

Collective Diffusion Model for Water Permeation through Microscopic Channels

Fangqiang Zhu, Emad Tajkhorshid, and Klaus Schulten

*Theoretical and Computational Biophysics Group, Beckman Institute, University of Illinois at Urbana-Champaign,
405 N. Mathews, Urbana, Illinois 61801, USA*

(Received 1 April 2004; published 24 November 2004)

Water permeation through nanometric channels, generally a coupled many-body process, is described in this study by a single collective coordinate. A collective diffusion model is proposed in which water movement at equilibrium is characterized as an unbiased diffusion along this coordinate and water transport in the presence of a chemical potential difference is described as one-dimensional diffusion in a linear potential. The model allows one to determine the osmotic permeability of a water channel from equilibrium simulations.

DOI: 10.1103/PhysRevLett.93.224501

PACS numbers: 47.60.+i, 83.10.Mj

Water channels are ubiquitous in all life forms. For some biological channels, such as aquaporins [1], permeation of water is the primary physiological function. Many other channels designed for conducting substrates such as ions are water filled and permeate water as well. Water permeation through these biological channels has been the topic of theoretical and experimental studies for many years [2]. Molecular dynamics (MD) simulations provide an ideal tool for investigating water transport through channels [3], since the movement of every single water molecule can be closely monitored in the simulations. Here we propose a *collective diffusion* model to characterize water movement in channels under equilibrium and nonequilibrium conditions. The model allows one to analyze MD simulations in terms of the major experimental quantity of a water channel, namely, the osmotic permeability.

The osmotic permeability of a channel, p_f (cm³/s), is defined through [2]

$$j_W = p_f \Delta C_S, \quad (1)$$

where ΔC_S (mol/cm³) is the concentration difference of an impermeant solute between the two reservoirs connected by the channel, and j_W (mol/s) is the net molar water flux through the channel. Water flows from the reservoir with the lower solute concentration to the other reservoir because the two have different chemical potentials (the difference denoted as $\Delta\mu$) for water. For dilute solutions $\Delta\mu$ is linearly proportional to ΔC_S [2]:

$$\Delta\mu = -k_B T V_W \Delta C_S, \quad (2)$$

where V_W (18 cm³/mol) is the molar volume of water. Therefore, p_f is determined by the linear coefficient relating water flux and $\Delta\mu$.

Following its definition, p_f is measured in experiments under nonequilibrium conditions, for systems with non-zero $\Delta\mu$. In principle, $\Delta\mu$ can be established in nonequilibrium MD simulations through an osmotic [4] or a hydrostatic [5,6] pressure difference across the channel.

Because of the presently limited (ns) time scale of MD simulations, however, one has to adopt a large $\Delta\mu$ to obtain sufficient statistics of water permeation. This leads to situations that are far from actual experimental conditions, and it is not clear whether the results represent the normal kinetics of the water channel under study. If one can establish a quantitative relationship between water conduction under equilibrium and nonequilibrium conditions, this problem can be circumvented. In fact, even equilibrium MD simulations ($\Delta\mu = 0$) can be used to characterize the osmotic permeability of a channel, as we demonstrate below.

Water permeation in a channel usually involves multiple water molecules whose movements are coupled to each other. As a result, a complicated multidimensional representation seems to be necessary to model this process. In the following, we introduce a collective coordinate, n , which offers a much simplified description of water translocation in channels.

Consider a channel (of length L) aligned along the z direction. The collective coordinate n is defined in its differential form as follows: Let $S(t)$ denote the set of water molecules in the channel at time t , and let us assume that the displacement of water molecule i in the z direction during dt is dz_i ; then we define

$$dn = \sum_{i \in S(t)} dz_i / L. \quad (3)$$

By demanding $n = 0$ at $t = 0$, $n(t)$ can be uniquely determined by integrating dn . Note that $S(t)$ changes with time, and that a water molecule i contributes to n only when it is in the channel, i.e., if $i \in S(t)$ at time t . We further note that every water molecule crossing the channel from one reservoir to the other contributes to n a total increment of exactly $+1$ or -1 . Therefore, n quantifies the net amount of water permeation, and the trajectory $n(t)$ describes the time evolution of the permeation.

An important scenario is the stationary state in which a steady water flux through the channel exists. In this case,

$n(t)$ on average grows linearly with t , and the water flux is given by

$$j_W = \frac{1}{N_A} j_n = \frac{1}{N_A} \langle n(t) \rangle / t, \quad (4)$$

where N_A is Avogadro's number, and $j_n = N_A j_W$ is the water flux in the unit of number of water molecules/second.

At equilibrium, the net amount of water permeation through the channel vanishes on average, i.e., $\langle n(t) \rangle = 0$. Spontaneous, random water transport, however, may occur due to thermal fluctuation. Such microscopic fluctuations may not be detectable in experiments, but can be observed readily in MD simulations through $n(t)$. At equilibrium, $n(t)$ can be described as a one-dimensional unbiased random walk. It can be shown that when t is much longer than the velocity correlation time of n , the mean square displacement (MSD) of n , $\langle n^2(t) \rangle$, obeys the Einstein relation [7]

$$\langle n^2(t) \rangle = 2D_n t, \quad (5)$$

where D_n is defined as the diffusion coefficient of n . D_n has dimension t^{-1} since n is dimensionless. All factors affecting water kinetics contribute to D_n and are effectively integrated into this single parameter.

In the presence of a chemical potential difference ($\Delta\mu$) of water between the two reservoirs, a water flux through the channel arises and adopts quickly a steady value, and the configurations of channel water reach a stationary state. We further assume that in such stationary state channel water adopts properties (such as density and order) that are essentially independent of $\Delta\mu$ and very close to the equilibrium properties. This assumption does not necessarily require $\Delta\mu$ to be small compared to $k_B T$; it can be satisfied for liquid water even when $\Delta\mu$ exceeds $k_B T$. In the stationary state, the change in free energy of the system is solely due to water transport between the two reservoirs, because the state of channel water does not change with time. Since the net transport of one water molecule from one reservoir to the other results in a change of $\pm\Delta\mu$ in the free energy, the total free energy change is proportional to the net amount of water transported, i.e.,

$$U(n) = \Delta\mu \cdot n. \quad (6)$$

Since the stationary configurations (in terms of relevant properties) of the channel water are assumed to be very close to the equilibrium condition, one may use the equilibrium value of D_n to characterize the biased random walk of n in this case. Since the evolution of n can be described as a one-dimensional diffusion in a linear potential [cf. Eq. (6)], n on average is drifting with a constant velocity [9]:

$$\langle n(t) \rangle = -\frac{\Delta\mu}{k_B T} D_n t. \quad (7)$$

Such behavior is indeed expected for a stationary condition. According to Eq. (4), the steady water flux is then given by

$$j_n = -\frac{\Delta\mu}{k_B T} D_n. \quad (8)$$

From Eqs. (1), (2), (4), and (8) one obtains then for the osmotic permeability of the channel

$$p_f = v_W D_n, \quad (9)$$

where $v_W = V_W/N_A$ is the average volume of a single water molecule. Equation (9) shows that one can determine p_f using the D_n value obtained from equilibrium MD simulations [cf. Eq. (5)].

Recently, a continuous-time random-walk (CTRW) model [10] was proposed for a special type of channel, namely, a single-file water channel. This model assumes that the whole water file moves in discrete hops simultaneously and concertedly, with rightward and leftward hopping rates measuring both k_0 at equilibrium, k_0 being the single kinetic parameter in the CTRW model. Since each hop changes the collective coordinate, n , by $+1$ or -1 , it holds $n(t) = m_r(t) - m_l(t)$, where $m_r(t)$ and $m_l(t)$ are the numbers of rightward and leftward hops during time t , respectively. Because $m_r(t)$ and $m_l(t)$ obey a Poisson distribution whose mean and variance are both $k_0 t$ at equilibrium [5], one obtains $\langle n^2(t) \rangle = 2k_0 t$. Comparison with Eq. (5) yields $D_n = k_0$. Therefore, for the discrete water movement described by the CTRW model, D_n is identical to the hopping rate k_0 , and the expression derived from the CTRW model, namely, $p_f = v_W k_0$ [5], is actually equivalent to Eq. (9) in our collective diffusion model. However, unlike the CTRW model, the applicability of the collective diffusion model is not limited to single-file water channels.

In order to test the validity of the collective diffusion model, we performed a series of MD simulations on two exemplary membrane channel systems, denoted as a and b , respectively (Fig. 1). In each system, two layers of carbon atoms mimicking a membrane were used to partition the bulk water; a carbon nanotube (CNT) was embedded in the membrane to serve as a water channel. The CNT in system a is of the (6, 6) armchair type with a C-C diameter of ~ 8 Å. Previous simulations [11,12] showed that this CNT conducts water strictly in a single-file manner. The CNT in system b is of the (15, 15) armchair type with a C-C diameter of ~ 20 Å, filled with disordered, bulklike water molecules. Systems a and b contain 276 (~ 5 in pore) and 1923 (~ 90 in pore) water molecules, respectively. The length of the channel is $L = 13.2$ Å in both systems.

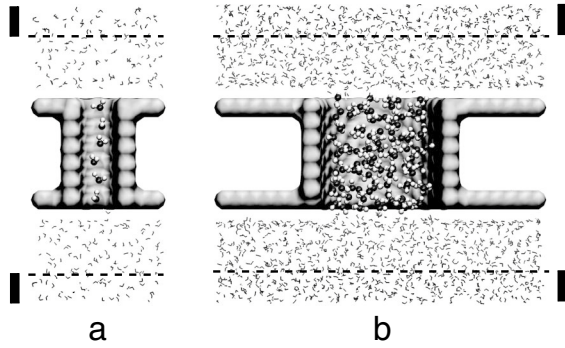


FIG. 1. Side view of the unit cells in systems *a* and *b*, with dimensions of $18.0 \text{ \AA} \times 18.0 \text{ \AA} \times 41.4 \text{ \AA}$ and $46.0 \text{ \AA} \times 42.1 \text{ \AA}$, respectively. Half of the CNT channels and the membranes are removed in order to reveal water molecules in the channels. The dashed lines and the bars indicate the layers where constant forces were applied to the water molecules in nonequilibrium simulations (see text). This figure was rendered by VMD [20].

All simulations were performed under periodic boundary conditions with constant volume. The temperature was kept constant ($T = 300 \text{ K}$) by Langevin dynamics with a damping coefficient of $5/\text{ps}$. The CNT and the membrane were fixed in all simulations. The TIP3P model [13] was used for water molecules. We employed the MD program NAMD2 [14] for the simulations, with full electrostatics calculated by the particle-mesh-Ewald method [15].

Equilibrium MD simulations of 40 and 20 ns were performed on systems *a* and *b*, respectively, with coordinates recorded every picosecond. We took the sum of one-dimensional displacements of all water molecules in the channel, divided by L , as the displacement Δn in each picosecond [cf. Eq. (3)]. If a water molecule enters or exits the channel within a picosecond, only the portion of its displacement within the channel contributes to the sum. The trajectories of $n(t)$, as shown in Fig. 2, were obtained by summing up (integrating) the Δn values as

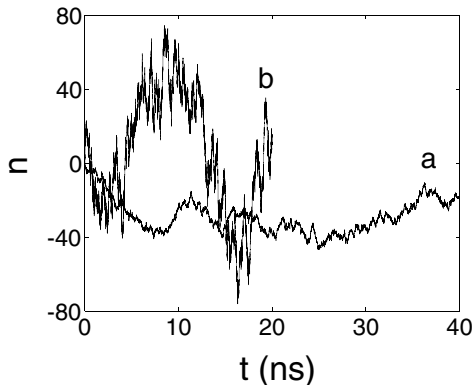


FIG. 2. Trajectories of n for equilibrium MD simulations of systems *a* and *b*.

explained above. The MSD of n for each system is presented in Fig. 3. According to Eq. (5), the diffusion coefficient D_n is one-half of the slope of the MSD- t curve. From the best-fit slopes, the D_n values were determined to be $(16.5 \pm 2.1)/\text{ns}$ and $(524 \pm 40)/\text{ns}$ for systems *a* and *b*, respectively.

To demonstrate the validity of Eq. (8) we performed nonequilibrium simulations in the presence of a chemical potential difference ($\Delta\mu$) of water across the membrane. This was achieved by using a method proposed recently [5], in which a constant force f along the z direction is exerted on all water molecules in a defined layer (with thickness d) of the bulk region. The application of f induces a hydrostatic pressure difference (ΔP), which corresponds to a chemical potential difference $\Delta\mu = fd$ across the membrane. Because of the equivalence between hydrostatic pressure and osmotic pressure, this setup can be used to calculate the osmotic permeability (p_f) of the channel. We refer the reader to [5] for a detailed explanation of the method. The defined layers in systems *a* and *b* are shown in Fig. 1, with thicknesses $d = 7.4 \text{ \AA}$ and $d = 8.1 \text{ \AA}$, respectively. By choosing a proper f , one can select any desired value for $\Delta\mu$. For each system, we performed six nonequilibrium simulations, with $\Delta\mu$ set to $0.2k_B T$, $0.5k_B T$, $1k_B T$, $2k_B T$, $5k_B T$, and $10k_B T$. The simulation times (1–40 ns) varied in different simulations, but were long enough to observe a net transport of at least 100 water molecules in each case.

Figure 4 shows both the predicted water flux (solid lines) from Eq. (8) and the observed water flux (squares) in the simulations, from which one can discern excellent agreement between predictions and simulations. It is remarkable that the water flux induced by a $\Delta\mu$ as large as $10k_B T$ can still be predicted by the D_n value determined

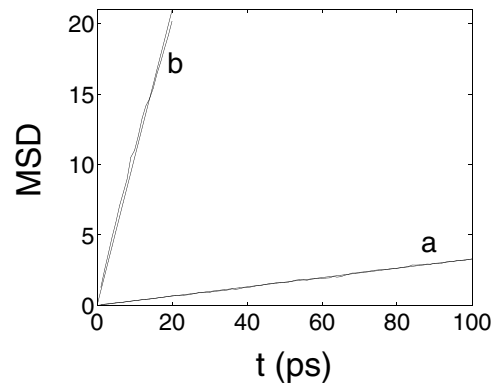


FIG. 3. Mean square displacements (MSDs) of n for systems *a* and *b*. For each system, the trajectory $n(t)$ shown in Fig. 2 was evenly divided into M (400 for system *a*, 1000 for system *b*) short time periods. $n(t)$ in each period was treated as an independent subtrajectory $n_i(t)$, and was shifted so that $n_i(t)|_{t=0} = 0$. The average over $n_i^2(t)$ ($i = 1, \dots, M$) was then taken as $\text{MSD}(t)$. A line with the best-fit slope was superimposed on each MSD curve.

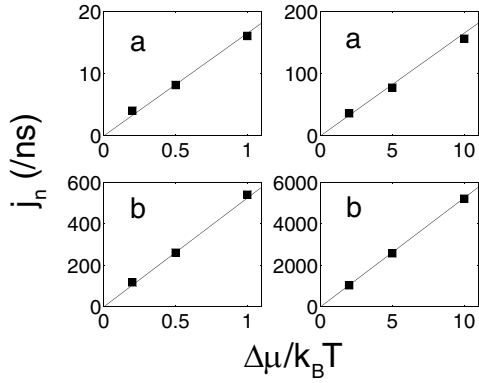


FIG. 4. The dependence of water flux (j_n) on the chemical potential difference ($\Delta\mu$) of water. Each data point (marked as a square) represents the j_n value obtained from a nonequilibrium simulation, by dividing the total displacement of n in the simulation by the simulation time [cf. Eq. (4)]. The solid lines show the j_n - $\Delta\mu$ relations predicted from Eq. (8), with $D_n = 16.5/\text{ns}$ for system *a* and $D_n = 524/\text{ns}$ for system *b*, both values being determined from the equilibrium simulations.

from equilibrium simulations. In light of this, one is not surprised that the reported osmotic permeability (p_f) of aquaporin-1 obtained from nonequilibrium simulations [5] agrees with experimental data despite the fact that the $\Delta\mu$ values ($\sim 1k_B T$) in the simulations were much larger than experimental values ($< 0.01k_B T$).

We note that Eq. (8) holds only under the assumed near-equilibrium condition. When $\Delta\mu$ is so large that the configurations of channel water deviate notably from equilibrium configurations, the evolution of n can no longer be described by D_n , and the dependence of j_n on $\Delta\mu$ will deviate from linearity. In the present case, due to the large incompressibility of liquid water, water configurations in the channel change insignificantly with ΔP (or $\Delta\mu$) even when $\Delta\mu$ exceeds $k_B T$. As a result, the equilibrium D_n describes water kinetics well and j_n is found to depend linearly on $\Delta\mu$ over a broad range. On the other hand, one expects a much smaller linear response range for gas transport through a channel, since gas configurations (such as occupancy) in the channel should be much more sensitive to $\Delta\mu$. Other factors, such as the conformation of the channel, should also be affected by $\Delta\mu$ and may in turn affect water configurations. For example, in previous simulations [5,6] we observed the blockage of aquaporin channels under large ΔP . Therefore, a real water channel may not have as large a linear response range as the fixed CNTs in the present case. In general, no universal conclusion on the j_n - $\Delta\mu$ relationship can be made for systems far from equilibrium.

In summary, the collective diffusion model establishes a quantitative relationship between the spontaneous water transport at equilibrium and the stationary water flux under nonequilibrium conditions. Using this model, the major experimental quantity for water channels, p_f , can

be determined readily from equilibrium MD simulations. In light of the many recent MD simulations of water transport through biological [3,16–18] and artificial [11,12,19] channels, the collective diffusion model should establish a valuable analysis tool.

We acknowledge NIH Grant No. P41-RR05969, NIH Grant No. 1 R01 GM067887-01, and NRAC Grant No. MCA93S028.

- [1] S. Hohmann, S. Nielsen, and P. Agre, *Aquaporins* (Academic Press, San Diego, 2001).
- [2] A. Finkelstein, *Water Movement through Lipid Bilayers, Pores, and Plasma Membranes* (John Wiley & Sons, New York, 1987).
- [3] E. Tajkhorshid, P. Nollert, M. Ø. Jensen, L. J.W. Miercke, J. O'Connell, R. M. Stroud, and K. Schulten, *Science* **296**, 525 (2002).
- [4] A. Kalra, S. Garde, and G. Hummer, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10175 (2003).
- [5] F. Zhu, E. Tajkhorshid, and K. Schulten, *Biophys. J.* **86**, 50 (2004).
- [6] F. Zhu, E. Tajkhorshid, and K. Schulten, *Biophys. J.* **83**, 154 (2002).
- [7] The MSD of a stochastic quantity, $n(t)$, and the autocorrelation function of its velocity ($v_n = dn/dt$) at equilibrium are related through [8] $\langle n^2(t) \rangle = 2 \int_0^t (t-t') \times \langle v_n(t') v_n(0) \rangle dt'$. This expression holds as long as the condition $\langle v_n(t') v_n(t'') \rangle = \langle v_n(t' - t'') v_n(0) \rangle$ is met. Equation (5) follows for t much longer than the velocity relaxation time. D_n is given by [8] $D_n = \int_0^\infty \langle v_n(t) v_n(0) \rangle dt$.
- [8] D. A. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976).
- [9] The probability distribution of n at time t , $p(n, t)$, obeys the Smoluchowski equation $\partial_t p(n, t) = (D_n \partial_n^2 + D_n \beta \Delta \mu \partial_n) p(n, t)$, whose solution is $p(n, t) = \frac{1}{\sqrt{4\pi D_n t}} \exp[-\frac{(n + D_n \beta \Delta \mu t)^2}{4D_n t}]$, where $\beta = 1/k_B T$.
- [10] A. Berezhkovskii and G. Hummer, *Phys. Rev. Lett.* **89**, 064503 (2002).
- [11] G. Hummer, J. C. Rasaiah, and J. P. Noworyta, *Nature (London)* **414**, 188 (2001).
- [12] F. Zhu and K. Schulten, *Biophys. J.* **85**, 236 (2003).
- [13] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- [14] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, *J. Comput. Phys.* **151**, 283 (1999).
- [15] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995).
- [16] B. L. de Groot, D. P. Tieleman, P. Pohl, and H. Grubmüller, *Biophys. J.* **82**, 2934 (2002).
- [17] B. L. de Groot and H. Grubmüller, *Science* **294**, 2353 (2001).
- [18] M. Ø. Jensen, E. Tajkhorshid, and K. Schulten, *Biophys. J.* **85**, 2884 (2003).
- [19] O. Beckstein and M. S. P. Sansom, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7063 (2003).
- [20] W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graphics* **14**, 33 (1996).