experiment. In other related work, Burykin *et al.*[10] also performed BD simulations on energy surfaces calculated using mesoscopic electrostatics calculations for the KcsA channel.

In a somewhat different, fascinating application, termed potential-of-mean-force-Poisson-Nernst-Planck (PMFPNP) theory, Mamonov *et al.*[5,11] modeled the flow of ions through a gramicidin channel as a mean-field diffusion in a self-consistent potential (such as standard PNP theory) but incorporating an external free energy (PMF) and diffusion constants derived from MD measurements. A novel methodology, combining MD and macroscopic electrostatics, is employed to calculate the PMF,[5,11,12] which, for narrow channels such as gramicidin, may be important, since PMFs calculated directly from MD in such cases tend to suffer from slow convergence[13] and artifacts due to the use of a small, periodic unit cell and nonpolarizable lipids.[14] In a subsequent paper, Mamonov *et al.*[15] investigated methods for calculation of diffusion constants. To deduce the current across the gramicidin pore, Allen *et al.*[14] also employed a continuum model in the form of one-dimensional (1D) Nernst–Planck theory using the PMF and diffusion coefficients calculated from MD simulations.

In this paper, we take as our point of departure the aim to reproduce the distribution of ion trajectories implicit in a MD simulation as closely as possible. Unlike other previous work, this leads us to employ non-Markovian generalized Langevin dynamics in three spatial dimensions rather than, say, Langevin dynamics or BD. Strictly speaking, the Markovian assumption is valid only when the solvent particles are much lighter than the Brownian particles being modeled; obviously this is not true for, say, ions in water. In practice, the assumption often works, but the interpretation of when and why this is the case is open to question. It is known that system memory may in some circumstances lead to different barrier crossing rates and other dynamics,[16,17] and we aim to develop a fairly general theory that will automatically take such considerations into account. The aim of reproducing the distribution of MD trajectories leads us to refer to this approach as "distributional MD." We discuss various approximations and assumptions in the context of this aim.

We model the full three-dimensional channel/reservoir system containing many anions and cations, so that the full dynamics of ion entry and exit from the channel is taken into account, and processes such as ion depletion under conditions of high current will be correctly dealt with. Our model also allows certain aspects of the macroscopic electrostatic calculations used in conventional BD simulations[6] to be incorporated, such as the voltage drop across the channel due to the membrane potential or dielectric mediated ion-ion interactions.

As a first test, we apply these methods to the gramicidin channel. Gramicidin channels are a useful benchmark example since, first, macroscopic approximations break down in channels containing single-file water[18] and second, gramicidin is a one, or at most two, ion channel, which simplifies the analysis. Several simplifying assumptions are made—for example, we assume that the ion obeys a generalized Langevin equation with an exponential friction kernel. We find that the PMF deduced using the CHARMM27 parameters is too high, and hence the conductance is much lower than that seen in experiment. This is perhaps not surprising, given the known difficulty in modeling the energetics of ions in gramicidin using MD and the fact that systematic errors have been identified that artificially raise the PMF.[13] By reverse engineering, we modify the PMF so that experimental results are reproduced. Note, however, that the resulting reverse engineered PMF is not necessarily unique.

II. THEORETICAL FORMULATION

The general procedure we follow comprises of three steps:

(1) A stochastic physical model is derived or assumed for the system, e.g., the generalized Langevin equation.
(2) The parameters governing the evolution of the system variables are estimated from MD simulations.
(3) The stochastic simulation is then carried out by numerically solving a stochastic dynamics equation, making use of the parameters derived in the previous step.

In the following sections, we describe each of these steps.

A. Derivation of a stochastic physical model

In this step, we choose a stochastic equation, such as the Langevin equation or the nonlinear generalized Langevin equation, to model the motion. We shall discuss the theoretical basis for this choice in order to gain some understanding of the validity and limits of the chosen approach.

Our system can be described by an indexed set of phase variables $\mathbf{\Gamma}=\{\Gamma_i\}=\{(\mathcal{Q}_i,\mathcal{P}_i)\}$. We shall use the following notational conventions: for an indexed quantity, such as p_i, boldface will indicate a column vector, so, for example, $p\equiv[p_1,p_2,\ldots,p_N]^T$. Similarly, for a quantity with two indices, the lack of indices will indicate the corresponding matrix. $\nabla_{\mathbf{q}}$ means the vector operator whose components are $\partial_{q_i}\equiv\partial/\partial q_i$. The pair $(\mathcal{Q}_i,\mathcal{P}_i)$ denotes the ith generalized coordinate and momentum, where $i\in[1,\ldots,3N]$ for three spatial dimensions and N particles. For example, here $\mathbf{\Gamma}$ describes the positions and momenta of all the atoms in a MD simulation of an ion channel, consisting of some ions of interest, a background solution, a channel protein, and a lipid membrane.

We are interested in a projection $\mathbf{\Gamma}_s$ that describes, for the ion channel example, the positions and momenta of all

ions of interest. The remaining phase space coordinates are of interest only through their effect on $\mathbf{\Gamma}_s$; they can be considered to be a "heat bath." Thus we partition the phase space into variables of interest, or "system" variables and "bath" variables,

$$\mathbf{\Gamma} = \mathbf{\Gamma}_s \times \mathbf{\Gamma}_b, \tag{1}$$

where

$$\mathbf{\Gamma}_s = \{\Gamma_{s,i}\} = \{(q_i, p_i)\}, \quad i = 1, \ldots, N_s, \tag{2}$$

are the system phase variables and

$$\mathbf{\Gamma}_b = \{\Gamma_{b,i}\} = \{(Q_i, P_i)\}, \quad i = 1, \ldots, N_b, \tag{3}$$

are the bath phase variables. Naturally, $N_s + N_b = 3N$.

We show that, under reasonable assumptions, the equations of motion can be expressed as a nonlinear (due to the presence of a position dependent force) generalized Langevin equation,[19–21] obeying a generalized fluctuation-dissipation theorem.[21] Our starting point is the fact that MD obeys Newton's laws of motion, and therefore will obey certain physical constraints. We wish to build these constraints into the model rather than simply hope that we have sampled the MD distribution well enough to ensure that they are satisfied. Using the partitioning into system and bath spaces, we can write the Hamiltonian as

$$H = H_s + H_b + H_{sb}, \tag{4}$$

where H_s describes terms arising purely from the system variables, H_b describes terms arising purely from the bath variables, and H_{sb} describes the interaction between the system and the bath.

$$H_s = \tfrac{1}{2}\mathbf{p}^{\mathrm{T}} m^{-1} \mathbf{p} + U_s(\mathbf{q}), \tag{5}$$

$$H_b = \tfrac{1}{2}\mathbf{P}^{\mathrm{T}} M^{-1} \mathbf{P} + U_b(\mathbf{Q}), \tag{6}$$

$$H_{sb} = U_{sb}(\mathbf{q}, \mathbf{Q}). \tag{7}$$

Here we define m and M to be the system and bath mass tensors; in Cartesian coordinates these are diagonal, with, for example m_{ii} being the mass of the ith particle.

Also we define the forces

$$\mathbf{F}_s^s = -\nabla_{\mathbf{q}} U_s(\mathbf{q}), \tag{8}$$

$$\mathbf{F}_b^b = -\nabla_{\mathbf{Q}} U_b(\mathbf{Q}), \tag{9}$$

$$\mathbf{F}_{sb}^s = -\nabla_{\mathbf{q}} U_{sb}(\mathbf{q}, \mathbf{Q}), \tag{10}$$

$$\mathbf{F}_{sb}^b = -\nabla_{\mathbf{Q}} U_{sb}(\mathbf{q}, \mathbf{Q}). \tag{11}$$

These are, respectively, the force exerted on the system by U_s, the force exerted on the bath by U_b, the force exerted on the system by U_{sb}, and the force exerted on the bath by U_{sb}.

1. The potential of mean force

At this point, we will reassign the force terms. Define the PMF (Ref. 22) as the potential whose gradient $\mathbf{F}_{\mathrm{PMF}}$ is the equilibrium average of $\mathbf{F}_{sb}^s$ at a given value of $\mathbf{q}$. As this is a function only of $\mathbf{q}$, we subtract it from $\mathbf{F}_{sb}$ and add it onto $\mathbf{F}_s$. Thus,

$$\mathbf{F}_{\mathrm{RF}} \equiv \mathbf{F}_{sb}^s - \mathbf{F}_{\mathrm{PMF}}, \tag{12}$$

$$\mathbf{F}_D \equiv \mathbf{F}_s^s + \mathbf{F}_{\mathrm{PMF}}, \tag{13}$$

where the subscripts RF and D denote, respectively, "random plus frictional" and "deterministic." The total force on the system variables can now be written as

$$\mathbf{F}_{\mathrm{tot}}^s = \mathbf{F}_D + \mathbf{F}_{\mathrm{RF}}. \tag{14}$$

2. The generalized Langevin equation

In much of the literature,[23] it is assumed that the evolution of the projection of the system plus bath onto the system variables can adequately be described by the generalized Langevin equation. It consists of (i) Newton's laws of motion for the system variables, (ii) a frictional force, multilinear in the velocity as a function of previous times, and (iii) a random force, related to the frictional force through a fluctuation-dissipation theorem. Often, one goes further and assumes that the friction depends only on the velocity at the present time, therefore giving the Langevin equation.

The existence of a nonlinear position dependent force $\mathbf{F}_D(\mathbf{q})$ means that we require a nonlinear generalized Langevin equation rather than the usual generalized Langevin equation which does not adequately describe any significant deviations from equilibrium. The nonlinear generalized Langevin equation is

$$\begin{aligned} \partial_t \mathbf{q}(t) &= m^{-1}\mathbf{p}(t), \\ \partial_t \mathbf{p}(t) &= \mathbf{F}_D(\mathbf{q}(t)) - \int_0^t dt' K(t')\mathbf{p}(t-t') + \mathbf{F}_R(t), \end{aligned} \tag{15}$$

where the sum of the second and third terms in Eq. (15) corresponds to $\mathbf{F}_{\mathrm{RF}}$ defined in Eq. (12) above. Apart from the presence of the nonlinear force term $\mathbf{F}_D(\mathbf{q})$, this equation is the same as the regular (linear) Langevin equation.[24,25] The following fluctuation-dissipation theorem is also assumed:[19,20]

$$\langle \mathbf{F}_R(0)\mathbf{F}_R^T(t)\rangle = kTK(t)m. \tag{16}$$

Here, the function $K(t)$ denotes the friction kernel and is intrinsic to the system. Finally, we assume that $\mathbf{F}_R$ is a Gaussian random process.[20] This can be empirically verified from MD simulations using the Kolmogorov–Smirnov or Andersen–Darling tests on the joint distribution over multiple time periods (see Sec. IV).

There is no theoretical guarantee that Eq. (15) is a reasonable approximation to the system motion, although this seems to be a reasonable assumption. Zwanzig[19,20] shows that this nonlinear generalized Langevin equation can be derived from first principles if we assume a bath potential that is quadratic and a linear system-bath coupling. Cicotti and Ryckaert[21] have analyzed the case of general (not necessarily quadratic) system-bath coupling. They find that Eq. (15) and associated fluctuation-dissipation results do not hold in gen-

eral. Instead, they derive a more complicated nonlinear generalized Langevin equation and associated fluctuation-dissipation theorems. We have applied this equation to MD simulations but found it to be numerically less stable and generally harder to interpret. In this paper, we use the simpler nonlinear generalized Langevin equation Eq. (15).

3. Relationship between the friction kernel and the momentum autocorrelation function

By multiplying Eq. (15) on the right by $\mathbf{p}^T(0)$ and taking an ensemble average [conditioned on $\mathbf{q}(t=0)$], we can derive an equation for the momentum autocorrelation function $C(t) \equiv \langle \mathbf{p}(t)\mathbf{p}^T(0)\rangle$:

$$\partial_t C = \langle \mathbf{F}_D(t)\mathbf{p}^T(0)\rangle - \int_0^t K(t-t')C(t')dt'. \tag{17}$$

Note that the correlation function on the right hand side of this equation can be estimated from the simulation data provided the PMF is known.

If we approximate $\mathbf{F}_D$ by a harmonic potential $U(\mathbf{q}) = k(\mathbf{q}-\mathbf{q}_0)^2/2$, as was done by Straub *et al.*,[23] then Eq. (17) simplifies to

$$\begin{aligned}\partial_t C &= -\int_0^t \left(K(t-t') + \frac{k}{m}\right)C(t')dt' \\ &= -\int_0^t K'(t-t')C(t')dt',\end{aligned} \tag{18}$$

where $K' = K + k/m$. Because, in realistic systems, $K(t)$ should decay to zero as $t \to \infty$, k/m can be estimated as the limit of K'. Therefore, if we make this harmonic approximation, it is not necessary to have knowledge of the PMF in order to derive K.

Equations (17) and (18) are Volterra equations. Given C, they can be solved for the friction kernel $K(t)$ by numerical integration.[26] Alternatively, an explicit expression for $K(t)$ can be derived using Laplace transform methods,[14,15,27] but this method, while being theoretically neater, tends to be numerically unstable in practice.[28]

To summarize, using data derived from MD simulation, we calculate $C(t)$. Solving Eq. (18) then yields the friction kernel $K(t)$. This, along with the PMF, enables us to fully characterize the nonlinear generalized Langevin equation, given the additional assumptions that fluctuation-dissipation theorem equation (16) holds and that $\mathbf{F}_R$ is Gaussian.

III. METHODS

A. Estimating the parameters of the generalized Langevin equation using molecular dynamics simulations

In this section, we will discuss estimation of the parameters of the generalized Langevin equation. As we have just showed, these are as follows:

(1) $\mathbf{F}_D(\mathbf{q})$, the deterministic force (gradient of the PMF +system force), and
(2) $K(t)$, the friction kernel.

It should be noted that the accuracy of the measured parameters will depend, among other factors, on the degree to which the system attains stationarity (ergodicity) over the measurement time horizon. It is an interesting question what the timescale of ergodicity is, especially for single-file water channels such as the gramicidin channel considered later as a numerical example.

1. Estimating the potential of mean force

The process of estimating the PMF has been extensively discussed in the literature.[22,29–32] Single particle PMFs can be 1D functions (e.g., the order parameter is z, the distance along the axis of the channel), two-dimensional functions [with order parameters (z,r)] for channels that are approximately radially symmetric, or three-dimensional functions (z,r,θ). In addition, we may have to consider the joint PMF for two or more particles—so, for example, we would have a four-dimensional order parameter $(\mathbf{q}_1, r_1, \mathbf{q}_2, r_2)$ for two particles where cylindrical symmetry is assumed.

So long as the nonlinearity is not too high, in the case where there are more than two ions in the channel, we need only calculate a two-ion PMF, since the force on one ion due to all the others can be expressed as a sum of ion-ion interactions. So at worst, we can expect to have to calculate a six-dimensional PMF (three spatial dimensions and two ions). However, even this can be computationally prohibitive. In this paper, we use a three-dimensional, radially symmetric PMF for a single ion. This choice is reasonable for a narrow channel such as gramicidin. Ion-ion interactions are modeled via force calculations derived from macroscopic electrostatics, although in future work, we intend to make two-ion PMF measurements.

In order to calculate the gramicidin PMF, we employ umbrella sampling techniques. For details see Sec. III C below. An interesting alternative that may overcome some of the problems of slow convergence and simulation artifacts of MD combines data taken from MD simulations with macroscopic electrostatics.[5,11,12]

2. Estimating the frictional and random forces from molecular dynamics simulation

As discussed above, the generalized Langevin equation and the fluctuation-dissipation theorem imply that both the frictional and random forces can be calculated from the friction kernel $K(t)$. Our problem is then how to sample $K(t)$ in cases where the starting point is far from (system) equilibrium. For example, for an ion channel, how do we sample paths that start at or near the top of the potential barrier given that the ion will rarely be observed in such a position?

We discuss two methods: "harmonic bias potential" and "clamp and release." The harmonic bias potential method is computationally more efficient but relies on modifying the system potential in a way that may lead to systematic error in the results. The clamp and release method is computationally more expensive and prone to methodological errors but is more rigorous from a theoretical point of view. In the numerical studies on the gramicidin channel presented in this

paper, we used the harmonic bias potential method, although we have also performed some tests using clamp and release.

3. Estimating the friction kernel $K(t)$ using a harmonic bias potential

In this method, a harmonic potential is applied to the ion, centered at the point in the channel at which the friction kernel is to be measured and using a spring constant strong enough to prevent the ion from wandering too far from this point. The trajectory is gathered and analyzed using the results of Sec. III A 2. Data gathered from the umbrella sampling used to estimate the PMF can be used in place of a separate simulation.

There is no guarantee that the presence of the harmonic bias potential will not distort the friction kernel. This would be a second order effect [since the first-order effect of any quadratic potential is compensated for in our solution, see Eq. (18)]. Furthermore, supposing the friction experienced inside the channel did not depend strongly on the ion's location in the channel, we would have further grounds for confidence in the results. Mamonov *et al.*[15] and Allen *et al.*[14] employed similar methods (that is, harmonically restraining the particle inside the channel while conducting an analysis based on the generalized Langevin equation) in order to determine the position dependent diffusion coefficient in gramicidin.

4. Estimating the friction kernel $K(t)$ using clamp and release

It can be shown using thermodynamics that we can sample from the equilibrium distribution contingent on $\mathbf{q}(0)=\mathbf{q}_0$ simply by sampling from the equilibrium distribution of the system where $\mathbf{q}(0)$ is fixed to $\mathbf{q}_0$ and sampling $\mathbf{p}(0)$ from a Boltzmann distribution. We thereby arrive at the following clamp and release procedure. First we perform the clamp phase: The system variables are clamped to $\mathbf{q}=\mathbf{q}_0$. The initial value of $\mathbf{p}(t=t_0)$ is sampled from a Boltzmann distribution. The clamped system is then equilibrated. Alternatively, instead of fixing $\mathbf{q}_0$, we can restrain it with a tight harmonic potential centered on $\mathbf{q}_0$, in which case there will be no need to sample $\mathbf{p}(t=t_0)$. Second, we perform the release phase. We remove the constraint on $\mathbf{q}$ and evolve the system while collecting the trajectory $\mathbf{q}(t)$.

At the end of the release phase, we decide whether to stop collecting data. The condition for stopping can be decided by testing disjoint subsets of the collected data; the observed variation in the resulting friction kernel will give an estimate of an upper bound on the statistical variation. If the stopping criterion is not met, we repeat the clamp and release cycle. The system is reverted back to its state immediately after the clamped equilibration, some further equilibration is performed, and a new release phase is started.

The procedure above is schematically illustrated in Fig. 1. This scheme is similar to that proposed by Adib,[32] although our aim is to measure the statistics of the ion by measuring correlation functions rather than measuring the PMF.

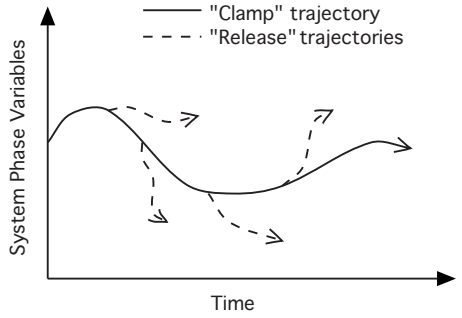

FIG. 1. A schematic diagram showing the evolution of the system during clamp and release. The solid line represents the trajectory of the system during the clamp phases, and the dashed side branches represent the release trajectories.

Now suppose we obtain a set of such trajectories, taken from the equilibrium distribution subject to $\mathbf{q}(0)=\mathbf{q}_0$. Using the results of Sec. III A 3, we can use these trajectories to compute $K(t)$, contingent on the initial value of $\mathbf{q}$, $\mathbf{q}(t=0)=\mathbf{q}_0$. However, microscopic reversibility ensures that the same autocorrelation function will apply contingent on the final value of $\mathbf{q}(t=t_f)=\mathbf{q}_0$. This suits our aim, which is to propagate $\mathbf{q}(t)$ forward in time, subject to its current position $\mathbf{q}(t)$ and information about its previous trajectory, $\mathbf{q}(t')$, $t'<t$.

Note also that Eq. (17) can be shown to apply, even when the autocorrelation is taken contingent on the initial or final value of $\mathbf{q}$, as it is here. Because we know $C(t)$, we can now apply Eq. (18) to determine the friction kernel $K(t)$ contingent on the final value of $\mathbf{q}(t=t_f)$.

B. Carrying out the distributional MD simulation

In this step, the simulation is carried out by solving the nonlinear generalized Langevin equation, as described in other papers.[17,33,34] However, there are several complications, for example, involving the boundary conditions for simulation space, that also need to be dealt with.[35]

Our stochastic dynamics setup for ion channels is shown schematically in Fig. 2. We divide the simulation space into two regions: a region consisting of two reservoirs, representing the bulk solution, and a channel region.

A fixed dielectric boundary between the protein/lipid system and the bulk/channel region is defined by embedding a cylindrically symmetric idealization of the shape of the channel within a dielectric slab that represents the lipid bilayer. In the bulk region, normal BD is carried out[35] with a long timestep, typically 100 fs. Boundaries are either treated as being perfectly elastic at the edges of the reservoirs, with a stochastic boundary being used to maintain concentrations in the top and bottom reservoirs or, where the protein/lipid enters the bulk region, using the macroscopic electrostatics that result from the solution of Poisson's equation with the dielectric boundary mentioned above. Ion-ion interactions are given as the sum of Coulomb forces, short range van der Waals forces, and two-body image forces resulting from the solution to Poisson's equation with the dielectric boundary.

In the channel region, the generalized Langevin equation is solved[17,33,34] using the PMF and friction kernel derived

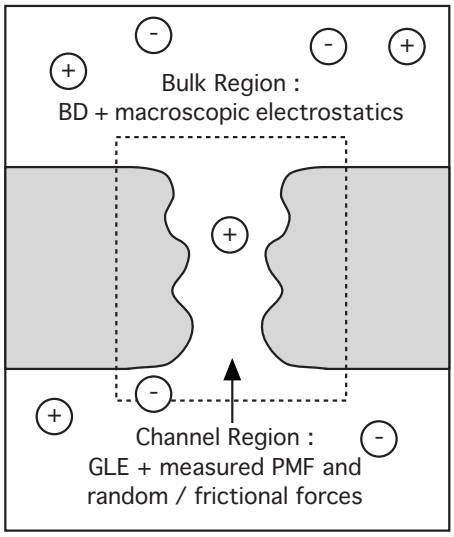

FIG. 2. A schematic diagram showing the setup used for stochastic dynamics simulations. We assume cylindrical symmetry. The simulation space is divided into a channel region and two reservoirs. In the channel region, the generalized Langevin equation is applied with a short timestep (2 fs) using a PMF and friction kernel derived from MD measurements. In the reservoir regions, the Langevin equation is solved with a long timestep (100 fs) using standard values for friction coefficients and with a force field derived from macroscopic electrostatics.

from MD measurements and using a shorter timestep, typically 2 fs. The option exists to apply external forces using macroscopic electrostatics, where the relevant data are not available from the MD measurements. For example, if we have not obtained a two-ion PMF, we can choose to treat ion-ion interactions in the same way as they are treated in the reservoir regions: using Coulomb and van der Waals forces plus two-body image forces, calculated using the dielectric boundary.

C. Details of the method used for the numerical example

We apply the theory and techniques outlined above to the gramicidin ion channel, an antibiotic polypeptide that has been extensively studied previously.[36–40] There are several reasons for choosing this channel as a test case: first, channels containing single-file water are not amenable to macroscopic electrostatics modeling,[18] and second, gramicidin is a one, or at most two, ion channel, which simplifies the analysis.

A three-dimensional, cylindrically symmetric single particle PMF is used for the single particle potential that is applied to each ion. This is likely to be a good approximation to the real non-cylindrically symmetric PMF.

The friction kernel is approximated using an exponential model:[17,33]

$$K(t) = \gamma\kappa \exp(-\kappa t). \tag{20}$$

The statistics that determine the ion motion are therefore governed by two parameters: γ, the inverse velocity decay time due to friction, and κ, the inverse decay time of the friction memory kernel. These parameters depend on the position of the ion in the channel. While this exponential model for the friction kernel ignores features seen in real measured friction kernels, we use it in this proof-of-principle study first because it provides the simplest possible way to estimate the general effect of a decaying friction kernel using only two parameters and second because it is guaranteed to be at least as good a representation of the motion as the more commonly used BD, since it can be made to approach the effect of a delta function as $\kappa \to \infty$. More general models for the friction kernel could be implemented using, e.g., the algorithm of Nilsson and Padró.[34]

In this initial study, the ion-ion potential consists of three terms: the usual Coulomb force, short range van der Waals forces consisting of a 6-12 Lennard-Jones potential, and an interaction, calculated by solving Poisson's equation, that results from the interaction of one ion with the surface charges induced by other ions on the fixed boundary of the protein-lipid system. In single-file channels, such as gramicidin, it is likely that this may not accurately capture some features of the ion-ion potential. For example, the mutual interaction between two ions separated by a single-file chain of water molecules will probably not be well described by Poisson's equation as applied here. Therefore, in future studies, we wish to model the ion-ion interaction using a two-ion PMF that is measured directly from MD. Note, however, that, since gramicidin is predominantly a one or two-ion channel, the approximations used here may adequately capture the dynamics of ion conduction, at least in cases where the concentration is not high enough to cause strong ion-ion interactions to take place inside the channel.

Timesteps of 100 fs in the bulk region and 2 fs in the channel region are used. We obtain the PMF and the friction kernel for an ion located at various points in the channel using the MD protocol outlined below: We begin with a model of the gramicidin dimer obtained from the protein data bank (PDB), PDB accession number 1JNO.[41] Water is placed inside the channel to speed up the equilibration process. A dimyristoylphosphatidylcholine (DMPC) membrane is then built around the hydrated channel using the "membrane builder" procedure developed by Woolf and Roux[42] and CHARMM,[43] and the membrane/protein system is then hydrated.

We run our simulations at zero concentration. This is done primarily because the interactions between the ions in solution will later be explicitly included in our generalized Langevin simulations, and we wish to avoid any minor effects that may arise from double counting of interactions between an ion in the channel and the bulk solution.

Various restraints are applied to the system. First, weak restraints are applied to stop the protein from drifting or rotating too far from its initial position: the z motion of the center of mass is weakly constrained by applying equal z-accelerations to each protein atom to avoid straining the bonds in the protein, and the C_α atoms are in addition weakly constrained in the xy plane to prevent the protein from tilting too much or drifting from the center of the simulation cell. Second, the Trp-9 χ_1 and χ_2 dihedrals are constrained[13] in order to prevent slow timescale dynamics that otherwise might occur as these residues drift to other metastable rotamers. Third, a cylindrical boundary of radius of 8 Å is ap-

plied to the test ion to prevent it from moving too far from the central axis in cases where it is no longer restrained by the channel.

The system, consisting of the gramicidin dimer, the DMPC membrane, water inside the channel, bulk water, and a test ion, is equilibrated using nanoscale molecular dynamics (NAMD) (Ref. 44) for 2 ns. Various other preprocessing steps are carried out with the help of NAMD and visual molecular dynamics (VMD).[45] The three-dimensional cylindrically symmetric PMF is then calculated using the WHAM procedure.[46] A total of 101 windows are used, spaced at 0.5 Å intervals, from $z=-25$ Å to $z=25$ Å. A harmonic constraining force with force constant of 12.5 kcal mol^{-1} Å^{-2} is applied for each window, centered at the location of the window. 1.5 ns of data are collected for each window. After preprocessing the data to ensure a straight channel centered on the z-axis, a two-dimensional free energy profile is then calculated,[47] with the coordinates being z and r. We idealize the potential by symmetrizing it about the $z=0$ plane.

In order to produce starting configurations with the ion located at the center of each weighted histogram analysis method (WHAM) window, we employ a procedure, implemented using NAMD's free energy perturbation (FEP) feature, where an ion in the channel is gradually "turned on"—in other words, its interactions with other atoms are gradually increased from zero to their usual values. This allows us to avoid the hysteresis effects that occur when an ion is "pushed" through the channel. An alternative, which we have also previously used, is to exchange the position of the ion and a water molecule located near the desired position of the ion. Further equilibration of 50 ps is carried out on each such starting configuration.

The data collected for the WHAM analysis are then analyzed to determine the friction kernel using the harmonic bias potential method explained previously. For each ion trajectory, we calculate the momentum autocorrelation and derive the kernel $K(t)$ using the methods detailed in Sec. II. The reciprocal of the relaxation time constant, γ, is then calculated as the time integral of $K(t)$ from $t=0$ to $t=\infty$. In reality, we examine the cumulative integral as a function of t and note that it saturates after about $t\approx 2$ ps [see Fig. 5(b)]. We then take the average value of the cumulative integral in the plateau region. Given γ, we then find κ in Eq. (20) by fitting an exponential function to $K(t)$.

IV. RESULTS FOR THE NUMERICAL EXAMPLE

A. Components of forces acting on the ions

The one- and three-dimensional PMF profiles constructed from the gramicidin pore using the CHARMM27 force field are shown in Figs. 3(a) and 3(b). The PMF profile we calculated is broadly similar to that reported by Allen *et al.*[13,14,48] In our PMF, outer and inner energy binding sites, each about $2kT$ ($T=310$ K) in depth, are visible at $z=\pm 11$ Å and $z=\pm 9$ Å, respectively. In the center of the pore is an energy barrier that rises about $17kT$ higher than the bottom of the wells. Several local minima are also visible along the length of the channel. In the lower panel, we show a two-dimensional slice through the three-dimensional, cylindrically symmetric profile encountered by a K^+ ion in the pore. As the ion moves away from the channel axis, the central barrier increases. Thus, the ion will predominantly dwell close to the axis of the pore.

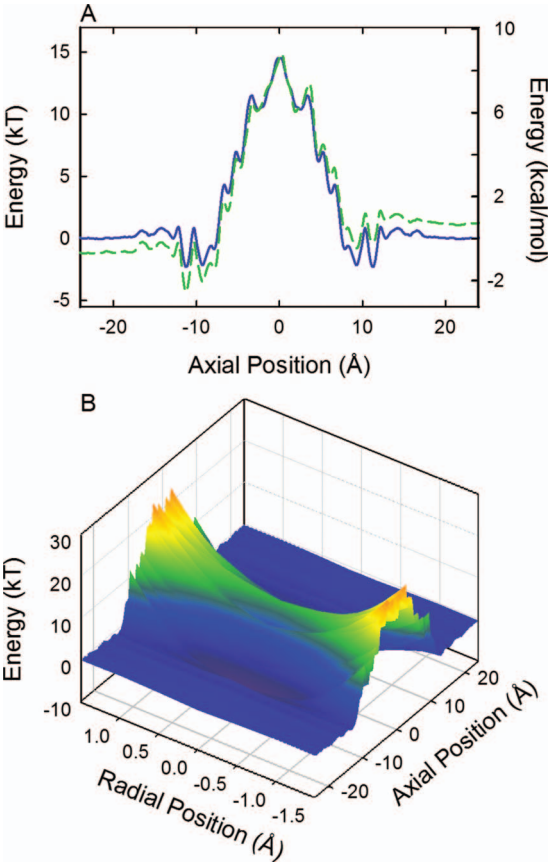

FIG. 3. PMFs of a K^+ ion in the gramicidin channel. (a) 1D PMF is obtained using CHARMM27 parameters from a gramicidin dimer (PDB No. 1JNO) embedded in a DMPC bilayer. The dashed line is the unsymmetrized PMF, and the solid line is the symmetrized PMF. (b) A two-dimensional cross section through the radially symmetrized three-dimensional PMF for the curve shown in (a). The energy is shown as a function of the ion's position along the channel axis and its coordinate along an axis perpendicular to the channel axis, assuming a cylindrically symmetric PMF.

Also shown in Fig. 3(a) is the unsymmetrized 1D PMF, as calculated directly from the raw data. There is an asymmetry in moving from one side of the membrane to the other of about $2.5kT$. This is consistent with the kind of error that might be expected from simulations on the order of 1–2 ns (Ref. 22) and gives some indication of the order of error that might be present in our symmetrized profile. As this study is primarily intended to test the theoretical techniques in this paper and given the other more serious uncertainties due, for example, to the force field parameters, we have limited our calculation to 1.5 ns per WHAM data collection window.

In order to verify that our stochastic simulations really are governed by the PMF shown in Figs. 3(a) and 3(b), we present, in Fig. 4, a comparison between the 1D PMF as derived from MD and a corresponding PMF calculated from our stochastic simulation under similar conditions to the original MD simulation. The match is excellent, except for a small discrepancy of around $0.5kT$ in the level of the bulk, which could be due either to a simulation artifact or to the

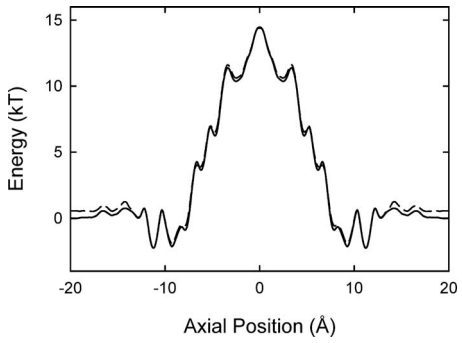

FIG. 4. A comparison between the 1D PMF seen in Fig. 3(a) and the corresponding PMF obtained for the stochastic simulation. There is a satisfactory match, accurate to around $0.5kT$, showing that our stochastic simulation should mimic the MD statistical mechanics reasonably well at lower concentrations.

difficulty of sampling the region around the transition to the bulk. We have performed this comparison both with (shown) and without background ions, finding more or less the same result in either case, which shows that the background solution has little effect on the one-ion PMF for the case studied here.

The friction kernels $K(t)$ measured inside ($z=1.5$ Å) and outside ($z=20$ Å) of the channel and their cumulative integrals are illustrated in Fig. 5. From the cumulative integral of $K(t)$ versus t, we can see that, within the channel, γ is increased by the presence of a long tail on K. Thus the motion of the ion in the channel is significantly retarded by an "echo" effect that takes effect over a timescale some ten times longer than that of the initial decay of the friction kernel. Using an exponential friction kernel to represent

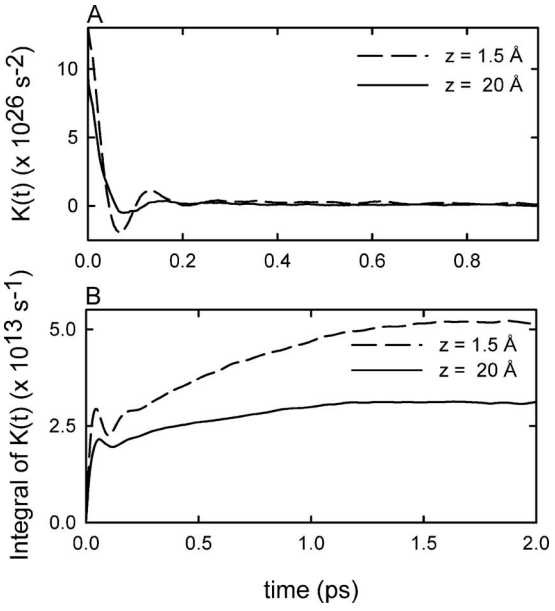

FIG. 5. The measured friction kernel $K(t)$ and its cumulative integral. (a) The friction kernel $K(t)$ is measured by using the techniques discussed in the text at $z=20$ Å and at $z=1.5$ Å. (b) The cumulative integral of $K(t)$, $\int_0^t K(t')dt'$, converges at $t\to\infty$ to γ.

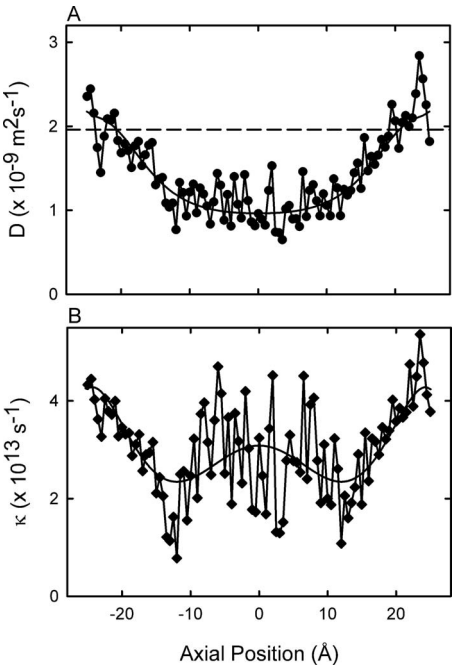

FIG. 6. (a) The diffusion coefficient and (b) the inverse decay time κ as a function of the ion's position along the channel axis. The diffusion coefficient is calculated from the measured values of γ using the relationship $D=kT/m\gamma$.

these dynamics will not be sufficient to fully represent the subtle details of the motion of the ion but nonetheless should be a reasonable approximation for most purposes.

Measured values of the diffusion coefficient $D=kT/m\gamma$ and κ are shown at various positions along the channel in Fig. 6. Note that the friction increases and the diffusion coefficient decreases from their bulk values as we move into the channel. The bulk diffusion coefficient for potassium ions, 1.96×10^{-9} m^2 s^{-1}, is indicated as a broken line in Fig. 6(a). On average, the diffusion coefficient inside the pore is reduced to 54% that of the bulk value. In our generalized Langevin equation simulations, we take account of the variation with z by dividing the channel up into nine regions and use average values over each region. These values enter only parametrically into the equation of motion. This parametric dependence would not be expected to be valid if there was an extreme variation of these parameters over a small distance: in this case, a new equation of motion would need to be derived in order to adequately represent the physical situation. Although a reasonably large variation can be seen in Fig. 6, particularly for the case of κ, we expect the generalized Langevin equation to be valid in the current case and do not pursue this question further.

We test whether the distribution of the random force can be approximated by a Gaussian distribution by applying the Kolmogorov–Smirnov and Anderson–Darling tests for normality to the data. We find that, given several hundred force samples, it is valid to assume a Gaussian random force. While the question of non-Gaussian force is somewhat subtle and quite interesting, we believe, based on these observations and other tests, that we are justified in modeling the random force as a Gaussian random variable.

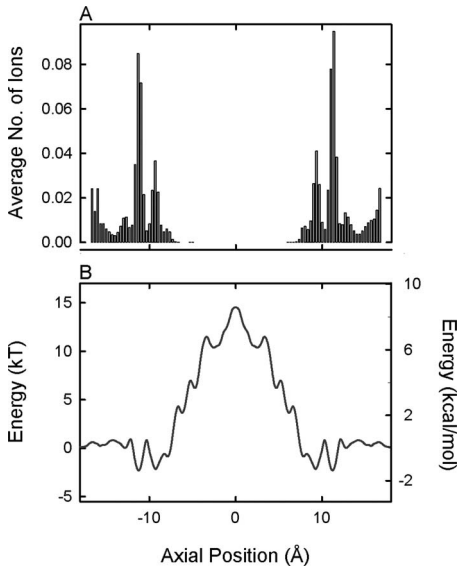

FIG. 7. A dwell histogram (a) obtained with the original PMF illustrated in Fig. 3(a) and reproduced in (b).

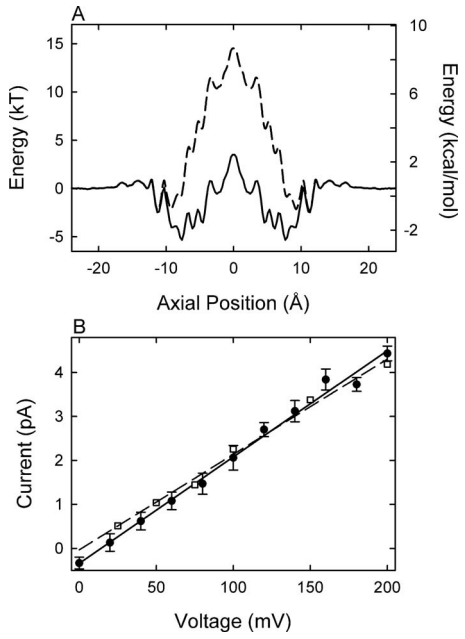

FIG. 8. Reverse engineered PMF. (a) A potential well of the form $E=A\exp(-(z/\sigma)^4)$ is subtracted from the PMF seen in Fig. 3. A and σ are chosen to give the good agreement seen in the current-voltage curve: $A=11kT$ and $\sigma=8$ Å. (b). The current-voltage curve. Superimposed on the curve is the experimental current-voltage relationship obtained by Andersen (Ref. 52). Both experimental and simulated currents are obtained with the KCl concentration of 500 mM.

B. Current versus voltage curves

In order to obtain current versus voltage curves in this numerical example, we apply a voltage drop that is linear over the length of the channel. Simulations are carried out at 500 mM concentration.

We incorporate the PMF illustrated in Fig. 3, along with the measured profiles for D and κ, into our stochastic dynamics algorithm and measure the current flowing across the pore. A simple estimate using the Boltzmann factor suggests that a barrier of this magnitude would suppress the current by more than six orders of magnitude. Clearly, we would not expect to see any ion surmounting this barrier during a feasible simulation period. With no applied potential, ions drift in and out of both binding sites. To determine the region in the channel where ions dwell preferentially, we divide the channel into 100 thin sections and count the number of ions in each slice during the simulation period of 0.8 μs. A dwell histogram shown in Fig. 7(a) reveals two sharp peaks, corresponding to the binding sites in the PMF. We see a 36% occupancy for each binding site (counting the inner and outer binding sites as a single large binding site).

As we have noted, the central barrier in the PMF illustrated in Fig. 3(a) is much too high to replicate the experimental ionic currents. There is a great deal of debate and uncertainty regarding the effectiveness of various MD force fields for describing processes such as ion conduction in gramicidin and similar channels. For example, small changes in the Lennard-Jones parameters describing the ion-protein interaction can lead to very large changes in the PMF.[13] Despite the many possible sources of uncertainty, attempts have been made to elucidate some of the systematic sources of error that may be present. Åqvist and Warshel[49] noted that the central barrier of the gramicidin PMF for Na^+ is reduced by $6kT-7kT$ due to the polarizability of the lipid and protein environments, an effect that is neglected in the standard CHARMM force field. They also correct for the artifacts introduced when using a finite simulation cell by the use of a Born-type formula. These effects are discussed in detail by Warshel *et al.*[50] Allen *et al.*,[13,14] in a later study, performed similar corrections on a MD system that is very similar to ours. They noted that the height of the central barrier is reduced by approximately $6.6kT$ once contributions resulting from various simulation artifacts are corrected for. Among the factors that contribute to the high central barrier, in this case, are the periodic boundary conditions used for MD simulations and the nonpolarizability of the membrane and protein hydrocarbon chains.

In order to obtain current-voltage curves under conditions more representative of the experimental data for gramicidin, we systematically lower the barrier height and measure the currents across the channel. From the original PMF, we subtract a potential well of the form $E=A\exp(-(z/\sigma)^4)$ so that the central barrier is low enough to permit conduction to take place. This modification to the potential is to a large extent arbitrary. However, it may be able to adequately represent the situation in real gramicidin, since one might suppose that, whatever the correct modification to the MD PMF might be, it should act systematically over the whole length of the channel and preserve features such as the outer binding sites and the central barrier.

The PMF that produced a close match to the experimental current-voltage relationship is shown in Fig. 8(a). The height A and half-width σ of the well subtracted from the original PMF are, respectively, $11kT$ and 8 Å. The depth of the well at each end of the pore is about $5kT$ and a barrier of about $9kT$ interposes between the two wells. In order to produce the best possible match to experimental data, we also found it necessary to artificially narrow the pore somewhat outside the channel mouth, at around $z=14-16$ Å.

The current-voltage curve illustrated in Fig. 8(b) is linear, with a conductance of 23 pS. Each point is determined from a simulation period of 25.6 μs. Superimposed on the simulated data in open squares are the experimental measurements obtained by Andersen. The match between the numerical and experimental results is seen to be excellent.

This may give an indication of what kind of PMF would be needed in order to replicate experimental observations. However, the simplifications made in our treatment of ion-ion interactions here mean that more work will be needed in order to come to a firmer conclusion.

V. CONCLUSIONS

In this paper we have given a framework for a distributional MD methodology: It uses MD simulations to inform stochastic dynamics. The resulting algorithm enables simulations that are two to three orders of magnitude faster than MD simulations while still sampling from approximately the same distribution of trajectories as does MD. For the example discussed here, stochastic dynamics is 500 times faster on a single processor. We have discussed the theory and assumptions involved in going from a deterministic MD to a stochastic dynamics simulation. We have then described details of how to estimate the parameters for the generalized Langevin equation from MD simulations: namely, the PMF and the friction kernel.

The new methodology presented here is then tested by modeling the permeation of ions across the gramicidin pore. Given the known difficulty of modeling ion conduction in gramicidin using classical MD,[13] it is perhaps not surprising that the PMF we have calculated using the conventional CHARMM27 force field is unable to reproduce the experimentally determined conductance. This may be due to artifacts in the MD simulations or possibly the sensitivity of the energy profile to small changes in the force field[13] or a combination of these two factors. As real-life gramicidin is known to conduct ions, we systematically modify the potential in order to match the experimental current-voltage relationship.

In the future, we will refine the technique, for example, by implementing more accurate ion-ion interactions, a more accurate modeling of the friction kernel in stochastic dynamics simulations, and a more careful treatment of the bulk/channel interface. The technique could then be fruitfully applied to more accurately measure macroscopic observables such as conductances and selectivity sequences for various ion channels under a range of conditions. Carbon or boron-nitride nanotubes,[51] due to their simplicity and rigidity, would be another particularly good candidate for study using this technique. Ultimately, we hope to enable a direct link to be made between aspects of channel structure and function, a task that would be nearly intractable using, for example, conventional MD modeling.

ACKNOWLEDGMENTS

This work was supported by grants from the National Health and Medical Research Council of Australia. The calculations upon which this work is based were carried out using the SGI Altix cluster of the Australian National University Supercomputer Facility. NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. We thank Dr. Tamsyn Hilder and Dr. Matthew Hoyles for their helpful comments on the manuscript.

[1] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon, Science **280**, 69 (1998).
[2] P. S. Crozier and R. L. Rowley, Phys. Rev. Lett. **86**, 2467 (2001).
[3] P. S. Crozier, D. Henderson, R. L. Rowley, and D. D. Busath, Biophys. J. **81**, 3077 (2001).
[4] A. Aksimentiev and K. Schulten, Biophys. J. **88**, 3745 (2005).
[5] R. D. Coalson and M. G. Kurnikova, in *Biological Membrane Ion Channels: Dynamics, Structure and Applications*, edited by S. H. Chung, O. S. Andersen, and V. Krishnamurthy (Springer, New York, 2007).
[6] S. H. Chung, T. W. Allen, M. Hoyles, and S. Kuyucak, Biophys. J. **77**, 2517 (1999).
[7] B. Hille, *Ion Channels of Excitable Membranes* (Sinauer Associates, Sunderland, MA, 2001).
[8] M. F. Schumaker, R. Pomès, and B. Roux, Biophys. J. **80**, 12 (2001).
[9] S. Bernèche and B. Roux, Proc. Natl. Acad. Sci. U.S.A. **100**, 8644 (2003).
[10] A. Burykin, C. Schutz, J. Villà, and A. Washel, Proteins: Struct., Funct., Genet. **47**, 265 (2002).
[11] A. B. Mamonov, R. D. Coalson, A. Nitzan, and M. G. Kurnikova, Biophys. J. **84**, 3646 (2003).
[12] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham, Acc. Chem. Res. **33**, 889 (2000).
[13] T. W. Allen, O. S. Andersen, and B. Roux, Biophys. J. **90**, 3447 (2006).
[14] T. W. Allen, O. S. Andersen, and B. Roux, Proc. Natl. Acad. Sci. U.S.A. **101**, 117 (2004).
[15] A. B. Mamonov, M. G. Kurnikova, and R. D. Coalson, Biophys. Chem. **124**, 268 (2006).
[16] T. Baştuğ and S. Kuyucak, Chem. Phys. Lett. **401**, 175 (2005).
[17] D. Gordon, V. Krishnamurthy, and S. H. Chung, Mol. Phys. **106**, 1353 (2008).
[18] S. Edwards, B. Corry, S. Kuyucak, and S. H. Chung, Biophys. J. **83**, 1348 (2002).
[19] R. Zwanzig, J. Stat. Phys. **9**, 215 (1973).
[20] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, New York, 2001).
[21] G. Ciccotti and J.-P. Ryckaert, J. Stat. Phys. **26**, 73 (1981).
[22] T. W. Allen, O. S. Andersen, and B. Roux, Biophys. Chem. **124**, 251 (2006).
[23] J. E. Straub, M. Borkovec, and B. J. Berne, J. Phys. Chem. **91**, 4995 (1987).
[24] H. Mori, Prog. Theor. Phys. **33**, 423 (1965).
[25] H. Mori, Prog. Theor. Phys. **34**, 399 (1965).
[26] B. J. Berne and G. D. Harp, Adv. Chem. Phys. **17**, 63 (1970).
[27] D. J. Evans and G. Morriss, *Statistical Mechanics of Nonequilibrium Liquids* (Cambridge University Press, Cambridge, 2008).
[28] A. M. Cohen, *Numerical Methods for Laplace Transform Inversion* (Springer, New York, 2007).
[29] M. Kato and A. Warshel, J. Phys. Chem. B **109**, 19516 (2005).
[30] C. Chipot, *Lecture Notes in Computational Science and Engineering* (Springer, New York, 2005), Vol. 49.
[31] J. Hénin and C. Chipot, J. Chem. Phys. **121**, 2904 (2004).
[32] A. B. Adib, J. Chem. Phys. **124**, 144111 (2006).
[33] S. Z. Wan, C. X. Wang, and Y. Y. Shi, Mol. Phys. **93**, 901 (1998).
[34] L. G. Nilsson and J. A. Padró, Mol. Phys. **71**, 355 (1990).
[35] V. Krishnamurthy and S. H. Chung, Proc. IEEE **95**, 853 (2007).
[36] O. S. Andersen and R. E. Koeppe, Physiol. Rev. **72**, 89 (1992).
[37] D. D. Busath, Annu. Rev. Physiol. **55**, 473 (1993).
[38] R. E. Koeppe and O. S. Andersen, Annu. Rev. Biophys. Biomol. Struct. **25**, 231 (1996).
[39] B. A. Wallace, J. Struct. Biol. **121**, 123 (1998).
[40] O. S. Andersen, R. E. Koeppe, and B. Roux, in *Biological Membrane Ion Channels: Dynamics, Structure and Application*, edited by S. H. Chung, O. S. Andersen, and V. Krishnamurthy (Springer, New York, 2007).
[41] L. E. Townsley, W. A. Tucker, S. Sham, and J. F. Hinton, Biochemistry **40**, 11676 (2001).

[42] Membrane builder (http://thallium.bsd.uchicago.edu/RouxLab/membrane.html).
[43] B. Brooks, R. Bruccoleri, D. Olafson, D. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[44] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, J. Comput. Chem. **26**, 1781 (2005).
[45] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).
[46] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, J. Comput. Chem. **13**, 1011 (1992).
[47] A. Grossfield, weighted histogram analysis method (http://www.dasher.wustl.edu/alan).
[48] T. W. Allen, T. Bastug, S. Kuyucak, and S. H. Chung, Biophys. J. **84**, 2159 (2003).
[49] J. Åqvist and A. Warshel, Biophys. J. **56**, 171 (1989).
[50] A. Warshel, P. K. Sharma, M. Kato, and W. W. Parson, Biochim. Biophys. Acta **1764**, 1647 (2006).
[51] T. A. Hilder, D. Gordon, and S. H. Chung, Small, DOI:10.1002/smll.209900349 (2009).
[52] O. S. Andersen (personal communication).