FULL PAPER

On the Cavitation Energy of Water

Siegfried Höfinger*[a, b] and Francesco Zerbetto*[a]

Abstract: Free-energy-perturbation theory from molecular dynamics calculations has been used to obtain the ΔG of adjoining cavities' formation in water. The ΔGs for systems with three, five and seven cavities are compared with that of a single cavity of the same volume, and found to be in good agreement. The conditions under which the analytical formulation of the energy of cavity formation proposed by Pierotti holds are discussed. The data for a single cavity have been tabulated and can lend themselves to a simple numerical implementation in standard quantum chemical packages, which can be used when high accuracy for ΔG^{cav} is required.

Introduction

Intermolecular interactions are the sum of several terms, which include electrostatic, repulsion, induction and dispersion terms. In explicit solvent models, solvent–solute interactions are described in this way; however, the introduction of implicit solvent models has brought in a new term, namely the free energy of cavity formation, ΔG^{cav}, or *FEC*.[1] This is the reversible work spent in the process of making a hole or a cavity in a solvent. To a first approximation, it depends on how many particles must be displaced, their size, the interactions present in the solvent, the pressure and the temperature. Pierotti developed a general analytical equation, PE, to obtain the cavitation energy as a function of the cavity's radius, which contains explicitly the reduced number density (that is, how many particles are present in a volume element—which, in turn, depends on the temperature), the solvent hard-sphere radius, the pressure and the temperature. Because of the way used by Pierotti to develop his approach (i.e., scaled particle theory), entropic contributions are not well accounted for.[1] This is especially true in rather extreme cases, such that of water, where its strong and dynamical hydrogen bonding may produce sizeable energy contribution when a molecule protrudes inside the cavity. So far, the best independent support for PE was given by Postma, Berendsen and Haak, PBH, who compared its values with the results of free-energy-

perturbation, FEP, calculations[2] obtained from molecular dynamics simulations. In the PBH approach, a number of calculations were run with a repulsive potential set inside a box of 216 water molecules in order to create a cavity. Modification of the parameters of the potential energy functions, changed the size of the cavity. This pioneering work was one of the first exhaustive free-energy-perturbation calculations on solvent effects. In practice, five radii of cavities were considered and allowed PBH to conclude that 1) it is possible to obtain the free energy of cavity formation from molecular dynamics calculations and 2) PE agrees well with the results of the simulations.

The Pierotti equation has now become so popular that it has been incorporated in a number of implicit solvent models of routine use in quantum chemistry. Importantly, when using PE, the free energy of cavity formation for a solute molecule is usually calculated with a set of scaled contributions from individual interpenetrated spheres and not with the molecular contour. One might, however, wonder if the standard approach could be replaced by considering a single cavity. Furthermore, given, the widespread computer power presently available, one could also consider a finite-differences approach that would extend the approach pioneered by PBH and obtain FEC in the form of a grid of points. The numerical approach of tabulating the molecular dynamics data would extend and generalise PE to cases in which it does not hold.

Here, we calculate the water-cavity energy for one, three, five and seven adjacent spheres and tabulate the FEC as a function of the radius using a recent and accurate force field. The conditions for the best agreement between PE and FEP calculations are discussed.

[a] Dr. S. Höfinger, Prof. F. Zerbetto
Dipartimento di Chimica "G. Ciamician", Università di Bologna
V. F. Selmi 2, 40126 Bologna (Italy)
Fax (+39) 051-2099473
E-mail: gatto@ciam.unibo.it

[b] Dr. S. Höfinger
Present address:
Novartis Research Center, Vienna
Brunnerstrasse 69, 1235 Vienna (Austria)
Fax (+43) 1-866-34-727
E-mail: siegfried.hoefinger@pharma.novartis.com

Computational Methods

The simulations were run with the TINKER 3.8 program,[3] which has found several satisfactory applications in our laboratory.[4] Water was

simulated with Ponder's multipole-based force field.[5] The box size was dimensioned to reproduce the experimental density at $T = 300$ K. Changes in the size of the box occur because of the variation of the cavity radius and the pressure coupling.[6] Internal degrees of freedom were allowed to relax. Trajectories were calculated for 100 ps with a time step of 1.0 fs, the first 25 ps were used for equilibration. To create the cavity, a potential of the type

$$V_{c,w} = \lambda \left(\frac{B}{r} \right)^{12} \tag{1}$$

was introduced at the centre of a box of 216 molecules of water. B defines the cavity radius and all the intermediate cavity radii B_i correspond to a well-defined value of $0 \leq \lambda \leq 1$ (note that PBH use thermal radii; to obtain the corresponding B_i values they must be divided by 1.0044047). As the cavity repulsive radius tends to zero, the potential introduces a discontinuity that can be avoided by the use of a softer potential[7]

$$V_{c,w} = \frac{\lambda^n}{\left[0.3(1 - \lambda)^2 + \left(\frac{r}{B} \right)^6 \right]^2} \tag{2}$$

between 0 and 1 Å, in which $n = 12$, although other choices can be made. Figure 1 illustrates this soft potential for selected B_i values of the cavity
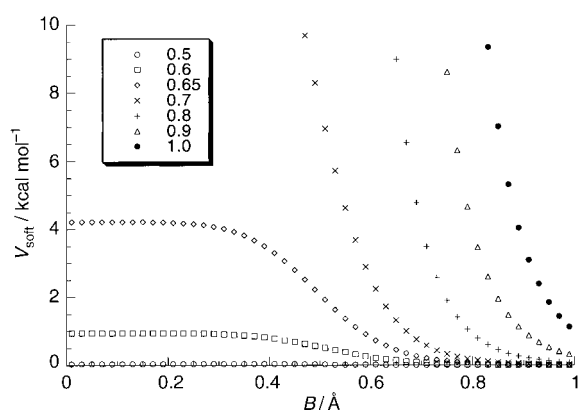

Figure 1. Potential energy curves from Equation (2) for selected values of the cavity radius, B_i. In the inset, the values of λ.

radius. In analogy to the overlapping-spheres technique used by PBH, the potential $V_{c,w}(\lambda)$ is expressed in terms of B_i through two steps that exploit Equations (1) and (2), in which the subscript c stands for cavity, the subscript w stands for water, and r is the smallest distance between any atom of a water molecule and the centre of the cavity. In their work, PBH set B_i to 0.96, 1.71, 2.28, 2.86, and 3.04 Å. Here, a total of 24 points are used. They are 0.5, 0.6, 0.65, 0.67, 0.7, 0.8, 0.9, 1.00 Å, which were chosen because they smooth out the singularity at the origin and make it disappear, and then 15 equally spaced points between 1 and 2.28 Å with increments of 0.08 Å.

Small variations of B_i around the selected cavity value gave the free energy, which is calculated as

$$\Delta G^{cav} = -k_B T \ln \left\langle e^{-\left[\frac{1}{k_B T} \sum_k^{atoms} \frac{(B_i + \delta B)^{12} - B_i^{12}}{r_k^{12}} \right]} \right\rangle_i \tag{3}$$

Here the energy is calculated at one of the reference B_i values over the ensemble of particles in the unperturbed state i with small perturbations δB along the trajectory, k_B is the Boltzmann constant, and the perturbation V_p is caused by the change from B_i to $B_{ref} = B_i \pm \delta B$

$$V_p(B_{ref}) = \sum_k \Delta V_{c,w,k} = \sum_k^{atoms} \frac{B_{ref}^{12} - B_i^{12}}{r_k^{12}} \tag{4}$$

which was calculated at one of the reference values given above, with k that runs over the water atoms. The various $V_p(B_{ref})$ must connect adiabatically for δB values both positive and negative. This requirement is implicit in the definition of the cavity energy as reversible work, and in the free-energy-perturbation model. Moreover, its value must be at most of the order of

$k_B T$ for perturbation theory to hold. The δB were set to multiples of 0.0032 Å, which generated 25 perturbation points between pairs of trajectories in the region where the potential of Equation (1) was used. In the initial step, in which Equation (2) was used δB was scaled down as needed. Because of the nature of Equation (3), there is a redundancy of values of V_p, and of the associated ΔG^{cav}, around a given B_i. As an example, and only as an example, $V(1.8384)$, can be obtained from $B_i = 1.8$ Å by setting $\delta B = 0.0384$ Å, or from $B_i = 1.88$ Å by setting $\delta B = -0.0416$; or $V(1.8192)$, can be obtained from $B_i = 1.8$ Å by setting $\delta B = 0.0192$ Å, or from $B_i = 1.88$ Å by setting $\delta B = -0.0608$ Å. The sum of the first two should be the same as the sum of the last two, and should also be equal to all the sums of the other pairs that connect $B_i = 1.8$ Å with $B_i = 1.88$ Å (this is the principle of overlapping-spheres techniques of ref. [2]). In reality, this does not happen all the time. The differences provide a way of estimating the error in the procedure (as the largest difference between the points). Importantly, if some values do not fulfil the $k_B T$ criterion, they are discarded. The final $\Delta G(B)$ is the sum over the successive incremental ΔG values with cavity radii smaller than B.

The analytical expression of the free energy of cavity formation given by Pierotti reads

$$\Delta G_c = K_0 + K_1 r_B + K_2 r_B^2 + K_3 r_B^3 \tag{5}$$

in which

$$K_0 = RT \left[-\ln(1 - y) + \frac{9}{2} \left(\frac{y}{1 - y} \right)^2 \right] - \frac{\pi p a^3}{6} \tag{6a}$$

$$K_1 = -\frac{RT}{a} \left[\frac{6y}{1 - y} + 18 \left(\frac{y}{1 - y} \right)^2 \right] + \pi p a^2 \tag{6b}$$

$$K_2 = \frac{RT}{a^2} \left[\frac{12y}{1 - y} + 18 \left(\frac{y}{1 - y} \right)^2 \right] - 2\pi p a \tag{6c}$$

$$K_3 = \frac{4}{3} \pi p \tag{6d}$$

here $r_B = (a + a_2)/2$ is the cavity radius given by half the sum of the solvent hard-sphere diameter, a, and the solute hard-sphere diameter, a_2, $y = \pi a^3 \rho / 6$ is the volume fraction of the solvent spheres, ρ is the number density of the pure solvent, p is the pressure. At atmospheric pressure, the last addends, linear in p, are negligible. Notice that r_B and B_i coincide for many practical purposes; however, the former describes the cavity when the solute is present, while the latter is used in the FEP calculations where a "bubble" of vacuum is blown up inside water. For clarity's sake, we feel that it is important to retain the distinction.

Results and Discussion

Figure 2a shows the free energy of formation of a single cavity as a function of its radius. The trend is smooth and the error bars grow with the radius of the cavity because, as explained in the previous section, the error in the procedure is additive and increases with the size of the cavity. (Comparison with the values calculated by PBH is shown later along with other FEC data.)

Two important issues must be addressed:
1) The agreement between the data in Figure 2a and those predicted by Pierotti's formula,
2) The approximation caused by replacing several adjoining cavities with a single cavity of the same volume, that is by replacing the molecular contour with a single sphere.

To investigate the first point, one must consider that in Equation (5) $r_B = (a + a_2)/2$ is the sum of the hard-sphere radii of solvent and solute. From Equations (5) and (6) it is
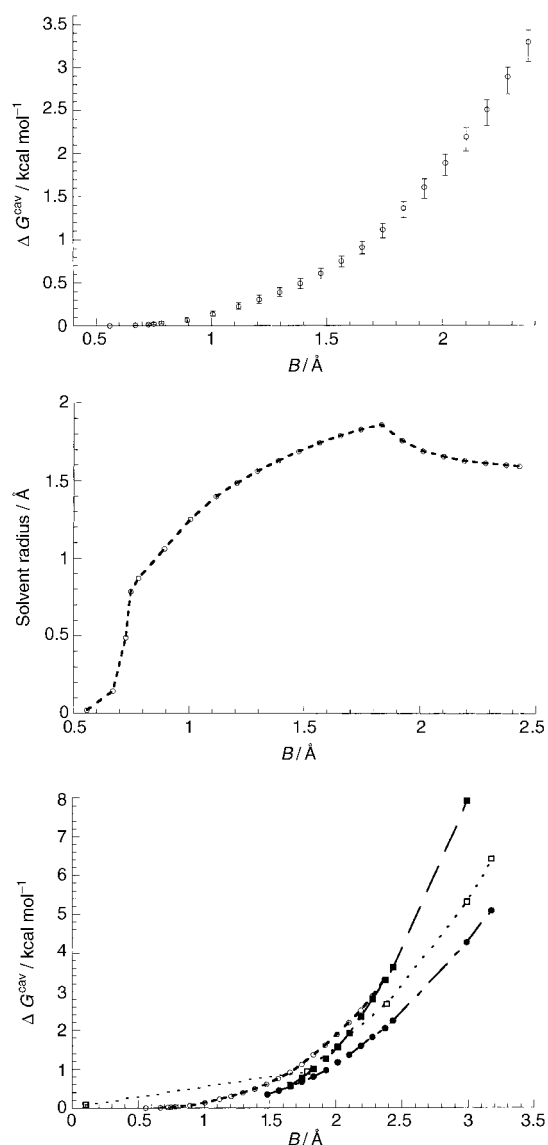
Figure 2. a) Free energy of cavity formation from the present *FEP* calculations, b) solvent hard-sphere radius that gives the best agreement between the *FEP* calculations and Equations (5) and (6), c) Comparison between *FEP* (○: present calculations, □: PBH) and PE by using Gaussian 98 (The Gaussian 98 cavitation energy was obtained by setting the solvent radius to 1.40 Å (●) and 1.6 Å (■) and then performing HF/3-21G PCM calculations on Cl⁻ anions with a growing size, comparable to the hard-sphere cavities of the FEP calculations.

apparent that the data of Figure 2a can be obtained by an infinite number of combinations of values of a and a_2. It is, however, simple to invert Equation (5) by using the present FEP results and determine which values of a and a_2 best fit the calculated data. Because of the applications in quantum chemical packages, the knowledge of the diameter, a, of water is important. For each of the overall radii of Figure 2a, Figure 2b shows the plot of the solvent radius that gives the best agreement between the FEP calculations and PE. The result is a curve that deserves closer inspection. Two regions can be noted:

1) When B is small, that is, less than 1.6 Å, PE is most successful if the cavity is entirely due to the solvent, a feature that is not unexpected since the solvent contributes

to the radius of the cavity. The trend is not perfectly linear and some irregularities are present. They are ascribed to the numerical procedure that is used to obtain the curve and should not have any physical implication,

2) At larger values of B, that is, $B \geq 1.6$ Å, the water radius that best satisfies PE tends towards the limiting value of 1.6 Å.

In practice, since in an implicit solvent model the cavity is defined by the solvent and solute radii, only the second region can be adequately treated by Pierotti's approach. For regions with small cavity diameter, the use of Equation (5) should therefore be discouraged.

Figure 2c compares the present results with the PBH data and with those obtained by a standard quantum chemical package when the two values of 1.4 and 1.6 Å are used in Equation (5). Note that 1.4 Å is the default value for water in the program.[8] When the cavity is not too large, the *FEP* calculations lead to the largest ΔG^{cav} of the four shown here, and they are also close to the results of PE when the water radius is set equal to 1.6 Å. The satisfactory similarity with the PBH data implies that the calculations are not a strong function of the force-field model used to simulate water. The default value of 1.4 Å set in the quantum chemical package leads to the lowest ΔG^{cav}. Because of the variability of the radii that one would have to use in Pierotti's formula, it is suggested that when high accuracy is required, the FEP calculations could be directly looked up in Table 1. These values can be interpolated or extrapolated rather easily once the solute hard-sphere radius is determined.

The second issue raised at the beginning of the section concerns the comparison of the ΔG^{cav} of several spherical systems in contact through a single point with that of a single

Table 1. Free energy of cavity formation, ΔG [kcal mol⁻¹], from the free-energy-perturbation calculations. In agreement with PBH, each cavity radius [Å] is multiplied by 1.044 to obtain the thermal radius and then by 1.0704 to obtain the hard-sphere radius. The origin of the errors is discussed in the text.

B	ΔG^{cav}	Lower ΔG^{cav} bound	Upper ΔG^{cav} bound
0.558774	0.0000	0.0000	0.0000
0.670529	0.0063	0.0041	0.0080
0.726406	0.0151	0.0089	0.0265
0.748757	0.0213	0.0147	0.0364
0.782284	0.0300	0.0214	0.0534
0.894039	0.0633	0.0476	0.0983
1.005793	0.1333	0.1122	0.1694
1.117548	0.2243	0.1913	0.2650
1.206952	0.3022	0.2580	0.3535
1.296356	0.3890	0.3386	0.4427
1.38576	0.4894	0.4316	0.5472
1.475164	0.6073	0.5450	0.6653
1.564568	0.7507	0.6828	0.8087
1.653971	0.9116	0.8312	0.9782
1.743375	1.1136	1.0181	1.1847
1.832779	1.3607	1.2523	1.4353
1.922183	1.6053	1.4710	1.7006
2.011587	1.8873	1.7401	1.9853
2.100991	2.1909	2.0238	2.2988
2.190395	2.5079	2.3251	2.6179
2.279799	2.8874	2.6867	3.0000
2.369202	3.2967	3.0646	3.4293

cavity with the same volume. Figure 3 shows the same *FEP* data reported in Figure 2a along with the ΔG^{cav} for 3, 5 or 7 small spheres, which are in contact by one point. The arrangement of the cavities is shown in Figure 4. If one considers the errors intrinsic to the free-energy-perturbation approach (see above), the multiple cavity data are in good agreement with the data of *FEC* for a single cavity of the same volume. The use of a "representative spherical" cavity is therefore justified even for systems that are far from perfectly spherical.
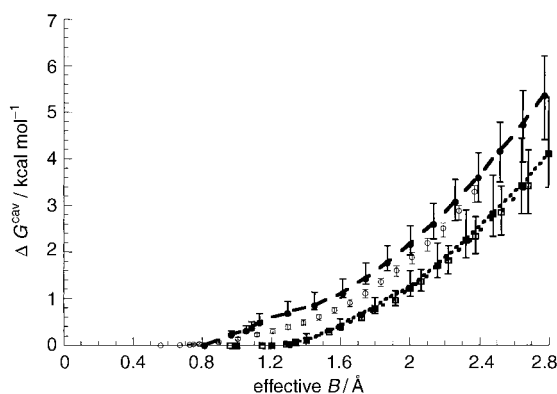

Figure 3. Comparison of the free energies for a single cavity (○) and a cavity with the same volume built from 3 (●), 5 (□) or 7 (■) adjacent cavities.
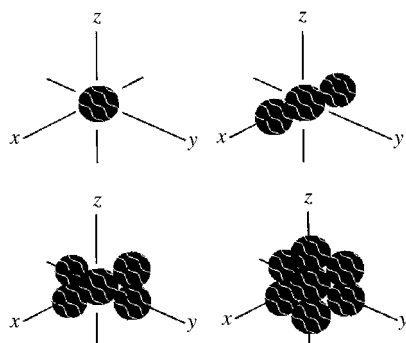

Figure 4. Water cavity arrangement for the data of Figure 3.

Conclusion

The free energy for the formation of a cavity in water has been calculated from free-energy-perturbation simulations. There is good agreement with previous similar calculations; this shows that the results are not a strong function of the force field used to describe water. The calculations also agree reasonably well with the analytical approach of Pierotti. The water hard-sphere diameter was derived in a parameter-free way (of course, the force field contains parameters but the *FEP* procedure does not introduce new parameters after the molecular mechanics model has been chosen). It is found that the water radius is not constant and tends to reach a limiting value of about 1.6 Å. Interestingly, in a recent study[9] one of us showed that the most appropriate water volume for the Connolly surface is 16.4 Å^3, which results in a radius of 1.58 Å in noteworthy agreement with the value of 1.6 Å that comes out of the present work. To account for the variability of the

radius, it is suggested that *FEP* calculations are tabulated and used numerically when high accuracy is required. It is also found that the energy of cavitation is nearly additive; this would justify the use of a single spherical "representative" cavity to describe highly nonspherical molecules of equal molecular volume. The results reported here are relevant not only to the extensive modelling of the solvent effect that is now being carried out with computer simulations, but may also be used to understand hydrophobic interactions in which the energy of cavity formation plays an important role.[10]

[1] R. A. Pierotti, *Chem. Rev.* **1976**, *76*, 717 – 726.
[2] P. M. Postma, H. J. C. Berendsen, J. R. Haak, *Faraday Symp. Chem. Soc.* **1982**, *17*, 55 – 67.
[3] a) J. W. Ponder, F. Richards, *J. Comput. Chem.* **1987**, *8*, 1016 – 1024; b) C. Kundrot, J. W. Ponder, F. Richards, *J. Comput. Chem.* **1991**, *12*, 402 – 409; c) M. J. Dudek, J. W. Ponder, *J. Comput. Chem.* **1995**, *16*, 791 – 816.
[4] V. Bermudez, N. Capron, T. Gase, F. G. Gatti, F. Kajzar, D. A. Leigh, F. Zerbetto, S. Zhang, *Nature* **2000**, *406*, 608 – 611; M. Cavallini, R. Lazzaroni, R. Zamboni, F. Biscarini, D. Timpel, F. Zerbetto, G. J. Clarkson, D. A. Leigh, *J. Phys. Chem. B* **2001**, *105*, 10826 – 10830; F. Biscarini, M. Cavallini, D. A. Leigh, S. León, S. J. Teat, J. K. W. Wong, F. Zerbetto, *J. Am. Chem. Soc.* **2002**, *124*, 225 – 233.
[5] Y. Kong, J. W. Ponder, *J. Chem. Phys.* **1997**, *107*, 481 – 492.
[6] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684 – 3690.
[7] a) T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, W. F. van Gunsteren, *Chem. Phys. Letters* **1994**, *222*, 529 – 539; b) T. Simonson, *Mol. Phys.* **1993**, *80*, 441 – 447.
[8] Gaussian 98 (Revision A.7), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh, PA, **1998**.
[9] S. Höfinger, O.Steinhauser *J. Chem. Phys.*, **2001**, *115*, 10636 – 10646.
[10] a) E. Gallicchio, M. M. Kubo, R. M. Levy, *J. Phys. Chem. B* **2000**, *104*, 6271 – 6285; b) M. A. Gomez, L. R. Pratt, G. Hummer, S. Garde, *J. Phys. Chem. B* **1999**, *103*, 3520 – 3523; c) H. S. Ashbaugh, S. Garde, G. Hummer, E. W. Kaler, M. E. Paulaitis, *Biophys. J.* **1999**, *77*, 645 – 654; d) S. Garde, G. Hummer, M. E. Paulaitis, *Faraday Discuss.* **1996**, *103*, 125 – 139; e) G. Hummer, S. Garde, A. E. Garcia, M. E. Paulaitis, L. R. Pratt, *J. Phys. Chem. B* **1998**, *102*, 10469 – 10482; f) G. Hummer, S. Garde, A. E. Garcia, A. Pohorille, L. R. Pratt, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8951 – 8955; g) M. Prevost, D. Vanbelle, G. Lippens, S. Wodak, *Mol. Phys.* **1990**, *71*, 587 – 603; h) L. R. Pratt, G. Hummer, A. E. Garcia, *Biophys. J.* **1994**, *51*, 147 – 165; i) G. Hummer, S. Garde, A. E. Garcia, L. R. Pratt, *Chem. Phys.* **2000**, *258*, 349 – 370; j) A. Pohorille, L. R. Pratt, *J. Am. Chem. Soc.* **1990**, *112*, 5066 – 5074; L. Pratt, A. Pohorille, *Chem. Rev.* **2002**, *102*, 2671 – 2691.