Permeation and gating in proteins: Kinetic Monte Carlo reaction path following

Gennady V. Miloshevsky and Peter C. Jordan[a)]
Department of Chemistry, MS-015 Brandeis University, P.O. Box 549110, Waltham, Massachusetts 02454-9110

(Received 16 November 2004; accepted 1 April 2005; published online 31 May 2005)

We present a new Monte Carlo technique, kinetic Monte Carlo reaction path following (kMCRPF), for the computer simulation of permeation and large-scale gating transitions in protein channels. It combines ideas from Metropolis Monte Carlo (MMC) and kinetic Monte Carlo (kMC) algorithms, and is particularly suitable when a reaction coordinate is well defined. Evolution of transition proceeds on the reaction coordinate by small jumps (kMC technique) toward the nearest lowest-energy uphill or downhill states, with the jumps thermally activated (constrained MMC). This approach permits navigation among potential minima on an energy surface, finding the minimum-energy paths and determining their associated free-energy profiles. The methodological and algorithmic strategies underlying the kMCRPF method are described. We have tested it using an analytical model and applied it to study permeation through the curvilinear ClC chloride and aquaporin pores and to gating in the gramicidin A channel. These studies of permeation and gating in real proteins provide extensive procedural tests of the method. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1924501]

I. INTRODUCTION

Many physical systems spend long periods in well-defined phase-space regions, characterized by minima on the potential-energy surface (PES). Transitions between these long-lived states occur as rare events of very short duration. Common examples are chemical reactions, permeation, or gating motions in proteins. Recent protein channel x-ray structures[1–7] open the possibility of atomic level study of both permeation and gating.[8] Permeation is commonly characterized by the free-energy profile along the channel axis normal to the bilayer plane (a well-defined reaction coordinate). With the reaction coordinate in hand, slow growth, umbrella sampling, thermodynamic perturbation and integration,[9] or the new technique of nonequilibrium work[10] can determine both axial free-energy profiles and the associated transition rates. However, the x-ray structures do not characterize conformational transitions of functioning, membrane-bound proteins. Proteins usually have multiple stable macrostates, associated with different aspects of their biological function. The study of gating dynamics involves the identification of the transition trajectories between macrostates. The trajectory along the minimum-energy path (MEP) is known as the reaction coordinate. Gating reaction coordinate(s) are then simply measures of the protein's conformation in its changes among the various closed and open states, processes typically of millisecond duration. As transition among them depends exponentially on the activation energies, these cannot be simulated by conventional methods such as molecular dynamics (MD), where behavior is typically observed for $\leqslant$100 ns. In MD simulations, Newton's equations of motion are integrated numerically and monitor behavior attributable to fast vibrations within local minima. Thus, simulating gating requires methods that can monitor rare transitions of short duration between stable states rather than sample events occurring within local minima. The existing computational tools need to be augmented to be able to predict large-scale conformational transitions in proteins.

Transition pathways in condensed-matter physics and materials science are found in many ways. The key idea is to focus directly on transition events by avoiding the sampling of fast vibrations within the stable states.[11] Therefore, some approaches accelerate computation by introducing a bias "boost energy" potential[12] to increase the probability of finding the system at a transition state or by heating the system[13] to increase the probability of rare events. These methods either modify the energy surface or sample at unrealistically high temperature. In the activation-relaxation technique[14] the system is driven from one potential-energy basin to another by inverting the force component acting on the system along a line drawn from the instantaneous configuration to the initial configuration, in effect the anti-force-bias method originally suggested by Cao and Berne.[15] As noted by Henkelman and Jónsson[16] this does not guarantee finding the transition with the lowest saddle point, making activation energy estimation uncertain. Transition path sampling[17] collects trajectories connecting two stable states using a Metropolis Monte Carlo (MMC) method, and monitors an ensemble of transition paths in trajectory space rather than in configuration space. It has been applied to biosystems, but none is larger than alanine dipeptide (22 atoms) solvated by 198 water molecules.[18] The non-Metropolis stochastic method, the Langevin equation as applied to Brownian motion, creates a set of dynamic trajectories through application of weighting

[a)]Author to whom correspondence should be addressed. Fax: 781-736-2516. Electronic mail: jordan@brandeis.edu

functions that bias a random walk.[19,20] Woolf *et al.* have described the details and challenges of this dynamic importance sampling method for analyzing gating transitions and permeation in ion channels.[21] The standard approach to climbing up the PES is normal-mode following,[22] but this requires the expensive determination and inversion of the second derivative Hessian matrix. Less expensive (requiring only first derivatives of the potential energy) is the dimer method,[16] which captures important features of normal-mode following. Some of these techniques are applicable if only the initial stable state is known, but others require advance determination of both potential-energy minima. The most effective method depends on the nature of the system under study; due to their individual limitations, none is universally applicable. Artificially constructed potential surfaces,[15,17,19,20] simple model systems,[12–14] or a combination thereof[16] were used for methodological testing. Transcribing these techniques, developed for ordered and disordered model systems, to real proteins, with their highly organized secondary and massively nonuniform tertiary structural elements, is not straightforward since complex conformational changes (rearrangement of the side chains and reconfiguration of limited portions of the backbone folds) affecting the reaction pathway occur during the gating transition.

We have combined ideas from the MMC [23] and kinetic Monte Carlo (kMC) techniques[24] and developed a practicable way to find and determine the MEP connecting local configurational minima. The resulting kinetic Monte Carlo reaction path following (kMCRPF) method samples both permeant motions and large-scale conformational transitions along the predefined degree of freedom (reaction coordinate) rapidly and correctly. Here we describe the computational strategy and the methodological and algorithmic details for applying the kMCRPF method and focus on real-life biomolecular applications where only the initial state of the transition (binding site, closed or open protein conformation) is known. We have tested kMCRPF using an analytical model and applied it to study permeation through the curvilinear ClC chloride and aquaporin pores[25,26] and gating in gramicidin A (gA).[27]

Our predictions provide new knowledge and they are in good agreement with much experimental data. These studies of permeation and gating in real proteins also provided extensive procedural tests of the kMCRPF technique, outlined in what follows.

II. METHODS

A. The constrained MMC method

For illustration, we particularize to the case of an ion permeating through a protein pore. Ionic motions perpendicular to the channel axis equilibrate rapidly; the axial coordinate Z is the only relevant slow variable and is thus a well-defined reaction coordinate. The Markov master equation[28] describes the time course of the probability $p(x,t)$ of the system being in state x at time t

$$\frac{\partial p(x,t)}{\partial t} = -\sum_{x'} w_{x\to x'} p(x,t) + \sum_{x} w_{x'\to x} p(x',t), \tag{1}$$

here $w_{x\to x'}$ is the transition probability per unit time that the system undergoes a transition from state x to x'. At equilibrium (steady state)[29] the left-hand side of Eq. (1) is zero and the sum of all transitions into a particular state x equals the sum of all transitions out of the state. This leads to the principle of detailed balance,[29] where the probability of starting from configuration x and going to configuration x' at a rate $w_{x\to x'}$ is matched by the reverse process

$$\frac{w_{x\to x'}}{w_{x'\to x}} = \frac{p(x')}{p(x)}.$$

We assume the system relaxes to equilibrium. Therefore, for an ergodic system, equilibrium is described by the stationary solution of the Markov master equation.[29] At equilibrium the ratio $p(x')/p(x)$ is given by the Boltzmann factor $\exp\{-[U(x')-U(x)]/kT\}$; thus a possible choice for $w_{x\to x'}$, which satisfies detailed balance, is the Metropolis walk,[23] for which

$$w_{x\to x'}\tau_{\mathrm{MC}} = \min\{1, p(x')/p(x)\}, \tag{2}$$

here $U(x)$ is the potential energy of configuration x and τ_{MC} is the Monte Carlo (MC) time step. The MMC algorithm[23] samples a Markov chain in state space so that the resulting (limiting/stationary/equilibrium) state distribution is equilibrated.[30] Detailed balance is used since it does not require knowing the transition matrix for the Markov walk and it guarantees Boltzmann–type behavior. Configurations are chosen with a probability $\exp[-U(x)/kT]$ and weighted evenly.[23] A new configuration x' is generated by displacing the coordinates of one atom (or an atomic group) by a randomly chosen small step, δ, followed by rejection or acceptance of x'. In standard MMC,[23] transitions to states of lower or equal energy have unit probability [the first argument of the min function in Eq. (2)] regardless of the energy barrier, the physical process, or the transition rate; this precludes establishing a temporal sequence of states. Moves that increase energy are rate limited and accepted with a probability $p(x')/p(x)$ [the second argument of the min function in Eq. (2)]. MC moves based on this algorithm cannot be interpreted dynamically as stochastic processes simulating random motion in time and consequently do not define a kinetic trajectory. However, a kinetic interpretation is possible if the sequential Markov states can be interpreted as dynamically correlated.

MMC permits Boltzmann-weighted moves to higher-energy states; energy barriers $\sim kT$ do not hinder the random walk. We focus on this property, constrain the ionic Z coordinate, and allow only unidirectional Z increments (constrained MMC for the ion). This approach is reminiscent of the nonreversal random-walk approach in statistical physics,[31] where the "walker" is constrained on a lattice where each step must be "forward;" immediate reversal is forbidden. This recollection feature limits the walker's moves; it can only access the nearest-neighbor forward sites on the lattice. However, our constrained MMC is designed

for off-lattice systems. The unidirectional constraint effectively provides an external driving force (voltage, concentration gradient, osmotic pressure, etc.), steering a permeant along the MEP. All other variable degrees of freedom (e.g., ionic X and Y coordinates, positions, and orientations of explicit waters and protein atoms) are unconstrained, relative to which the ionic Z coordinate evolves slowly. This accounts for a major aspect of ionic motion along the reaction pathway. To ensure that for all accepted ionic moves the low-energy pathway is equilibrated we make many MC trials, relaxing the ion and its environment while fixing the new ionic Z coordinate. Thus the other degrees of freedom relax fully in response to small ionic movement along the reaction coordinate; this is especially important in downhill evolution, since new, lower-energy configurations are always accepted. Permeation is sufficiently slow that a Boltzmann distribution is maintained, with the monitored ion always in equilibrium with the rest of the system. As the unidirectional Z constraint guides the ion along the MEP, it finds the lowest-energy trajectory along the pore and can overcome steep energy barriers.

Constrained MMC, essentially generalized diffusion, drives ion motion along the MEP producing a "dynamical hierarchy"[28] of distinct transitions separated by various MC time steps τ_{MC}. It establishes a chronological sequence of distinct ionic moves and permits dynamical interpretation since MC time is resolved on a scale for which no two events occur simultaneously. This chronological sequence (in this application, the ionic jumps), establishing a dynamical interpretation for MMC, differs from that used in standard kMC. The correspondence of MC time to physical time can be established in two ways: (1) Knowing the free energy barriers along the reaction coordinate, transition state theory determines the transition rates and the average duration of ionic occupancy of distinct stable states. (2) Mapping "kMC simulation time" onto physical time may be incorporated with an appropriately defined MC time scale. For instance, kMC simulation time may be expressed as the number of sweeps (displacements per ion) performed. As permeation and gating are slow processes with characteristic times of microsecond and millisecond (or longer), respectively, the physical time per MC sweep may be estimated by scaling measured permeation (duration of a channel crossing) or gating (lifetime of a closed or open state) times. This identifies the "time-consuming" sites on a permeation pathway or "hard" conformations in the dynamics of protein gating.

Clearly, constrained MMC for the ion does not satisfy detailed balance as reverse ionic moves never occur. However, the rest of the biomolecular system obeys this principle. Recent work shows, quite generally, that detailed balance is too strict a condition; any MC method that leaves the Boltzmann distribution invariant is correct.[32]

B. MC trial moves

In standard MMC an atom is chosen at random and displaced to any point inside a simulation box with equal probability. The maximum allowed displacement governs the trial move size, typically targeting acceptance ratios of $\sim$50%, well suited to simple homogeneous systems with identical atoms or molecules. However, an ion channel protein, with its pore and mouths filled with water, is a close-packed and highly organized inhomogeneous structure. It is very difficult to find a "hole" and successfully insert an ion or water in either the pore or its mouths. Standard methods fail as nearly all trials are energetically rejected. We thus limit displacement to immediate neighbors of the mobile ion or water, somewhat like the kMC technique[24] for simulating kinetic evolution processes on lattice systems (crystals) with fixed atomic sites. kMC computations involve single jumps of atoms to vacant neighboring positions. However, our monitored ion(s), water molecules, and protein atoms are not at fixed loci. Protein atoms vibrate freely and water molecules are not constrained at all. For the tagged ion, our kMC algorithm involves four steps: (1) sample the direction of ion transport randomly, but constrained to a forward or backward hemisphere (unidirectional motion along Z); (2) find the closest water or protein atom on this vector and calculate the distance l from the ion to its putative collision partner; (3) calculate a new ion coordinate, r', using $r'=r+l\xi$, with ξ randomly chosen between 0 and 1; and (4) move the ion to r'. Thus, the Z increment is not predetermined, but sampled randomly corresponding to accepted Boltzmann-weighted configurations. As the ion approaches a saddle point with rapidly increasing energy, only small changes in the reaction coordinate are likely to be accepted. This kMC approach is also used to translate water molecules and other ions in the system, but with no directional constraint.

C. The kMCRPF method

Standard kMC takes precomputed transition rates and creates a Markov-chain process mimicking the evolution of the real system. It neglects fast vibrations around equilibrium and follows slow configurational changes governed by the physical mechanisms driving each of the relevant transitions from a given initial state. Unlike MMC, both downhill and uphill kMC moves are rate limited; the energy barriers, physical processes, and transition rates are directly considered. Since Newton's equations of motion need not be solved, changes requiring long simulation times can be effected. To produce stochastic dynamics and map it to real times a sequential order in the transition probabilities obeying detailed balance is most easily created by assuming that the transition probability of the maximum rate process is one.[28] The time scale is incorporated via the transition rates.

Requiring prior characterization limits traditional kMC to simple lattice systems. However, ions, water molecules, and protein atoms do not sit on lattice sites. All possible permeation or gating pathways cannot be cataloged and specified in advance. The relevant transition probabilities must be found on the fly during simulation. The point is to find a way for evolution to occur slowly along the MEP (without a lattice approximation, with no predefined probability table). Continuing to focus on ion translocation, we want to traverse the MEP under constraints that either push the ion away from an initial state along the reaction path, or pull it toward a final state. However, our aim is also to simu-

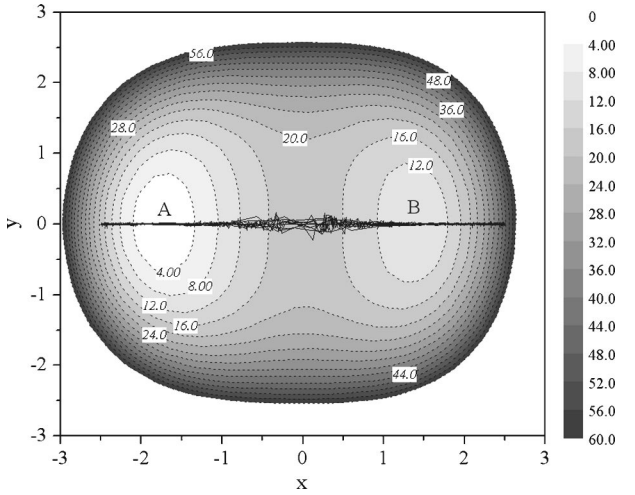

FIG. 1. Contour lines of the two-dimensional model potential surface, $V(x,y)$, used in the kMCRPF simulations. Contour lines are separated by 4 potential units. The labels on the map show the potential values. *A* and *B* indicate the locations of stable potential minima. The line at $y=0$ is a transition path. The curves at $y=0$ are particle trajectories from ten kMCRPF runs.

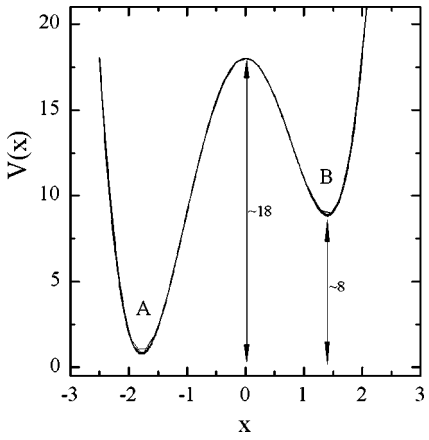

FIG. 2. Potential profiles along particle trajectories from ten kMCRPF runs. For all simulations $\beta=10$.

late long-time system evolution as accurately as possible without suppressing important underlying short-time processes. Our strategy is to meld ideas from both MMC and kMC. Evolution proceeds on the reaction coordinate by small ionic jumps (kMC technique) toward the nearest lowest-energy uphill or downhill states, with the jumps thermally activated (constrained MMC). The resulting kMCRPF method is new, with transition state search on the fly. The energy function is not biased; sampling is Boltzmann weighted. As the system evolves slowly, energy barriers need not be predetermined. Neither periodicity nor enumeration of all possible states is needed. In evolution, the ion moves among potential-energy minima on the PES; ionic energy fluctuates within the lowest-energy groove. Thermal fluctuations ($\sim kT$) will, in general, spread the set of transition trajectories. Carrying out several MC runs samples a large set of closely related transition trajectories.

III. TESTING THE KMCRPF METHOD

To test the method's accuracy and practicality we consider one particle in the two-dimensional potential:

$$V(x,y) = (x^2+y^2)^2 + x^3 + (x^2-5)^2 - 7, \tag{3}$$

with energy in reduced units. As seen in Fig. 1, $V(x,y)$ has two wells of different depths, *A* and *B*, located at $(x=-1.78, y=0)$ and $(x=1.4, y=0)$, respectively, and separated by a saddle at $(x=0, y=0)$. The profile $V(x,0)$ is given in Fig. 2, where the line $y=0$ corresponds to the MEP and x is the reaction coordinate.

Setting $y=0$ in Eq. (3) yields an analytical expression for the reaction coordinate. Well *A* is deeper than *B* by ~ 8. The barrier from the bottom of *A* to the saddle point at $x=0$ is ~ 18. A particle placed in either well *B* or *A* has a specific probability of moving to the other well along the MEP. We are interested in applying our kMCRPF method to reaction pathways in biomolecules that correspond to high friction, overdamped, or diffusive transitions. Therefore, motion along the reaction coordinate has to be diffusive, not simple ballistic jumps between wells. We can control particle behavior by adjusting the reduced temperature. At low temperatures, particle moves are accepted according to a Metropolis criterion if the change in potential is small. At high temperatures acceptable potential changes associated with particle motion can be large and the particle can jump over the barrier. In the simulations discussed here the reciprocal-reduced temperature was $\beta=10$. For this β only moves with a potential increase less than $\Delta V \sim 0.4$ have an appreciable nonzero acceptance probability $\exp(-\beta \Delta V)$. The maximum allowed particle displacement, δ, was adjusted to $\delta=0.25$ so that about 50% of the moves were accepted. Under these conditions a particle cannot traverse the transition region in a single attempt due to a large barrier (see Fig. 2), and it carries out random diffusive motion on the potential surface. We performed kMCRPF simulations by initially placing a particle at $x=\mp 2.5$ and then following it along the MEP either in a positive or negative direction along the x axis.

The direction of particle transport along the x axis is sampled in a forward or backward semicircle on the (x,y) plane (unidirectional motion). The y coordinate of a particle is unconstrained. For accepted moves on either semicircle, many MC trials are performed to relax the y coordinate while constraining the new x coordinate. A particle finds the MEP on the potential surface and travels from one well to the other. Ten transition trajectories of a particle on the (x,y) plane are illustrated in Fig. 1.

First five trajectories were started at $x=-2.5$ and they correspond to particle transport in the positive x direction; other five trajectories start at $x=2.5$ and follow in the negative x direction. It is seen in Fig. 1 that within each potential well particle trajectories remain very close to the line $y=0$. As is to be expected, transition trajectories are somewhat spread in the saddle region. The corresponding ten potential profiles are shown in Fig. 2. There is excellent agreement with the analytical potential profile $V(x,0)$.

To further test the method and demonstrate the effectiveness of our procedure we computed the nonbonded ion-protein interaction energy between a chloride ion and a ClC chloride pore, using a grid search,[9] and determined the en-

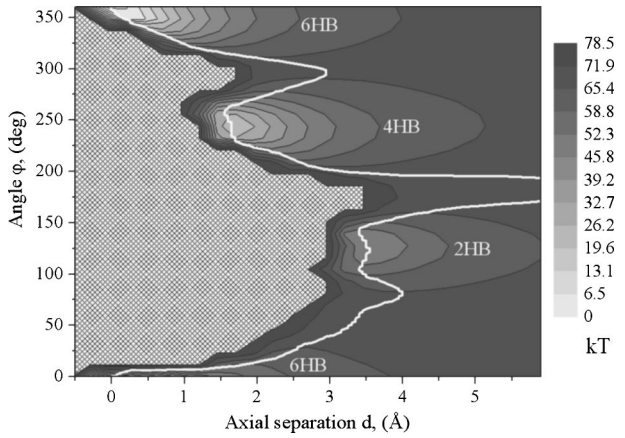

FIG. 3. Contour (d,φ) map of the nonbonded [electrostatic +van der Waals (vdW)] energy as a function of the separation distance d and rotation angle φ between gramicidin A monomers. The white trajectory is the gating pathway. The hatched part of the energy map is the region sterically inaccessible at small intermonomer separations.

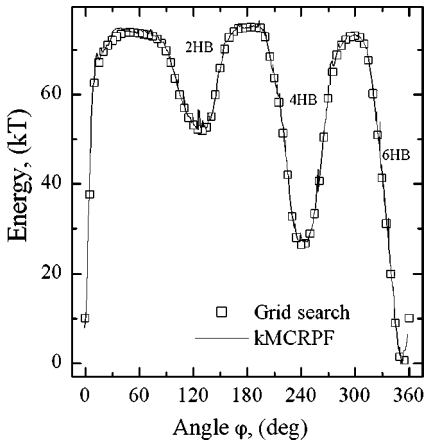

FIG. 4. Comparison of the minimum-energy profiles along the gramicidin A gating pathway from grid search and kMCRPF methods.

ergy at cubic lattice points superposed on the protein.[25] The grid points of lowest energy were identified with the paths of interest. For all values of Z the two energy profiles were in good agreement, both going uphill and downhill. A study of gA gating[27] provided a second test. The gA channel is a dimer formed by head-to-head association of two monomers at their amino termini, one from each bilayer leaflet,[33] a dimeric assembly stabilized by six junctional hydrogen bonds. Gating involves junctional dissociation and association, breaking and restoring the dimer's stabilizing hydrogen bonds. Since grid searching must be limited to just a few independent variables, gA monomers were only permitted two degrees of freedom, the intermonomer axial separation distance d and rotation angle φ. Figure 3 presents the grid-calculated contour (d,φ) map of the monomer-monomer interaction energy as a function of d and φ. As the monomers rotate relative to one another, their ability to hydrogen bond (HB) changes, with 6HB, 4HB, and 2HB states forming at 360°, 240° and 120°. Distinct, well-separated energy wells corresponding to these states are clearly evident. The surface's significant asymmetry provides a nontrivial test for the kMCRPF algorithm. The white solid line is the minimum-energy trajectory from grid searching. A kMCRPF simulation was also carried out to determine this trajectory and the corresponding energy profile, starting from the 6HB potential minimum. The two minimum-energy profiles, as functions of the rotation angle φ, are compared in Fig. 4. Both are similar over the whole φ range, with good agreement in both uphill and downhill directions. Strikingly, even along downhill paths, kMCRPF does not stray from the "exact" reaction coordinate; naturally agreement is imperfect, as the energy fluctuates thermally on the reaction pathway.

IV. APPLICATIONS

A. Permeation

We used kMCRPF to analyze anion translocation through a bacterial ClC chloride pore[25] and the translocation of water and charged species through the aquaporin-1 and glycerol facilitator channels,[26] all treated as rigid. Our focus was on the essential degrees of freedom (the permeating ion and its surroundings), excluding fast protein vibrations. In the absence of a gating rearrangement, we expected the average positions of the pore-lining protein atoms to be near those of the crystal structures. Determining ion-protein or water-protein interaction energy along the curvilinear pores in these proteins is not straightforward, as coordination is mainly with the side chains and the coordinates of the MEPs are not known in advance. We established the coordinates of the lowest-energy pathways, the location of the binding sites, the coordinating amino acids, and the nonbonded and free-energy profiles. Unlike condensed-matter systems (e.g., adatom diffusion on surfaces or crystal growth) that are usually homogeneous or periodic, the highly structured and inhomogeneous biomolecular systems have functionally important "special subsystems of interest" such as the permeating entity. Our implementation of kMCRPF stresses these "special degrees of freedom" by separating them from the rest of biomolecular system, thus highlighting the major determinant of the process being studied. Preferential sampling,[34] initially developed to efficiently simulate dilute aqueous solutions, was adapted and incorporated in kMCRPF to move the permeating species and its near neighbors more frequently than those farther away. This efficiently models complex biomolecular systems without loss of structural and kinetic details since its focus is on the permeant and its immediate neighborhood, avoiding much unnecessary calculation. This decomposition does not influence the rest of biomolecular system or its effect on the permeating species. With free-energy perturbation theory,[9] incorporating a backward sampling into the kMCRPF protocol, we calculated the potential of mean force (PMF) along a reaction coordinate,[26] so that a single complete kMCRPF run *simultaneously* determines the MEP and PMF. Our path following technique differs significantly from both umbrella sampling and solvation free-energy computations. Those techniques use arbitrary and equally spaced locations (windows) along a reaction coordinate. Our approach provides a realistic way of dynamically following the diffusing particle during permeation.

B. Gating

We applied kMCRPF to the simplest channel gating process: the dissociation of a gA dimer to monomers,[27] with the rotation angle φ between gA monomers chosen as the reaction coordinate, relying on our earlier determination of the contour (d,φ) map for the monomer-monomer interaction energy (Fig. 3). Intensive tests showed that φ is a proper reaction coordinate; recrossings do not occur when the monomer's lateral displacement and tilt are accounted for. We held the monomers effectively rigid,[27] an assumption justified by the experimental finding that the monomeric closed state of the channel is a half-dimer in the bilayer leaflet.[35] The kMCRPF method was able to identify multiple gating pathways illustrating possible ways in which the gA dimer may dissociate. kMCRPF predictions of the details of the gating process in gA channels[27] agreed fully with recent experiments,[36] based on a new single-molecule measurement technique, patch-clamp fluorescence microscopy.[37] Complementarities between theory and experiment are as follows: (1) Both concluded that multiple open and closed states result from fluctuations at the intermonomer junction, not from pore blockage by side chains occluding the mouth. Theory shows that these fluctuations arise from transitions between stable states with different numbers of hydrogen bonds passing through dimeric intermediates corresponding to closed and nonconductive pores; (2) Experiment finds considerable variation in channel conductance with no significant change in intermonomer separation. Theory explains the *mechanism* of this channel open-closed activity as due to coupled monomer rotation and lateral displacement interrupting and reforming the conductance pathway. Coupling rotation with axial separation is energetically far more costly; (3) Theory predicts and experiment finds conformational fluctuations in the open states. The theory explains the origin of flickers[38] in the gA channel as the result of such fluctuations; (4) Experiment shows that dynamics and kinetics of the gramicidin dimer are inhomogeneous and spatially confined. Theory reveals that gramicidin's helical structure makes specific rotational and spatial demands on the relative monomer arrangement along the most favorable dissociation path. It reproduces the experimental observation that the dimer rotates about its axis[39] and predicts that it undergoes translational diffusion in the membrane plane, ascribing both phenomena to restricted translational and rotational monomer motions.

V. DISCUSSION AND CONCLUSIONS

We have described the kMCRPF method for analyzing permeation and conformational changes. Here a predefined degree of freedom is identified as the reaction coordinate. For ion permeation through a pore a normal to the bilayer (Z coordinate) provides a natural reaction coordinate. The kMCRPF method enables a single MC run to *simultaneously* determine the minimum-energy trajectory along a curvilinear pore and the associated free-energy profile. Slow evolution along the reaction coordinate is essentially the same as ionic diffusion in that particular direction. The height and shape of the energy barriers need not be known in advance. The ion always equilibrates its surroundings. An ensemble of permeation trajectories through the pore can be collected from a set of MC runs. The resulting data then establish the optimal permeation trajectories and the averaged free-energy profiles. Simulation efficiency depends crucially on the computer time needed to relax the other degrees of freedom of the biomolecular system. Our implementation of preferential sampling techniques,[34] focusing the sampling on the permeating species and its immediate surroundings, greatly reduces the computational cost.

Understanding gating transitions in proteins is more difficult as the reaction coordinates cannot be identified by inspection. The PES must be carefully scrutinized to determine a reasonable initial guess as to likely degree(s) of freedom to be identified with the reaction coordinate. This requires detailed prior knowledge of the biomolecular system, an approach suitable for small peptides. Our exhaustive analysis of the gA PES [27] revealed that the relative rotation angle φ between the gA monomers was an appropriate reaction coordinate for which recrossings do not occur. For large proteins[6,7] identifying a gating degree of freedom (or a set of such degrees) by detailed study of the PES is very difficult as the energy surface is highly complex. A reaction coordinate is most unlikely to be described by a single degree of freedom; more probably it involves intricate coupling of many of them, with multiple recrossings. During gating the reaction coordinate switches among these degrees of freedom, which poses two problems. How to establish which of them contributes to the reaction coordinate? How to properly switch among them? Our kMCRPF method must be supplemented with another technique in order to determine a reaction coordinate. An immediately applicable procedure is not self-evident. Following local normal modes[22] to find transition pathways or using the dimer method[16] may be effective strategies. These approaches, since they also involve climbing the PES, share an important feature with kMCRPF. Because all motions orthogonal to the reaction coordinate lead to higher-energy states, the system typically oscillates (damps) rapidly in these directions as it evolves. kMCRPF may be modified to take small random MMC steps up or down the potential surface along the lowest-frequency normal mode (eigenmode), followed by relaxation of the energy along all other normal modes. However, this procedure may be quite tricky in its implementation since, *inter alia*, the transition is highly anharmonic, involving the protein in substantial structural rearrangements during which its normal modes and spectral density function will change markedly and repeatedly.

Further development requires determining an appropriate procedure to establish reaction coordinates in large protein systems. Much experimental and theoretical evidence demonstrates that slow rigid-body motions of secondary structural elements correspond to the low-frequency normal modes of proteins. Our rigid protein group decomposition approach[25] is ideal for exploiting the rotation translation block method[40] to carry out normal-mode analysis on large proteins, introducing protein flexibility using the scaled collective variables approach.[41] This path-finding strategy will be incorporated in kMCRPF. The resulting kMCRPF method

will *simultaneously* define a reaction coordinate, progress along it, and determine the free-energy profile. The transition rate and kinetics can then be evaluated using other special purpose algorithms. We plan to apply these modifications to the study of gating transitions in the bacterial ClC chloride system[4,6] and in the voltage-dependent K^+ channel from Aeropyrum pernix (KvAP).[7]

ACKNOWLEDGMENT

This research was supported by a NIH grant, Grant No. GM-28643.

[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]G. Chang, R. H. Spencer, A. T. Lee, M. T. Barclay, and D. C. Rees, Science **282**, 2220 (1998).
[3]Y. Jiang, A. Lee, J. Chen, M. Cadene, B. T. Chait, and R. MacKinnon, Nature (London) **417**, 523 (2002).
[4]R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, and R. MacKinnon, Nature (London) **415**, 287 (2002).
[5]R. B. Bass, P. Strop, M. Barclay, and D. C. Rees, Science **298**, 1582 (2002).
[6]R. Dutzler, E. B. Campbell, and R. MacKinnon, Science **300**, 108 (2003).
[7]Y. Jiang, A. Lee, J. Chen, V. Ruta, M. Cadene, B. T. Chait, and R. MacKinnon, Nature (London) **423**, 33 (2003).
[8]G. V. Miloshevsky and P. C. Jordan, Trends Neurosci. **27**, 308 (2004).
[9]R. Leach, *Molecular Modelling: Principles and Applications* (Harlow, England, 2001).
[10]C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[11]L. R. Pratt, J. Chem. Phys. **85**, 5045 (1986).
[12]F. Voter, Phys. Rev. Lett. **78**, 3908 (1997).
[13]A. Sorensen and A. F. Voter, J. Chem. Phys. **112**, 9599 (2000).
[14]G. T. Barkema and N. Mousseau, Phys. Rev. Lett. **77**, 4358 (1996).
[15]J. Cao and B. J. Berne, J. Chem. Phys. **92**, 1980 (1990).
[16]G. Henkelman and H. Jónsson, J. Chem. Phys. **111**, 7010 (1999).
[17]C. Dellago, P. Bolhuis, F. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[18]P. G. Bolhuis, C. Dellago, and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **97**, 5877 (2000).
[19]T. B. Woolf, Chem. Phys. Lett. **289**, 433 (1998).
[20]D. M. Zuckerman and T. B. Woolf, J. Chem. Phys. **111**, 9475 (1999).
[21]T. B. Woolf, D. M. Zuckerman, N. Lu, and H. Jang, J. Mol. Graphics Modell. **22**, 359 (2004).
[22]J. Baker, J. Comput. Chem. **7**, 385 (1986).
[23]N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[24]K. Binder, *The Monte Carlo Method in Condensed Matter Physics* (Springer, Berlin, 1992).
[25]G. V. Miloshevsky and P. C. Jordan, Biophys. J. **86**, 825 (2004).
[26]G. V. Miloshevsky and P. C. Jordan, Biophys. J. **87**, 3690 (2004).
[27]G. V. Miloshevsky and P. C. Jordan, Biophys. J. **86**, 92 (2004).
[28]K. A. Fichthorn and W. H. Weinberg, J. Chem. Phys. **95**, 1090 (1991).
[29]D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000).
[30]J. S. Liu, *Monte Carlo Strategies in Scientific Computing* (Springer, New York, 2001).
[31]K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics. An Introduction* (Springer, Berlin, 2002).
[32]V. I. Manousiouthakis and M. W. Deem, J. Chem. Phys. **110**, 2753 (1999).
[33]L. E. Townsley, W. A. Tucker, S. Sham, and J. F. Hinton, Biochemistry **40**, 11676 (2001).
[34]J. C. Owicki and H. A. Scheraga, Chem. Phys. Lett. **47**, 600 (1977).
[35]K. He, S. J. Ludtke, Y. Wu, H. W. Huang, O. S. Andersen, D. Greathouse, and R. E. Koeppe, Biophys. Chem. **49**, 83 (1994).
[36]G. S. Harms, G. Orr, M. Montal, B. D. Thrall, S. D. Colson, and H. P. Lu, Biophys. J. **85**, 1826 (2003).
[37]G. Harms, G. Orr, and H. P. Lu, Appl. Phys. Lett. **84**, 1792 (2004).
[38]R. L. Goforth, K. C. Aung, V. G. Denise, L. P. Lyndon, R. E. Koeppe II, and O. S. Andersen, J. Gen. Physiol. **121**, 477 (2003).
[39]K. -C. Lee, W. Hu, and T. A. Cross, Biophys. J. **65**, 1162 (1993).
[40]F. Tama, F. X. Gadea, O. Marques, and Y. H. Sanejouand, Proteins **41**, 1 (2000).
[41]T. Noguti and N. Go, Biopolymers **24**, 527 (1985).