Published in final edited form as:

J Phys Chem B. 2010 October 7; 114(39): 12720–12728. doi:10.1021/jp1056122.

Renormalizing SMD: The Renormalization Approach and Its Use in Long Time Simulations and Accelerated PMF Calculations of Macromolecules

Anatoly Dryga and Arie Warshel*

Department of Chemistry, University of Southern California, Los Angeles, California 90089-1062

Abstract

Simulations of long time process in condensed phases in general and in biomolecules in particular, presents a major challenge that cannot be overcome at present by brute force molecular dynamics (MD) approaches. This work takes the renormalization method, introduced by us sometime ago, and establishes its reliability and potential in extending the time scale of molecular simulations. The validation involves a truncated gramicidin system in the gas phase that is small enough to allow very long explicit simulation and sufficiently complex to present the physics of realistic ion channels. The renormalization approach is found to be reliable and arguably presents the first approach that allows one to exploit the otherwise problematic steered molecular dynamics (SMD) treatments in quantitative and meaningful studies. It is established that we can reproduce the long time behavior of large systems by using Langevin dynamics (LD) simulations of a renormalized implicit model. This is done without spending the enormous time needed to obtain such trajectories in the explicit system. The present study also provides a promising advance in accelerated evaluation of free energy barriers. This is done by adjusting the effective potential in the implicit model to reproduce the same passage time as that obtained in the explicit model, under the influence of an external force. Here having a reasonable effective friction provides a way to extract the potential of mean force (PMF) without investing the time needed for regular PMF calculations. The renormalization approach, which is illustrated here in realistic calculations, is expected to provide a major help in studies of complex landscapes and in exploring long time dynamics of biomolecules.

I. Introduction

One of the long-standing problems in biomolecular simulations is getting meaningful time dependent information on long time functional processes. That is, while meaningful simulations of the function of biological systems on the picoseconds time range has been feasible for long time^{1,2}, simulating processes in the milliseconds and seconds time range has been a remarkable challenge.

Of course, using free energy perturbation (FEP) and umbrella sampling (US) approaches (e.g. ^{3,4}) has been extremely useful for evaluating activation barriers and then using transition state theory, but exploring true dynamical behavior on long time scale has remain a major challenge.

Another long standing challenge is involved in the effort to obtaining converging free energy profiles by FEP and US approaches within a reasonable computer time (e.g. see our

*Correspondence to: Arie Warshel, Department of Chemistry, 418 SGM Building, University of Southern California, 3620 McClintock Avenue, Los Angeles, CA 90089-1062. warshel@usc.edu.

discussion in ref ⁵). Here the emergence of new clever approaches has not yet provided practical advances. For example, the insightful work of Jarzynski^{6,7}, that has been examined in exploring alternative to FEP approaches (using many short time simulations) is considered by some as a providing somehow a way to circumvent the fundamental convergence problems. However, although the Jarzynski's identity⁶ is formally correct, it does not mean that the corresponding expression converges faster than standard FEP approaches⁷. Furthermore, although related strategies⁸ provided impressive analysis and formal prescriptions for using non equilibrium simulations, we are not aware of quantitative demonstrations that such approaches provide major convergence acceleration. We are also not aware of clear analysis that would provide a guide for choosing an effective pulling force. Similarly, the promising replica exchange approach^{9,10} has not been demonstrated to be much more effective than using repeated FEP calculations with random initial conditions¹¹. The difficulties with different versions of the Jarzynski and related approaches will be further considered below.

As much as the estimate of free energy profiles is concerned it is tempting to examine so called steered molecular dynamics (SMD) approaches^{12,13}. In fact, this type of approach has been used by us¹⁴ in a preliminary attempt to explore ion transport within the linear response approximation, but that work has not been advanced in a more systematic way until the work of ref ¹⁵. At present the problem with SMD is that it is not clear how to convert the results obtained with large external force forces to the corresponding results that would be obtained without external force at equilibrium (see below). Thus, while such an approach may tell us about the path between two known structures¹⁶, it is hard to determine what has been the effect of the bias on the path. The use of SMD for estimating activation barriers is even more problematic.

Before describing our strategy for obtaining long time properties and for estimating free energy barrier it is useful to clarify our general concern about current approaches that use SMD or related nonequilibrium approaches. Our point is best illustrate with the help of Fig. 1, which presents a typical case where we like to know the PMF for moving from the initial point (a) to the final point (b) and to do so in an efficient way. Now, a formal examination of this problem, from the perspective of nonequilibrium approaches (e.g. approaches that are related to the Jarzynski equality), is that we can get $\Delta G(t)$ for the given move by considering the corresponding nonequilibrium work. Unfortunately, if we use a strong force we will go directly through path 3 which will very rarely or never sample points along the least energy path (path 1) and thus will never give the correct PMF, unless we have infinitely long run. This indicates that we will have to follow the same strategy as in regular PMF calculations and to spend the same amount of computer time. Of course, one can argue that this is well known, but we are not aware of such discussions in the contest of non equilibrium approaches. Furthermore, and more importantly, we are not aware of computational prescriptions that show what is the advantage in using such approaches in evaluation of equilibrium PMF, nor we are aware of clear discussions of the optimal time and force for performing such calculations. The problem is that in the simulation field it is important to have unique validation studies that establish the advantage and reliability of different strategies and such studies are rarely available (see Discussion Section).

Overall, it seems to us that the fascination with the term nonequilibrium may have lead many to assume that FEP or umbrella sampling approaches requires "equilibration" and that somehow this is an advantage over what are presumably non equilibrium approaches. This possible belief is problematic, since in a typical FEP calculation all what is needed is to obtain converging results by sampling the system for a "sufficiently" long time. The simulation has to be average over the Boltzmann distribution, using terms of the form $\langle \exp(-(v_m - I - v_m)\beta) \rangle_{v_m}$ where $\langle \rangle$ designates an average over the given sampling potential (see

e.g. ref³) and this is automatically satisfied by running MD at the given temperature and has no formal relationship to the equilibration time. Of course, the given procedure may not converge if the system has not fully relaxed due to incomplete sampling in the proper region, but exactly the same problem would exist in all the so-called nonequilibrium approaches. Obviously, umbrella sampling approaches converge better when the segments are evaluated from corresponding equilibrium points, but the same is absolutely true for the so called nonequilibrium methods, since being at the lowest points of the least energy path lead to the fastest convergence.

One of the appealing options of advancing the use of nonequilibrium approaches has been the use of the time dependant linear response approximation (LRA) formulation¹⁷ in the way used by us in ref¹⁸ (see method section). With this formulation it should be possible to use different external forces and to see the relationship between the applied force and the average reaction coordinate and then to build the response function. This might allow one to obtain a general response function that can be used to obtain the equilibrium properties. However, experimenting with this approach we realized that a better alternative is provided by our renormalization method^{19–22}. This approach tries to get the best correspondence between the free energy and dynamics of a full explicit model (all-atom molecular dynamics) and an implicit Langevin Dynamics (LD) model of reduced dimensions. This is done by using constraints of different strengths in both the implicit and the explicit models and then adjusting the friction in LD treatment to maximize the agreement between the time dependence responses of both models (see Method section). Using external constraint allows us to run all-atom molecular dynamics simulation in a reasonable computer time (from several hours to several days using serial code), and then obtaining key information for LD simulations of long time processes. The renormalization approach had already being used very effectively in studies of the selectivity of ion channels¹⁹, in simulating proton transport²³ and exploring protein dynamics²¹. In fact, this approach has been validated on some level in ref²³. However, up to now we have not provide a systematic validation and a clear rational for our strategy and this is done in the present work.

Another new aspect of the present work is the realization that the renormalization approach provides a powerful way for evaluating PMFs. This “non-equilibrium” strategy is also developed and demonstrated here. Finally the relationship to some alternative approaches is discussed.

Overall we point out in the present work that the ultimate validation of a simulation approach is not the formal elegance but the actual ability to obtain converging results for the systems of interest. Considering the difficulty with other approaches that try to estimate effective frictions and PMFs, we start from the idea that there must be some optimal correspondence between the full explicit and the implicit LD model. Thus the best way to get this correspondence is to adjust the friction and PMF of the LD model simultaneously, until the best agreement between the two models is obtained. We believe that this seemingly pedestrian approach is far more promising than any other sophisticated formulation.

II. Methods

Our overall aim is to find a practical way for performing long time simulations (in micro or milliseconds time range) of complex systems, while still capturing the correct physics. A reasonable starting point would be LRA treatment of the time dependant response of an average property to a perturbation which is switched on at time $t=0$:

$$H=H^0+H'=H^0+KQ \quad (1)$$

Where H' is a perturbation, H^0 is Hamiltonian of the system without perturbation, Q can be a set of generalized coordinates (in the case of system described in present work it is the z component of the ion coordinate), and K is a constant in the present case. According to the LRA the response is given by²⁴

$$\langle \delta Q(t) \rangle = -K\beta \int_0^t \phi(Q, t') dt' = -\beta \int_0^t \langle \dot{Q}(0) Q(t') \rangle K dt' \quad (2)$$

where $\langle \delta Q(t) \rangle$ is the time-dependent system response to perturbation, $-\phi(Q, t')$ is a kernel that is given by the derivative of the time-correlation function of Q , and $\beta = (kT)^{-1}$. This approach turn out to be very effective in exploring fast downhill processes^{18,25,26}, but using it for high barrier cases is much more problematic. That is, in the case of systems with high barriers it is not clear that just evaluating the equilibrium $-\phi(Q, t')$ will provide the desired results even if we have the long time behavior of the relevant correlation time. A much more promising approach would be to consider $\phi(Q, t')$ as some general function with several decay times

$$\phi(Q, t') = \sum_{i=1}^n A_i(Q) e^{-t'/\tau_i} \quad (3)$$

Now we can evaluate the parameters A_i and τ_i by applying different values of K to the actual simulation system, finding the corresponding $\langle \delta Q(t) \rangle$ and then fitting Eq. (2) to the resulting behavior. Such an approach can be used effectively for predicting the response of an ion channel to an external potential but this of course requires one to have a reasonable estimate of the long time terms in Eq. (3)

The resulting approach would be similar in some ways to the strategy used by electrical engineers in determining the resistance (or impedance) by using different potentials. However, it is not so clear what is the most effective way of implementing such a treatment and how to benefit from a prior understanding of the shape of the landscape. Thus we look for a retaliated alternative that still considers the explicit response of Q to different forces (different K), but then generates the corresponding friction and effective potential in an LD treatment. In other words we can write following our renormalization approach^{15,19,23 21}, and represent the full system by a simplified system which is propagated by a LD simulation, using

$$m\ddot{q}_\alpha = -m\gamma\dot{q}_\alpha - \frac{\partial \Delta G}{\partial q_\alpha} - \frac{\partial H'}{\partial q_\alpha} + A'(t) \quad (4)$$

where α runs over the x, y and z coordinates of the ion, m and γ are the mass and friction of the ion and G is the effective free energy function that includes mainly the electrostatic profile, but also the steric restriction term used in ref¹⁵. $A'(t)$ is a random force that satisfies the fluctuation dissipation theorem²⁴. Integrating the LD equation we can obtain $\langle Q'(t) \rangle_K$ and compare it to the corresponding coordinate in the full system (with the corresponding constraint). Such a procedure allows us to find the best friction constant (or if needed to use an even more sophisticated memory kernel) and use it in long time LD simulations, instead of Eq. 2. This philosophy is described in more details below. Our challenge is to simulate the dynamics in the space defined by the effective coordinates in a way that it represents the

long-time behavior of the real system. This is done by adopting an hierarchical renormalization model, where we force a model of a limited dimension with a coordinate set Q (which will be described here by three coordinates) to represent an explicit all atom model with a coordinate set r .

We start by mapping the full model along the reaction coordinate and generating a one dimensional effective free energy surface. The next step is to force the two models to have an equivalent dynamics. This is done by starting with a regular MD simulations on model A (explicit) and evaluating the time dependence of moving between two sets of coordinates (z_1 and z_2) under the influence of a constraint or pulling potential:

$$U_{con}=K(z-z_0)^2 \quad (5)$$

where K is a force constant and where a large value of K is needed in order to obtain this transition in a reasonable time. The value of z_0 ($z_0=20000\text{\AA}$) was chosen to ensure that force ($|F| \approx 2Kz_0$) on the ion is (almost) constant during pulling process, where the constraint potential affects directly only the ion in the simulation system. Next we move to the implicit model and evaluate the time dependence of moving between two coordinate sets Q_1 and Q_2 (that correspond to z_1 and z_2) using a LD treatment as outlined in Eq.(4):

$$\begin{aligned} m\ddot{q}_z &= -m\gamma\dot{q}_z - \frac{\partial\Delta G}{\partial q_z} - 2Kz_0 + A'(t) \\ m\ddot{q}_\beta &= -m\gamma\dot{q}_\beta - \frac{\partial\Delta G_{steric}}{\partial q_\beta} + A'(t) \end{aligned} \quad (6)$$

where q_z is the z coordinate of the ion, β runs over the x and y coordinates of the ion, ΔG_{steric} is the change in free energy associated with the shift from the channel z axis (see ref 15 for details).

The key issue here is to force the dynamics of the implicit (simplified) model to correspond to that of the explicit (full atomistic) model. Thus we force the time dependence for moving between the two generated sets of coordinates in each model (e.g. the time required to move between two significantly different coordinate sets r_1 and r_2 in the explicit model should correspond to the time of moving between the equivalent coordinates in implicit model) for different values of the applied constraint. We then refine the friction, γ , by comparing the passage time and overall time dependence obtained from dynamics in different models.

It is also possible to validate and “fine-tuned” the dynamical properties at different timescales) of our renormalization approach by requiring that the velocity autocorrelation of $z(t)$ or other related properties will have similar behavior in the two models.

Once the renormalization model is calibrated by the above procedure we can use it for exploring the behavior of the system in very long time scale by simply, running the calibrated implicit model for a long time and exploring long time properties, as was done for example in our study of the coupling between conformational and chemical dynamics²¹.

In addition to the above application of the renormalization approach, we can use it in an accelerated evaluation of PMFs. That is, after calibrating the friction term in the simplified LD equation we can consider cases when the potential of the explicit model is changed (let say because of a mutation) and then try to get the same passage time and $\langle z(t) \rangle$ for the explicit modified surface and the LD effective potential. This is done while keeping the friction unchanged and adjusts the PMF along the z direction until we reproduce the same

passage time and $\langle z(t) \rangle$ of the explicit system. Our strategy will be clarified and demonstrated below.

The system studied is described in the next section, where we give the rationale for choosing such system. All the simulations for the explicit model were performed with the MOLARIS program package. The parameters used in the present work are exactly the same as in ref⁵. The LD simulations were done with the CHANELIX¹⁹ module of MOLARIS.

III. Getting a relevant benchmark is not simple

In order to be able to examine the validity of our approach (or, in fact, any other approach) it is crucial to have a relevant benchmark. That is, as we argued repeatedly (e.g. ref⁵) it is crucial to have a benchmark that reflects the actual problem of interest. For example, in studies of approaches for PMF calculations of ion channels one cannot assume that models validated in studies of torsional rotation in dipeptide have been properly validated for its ability to obtain PMF for ion conductance. Apparently, such models do not encounter the relevant problems of solvation in a multidimensional heterogeneous environment (see discussion in ref.⁵). In other words, many of the key issues in biophysical simulations cannot be resolved by some formal considerations (or by using models that do not reflect the proper complexity), but must be based on actual validation on the systems of interest. For example, in the case of ions in polar proteins or ion channels the relevant issues are the time of response of the protein dipoles to the movement of the ion and this depends on the reorganization time of these dipoles. Thus it is crucial to have a model that reflects relevant biophysical complexity and also allows for reasonably long simulation time. For this purpose we use the truncated gramicidin A (gR) in a gas phase model depicted in Fig. 2. The initial structure (pdb 1JNO) was mutated to all glycine residues and position constraint of 10 kcal/Å² on all C- α atoms were used to create a stable benchmark (this was done since the membrane is removed and the simplified model channel cannot keep a reasonable shape by itself). The system used in the current work, as well as the direction of the pulling force, can also be seen on Fig. 2.

It is also useful to point out that even the use of decent PMFs of the full solvated gR does not provide, in our view, an optimal validation system in view of the uncertainties with the calculations and their convergence (see ref⁵). Thus the use of a system where we basically know the “exact” result is highly recommended.

IV. Results

The initial step in our renormalization approach requires one to obtain the potential of mean force (PMF). In the present case we already have a reliable PMF for the simulation system, since the specific PMF had been already obtained in ref⁵. We also considered changes in the PMF and these were obtained by the same procedure used in ref⁵. At any rate, the PMF obtained for the explicit model (Fig. 3) was digitized and used as the effective free energy profile in the LD model. We also generated a model that allowed us to modify the PMF and to be able to see how the ion moves through the channel without a pulling potential. To achieve this goal we altered our simulation system adding four fixed charges outside the ion channel with the distance of 8 Å from the center of the symmetry of the channel as shown in Fig. 4. Changing the charge on the atoms allowed us to increase the activation barrier in a parametric way without modifying the friction. The low barrier system has slightly different minimum along the PMF profile and therefore the average time required for the ion to pass the channel was defined as the time for moving the ion through the barrier, from $z = -6$ Å to $z = 6$ Å.

Next we started to generate the $\langle z(t) \rangle$ as well as the average time required for the ion to pass through the channel (this is defined as the time for moving ion from $z=-8\text{\AA}$ to $z=8\text{\AA}$ through the barrier of the explicit model). This was done by using of the gR models with different K values. Typical trajectories of the explicit model for different force constants are depicted in Fig. 5, while Fig. 6 compares trajectories and autocorrelation functions for the explicit and implicit models. As shown in Fig. 6 we could easily force the LD model (by adjusting the friction coefficient) to have similar trajectories and similar velocity autocorrelation function to that of the explicit model. Obviously, using the same friction for all force constants is more challenging but this seems to be feasible for the model studied to date. At any rate, the results obtained for our systems are summarized in Fig. 7, where we provides the overall dependence of the first passage time on the force constant for the explicit and implicit models. In this case we determine the average time by using 16 or 32 trajectories for each force constant. Typically the total time is 5–10 times larger than the average time required for the ion to pass the channel (first passage time) for a given force constant. As seen from the figure we obtained an excellent agreement between the behaviors of the implicit and explicit models over a very range of first passage time.

After establishing our ability to obtaining the same fist passage times in the two models, we also evaluated a very long time trajectory in the absence of any force (but with a constraint that prevents the ion from going out of the channel). The simulation of the corresponding test system, with a low barrier without the pulling potential, took several CPU days, and yield reasonable agreement between the explicit and implicit models (see Fig. 8). It should be noted that the parameterization for the LD model did not use data for a system with the low barrier. That is, the value obtained in the study of the system with the high barrier (namely a friction coefficient of 8 ps^{-1}). Considering the above results, we believe that we establish that a renormalized ion channel system can be used for reliable LD simulations to investigate ion current and selectivity

Next we examined options for accelerating calculations of activation free energy. In order to do so we looked for a simple and controllable way of changing the activation barrier in the explicit model without drastic change of the factors that control the friction. This was done by using the system of Fig. 4. Here we considered several cases with known change in the PMF (shown in Fig. 9) and evaluated the effect of strong pulling force on the passage time of these test systems, pretending that we do not know the activation barriers. We then forced the LD simulations, with a friction coefficient that corresponds to the system without 4 fixed charges, to reproduce the passage time of the explicit system by adjusting the effective potential for the LD runs. A Gaussian shape potential with width of $2\text{\AA}$ (see Fig. 9) were added to the initial known potential⁵. As shown in Fig. 9,10 and 11 the potentials that allowed us to obtain the best match between the implicit and explicit passage times reproduce the actual activation barriers in the explicit system. $\Delta\Delta g^\ddagger$ defined as a difference in the barrier height of the system with 4 fixed charges and the high barrier system(without 4 fixed charges).

The next challenge is to determine the complete PMF when the shape is not known. In this case we have to determine the PMF from the several segments and such an approach was explored using the system with the high barrier (see Fig. 12). The actual simulations were done on two segments and then extended by exploiting the symmetry of the system. The determination of the PMF from separate segments required the use of different force constant, in order to achieve a passage time that is not too fast and not too long (passage time in hundreds of picoseconds range were used). It should also be noted that our approach requires one to obtain the PMF of the segments by running uphill trajectories, since downhill trajectories with and without external force have similar passage time. However, this limitation can be easily overcome by applying pulling potential in the opposite direction.

We also explored different strategies for estimating the barrier height for the explicit system. The starting point in these studies was the crucial need of having reasonable effective friction (see Discussion section). In principle we can obtain a reasonable approximation from a renormalization study of regions with low activation barriers, but here we used the already known optimal friction and tried first to estimate the barrier height by using effective potentials with a Gaussian shape. The corresponding results are presented in Fig. 12 and as seen from the figure these results are only qualitatively correct (we obtain a deviation of 5 kcal/mol from the correct barrier of the explicit model). Thus we adopted a second approach where we “know” the shape of the potential (this can be done with the approach considered below). In this strategy we generated potentials that have the same minima as the explicit potential and the barrier between these two minima and systematically scaled the barriers by a factor (in the range of 0.6 – 1.6). Now we obtained excellent agreement between the renormalized PMF and the corresponding explicit result (see Fig. 12).

The above validation emphasized the advantage of knowing the general shape of the potential in conducting a renormalization process. This highlighted the advantage of simple model such as the PDL/D/S-LRA²⁷ in obtaining the general shape of the potential or to use the above renormalization approach in studies of mutational effects when the changes in the potential are small. Of course, we also have to have a reasonable idea of the best effective friction and this is another crucial advantage of the renormalization idea

In this respect we would like to note that our approach depends on the knowledge of the friction. However, as mentioned above the optimal friction can be usually estimated without the full renormalization treatment, by performing renormalization studies of regions with low activation barriers of the given system. This issue will be further explored in subsequent works.

At any rate, the advantage of the present approach is most apparent when we can use a relatively strong force constant and short simulation time. This can be extremely useful in assessing the rate of biological processes with large barriers. Of course, if we like to evaluate the exact details of the PMF in particularly in segments where the PMF is shallow and where the reaction coordinate does not follow a simple path, we need to use smaller constraint forces and longer simulation times.

At this point it is useful to discuss the simulation times of the different models. Obviously the simulation time depends on the force constant. The longest simulations reported for model with the high barrier (from ⁵) were about 0.2 μ s, which correspond to several days of computer time (using one core of Dual Quadcore AMD Opteron 2.3 GHz) for the explicit model and about an hour for the implicit model. The PMF simulations for the model with the high barrier (from ⁵) took, with averaging for backward and forward runs, about 16 hours, while evaluating the PMF with the renormalization approach requires 2 hours when a large force constant was used for determination of the barrier height. We expect greater speedup in solution and this issue is under investigation.

Of course, the main time saving in our model is not the PMF calculations but the ability to run long time simulations on the implicit model. Here we can obtain a millisecond (0.1 μ s) simulation in Fig. 8 with the implicit model (in several hours), whereas we will need several days with the explicit model (even for our small simulation system). For larger systems the saving is much more remarkable.

V. Some Challenging Problems

While the renormalization approach has been shown to be very effective in the case of ion channels, we are also interested in the performance of this approach in other cases like conformational transitions in proteins. In such cases we are dealing with more complex landscape and the renormalization procedure can be more demanding. As an example we consider the conformational change of Adenylate Kinase (ADK) that has been used as a generic model in our study of the coupling between chemical and conformational changes²¹. In this case we are dealing with a large conformational change that has been explored by several workers (e.g. ref ^{28,29}). In our preliminary study²¹ we considered the exact details of the conformational transition to be highly *irrelevant*, since our study was about the general coupling issue and not about any specific enzyme (see also below). Here we considered the ADK type renormalization problem in more details, examining (as before) three models; the full explicit model, the CG model where each side chains is replaced by single interaction centers³⁰. and a 2-D model with one conformational coordinate and one chemical coordinate²¹. The renormalization started by applying the same forces to the three models, while adjusting the corresponding frictions in CG and 2-D model. This was done, however, by using a simple single barrier potential in the 2-D model (rather than by evaluating first the PMF in the full or CG model, since the work of ref ²¹ meant to give conclusions that are independent of the given conformational PMF). In the current case we applied weaker forces than those used in Ref.²¹ and the corresponding results are presented in Fig. 13.

As seen from this figure we can find friction constants that satisfies the renormalization requirement, however, it is not yet clear how unique are the results. Furthermore, calculation of the autocorrelation of the conformational coordinate shows that we can reproduce the autocorrelation of the explicit model with very different frictions (namely 0.5 and 100 ps⁻¹). In fact, attempts to simulate small peptides by LD approaches (e.g. ³¹) seem to give similar results with frictions ranging from 0.5 to 50 ps⁻¹.

In our opinion the renormalization approach should provide unique results, or at least a consistent way for modeling the long time motions and free energy profiles in studies of protein conformational changes. However, this should be demonstrated in studies of relevant systems. The most obvious first step would be to estimate the actual PMF of the protein conformational change and to refine the renormalization results. Such a study can start by considering the correspondence between the CG and the 2-D models, rather than starting from the full model (the CG model, in fact, may provide a more reliable PMF than the full model). Another possible strategy can be provided by using relatively weak forces that would allows one the pull the system along small segment one in a time. This approach can be used also to simultaneously construct the PMF. The performance of the renormalization in such cases would be validated in subsequent studies and it clearly presents an interesting challenge. However, we also like to emphasize that we do not see any fundamental problem in using the renormalization in studies of protein conformational changes, except the likely requirement to perform the renormalization on segments of the overall change in order to obtain more reliable results

Regardless of the needed refinement and validation in modeling large conformational changes, we must clarify that the finding of the absence of dynamical coupling between chemical and conformational coordinates in enzyme is completely valid. That is, our conclusions were obtained by examining all the range of reasonable frictions including changing corrugation of the 2-D model. The same results were also obtained by using the CG model with all the protein features for limited changes of the conformational changes (Fig. 5 of ref ²¹).

VI. Discussion

This work presents practical and powerful renormalization method and validated this approach on a system that reflects the main complexity of realistic ion channels. The renormalization approach is found to be reliable and arguably presents what is perhaps the first approach that allows one to exploit SMD in quantitative and more meaningful studies.

The renormalization approach has already been used in studies of several important problems^{15,19–21}, but the present work provided the first systematic validation of this method. It is established that we can reproduce the long time behavior of large complex systems by using LD simulations of a renormalized implicit model. This is done without spending the enormous time needed to obtain such trajectories in the explicit system. The present study also provided a promising breakthrough in accelerated evaluation of PMFs. This is done by adjusting the effective potential in the implicit model to reproduce the same passage time as that obtained in the explicit model under the influence of an external force provides a way to extract the PMF without inverting the time need from regular PMF calculations. This approach, which is illustrated here in realistic calculations, is expected to provide a major help in studies of complex landscapes. It also should provide a very useful way for assessing activation free energies of macromolecular processes.

The main point in our approach is the idea that there must be some correspondence between the full and the implicit models and the best way to get this correspondence is to adjust the friction and PMF of the implicit model simultaneously, until the best agreement between the two models is obtained. We believe that this seemingly simple strategy is more promising than any other formulation with clever and sophisticated formulations.

The use of the renormalization strategy for estimating PMFs may look similar to the approach proposed by Hummer and Szabo⁸ for the exploitation of Jarzynski equality. However, the resulting approach has not been validated by comparing its estimate to the actual PMF of a realistic model system (with known PMF). In fact the instructive attempt of ref³² to reproduce the observed energetics of RNA unfolding have resulted in a major discrepancy that was attributed to the complexity of the real system. The above problem also appears to exist in the attempts of Schulten and coworkers^{13,33} to construct PMFs. Here we have again a formally appealing treatment, but we do not have a realistic validation study on a complex system. That is, in the study of ref³³ (as well as in most other studies in the field) we see sophisticated derivations but the validation does not consider a system with *both* the relevant complexity and reliably determined PMF. For example ref³³ uses simply a one dimensional potential, which cannot be used in validating approaches for free energy calculations in complex systems. Ref¹³ considered the challenging evaluation of the PMF of the end to end distance of deca-alanine, but unfortunately it is not clear what is the exact converging result for this PMF. More importantly, despite the formal sophistication of the above approaches they present a clear problem; for example, the corresponding formulation does not tell us what is the magnitude of the driving force that would give reliable results for the given simulation time. The same is true with regards to the selection of the optimal friction in attempts to use SMD in PMF calculations. Here the renormalization approach provides physically consistent and a clear prescription for obtaining simulation-based friction, which is then used to guide the selection of the optimal simulation and convergence conditions.

In our view, in the current stage of the field it is crucial to have a numerical rather than formal validation as the issue is really how much computer effort is needed to get correctly converging results.

The approach that is probably the closest in spirit to our approach is the so called surrogate process approximation (SPA) of Calderon and coworkers^{34,35}. This approach that appeared recently independently from our earlier renormalization approach (e.g. ref ¹⁹), uses SMD to construct PMFs. This approach was actually validated in calculating PMF of the gR channel and obtained encouraging results (see however above). At any rate, the SPA approach is focused at evaluating PMF whereas our approach is more directed towards providing the long time properties of the system, which cannot be estimated by other approaches.

One of the arguments in support of Jarzynski type approaches is the option of running many short simulations in parallel, which classifies them among the “embarrassingly parallel” type of problems. While this can be an advantage, there is nothing that would prevent the renormalization strategy (or standard FEP calculations) from being performed by running many simultaneous runs (where we again have an embarrassingly parallel type of problem) with both the averaging on several runs and performing the calculations on many short segments of the reaction coordinate.

In discussing alternative strategies it is important to state that several approaches outlined effective ways for evaluating PMFs with a known friction constant. The problem is that the proper friction constant is not known. Here our idea, of obtaining both the friction and PMF from SMD type treatments, presents a crucial advantage. In this respect we like to reemphasize that the friction obtained from standard treatments (namely the Einstein’s or related formulation) may not provide the optimal effective friction. Here it is instructive to consider the recent work of ref ³⁶ where the approach of ref ¹³ was used in a study of a single ion PMF in the pore of Vpu from HIV -1. It was found as expected that the results depend drastically on the friction used making the approach rather unreliable without rather arbitrary scaling of the friction.

It must be clarified here that although the nature of frictional constant has been understood by giants like Chandrasekhar³⁷, Kubo²⁴ and others, the actual value depends on the model used (e.g. ref ³⁸). Furthermore, the friction in the effective LD model may reflect factors that are only considered implicitly. For example if the PMF missed some corrugation it will be reflected in change of the optimal friction.

The gR system chosen might be considered by some as a simple one dimensional system. However, in fact, we have here a system with the main complexity of ion channels. That is, the reorganization of the channel dipoles upon transfer of the cation provides a complex “solvent” coordinate, with the typical solvent relaxation dynamics. Capturing the coupling between the solvent and solute motions is a challenge that has to be addressed by the renormalization approach. Here the friction and the PMF must reflect behavior of the correct solute/solvent system with its effective two dimensional features.

It must be emphasized, despite our intensive discussion of PMF calculations, that the main innovation and aim of our renormalization approach is the evaluation of long time dynamics. We do have many strategies for estimating PMFs in macromolecules, in particular with electrostatic models, but where we feel the need for innovative approaches is the calculations of long time dynamics and this is where the renormalization approach provides a powerful and effective tool. It is also useful to point out in this respect that some of the suggestions to use path sampling and related approaches for long time simulations^{19,39–41} do not in fact offer a way for long time simulations and that the results reported by ref ³⁹ reflect ns simulations and the assumption that the ms processes follow transition state theory.

One of the main motivations for the development of the renormalization approach has been the realization that (unless one believes in the absolute reliability of computer simulations)

we cannot judge with certainty which simulation results provides the most consistent description of the available experiments. A case in point is the use of voltage- current experiments in ion channels in assessing the quality of simulation results. Here, for example we have the study of ref ⁴² who compared the PMF produced by Aqvist and Warshel⁴³ to that produced by Roux and Karplus ⁴⁴ and concluded that the PMF of ref ⁴⁴ is more reliable (forgetting to point out that this conclusion had been obtained only after scaling the potential of ref ⁴⁴ by 0.3 and thus obtaining a similar low barrier to that obtained in ref⁴³). Another interesting related work is that of ref ⁴⁵, who tried to judge the reliability of different PMFs of gR. The problem is that the conclusions of the above studies are not unique and that even current simulations of PMFs suffer from convergence problems and from problems in treating long range effects and induce dipoles. Of course, the problems are much more serious in the case of biological ion channels. Thus we view the renormalization approach as probably the best way of projecting the explicit results on a simpler dimensionality and allowing for consistent comparison with the relevant experimental information (e.g. voltage-current relationships). In fact, our point can be best understood by considering the fact that the conclusions of the above works are based on using a friction constants that are quite different from those obtained by our renormalization approach and thus are unlikely to produce the correct current with the correct PMF.

The demonstration of the power of the renormalization approach is likely to open new applications. For example, this very promising approach is likely to advance in simulating the selectivity and activation of voltage activated ion channel in a way that is still reliably and yet feasible with the available computer time.

Acknowledgments

This work was supported by NIH grants GM 24492 and GM40283. We gratefully acknowledge the University of Southern California's High Performance Computing and Communications Center for computer time.

References

1. Warshel A. *Nature*. 1976; 260:679–683. [PubMed: 1264239]
2. Warshel A. *Acc Chem Res*. 2002; 35:385–395. [PubMed: 12069623]
3. Warshel, A. *Computer modeling of chemical reactions in enzymes and solutions*. Wiley; New York: 1991.
4. Kollman P. *Chem Rev*. 1993; 93:2395–2417.
5. Kato M, Warshel A. *J Phys Chem B*. 2005; 109:19516–19522. [PubMed: 16853521]
6. Jarzynski C. *Phys Rev Lett*. 1997; 78:2690–2693.
7. Hendrix DA, Jarzynski C. *J Chem Phys*. 2001; 114:5974–5981.
8. Hummer G, Szabo A. *Proc Natl Acad Sci USA*. 2001; 98:3658–3661. [PubMed: 11274384]
9. Hukushima K, Nemoto K. *J Phys Soc Jpn*. 1996; 65:1604–1608.
10. Sugita Y, Okamoto Y. *Chem Phys Lett*. 1999; 314:141–151.
11. Liu H, Warshel A. *J Phys Chem B*. 2007; 111:7852–7861. [PubMed: 17571875]
12. Isralewitz B, Gao M, Schulten K. *Curr Opin Struct Biol*. 2001; 11:224–230. [PubMed: 11297932]
13. Park S, Schulten K. *J Chem Phys*. 2004; 120:5946–5961. [PubMed: 15267476]
14. Warshel, A. *Abstracts of Papers of the International Biophysics Congress*. 1990.
15. Burykin A, Schutz CN, Villa J, Warshel A. *Proteins: Struct, Funct, and Genet*. 2002; 47:265–280. [PubMed: 11948781]
16. Xiang Y, Goodman MF, Beard WA, Wilson SH, Warshel A. *Proteins: Struct Funct Bioinf*. 2008; 70:231–247.
17. McQuarrie, DA. *Statistical mechanics*. University Science Books; Sausalito, Calif: 2000.
18. Hwang JK, King G, Creighton S, Warshel A. *J Am Chem Soc*. 1988; 110:5297–5311.

19. Burykin A, Kato M, Warshel A. *Proteins: Struct, Funct, and Genet.* 2003; 52:412–426. [PubMed: 12866052]
20. Liu H, Shi Y, Chen XS, Warshel A. *Proc Natl Acad Sci USA.* 2009; 106:7449–7454. [PubMed: 19383795]
21. Pislakov AV, Cao J, Kamerlin SCL, Warshel A. *Proc Natl Acad Sci USA.* 2009; 106:17359–17364. [PubMed: 19805169]
22. Kamerlin SCL, Vicatos S, Dryga A, Warshel A. *Ann Rev Phys Chem.* 2010 in press.
23. Braun-Sand S, Warshel A. *Biophys J.* 2005; 88:507A–507A.
24. Kubo R. *Rep Prog Phys.* 1966; 29:255–284.
25. Warshel A, Parson WW. *Q Rev Biophys.* 2001; 34:563–679. [PubMed: 11852595]
26. Parson WW, Warshel A. *J Phys Chem B.* 2004; 108:10474–10483.
27. Warshel A, Sharma PK, Kato M, Parson WW. *BBA-Proteins and Proteomics.* 2006; 1764:1647–1676. [PubMed: 17049320]
28. Lu Q, Wang J. *J Am Chem Soc.* 2008; 130:4772–4783. [PubMed: 18338887]
29. Whitford PC, Gosavi S, Onuchic JN. *J Biol Chem.* 2008; 283:2042–2048. [PubMed: 17998210]
30. Messer BM, Roca M, Chu ZT, Vicatos S, Kilshtain AV, Warshel A. *Proteins: Struct Funct Bioinf.* 2010; 78:1212–1227.
31. Feig M. *J Chem Theory Comput.* 2007; 3:1734–1748.
32. Hummer G, Szabo A. *Acc Chem Res.* 2005; 38:504–513. [PubMed: 16028884]
33. Gullingsrud JR, Braun R, Schulten K. *J Comput Phys.* 1999; 151:190–211.
34. Calderon CP, Janosi L, Kosztin I. *J Chem Phys.* 2009; 130:130–142.
35. Calderon CP. *J Chem Phys.* 2007; 126:126–136.
36. Patargias G, Martay H, Fischer WB. *J Biomol Struct Dyn.* 2009; 27:1–12. [PubMed: 19492858]
37. Chandrasekhar S. *Rev Mod Phys.* 1943; 15:1–89.
38. Uhlenbeck GE, Ornstein LS. *PhRv.* 1930; 36:823–841.
39. Vreede J, Juraszek J, Bolhuis PG. *Proc Natl Acad Sci USA.* 2010; 107:2397–2402. [PubMed: 20133754]
40. Karplus M. *Proc Natl Acad Sci USA.* 2010; 107:E71–E71. [PubMed: 20424124]
41. Hummer G. *Proc Natl Acad Sci USA.* 2010; 107:2381–2382. [PubMed: 20133629]
42. McGill P, Schumaker MF. *Biophys J.* 1996; 71:1723–1742. [PubMed: 8889150]
43. Aqvist J, Warshel A. *Biophys J.* 1989; 56:171–182. [PubMed: 2473789]
44. Roux B, Karplus M. *J Am Chem Soc.* 1993; 115:3250–3262.
45. Gordon D, Krishnamurthy V, Chung SH. *J Chem Phys.* 2009; 131

Figure 1.

Demonstrating the nature of response to a one dimensional force in two dimensional system. The figure is aimed at clarifying the fact that nonequilibrium type approaches can never provide reliable PMF with short simulations and large forces. In such case the system will pass through path 3 and will not collect information about the actual least energy path. Of course running infinite number of short trajectories with strong force will also rarely sample path 1 but this would not offer an efficient sampling strategy.

Figure 2.
The benchmark system of a truncated gR with a Sodium ion

Figure 3.

The PMFs for the high barrier model (from ref ⁵) and the Low barrier (9.3 kcal/mol) model of Fig. 4 with charges $q = -0.3$. The PMF of the system with the Low barrier is shifted in order to have the same minimum as in the system with the high barrier.

Figure 4. Truncated gR with 4 additional charges at the center of the system. The change of the charges provides a systematical way to alter the original PMF.

Figure 5. Two trajectories of the explicit model, with a high barrier, for different pulling forces, K in kcal/(mol·Å²).

Figure 6.

Typical MD and LD ion trajectories for the explicit and implicit models, respectively (A) as well as the velocity autocorrelation functions (B) for the high barrier model and for $K=0.0001$ kcal/mol/Å².

Figure 7.

The first passage time for explicit and implicit models for different force constants. The data shown are for two model: A high barrier model (from ref⁵) and a low barrier model (truncated gR with 4 negative charges).

Figure 8. Long time trajectories of the explicit and implicit models for the system with a low barrier (9.3 kcal/mol) and without a pulling force. The figure demonstrates that we can reproduce the long time behavior of the explicit model.

Figure 9.
The original PMF for the explicit model and the modified PMFs for the system with four additional positive charges

Figure 10.

Relative first passage time as a function of the $\Delta\Delta g^\ddagger$ of the explicit model for three different pulling forces. The estimated $\Delta\Delta g^\ddagger$ is independent of the magnitude of the pulling force. The relative time is defined as the ratio of the first passage time (with a given pulling constant) for the system with a modified PMF to first passage time for the original high barrier model.

Figure 11.

A Comparison of the $\Delta\Delta g^\ddagger$ obtained by the renormalization approach and the exact results obtained with the careful analysis of the explicit model.

Figure 12.

A comparison of the PMF of the explicit model (black) and the PMF (red) obtained by using the renormalization approach (using LD simulations that reproduce the first passage time for different segments of the explicit model) as well as estimation of the barrier height.

Figure 13.

A qualitative renormalization study of the conformational change in ADK for the full-atom, Coarse-Grain and 2-D LD models. It describes the timed appearance of the response to force of the three models.