THE JOURNAL OF PHYSICAL CHEMISTRY B

COMMENT

pubs.acs.org/JPCB

Reply to "Comment on 'Molecular Selectivity in Aquaporin Channels Studied by the 3D- RISM Theory'"

Saree Phongphanphanee, Norio Yoshida, and Fumio Hirata*

Department of Theoretical and Computational Molecular Science, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan

In our recent paper entitled "Molecular Selectivity in Aquaporin Channels Studied by the 3D- RISM Theory", we have compared our results for the potential of mean force (PMF) of several ligand molecules inside of AQP1 and GlpF in their open states with those obtained from the MD simulation carried out by de Groot and co-workers.[1,2] The results from the two methods were markedly different; for example, PMF of glycerol in GlpF calculated from 3D-RISM is essentially negative along the channel axis, except for the small peak at the selective filter region, while that due to de Groot and co-workers has large positive peaks reaching to ~12 kJ/mol. In light of the standard definition of PMF based on the statistical mechanics, this behavior seems quite unnatural because glycerol should be comfortably accommodated inside of the pore of GlpF, or PMF should be negative or at least nearly 0.

In the Comment to our paper, which is to be published side-by-side with the present reply,[3] de Groot and co-workers claimed that the difference between the two results originates essentially from the definition of PMF in the two papers. According to them, their definition of PMF includes what they call "entropy penalty" to take into account the entropy cost that should be paid by a ligand to be confined at the small space around the channel surface. They also claim that our definition of PMF based on 3D-RISM fails in comparison with experiments because we do not include their "entropy penalty" in our definition of PMF. In this rebuttal to their comment, we clarify that their PMF is ill-defined due to inclusion of the "entropy penalty", and that it does not represent the thermodynamic stability of a ligand in the channel pore. We also clarify that their definition of PMF does not include the *dehydration penalty or gain*, which takes account of the free energy cost or gain upon both the ligand and channel pore being dehydrated.

Apparently, an important part of the difference between the two results originates from the definition of PMF. Therefore, it is natural to begin with the definition of PMF in order to clarify the problem. The definition of PMF that we have employed is what appears in standard text books of statistical mechanics[4]

$$W_\alpha(\mathbf{r}) = -kT \ln g_\alpha(\mathbf{r}) \tag{1}$$

where $g_\alpha(\mathbf{r})$ is the spatial distribution function (SDF) of solvent site (atom) α around a protein, normalized by the solvent density. The SDF is defined in terms of the configuration integral as

$$g_\alpha(\mathbf{r}) = \frac{V\int_V \dots \int_V \exp(-U(\mathbf{R},\mathbf{r},\mathbf{r}_2,\mathbf{r}_3,\dots)/kT)\, d\mathbf{r}_2\, d\mathbf{r}_3 \dots d\mathbf{r}_{N-1}}{\int_V \dots \int_V \exp(-U(\mathbf{R},\mathbf{r},\mathbf{r}_2,\mathbf{r}_3,\dots)/kT)\, d\mathbf{r}_1\, d\mathbf{r}_2 \dots d\mathbf{r}_N} \tag{2}$$

where $\mathbf{R}$ and $\mathbf{r}$ represent the atomic coordinates of a protein molecule and solvent molecules including those of ligands, respectively, and $U(\mathbf{R},\mathbf{r}_1,\mathbf{r}_2,\dots)$ represents the potential energy of the system. (In eq 2, $\mathbf{R}$ is fixed around the origin; therefore, it does not appear in the expression in our original paper. However, here, we show it explicitly for clarity.) The reason why $W(\mathbf{r})$ is called the potential of mean force is because the derivative of $W(\mathbf{r})$ with respect to the position $\mathbf{r}$ of a solvent molecule becomes a *mean force* acting on the solvent molecule from other solvent molecules as well as that from the protein. The multifold integrals with respect to the solvent coordinates, appearing in eq 2, extend over the entire box V containing the protein, ligand, and water, which is essentially infinitely large compared to the scale of a molecule. In the 3D-RISM theory, we evaluate the multifold integral over the infinitely large space *analytically* by means of the statistical mechanics.[5] By virtue of the analytical integral, we are sampling the entire configuration space of solvent (water and ligand), including that inside of the protein or channel pore. Therefore, the PMF so calculated from the 3D-RISM naturally includes the contribution from the so-called *dehydration penalty* (or gain). (The free-energy cost or gain includes the contribution from the entropy. However, the entropy cost or gain is entirely different from the "entropy penalty" defined by the Comment authors, as we clarify in a moment.) It is important to realize that PMF can be calculated entirely from the microscopic viewpoint of the molecular interactions.

In order to compare the PMF from 3D-RISM with that from the MD simulation, we have projected $g(\mathbf{r})$ onto the z-axis of the channel using the following equation[6]

$$g(z) = \frac{\int_{\text{area}\, g(r) > 10^{-4}} g(\mathbf{r})\, dx\, dy}{\int_{\text{area}\, g(r) > 10^{-4}} dx\, dy} \tag{3}$$

The integral in eq 3 is carried out over the space inside of the channel, where $g(\mathbf{r})$ is greater than 0 (where we regarded $g(\mathbf{r}) < 10^{-4}$ as 0). The quantity is normalized by the volume in the pore, where a solute can access; it is not concerned with any confined solute in a cavity outside of the pore. The *dehydration penalty* or gain is naturally included in the definition of PMF (eq 2). Our results are the most faithful realization of the PMF that has a solid foundation of the statistical mechanics. There is no ambiguity in our definition of PMF of a ligand in a channel.

On the other hand, the definition of PMF by the Comment authors is quite ambiguous. The most important point with wich the authors are concerned is what they call the "entropy penalty",

Received: April 19, 2011
Published: June 06, 2011

defined by the following equation[3]

$$\Delta G(z) = -kT \ln(A_{\text{pore}}(z)/A_{\text{mono}}) \quad (4)$$

where A_{mono} (=A_{ref}) is the cross section area of the channel monomer at the membrane surface and A_{pore} is the cross section area of the pore region of the channel. According to their explanation, this quantity takes account of the entropy cost that should be paid by a ligand to be confined at the small space around the channel surface. The authors attributed the large positive value in PMF of a ligand inside of the channel to the "entropy penalty" estimated based on the definition.

Unfortunately, this definition of the "entropy penalty" does not make sense in many ways. First of all, $-kT \ln A_{\text{mono}}$ does not represent the free energy (or entropy) of a solute in the bulk. Second, the definition of A_{mono} is ambiguous, as is indicated by the authors themselves: "A_{ref} can be chosen freely." Third, one may produce any number for the "entropy penalty" that he wishes.

The ambiguity in their definition apparently originates from their confusion of the microscopic PMF defined by eq 1 with an "equivalent" or "effective" activation energy that determines the permeation rate of the ligand through a channel. In order to clarify this point, let us reproduce the procedure they they used to reach the expression (eq 4) for their "entropy penalty". (The derivation may not be exactly the same as what they have done because they did not provide any derivation to reach the equation. However, our derivation is sufficient to clarify the essential physics of their "entropy penalty".) The experimental flux of a solute (ligand) through a membrane can be expressed as

$$j = \sigma \Delta \Pi \quad (5)$$

where $\Delta\Pi$ is the difference of the osmotic pressure between a lipid membrane and σ is the conductivity or permeability. It is common practice to use the Arrhenius-type equation of the activation barier crossing to analyze the conductivity

$$\sigma = f \exp(-\Delta E_a/kT) \quad (6)$$

where f and ΔE_a are the frequency factor and the activation energy. The flux j is an overall flux that consists of the contributions from the two different regions of membrane surfaces, one through the lipid bilayer, j_L, and the other through the channel, j_C. The flux j_L through the lipid bilayer is extracted by experimentalists by performing a separate experiment that is just concerned with the membrane without channels. (The entire surface area of the membrane can be attributed to the lipid bilayer.) In order to extract the flux j_C through the channel region, the information concerning the ratio of the membrane surface area between the two regions is required because $Aj = A_L j_L + A_C j_C$, where A, A_L, and A_C denote the surface area of the entire membrane, that of the lipid bilayer, and that of the channel, respectively. However, such information is not available from experiments. Therefore, the Comment authors have introduced a quite ambiguous concept, or "reference area", A_{ref}, and equated it with the cross section area of a channel monomer. Further, they equated the flux through A_{mono} with that through the cross section area of channel pore, or

$$j_{\text{mono}} A_{\text{mono}} = j_{\text{pore}} A_{\text{pore}} \quad (7)$$

This assumption is reasonable because solute can only be permeated through the pore region of the channel. If one expresses j_{mono} and j_{pore} in terms of the eqs 5 and 6 and puts that into eq 7, one gets, by assuming that the frequency factor f is the same

$$\Delta G = E_{\text{pore}} - E_{\text{mono}} = -kT \ln[A_{\text{pore}}/A_{\text{mono}}]$$

This expression of a phenomenological activation barrier is already extremely ambiguous due to the ad hoc definition of A_{ref} or A_{mono}. However, it is not the end of the story. They extended the phenomenological quantity to the entire channel region by making A_{pore} dependent on the channel coordinates z, or $A_{\text{pore}}(z)$, without any theoretical and physical foundations.

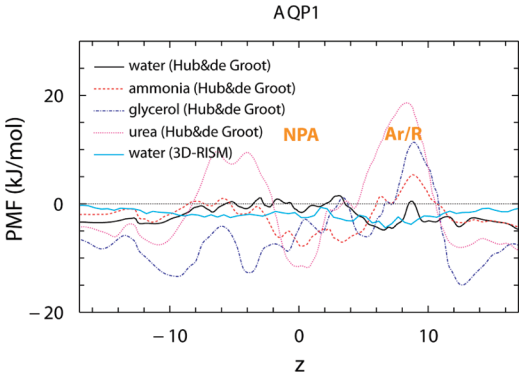

Figure 1. PMF corresponding to the pore area of solutes in AQP1.

We have clarified so far that the large positive peaks in their "PMF" are caused by the inclusion of an "entropy penalty", which has an unphysical origin. Therefore, it may be of interest to compare our PMF obtained from the 3D-RISM theory with their PMF calculated from the umbrella sampling without the "entropy penalty" correction. However, there is one problem for such comparison, which is the *dehydration penalty* or gain. The PMF calculated from the umbrella sampling does not account for the quantity, while that from 3D-RISM naturally includes it. However, there is a case in which the *dehydration penalty* and gain do not make an important contribution to PMF. It is the case where pure water is only a species permeated through the channel. In this case, it can be verified by a simple thought-experiment that the *dehydration penalty* and gain largely cancel out. During the MD simulation, water molecules are changing their positions contiguously. When one water molecule gets into the channel pore from one side of the channel, another water molecule should be getting out of the other side of the channel in order to keep the chemical potential of water molecules inside of the channel constant. The free-energy penalty spent by a water molecule entering the channel should be largely canceled by the free-energy gain of a water molecule that is exiting. Therefore, in the case of water permeated through the channel, their results of PMF calculated from the umbrella sampling, without the "entropy penalty", are expected to be in better agreement with our results of PMF from the 3D-RISM theory. In Figures 1 and 2, PMFs of water and other ligands inside of AQP1 and GlpF pores are plotted, which are obtained by subtracting the contribution due to eq 4 (or the "entropy penalty") from the corresponding results in Figures 1 and 2 of ref 2. (In the plot, we assumed that the accessible area of a solute is equal to the area of the pore that was generated by HOLE.)[7] Our PMF results for water are also plotted in the figures. As can be seen, their results for PMF of water in the channel agrees reasonably well with our results, which are in accord with our physical intuition that PMF of water inside of the aquaporin channel should be negative or 0.

dx.doi.org/10.1021/jp203661z |J. Phys. Chem. B 2011, 115, 8367–8369

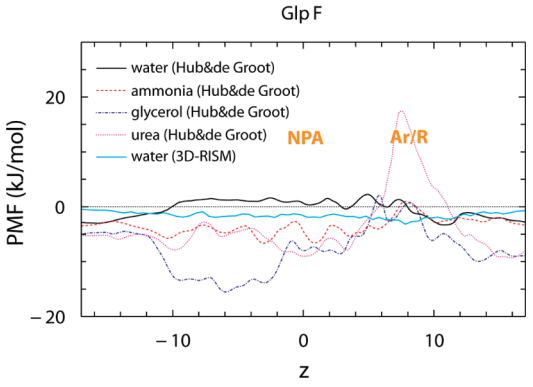

Figure 2. PMF corresponding to the pore area of solutes in GlpF.

The situation is a little different in the case of ammonia, glycerol, and urea. For those cases, one has to take the "dehydration" penalty or gain of the ligand and of channel into consideration explicitly because dehydration penalties of ligands and of the channel pore will not cancel each other in these cases. Nevertheless, we have plotted in Figure 2 their results of PMF for NH_3, glycerol, and urea in GlpF after subtracting the "entropy penalty" contribution from their results in Figures 1 and 2 of ref 2. The results look very reasonable in light of our physical intuition and are in much better agreement with our results from the 3D-RISM theory, although they show some odd behaviors that might be attributed to either disregarding the hydration penalty or insufficient sampling. (We are still concerned with the convergence of their umbrella sampling because the values of PMF at the channel entrance and exit are significantly different, according to their Supporting Information.)

CONSISTENCY WITH EXPERIMENTS

Other points made by the authors of the Comment on our paper are concerned with the consistency between the theoretical and experimental results.[3] They refer to two cases, one concerning ammonia permeation through AQP1 and GlpF and the other the relative permeability of urea versus glycerol for GlpF. In the ammonia case, they claim that their results are consistent with the experimental results because the conduction rate of GlpF is ~10 times higher than that of AQP1.[3] In the case of urea versus glycerol, they claimed that the conductivity of glycerol through GlpF, converted from their "PMF", is ~70 times greater than that of urea and that the results are "in line" with the experimental results of Agre.[3,8]

However, the agreement of their results with the experiment's is senseless. Their "PMFs" to be compared with the activation barrier from the experiments are largely contaminated by what they call an "entropy penalty", calculated based on eq 4 as we clarified above. The value of the quantity can be adjusted arbitrarily by choosing A_{ref} as they wish. Therefore, their results are not free from suspicion that they used A_{ref} as an adjustable parameter to bring their results "in line" with the experiments.

We believe that the results regarding conductivity produced by any theoretical methods are not ready to be compared with experiments in any serious manner at this moment. As is well documented in the experimental papers concerning aquaporins, the conductivity of a ligand is measured by phenomenological methods, such as the rate of volume increase of the liposome due to the flux of ligands permeated through the membrane, responding to the osmotic pressure. They analyze the overall rate constant in terms of an Arrhenius-type rate equation or of the activation barrier crossing mechanism. The rate of ligand permeation through the membrane or the activation barrier so determined has many different physical origins, the probability of ligands to find the entrance of the channel, PMF including the *dehydration penalty* (or gain) for both the channel pore and ligands, structural fluctuations of protein that are related to the gating mechanism of the channel, and so forth. The gating mechanism, especially, is supposed to make an important contribution to the activation barrier because it is the most distinct characteristics of molecular channels to determine the conduction rate of ligands through a molecular channel. In our case, the models of the channels are only those in "open" states, which do not include any gating mechanism to the closed sate. Therefore, we are not ready to compare our results with experiments of the conductivity of ligands through the channel. Thereby, we have confined our discussion just within the distribution and the PMF of ligands inside of the channel pore. When we said "agreement with experiments" in our original paper, we meant "qualitative agreement with the experiments concerning the distribution of ligands inside of the channels", which can be measured directly by X-ray and neutron diffractions in some cases or be deduced intuitively from the fact that the channel permeated the ligands.

AUTHOR INFORMATION

Corresponding Author
*E-mail: hirata@ims.ac.jp.

REFERENCES

(1) Phongphanphanee, S.; Yoshida, N.; Hirata, F. *J. Phys. Chem. B* **2010**, *114*, 7967–7973.
(2) Hub, J. S.; de Groot, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1198–1203.
(3) Hub, J. S.; de Groot, B. L. *Comment on "Molecular Selectivity in Aquaporin Channels Studied by the 3D- RISM Theory"* **2011**10.1021/jp2022242.
(4) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000.
(5) Kovalenko, A. In *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer: Dordrecht, The Netherlands, 2003; Vol. 169, pp 169−275.
(6) Phongphanphanee, S.; Yoshida, N.; Hirata, F. *J. Mol. Liq.* **2009**, *147*, 107–111.
(7) Smart, O. S.; Neduvelil, J. G.; Wang, X.; Wallace, B. A.; Sansom, M. S. P. *J. Mol. Graphics* **1996**, *14*, 354–360.
(8) Borgnia, M.; Agre, P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2888–2893.